

The oligomeric Rep protein of Mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase

Nirupam Roy Choudhury, Punjab Singh Malik, Dharmendra Kumar Singh, Mohammad Nurul Islam, Kosalai Kaliappan and Sunil Kumar Mukherjee*

Plant Molecular Biology Group, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi-110 067, India

Received July 31, 2006; Revised October 11, 2006; Accepted October 13, 2006

ABSTRACT

Geminiviruses replicate by rolling circle mode of replication (RCR) and the viral Rep protein initiates RCR by the site-specific nicking at a conserved nonamer (TAATATT↓AC) sequence. The mechanism of subsequent steps of the replication process, e.g. helicase activity to drive fork-elongation, etc. has largely remained obscure. Here we show that Rep of a geminivirus, namely, Mungbean yellow mosaic India virus (MYMIV), acts as a replicative helicase. The Rep-helicase, requiring ≥ 6 nt space for its efficient activity, translocates in the 3'→5' direction, and the presence of forked junction in the substrate does not influence the activity to any great extent. Rep forms a large oligomeric complex and the helicase activity is dependent on the oligomeric conformation (~24mer). The role of Rep as a replicative helicase has been demonstrated through *ex vivo* studies in *Saccharomyces cerevisiae* and *in planta* analyses in *Nicotiana tabacum*. We also establish that such helicase activity is not confined to the MYMIV system alone, but is also true with at least two other begomoviruses, viz., Mungbean yellow mosaic virus (MYMV) and Indian cassava mosaic virus (ICMV).

INTRODUCTION

Geminiviruses infect a wide variety of crop plants causing great economic losses, and are subjects of immense concern worldwide. They replicate in plant nuclei by a rolling-circle mechanism. Geminiviruses encode only a few proteins for their replication and recruit most of their replication machinery from their plant hosts (1). Our attention is focused in unraveling the mechanism of DNA replication of a model geminivirus, namely Mungbean yellow mosaic India virus

(MYMIV). MYMIV belongs to the genus Begomoviridae and infects a number of leguminous plants using whitefly (*Bemisia tabaci*). MYMIV possesses bipartite ssDNA genomes named as DNA-A and DNA-B, both being ~2.7 kb in size (2). Both components share a common region (CR) of about 200 bp containing the important *cis*-elements for viral DNA transcription and rolling circle replication (RCR) (3).

The replication initiator protein Rep (or AC1), encoded by DNA-A, is the major protein required for replication of any geminiviral DNA genome (4–6). Rep is a multifunctional protein having site-specific nicking and ligation, site-specific DNA-binding and ATPase activities (3,7,8). It regulates its own expression at transcriptional level and is also known to induce host replication machinery, presumably to enable the virus to replicate in differentiated cells (9). The *ori* of MYMIV DNA-A contains three direct repeats (CGGTGT) that flank both sides of the stem-loop region of the *ori*. During RCR, Rep binds to these repeat elements and eventually makes a site-specific nick at TAATATT↓AC of the loop region of the plus strand to initiate replication. The 3'-OH end of the nicked DNA is probably then used as a primer for synthesis of the viral DNA (10). It is possible that other viral and host proteins assist Rep in the above-mentioned RCR-initiation process.

Despite the considerable available knowledge about the RCR-initiation step, very little is known about the subsequent steps. Following the Rep-mediated nicking of the viral strand, a helicase activity will be required to dislodge the parental strand, assemble the fork-proteins at the primer-end and move the replication fork forward (5). Hence, identification and characterization of helicase activity for geminivirus RCR are of utmost significance. Considerable effort in search of the helicase responsible for unwinding the replicating geminiviral DNA has remained elusive so far. The Rep proteins isolated from various geminiviruses have been shown to possess the ATPase activity and their helicase activities have been speculated *in silico* (8,11–13). However, no helicase activity of any of the geminiviral Rep proteins has been demonstrated so far.

*To whom correspondence should be addressed. Tel: +91 11 26189358; Fax: +91 11 26162316; Email: sunilm@icgeb.res.in

Though the detailed structure and conformation of full-size Rep is yet to be established, Rep is known to possess modular functions (14,15). The N-terminal part of Rep largely contains DNA-binding, nicking-ligation and oligomerization domains, while the C-terminal half contains ATP-binding and ATPase activity domains. Most of the helicases of bacterial and viral origin contain ATP-binding (G-Xaa-Xaa-Xaa-Xaa-GKT/S) and ATP hydrolysis motifs (DXXD), known as Walker A and Walker B motifs, respectively. However, the oligomerization status of these helicases seems to be diverse. The DnaB protein, the replicative helicase of *Escherichia coli*, is a member of the hexameric helicase family, which includes the T4 and T7 DNA helicases, plasmid RSF1010-encoded RepA and human MCM proteins (16). On the other hand, SV-40 large T-antigen of simian virus has been shown to be active only in the double hexameric form (17,18). In contrast, the PcrA from *Bacillus stearothermophilus* and the Rep-helicases from *E.coli* act in the monomeric and dimeric forms, respectively (19,20). The mechanism by which a helicase unwinds DNA duplex structure is strongly dependent on its oligomerization status and thus the understanding of the helicase activity requires a detailed study of the oligomerization states of the enzyme. Interestingly, the middle portion of Rep protein of all begomoviruses harbors a domain for oligomerization. As the MYMIV-Rep contains domains for ATP-binding, ATP hydrolysis and oligomerization, we examined if the C-terminal- or full Rep protein could possess any helicase activity.

Here we report efficient helicase activity of recombinant, full-length MYMIV-Rep (or AC1) protein. Various deletion mutants of Rep showed the minimal region required for such activity. The native molecular mass of Rep was determined by sucrose gradient fractionation and nondenaturing gel electrophoresis, and the helicase-active fraction was found to be oligomeric in nature. Such helicase activity was not confined to MYMIV-Rep alone, but was also exhibited by the Rep proteins of other begomoviruses. We also conducted *ex vivo* and *in planta* experiments to establish that the MYMIV-Rep protein indeed acted as a replicative helicase underscoring the important role played by Rep during the elongation phase of viral DNA replication. This is the first report of replicative helicase activity of a geminiviral Rep protein.

MATERIALS AND METHODS

Construction of various mutants of Rep, isolation and purification of proteins

Construction of His-Rep 1–362, His-Rep 1–182, His-Rep 183–362 and His-Rep 120–362 has been discussed elsewhere (11). Proteins were isolated from overproducing *E.coli* BL21 (DE3) cells and purified following the standard procedures. The induction process was carried out at 18°C for 16 h with 0.1 mM isopropyl- β -D-thiogalactopyranoside (IPTG). Induction under such condition resulted in accumulation of considerable amount of soluble Rep proteins. The soluble Rep was chromatographed through relevant affinity matrices as per manufacturer's protocol. The eluted proteins were further purified through Heparin–Sepharose CL6B and

Table 1. List of the sense oligonucleotides for construction of Rep mutants

Mutation	Sequence of sense primer
Rep G (226) A:	(5'→3') CGTGCCACATGGTTTACCGTGCAC
Rep K (227) E:	(5'→3') GCACGTGCCACATGGTTTACCCG-TGCG
Rep D (261) K:	(5'→3') CAAATAATGTGGATCAACGTCCTT-GATGACGTTGTACC
ODM1:	(5'→3') CAATTCATAACTTAAATTGTATATTCACAGACCTACACAG
ODM2:	(5'→3') GGCTCCTAAGGATTCATTTT-AGCATTCATAACTTAAATTGTAATTTG

Q-Sepharose 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 phenylmethylsulfonyl fluoride (PMSF) and benzamide, and 40% glycerol. The purified Rep protein was active for at least two weeks when stored at the concentration of about 750 ng/ μ l at –20°C. The Rep proteins were also purified with glutathione *S*-transferase (GST)-tags and were stored at concentrations similar to the His-tagged ones. 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.

Construction of Rep FF (75,76) SS mutation, and the isolation as well as purification of the protein has been described earlier (11). Mutations in the ATPase domains [Rep G (226) A, Rep K (227) E, Rep D (261) K] and in the oligomerization domain (ODM1 and ODM2) were introduced by using Quick-change site directed mutagenesis kit (Stratagene) following the manufacturer's protocol. PCR amplification was carried out using the sense primers presented in Table 1 and corresponding complementary oligonucleotide (anti-sense) primers (data not shown). The full-length wild-type Rep-pET28a DNA was used as template in each mutagenesis. The mutants were confirmed by manual sequencing.

ATP-binding analysis

ATP-binding analysis was carried out in a 10 μ l reaction mixture containing 12 mM HEPES (pH 7.9), 1 mM MgCl₂, 60 mM KCl, 1 mM DTT, 5 μ Ci of [α -³²P]ATP (3000 Ci/mmol, Perkin–Elmer Life Sciences, USA) and required amount of protein. The mixture was incubated on ice for 15 min followed by cross-linking by ultraviolet (UV) irradiation (254 nm) for 45 min on ice. The reaction was terminated by addition of 0.5 μ l of 100 mM non-labeled ATP. Proteins were separated on a 10% SDS–PAGE and the gel was stained with Coomassie blue. The radiolabeled bands were visualized by autoradiography.

ATPase assay

ATPase assay was carried out essentially as described earlier (11). Briefly, the desired amounts of proteins were incubated with 0.2 μ Ci of [γ -³²P]ATP (6000 Ci/ mmol, Perkin–Elmer Life Sciences, USA) in a buffer containing 20 mM Tris–HCl (pH 8.0), 1 mM MgCl₂, 100 mM KCl, 8 mM DTT and 80 μ g/ml BSA in a total reaction volume of 10 μ l at 37°C for 30 min. At the end of the reaction, 1 μ l of the reaction mixture was spotted on a polyethyleneimine thin layer

Table 2. List of oligonucleotides used for preparation of helicase substrates

Designation	Sequence (5'→3')
M13FWD17	GTTTTCCAGTCACGAC
M13FWD23	CCCAGTCACGACGTTGTAACG
HELPO1	GACTCTAGAGGATCCCCGGGTA- CCGAGCTCG
HELPO60	(T) ₈ CCTGCAGGTCGACTCTAGAGGATC- CCCCGGTACCGAGCTCG (T) ₁₁
HELFOK8	CCAAAACCCAGTCACGACGTTGT- AAAACG
HELFOK9	CCCAGTCACGACGTTGTAACG- TGCCGG
Helicase 0 T7	GTAATACGACTCACTATAGGGC
Helicase 0 T7F4	GCTG GTAATACGACTCACTATAGGGC
Helicase 0 bottom strand	GCCCTATAGTGAGTCGTATTACGT- CGTGACTGGGAAAAC
Helicase 3 bottom strand	GCCCTATAGTGAGTCGTATTACGG- TGTCGTGACTGGGAAAAC
Helicase 5 bottom strand	GCCCTATAGTGAGTCGTATTACGG- TGGGTCGTGACTGGGAAAAC
Helicase 6 bottom strand	GCCCTATAGTGAGTCGTATTACGG- TGTTGTCGTGACTGGGAAAAC
Helicase 9 bottom strand	GCCCTATAGTGAGTCGTATTACG- GTGGTGGTGTGCTGACTGGGAAAAC
Helicase 12 bottom strand	GCCCTATAGTGAGTCGTATTACGG- TGTTGTTGGTGTGCTGACTGGGAAAAC

chromatography (TLC) plate (Sigma–Aldrich, USA) and air-dried. Chromatography was performed using 0.5 M LiCl and 1 M HCOOH as the running solvent. Following completion of the chromatography, the TLC paper was air dried and autoradiographed.

