MOLECULAR PLANT-MICROBE INTERACTIONS
Vol. 2, No. 2, pp. 43-52, 1989

Expression of *Agrobacterium* Nopaline-Specific VirD1, VirD2, and VirC1 Proteins and Their Requirement for T-strand Production in *E. coli*

Guido De Vos and Patricia Zambryski

Division of Molecular Plant Biology, University of California, Berkeley 94720 U.S.A.
Received 16 December 1988. Accepted 22 December 1988.

Induction of Ti plasmid virulence (*vir*) genes during early stages of the genetic transformation of plant cells by *Agrobacterium tumefaciens* results in several molecular events that are involved in generating a transferable T-DNA copy. These events include site-specific nicking at the T-DNA borders and synthesis of free, unipolar, linear, single-stranded copies of the T-DNA (T-strands). Here *E. coli* was used as a heterologous cell to assay the requirements for T-strand synthesis. Cells of *E. coli* harbored two compatible plasmids, one containing coding sequences overlapping the *virC* and *virD* regions of the nopaline Ti plasmid, and a second plasmid containing a T-DNA region. The amount of *vir* proteins produced was varied by placing their expression under the control of either native *Agrobacterium*, *tac*, or T7 promoters. The data show that VirD1 and VirD2 proteins are absolutely essential for T-strand production in *E. coli*, and the relative amounts of these polypeptides produced correlate with the amounts of T-strand observed. When VirD1 and VirD2 products are limiting, the VirC1 protein increases T-strand production. The yield of T-strands also varies as a function of the plasmid vector used to clone the T-DNA region substrate; the same T-DNA cloned into pLAFR1 produces more T-strands than that cloned into the higher copy number plasmid pACYC184. In summary, VirD1 and VirD2 proteins are the minimal requirements for T-strand production; however, other factors such as VirC1, the relative concentration of VirD1, VirD2, and the T-DNA substrate, and possibly additional functions (e.g., those specified by pLAFR1) influence the efficiency of T-strand production. Additional results regarding the requirements for expression of VirD1 and VirD2 polypeptides are presented.

Additional keywords: *Agrobacterium*-plant cell interaction, *Agrobacterium vir* genes, *Agrobacterium vir* gene function.

Agrobacterium tumefaciens genetically transforms plant cells by transferring a particular DNA segment, the T-DNA, from its large tumor-inducing (Ti) plasmid into the plant nuclear genome (reviewed in Zambryski 1988). The T-DNA is flanked on the Ti plasmid by two 25-bp direct repeats. These sequences fully define the T-DNA because any DNA located between these T-DNA is efficiently transferred and integrated into the plant nuclear genome. However, the T-DNA does not encode the products that mediate its transfer. Instead, another region of the Ti plasmid, the 35-kbp virulence (*vir*) region including six loci, *virA, virB, virC, virD, virE*, and *virG*, provides most of the *trans* acting functions for T-DNA transit (Stachel and Nester 1986). Induction of *vir* gene expression is the control switch for T-DNA transfer to plant cells and occurs only when *A. tumefaciens* is in the presence of wounded susceptible plant cells; these plant cells excrete low molecular weight phenolic compounds that act as *vir*-inducing signal molecules (Stachel *et al.* 1985).

Induction of *vir* gene expression initiates several molecular reactions that lead to T-DNA transfer. First, single-stranded (ss) nicks occur on the bottom strand of the 25-bp border repeat sequences (Stachel *et al.* 1986, 1987; Yanofsky *et al.* 1986; Wang *et al.* 1987; Albright *et al.* 1987). Second, a single-stranded copy of the T-DNA corresponding to the bottom strand of the T-DNA region on the Ti plasmid is generated (Stachel *et al.* 1986, 1987; Veluthambi *et al.* 1988). This novel molecule, designated T-strand, is produced at about one copy per *vir*-induced *Agrobacterium* and has been proposed to represent the T-DNA intermediate molecule that is transferred to the plant cell (Stachel *et al.* 1986). Characterization of the processes resulting in T-strand production and possible transfer to a susceptible plant cell requires an investigation of the *vir* gene products that specify these events. Several laboratories have demonstrated that two polypeptides, VirD1 and VirD2, from the *virD* complementation group are essential for T-DNA border nicking to occur (Yanofsky *et al.* 1986; Stachel *et al.* 1987; Veluthambi *et al.* 1987). Mutations in VirD1 and VirD2 also block T-strand production because ss border nicks are likely used as initiation and termination sites for T-strand generation (Stachel *et al.* 1987). Saturation mutagenesis of the *vir* region has not revealed any additional *vir* functions necessary for T-strand synthesis; presumably other required functions such as helicases and polymerases must be essential bacterial functions encoded by the chromosome of *A. tumefaciens*. One other *vir* function, encoded by the largest open-reading frame (ORF) of the *virE* complementation group, VirE2, has been identified as an ssDNA binding protein (Gietl *et al.* 1987; Citovsky *et al.* 1988; Das 1988; Christie *et al.* 1988). The tight, stoichiometric, and cooperative binding of the VirE2 polypeptide to ssDNA has suggested that it may be involved in a later step of T-strand transfer, potentially forming a T-strand protein transfer complex (Citovsky *et al.* 1989).

To better define the requirements for T-strand production, a two-plasmid T-strand producing system was constructed in a heterologous host cell of *E. coli*: one plasmid contained *vir*-specific coding sequences, and the other plasmid contained a two-border T-DNA region of defined length. Several *vir* gene expression plasmids

carrying promoters of different strengths were tested, and the T-DNA plasmid substrate was varied by changing its relative copy number. The results confirm that the VirD1 and VirD2 polypeptides are essential for T-strand production. However, efficient T-strand production in *E. coli* is directly correlated with the amount of VirD1 and VirD2 polypeptides produced. Furthermore, expression of the VirC1 polypeptide, which maps immediately upstream of the VirD1 coding region *in vivo*, increases the levels of T-strand produced. During our attempts to produce abundant amounts of VirD1 and VirD2 polypeptides, we observed several novel features that affect their translation and functionality. Whereas Jayaswal *et al.* (1987) have previously shown T-strand production in *E. coli*, the present studies examine several parameters that influence this process.

