The capsid protein of tomato yellow leaf curl virus binds cooperatively to single-stranded DNA

K. Palanichelvam,† Talya Kunik, Vitaly Citovsky and Yedidya Gafni

1 Department of Genetics, Agricultural Research Organization, The Volcani Center, PO Box 6, Bet Dagan 50250, Israel
2 Department of Biochemistry and Cell Biology, State University of New York, Stony Brook, NY 11794-5215, USA

The capsid protein (CP) of tomato yellow leaf curl virus (TYLCV) is the only known component of the virus coat. Here, we identify TYLCV CP as a single-stranded (ss) DNA binding protein. Purified TYLCV CP bound ssDNA in a highly cooperative and sequence-nonspecific fashion. TYLCV CP–ssDNA complexes were resistant to nucleolytic digestion and remained stable at relatively high salt concentrations. Because TYLCV CP is known to contain an active nuclear targeting signal, we propose that its association with the viral genomic ssDNA mediates TYLCV entry into the host cell nucleus during the infection process.

Introduction

Tomato yellow leaf curl virus (TYLCV), a major pathogen of tomato crops in the Mediterranean basin, southern Asia, Africa and Central America, is a geminivirus transmitted by the whitefly Bemisia tabaci (Cohen & Antignus, 1994). In contrast to most other known whitefly-transmitted geminiviruses, whose genomes are split between two single-stranded (ss) DNA components (DNA A and DNA B), the TYLCV genome consists of a single circular ssDNA molecule of 2787 nucleotides (Khey-Pour et al., 1991; Navot et al., 1991; Noris et al., 1994). The genome of TYLCV encodes six open reading frames (Fig. 1), one of which, V1, encodes the capsid protein (CP), which represents a building block of the virus particle. Unlike most other plant viruses, both mono- and bipartite geminiviruses need to be imported into the host-plant cell nucleus for their replication (Timmermans et al., 1994). Studies of the bipartite geminivirus squash leaf curl virus (SqLCV) indicated that a protein (BR1) encoded by the DNA B component binds the genomic viral ssDNA and shuttles it into and out of the host cell nucleus (Sanderfoot & Lazarowitz, 1995; Sanderfoot et al., 1996). Also, Noueiry et al. (1994) have shown that the BR1 protein of another bipartite geminivirus, bean dwarf mosaic geminivirus (BDMV), is involved in the nuclear export of viral DNA. Monopartite geminiviruses, such as TYLCV, lack the DNA B genomic component and do not encode a BR1 protein homologue. Thus, the mechanism(s) by which TYLCV enters and exits the host cell nucleus is largely unknown.

Recently, we have shown that the TYLCV CP contains at least one functional nuclear localization signal (NLS) sequence which mediates the import of this protein into the nuclei of plant and insect cells (Gafni et al., 1997; Kunik et al., 1998). Based on these results, we suggested that TYLCV CP may function similarly to the BR1 protein of SqLCV and BDMV, transporting the viral genomes across the nuclear envelope.

Fig. 1. Genome organization of TYLCV. Arrows indicate all open reading frames within the TYLCV genomic DNA that have a capacity to encode proteins in excess of 10 kDa. The V1 gene encodes the 30-3 kDa TYLCV CP.

Author for correspondence: Yedidya Gafni.
Fax +972 3 966 9642. e-mail vcgafni@netvision.net.il

† Present address: Department of Microbiology, University of Illinois, Urbana–Champaign, Illinois, USA.
This model implies that TYLCV CP must associate directly with ssDNA. In the present study, we tested this idea and demonstrated that TYLCV CP binds ssDNA but not double-stranded (ds) DNA in vitro. TYLCV CP binding to ssDNA was cooperative and resulted in the formation of protein–DNA complexes resistant to nucleolytic digestion.