Helicase assay

The list of oligonucleotides (Microsynth) used for preparation of various substrates is presented in Table 2. The oligonucleotides (~2 pmols) were ³²P-radiolabeled at the 5' end using T4 polynucleotide kinase and [γ-³²P]ATP (6000 Ci/mmol, Perkin–Elmer Life Sciences, USA) according to manufacturer's protocol. The radiolabeled oligonucleotides were annealed to M13mp18 ssDNA (5 pmol) or other specified oligonucleotides by heating the reaction mixture at 98°C and slowly cooling to room temperature. For preparation of non-forked substrates containing 17, 23 and 31mer and forked (both at 5' and 3' ends), 60mer oligonucleotides, the oligonucleotides designated as M13FWD17, M13FWD23, HELPO1 and HELPO60, respectively were used. Similarly, for substrates containing overhang of 6 nt at the 5' and 3' ends, oligonucleotides designated as HELFOK8 and HELFOK9, respectively were used. The substrates were incubated with the desired quantities of proteins in a buffer containing 20 mM Tris–HCl (pH 8.0), 1 mM MgCl₂, 100 mM KCl, 8 mM DTT, 5 mM ATP and 80 μg/ml BSA in a total reaction volume of 20 μl at 37°C for 30 min. Following completion, the reaction was terminated with 0.1% SDS and the products were separated on a 15% polyacrylamide gel. The gels were dried, autoradiographed and the results analyzed by densitometric scanning using Typhoon 9210 scanner and ImageQuant TL software (Amersham Biosciences, USA).

The partially double-stranded extension products, containing the oligonucleotides of sizes varying from 31 to >450 nt,

annealed to M13mp18 ssDNA (Figure 3B), were prepared by primer extension of the 31mer substrate using standard Sanger's dideoxy sequencing protocol with following modifications: the extension reaction was carried out on ice for <1 min and the reaction was terminated by addition of 2 μM ddGTP. The molar ratio of M13mp18 ssDNA to 31mer primer was kept at 5:1 during annealing. The extended product was used for helicase assay without any further purification. The substrate thus prepared contained a population of oligonucleotides of different size ranging from 31 to >450 nt.

The substrates used for nucleotide gap-length requirement study (Figure 3F–H) were as follows: for no gap, oligonucleotide designated as Helicase 0 bottom strand (Table 2) was annealed to M13FWD17 and Helicase0T7. Similarly, for creation of 3, 5, 6, 9 and 12 nt gaps in the substrates, M13FWD17 and Helicase0T7 oligonucleotides were annealed to oligonucleotides designated as Helicase3, Helicase5, Helicase6, Helicase9 and Helicase12 bottom strands (Table 2), respectively. In all these cases, Helicase0T7 was 5'-³²P-radiolabeled prior to the annealing step as described above. For similar experiments with forked structure at the top strand (designated as Helicase3F4, Helicase6F4), M13FWD17 and 5'-³²P-radiolabeled oligonucleotide designated as Helicase0T7F4 were annealed to the oligonucleotides Helicase3 and Helicase6, respectively. The substrates were purified through Bio-Gel Hydroxyapatite columns (BioRad, USA) according to manufacturer's protocol and used for helicase assays.

Sucrose gradient centrifugation

About 250 μg of each purified proteins was layered directly on a 10.5 ml of 10–40% (w/v) sucrose step gradient in a buffer containing 25 mM Tris–HCl (pH 8.0), 250 mM NaCl, 2 mM sodium bisulphite and 0.05% Triton X-100. Gradients were centrifuged in a Beckman SW41Ti rotor at 35 000 r.p.m. for 20 h at 4°C. Fractions (250 μl) were collected and subjected to 10% SDS–PAGE. The protein bands were visualized by silver staining. Protein molecular mass markers were run in parallel gradients. Standard curve was generated using the Microsoft Excel application program by plotting sedimentation distance as a function of molecular mass for various marker proteins, viz., chymotrypsin (25 kDa), ovalbumin (43 kDa), catalase (232 kDa), thyroglobulin (670 kDa) (Amersham Biosciences, USA) and IgM (980 kDa; a kind gift from Dr K. Rao, ICGB, New Delhi, India). Each fraction of 250 μl represented a sedimentation distance of 2.12 mm as an 11 ml solution filled up an axial length of 89 mm in the centrifuge tube. The sedimentation distance (*y* in mm) corresponding to a fraction '*f*' was represented by the equation $y = 67 + 2.12 \times 'f'$, where 67 = distance from the axis of rotation to the top of the centrifuge tube. Regression analysis using the Microsoft Excel application program yielded the equation: $y = 35.490 + 29.754 \times \text{Log}(x)$; $R^2 = 0.997$, where *y* represents the sedimentation distance (in mm) and *x* represents the molecular mass (in kDa). The sedimentation distances for Rep, Rep 120–362 and Rep C were fitted into the standard curve and their native molecular masses were estimated.

Nondenaturing PAGE

Proteins were separated on a native 7.5% polyacrylamide gel. Electrophoresis was carried out using Tris-Glycine buffer [25 mM Tris and 250 mM glycine (pH 8.3)] followed by staining with Coomassie blue. Standard curve was generated using the Microsoft Excel application program by plotting molecular mass as a function of migration distance for the marker proteins, ferritin (440 kDa) and thyroglobulin (670 kDa) (Amersham Biosciences, USA). Regression analysis using the Microsoft Excel application program yielded the equation: $y = 824.54 - 1193.4 \times \text{Log}(x)$; $R^2 = 1.000$, where y represents the molecular mass (in kDa) and x represent the migration distance (in cm). The migration distance for Rep was fitted into the standard curve and its native molecular mass was estimated.

Replication assay of the plasmid harboring CR-AC3 segment from the DNA-A component of MYMIV in *Saccharomyces cerevisiae* (ex vivo analysis)

The strain of *S.cerevisiae* (W303a) and various constructs viz., YCPO⁻-2A, YCPO⁻ (CR-AC3) F, YCPO⁻ (CR-AC3) R and YCPO⁻ 2 (CR-AC3) F, used for the assay, have been described elsewhere (V. Raghavan unpublished PhD Thesis). The notations F, R and 2 represented the forward, reverse and dimer constructs of CR-AC3 segment, respectively. We carried out the replication assay essentially by using the yeast transformation protocol as described (3). Briefly, equimolar quantities of plasmids containing various constructs were mixed separately with competent yeast cells in transformation mix, vortexed and incubated at 30°C for 30 min. The cells were subjected to heat shock at 42°C for 20 min and harvested at 8000 r.p.m. for 15 s in a microfuge tube. The supernatant was discarded and the cell pellet was resuspended in 1 ml of sterile water. Aliquots of 200 µl of resuspended cells were plated on suitable (Ura⁻) selection media. The plates were incubated at 30°C for 2–4 days to recover transformants. The number of colonies was scored, which essentially served as an index of plasmid replication. The relative replication efficiencies of various constructs were calculated taking the replication efficiency (measured as the number of transformants) of YCPO⁻-CR-AC3F construct to be 100%. Each experiment was conducted in triplicate and the average value was taken. Similar experiments were carried out with constructs carrying CR-AC3 (G226A) and CR-AC3 (K227E) mutations.

In planta assay

In order to study MYMIV-Rep (AC1)-dependent transient replication, we have constructed a binary vector (VA/pCAMBIA1391Z) based on the minimal replicon (CR-AC3) of the MYMIV DNA-A. As shown in the Figure 6C–F, the construct includes a reporter GFP flanked by two copies of CR from MYMIV DNA-A. The term VA essentially signifies virus-based amplicon. When the tobacco leaves are agroinfiltrated with the agrobacteria harboring the binary construct, a circular episome of 2.4 kb is released following Rep-dependent site-specific nicking action at CR and subsequent DNA replication. The vector was constructed in such a manner that the episome is supposed to contain one copy of CR,

the region of DNA-A spanning AC1–AC3 and the reporter green fluorescent protein (GFP) gene along with the promoter. The episome formation was found to be associated with GFP production at the site of agro-infiltration. The episome formation was also detectable by Southern blot of total DNA of the infiltrated leaf-region using GFP as probe. The isolated total DNA was digested with ScaI enzyme before resolving in agarose gel and the Southern band was detected at 3.1 kb. The unreplicated binary vector gave rise to 4.8 kb DNA band under similar circumstances because of the location of ScaI sites.

The mutations in various motifs associated with the helicase function of Rep (AC1) protein were created by PCR mutagenesis in the replicon template following manufacturer's protocol producing CR-AC3 (K227E) or CR-AC3 (G226A) mutants. Using these Rep mutated replicons the episomal vector was reconstituted as VA RepK227E/pCAMBIA1391 and VA Rep G226A/pCAMBIA1391. Such binary vectors were used to transform the LBA 4404 agrobacterium strain and the transformants were infiltrated in tobacco leaves. The GFP expression in the infiltrated zones was observed under UV-light at 10 days post infiltration. The genomic DNA was also isolated for analysis of released episomes by Southern hybridization using the GFP specific probe.

RESULTS

ATP-binding, ATPase activities of Rep and its deletion mutants

The C-terminal part of MYMIV-Rep contains the Walker A/B motifs (Figure 1A) that are supposedly responsible for ATP-binding and hydrolysis. To examine the functionality of these motifs, we constructed the Rep 1–182 (Rep N), Rep 120–362 and Rep 183–362 (Rep C) truncation mutants and over expressed the proteins as His6X- (Rep 120–362, Rep 183–362 [Rep C] and Rep 1–182 [Rep N]), recombinant proteins. Rep protein (1 µg) was incubated with 5 µCi of [α -³²P]ATP (3000 Ci/mmol) and cross-linked with UV (254 nm). The autoradiogram (Figure 1C, lane 2) reflects the strong binding of Rep to ATP. About 1.5 µg of *E.coli* DNA polymerase I (lane 1) was used as a positive control. As expected, there was no significant ATP-binding activity with 1 µg of His6X-tagged Rep N (lane 3). The results demonstrated the presence of functional either ATP-binding domain in the C-terminal part of Rep. All the Rep fragments as well as the full-length Rep were expressed as His- or GST- tagged recombinant proteins.

Next, we investigated the ATPase activities of Rep and its various mutant forms. The autoradiogram presented in Figure 1D shows that the Rep, Rep 120–362 and Rep C showed strong ATPase activities at various concentrations of the proteins while Rep N lacked any such activity. The fact that the ATPase activities of Rep, Rep 120–362 and Rep C were comparable suggests that the ATPase activities are located in the C-terminal domain of Rep. Furthermore, the ATPase activity of Rep increased in a dose dependent manner (Figure 1E), and the activity increased several-fold in presence of ssDNA (Figure 1F and G).

