Agrobacterium tumefaciens virE operon encodes a single-stranded DNA-binding protein

Anath Das

Department of Biochemistry and Plant Molecular Genetics Institute, 1479 Gortner Avenue, University of Minnesota, St. Paul, MN 55108

Communicated by Donald R. Helenki, December 4, 1987 (received for review August 10, 1987)

ABSTRACT The virulence (vir) genes of Agrobacterium tumefaciens Ti plasmids are essential for transformation of plant cells. Overproduction of a virE-encoded gene product in Escherichia coli was achieved by construction of an operon fusion with the E. coli tryptophan (trp) operon. The virE2 gene product in E. coli partitioned into the insoluble membrane fraction. The protein was solubilized by treatment with 4 M urea at 0°C. DNA–protein binding experiments showed that a strong single-stranded (ss) DNA-binding activity was present in protein fractions containing the virE2 gene product. The binding was highly specific with little or no binding observed with either double-stranded DNA or ssRNA. No significant binding to Ti plasmid DNA sequences was observed. Protein blotting studies indicated that the ssDNA-binding activity was associated with the 68-kDa virE2 polypeptide.

Agrobacterium tumefaciens, a phytopathogenic soil bacterium, causes crown gall tumors on most dicotyledonous plants (1). During infection, agrobacteria donate a 15- to 20-kilobase (kb) segment of Ti-plasmid-borne DNA to plant cells. The transferred DNA (T-DNA) becomes integrated into the plant nuclear genome and is maintained stably thereafter (reviewed in ref. 2). Two cis-acting border sequences consisting of 24-base-pair (bp) imperfect direct repeats are required for Agrobacterium-mediated DNA transfer to plant cells (3–7). Recent studies demonstrate that under induction conditions, site-specific and strand-specific, cleavages occur within these border sequences. These cleavages lead to the formation of a single-stranded (ss) DNA molecule composed of the T (transferred) strand of T-DNA (8, 9). The ssDNA molecules are likely intermediates in the DNA transfer process.

In addition to the T-DNA segment, two other regions are absolutely essential for tumor formation (2). One of these is encoded in the bacterial chromosome, while the other is Ti-plasmid encoded (10–12). The Ti-plasmid encoded virulence (vir) region is located left of the T-DNA, encompasses ~35 kb of DNA, and contains six complementation groups (2, 13). Two loci, virA and virG, are regulatory in nature and are transcriptionally active in free-living bacteria. The other four loci, virB, -C, -D, and -E, are strongly induced when bacteria and plant cells are cocultivated together (13, 14). These loci are believed to encode enzymes that catalyze the transfer and integration of T-DNA into the plant nuclear genome. The role of most of these gene products in the DNA transfer process is not clear at present. Recent studies indicate that one or two of the virD-encoded polypeptides catalyze site-specific cleavages at the T-DNA borders (15).

A second operon, virE, has been investigated in some detail. DNA sequence analysis showed that the virE operon can encode two polypeptides with apparent masses of 7.0 and 60.5 kDa, respectively (16). Genetic complementation studies revealed that a strain bearing mutation in the virE locus can be complemented to form tumors on test plants if coinfected with a second Agrobacterium carrying wild-type vir loci but lacking T-DNA (17). The complementing strain must contain four other vir loci—namely, virA, virB, virD, and virG—in addition to the virE locus. In transient assays, where integration of T-DNA was not obligatory, virE was found to be a nonessential locus (18). The possibility that a low frequency of integration events leads to the observations of Gardner and Knauf (18) cannot be completely ruled out. These findings indicate that while virE is not absolutely necessary for DNA transfer it is essential for T-DNA integration. These studies, in conjunction with the possibility that ssDNA may be an intermediate in DNA transfer, prompted this investigation of the possibility that virE encodes a DNA-binding protein. In this report, I demonstrate that the larger polypeptide (virE2 in ref. 16) encoded by the virE operon specifically binds ssDNA.

