Journal of General Virology (2001), **82**, 2559–2567. Printed in Great Britain

Molecular characterization of the Rep protein of the blackgram isolate of Indian mungbean yellow mosaic virus

V. Pant,[1] D. Gupta,[1] N. Roy Choudhury,[1] V. G. Malathi,[2] A. Varma[2] and S. K. Mukherjee[1]

[1] Plant Molecular Biology, International Centre for Genetic Engineering & Biotechnology, Aruna Asaf Ali Marg, New Delhi 110067, India
[2] Division of Plant Pathology, Indian Agricultural Research Institute, New Delhi 110012, India

The complete nucleotide sequence of the blackgram isolate of mungbean yellow mosaic virus, IMYMV-Bg, which infects legumes in India, was determined and compared at the amino acid level with those of other whitefly-transmitted geminiviruses. The genome organization of IMYMV-Bg was similar to that of the begomoviruses. A unique feature of the genome organization was the sequence divergence of the common region (CR) between DNA-A and DNA-B. In order to understand the mechanism of viral DNA replication, the replication initiator protein, Rep, of IMYMV-Bg was overexpressed in *E. coli*. The recombinant and refolded Rep bound to CR-sequences of IMYMV-Bg in a specific manner. In this study, evidence is presented for ATP-upregulated cleavage function and ATP-mediated conformational change of Rep. It is hypothesized that, although ATP is not required for cleavage, ATP-mediated conformational changes may result in better access of Rep to the DNA-cleavage site. Evidence is also presented for a site-specific topoisomerase function of Rep, which has not been demonstrated before. The Rep protein can be classified as a type-I topoisomerase because of its nicking activity and sensitivity towards camptothecin, a topoisomerase type-I inhibitor.

Introduction

Pulses and grain legumes, a major source of dietary protein, are subject to yellow mosaic and golden mosaic diseases caused by whitefly-transmitted geminiviruses (WTGs or begomoviruses) in tropical and subtropical countries. Of these viruses, mungbean yellow mosaic virus (MYMV) is important as it infects five major leguminous species, blackgram, mungbean, French bean, pigeonpea and soybean, causing an annual loss of yield of about $300 million (Varma *et al.*, 1992). MYMV was first observed in Delhi in the late fifties (Nariani, 1960). It produces typical yellow mosaic symptoms and is easily transmitted by the whitefly vector *Bemisia tabaci* (Nene, 1973).

Most of the yellow mosaic viruses infecting legumes in India are not sap-transmissible, share a very narrow host range within legumes and cause biologically indistinguishable symptoms, making specific identification of the viruses difficult.

Author for correspondence: Sunil Mukherjee.
Fax +91 11 616 2316. e-mail sunilm@icgeb.res.in

The GenBank accession numbers for the DNA-A and DNA-B sequences of IMYMV-Bg are AF126406 and AF142440.

Based on epitope-profile studies using panels of monoclonal antibodies to Indian and African cassava mosaic geminivirus (ICMV and ACMV), Swanson *et al.* (1992) classified legume geminiviruses in India broadly into two groups. One group comprises dolichos yellow mosaic virus and the other group includes the yellow mosaic viruses that infect blackgram, cowpea, French bean, horsegram, pigeonpea, soybean and mungbean etc. A WTG, *Vigna mungo* yellow mosaic virus (VMYMV), infecting blackgram in South India has been described recently (Vanitha Rani *et al.*, 1996). In addition to India, the virus is widely prevalent in other countries of the Indian subcontinent, namely Sri Lanka, Bangladesh and Pakistan. An epidemic outbreak of yellow mosaic disease of mungbean was also identified in Thailand in the 1980s, which was caused by a sap-transmissible WTG, Thailand MYMV (TMYMV) (Honda *et al.*, 1983). TMYMV has been well characterized (Morinaga *et al.*, 1990) and found to be different from other WTGs derived from India in its epitope profile (Harrison *et al.*, 1991).

In an attempt to unravel the complexity of WTGs infecting legumes in India, genomic components of the isolates of MYMV infecting blackgram, mungbean, mothbean and pigeonpea have been cloned and the infectivity of the clones

was demonstrated by using agroinoculation techniques (Mandal *et al.*, 1997). The nucleotide sequence of the blackgram isolate of MYMV, hereafter referred to as IMYMV-Bg, was determined. The genome organization of a full-length, infectious, insect-transmissible clone of IMYMV-Bg confirmed its inclusion as a member of the genus *Begomovirus* (Harrison & Robinson, 1999) of the family *Geminiviridae*. In this report, we touch upon the relationship of IMYMV to other begomoviruses (Padidam *et al.*, 1995; Rybicki, 1994) and examine the functions of the IMYMV-Bg Rep protein.

The majority of members of the begomovirus subgroup have bipartite, single-stranded circular genomes, known as DNA-A and DNA-B. In general, DNA-A has one ORF (AR1) in the virus sense and three in the complementary sense (AL1, AL2 and AL3) (Stanley, 1991; Lazarowitz, 1992). In DNA-B, there is one ORF in each sense (BR1 and BL1). In both genome components, the ORFs extend in opposite directions from a 180–200 nt region common to both components. This region is called the common region (CR) and contains the viral DNA-replication origin, as well as the promoter for the leftward ORFs (ALs). Within this region, there is a stretch of 31 bp that includes the characteristic invariant nonamer TAATATTAC, which is found in all geminivirus DNA-replication origins. This region has a characteristic secondary structure containing a GC-rich stem and an AT-rich loop.

There are two kinds of proteins encoded by begomoviruses, namely the replication initiator (Rep or AL1) and replication enhancer (Ren or AL3), that are required for viral DNA replication. Various properties of Rep proteins, especially those of tomato golden mosaic virus (TGMV), have been reported previously and the mode of function of Rep as the rolling-circle replication-initiation protein is emerging (Bisaro, 1996; Hanley-Bowdoin *et al.*, 1999). During rolling-circle replication, Rep binds to specific sequences (repeats/iterons) present in the CR and hydrolyses the phosphodiester bond between the seventh and eighth residues of the invariant nonamer 5′ TAATATT↓AC 3′ (Stanley, 1995; Laufs *et al.*, 1995). Rep remains bound covalently to the 5′-phosphate end, and the 3′-hydroxyl end thus generated becomes available for rolling-circle replication. After a full cycle of replication, the new origin sequence is generated, which is again hydrolysed by Rep. Subsequently Rep ligates the nascent 3′ end of DNA with the previously generated 5′ end. In this way, a unit-genome length, circular, single-stranded DNA molecule, the mature viral genome, is processed.