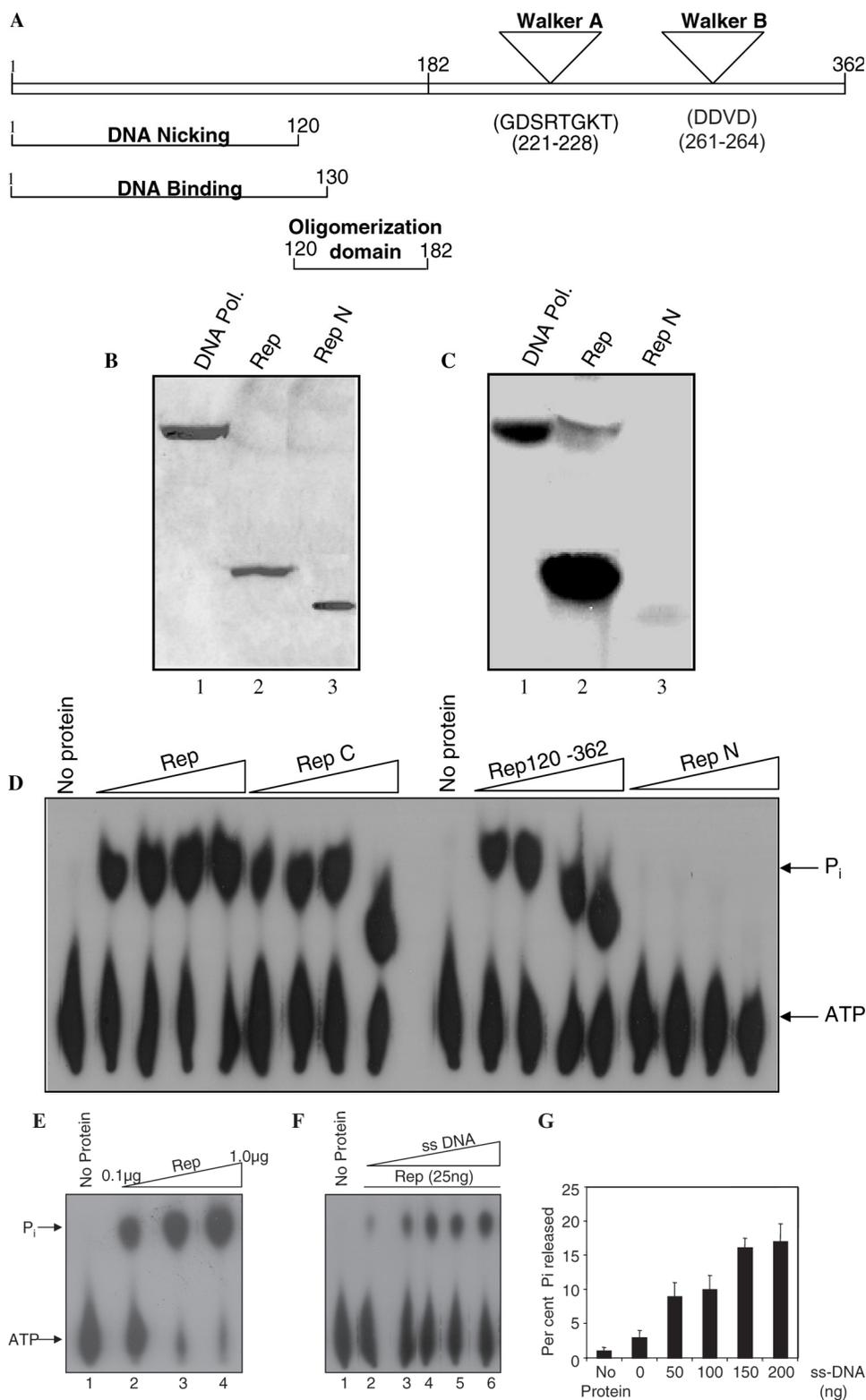


Figure 1. ATP-binding and ATPase activities of Rep. (A) Rectangular box represents full-length Rep (AC1) gene. The locations of various domains and the amino acid positions are indicated. (B and C) Rep binds to ATP. Purified Rep and Rep 1–182 (Rep N) proteins were cross-linked with [α - 32 P]ATP for 30 min and electrophoresed on 10% SDS-PAGE. The gel was stained with Coomassie blue (B) and autoradiographed (C). *E. coli* DNA polymerase I was taken as a positive control. The respective proteins are marked. (D) Autoradiograph showing the ATPase activity of Rep and its deletion derivatives. Respective proteins are marked on top of the lanes. Positions of ATP and released P_i are indicated. The amounts of the proteins used were as follows: Rep: 0.5, 1.0, 2.0, 3.0 μ g; Rep C: 0.1, 0.3, 0.6, 3.0 μ g; Rep 120–362: 0.4, 0.8, 1.5, 3.0 μ g; Rep N: 2.0, 4.0, 6.0, 8.0 μ g. (E–G) Autoradiograph showing the ATPase activity at different concentrations of Rep (E) and the enhancement of ATPase activity of Rep protein (25 ng) in presence of increasing amounts of ssDNA (F). The quantified values (fraction percent) of released P_i in (F) are presented in (G).

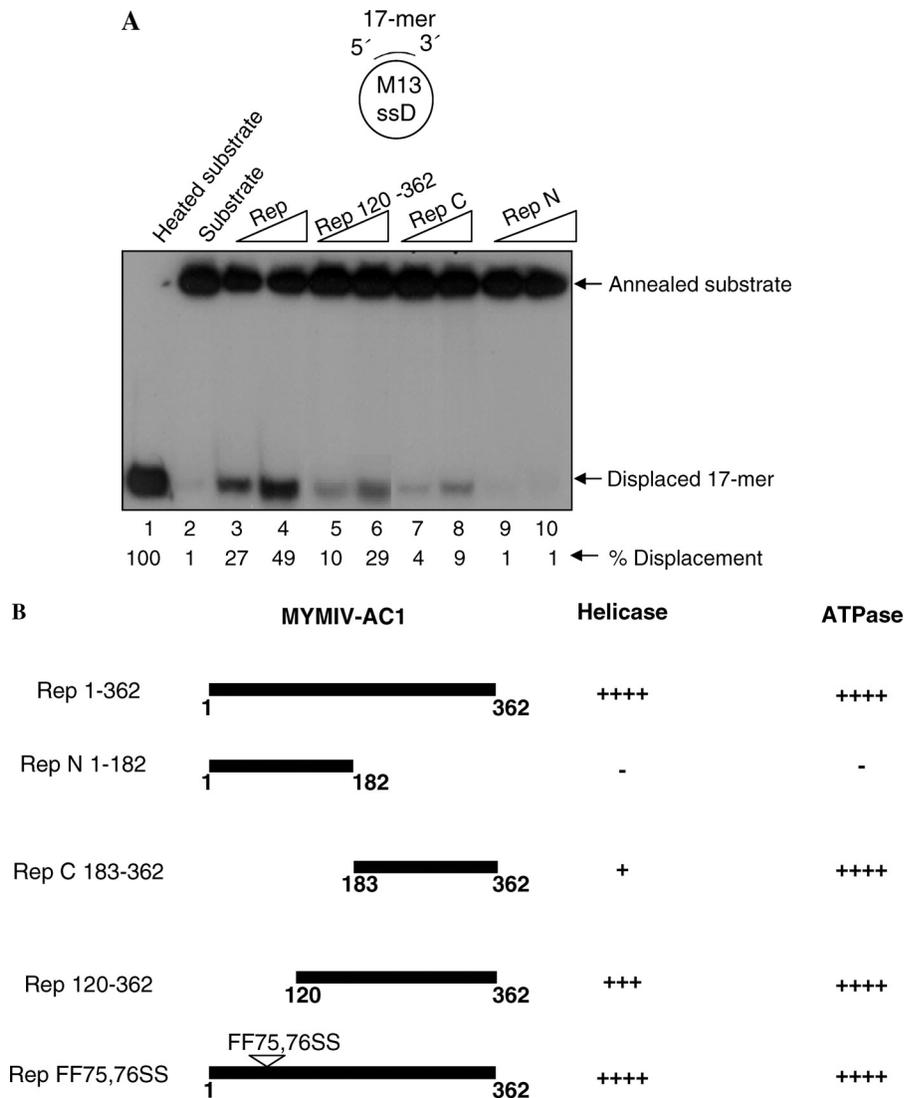


Figure 2. Helicase activity of Rep. (A) Autoradiograph showing the helicase activities of Rep and its derivatives with a 17mer substrate. About 0.5 and 1.0 μg of Rep (lanes 3 and 4), 0.5 and 1.0 μg of Rep 120–362 (lanes 5 and 6), 1.0 and 2.0 μg of Rep C (lanes 7 and 8), 4.0 and 8.0 μg of Rep N (lanes 9 and 10) were used, respectively. The relative unwinding (displacement) values are presented. (B) Summary of ATPase and helicase activities data of Rep and its other mutants as indicated (++++, high; +++, moderately high; +, low; –, negligible).

Helicase activities of Rep and its derivatives

Since enhancement of the ATP hydrolysis in presence of ssDNA is indicative of the helicase activity, we examined the helicase activities of Rep protein and its derivatives using the standard strand displacement assay (21). The helicase substrate containing a 17mer oligonucleotide (M13FWD17), annealed to M13mp18 ssDNA, was incubated with desired amounts of proteins at 37°C in presence of 1 mM MgCl_2 and 5 mM ATP. Figure 2A shows the presence of helicase activities of Rep and its other mutants. Both Rep and Rep 120–362 displaced oligonucleotide from the substrate with high efficiency and in a dose dependent manner, while Rep C did not reveal any appreciable helicase activity. Rep N, on the other hand, did not show any helicase activity at all. The lack of helicase activity of Rep N is not surprising in view of the absence of ATP-binding and ATPase domains in it. The results of ATPase and helicase activities of various

Rep-derivatives are summarized in Figure 2B. A comparison of the helicase activities between Rep 120–362 and Rep C immediately suggests that the oligomerization domain (120–183 amino acid) could be the crucial factor for helicase activity. We have addressed this possibility in details in later sections. Two point mutations in Rep at 75 and 76 residues [Rep FF (75,76) SS] were introduced to check if the mutations at the putative PCNA-binding site could also alter the Rep's ATPase and helicase activities. However, Rep's ATPase and helicase activities remained unchanged with such mutations (Figure 2B).

Characterization of the helicase activity

- (i) Fork-structure requirement: Supplementary Figure S1A reveals that 1.5 μg Rep was able to displace ~87% of a 17mer oligonucleotide (lane 6) while similar amount of

protein displaced the 23mer oligonucleotide (designated as M13FWD23) to a lesser degree (~50%, Figure 3A, lane 4). Figure 3A also reflects that the helicase activity remained mostly unaltered when a substrate with 23mer duplex and 6mer 5'-forked region (designated as HELFOK8) was used. Since similar amount of Rep (0.75 μ g) was used in lanes 3 and 7, it appears that the 5'-forked region improved the helicase activity to some extent. Similar results were obtained when a 6mer 3'-forked region (designated as HELFOK9) containing substrate was used (data not shown). This suggests that the forked structure of the substrate is not a requirement for Rep's helicase activity, but could act as a facilitator.