MATERIALS AND METHODS

Plasmids and Strains. The virE operon of Agrobacterium Ti-plasmid pTiA6 is encoded within a 3.2-kb XhoI fragment (16). Plasmid pSW108 is a pUC7 derivative containing this XhoI fragment cloned into a SalI site of the vector (16). To isolate the virE2 coding segment, plasmid pSW108 was first digested with the enzyme ThII. After filling in with Klenow enzyme, the DNA sample was ligated with BamHI linkers (dCCGGATCCGG). After heat inactivation of DNA ligase, the reaction mixture was digested with the enzyme BamHI. The DNA was extracted with phenol and chloroform, precipitated with ethanol, dried, and resuspended in water. The BamHI fragment containing the virE2 coding region was then cloned into the BamHI site of plasmid pATH2 to give plasmid pAD1075. Plasmid vector pATH2, an expression vector, contains the Escherichia coli tryptophan (trp) operon promoter-operator region, and a large segment of the trpE coding sequence, followed by multiple cloning sites (19).

Plasmid pAD1082 was constructed by linearizing plasmid pAD1075 at its unique BglII site followed by a filling-in reaction with Klenow enzyme and recircularization with T4 DNA ligase. Plasmids pAD1075, pAD1082, and pATH2 were introduced into E. coli strain MO412 (recA le1392) by transformation to yield strains AD1075, AD1082, and ATH2, respectively.

Plasmid pAD1012 was constructed by cloning a 476-bp HpaI/NruI restriction fragment (bp 13,800–14,276 of Barker et al. (20)) containing the right border sequences of pTiA6 into the HincII site of vector pUC18 (21).

Overproduction of virE2 Gene Product in E. coli. Cells were grown in M9 media containing 0.2% glucose, 0.5% acid-
hydrolyzed casein, L-tryptophan (20 μg/ml), and ampicillin (100 μg/ml). To induce transcription from the trp promoter, an overnight culture was grown in medium containing tryptophan. The cells were collected by centrifugation, washed twice with M9 medium, and resuspended in an equal volume of the same medium. One milliliter of culture was then used to inoculate 100 ml of medium without tryptophan. After growth of cells to A900 0.5–0.6 unit, 100 μl of indole acetic acid (20 mg per ml of stock solution in alcohol) was added to the flask and cells were grown for an additional 2 hr. Uninduced cells were grown in the presence of L-tryptophan (20 μg/ml) at all times. After growth, cells were cooled in ice water, harvested by centrifugation at 5000 × g for 10 min, washed twice with cold 0.8% sodium chloride, and stored frozen at −70°C until further use.

**Partial Purification of the virE2 Gene Product.** Induced AD1075 cells were thawed and resuspended (2 ml per g of cells) in buffer A [20 mM Tris-HCl, pH 8.0/50 mM NaCl/10% (vol/vol) glycerol/1 mM dithiothreitol/50 μM EDTA (DETA)]. Cells were lysed in a French pressure cell at 10,000 psi (1 psi = 6.89 kPa). The lysate was adjusted to 0.5 M NaCl by addition of 5 M NaCl and incubated in ice for 45 min. The cell lysate was centrifuged at 5000 rpm for 5 min in a JA-14 rotor to pelleted cells. The supernatant was transferred to another tube and centrifuged at 12,000 rpm for 15 min. The pellet containing virE2 protein was washed twice with buffer A and then resuspended in 1 ml of buffer A. An equal volume of buffer A containing 8 M urea was added and mixed briefly. The mixture was incubated on ice for 15 min with occasional swirling. After centrifugation at 12,000 rpm for 15 min, the supernatant was collected and dialyzed overnight against buffer A with one change of buffer. The sample was centrifuged to clarify, made 50% in glycerol, and stored at −20°C (fraction II preparation). Most of the virE2 protein was solubilized in this manner. In control experiments, uninduced AD1075 and induced AT2H cells were processed in similar fashion.