In order to understand the mechanisms of geminivirus pathogenesis so as to interfere with the virus growth process, it is imperative to understand the biochemical steps of viral DNA replication within the nucleus of the infected plant cell. We wish to address questions related to DNA replication using the genome of IMYMV-Bg, because of its prevalence in the agricultural fields of India. As a first step, the Rep of IMYMV-Bg was cloned and overproduced in bacteria. Here, we show that the recombinant Rep can be refolded to a functionally active form and its nicking activity is modulated by ATP. We also report for the first time that Rep can act as a site-specific type-I topoisomerase.

Methods

■ **Virus source.** Cloning and characterization of IMYMV-Bg genomic components have been described previously (Mandal *et al.*, 1997). Infectious clones of DNA-A (pMYAH13) and DNA-B (pMYBB12) in pUC18, maintained in *E. coli* strain DH5α, were used for the study.

■ **Phylogenetic analysis.** Phylogenetic analyses were carried out by the neighbour-joining method and distances were calculated using the PROTDIST program based on the PAM001 model. To enhance the robustness of the branching, a total of 1000 bootstrap replicates of the data were performed for phylogenetic reconstructions using SEQBOOT. Distances calculated from PROTDIST were used as input for the distance-matrix program FITCH (Fitch–Margoliash and least-squares distance method). CONSENSE was used to build a strict consensus bootstrap tree. TREEVIEW roderic DM 2000, procured from http://taxonomy.zoology.gla.ac.uk/rod/rod.html, was used to display the phylogenetic trees (not shown).

■ **Expression and purification of IMYMV-Bg Rep.** The entire ORF of IMYMV Rep (AL1) was amplified by PCR using the full-length DNA-A clone. The following primers were employed in the PCR: MYMV-AL1 forward, 5′ ATGGATCCATGCCAAGGGAAGGTCGT 3′; and MYMV-AL1 reverse, 5′ TGAAAGCTTTCAATTCGAGATC-GTCGA 3′. The calculated length of the amplified fragment was 1098 bp and it was cloned between the *Bam*HI and *Hin*dIII sites of the overexpression vector pET28a. *E. coli* BLR (DE3) cells were transformed with the recombinant plasmid and induced to express Rep with 1 mM IPTG. The crude lysate of the bacteria contained ~ 43 kDa induced Rep protein. The majority of the Rep localized within inclusion bodies and was solubilized and refolded by using standard procedures (Arora & Khanna, 1996). Briefly, the inclusion bodies were isolated from 800 ml culture induced with 1 mM IPTG and solubilized in 40 ml S buffer (10 mM Tris–HCl, pH 8, 100 mM NaH_2PO_4, 8 M guanidine hydrochloride, 10 mM DTT) at 4 °C. This solution was mixed gently with 1 l refolding buffer R [containing 60·6 g urea, 105 g arginine, 50 ml 1 M Tris–HCl, pH 8, 2 ml 0·5 M EDTA, 614 mg reduced glutathione (GSH) and 122 mg oxidized glutathione (GSSG)] and kept at 10 °C for 30 h. This preparation was dialysed against 20 l standard PBS buffer, pH 8, for 10 h with three buffer changes. The dialysate was then passaged through a Ni–NTA column and the bound protein was eluted with a gradient of 100–400 mM imidazole in E buffer (50 mM Tris–HCl, pH 8, 50 mM NaCl). The Rep protein (250–300 mM imidazole) was finally dialysed against 40% glycerol in E buffer, concentrated by ultrafiltration and stored in the presence of a set of six protease inhibitors (1 μg/ml aprotinin, 1 μg/ml leupeptin, 2 μg/ml antipain, 100 μg/ml PMSF, 1 μg/ml pepstatin, 3 mg/ml benzamidine) and 10 mM DTT.

■ **Nitrocellulose filter-binding assay.** The ability of the refolded Rep protein to bind DNA in a sequence-specific manner was tested by a competitive filter-binding assay (Hop *et al.*, 1999). The template DNAs used for binding were CR A (190 bp), CR B (192 bp) and CP_f (500 bp). These DNA fragments were amplified by PCR employing the primers mentioned below. The fragments were cloned in pGEM-T vector and suitable restriction fragments containing the DNA were excised for the purpose of 3′-end labelling.

Primer sequences were: CR A Forward, 5′ TAGCAAAACGACC-TTCCCTTGG 3′; CR A Reverse, 5′ GCAGAAAAGTTAAAGTAAC-CCC 3′; CR B Forward, 5′ GCTAGCCAAAAGAGCGTGTCG 3′; CR

B Reverse, 5′ ACTCCATAGTGCCCACGTATT 3′; CP_r Forward: 5′ GGTCCCCTGATGTCCCTCGTG 3′; CP_r Reverse, 5′ ATGCGTTCT-CAGTATGGTTCT 3′.

Usually about 10 ng radiolabelled DNA was incubated with different concentrations of Rep in 50 µl reaction mixture containing 50 mM Tris–HCl, pH 8·0, 1 mM EDTA, 100 mM NaCl and 5 mM $MgCl_2$ for 20 min at 30 °C. Incubations were carried out with or without unlabelled competitor DNA. After incubation, the DNA–protein complexes were trapped on Millipore nitrocellulose filters (DAWP, pore size 65 µm). Filters were washed with 3 ml binding buffer, dried and subjected to scintillation counting to determine retained radioactivity. The ratio of the amount of radioactivity retained in the presence of Rep and the amount of input label was expressed as the fraction of DNA bound to the filter. Values given in Fig. 2(*a*, *b*) represent the means of five tests.

■ **Cleavage and ligation function of Rep.** Oligonucleotides corresponding to sequences in the hairpin region of IMYMV-Bg DNA-A, i.e. T1, and other unrelated sequences (T3) were 5′-end-labelled using T4 polynucleotide kinase. The primer T1 used for the cleavage test was a 26-mer bearing the sequence 5′ CGACTCAGCTATAATATTACC-TGAGT 3′. The Rep-mediated cleavage product that could be visualized was a 5′-labelled 18-mer. The primer T2 used for the ligation test was an unlabelled 26-mer with the sequence 5′ CTATAATATTACCTGAGT-GCCCCGCG 3′. The ligation product following the cleavage of T1 and T2 was visualized as a labelled 34-mer. T3 was a forward M13 primer and was used as a non-specific control substrate. Approximately 1 ng (spec. act. $\sim 1{\cdot}5 \times 10^9$ c.p.m./µg) of either T1 or T3 was incubated with 5 µg purified Rep protein in 10 µl cleavage buffer (Orozco *et al.*, 1997) at 37 °C for 30 min. The reaction was terminated by adding 6 µl loading buffer (1% SDS, 25 mM EDTA, 10% glycerol) and heating to 90 °C for 2 min. The products were resolved on a 15% acrylamide–urea gel and analysed by autoradiography.