- (ii) Power of displacement: Supplementary Figure S1B reveals that Rep unwound a duplex region of about 30mer with or without the forked structure. Supplementary Figure S1C shows that the efficiency of displacement decreased with the increased length of the displacing oligonucleotide. In order to gain insight about the processivity, we varied the length of oligonucleotide from 17 to ~300 nt as described in the Materials and Methods section. In Figure 3B, the double-stranded extension products were used. By visual inspection, the fraction of substrate populations with 31mer duplex was found to be higher than that with >300mer duplex which, in turn, was also considerable (lane 1). The lanes 3–5 (Figure 3B) reveal that Rep was capable of displacing oligonucleotides of at least 200 nt. These results point to the possible role of Rep as a replicative helicase.
- (iii) 3'→5' translocation: To determine the directionality of Rep-helicase, the employed experimental strategy is schematically shown in Figure 3C. We prepared a substrate by annealing an unlabeled 31mer oligonucleotide having an internal SmaI site to M13mp18 ssDNA. The resulting substrate was treated with *Pfu* polymerase enzyme along with dATP, dTTP and 50 μ Ci of [α -³²P]dCTP for 3' end radiolabeling. Since the dNTPs supply lacked dGTP, the extension of synthesis preceded only up to 5 nt till the polymerase reached the nucleotide C of the template. The extended product was then digested with SmaI generating a 3'-labeled substrate with 19mer duplex. In a parallel experiment, the 5'-radiolabeled 31mer was annealed to M13mp18 ssDNA; 3' end extended with unlabeled dNTPs lacking dGTP and finally restricted with SmaI. The digested product was used as a substrate for helicase assay. The results of the helicase assay are presented in Figure 3D and E, respectively. Rep was able to displace oligonucleotide only from the 5'-labeled substrate, thus establishing the direction of translocation in the 3'→5' direction. The displacement activity of the 2.0 μ g Rep with the 5'-labeled 31mer is shown in lane 8 of Figure 3E as a control.
- (iv) Inhibitors: The Rep-helicase activity with the 17mer containing substrate was strongly inhibited in presence of either ADP (2.0 mM) or nogalamycin (5.0 μ M), a known helicase specific inhibitor (Supplementary Figure S2A and S2B). It was interesting to note that *E.coli* SSB, too, inhibited Rep's helicase activity (Supplementary

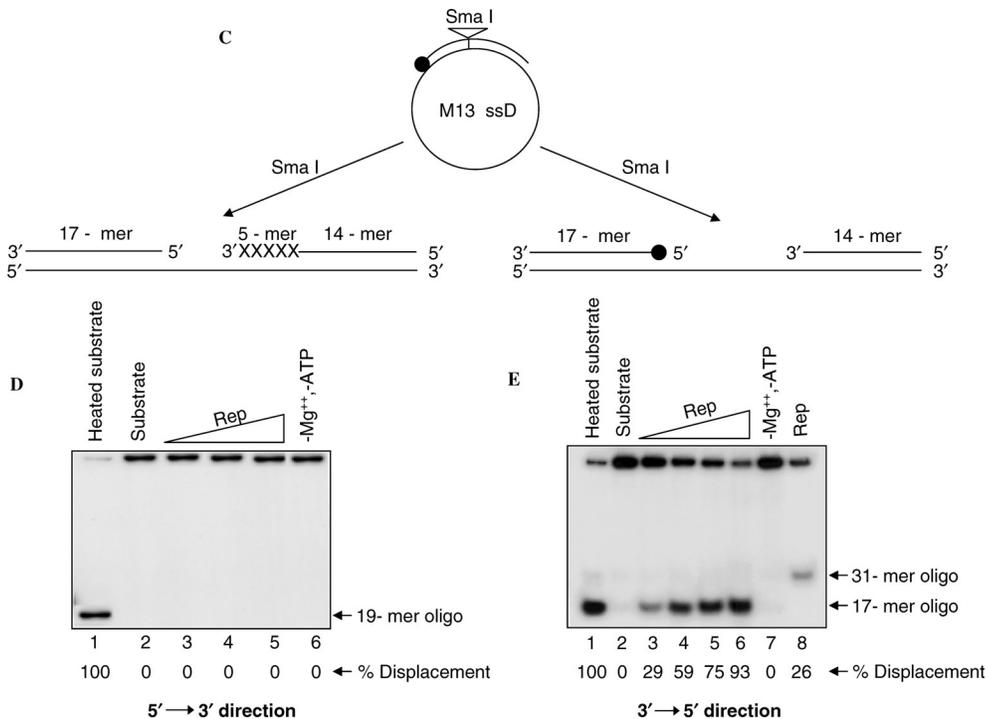
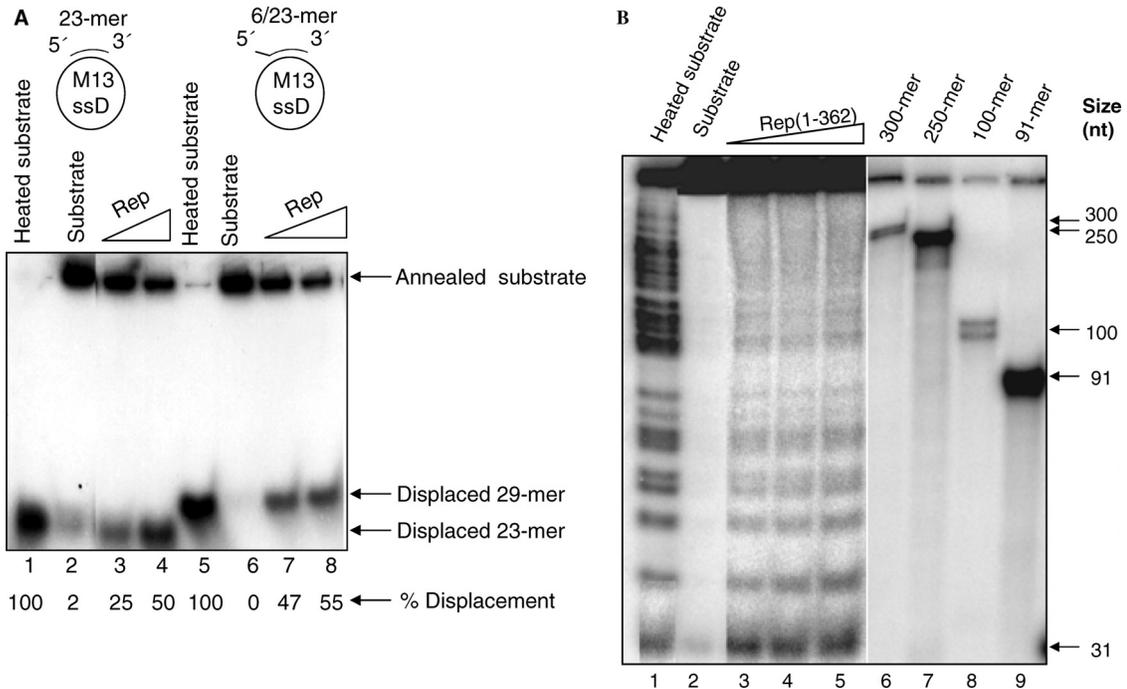
Figure S2C) indicating thereby that MYMIV-Rep acted mostly as an eukaryotic helicase and not as a prokaryotic one (*E.coli* SSB does not inhibit the *E.coli* helicase). The inhibition was similar when SSB was present either in the pre-bound form with the substrate (data not shown) or along with Rep during the helicase reaction (Supplementary Figure S2C).

- (v) Space requirement: During the biosynthesis of viral DNA, Rep makes a site-specific nick at the viral strand of the replicative form (RF) to initiate RCR. Hence, we wanted to explore whether Rep could use this nick alone or would require a gap of few nucleotides in order to carry out the helicase activity properly. To determine the minimum gap-length required for helicase activity post-nicking, we designed various substrates, four of which are depicted in Figure 3F. The substrates have been designed to mimic the structure of the RF of DNA at the initiation phase of replication. In each case, three oligonucleotides (designated as Oligo1, Oligo2 and Oligo3, respectively in the top part of Figure 3F) were combined to generate the helicase substrate containing the desired lengths of gap that ranged from zero to 12 nt. The Oligo2 was 5'-radiolabeled and its displacement was monitored. For example, the Helicase 3 and Helicase 6 substrates in Figure 3G and H contained gaps of 3 and 6 nt, respectively. Similar substrates were also prepared containing 4 nt overhang at the 5' end of one of the combining oligonucleotide (i.e. in Oligo2) to follow the enhancement of helicase activity, if any (e.g. Helicase 3F4 and Helicase 6F4 in Figure 3F). The results of helicase activity of Rep with two such representative substrates are shown in Figure 3G and H, and the same for all the substrates are summarized in Figure 3I. A minimum length of 6 nt gap was essential for Rep's efficient helicase activity. However, the 5'-flap structure of Oligo2 (or fork-like structure) did not improve the helicase activity substantially (data not shown), which corroborated with the earlier findings (Figure 3A and Supplementary Figure S1B and S1C). These data shed light on some basic requirements for Rep to act as a helicase. However, a detailed knowledge about Rep's architecture in the replication complex is needed to understand the full significance of the above data.
- (vi) Loss of function mutants of Rep: To correlate the ATPase and helicase activities of Rep, we constructed three different Rep mutants with altered Walker A and Walker B motifs separately, and compared their ATPase versus helicase activities. In the first two mutants, conserved G residue at 226 amino acid position and conserved K residue at 227 amino acid position were mutated to A and E residues, respectively, while in the third mutant, highly conserved D residue at 261 amino acid position was mutated to K residue. The mutant proteins are shown in Supplementary Figure S3A and were sufficiently pure. The results of the ATPase and helicase assays with the purified mutant proteins are presented in Figure 4A–C and Supplementary Figure S3B, respectively. The ATPase activity of Rep G (226) A protein was reduced by >80% while for other mutants the activity reduced by >90% compared to that

of the wild-type. All the mutants lost helicase activities to near completion (Supplementary Figure S3B). For example, while the wild-type Rep protein unwound the substrate to the extent of ~30%, none of the mutant proteins could unwind the same substrate by >3%. These data clearly suggest that the helicase activity of Rep was tightly coupled to the ATPase activity. Similar inhibition in the helicase activity has also been reported with various MCM mutants (22).

Oligomerization status of Rep and its mutants

Since Figure 2B reveals that the helicase activity of Rep was dependent on oligomerization domain, Rep might need a particular oligomeric conformation for helicase activity. The studies in Rep-oligomerization might also provide insight into the mechanism of the helicase activity. Our earlier data (3,11,23) clearly demonstrated that the Rep protein is capable of oligomerization both *in vitro* and *ex vivo* conditions. Here we estimated the native size of