**DNA–Protein Binding Assay.** The assay mixture (20 μl) contained 20 mM Tris-HCl (pH 8.0), 50 mM NaCl, 50 μM EDTA (pH 8.0), 1 mM dithiothreitol, bovine serum albumin (50 μg/ml), 2–3 ng of heat-denatured [32P]DNA (see below), 100 ng of sonicated double-stranded (ds) E. coli DNA, and protein fraction, as indicated. After incubation for 15 min at 26°C, 2 μl of dye mix containing 50% (vol/vol) glycerol and bromophenol blue marker dye was added to the reactions, mixed, and loaded immediately onto 5% polyacrylamide gel (in 0.5 × TBE) (22). The gels were electrophoresed for ~2 hr at 400 V, dried, and autoradiographed. The radiolabeled DNA used in these assays was prepared by digestion of plasmid pAD1012 DNA with EcoRI, HindIII, and Pvu II, followed by filling in with [γ-32P]ATP and Klenow enzyme.

**NaDodSO4/PAGE and Protein Blotting.** Proteins were separated on 12.5% NaDodSO4/polyacrylamide gels according to Laemmli (23). Where indicated, gels were stained with Coomassie brilliant blue R. To identify DNA-binding proteins, after separation on NaDodSO4/polyacrylamide gels, the proteins were blotted onto nitrocellulose filters by electroblotting procedures (24; E-C Apparatus, Saint Petersburg, FL). The filters were prehybridized with 5% nonfat dry milk in 10 mM Tris-HCl, pH 8.0/1 mM dithiothreitol for 1 hr at room temperature. Hybridization was for 1 hr at room temperature in 0.25% nonfat dry milk, 20 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM dithiothreitol, 50 μM EDTA, and 102 cpm (Cerenkov) per ml of 32P-labeled probe (specific activity 1–4 × 106 cpm/μg) (25). The filters were washed twice for 15 min in the hybridization buffer, dried, and autoradiographed.

**Other Methods.** Antibody against virE2 polypeptide was raised in rabbits by subcutaneous injection with 100 μg of protein followed by a second injection after 2 weeks. Ten days later, the rabbit was bled and serum was collected. The protein used for injection was isolated by excision of the band from NaDodSO4/polyacrylamide gels.

DNA probe was prepared by the method of Feinberg and Vogelstein (26). RNA probe was prepared by transcription from Sph I-digested pGEM1 DNA in the presence of T7 RNA polymerase (Promega Biotec, Madison, WI). Protein concentrations were determined by the Bradford assays (27). E. coli strain MO412 was a gift of Marc Orbach (Stanford University, Stanford, CA).

## RESULTS

**Overproduction of virE2 Gene Product.** To facilitate studies on the function of the virE gene products in Ti-mediated gene transfer, a recombinant plasmid pAD1082 was constructed that contains a trpE′-virE2 operon fusion (Fig. 1). In pAD1075, the coding region of the virE2 gene was cloned downstream from the strong, regulatable trp promoter in a manner such that the trpE′ operon reading frame terminates translation 4 bp upstream of the virE2 translation initiation site. When induced with indole acetic acid, cells containing plasmid pAD1075 synthesized two polypeptides with apparent masses of 68 and 38 kDa (compare lanes 1 and 3 in Fig. 2A). The 38-kDa protein is the truncated trpE′ gene product, which serves as an internal marker to monitor induction of the trp operon. The 68-kDa band is presumably the virE2 gene product. This size agrees well with data obtained from maxi-cell experiments (16). Neither protein was present in significant amounts in extracts of cells induced for 30 min (lanes 2 and 5) but was easily visible after 90 min of induction (lanes 3 and 6). To ensure that the 68-kDa band is indeed synthesized from the virE coding segment, plasmid pAD1082 was constructed by filling in a unique Bgl II site located within the virE2 coding region (Fig. 1). Alteration of the Bgl II site results in a frame-shift mutation. The truncated virE2 polypeptide (virE′) in this mutant should be 196 amino acid residues in length. Cell extracts prepared from induced cells containing plasmid pAD1082 did not synthesize the 68-kDa polypeptide product (Fig. 2A, lane 6) but instead synthesized a band of ~22 kDa, which is the expected molecular mass of virE′ if translation initiated at the virE2 translation initiation site.