MATERIALS AND METHODS

Bacterial strains, plasmids, and general procedures. Bacterial strains and plasmids are listed in Table 1. Strains of *E. coli* were grown in either M9-minimal medium or Luria-Bertani medium (Miller 1972), supplemented with antibiotics (carbenicillin, 50 μg/ml, or tetracycline, 10 μg/ml) when appropriate. Enzymes were purchased from New England Biolabs and used according to the manufacturer's specifications. Labeling of plasmid-encoded proteins in maxicells of *E. coli* was performed essentially as described by Sancar *et al.* (1981) with strain RB901, irradiated with 35 J/m^2. Constructs that use the *E. coli* tacII promoter in pGS351 were induced with IPTG (1 mM) 30 min before the labeling. Alternatively, constructs that express proteins under control of the T7 promoter were induced in BL21DE3 with IPTG (100 μM) for 30 min, after which rifampicin was added (100 μg/ml) for another 30 min, followed by radioactive label (10 μCi of ^{35}S-methionine per milliliter [1,100 Ci/mM] for 30 min). The latter strain contains a lysogenic phage lambda derivative carrying the *tac*-inducible T7 RNA-polymerase gene (Rosenberg *et al.* 1987). Other procedures were essentially as described by Maniatis *et al.* (1982).

Construction of hybrid plasmids. We constructed two different plasmids that contain a T-DNA structure by using an *Eco*RI fragment isolated from pAGS111, which carries a copy of the NPTII gene under the *nos* promoter, flanked by octopine Ti-plasmid border sequences (van den Elzen *et al.* 1985). This fragment was cloned into pACYC184 (Chang and Cohen 1978) and pLAFR1 (Friedman *et al.* 1982), and the constructs were designated pGS111 and pGS112 (Table 1).

Three classes of plasmids were used to provide VirD proteins. The first set of constructs express *virD* under its native promoter. The nopaline C58 Ti plasmid *Eco*RI fragment 29 was cloned in both orientations in pUC118 (Vieira and Messing 1987), yielding plasmids pGS400 and pGS401. In pGS408, a derivative of pGS400, the single *Bgl*II site was subsequently filled in by using Klenow DNA polymerase. A VirC1 deletion derivative of pGS400, pGS409 was obtained after restriction with *Nae*I and recircularization, deleting sequences from the *Nae*I site within the pUC vector plasmid (approximately 500 bp left of the polylinker) through the C-terminus of VirC1 up to a second *Nae*I site in the 14th codon of VirC1 (Close *et al.* 1987).

Table 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source or reference
MM294A	*pro*-82 *thi*-1 *end*A1 *hsd*R17 *sup*E44	G. Walker
RR1	*hsd*S20 *sup*E44 *ara*-14 *gal*K2 *lac*Y1 *pro*A2 *rps*L20 *xyl*-5 *mt*-1	R. Brent
BL21DE3	F$^-$ *hsd*S *gal* with T7-RNA polymerase under control of the lacUV5 promoter in lysogenic λ in strain DE3	Studier and Moffatt 1986
RB901	F$^-$ *thr*-1 *leu*-6 *lac*Y? *rps*L31 *mtl*AΔ(*rec*A-*srl*)21 *lex*-A51(Def) *lex*A3(Ind$^-$) *sul*A11	Sancar *et al.* 1981
pAGS111	pUC8 containing a T-DNA *Eco*RI insert	van den Elzen *et al.* 1985
pGV0361	Restriction fragment of pTiC58 cloned in pBR322 containing *virD*, *virE*, and part of *virC*	Depicker *et al.* 1980
ptac5	pBR322 carrying tacII promoter and a polylinker	Hallewell *et al.* 1985
pGS111	pACYC184 with T-DNA *Eco*RI fragment of pAGS111	This work
pGS112	pLAFR1 with T-DNA *Eco*RI fragment of pAGS111	This work
pGS351	ptac5 with *lac*I^q, mob$^-$	This work
pGS360	Partially restricted *Bgl*II fragment (6.1 kb) containing the *virD* gene cloned in the *Bgl*II site of pGS351	This work
pGS361	*Eco*RI complete deletion derivative of pGS360 (Fig. 2)	This work
pGS362	*Bam*HI complete deletion derivative of pGS360 (Fig. 2)	This work
pGS363	*Eco*RI partial deletion derivative of pGS360 (Fig. 2)	This work
pGS364	*Hin*dIII complete deletion derivative of pGS360 (Fig. 2)	This work
pGS366	*Pst*I partial deletion derivative of pGS360 (Fig. 2)	This work
pGS368	*Sma*I complete deletion derivative of pGS360 (Fig. 2)	This work
pGS369	*Mae*I-*Bam*HI fragment of pGS400 (*Bam*HI in polylinker) cloned in pET3c restricted with *Nde*I-*Bam*HI (see Fig. 7)	This work
pGS371	*Bgl*II fragment of pTiC58 containing the coding region of *virD*1 and two proteins cloned in pET3a (protein fusion in GGA reading frame)	This work
pGS373	Same as above in pET3b (GAT frame)	This work
pGS375	Same as above in pET3c (ATC frame)	This work
pGS378	*Bgl*II-*Xmn*I deletion derivative of pGS361 (Fig. 2)	This work
pGS380	*Bgl*II-*Xmn*I deletion derivative of pGS363 (Fig. 2)	This work
pGS400	pTiC58 *Eco*RI fragment 29 cloned in pUC118, *virC*1 coding region toward *lac* promoter (Fig. 2)	This work
pGS401	Same as pGS400 but fragment in opposite orientation	This work
pGS408	Derivative of pGS400 that has the unique *Bgl*II site filled using Klenow *E. coli* polymerase	This work
pGS409	Derivative of pGS400 that has an *Nae*I deletion (*virC*1)	This work
pET3a	pBR322-derived vector to construct protein fusions under control of the T7 promoter across *Bam*HI site in GGA reading frame	Rosenberg *et al.* 1987
pET3b	Same as above for fusions in GAT frame	Rosenberg *et al.* 1987
pET3c	Same as above for fusions in ATC frame	Rosenberg *et at.* 1987

The second set of constructs uses the *tac*-inducible promoter of *E. coli* (deBoer *et al.* 1983). First, we modified the pBR322 derived *tac*-promoter vector ptac5 (Hallewell *et al.* 1985) to include the *lacI*^{q} gene of *E. coli*. The area of ptac5 containing the nick/bom site was excised by using *Tth*111I and *Nde*I, and after blunt ending with Klenow DNA polymerase, ligated to a blunt-ended *Hin*dIII fragment containing *lacI*^{q} (Bagdasarian *et al.* 1983). The single *Bgl*II site of this vector, pGS351 (Fig. 1), was used to insert the *vir*D region of pTi C58 as a partially restricted *Bgl*II fragment of 6.1 kb (see Hagiya *et al.* 1985), isolated from pGV0361 (Depicker *et al.* 1980). 3′ deletion derivatives of this construct, pGS360, were obtained by complete or partial restriction with different enzymes, followed by recircularization (Fig. 2). A deletion of the first ORF, VirD1, was introduced in plasmids pGS361 and pGS363 after double digest with *Xmn*I and *Bgl*II, blunt ending of the *Bgl*II end with Klenow DNA polymerase, and recircularization, yielding pGS378 and pGS380, respectively (Fig. 2).