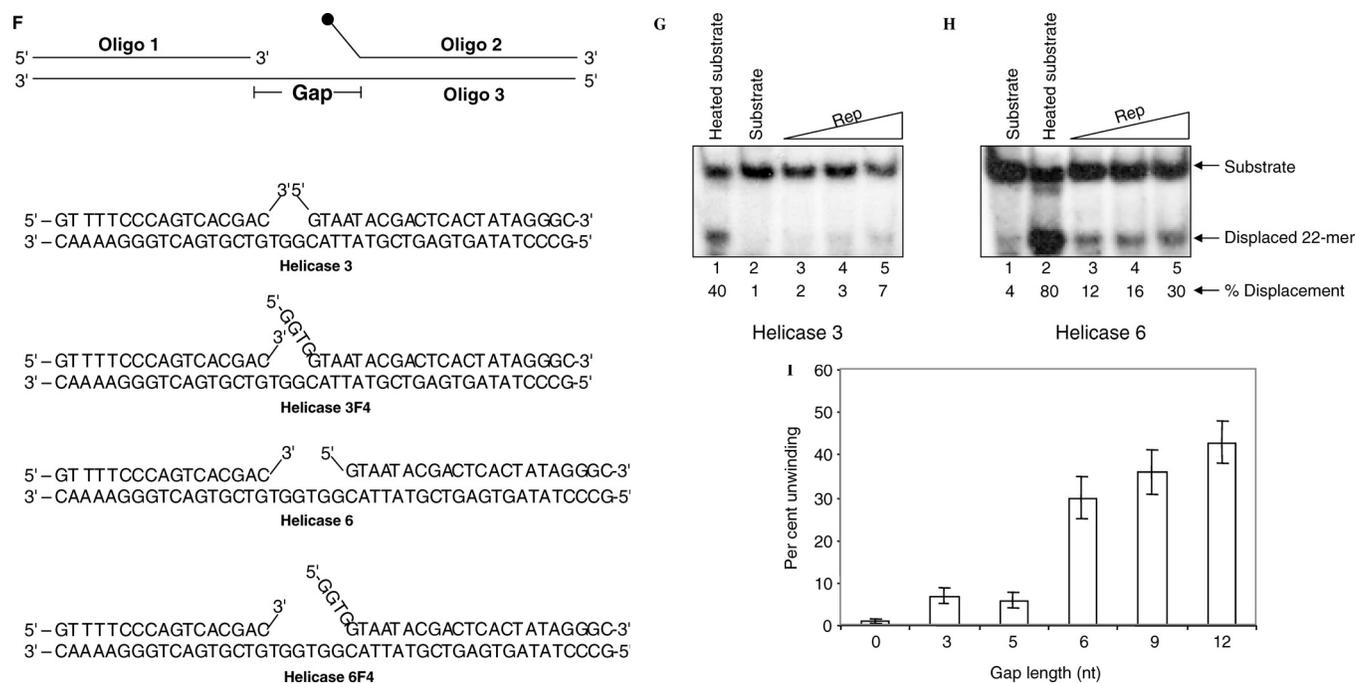


Figure 3. Characteristics of helicase activity of Rep. (A and B) Autoradiograph showing helicase activity of Rep using various substrates. (A) The helicase activity at increasing concentrations of Rep with 23mer unforked substrate (lanes 1–4), and with 29mer substrate containing 6 nt 5'-overhang (lanes 5–8), respectively. Respective proteins (1.0 μ g each for lanes 3, 7, and 2.0 μ g each for lanes 4 and 8) are marked on top of the lanes. The relative unwinding (displacement) values are presented. (B) Autoradiograph showing the helicase activity at increasing concentrations of Rep (0.5, 1.0 and 1.5 μ g in lanes 3, 4 and 5, respectively) using the substrate containing the partially double-stranded extension products (31 to >300 nt long oligonucleotides annealed to M13mp18 ssDNA). Lane 1 denotes the boiled substrate, showing the presence of 31 to >300 nt long oligonucleotides that are released on boiling the substrate. Lane 2 denotes the native substrate. In lanes 6–9, different radiolabeled oligonucleotides of 300, 250, 100 and 91 nt length, respectively were loaded as size markers. (C) Diagram showing the design for determination of polarity of Rep-helicase. X and \bullet denote 3' and 5' labels, respectively. Products shown at the left and the right panels following SmaI digestions were used as the substrates for (D and E), respectively. (D) Autoradiograph showing the helicase activity at increasing concentrations of Rep using the substrate with unlabeled 5' end and labeled 3' end. Absence of any helicase activity in this case indicates that Rep did not translocate in the 5'→3' direction. (E) Autoradiograph showing the helicase activity at increasing concentrations of Rep (1.0, 2.0, 2.5 and 3.0 μ g for lanes 3–6, respectively) using the substrate with 5'-radiolabeled 31mer. Lane 8 represents the control using the same substrate as used in Supplementary Figure S1B. High helicase activity suggested the polarity of Rep to be in the 3'→5' direction. (F) Representative substrates for determination of space requirement for Rep. Various oligonucleotides designed so as to create 3 nt gap (Helicase 3), 3 nt gap containing 4 nt 5'-overhang (Helicase 3F4), 6 nt gap (Helicase 6), 6 nt gap containing 4 nt 5'-overhang (Helicase 6F4), are shown. Other substrates used for the experiment have been presented in the Materials and Methods section. At the top row, the three oligonucleotides constituting each substrate have been schematized as Oligo1, Oligo2 and Oligo3, respectively. The 5' end of Oligo2 was radiolabeled (black dot) in all the cases. (G and H) Autoradiographs demonstrating the helicase activity with 3 nt-gapped (G) and 6 nt-gapped (H) substrates, denoted as Helicase 3 and Helicase 6, respectively. The amounts of Rep used were 1.0, 1.5 and 2.0 μ g (for lanes 3–5 of both the panels). Rep could displace oligonucleotide from Helicase 6 substrate, but not from Helicase 3. (I) Column graph representing the helicase activity of Rep (percent unwinding) with substrates containing 0, 3, 5, 6, 9 and 12 nt gaps. The graph reflects that the minimum space requirement of loading of Rep is 6 nt.

MYMIV-Rep and its various derivatives by sucrose gradient centrifugation (velocity sedimentation) technique. Besides the standard markers, namely, chymotrypsin (25 kDa), ovalbumin (43 kDa), catalase (232 kDa) and thyroglobulin (670 kDa) (all procured from Amersham Biosciences, USA), another pentameric protein, IgM (a kind gift from Dr K. Rao, ICGEB, New Delhi, India) of 980 kDa, was included in the sedimentation experiment (Figure 5A and B). The molecular masses of Rep, Rep 120–362 and Rep C were estimated to be ~1020, 740 and 48 kDa, respectively (Figure 5A and B, and Supplementary Figure S4B). A comparison of these data with those obtained by SDS-PAGE analysis reflects that both Rep and Rep 120–362 most probably existed as 24mer, while the Rep C was a dimer.

To investigate if the helicase activity of the protein was related to its oligomerization, we carried out helicase assay with each of the sedimentation fractions. In case of both Rep and Rep 120–362, the fractions rich in the oligomeric

structure (within the limits of estimation) exhibited the maximal helicase activity (Figure 5A and B). All other fractions showed much reduced or negligible helicase activity. Rep C displayed very weak helicase activity (Figure 5A), which corroborated with the data mentioned earlier (Figure 2B). Thus, the helicase activity of Rep or Rep 120–362 was dependent on its oligomerization status.

Similar experiment was carried out with a sample of Rep protein that was rendered inactive by storage at 4°C for 60 days. This sample lost significantly only its helicase activity but retained the nicking-closing activity (data not shown). The velocity sedimentation data shown in Supplementary Figure S4A reveal that the sample contained a wide distribution of population of varying molecular sizes, lacking any well-defined peak. Interestingly only the few fractions of high oligomeric states were capable of displaying the helicase activity. This observation strengthened the notion that the helicase activity was owing to a particular conformation of Rep.

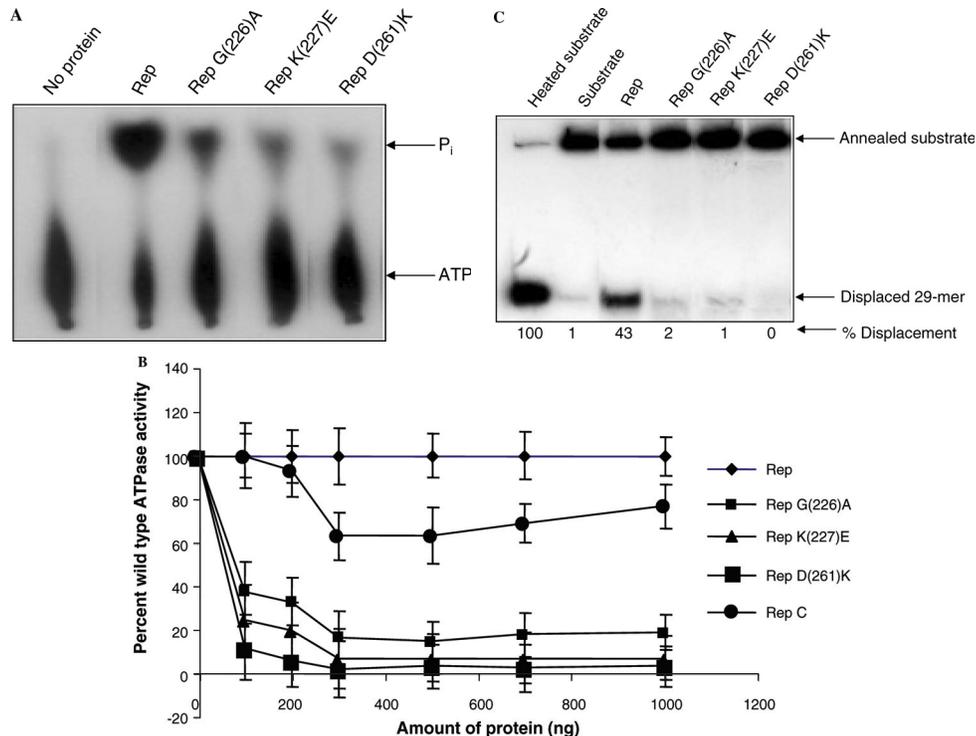


Figure 4. ATPase and Helicase activities of Rep mutated in Walker A and Walker B motifs. (A) Autoradiograph showing the ATPase activities of 0.5 μ g Rep, 1.0 μ g Rep G (226) A, 1.0 μ g Rep K (227) E, 1.0 μ g Rep D (261) K and (B) plot of the relative ATPase activities of the proteins at different concentrations. The ATPase activity of wild-type Rep was taken as 100% and those of the mutants were expressed as percent wild-type activity. Data for Rep C is included in the plot for comparison. (C) Autoradiograph showing the helicase activities of the indicated proteins (at concentrations as used in (A)), using the oligonucleotide HELFOK9 (29mer) annealed to M13mp18 ssDNA as substrate.

To estimate the native molecular mass of Rep by another independent technique, Rep was electrophoresed in a non-denaturing 7.5% polyacrylamide gel where thyroglobulin (670 kDa) and ferritin (440 kDa) were used as molecular mass markers (Supplementary Figure S4C). The native molecular mass of Rep was estimated by regression analysis and was found to be \sim 1135 kDa (Supplementary Figure S4D), which is in agreement with that found from the sucrose gradient centrifugation studies.

Two different mutations in Rep's oligomerization domain were introduced to study further the consequences of loss of oligomerization. In one mutation (ODM1), the 168 NDLS 171 amino acid residues (Supplementary Figure S4E) were deleted. These residues are highly conserved among geminiviruses and are present in the oligomerization domain of SV-40 large T-antigen. In the second mutation (ODM2), glutamine (Q) residue at position 161, which is also highly conserved amongst geminiviruses, was changed to alanine (A). The analyses with the purified mutant proteins (profile in Supplementary Figure S4F) revealed that, while Q161A mutation caused \sim 40% reduction in the helicase activity, Δ NDLS 168–171 mutation resulted in severe loss in the helicase function (Figure 5C). The sucrose gradient sedimentation revealed that the ODM1 mutant protein was split in populations having various different oligomeric states while the ODM2 mutant protein was similar to the wild-type Rep (Figure 5D). However, the oligomeric states of ODM2 were not exactly identical to those of wild-type Rep. The loss in helicase activity of ODM1 protein can thus be explained in terms of the loss in its oligomerization property, lending

further support to our view that attainment of proper oligomeric conformation is a prerequisite to the Rep-mediated helicase activity.