**Localization of virE2 in E. coli.** In induced E. coli cells, the virE2 protein was always found associated with the insoluble fraction (Fig. 2B, lane 4). None was detectable in the soluble fraction (lane 3). Attempts to solubilize the protein by using nonionic detergents—e.g., deoxycholate, Triton X-100, Nonidet P-40, etc.—were unsuccessful (data not shown). Incubation of the pellet fraction with 4 M urea at 0°C resulted in solubilization of significant amounts of the virE2 polypeptide. After removal of urea by dialysis, the virE2 polypeptide remained in the soluble fraction (lane 5). When antibody raised against the purified virE2 polypeptide was used to probe total Agrobacterium proteins in an immunoblot experiment (28), a single protein band that comigrated with the protein produced in E. coli cells containing the cloned virE gene was recognized by the antiserum (see Fig. 5C and data

**Fig. 1.** Physical map of trpE′–virE2 operon fusion in pAD1075. The virE2 coding sequence is shown by shaded area (not to scale). Translation of trpE′ protein terminates within the insert DNA 4 bp preceding the virE2 start codon. Arrows indicate direction of transcription. Location of the trp promoter–operator (P/O), trpE′, ampicillin resistance gene (amp), and recognition sites of restriction enzymes BamHI and BglII are also indicated. The Bgl II site is located 0.56 kb downstream of the virE2 start codon (16).
Fig. 2. Analysis of proteins by NaDodSO₄/PAGE. (A) Total proteins (50 μg) from E. coli AD1075 (lanes 1-3) or AD1082 (lanes 4-6) were analyzed on 12.5% NaDodSO₄/polyacrylamide gels. Samples in lanes 1 and 4 were from uninduced cells, while others were from cultures induced with indole acrylic acid either for 30 min (lanes 2 and 5) or for 90 min (lanes 3 and 6). The location of the virE, trpE, and virE' bands are indicated. Numbers (in kDa) indicate the position of marker proteins. (B) Purification of virE2 protein: All samples were from cells induced for 2 hr with indole acrylic acid except that from lane 1 was from uninduced cells. Lanes 2-5, crude extract, cytosol fraction, pellet fraction, and fraction II preparation, respectively (see text for details).

not shown). The antigen was present only in the cytosolic fraction of a Ti-plasmid-containing strain A348 that was induced with plant cells. Thus, it appears that fractionation of the virE2 polypeptide into the E. coli membrane may be an artifact of overproduction as has been observed in a number of cases (29).

**DNA–Protein-Binding Studies.** To assess whether the induced AD1075 cell extract contained any nucleic acid-binding activity, gel retardation experiments were performed (Fig. 3). Fraction II from induced AD1075 cells showed a strong ssDNA-binding activity (lanes 4–7). In the presence of induced AD1075 extract, all radiolabeled DNA fragments (bands a–d) bound to proteins and migrated very slowly at the top of the gel. No binding to either dsDNA or ssRNA was observed in similar experiments (data not shown). In control experiments, little or no DNA-binding activity was observed in fraction II preparations from uninduced AD1075 cells (lanes 2 and 3) or induced ATH2 cells (lanes 8 and 9).

Of the four radiolabeled fragments, b is a 510-bp restriction fragment containing a 476-residue T-DNA sequence (coordinates 13,800–14,276 in ref. 20) that includes the right border region of the octopine-type Ti-plasmid pTiA6, and fragments a, c, and d are sequences derived from the plasmid vector pUC18. In the binding experiment, all four fragments bound to proteins as apparent from the significant loss in intensity of these bands and the appearance of slower-moving bands (Fig. 3, lanes 4–7).

To determine whether the ssDNA-binding protein present in induced AD1075 extract had any sequence specificity, different amounts of proteins were used in binding experiments (Fig. 3). An increase in binding with increasing amounts of protein was apparent from the appearance of a ladder-like pattern and the retention of all radioactivity at the top of the gel (lanes 4–7). In all cases, both Ti-plasmid-specific sequence (band b) and nonspecific vector sequences (bands a, c, and d) were retarded in a similar fashion. No preferential retardation of a fragment was apparent in these studies. Binding to all fragments was found to be very strong, as no significant reduction in binding was observed even at high salt concentration (0.3 M NaCl) (data not shown).