■ **Topoisomer assay.** The topoisomerase function was assayed by measuring the ability to relax a negatively supercoiled DNA molecule and generate a topoisomer ladder. One µg negatively supercoiled recombinant plasmid IMYMV DNA-A or DNA-B was incubated with the indicated amount of Rep protein in 20 µl reaction mixture containing 50 mM Tris–HCl, pH 7·5, 50 mM KCl and other co-factors, if necessary. The reaction mixture was incubated at 37 °C for 30 min and the topoisomers were resolved in a 1% agarose gel by electrophoresis at 2 V/cm for 16 h. The DNA bands were visualized by staining with ethidium bromide.

Results

Genome organization of IMYMV-Bg

The virus clones of IMYMV-Bg DNA-A and DNA-B were sequenced (accession nos. AF126406 and AF142440) and analysis of the sequences indicated that IMYMV-Bg is a member of the genus *Begomovirus*. Base composition, gene ordering and disposition of the common regions (CR) between the bipartite genomes confirmed the classification of IMYMV-Bg.

IMYMV Rep (or AL1) showed a high degree of similarity in sequence and domain organization to the Rep proteins of other begomoviruses [62% to bean golden mosaic virus (BGMV) Rep and about 84% to the Reps of TMYMV and VMYMV]. The overall similarity of the genome sequences and the disease symptoms shared by the three viruses IMYMV-Bg, TMYMV and VMYMV suggest that they may form a subfamily within the family of begomoviruses. The coat protein (CP/AR1) sequences of this subfamily displayed the maximum degree of similarity ($\geqslant$ 92%) in a C-terminal domain that constituted about 80% of the total protein. The DNA-B component carried one ORF in each of the virus and complementary senses, BL1 and BR1. The BL1 protein of IMYMV showed high conservation relative to the corresponding proteins of other begomoviruses ($>$ 93% with BL1 of each of VMYMV and TMYMV). However, IMYMV BR1 showed only 86% similarity to the counterparts from VMYMV and TMYMV. The replication enhancer (Ren/AL3) and the transcription-activator (Trap/AL2) of IMYMV-Bg also seemed to be different from the similar proteins of other begomoviruses. Another interesting region of the IMYMV genome was its CR. The intergenic region (IR) of IMYMV DNA-B is larger than that of DNA-A and the level of sequence similarity between the two IRs is lower than that generally found in other begomoviruses, including the MYMV subfamily. The CRs of the IMYMV genomes spanned about 135 nt with 43 point changes.

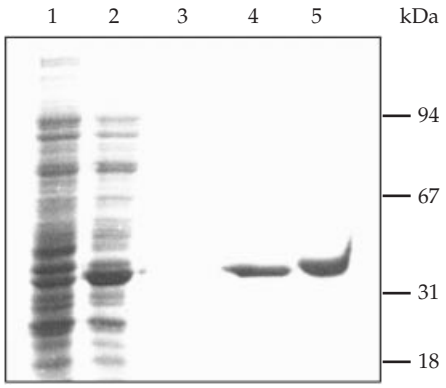

Fig. 1. Purification profile of recombinant Rep. Rep was overexpressed in *E. coli* BLR (DE3) cells as a 6 × His-tagged protein and purified by using a Ni–NTA agarose column. Proteins were resolved by 12% SDS–PAGE and visualized by staining with Coomassie blue. Lanes: 1, 60 µg protein from uninduced *E. coli* cells; 2, 60 µg protein from cells induced with 1 mM IPTG for 3 h; 3, proteins from clear supernatant of *E. coli* sonic lysate prepared following induction; 4 and 5, early and late protein fractions eluted from the Ni–NTA column with a gradient of 100–400 mM imidazole.

Cloning and overexpression of IMYMV Rep

Since the template DNA, the sequence of which is reported here, causes virus infection following agroinoculation, its functionality is proven. In order to study function at the molecular level, it was necessary to examine the putative activities of the predicted virus genes. Towards this end, the virus Rep protein was overproduced in bacteria and the biochemical properties of the recombinant Rep protein were examined.

A 1098 bp long DNA encoding the IMYMV Rep was amplified by PCR and recloned in pET28a for overexpression in *E. coli*. Fig. 1 shows that the His-tagged recombinant Rep of