Rep-helicase mutants are defective in DNA replication *ex vivo* and *in planta*

The various *in vitro* data presented above, namely, 3' \rightarrow 5' translocation, high processivity etc. point to a possible role for Rep (AC1) as a replicative helicase. To investigate if Rep functions truly as a replicative helicase *in vivo*, the following experiments were carried out. Earlier, we had established that an ARS-less plasmid bearing two tandem copies of DNA-A of the geminivirus MYMIV could replicate well in *S.cerevisiae*, based on the colony forming abilities of various constructs in the selective (Ura drop out) medium (3). Since the same ARS less vector bearing the sub genomic fragment CR-AC3 of DNA-A retained the ability to transform yeast, the CR-AC3 region seemed to be the minimal replicon of DNA-A of MYMIV (Table 3). Any truncation in ORF-AC3 would lead to a serious defect in the colony forming ability of the resultant plasmid (data not shown). The sub genomic construct containing CR-AC3 in either orientations showed about 50% relative growth efficiency compared to the YCPO⁻2A construct, while the dimer [2(CR-AC3)] construct showed \sim 60% relative growth efficiency. Since the CR-AC3 region functions as the minimal fragment essential for replication, various mutations related to the helicase activity were introduced in Rep to test their roles in replication *in vivo*. The two mutations, namely

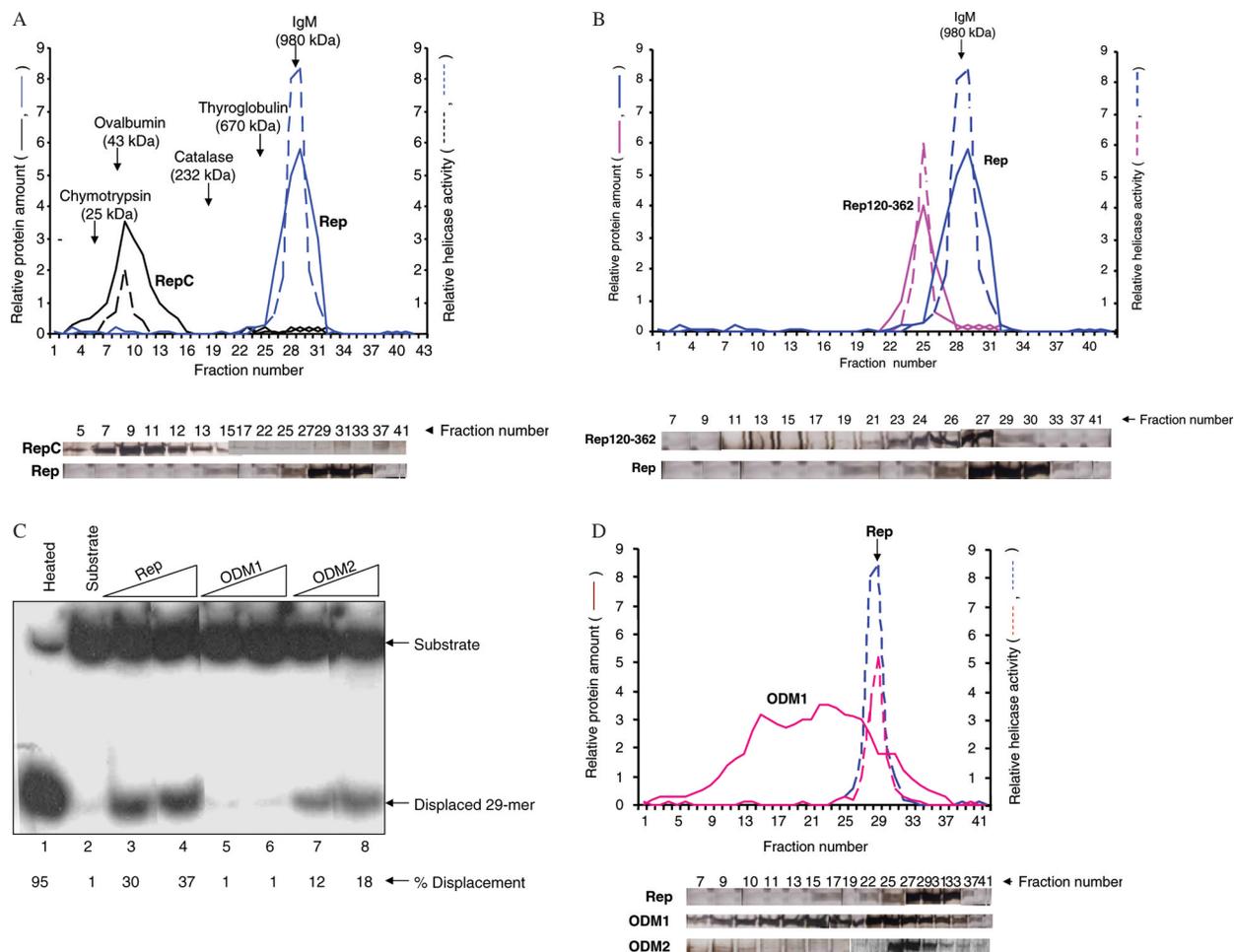


Figure 5. Estimation of native molecular masses of Rep and its mutants by sucrose gradient centrifugation analysis. (A and B) Relative protein content (solid line) and helicase activity (dashed line) of individual fractions is plotted against fraction number for Rep and Rep C (A) and for Rep120–362 (B). The vertically down arrows indicate the positions of the peaks for the markers. The silver stained gel pictures of each fraction, resolved on a 10% SDS–PAGE, are shown at the lower panels of each Figure. (C and D) Helicase activities and estimated sizes of the oligomerization mutants of Rep. The helicase activities of the wild-type as well as mutant proteins using HELFOK9 (29mer) substrate are shown in (C). About 1.5, 2.0 μ g of Rep; 2.0, 3.0 μ g of ODM1 and 1.0, 2.0 μ g of ODM2 were used in lanes 3–8, respectively. The results of sucrose gradient fractionation and helicase activities of each fraction for ODM1 are given in (D). The corresponding helicase activities for Rep are included in the figure for comparison. The lower panel shows the silver stained gel of various fractions of Rep and the mutants of oligomerization domains (ODM1 and ODM2). The helicase activities for ODM2 were similar to those obtained for Rep and are not included in the plot.

G226A and K227E of Rep, which affected the helicase activities of Rep (Figure 4C), were placed in the CR-AC3 background. The mutant constructs were used to transform yeast, and their replication efficiencies compared to that of the wild-type CR-AC3 were calculated. Though both CR-AC3 (G226A) and CR-AC3 (K227E) mutants showed drastic reduction in replication efficiency, the CR-AC3 (K227E) mutation reduced the replication efficiency to almost zero level (Figure 6B, Table 3). Since these mutations affect only the ATPase and helicase activities of Rep, it may be safe to correlate the reduction of replication efficiency with the loss in helicase activity. This clearly suggests that Rep indeed acted as a replicative helicase, and it is noteworthy that no other host helicase could substitute the activity of Rep during viral DNA replication in yeast.

To further confirm the role of Rep as a replicative helicase *in planta*, tobacco leaves were agroinfiltrated with wild-type (wt) and mutated viral amplicon (VA) vectors [VA/Rep G (226) A or Rep K (227) E] carrying GFP as a tractable

marker gene (Figure 6D). When the wt VA vector was agroinfiltrated in tobacco, a 3.2 kb episomal circular DNA was released from the vector following RCR as schematized in Figure 6C. The presence of this replicated DNA was detected following restriction of the isolated genomic DNA post agro-inoculation with the *ScaI* enzyme and Southern analysis using the GFP probe. The unreplicated and the replicated DNA would reveal as the 4.8 and 3.2 kb bands, respectively in a typical Southern analysis (N. I. Mohammad *et al.*, unpublished data). *DpnI* digestion of the genomic DNA was also carried out to distinguish between the replicated and unreplicated DNA (data not shown). We argued that if replication (and consequent formation of 3.2 kb episomal DNA) were dependent on Rep-helicase, the helicase-motif mutants would render the replication inefficient and cause lower accumulation of GFP compared to the wild-type VA vector. At 10 days post infiltration, GFP expression in the inoculated leaf samples was greatly reduced with mutated vectors (Figure 6E). The reduced level of GFP expression

Table 3. Survival of the helicase mutants

Construct		Cfu(Ura)/pmol	Percent relative growth
A A →→	YCPO ⁻ -2A	15 355 ± 450	181.8
CR-AC3 →	YCPO ⁻ -CR-AC3F	8445 ± 265	100
CR-AC3 ←	YCPO ⁻ -CR-AC3R	8000 ± 275	94.7
2(CR-AC3) →→	YCPO ⁻ -2(CR-AC3)F	10900 ± 150	129.1
CR-AC3 (G226A) →	YCPO ⁻ -CR-AC3F(G226A)	380 ± 20	4.5
CR-AC3 (K227E) →	YCPO ⁻ -CR-AC3F(K227E)	17 ± 15	0.2
Vector	YCPO ⁻	No growth	0

The results of colony forming experiments with the constructs carrying CR-AC3 region or the mutants, *viz.*, CR-AC3 (RepG226A) and CR-AC3 (Rep K227E) are presented in the form of relative numerical values. The values presented are the average of six independent readings. The standard deviations are indicated.

was also corroborated by the loss of the 3.2 kb band in the Southern blot (Figure 6F, lanes 2–4). Here also, the M1 mutation (K227E) turned out to be more severe than the M2 mutation (G226A). The level of actin signal in the blot reprobed with the actin gene showed the uniformity in loading control (bottom panel of Figure 6F). Both of these biological assays very strongly point to the role played by Rep as a replicative helicase.

Rep-helicase activity of other begomoviruses

We wondered whether the observed helicase activity is peculiar to MYMIV-Rep alone or a common feature of Rep of all geminiviruses. Here we have examined only two geminiviruses that belong to the same genus (begomovirus) of MYMIV. Rep of one of these, *viz.*, MYMV (Mung bean yellow mosaic virus, prevalent in Southern India and in Thailand), is ~80% similar to MYMIV-Rep, while the Rep of the other, *i.e.* Indian cassava mosaic virus (ICMV) is phylogenetically different from MYMIV-Rep. The degree of purity of Rep proteins isolated from these three sources is shown in Figure 7A. Figure 7B–D show that all these three proteins displayed comparable level of helicase activities. Hence it appears that the Rep proteins of most of the geminiviruses might show the helicase activity.

DISCUSSION

Of late, geminiviruses are devastating the crop-cultivation worldwide and the virus MYMIV alone causes an estimated annual loss of about 300 million dollars in legume cultivation (24). The nature of RCR of viral DNA and the recombinogenic property of the single-stranded viral DNA might contribute greatly to the cause of virus variability (25). Though some host and viral factors are required, the geminiviral Rep protein is, by far, the most important one for viral DNA replication. The MYMIV-Rep protein initiates RCR and plays the role as a replicative helicase. The findings of this report would thus help in understanding the mechanism(s) pertaining to fork-extension during the replication of MYMIV–DNA.

Various motifs and oligomerization of Rep

In search of a helicase protein for viral DNA replication, we observed that the ATPase activity of Rep was up regulated by exogenous ssDNA (Figure 1F). Since such up-regulation is a characteristic feature of a helicase protein, we were led to examining the novel helicase activity of the recombinant

MYMIV-Rep. The chemical ‘nogalamycin’, a known specific inhibitor of the helicase activity (26), also turned out to be a potent inhibitor of the Rep-mediated helicase activity. Thus the purified Rep, despite lacking several other motif-sequences that are characteristics of most of the helicases, was functionally similar to other helicases. The complete inhibition of Rep’s helicase activity in presence of ADP further indicated that Rep was unable to utilize ADP as the energy source. All the helicases reported so far has been shown to utilize either nucleotide or deoxynucleotide triphosphates, but not diphosphates (27), which is consistent with our findings.