A third set of constructs expresses both VirD1 and VirD2 under control of the T7 promoter. Protein fusions with VirD1 were made by using the T7 expression vectors pET3a, pET3b, and pET3c (Rosenberg *et al.* 1987), which generate fusions in all three ORFs across a unique *Bam*HI site. The promoter proximal *Bgl*II fragment of pGS360 was subcloned into each pET vector to generate three plasmids, pGS371, pGS373, and pGS375, respectively (Table 1, Fig. 3). Another construct, pGS369, was made by cloning a 1,250-bp *Mae*I-*Bam*HI fragment of pGS400 (see above) into pET3c after restriction with *Nde*I and *Bam*HI (Fig. 3).

Induction of border nicks and T-strands in *E. coli*. Plasmids pGS111 and pGS112, which contain a T-DNA structure, were transformed into strains RR1 and BL21DE3 of *E. coli* by using selection for tetracycline resistance, and these strains were subsequently transformed with the appropriate *virD* constructs, selecting for carbenicillin

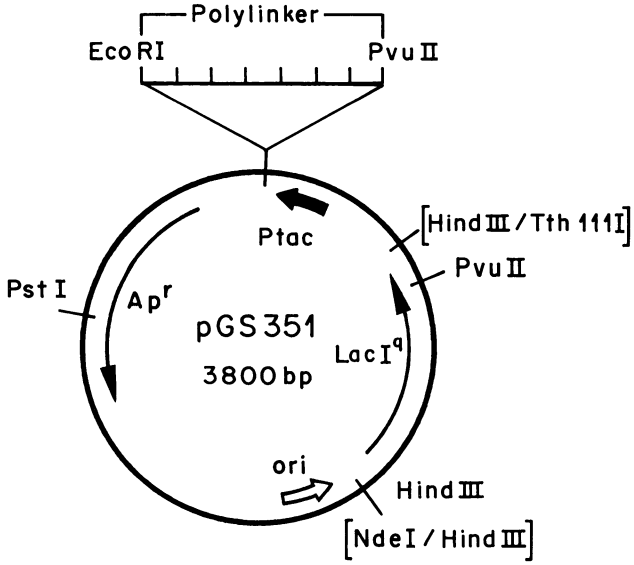

Fig. 1. Restriction map of the *tac* expression vector pGS351. Restriction sites in the polylinker are (left to right): *Eco*RI, *Ava*I, *Sma*I, *Sal*I, *Pst*I, *Bgl*II, *Xba*I, and *Pvu*II. Ligation of the blunted *Hin*dIII fragment containing *lacI*^{q} into the blunted *Nde*I-*Tth*111I vector plasmid resulted in regeneration of the *Hin*dIII site at the *Nde*I/*Hin*dIII junction.

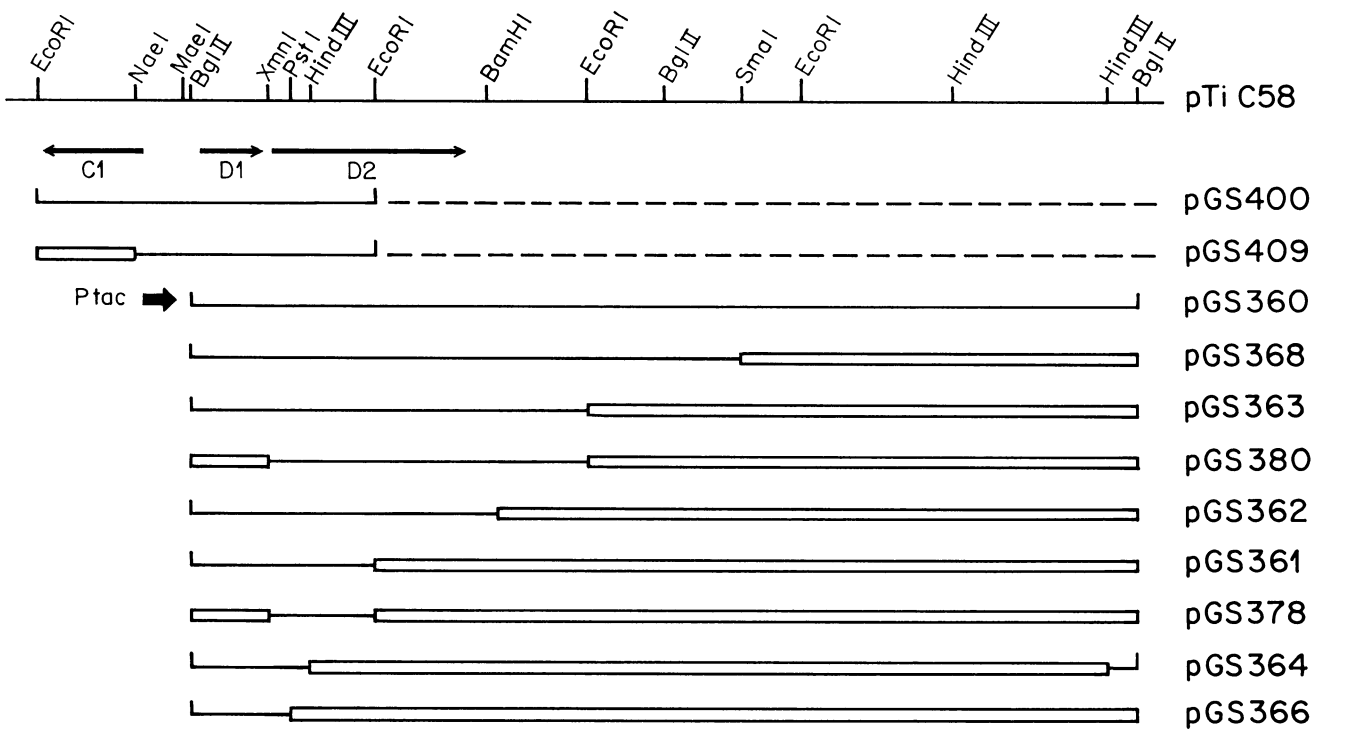

Fig. 2. Restriction map of the *virC-virD* region of pTiC58 and derived recombinant plasmids. pGS400 contains *Eco*RI fragment 29 (Depicker *et al.* 1980) cloned in pUC118, and pGS409 is an *Nae*I deletion derivative of pGS400. pGS360 contains the entire *virD* region as a 6.1-kbp partial *Bgl*II fragment cloned into the *Bgl*II site of the *tac* vector plasmid pGS351. The plasmids shown are all derived from pGS360 and represent deletions obtained by using either partial of complete digests of pGS360, and their limits can be seen with reference to the restriction map. Deleted areas are shown as open bars. The right end of all deletions, except pGS364, are within the relevant site of the polylinker of pGS351.

resistance. DNA was prepared and analyzed for border nicks and T-strands as in Stachel *et al.* 1986.