Methods

Expression of the TYLCV CP gene in E. coli. To produce the large amounts of TYLCV CP required for in vitro DNA binding experiments, the CP gene, V1, was subcloned from the full-length TYLCV genome (pTYH 19; Kunik et al., 1994) into the bacterial expression vector pET28 (Novagen). To this end, TYLCV genome (pTYH 19; Kunik et al., 1994) was subcloned into the upstream region of the V1 gene of the expression vector pET28 (Novagen). At this point, V1 was cloned as a HindIII–Bgl II restriction fragment into pBR328 (Boehringer). An Ndel restriction site was introduced into the upstream region of the V1 gene by PCR. To minimize errors during amplification, PCR primers were designed to amplify only a 589 bp region between two Apal restriction sites in the V1 nucleotide sequence. The sense primer corresponded to the V1 sequence from nucleotides 458 to 476 (5′ GATAAAGGGCGCTGGACATATG 3′) and the antisense primer was complementary to V1 nucleotides 1035 to 1054 (5′ CCACAGGCCCACCAATAAC 3′) (Navot et al., 1991). In the sense primer, a single adenine insertion (double-underlined) formed a new Ndel restriction site adjacent to the translation initiation codon of the TYLCV CP (underlined); both primers included Apal restriction sites (dotted underline).

The resulting PCR product was subcloned into pGEMT with the TA cloning kit (Promega), and its nucleotide sequence was verified by dideoxynucleotide DNA sequencing. Next, this PCR-derived portion of the V1 gene containing the introduced Ndel restriction site was removed from pGEMT as an Apal restriction fragment and exchanged with the wild-type V1 Apal restriction fragment in pBR328. From this construct, an Ndel–Bgl II restriction fragment carrying the full-length V1 gene was removed and inserted between the Ndel and BamHI restriction sites of pET28 to generate the plasmid pETCP.

TYLCV CP was produced in E. coli using the T7 RNA polymerase over-expression system. In this approach, pETCP was transfected into E. coli strain BL21(DE3) carrying the plasmid pLysE (Novagen), essential for pETCP stability. T7 RNA polymerase expression was then induced by addition of IPTG to liquid cultures of cells harbouring the pETCP plasmid, as described by Citovsky et al. (1990). Purification of the over-expressed TYLCV CP and its subsequent analysis by SDS–PAGE were performed as described by Citovsky et al. (1990). Purified TYLCV CP was stored in binding buffer (10 mM Tris–HCl, pH 8.0, 10% glycerol, 30 mM NaCl) at −70 °C until required.

Preparation of radioactively labelled DNA probe. The virus-specific DNA probe included most of the intergenic region and part of the V1 gene of the TYLCV genome (Navot et al., 1991). This 422 nucleotide DNA fragment was synthesized by PCR using sense (5′ TGCCTTACTTATATCGGAC 3′) and antisense (5′ TAGACACATTCCCCGCGACTGA 3′) primers corresponding to the TYLCV nucleotide sequences between positions 51–70 and the complement of 451–473, respectively (Navot et al., 1991). The agarose gel-purified DNA fragments were end-labelled with [γ-32P]ATP by T4 polynucleotide kinase (Promega), followed by removal of unincorporated radio-nucleotides, as described by Sambrook et al. (1989).

Gel mobility shift assay. Binding of TYLCV CP to ssDNA and dsDNA was assayed by its ability to decrease the electrophoretic mobility of the DNA probe. Various amounts of TYLCV CP were incubated for 30 min at 4 °C in 50 µl binding buffer with 5 ng ds or heat-denatured, ss probe DNA. After incubation, samples were electrophoresed on a 4% native polyacrylamide gel and then characterized by autoradiography (Citovsky et al., 1990). TYLCV CP–ssDNA binding was quantified by analysis of the gel area corresponding to the protein–ssDNA complex with a Molecular Dynamics phosphorimager. For proteinase K treatment, 1 mg/ml enzyme was added to the TYLCV CP–DNA mixture, and the mixture was incubated for 30 min at 37 °C prior to gel electrophoresis. The effect of salt concentration on TYLCV CP–ssDNA binding was assayed by supplementing the binding mixtures with increasing amounts of NaCl.