Affinity for specific DNA sequences was also assessed by competition experiments with a homologous right-border-specific DNA fragment (band b) (Fig. 3, lanes 11–14) and nonspecific E. coli DNA (lanes 15–18) as competitor DNAs in a gel retardation experiment. Both T-DNA and E. coli DNA competed effectively, as apparent from the reappearance of bands a–d with increasing concentration of competitor DNA (lanes 11–18). A 2- to 3-fold mass excess of E. coli DNA was necessary to compete as effectively as the homologous T-DNA fragment.

**Protein Blotting.** Protein blotting procedures were used for the identification of DNA-binding protein(s) present in induced AD1075 cell extract (Fig. 4). Total proteins from induced and uninduced cells were separated on 12.5% NaDodSO₄/polyacrylamide gels and transferred onto nitrocellulose filters. When heat-denatured ssRNA was used to probe the filters, no binding activity was detected in uninduced cells (Fig. 4A, lane 1). Induced cells contained a 68-kDa ssDNA binding protein (lane 2). This band has previously been identified as the virE2 polypeptide (Fig. 2). No binding to this (or other protein) was observed when ssRNA was used as a probe (data not shown). The DNA probe used in these studies contained both T-DNA and vector sequences. To determine whether the same protein in fraction II preparations used in DNA–protein-binding experiments was binding to ssDNA, and to further assess the specificity of the ssDNA-binding activity, duplicate blots containing fraction II preparations from uninduced AD1075 cells, induced AD1075 cells, and induced ATH2 cells were analyzed by DNA blotting experiments (Fig. 4B and C). Both T-DNA probe and vector DNA probe bound to the virE2 polypeptide present only in the induced AD1075 cells (lanes 2). The two control extracts failed to show any DNA-binding activity in these assays (lanes 1 and 3). When visualized with Coomassie brilliant blue protein stain (Fig. 4D), all three preparations showed similar protein patterns except that
induced AD1075 cells contained an additional virE2 polypeptide band (lane 2), and induced ATH2 cells contained β-lactamase gene product (bla) (lane 3). The latter is due to overproduction of bla as a result of transcription from the trp promoter under induction conditions; the amp gene of the plasmid is located immediately downstream from the trp sequences (Fig. 1).

Protein blotting procedures were also used to identify DNA binding activity, if any, present in agrobacteria cocultivated in the presence or absence of plant cells (Fig. 5). At least two ssDNA-binding proteins were present in induced agrobacteria (Fig. 5B, lane 4). Of these, the smaller polypeptide comigrated with the virE2 polypeptide overproduced in E. coli (lane 2). The virE2 gene product, isolated either from E. coli or from Agrobacterium, did not bind to dsDNA (Fig. 5A, lanes 2 and 4). The larger polypeptide bound to both dsDNA and ssDNA (Fig. 5A and B, lanes 3 and 4). The affinity for ssDNA, however, was significantly higher as apparent from the intensity of the bands (lanes 4). This protein was not encoded in the Ti plasmid, and its synthesis was not induced by cocultivation of bacteria with plant cells (Fig. 5B and data not shown).

To further ensure that the Agrobacterium protein comigrating with the ssDNA-binding activity present in induced AD1075 cells was the virE2 polypeptide, immunological procedures were used. The filter used in experiments described in Fig. 5B was stripped of radioactivity by high salt (0.5 M NaCl) treatment. It was then probed with anti-virE2 antibody raised against virE2 polypeptide produced in E. coli. Several E. coli proteins present in both induced and uninduced cells reacted with both immune and control serum (Fig. 5C, lanes 1 and 2, and data not shown). With immune serum an additional band that corresponded to the virE2 polypeptide was present only in the induced cell extract (lane 2). In agrobacteria, only a single protein reacted with the antibody (lane 4). This protein was present only in cells induced with plant cells and was absent in agrobacteria grown in culture (lanes 3 and 4).