RESULTS

Construction of *vir* gene expression plasmids. If indeed the only specific protein requirements for T-strand production are the VirD1 and VirD2 products, then synthesis of these polypeptides in a heterologous host cell should be sufficient to direct T-strand production; we assume other requisite enzymes of DNA metabolism are constitutive chromosomally encoded functions. Here the production of T-strands was measured by using cells of *E. coli* containing two types of plasmids, one to produce *virD* products and a second to provide a T-DNA template.

To correlate the production of T-strand with the amounts of *virD* products made, several sets of constructs (Table 1) differing in efficiency to promote transcription of *virD* coding sequences were utilized. The first set of constructs carries *Agrobacterium virD* sequences under control of their natural promoter. Plasmid pGV0361 (Depicker *et al.* 1980) carries *Agrobacterium* pTiC58 *virC*, *virD*, and *virE* coding sequences (with reference to Fig. 2, this plasmid contains 850 bp to the left of the *Eco*RI site and 7.3 kbp to the right of the *Bgl*II site). A derivative of this plasmid, pGS400 contains *Eco*RI fragment 29, carrying the promoter regions and ORFs of VirC1, VirD1, and approximately half of VirD2 (Fig. 2).

A second set of constructs carries different parts of the *virD* operon under control of the *tac*-inducible promoter of *E. coli*. Plasmid pGS360 contains the entire *Agrobacterium virD* operon as a partially restricted 6.1-kbp *Bgl*II fragment (Fig. 2); the promoter proximal *Bgl*II site of this construct is about 25 bp upstream from the start of the *Agrobacterium* VirD1 ORF (Hagiya *et al.* 1986; Fig. 3). Several deletion derivatives of pGS360 were made (Fig. 2), including progressive 3′ deletions and deletions of the first (VirD1) ORF.

A third set of constructs contains the VirD1 and VirD2 coding region under control of the strong T7 promoter (Rosenberg *et al.* 1987). Three plasmids, pGS371, pGS373, and pGS375, were constructed to make protein fusions in all three ORFs across the *Bgl*II site (Table 1, Fig. 3) upstream of the VirD1 coding sequence. A fourth plasmid, pGS369, fuses the *Nde*I site located at the T7 AUG start codon to the *Mae*I site 70 bp upstream from the virD1 ORF, combining the T7 promoter with *Agrobacterium* translational control sequences.

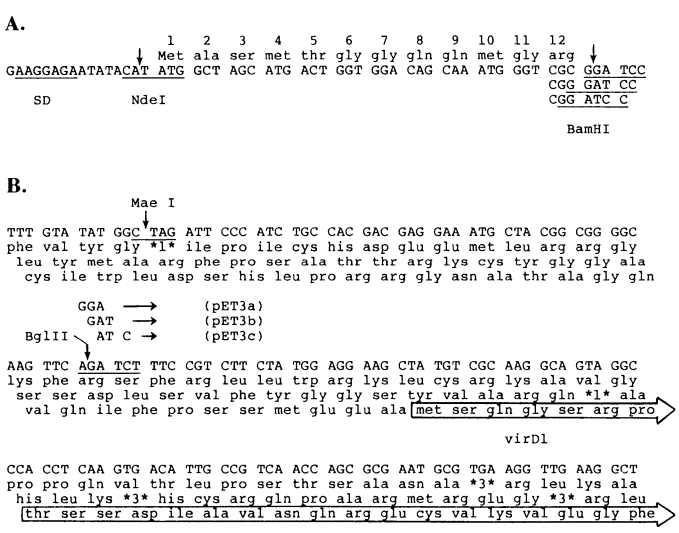

Fig. 3. Nucleotide sequence of T7 vectors and 5 *VirD* region. **A,** The nucleotide sequence of the fusion protein starts in pET3a, pET3b, and pET3c T7-protein fusion vectors (Rosenberg *et al.* 1987). **B**, The *virD*1 upstream sequence (Hagiya *et al.* 1985). Constructs pGS371, pGS373, and pGS375 express fusions with GGA, GAT, and ATC, respectively, as the in-frame codons through fusion of the *Bam*HI site (**A**) with the *Bgl*II site (**B**). The *Nde*I and *Mae*I sites used in the cloning of pGS369 are shown. The vertical arrows show points of cleavage by restriction enzymes. The ORF of VirD1 is enclosed in a large arrow. SD, Shine-Dalgarno ribosome binding site.

***Agrobacterium vir* protein expressed in *E. coli*.** The three groups of constructs were tested for their ability to produce the 15-kDa VirD1 and the 56-kDa VirD2 polypeptides (although the nucleotide sequence of VirD2 predicts a polypeptide of 47 kDa, its mobility on sodium dodecyl sulfate-protein gels is 56 kDa [Yanofsky *et al.* 1986]). The proteins shown in Figure 4 are plasmid specific, either synthesized in maxicells or following induction of T7 RNA polymerase. Lane a of Figure 4 shows that plasmid pGV0361 produces no detectible levels of VirD proteins, although an intense band at 31 kDa corresponding in size to VirC1 (Yanofsky and Nester 1986; Close *et al.* 1987) is present. Plasmid pGS400, containing an *Eco*RI subfragment of pGV0361, also produces low levels of the 15-kDa VirD1 polypeptide. This construct should also produce a polypeptide migrating at about 30 kDa, the expected size for the deletion protein of VirD2; its visualization in lanes c and d of Figure 4 is obscured by VirC1, but not in a VirC1 deletion mutant (data not shown). Production of *vir* polypeptides by the *Eco*RI subfragment is independent of its orientation in the vector (Fig. 4, lane d).

The *E. coli tac* promoter construct, pGS360, produces high levels of the 15-kDa VirD1 protein and lesser amounts of the 56-kDa VirD2 polypeptide (Fig. 4, lanes b and e). Another less abundant polypeptide migrating at approximately 43 kDa (seen also in lanes g, h, and i) is presumably a breakdown or processing product of VirD2 because it reacts strongly with VirD2 antibody (data not shown). The prominent band at 40 kDa is synthesized in maxicells containing the vector plasmid alone (data not shown). Because pGS360 contains the entire 4.5-kbp *vir*D region, it also includes sequences encoding additional *virD*-specific polypeptides. In the octopine Ti plasmid, sequences downstream of VirD2 predict two ORFs of 21.3 and 75.8 kDa (Porter *et al.* 1987; Jayaswal *et al.* 1987). There are bands migrating at mobilities corresponding to these molecular weights that may represent the downstream nopaline *virD* polypeptides; however, these downstream polypeptides play no direct role in the synthesis of T-strands (Yanofsky *et al.* 1986; Stachel *et al.* 1987; Veluthambi *et al.* 1987; and see below).