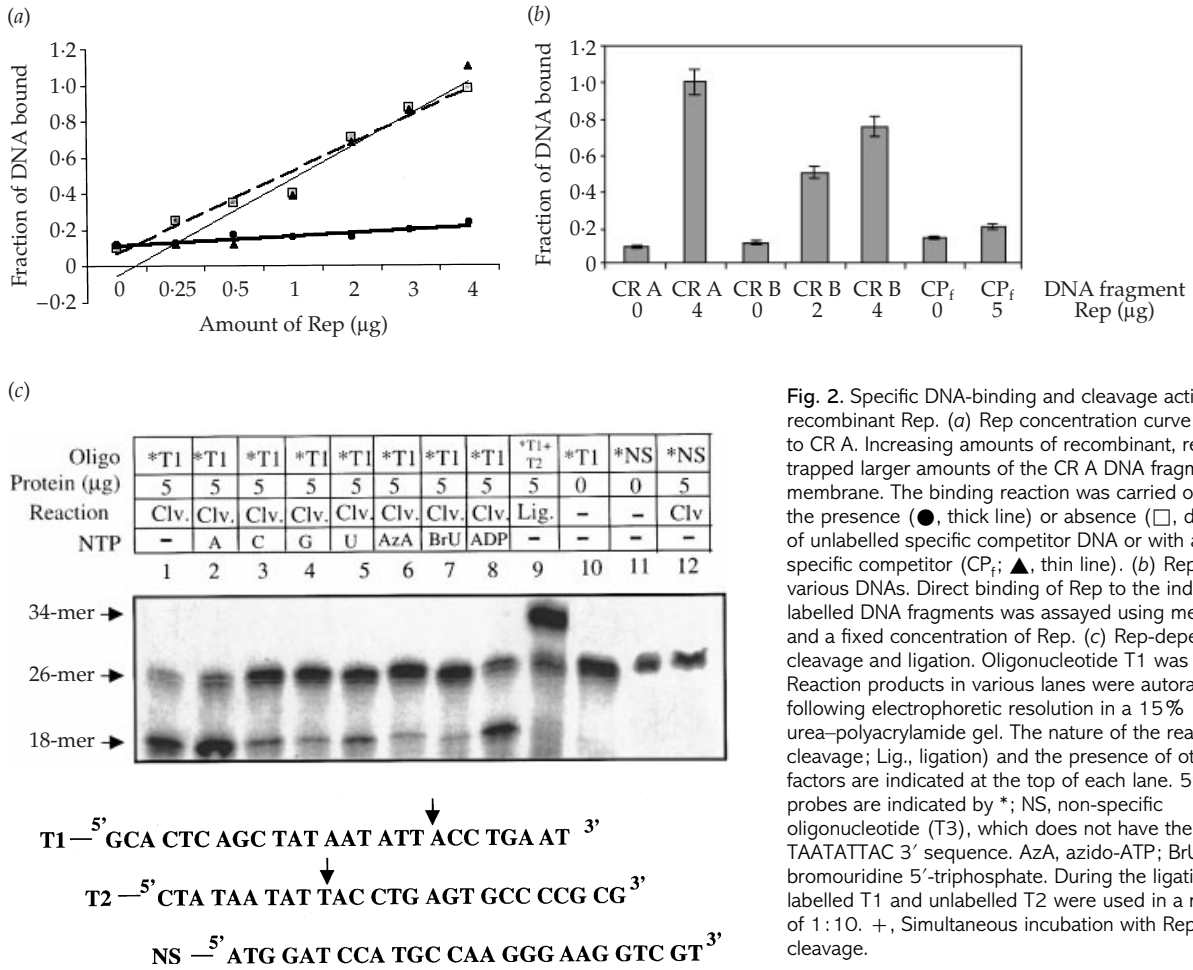

Fig. 2. Specific DNA-binding and cleavage activity of recombinant Rep. (*a*) Rep concentration curve of binding to CR A. Increasing amounts of recombinant, refolded Rep trapped larger amounts of the CR A DNA fragment on the membrane. The binding reaction was carried out in either the presence (●, thick line) or absence (□, dashed line) of unlabelled specific competitor DNA or with a non-specific competitor (CP$_f$; ▲, thin line). (*b*) Rep binding to various DNAs. Direct binding of Rep to the indicated labelled DNA fragments was assayed using membranes and a fixed concentration of Rep. (*c*) Rep-dependent cleavage and ligation. Oligonucleotide T1 was 5′-labelled. Reaction products in various lanes were autoradiographed following electrophoretic resolution in a 15% urea–polyacrylamide gel. The nature of the reaction (Clv., cleavage; Lig., ligation) and the presence of other co-factors are indicated at the top of each lane. 5′-Labelled probes are indicated by *; NS, non-specific oligonucleotide (T3), which does not have the crucial 5′ TAATATTAC 3′ sequence. AzA, azido-ATP; BrU, 5-bromouridine 5′-triphosphate. During the ligation reaction, labelled T1 and unlabelled T2 were used in a molar ratio of 1:10. +, Simultaneous incubation with Rep; ↓, site of cleavage.

43 kDa was overproduced upon induction with IPTG, but most of the induced Rep accumulated in inclusion bodies. The refolded molecules of Rep were purified using a Ni–NTA affinity matrix and these were biochemically active, as shown below.

Binding of recombinant Rep to the CR of the bipartite genome

The refolded Rep bound to the CRs of the genome (i.e. CR A and CR B) in a sequence-specific manner. To show this, 190 (CR A) and 192 (CR B) bp CRs were amplified from DNA-A and DNA-B of IMYMV, respectively. The fragments were respectively radiolabelled at the 3′ ends to specific activities of 3×10^8 and 5×10^8 c.p.m./μg. The labelled fragments were used as probes for DNA binding, employing the nitrocellulose membrane-binding technique. As a non-specific DNA-binding control, a fragment CP$_f$ (500 bp, sp. act. $\sim 8 \times 10^8$ c.p.m./μg) from the coat-protein gene (AR1) was also used.

Fig. 2(*a*) shows that gradually larger amounts of CR A were bound to the membrane with increasing concentrations of the refolded Rep protein. When the binding mixture contained a 25-fold molar excess of unlabelled CR A, there was a drastic reduction in the retention of radioactivity. However, there was no significant reduction of membrane-bound radioactivity when the binding mixtures included a 50-fold molar excess of the unlabelled coat-protein gene fragment CP$_f$. It is noteworthy that the Rep-binding curve in the presence of CP$_f$ departed from linearity, especially at low Rep concentrations. This might mean that the non-specific binding of Rep to the CP$_f$ DNA has some unusual characteristics which break down at high Rep concentrations. Fig. 2(*b*) shows that about 4 μg refolded Rep could bind 9·5 and 7·5 ng CR A and CR B, respectively. Hence, the specific binding efficiency of CR A was probably higher than that of CR B. Since there are about 72 point mismatches between CR A and CR B, the differential Rep binding of the CRs may indicate a role for flanking nucleotides around the core binding region.

Site-specific cutting and ligation by the recombinant Rep

Geminivirus Rep proteins generally cleave the phosphodiester chain between the seventh and eighth nucleotides (Fig. 2*c*; vertical arrow) of the nonamer that is conserved amongst

Table 1. Cleavage efficiency in the presence of various modulators

The cleavage efficiency in the absence of added modulators was set as 100%. Values are means ± standard deviations. BrUTP, 5-Bromouridine 5′-triphosphate.

Reagent (10 mM)	Degree of cutting (%)
None	100
ATP	> 200 ± 15
CTP	10 ± 2
GTP	5 ± 1
UTP	15 ± 2
dATP	5 ± 1
dGTP	4 ± 1
dCTP	5 ± 1
dTTP	5 ± 1
Azido-ATP	5 ± 1
BrUTP	5 ± 1
ADP	80 ± 10

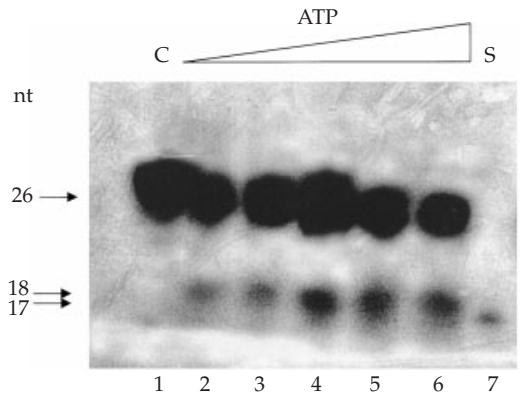

Fig. 3. ATP dependence of the cleavage reaction. Autoradiographic visualization of the cleavage product. T1 oligonucleotide was allowed to be cleaved by refolded Rep in the presence of various concentrations of ATP. C, Uncleaved T1 26-mer (lane 1). The concentrations of ATP used for the reactions shown in lanes 2–6 were 0, 5, 10, 15 and 20 mM. S, 17-mer molecular size marker (lane 7). The wet gel containing the products was autoradiographed.

all geminivirus DNA-replication origins. Besides this cleavage, Rep-mediated ligation at the same site is also required for the maturation of single-stranded viral DNA. Therefore, the cutting and ligation activities of the refolded Rep were examined separately.