**DISCUSSION**

The overproduction of A. tumefaciens virE2 polypeptide in E. coli was achieved by construction of a trpE'-virE operon fusion (Fig. 1). A protein fraction containing virE2 gene product strongly bound to ssDNA both in gel retardation and in protein blotting experiments (Figs. 3–5). This protein exhibited no significant activity for either dsDNA or ssRNA (Fig. 5 and data not shown). Using ssDNA probes specific for right border sequences, left border sequences, internal T-DNA sequences, and vector sequences, no specificity for T-DNA sequences was evident (Figs. 3–5 and data not shown). Agrobacterium induced with plant cells synthesized a protein that binds to ssDNA and reacted with anti-virE2 antibody (Fig. 5). These results indicate that virE encodes a ssDNA-binding protein.

The indiscriminatory DNA-binding property of the virE gene product provides indirect support for the hypothesis that T-DNA transfer from Agrobacterium to plant cells occurs via a ssDNA intermediate. Studies of Stachel et al. (8) and Albright et al. (9) have demonstrated that in agrobacteria, induced either with the chemical inducer acetosyringone or with plant cells, specific cleavages occur at the T-DNA border sequences. These cleavages may lead to the production of ssDNA molecules composed of the bottom strand of the T-DNA (8). It is well documented that the sites of integration of T-DNA in various tumor lines occur within or close to the two border sequences (3–6). Since site-specific cleavages occur within the border sequence, which in turn are predominantly the site of integration in the plant genome, it has been postulated that the cleavage is a prerequisite for DNA transfer (8, 9). It was also observed that the T-DNA is present in plant tumors mostly as a contiguous segment of DNA, although some rearrangements were noted (2).

If the ss T-DNA is indeed an intermediate in the DNA transfer process, it probably is necessary to protect the DNA from nuclease prior to its integration into the plant genome. The virE protein, because of its strong affinity for ssDNA, probably serves this function. This protective role of virE gene product can account for earlier observations that virE mutants transfer DNA into plant cells but fail to form tumors (17, 18). In gene transfer studies, a low efficiency of DNA transfer will suffice because of a large amplification of the signal (usually a viral sequence that can replicate autonomously in plant cells (18). In virE mutants, in the absence of an efficient protection function, a small fraction of T-DNA molecules is expected to survive endogenous nuclease ac-
tion. This will significantly reduce the efficiency of DNA transfer. Tumor formation, a process involving integration of T-DNA into the plant genome, may require a higher efficiency of DNA transfer. The latter, therefore, requires the presence of an active virE gene product. In addition, if the virE protein serves an additional function in the plant cell, as suggested by Gardner and Knauf (18), the protein must be transported into the plant cells. An economic and efficient way of protein transfer across a phylogenetic barrier will be to use the DNA transfer mechanism the bacteria have already established for plant cell transformation. The protein can be transferred by piggybacking on the T-DNA.

A significant amount of information is available on the structure-function relationship of ssDNA-binding proteins (for review, see ref. 32). Biochemical and mutational analyses of several ssDNA-binding proteins have identified the functional domains involved in DNA binding, protein–protein interaction, etc. For both E. coli ssDNA-binding proteins and T4 gp32 proteins, the N-terminal regions encode the DNA-binding function while the same function is encoded in the C-terminal region of adenovirus DBP (32–34). Like DBP, the DNA-binding domain of virE is probably encoded within the C-terminal two-thirds of the protein because the truncated protein produced in AD1082 had no DNA-binding activity (unpublished data). The possibility that an altered tertiary structure of the truncated protein leads to the loss of DNA-binding activity cannot be ruled out at present. At the primary structure level, these proteins do not share any significant homology. Comparison of the virE2 protein sequence with sequences in the National Biomedical Research Foundation Protein Database identified no highly homologous sequences.

Note Added in the Proof. While this paper was in press, Gietl et al. (35) also demonstrated that the virE operon of Agrobacterium encodes a ssDNA-binding protein.

I thank T. Glover and B. Allen for typing the manuscript and J. Carpenter, C. DeFranco, J. Fuchs, J. Ghai, G. Pazour, and J. Schottel for critical readings of the manuscript. The excellent technical assistance of Grace Chiou is gratefully acknowledged. This work was supported by National Institutes of Health Grant GM 37555, University of Minnesota Graduate School Grant-in-Aid, American Cancer Society Institutional Grants, and an American Cancer Society Junior Faculty Research Award (JFRA 170).