The protein patterns of the plasmids utilizing the T7 promoter show the surprising result that two out of the three ORFs of the fusion constructs synthesize polypeptides migrating at positions expected for VirD1 or T7-VirD1 fusion polypeptides. The pGS375 construct that makes a fusion in the ATC reading frame (across the *Bam*HI-*Bgl*II vector/*virD* junction; Fig. 3) was the only one expected to produce a VirD1 fusion polypeptide 21 amino acids longer than the native VirD1 polypeptide; one of the three polypeptides migrating between 16.5 and 17.5 is likely the

expected fusion polypeptide (Fig. 4, lane g); why there are three fusion polypeptides synthesized is unknown. As expected, no VirD1 protein band is produced in the construct (pGS371), which makes a protein fusion in the GGA reading frame that is out of frame with VirD1 (Fig. 4, lane i). However, the GAT reading frame construct (pGS373) is also out of frame with the VirD1 ORF. This construct does not direct the synthesis of VirD1 fusion polypeptide(s); instead, it directs the synthesis of a protein band migrating at the position of the VirD1 polypeptide, and an additional, slightly lower molecular weight band (Fig. 4, lane h). The pGS369 construct directs the synthesis of the expected 15-kDa VirD1 polypeptide and a larger 18-kDa VirD1 fusion polypeptide. In a later section we will describe hypotheses for these unexpected VirD1-related bands in the different T7 promoter constructs. Note that the aberrant translation products are specific to the smaller, that is, VirD1 polypeptides, because all the T7 constructs direct the synthesis of the expected VirD2 polypeptides, either the 56-kDa polypeptide or the 32-kDa deletion polypeptide in pGS369 (pGS369 only contains VirD2 sequences up to the *Eco*RI site as in pGS400).

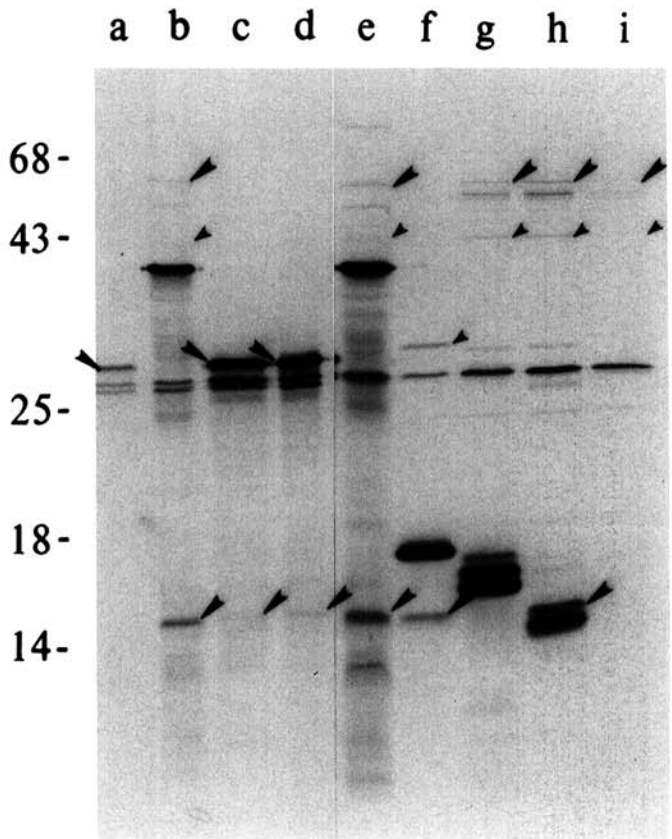

Fig. 4. *Vir* protein expression in *E. coli*. Lanes a–e show the protein patterns obtained in maxicells of *E. coli* for plasmids pGV0361, pGS360, pGS400, pGS401, and pGS360. Lanes f–i show protein patterns obtained after induction of the T7 polymerase for plasmids pGS369, pGS375 (ATC reading frame), pGS373 (GAT reading frame), and pGS371 (GGA reading frame). The separation was obtained following electrophoresis through a 12.5% acrylamide gel, and the numbers to the left refer to the apparent molecular weights ($\times 10^{-3}$) of standard protein markers. The larger arrows indicate the positions of VirD1 (15kDa), VirD2 (56 kDa), or VirC1 (31 kDa), and the smaller arrows indicate the positions of a VirD1-derived polypeptide (43 kDa) and VirD2-derived polypeptides (32 kDa in lane f).

Vir-mediated production of T-strands in *E. coli*. Each construct described above was tested for its ability to promote T-strand production in *E. coli*. Either one of two compatible plasmids, pGS111 or pGS112, containing the same T-DNA region (Fig. 5A), was provided in *trans*. Plasmid pGS111 contains the T-DNA region cloned in pACYC184 (4.2 kbp), and pGS112 is cloned in pLAFR1 (22.6 kbp); pACYC-type plasmids are maintained in the bacterial cell at five- to 10-fold higher copy number than pLAFR-type plasmids.

In a first experiment, the *virD* region cloned under the *tac* promoter (pGS360) was assayed for its ability to produce the known *Agrobacterium vir*-induced T-DNA-associated molecular reactions, that is, ss nicks at the T-DNA borders,