S1 nuclease protection assay. The DNA probe was labelled with [32P]dCTP by using random hexanucleotides (Pharmacia) to prime DNA synthesis on denatured DNA strands. TYLCV CP (15 ng) was incubated with the radioactively labelled probe (5 ng), as described for the gel mobility shift assay. The incubation mixture was then diluted and supplemented with 250 U S1 nuclease and 10× S1 reaction buffer (300 mM sodium acetate, pH 4.5, 2.5 M NaCl, 10 mM ZnSO4, 50% v/v glycerol). After S1 nuclease digestion at 25 °C for 30 min, the reaction was terminated by addition of 10 µl 0.5 EDTA. The reaction products were analysed by the gel mobility shift assay, followed by phosphor-imager analysis.

Results

Expression and purification of TYLCV CP

To test its possible nucleic acid binding properties, TYLCV CP was over-expressed in E. coli cells carrying a plasmid, pETCP, which contains the strong T7 RNA polymerase promoter linked to the coding sequence of TYLCV CP. Induction of T7 RNA polymerase expression by IPTG resulted in formation of TYLCV CP inclusion bodies which could be solubilized in 4 M urea and which remained soluble after the urea had been removed (data not shown). Similar solubilization of proteins produced in the T7 RNA polymerase system has been reported for the cell-to-cell movement proteins of tobacco mosaic virus (Citovsky et al., 1990) and cauliflower mosaic virus (Citovsky et al., 1991). The urea solubilization protocol (Citovsky et al., 1990) resulted in preparation of a protein that was soluble and more than 90% pure (Fig. 2, lane 2), yielding about 0.8 mg TYLCV CP per 20 ml bacterial culture.

DNA binding properties of TYLCV CP

Binding of TYLCV CP to ssDNA and dsDNA was assayed by its ability to retard a radioactively labelled DNA probe (a 422 bp fragment of TYLCV DNA containing nucleotides 51–473; Navot et al., 1991) on a native polyacrylamide gel. Fig. 3 shows that ssDNA (obtained by heat denaturation of the dsDNA probe) incubated with purified TYLCV CP was strongly retarded (lane 2, open arrowhead) compared with the free ssDNA probe (lane 1). The large difference in migration between the free and protein-bound radioactive probes suggested that more than one molecule of TYLCV CP was bound per molecule of ssDNA (see also below). Treatment of the reaction mixture with proteinase K abolished the binding.
ssDNA binding activity of TYLCV capsid protein

Fig. 2. SDS-PAGE of purified TYLCV CP. TYLCV CP was expressed in E. coli using the T7 RNA polymerase system as described in Methods, and analysed on a 12.5% SDS–polyacrylamide gel followed by Coomassie Blue staining. M, molecular mass standards; CP, purified TYLCV CP. Numbers to the left indicate sizes of molecular mass markers (kDa).

Fig. 3. Interaction of TYLCV CP with ssDNA and dsDNA. Purified TYLCV CP (100 ng) was incubated with ssDNA or dsDNA probes and protein–DNA binding was analysed by gel mobility shift assay as described in Methods. Lane 1, ssDNA probe incubated in binding buffer alone; lane 2, ssDNA probe incubated with TYLCV CP; lane 3, ssDNA probe incubated with TYLCV CP in the presence of a 1000-fold molar excess of unlabelled ssDNA probe; lane 4, ssDNA probe incubated with TYLCV CP in the presence of a 1000-fold molar excess of unlabelled ssDNA; lane 5, ssDNA probe incubated with TYLCV CP and treated with proteinase K as described in Methods; lane 6, dsDNA probe incubated in binding buffer alone; lane 7, dsDNA probe incubated with TYLCV CP. Open arrowhead indicates the position of the TYLCV CP–ssDNA complex; filled arrowheads indicate the positions of the free probes.