Fig. 2(*c*) shows a 26-mer oligonucleotide (T1) corresponding to the IMYMV-Bg replication-origin region spanning the invariant nonamer. The refolded Rep acted on T1 at the expected site to generate a 18-mer 5′-labelled fragment (lane 1). The cutting activity was enhanced about 2-fold in the presence of ATP (Fig. 2*c*, lane 2; Table 1). In the presence of other co-factors, namely azido-ATP, deoxynucleotides and ribonucleotides other than ATP, the cleavage activity was severely inhibited (Table 1). The refolded Rep did not show any activity with a 26-mer non-specific oligonucleotide, T3, the sequence of which was taken from the M13 bacteriophage genome (lane 12), thereby demonstrating the site-specificity of the IMYMV Rep-mediated cutting function. In order to demonstrate site-specific ligation activity, refolded Rep was incubated with 5′-labelled T1 along with an unlabelled 26-mer, T2. T2 overlaps with T1 in its N-terminal sequence and should yield a longer 3′ fragment (16-mer for T2 and 8-mer for T1) when cleaved with IMYMV Rep. Following the site-specific Rep-mediated ligation between the 5′ fragment of T1 and the 3′ fragment of T2, a 34-mer was generated (Fig. 2*c*, lane 9). Since the relative molar amounts of T2 and T1 were 10:1, the majority of the 5′ label of the T1 fragment was used up to generate the labelled 34-mer. All of the site-specific cleavage and ligation assays were carried out in the presence of Mg^{2+} ions and sodium ATP was used whenever necessary.

Since ATP-stimulated cleavage is reported here for the first time, we examined the ATP concentration-dependent stimulation of IMYMV Rep-mediated cleavage. Fig. 3 clearly shows that maximum stimulation was obtained with 10 mM ATP, and that higher concentrations of ATP were not further stimulatory. Quantification of the cutting activities revealed that the mean activities with 5, 10 and 15 mM ATP were stimulated 2-, 2·75- and 2·5-fold over that found in the absence of ATP.

Site-specific topoisomerization by the recombinant Rep

It was apparent from the data shown in Fig. 2(*c*) that the refolded IMYMV Rep displayed nicking and closing activity at the same site, a property characteristic of type I topoisomerases. Hence, we examined whether the refolded Rep protein could also cause topoisomerization of a supercoiled DNA template.

The IMYMV DNA-A and DNA-B genomes were cloned in pBluescript (pBS) vector and the recombinant plasmids (pBS/A) were used in a topoisomerization assay. Lane 1 of Fig. 4(*a*) shows typical ladder formation of the above-mentioned template when treated with a control type-I topoisomerase. Similar activity was exhibited by the refolded Rep (lanes 3 and 4). In contrast, the supercoiled vector DNA alone (pBS), which lacks the site for Rep activity, remained unaffected by the Rep protein (compare lanes 7 and 8). Thus, IMYMV Rep behaves as a site-specific type-I topoisomerase.

To provide further evidence for type-I topoisomerase activity of Rep, we carried out Rep-mediated pBS/A DNA relaxation experiments in the presence of camptothecin, a specific inhibitor of type-I topoisomerases. Lane 3 of Fig. 4(*b*) shows the partial inhibitory effect of 100 μM camptothecin on a control pea type-I topoisomerase. The data shown in Fig. 4(*b*) display about 85, 25 and 10% inhibition of topoisomerase