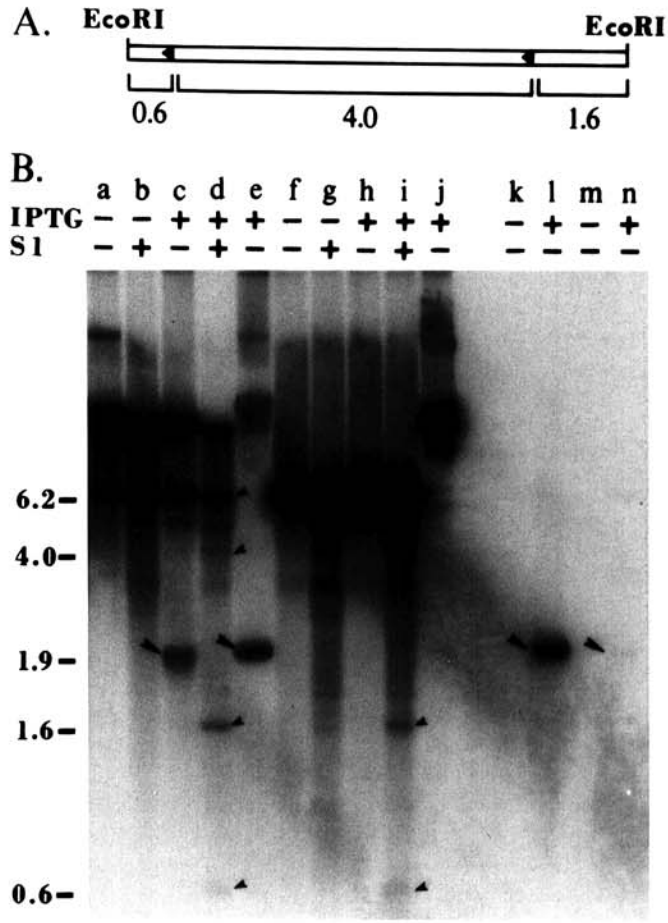

Fig. 5. Nicking and T-strand induction in *E. coli*. **A.** 6.2-kbp *Eco*RI restriction fragment containing the same T-DNA region cloned in pLAFR1 (pGS112) or in pACYC184 (pGS111). The arrows indicate the positions of the 25-bp T-DNA border repeat sequences. The sizes of fragments expected to arise following nicking at the T-DNA borders and subsequent treatment of DNA with *Eco*RI and S1 nuclease are shown below. **B**, Nicking and T-strand induction in cells of *E. coli* containing pGS360 (for *vir* protein synthesis) and either pGS112 (lanes a–e k, l) or pGS111 (f–j, m, n). *tac*-promoted *virD* protein synthesis was induced with IPTG in lanes c–e, h–j, l, and n. The DNAs in lanes b, d, g, and i were treated with S1 nuclease before electrophoresis. The left part of the gel (lanes a–j) was denatured before Southern blotting; the right part (lanes k–n) was not. Lanes a–d and f–i were restricted with *Eco*RI before electrophoresis. The sizes of the relevant hybridization signals are given in kbp to the left of the figure. The smaller arrows indicate the positions of border nick products, and the larger arrows indicate the position of the T-strand produced from either pGS112 or pGS111. Purified 6.2-kbp *Eco*RI fragment containing the entire T-DNA region was used as probe.

and ss T-strand molecules. The results in Figure 5B, lanes d and i, show the presence of bands expected to arise following nicking at the T-DNA borders; these bands appear specifically in DNA isolated following induction of the *virD* expression under the *tac* promoter. Border nicks are single stranded because their visualization depends on the treatment of induced DNA with S1 nuclease before electrophoresis and hybridization analysis. Note that the most abundant nicked products represent the bands corresponding to the fragments outside the left (0.6 kbp) and right (1.6 kbp) T-DNA borders (see Fig. 5A). The nicked product corresponding to the 4.0 kbp internal T-DNA region is observed to a lesser degree compared with the outside border fragments. Whereas the left or right nicked product can be observed independently, the internal nicked product is less abundant because it requires that both T-DNA borders are nicked synchronously on the same T-DNA plasmid molecule.

In the undigested DNA samples (Fig. 5, lanes e, j, l, n), there is a band migrating at 1.9 kb, the expected size for the T-strand produced from pGS111 or pGS112 (ssDNA migrates at approximately one-half the size of the expected 4.0 double-stranded T-DNA molecule [Stachel *et al.* 1986, 1987] under the conditions utilized). This molecule is single-stranded because it is sensitive to S1 nuclease (data not shown) and transfers to nitrocellulose under both denaturing (Fig. 5, lanes e and j) and nondenaturing (Fig. 5, lanes l and n) conditions. This ss 1.9-kb DNA is only produced when virD expression is induced (compare lanes k and l, or m and n, of Fig. 5).

Thus, both border nicks and T-strands can be produced in *E. coli* under the direction of the products of the *Agrobacterium virD* region. To determine which of the *virD* products are essential for these reactions, a series of deletion mutants was constructed from plasmid pGS360. Figure 6 shows that the region downstream of VirD2 (deleted in plasmids pGS368, pGS363, pGS362; Fig. 6, lanes b, c, and f) is not essential for T-strand production. Plasmid pGS361 (Fig. 6, lane d), containing the sequences for VirD1 and about 50% of VirD2, is the smallest plasmid capable of directing T-strand production; the shorter VirD2 polypeptide produced from this plasmid is equally efficient as plasmids carrying the entire VirD2 coding sequence. No T-strand is produced from 3′ deletions removing almost all of VirD2 (pGS364, pG366; Fig. 6, lanes e and g) or from 5′ deletions removing VirD1 (pGS380, pGS378; data not shown). These results confirm those of Yanofsky *et al.* (1986) and Jayaswal *et al.* (1987) that VirD1 and VirD2 polypeptides are required for border-nicking activity in *E. coli*. These workers also showed that about 40% of the C-terminal end of VirD2 coding sequence is unnecessary for these T-DNA-associated reactions. The present data further delimit that a maximum of 50% of the N-terminus of VirD2 is required for function.

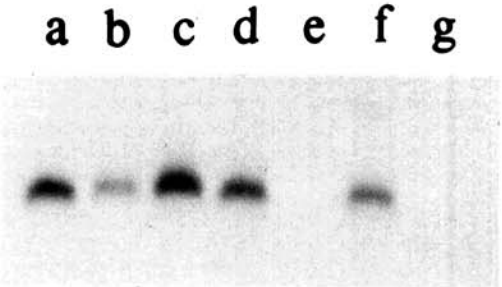

Fig. 6. T-strand production mediated by different portions of the *virD* region. Several *tac* promoter constructs depicted in Figure 2 were used to measure T-strand production from pGS112 (pLAFR derived). The constructs assayed are: pGS360 (a), pGS368 (b), pGS363 (c), pGS361 (d), pGS364 (e), pGS362 (f), and pGS366 (g). Only the region of the gel where the 1.9-kb T-strand migrates is shown. The DNA was analyzed following electrophoresis and transfer to nitrocellulose under nondenaturing conditions; the probe was as in Figure 5.

Interestingly, the data in Figure 5B indicate that whereas border nicks are efficiently produced from both T-DNA plasmids, T-strands are produced efficiently only from the pGS112 plasmid. This latter result may reflect either the copy number difference between the two T-DNA plasmids, that is, the *virD* products are limiting for T-strand production in the higher copy plasmid in pGS111, or there may be additional functions on the pLAFR vector plasmid in pGS112 that promote T-strand production. Most likely both of these possibilities may play a role in T-strand production in *E. coli*.