Rep C (183–362 amino acid) containing the Walker A/B motifs was as strong an ATPase as the wild-type full-length Rep but was a weak helicase. However, the Rep fragment (120–362 amino acid) behaved as an efficient helicase and ATPase. Thus, the Rep-peptide spanning the region from 120 to 182 amino acid plays a great role in the helicase activity of the Rep protein. This region of Rep, termed as the oligomerization domain, is generally similar in sequence among various begomoviruses. Though the presence of this domain makes the geminiviral Rep protein oligomeric in character, the stoichiometry of the subunits has not been studied earlier in details (15,28). The studies using sucrose density sedimentation and nondenaturing gel electrophoresis suggested that the recombinant MYMIV-Rep is highly oligomeric and perhaps existed as a large oligomeric complex (molecular weight > 1000 kDa) in solution and any loss of the oligomeric character resulted in loss of the helicase activity (Figure 5A, B and D, and Supplementary Figure S4A and S4B). To the best of our knowledge, this is the first report of estimation of the molecular mass of any geminiviral Rep protein in the purified form. The helicases like hexameric *E.coli* DNAB-type helicases, the double hexameric SV-40 large T-antigen etc. reveal oligomer-dependent replicative helicase activity but the oligomeric state as huge as that of the Rep-helicase has not yet been reported in the literature (18,29). The secondary structure prediction tools (PSIPRED Protein Structure Prediction Server, UCL, UK) reveal that the oligomerization domain of Rep largely (72%) consists of coiled structures and beta sheets but little alpha helices (28%); and this domain functionally controls self-interaction, interaction with a range of hosts and virus-encoded factors (3,11,23,30–33). The formation of the large oligomeric state might thus be dictated by several constraints of structure and function. Many different laboratories have purified geminiviral Rep proteins using various techniques, namely, baculovirus and *E.coli* expression systems using the affinity-tags but the helicase activities were not detectable with such

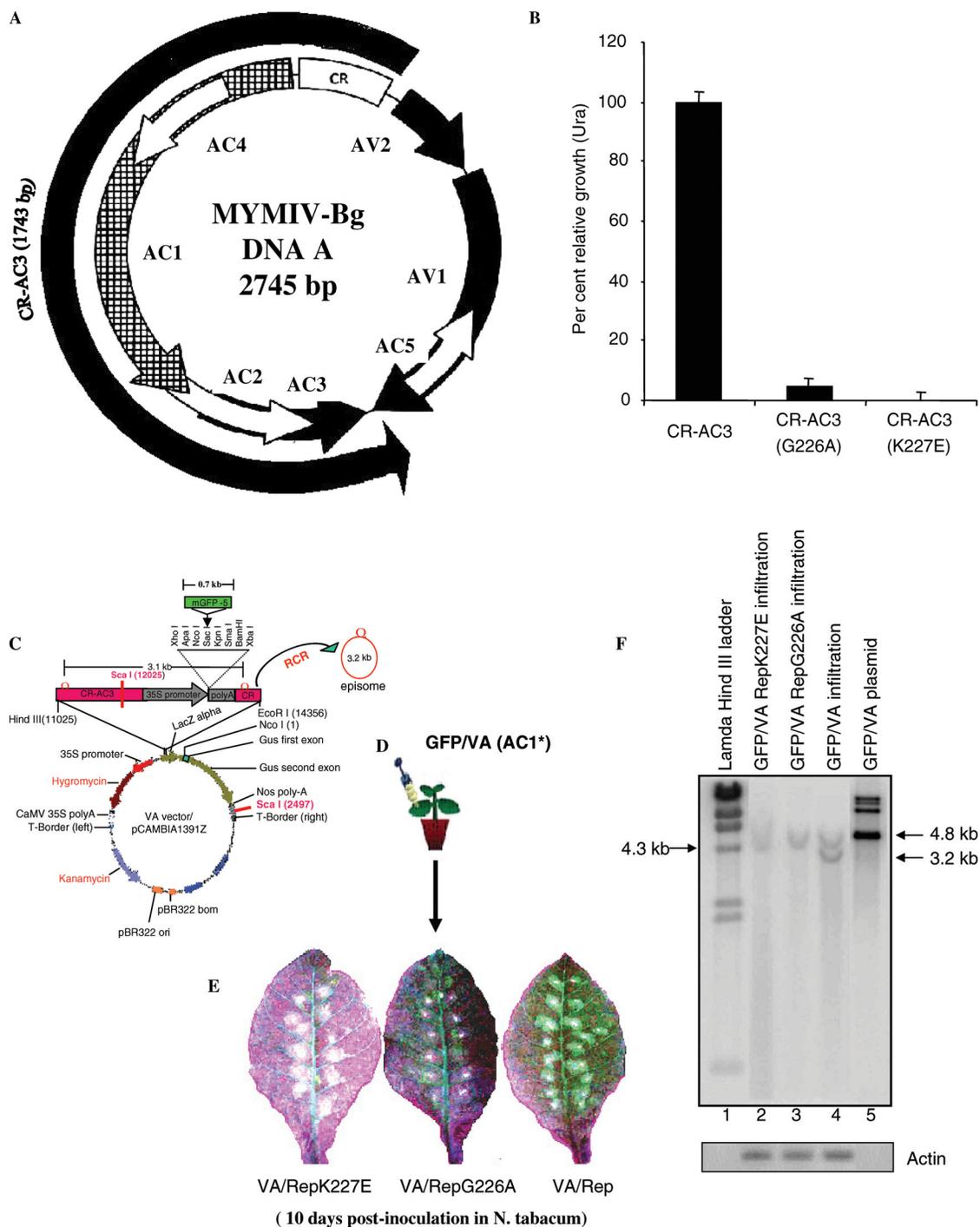


Figure 6. *Ex vivo* and transient *in planta* analysis of replication efficiency of Rep and its helicase mutants. (A) The map of MYMIV-DNA 'A' genome. The map shown on the left side of the panel depicts the positions of different ORFs and the construct CR-AC3 used for the study. The description and notations of various constructs were cloned in YCPO⁻ background and their colony forming abilities in *S. cerevisiae* are given in Table 3. (B) Survival of the helicase mutants. The results of similar colony forming experiments with the constructs carrying CR-AC3 region or the mutants, *viz.*, CR-AC3 (RepG226A) and CR-AC3 (RepK227E) are presented in the form of column graphs. The colony forming abilities of constructs carrying above-mentioned mutations of CR-AC3 were expressed taking the colony forming ability of construct carrying CR-AC3 region to be 100%. The values presented are the average of six independent readings. The standard deviations are indicated in the column graphs. (C) Map showing the VA vector and its cloning sites. The engineered viral insert containing either Rep or its mutants [G (226) A and K (227) E], each carrying GFP as a marker, was cloned in the pCambia1391Z background. The circular episome that formed in the plant-leaves due to replication-release following agro-inoculation is shown schematically using a curved arrow. (D) Non-transgenic tobacco leaves were agro-inoculated with the above-mentioned constructs. (E) The tobacco leaves were collected at 10 dpi and GFP expression was monitored by exposing the leaves to UV. The wt-Rep containing construct showed high expression of GFP, while the mutants exhibited much reduced expression of GFP. (F) Autoradiograph of Southern hybridization of genomic DNA isolated from the inoculated leaves using radiolabeled GFP DNA as the probe. The appearance of 3.2 kb fragment in case of wild-type Rep construct reflected the ability of Rep to cause replication in the plant, while the presence of 4.8 kb band revealed the unreplicated paternal DNA. In the bottom panel, actin amplification data are presented as the loading control, using the template DNA that were employed for the lanes 2–4 of the figure shown in the upper panel.

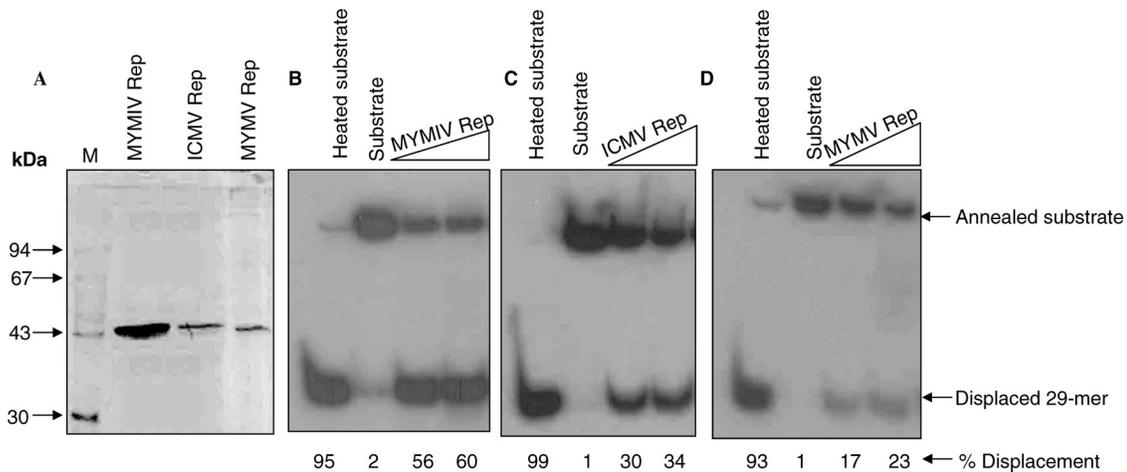


Figure 7. Helicase activities of Rep from various geminiviruses. (A) Coomassie blue stained gel showing the degree of purity of Rep proteins from MYMIV, ICMV and MYMV. About 4.5, 0.75 and 0.5 μ g of the proteins of MYMIV-Rep, ICMV-Rep and MYMV-Rep were loaded, respectively. The helicase activities at two different concentrations (1 and 2 μ g) of Rep proteins of MYMIV, ICMV and MYMV using oligonucleotide HELFOK9 (29mer) annealed to M13mp18 ssDNA as substrate are shown in (B–D), respectively. The data revealed the ability of all of these viral Rep proteins to act as helicase.

proteins (8,34). It is possible that the conditions of the expression systems, such as the temperature of induction, denaturation of the inclusion bodies followed by renaturation, post-translational modification, and the presence of fusion-tag etc. stood in the way of formation of correct oligomeric and helicase-active state of Rep. It is also worthwhile to note that two additional, putative motifs of helicase, namely the 9 amino acid containing Q motif with the invariant terminal Q (120GRSARGGQQ128) and the motif VI (117QIDGRSAR124) are localized in the oligomerization domain of MYMIV-Rep. The Q motif is supposed to regulate the ATPase activities and the motif VI probably provides the unwindase activities of any helicase (13,35). However, the functionality of these motifs needs to be demonstrated.

Rep as a replicative helicase

Studies on T-antigen and other helicases reveal that helicases require some DNA-space to load onto DNA. T-antigen has been observed to require 9–10 nt for binding to ssDNA portion, while quantitative binding studies with PriA and rho show that generally 7–8 nt are bound by each subunit. Co-crystal structures of the Rep, NS3 of hepatitis virus and PcrA helicases with ssDNA or partial duplexes also point to the similar requirements of 7–8 nt (18). The MYMIV-Rep needed 6–9 nt of ssDNA for proper loading and functioning (Figure 3F–I). This implies that melting of viral DNA strand might occur for loading of Rep-helicase at the site of initiation of RCR. Following loading, the Rep-helicase tracks in the 3'→5' direction along the non-viral strand (Figure 3C and E) to facilitate copying of non-viral strand, thus producing the nascent viral DNA strand. Such a scenario calls for the replicative helicase activity of the Rep protein.

Furthermore, the Rep-helicase showed processivity, unwinding at least up to 200 nt from the partial duplex substrate (Figure 3B). Such processivity is often associated with the activity of a replicative helicase. Finally, the replicative character of the Rep-helicase is evident from the set of experiments shown in Figure 6. Earlier, Raghavan *et al.*

had reported the ability of MYMIV genome to replicate efficiently in yeast system (3). Here we demonstrate that the minimal segment of MYMIV genome that could sustain replication in such a surrogate system is the CR-AC3 region (Figure 6A). While the MYMIV CR-AC3 region could replicate efficiently in *S.cerevesiae*, the G (226) A and K (227) E mutations of Rep did not support replication (Figure 6B, Table 3). Similar experiments *in planta* with tobacco led to identical results. These mutations of Rep had been shown *in vitro* to adversely affect only the ATPase and helicase activities of Rep but not the other activities, *viz.*, site-specific DNA-binding, nicking-closing, oligomerization, etc. that are important for initiation of RCR. The combined data immediately suggest the involvement of Rep as the replicative helicase. However, the roles of either the virus-encoded factors or the host proteins in enhancing the helicase activity of Rep cannot be underestimated. There are at least two MYMIV-DNA-A coded factors that up regulate the ATPase/helicase activity of Rep and some host proteins having helicase activities, e.g. Rad 54, Rad 51, etc., associate with Rep (S. K. Mukherjee *et al.*, unpublished data). It would be interesting to find the composition of the fork-proteins during the RCR of the MYMIV genome since the speed and processivity of the Rep-helicase might be higher in the replisome form than in its free native state.