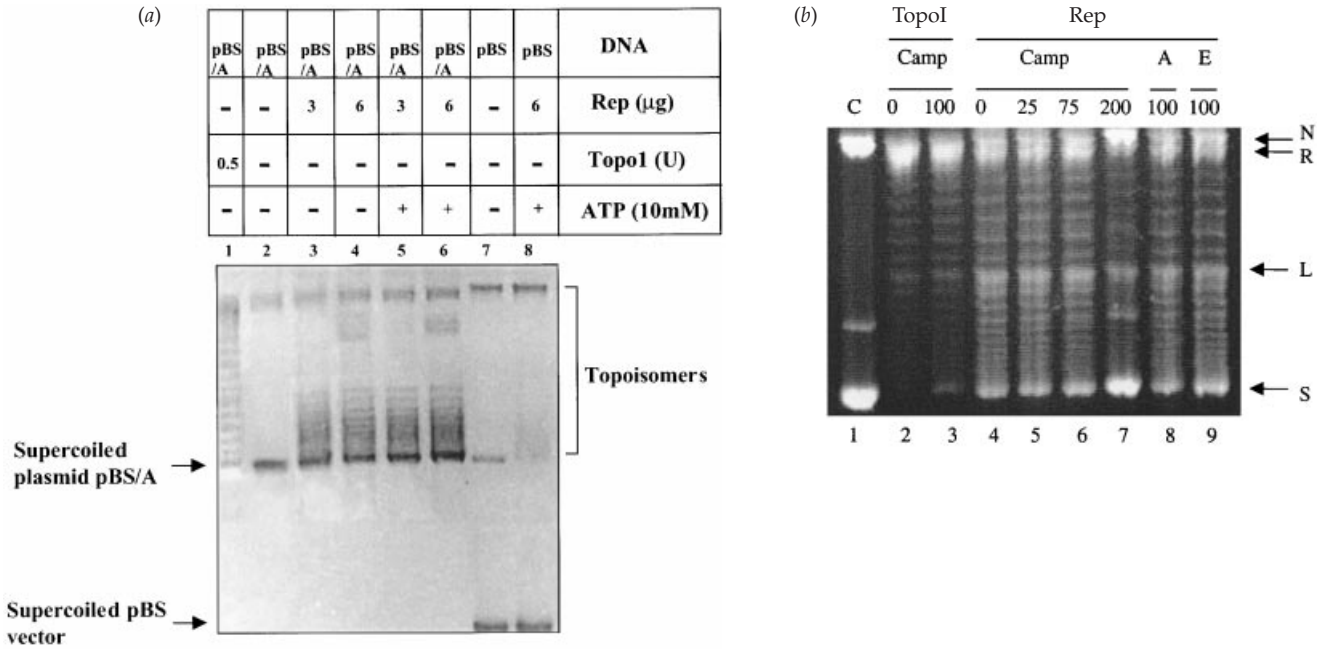

Fig. 4. Site-specific topoisomerization by Rep. (*a*) Relaxation of supercoiled DNA. The various components used in the relaxation reactions are indicated at the top of each lane. Control topoisomerase type I (TopoI) was recombinant pea nuclear topoisomerase, purified in our laboratory. The positions of the supercoiled plasmid DNA and the topoisomers are indicated. Supercoiled pBS DNA lacking the site of action of Rep was used for the reactions shown in lanes 7 and 8. (*b*) Camptothecin sensitivity of Rep-mediated topoisomerization. Topoisomerization reactions were carried out in the presence of camptothecin (Camp) or the known topoisomerase type-II inhibitors m-AMSA (A) and etiposide (E). DNA ladder formation was observed by agarose gel electrophoresis followed by staining with ethidium bromide. Inhibitor concentrations are indicated at the top (in μM). The positions of supercoiled (S), linear (L), relaxed (R) and nicked (N) forms of DNA are indicated. The various DNA forms were obtained by treating the supercoiled DNA with the appropriate DNA-modifying enzymes. Lane C, Control supercoiled pBS/A DNA.

activity of Rep in the presence of 200, 75 and 25 μM camptothecin. A comparison between lanes 7 and 4 reveals the progressive formation of nicked DNA as the concentration of camptothecin was gradually increased. Hence, the sum of the amount of nicked and supercoiled DNA was taken into account when calculating the strength of inhibition. Lanes 8 and 9 show that there was no effect on the DNA-relaxation activity of Rep in the presence of either m-AMSA (A) or etiposide (E), type II-specific topoisomerase inhibitors. The respective degrees of inhibition at 100 μM of each inhibitor were 35, 8 and 3 % for camptothecin, m-AMSA and etiposide. Generally, more linear DNA (L in Fig. 4*b*) accumulates in the presence of type-II topoisomerase and its inhibitor, but no such increment was visible with the Rep protein (compare lane 4 with either lane 8 or 9), thereby confirming its specific type-I topoisomerase activity. No further increase in linear DNA was observed when the concentration of etiposide was increased to 250 μM.

ATP-dependent conformational change

IMYMV Rep contains the GDSRTGKT^{230} motif corresponding to the P-loop of the phosphate-binding site of many ATP-hydrolysing proteins. The Rep proteins of TGMV and tomato yellow leaf curl virus (TYLCV) are known to hydrolyse ATP (Orozco *et al.*, 1997; Desbiez *et al.*, 1995). The data shown in Figs 2–4 indicate that the IMYMV Rep-mediated nicking and closing activities were altered in the presence of ATP. Hence, we wondered whether an ATP-induced conformational change of IMYMV Rep was responsible for alteration of these activities. Such ATP-dependent changes were examined using site-specific proteases as probes. The refolded protein was relatively more resistant to proteolytic digestion by either V8 proteinase (Fig. 5*a*) or trypsin (Fig. 5*b*) in the presence of ATP. The profile of protease resistance did not change when ATP was replaced by ATPγS (data not shown). Thus, it seems that ATP binding could cause a conformational change in IMYMV Rep. Addition of 10 mM ATP did not cause any change in the proteolytic pattern for BSA, used as a control protein (data not shown). Thus, the co-factor ATP did not affect proteolytic activities by itself. The profile of proteolytic cleavage also did not alter significantly with the various batches of Rep prepared. Since these Rep preparations differed from each other perhaps only in the degree of refolding, the most likely explanation for the general phenomenon of ATP-induced resistance towards proteolysis of Rep would be an ATP-mediated conformational change. However, we cannot rule out at present the possibility that ATP might aid the process of renaturation of Rep, resulting in

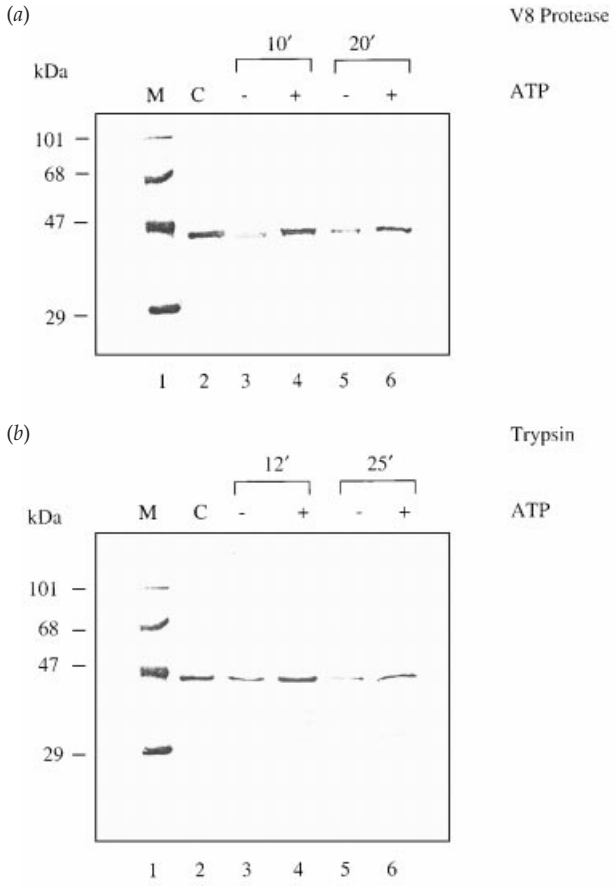

Fig. 5. Protease digestion profile of Rep. (*a*) Proteolysis with V8 protease. About 0·125 nM Rep was incubated with 25 ng/ml V8 proteinase in a 20 µl reaction at 37 °C for the indicated periods of time (in min) in either the presence or absence of 10 mM ATP. The reactions were terminated with SDS–PAGE loading dye and by boiling and the products were resolved by 12% SDS–PAGE. The proteins or proteolysed products were visualized by silver staining. M, Molecular mass standards; C, mock-treated Rep. The presence or absence of 10 mM ATP is indicated by + or −. (*b*) Proteolysis with trypsin. Assays for the reaction and the reaction products were similar to those for V8 protease. For each reaction, about 10 ng/ml trypsin was used in the presence of 20 mM $CaCl_2$. EGTA (50 mM) was used to terminate proteolysis.

the observed differences caused by ATP in the pattern of proteolysis.