Figure 7 shows that the amounts of T-strand produced from the pACYC or pLAFR containing T-DNA plasmids following *virD* protein synthesis controlled by either the *tac*, T7, or native *Agrobacterium* promoters. There are three general conclusions: all three types of promoters are capable of synthesizing sufficient *vir* products in *E. coli* to result in T-strand production; the amount of T-strand production is

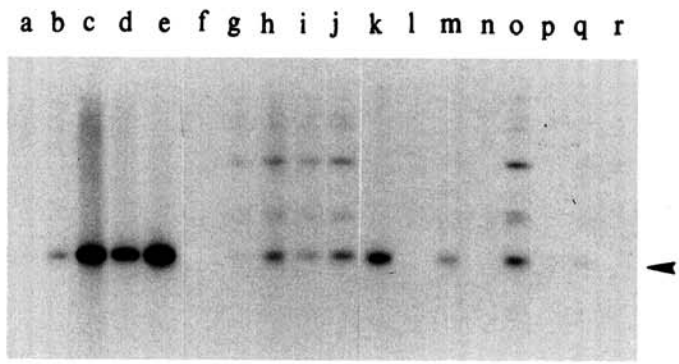

Fig. 7. T-strand production with different *virD* expression plasmids. Lanes a–e show the bands obtained from the pLAFR1-T-DNA containing plasmid pGS112 with pGV0361 (a), pGS409 (b), pGS400 (c), pGS361 (d), and pGS360 (e). Lanes f–j show the bands obtained from the pACYC T-DNA plasmid pGS111 with the same *virD* protein expression plasmids. Lanes k–n and o–r show T-strand hybridization patterns obtained with pGS112 and pGS111, respectively, with the T7-promoter constructs pGS369 (k and o), pGS375 (l and p), pGS373 (m and q), and pGS371 (n and r). The DNA was analyzed following nondenatured transfer to nitrocellulose and probing with T-DNA as in Figure 5. Note that the sizes of the T-strands produced from both pLAFR and pACYC plasmids were expected to be the same; indeed, the major ssDNA detected is 1.9 kb and corresponds to the T-strand produced if the left and right T-DNA borders were used to initiate and terminate its production. Whereas this is the only T-strand produced by using pLAFR T-DNA plasmid, two other relatively abundant T-strands are also produced from the pACYC T-DNA plasmid. The sizes of these latter ssDNAs are 2.4 and 4.0 kb (relative to dsDNA as molecular weight markers). These sizes are different from that expected for T-strands initiating at the left border and terminating at the right border (3.2 kb, i.e., half of 0.6 + 4.2 + 1.6) or a complete copy of the entire plasmid (5.2 kb, i.e., half of 6.2 + 4.2). Thus, these ssDNA molecules represent other unpredicted T-strands potentially using T-DNA border-related sequences on the pACYC vector plasmid. All these "extra" T-strands are dependent on VirD1 and VirD2.

always greater in strains carrying the T-DNA region on the pLAFR vector compared with the pACYC vector; there is a correlation between the amount of *virD* proteins produced and the amount of T-strand produced. The data with respect to the third point are described in more detail below.

In the native *Agrobacterium virD* promoter-driven constructs, the strain carrying plasmid pGS400 produces T-strand, and the strain carrying the pGV0361 plasmid produces barely detectable levels of T-strand (Fig. 7, e.g., lanes a and f vs. lanes c and h). The T-strand produced in these two strains correlates with the data for the amounts of *virD* protein produced in the two constructs, that is, pGS400 produces *virD* polypeptides, whereas pGV361 produces (little or) no detectable amounts of these polypeptides.

Comparison of the levels of *virD* proteins and T-strands in strains carrying the native promoter versus the *tac* promoter constructs is less straightforward. At first glance, the data suggest these two strains produce equivalent amounts of T-strand (Fig. 7, lanes c and h vs. lanes d and e or lanes i and j), although they differ in amounts of *virD* polypeptide synthesis (Fig. 4). However, pGS400 also contains sufficient sequences to encode a complete VirC1 polypeptide. Indeed, deletion of the VirC1 coding sequence from this construct substantially reduces T-strand production. If we now compare the levels of *virD* proteins or T-strands produced in the *tac* (pGS361) (Fig. 4, lanes d and e, i and j) versus the pGS409 (VirC1 minus) constructs (Fig. 4, lanes b and g), there is a direct correlation, that is, the *tac* construct produces more T-strands and more *virD* polypeptides. Thus, these data suggest a possible role for the VirC1 product to increase the efficiency of T-strand production.

The results with the T7 expression plasmids are surprising. Because the T7 promoter was the strongest promoter used, these constructs were expected to be extremely active in T-strand production. Instead, strains carrying these constructs produce slightly less T-strand than strains carrying the *tac* promoter construct (Fig. 7, lanes k through r). These results can be explained in two parts. First, we cannot compare the relative levels of VirD products produced from the T7 constucts to that from *tac* constructs; *tac*-controlled proteins are observed in maxicells, and T7-controlled proteins are observed directly following induction of their transcription by T7 polymerase. Second, careful examination of the data in Figure 4 reveals that although the T7 constructs produce abundant levels of low molecular weight polypeptides, none produces more 15-kDa VirD1 polypeptide than the *tac* construct.

That the construct in the ATC reading frame does not produce much T-strand suggests that the extra 21 amino acids at the N-terminus of the fusion protein interferes with VirD1 function; thus, the N-terminus of VirD1 must be also essential in the native protein. Significant levels of T-strand are produced only by using the construct carrying the GAT T7-*virD* fusion (pGS373, lanes m and q) or the "nonfusion" construct (pGS369, lanes k and o); both constructs produce a polypeptide comigrating with 15-kDa VirD1 and additional lower (pGS373) or higher (pGS369) molecular weight bands. That the pGS369 construct produces less T-strands (than the *tac* construct) may reflect that there is less 15-kDa VirD1 polypeptide produced; as suggested above, the larger 18-kDa polypeptide may be nonfunctional due to extra amino acids at the N-terminus of the fusion polypeptide. The GAT construct makes as much or more of the 15-kDa VirD1 polypeptide than the *tac* construct; potentially, cells carrying the GAT construct are less efficient in T-strand production due to interference by the lower molecular weight polypeptide in forming a complex (or in acting) with the VirD2 polypeptide to produce border endonuclease activity. Because the amounts of VirD2 polypeptide made in the T7 constructs producing VirD1 fusion products are approximately equivalent to that produced in the *tac* construct, the variability in T-strand production cannot be explained by differing amount of VirD2.

The complexity of the data from the T7 constructs neither supports nor contradicts the correlation between the amount of virD products and T-strand production. Instead, it points out features about the functionality of the N-terminus of VirD1 as well as some requirements for proper translation of VirD1. These data are relevant to any future attempts to overproduce VirD1 and VirD2 polypeptides. The next section looks more carefully at the data on the T7-*virD* expression plasmids, emphasizing the requirements for translation of VirD1.

Unusual translation of VirD1-related polypeptides when T7 promoter-derived expression vectors are used. The data from the T7-derived expressed plasmids reveals several unexpected results regarding the synthesis of VirD1 polypeptides. First, VirD1 and T7-VirD1 fusion polypeptides were generated by using two out of the three possible fusion vector plasmids. Figure 3 shows the relevant DNA and amino acid sequences predicted for the three possible fusion constructs. In the T7 expression plasmids used, the strong T7 promoter and the first 12 amino acids of the T7 gene 10 protein are fused via a *Bam*HI site to the *Bgl*II site nine amino acids upstream of the native VirD1 coding sequence. Three T7 vectors were used to make fusions in three possible ORFs for VirD1 coding sequences. Only the fusion in the ATC frame should generate a VirD1 fusion polypeptide; the data in Figure 4 show several polypeptides migrating at the higher molecular weight expected for VirD1 containing an extra 21 N-terminal amino acids. Why there are several fusion polypeptides using this reading frame is unknown. Even more confounding, the construct in the GAT frame produces two polypeptides migrating at and just below the 15-kDa VirD1. The sizes of these polypeptides preclude their having been generated as fusions with T7 or 5′ upstream *virD* sequences. The following explanation for how the GAT frame can produce a 15-kDa VirD1 is offered.