Characteristic helicase activity of begomovirus Rep proteins

To address the point whether the Rep-mediated helicase activity is a common occurrence among the geminiviruses, the Rep proteins from two independent begomoviruses, namely MYMV and ICMV, which are phylogenetically close to and distant from MYMIV, respectively, were examined. The Walker A/B motif-sequences of these three viruses are identical and the oligomerization domains of MYMV and ICMV share ~89 and 75% amino acid sequences identity respectively to that of MYMIV. The observation that the recombinant Rep proteins of ICMV, MYMV and MYMIV showed

helicase activity to a comparable degree might suggest that the geminivirus Rep proteins have evolved to possess the helicase activity in general. However, the broad-range validity of the above hypothesis remains to be tested.

The MYMIV-Rep and the Rep-78 of adeno-associated virus-type 2 (AAV-2) are closely analogous in their biological behavior as both the proteins serve as replication initiators by virtue of their nicking-closing and helicase activities. The structural as well as functional similarities of their NTP-binding domains classify them as members of a superfamily (SF3). SF3 includes the NTP-binding pattern containing conserved protein segments encoded by genomes of small DNA and RNA viruses (36). Though the SF3 family members are derived from dissimilar groups of ss- and ds-DNA viruses, namely papova-, parvo-, geminiviruses and P4 bacteriophage, they show helicase activities with characteristic oligomeric constitution (37,38). A majority of the listed members from the DNA viruses are also involved in the initiation of DNA replication. Keeping all these facts in mind, it is tempting to speculate that all geminiviral Rep proteins might act as replicative helicases using their domains of oligomerization.

While our manuscript was under preparation, the helicase activity of truncated Rep protein from a monopartite begomovirus, namely Tomato yellow leaf curl Sardinia virus (TYLCSV) was reported (39). Though the two helicases, namely MYMIV-Rep and TYLCSV-Rep, have lot of biochemical similarities, there are few important differences too. TYLCSV-Rep 122–359 did not show helicase activity with unforked substrates, while the MYMIV-Rep showed efficient helicase activity with such substrates. Additionally, in case of TYLCSV-Rep 122–359, the helicase activity was associated with dodecameric conformation, while the MYMIV-Rep showed helicase activity with much higher order oligomeric conformation (~24mer). The *in vivo* replicative characteristics of both the viruses might account for such differences of biological activities. The identification and characterization of Rep interacting proteins of both the viruses might be important in determining the biochemical activities of the two proteins.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

The authors thank Dr Basavaraj Bagewadi and Mr Vikash Kumar for their technical help. The gifts from Prof. K. Veluthambi, Madurai Kamraj University, Madurai, India, and Dr Indranil Dasgupta, University of Delhi, New Delhi, India of the Rep DNA clones of MYMV (Mungbean yellow mosaic virus) and ICMV (Indian cassava mosaic virus), respectively, are gratefully acknowledged. The Open Access publication charges for this article were waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

1. Hanley-Bowdoin, L., Settledge, S.B., Orozco, B.M., Nagar, S. and Robertson, D. (1999) Geminiviruses: models for plant DNA replication,

- transcription, and cell cycle regulation. *Crit. Rev. Plant Sci.*, **18**, 71–106.
2. Pant, V., Gupta, D., Roy Choudhury, N., Malathi, V.G., Varma, A. and Mukherjee, S.K. (2001) Molecular characterization of the Rep protein of the blackgram isolate of Indian mungbean yellow mosaic virus. *J. Gen. Virol.*, **82**, 2559–2567.
3. Raghavan, V., Malik, P.S., Roy Choudhury, N. and 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.
4. Elmer, J.S., Brand, L., Sunter, G., Gardiner, W.E., Bisaro, D.M. and Rogers, S.G. (1988) Genetic analysis of the tomato golden mosaic virus. II. The product of the AL1 coding sequence is required for replication. *Nucleic Acids Res.*, **16**, 7043–7060.
5. Gutierrez, C. (2000) DNA replication and cell cycle in plants: learning from geminiviruses. *EMBO J.*, **19**, 792–799.
6. Hong, Y., Stanley, J. and van Wezel, R. (2003) Novel system for the simultaneous analysis of geminivirus DNA replication and plant interactions in *Nicotiana benthamiana*. *J. Virol.*, **77**, 13315–13322.
7. Orozco, B.M. and Hanley-Bowdoin, L. (1996) A DNA structure is required for geminivirus replication origin function. *J. Virol.*, **70**, 148–158.
8. Desbiez, C., David, C., Mettouchi, A., Laufs, J. and Gronenborn, B. (1995) Rep protein of tomato yellow leaf curl geminivirus has an ATPase activity required for viral DNA replication. *Proc. Natl Acad. Sci. USA*, **92**, 5640–5644.
9. Egelkrout, E.M., Mariconti, L., Settledge, S.B., Cella, R., Robertson, D. and Hanley-Bowdoin, L. (2002) Two E2F elements regulate the proliferating cell nuclear antigen promoter differently during leaf development. *Plant Cell*, **14**, 3225–3236.
10. Khan, S.A. (2000) Plasmid rolling-circle replication: recent developments. *Mol. Microbiol.*, **37**, 477–484.
11. Bagewadi, B., Chen, S., Lal, S.K., Roy Choudhury, N. and Mukherjee, S.K. (2004) PCNA interacts with Indian Mung Bean Yellow Mosaic Virus Rep and downregulates Rep activity. *J. Virol.*, **78**, 11890–11903.
12. Hanson, S.F., Hoogstraten, R.A., Ahlquist, P., Gilbertson, R.L., Russell, D.R. and Maxwell, D.P. (1995) Mutational analysis of a putative NTP-binding domain in the replication-associated protein (AC1) of bean golden mosaic geminivirus. *Virology*, **211**, 1–9.
13. Gorbalenya, A.E. and Koonin, E.V. (1993) Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.*, **3**, 419–429.
14. Campos-Olivas, R., Louis, J.M., Clerot, D., Gronenborn, B. and Gronenborn, A.M. (2002) The structure of a replication initiation protein unites diverse aspects of nucleic acid metabolism. *Proc. Natl Acad. Sci. USA*, **99**, 10310–10315.
15. Orozco, B.M., Miller, A.B., Settledge, S.B. and Hanley-Bowdoin, L. (1997) Functional domains of a geminivirus replication protein. *J. Biol. Chem.*, **272**, 9840–9846.
16. Patel, S.S. and Picha, K.M. (2000) Structure and function of hexameric helicases. *Ann. Rev. Biochem.*, **69**, 651–697.
17. Konieczny, I. (2003) Strategies for helicase recruitment and loading in bacteria. *EMBO Rep.*, **4**, 37–41.
18. Alexandrov, A.I., Botchan, M.R. and Cozzarelli, N.R. (2002) Characterization of simian virus 40 T-antigen double hexamers bound to a replication fork. The active form of the helicase. *J. Biol. Chem.*, **277**, 44886–44897.
19. Soutanas, P., Dillingham, M.S., Wiley, P., Webb, M.R. and Wigley, D.B. (2000) Uncoupling DNA translocation and helicase activity in PcrA: direct evidence for an active mechanism. *EMBO J.*, **19**, 3799–3810.
20. Amaratunga, M. and Lohman, T.M. (1993) *Escherichia coli* rep helicase unwinds DNA by an active mechanism. *Biochemistry*, **32**, 6815–6820.
21. Kornberg, A. and Baker, T.A. (1992) *DNA Replication*. W. H. Freeman and Company, New York, USA.
22. You, Z., Ishimi, Y., Masai, H. and Hanaoka, F. (2002) Roles of Mcm7 and Mcm4 subunits in the DNA helicase activity of the mouse Mcm4/6/7 complex. *J. Biol. Chem.*, **277**, 42471–42479.
23. Malik, P.S., Kumar, V., Bagewadi, B. and Mukherjee, S.K. (2005) Interaction between coat protein and replication initiation protein of Mung bean yellow mosaic India virus might lead to control of viral DNA replication. *Virology*, **337**, 273–283.

24. Varma,A. and Malathi,V.G. (2003) Emerging geminivirus problems: a serious threat to crop production. *Ann. Appl. Biol.*, **142**, 145–164.
25. Jeske,H., Lutgemeier,M. and Preiss,W. (2001) DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. *EMBO J.*, **20**, 6158–6167.
26. George,J.W., Ghate,S., Matson,S.W. and Besterman,J.M. (1992) Inhibition of DNA helicase II unwinding and ATPase activities by DNA-interacting ligands. *J. Biol. Chem.*, **267**, 10683–10689.
27. Tuteja,N. and Tuteja,R. (2004) Unraveling DNA helicases: motif, structure, mechanism and function. *Eur. J. Biochem.*, **271**, 1849–1863.
28. Sanz-Burgos,A.P. and Gutierrez,C. (1998) Organization of the *cis*-acting element required for wheat dwarf geminivirus DNA replication and visualization of a rep protein-DNA complex. *Virology*, **243**, 119–129.
29. LeBowitz,J.H. and McMacken,R. (1986) The *Escherichia coli* dnaB replication protein is a DNA helicase. *J. Biol. Chem.*, **261**, 4738–4748.
30. Arguello-Astorga,G., Lopez-Ochoa,L., Kong,L.J., Orozco,B.M., Settlege,S.B. and Hanley-Bowdoin,L. (2004) A novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein. *J. Virol.*, **78**, 4817–4826.
31. Xie,Q., Sanz-Burgos,A.P., Guo,H., Garcia,J.A. and Gutierrez,C. (1999) GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. *Plant Mol. Biol.*, **39**, 647–656.
32. Kong,L.J. and Hanley-Bowdoin,L. (2002) A geminivirus replication protein interacts with a protein kinase and a motor protein that display different expression patterns during plant development and infection. *Plant Cell*, **14**, 1817–1832.
33. Settlege,S.B., See,R.G. and Hanley-Bowdoin,L. (2005) Geminivirus C3 protein: replication enhancement and protein interactions. *J. Virol.*, **79**, 9885–9895.
34. Castillo,A.G., Kong,L.J., Hanley-Bowdoin,L. and Bejarano,E.R. (2004) Interaction between a geminivirus replication protein and the plant sumoylation system. *J. Virol.*, **78**, 2758–2769.
35. Tanner,N.K., Cordin,O., Banroques,J., Doère,M. and Linder,P. (2003) The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol. Cell*, **11**, 127–138.
36. Gorbalenya,A.E., Koonin,E.V. and Wolf,Y.I. (1990) A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. *FEBS Lett.*, **262**, 145–148.
37. James,J.A., Escalante,C.R., Yoon-Robarts,M., Edwards,T.A., Linden,R.M. and Aggarwal,A.K. (2003) Crystal structure of the SF3 helicase from adeno-associated virus type-2. *Structure*, **11**, 1025–1035.
38. Hickman,A.B. and Dyda,F. (2005) Binding and unwinding: SF3 viral helicases. *Curr. Opin. Struct. Biol.*, **15**, 77–85.
39. Clerot,D. and Bernardi,F. (2006) DNA helicase activity is associated with the replication initiator protein Rep of Tomato yellow leaf curl geminivirus. *J. Virol*, **80**, 11322–11330.