(Fig. 3, lane 5), indicating that the retarded ssDNA probe indeed represented a protein–DNA complex. Unlike ssDNA, dsDNA was a poor substrate for TYLCV CP binding, since retardation of dsDNA incubated in the presence of TYLCV CP was hardly detectable (Fig. 3, lane 7) when compared with the protein-free dsDNA (Fig. 3, lane 6). Thus, DNA binding by TYLCV CP was specific for ssDNA.

TYLCV CP association with the radioactively labelled ssDNA probe was blocked by the addition of a 1000-fold molar excess of unlabelled probe (Fig. 3, lane 3). Interestingly, the presence of this competitor DNA increased the degree of probe renaturation, as illustrated by the elevated level of dsDNA in the reaction mixture (Fig. 3, lane 3, bottom band). That TYLCV CP–ssDNA binding was not sequence-specific was indicated by the finding that unlabelled ssDNA competitor derived from a non-viral source (a 305 bp EcoRI restriction fragment of pKK232-8; Pharmacia) efficiently inhibited TYLCV CP association with the radioactively labelled viral ssDNA probe (Fig. 3, lane 4).

TYLCV CP cooperatively binds ssDNA

Sequence-nonspecific binding of TYLCV CP and strong retardation of the ssDNA probe suggest that multiple copies of this protein may bind along the entire length of the DNA molecule. To provide further insight into the binding mode of TYLCV CP, we studied the effect of protein concentration on ssDNA binding. Gel mobility shift analysis of ssDNA probe incubated with increasing amounts of TYLCV CP suggested that the protein–ssDNA binding was cooperative (Fig. 4a). At sub-saturating concentrations of protein, i.e. between 10 and...
50 ng (Fig. 4a), only two species of ssDNA probe were detected: free probe and completely retarded probe. The absence of intermediate bands representing probe molecules only partially coated with TYLCV CP is characteristic of cooperative protein binding (Lohman et al., 1986; Citovsky et al., 1986).

Interaction of TYLCV CP with ssDNA was quantified by means of phosphorimagery analysis of the protein–ssDNA complexes resolved on native polyacrylamide gels. A saturation curve of TYLCV CP binding to ssDNA was obtained from the gel mobility shift experiments shown in Fig. 4(a). Fig. 4(b) shows that saturation of limiting ssDNA by increasing concentrations of TYLCV CP was achieved at a protein-to-ssDNA weight ratio of 10:1. Since the molecular mass of the probe (422 nucleotides) is approximately 126·6 kDa and that of TYLCV CP is 30·3 kDa, this weight ratio corresponds to 42 TYLCV CP monomers per ssDNA molecule. Thus, the minimum size of the binding site corresponding to this ratio is about 10 nucleotides per TYLCV CP monomer.

**TYLCV CP protects ssDNA from S1 nuclease digestion**

To examine the ability of TYLCV CP to protect its cognate ssDNA from nucleolytic digestion, we formed TYLCV CP–ssDNA complexes at a sub-saturating protein-to-ssDNA ratio and treated the resulting mixtures with S1 nuclease. Following separation on a native polyacrylamide gel, the amounts of free and protein-bound ssDNA were quantified using the phosphorimager. Fig. 5(a) shows that in the absence of S1 nuclease ssDNA incubated alone migrated as a non-retarded band (column A, open bar), with negligible signal at the shifted position (column A, filled bar). In contrast, ssDNA incubated with a subsaturating amount of TYLCV CP in the absence of S1 nuclease was partitioned between free, non-retarded probe (column B, open bar) and a shifted TYLCV CP–ssDNA complex (column B, filled bar). Under the same conditions, S1 treatment completely digested the free probe (columns C and D, open bars) whereas the protein–ssDNA complex remained intact (column D, filled bar). Protection of ssDNA by TYLCV CP is consistent with cooperative protein binding to coat the DNA molecule fully (reviewed by Chase & Williams, 1986).