Discussion

The IMYMV-Bg genome shows some unique characteristics

The genome organization of IMYMV-Bg resembled that of a typical begomovirus bipartite genome. DNA-A has five ORFs (AR1, AL1, AL2, AL3 and AL4) and DNA-B has two (BR1 and BL1). Besides these, DNA-A carries ORF AR2, as seen in monopartite and Old World geminiviruses. A small ORF, AL5, was also present within ORF AR1, but in the complementary sense, as has also been observed in other WTGs, namely TMYMV, pepper huasteco virus (Torres-Pacheco *et al.*, 1993), BGMV, barley dwarf mosaic virus and potato yellow mosaic virus. The functional importance of these two ORFs is yet to be ascertained.

A striking feature of the genome organization of IMYMV-Bg is the sequence divergence in the CRs between DNA-A and DNA-B. This feature is also shared by TMYMV, but not by VMYMV. The only example of similar divergence in CRs between components A and B comes from cabbage leaf curl virus (Hill *et al.*, 1998). This divergence might reflect the difference in intracellular copy number between DNA-A and DNA-B of IMYMV. It might also reflect the high level of variability of the viral genomes within the infected hosts.

The sequence similarity in CRs between VMYMV, TMYMV and IMYMV is interesting, considering that these genomes diverge by more than 60% in some of their corresponding gene products. Conservation at the 5′ end of the CR runs parallel with conservation in Rep sequences. The changes at the 3′ end of the CR and the N-terminal region of the coat protein were more significant than in other regions. Considering that they display less than 60% similarity in different ORFs (especially ORF AL3), these viruses could best be recognized as distinct viruses of a separate subfamily.

IMYMV Rep shows multifunctional characteristics

We have demonstrated the sequence-specific DNA-binding function of IMYMV-Bg Rep. About a 200-fold molar excess of protein over target DNA was required for saturation of binding. There are several possible explanations for this. Firstly, a fraction of the protein population could remain in an unfolded or DNA-binding-incompetent form. Secondly, the Rep protein might oligomerize following DNA binding. Thirdly, the binding template (CRs) might have multiple binding sites for Rep. In fact, there are as many as five putative Rep-binding sites within the CR of genome A. A combination of all these factors would require molar excesses of protein to bind to the target DNA, resulting in the low specific DNA-binding activity of Rep. The experiments reported in the Fig. 2 were carried out several times with the same batch of protein and the data were consistently reproducible. It is also noteworthy that the IMYMV-Bg Rep bound DNA-A more efficiently than DNA-B. The sequence differences observed between DNA-B and DNA-A may account for the relatively inefficient binding of DNA-B, though it may replicate efficiently *in vivo*.

IMYMV Rep showed site-specific nicking/closing and topoisomerase-like activities. An ATP-dependent conformational change of Rep was also suggested. The recombinant Rep also bound *in vitro* to many (at least six) factors present in nuclear extracts derived from the young leaves of Indian mungbean cultivars (data not shown). Taking all these facts together, it appears that IMYMV Rep behaves as a multifunctional protein.

The Rep proteins of all geminiviruses contain an NTP-binding domain (Hanley-Bowdoin *et al*., 1999). The purified Rep proteins of TYLCV and TGMV show ATPase activity (Desbiez *et al*., 1995; Orozco *et al*., 1997), but the ATP hydrolysis is DNA independent and is not linked to a cleavage/ligation function. Desbiez *et al*. (1995) showed that, for *in vitro* cleavage/ligation of single-stranded virus origin DNA, ATP binding or hydrolysis is not required. In our experiments, the cleavage/ligation function also did not require exogenous ATP. However, the presence of ATP enhanced Rep-mediated cleavage (Fig. 3). It is likely that ATP binding caused conformational changes that resulted in better access and binding of Rep to the substrates. Although an ATP-dependent conformational change is suggested from the data shown in Fig. 5, the mechanism of inhibition by other nucleotides (Table 1) needs to be studied in greater detail.

The nicking/closing sites of the Rep proteins of virus subgroup III are generally conserved (Bisaro, 1996; Hanley-Bowdoin *et al*., 1999). The crystal structures and consequently the mechanisms of action of nicking of some of the nicking/closing enzymes have also been reported in the literature (Stewart *et al*., 1998). Based on these comparative studies, it appears that the nicking/closing catalytic site of IMYMV Rep could be constituted by tyrosine, histidine and arginine residues (Y^{108}, H^{62}, R^{53} and R^{50}). Besides these, the substrate-binding sites could also involve some other conserved residues, such as Q^{68}, R^{79}, F^{81}, D^{82}, P^{94}, N^{95}, R^{97}, A^{99} and K^{100}. All of these important DNA-binding and nicking sites are far away from the putative ATPase domain of Rep, represented by ^{223}GDSRTGKT…DD^{276}. It is possible that there is 'cross-talk' between the DNA-nicking and ATPase domains, although they could both function independently to some extent. Following ATP binding, the geometry of the nicking and substrate-binding sites may alter in such a manner that more of the substrate DNA molecules could attach to a Rep protein and eventually get cleaved. Alternatively, turn-over of the nicking process may also be enhanced following ATP binding to IMYMV Rep. A comparison of the crystal structures of IMYMV Rep bound to and free of ATP will validate this hypothesis in future.

IMYMV Rep only showed topoisomerase activity with DNA templates containing the virus origin. Fig. 2 shows that Rep binds to and cleaves only at the CR of genome A. This suggests that the observed topoisomerase activity of Rep was probably due to the presence of the virus replication origin (or CR) in the DNA template. This notion is supported further by the fact that Rep protein could not relax supercoiled pBS DNA, which lacked the CR, at all (Fig. 4*a*). From a sequence alignment of various geminivirus Rep and rolling-circle-replication (RCR) initiator proteins, it appears that the active tyrosine residue (Y^{108}) of IMYMV Rep, which is involved in cleavage, might also be necessary for topoisomerase function. Although a topoisomerase function has been suggested previously for Rep (Bisaro, 1996), this has not previously been demonstrated *in vitro*. Thus, we present evidence for an important function of Rep for the first time.