The strength of TYLCV CP–ssDNA binding was also reflected by the concentration of salt required to inhibit formation of the protein–nucleic acid complex. Fig. 5(b) shows that 50% inhibition was achieved at 0·4 M NaCl whereas concentrations greater than 0·6 M almost completely prevented binding of TYLCV CP to ssDNA. These data suggest that the TYLCV CP–ssDNA complex is stable under physiological (0·1–0·2 M) salt concentrations.

**Discussion**

Geminiviruses are thought to replicate their ssDNA genomes in the nuclei of host cells (Goodman, 1981; Davies & Stanley, 1989). Bipartite geminiviruses have been shown to encode a specific nonstructural protein, BR1, that shuttles the viral genome into and out of the nucleus (Noueiry et al., 1994; Sanderfoot & Lazarowitz, 1995; Sanderfoot et al., 1996). Since TYLCV, a monopartite gemivirus, lacks BR1, another viral protein must mediate nuclear transport of the TYLCV genome. One candidate for this function is TYLCV CP. Indeed, we have shown that TYLCV CP is a karyophilic protein which is actively transported into the nuclei of plant cells (Gafni et al., 1997; Kunik et al., 1998).

The present results strongly support our idea of TYLCV CP-mediated nuclear import of the viral ssDNA genome. In this model, TYLCV CP is expected to bind to the transported ssDNA molecule, providing it with an NLS. Indeed, TYLCV CP exhibited a high binding affinity toward ssDNA molecules but displayed only negligible levels of association with dsDNA. Although we did not detect a specific interaction of TYLCV CP with TYLCV-derived ssDNA, this interaction cannot, as yet, be ruled out. It is possible that particular conditions exist in vivo that favour specific binding of TYLCV CP to the viral genomic DNA. Alternatively, the lack of sequence specificity may play a role in facilitating the nuclear import process by allowing the protein to coat the entire
ssDNA binding activity of TYLCV capsid protein

ssDNA molecule. Consistent with this notion, TYLCV CP binding to ssDNA was cooperative, probably resulting in a large number of NLS per TYLCV CP–ssDNA complex. Multiple NLS have been hypothesized to increase the probability of interaction of the transported molecule with the nuclear import machinery (Gerace & Burke, 1988).

In addition to targeting the bound ssDNA into the cell nucleus, TYLCV CP may function to protect the transported nucleic acid molecule from intracellular nucleases. In fact, TYLCV CP–ssDNA complexes were highly resistant to S1 nuclease activity in vitro. Collectively, the results presented in the present work suggest that TYLCV CP plays a direct role in viral nuclear entry by associating with the TYLCV genomic ssDNA, protecting it from nucleolytic degradation, and supplying it with NLS. Interestingly, nuclear import of DNA cooperatively coated with multiple copies of an NLS-containing protein may represent a generalized process of nuclear uptake of nucleic acid. For example, VirE2, a ssDNA binding protein of Agrobacterium tumefaciens, has been shown to bind ssDNA directly and to transport ssDNA into the cell nucleus via its NLS (Guralnick et al., 1996; Zupan et al., 1996). While this model describes how TYLCV ssDNA is imported into the host cell nucleus, the mechanism of export of the virus genome from the nucleus, following its replication, remains obscure. However, it is tempting to speculate, by analogy to the bipartite geminivirus BR1 nuclear shuttle protein, that TYLCV CP functions during both the import and the export processes.

This work was supported by the U.S.–Israel Binational Agricultural Research and Development Fund (BARD) grant US-2247-93 to Y. G. and V. C., and by grants from the U. S. National Institutes of Health (ROI-GM50224) and the U.S. Department of Agriculture (94-02564) to V. C. We also wish to thank the Israeli Foreign Ministry for a Postdoctoral Fellowship awarded to K.P. This paper is a contribution from the Agricultural Research Organization, The Volcani Center, Bet Dagan, Israel, No. 126/98, 1998 series.

References


Received 26 May 1998; Accepted 22 July 1998