Since the nicking action of Rep requires single-strandedness of the invariant nonamer region, the supercoiled, plasmid-borne nonamer region must be in the denatured state for the plasmid DNA to be relaxed by Rep. Denaturation *in vitro* could be influenced by a variety of factors, such as negative supercoiling of the plasmid, Rep-mediated DNA–protein interaction at the CR or even the conditions (temperature, pH etc.) of the relaxation reaction. The roles of various contributors to the denaturation process need to be assessed by employing a $KMnO_4$-footprinting technique (Mukhopadhyay & Chattoraj, 1993). Whether Rep shows topoisomerase-like activity *in vivo* and how this activity affects intracellular replication strategies and virus pathogenesis are open questions.

We thank Professors V. S. Chauhan and S. K. Sopory sincerely for their encouragement in this work. Professor S. K. Sopory made important suggestions in this work and modifications to the manuscript. We thank Ms R. Radha for her excellent secretarial assistance.

References

Arora, D. & Khanna, N. (1996). Method for increasing the yield of properly folded recombinant human gamma interferon from inclusion bodies. *Journal of Biotechnology* **52**, 127–133.

Bisaro, D. M. (1996). Geminivirus DNA replication. In *DNA Replication in Eukaryotic Cells*, pp. 833–844. Edited by M. L. DePamphilis. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Desbiez, C., David, C., Mettouchi, A., Laufs, J. & Gronenborn, B. (1995). Rep protein of tomato yellow leaf curl geminivirus has an ATPase activity required for viral DNA replication. *Proceedings of the National Academy of Sciences, USA* **92**, 5640–5644.

Hanley-Bowdoin, L., Settlage, S. B., Orozco, B. M., Nagar, S. & Robertson, D. (1999). Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. *Critical Reviews in Plant Sciences* **18**, 71–106.

Harrison, B. D. & Robinson, D. J. (1999). Natural genomic and antigenic variation in whitefly-transmitted geminiviruses (begomoviruses). *Annual Review of Phytopathology* **37**, 369–398.

Harrison, B. D., Muniyappa, V., Swanson, M. M., Roberts, I. M. & Robinson, D. J. (1991). Recognition and differentiation of seven whitefly-transmitted geminiviruses from India, and their relationships to African cassava mosaic and Thailand mung bean yellow mosaic viruses. *Annals of Applied Biology* **118**, 299–308.

Hill, J. E., Strandberg, J. O., Hiebert, E. & Lazarowitz, S. G. (1998). Asymmetric infectivity of pseudorecombinants of cabbage leaf curl virus and squash leaf curl virus: implications for bipartite geminivirus evolution and movement. *Virology* **250**, 283–292.

Honda, Y., Iwaki, M., Saito, Y., Thongmeearkom, P., Kittisak, K. & Demma, N. (1983). Mechanical transmission, purification and some properties of whitefly-borne mungbean yellow mosaic virus in Thailand. *Plant Disease* **67**, 801–804.

Hop, D. V., Gaikwad, A., Yadav, B. S., Reddy, M. K., Sopory, S. & Mukherjee, S. K. (1999). Suppression of pea nuclear topoisomerase I enzyme activity by pea PCNA. *Plant Journal* **19**, 153–162.

Laufs, J., Schumacher, S., Geisler, N., Jupin, I. & Gronenborn, B. (1995). Identification of the nicking tyrosine of geminivirus Rep protein. *FEBS Letters* **377**, 258–262.

Lazarowitz, S. G. (1992). Geminiviruses: genome structure and gene function. *Critical Reviews in Plant Sciences* **11**, 327–349.

Mandal, B., Varma, A. & Malathi, V. G. (1997). Systemic infection of *Vigna mungo* using the cloned DNAs of the blackgram isolate of mungbean yellow mosaic geminivirus through agroinoculation and transmission of the progeny virus by whiteflies. *Journal of Phytopathology* **145**, 505–510.

Morinaga, T., Ikegami, M. & Miura, K. (1990). Physical mapping and molecular cloning of mung bean yellow mosaic virus DNA. *Intervirology* **31**, 50–56.

Mukhopadhyay, G. & Chattoraj, D. K. (1993). Conformation of the origin of P1 plasmid DNA replication. Initiator protein induced wrapping and intrinsic unstacking. *Journal of Molecular Biology* **231**, 19–28.

Nariani, T. K. (1960). Yellow mosaic of mung (*Phaseolus aureus* L.). *Indian Phytopathology* **13**, 24–29.

Nene, Y. L. (1973). Viral disease of some warm weather pulse crops in India. *Plant Disease Reporter* **57**, 463–467.

Orozco, B. M., Miller, A. B., Settlage, S. B. & Hanley-Bowdoin, L. (1997). Functional domains of a geminivirus replication protein. *Journal of Biological Chemistry* **272**, 9840–9846.

Padidam, M., Beachy, R. N. & Fauquet, C. M. (1995). Classification and identification of geminiviruses using sequence comparisons. *Journal of General Virology* **76**, 249–263.

Rybicki, E. P. (1994). A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. *Archives of Virology* **139**, 49–77.

Stanley, J. (1991). The molecular determinants of geminivirus pathogenesis. *Seminars in Virology* **2**, 139–149.

Stanley, J. (1995). Analysis of African cassava mosaic virus recombinants suggests strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. *Virology* **206**, 707–712.

Stewart, L., Redinbo, M. R., Qiu, X., Hol, W. G. & Champoux, J. J. (1998). A model for the mechanism of human topoisomerase I. *Science* **279**, 1534–1541.

Swanson, M. M., Varma, A., Muniyappa, V. & Harrison, B. D. (1992). Comparative epitope profiles of the particle proteins of whitefly-transmitted geminiviruses from nine crop legumes in India. *Annals of Applied Biology* **120**, 425–433.

Torres-Pacheco, I., Garzón-Tiznado, J. A., Herrera-Estrella, L. & Rivera-Bustamante, R. F. (1993). Complete nucleotide sequence of pepper huasteco virus: analysis and comparison with bipartite geminiviruses. *Journal of General Virology* **74**, 2225–2231.

Vanitha Rani, R., Karthikeyan, A. S., Anuradha, S. & Veluthambi, K. (1996). Genome homologies among geminiviruses infecting vigna, cassava, acalypha, croton and vernonia. *Current Science* (*Bangalore*) **70**, 63–69.

Varma, A., Dhar, A. K. & Mandal, B. (1992). MYMV transmission and control in India. In *Mungbean Yellow Mosaic Disease*, pp. 8–27. Edited by S. K. Green & D. Kim. Taipei: Asian Vegetable Research and Development Centre.

Received 14 May 2001; Accepted 26 June 2001