We assume that the ribosome initiates at the T7 ATG codon, proceeds through the T7 leader and 5′ VirD1 sequences, and then terminates at the TAG codon just downstream, in the vicinity of the fifth (*arg*) codon, of the native VirD1 (Fig. 3). We then assume that the ribosome drifts back and reinitiates at the native VirD1 ATG. In this manner, the native 15-kDa VirD1 would be produced as observed. Ribosomal drifting is a mechanism sometimes observed in bacteria; for example, reinitiation of translation following a frame shift mutation can occur if there is a Shine-Dalgarno ribosome binding site close to the site of frame shift-induced termination (Shinedling *et al.* 1987). We have no explanation for the polypeptide migrating slightly smaller than the 15-kDa VirD1; this smaller polypeptide may be also present *in vivo*, but may require the higher

expression mediated by the T7 system to be detected. Two nopaline *virD*-specific bands migrating at about 15 kDa were also observed by Alt-Moerbe *et al.* (1986) in protein fusions equivalent to the ATC frame used here. Note that the GGA reading frame does not produce VirD1 because initiation at the T7 ATG followed by translation would terminate in the vicinity of the 20th amino acid of native VirD1, a distance presumably too great to allow ribosomal drift and reinitiation at the native ATG.

The results with the GAT T7 fusion (and the *tac* or native *virD* promoters) suggested that there is a strong translational site for VirD1 in the native upstream sequences. To potentially optimize the yield of VirD1 (and VirD2), we constructed a plasmid, pGS369, where the strong T7 promoter and Shine-Dalgarno ribosome binding site are fused via the T7 translational start (*Nde*I site) to a *Mae*I site 77 nucleotides upstream of the VirD1 initiation codon. This fusion results in the loss of the ATG initiation codon close to the T7 ribosome binding site; thus, translation of the fusion transcript should initiate at the downstream VirD1 native site. However, instead of only the expected 15-kDa VirD1, a larger 18-kDa polypeptide also was produced with this construct (Fig. 4). An 18-kDa polypeptide is approximately 27 amino acids longer than a 15-kDa polypeptide. This size implies that translation has started very close to the T7 Shine-Dalgarno sequence. ATG is the preferred translational start; GTG and occasionally TTG can be used; and ATA has been observed to allow inefficient initiation (reviewed in Gold 1988). Assuming that no ribosomal frame drifting has occurred, to make an 18-kDa protein requires translation in the same reading frame as native VirD1. There is an ATA sequence overlapping the T7 Shine-Dalgarno that is in frame with VirD1 that may be used as the start site to produce the 18-kDa polypeptide. Several other factors besides Shine-Dalgarno sequences and initiation codons, including other primary sequences, secondary codons, and secondary structure, facilitate translational initiation. Some of these other factors must play a role in the efficient and unusual initiation of the 18-kDa polypeptide observed for the pGS369 construct.

DISCUSSION

To systematically dissect the requirements for synthesis and processing of the *Agrobacterium* T-strand DNA transfer intermediate, we set out to study its formation in a heterologous host cell. Specifically, *E. coli* was used to define the *Agrobacterium* protein components necessary for synthesis of the T-strand. Because the VirD1 and VirD2 polypeptides were shown to be required for border nicks and T-strand synthesis in *Agrobacterium*, the analyses focused on whether these polypeptides were also sufficient to promote T-strand production in *E. coli*. A two-plasmid T-strand synthesizing system was designed consisting of one plasmid carrying *Agrobacterium virD* coding sequences and a second compatible plasmid carrying a T-DNA region. The results suggest several novel features in two general areas: the factors that influence T-strand production and the requirements for efficient expression of VirD1 and VirD2 polypeptides.

Three types of *virD* expression plasmids were constructed by utilizing three different promoters, either the native *Agrobacterium virD* promoter, the *tac* promoter, or the T7 polymerase-dependent promoter. All three types of constructs sometimes yielded surprising results. First, the native *Agrobacterium* promoter is capable of directing readily detectable levels of VirD1, VirD2, and VirC1 polypeptides. Smaller constructs (i.e., present in high copy number) containing only VirC1, VirD1, and VirD2 coding sequences (pGS400), give better expression compared with larger constructs (pGV0361) carrying these as well as several additional kbp of sequences spanning the nopaline *vir* region. In *Agrobacterium*, the expression of *vir* genes, including the *virC* and *virD* complementation groups, are under tight control dependent on the presence of plant signal molecules such as acetosyringone (Stachel *et al.* 1985). The activation of *Agrobacterium vir* gene expression, in the presence of plant signals, is mediated by two of the *vir* loci, *virA* and *virG* (Stachel and Zambryski 1986). *virA* is a membrane protein that is proposed to recognize plant signal molecules (Leroux *et al.* 1987), and *virG* is proposed to be a transcriptional activator of the other *vir* loci (Stachel *et al.* 1986; Winans *et al.* 1986). Neither pGS361 nor pGS400 contains *virA* or *virG* sequences, and the *E. coli* cells were not grown in the presence of plant signals. Thus, the expression of VirC and VirD polypeptides must reflect their transcription via *E. coli* factors. Because the promoters for the *vir* genes are not preceded by classical TATA boxes, *E. coli* factors resembling *virA* and *virG* may be involved in the observed expression. This suggestion is reasonable because the *virA* and *virG* gene products are highly homologous to other two component pairs of sensor/regulator genes (Ronson *et al.* 1987) present in bacterial cells. Because the nopaline VirC and VirD promoters are adjacent and diverging within a short 250-bp region of dyad symmetry (Tait and Kado 1988), in the simplest case, a factor stimulating transcription of one gene simultaneously may promote the transcription of the other.

The native *Agrobacterium* promoter in the pGS400 plasmid produces sufficient *vir* polypeptides to direct the production of T-strands in *E. coli*. However, this production depends also on the presence of the VirC1 product. Because the amount of VirD1 and VirD2 polypeptides synthesized by pGS400 is low relative to the *tac* or T7 constructs, we propose that the VirC1 polypeptide enhances the efficiency of T-strand production when VirD1 and VirD2 are limiting. Because the *tac* and T7 constructs make more abundant levels of VirD1 and VirD2, their requirement for VirC1 is not apparent. Exactly how VirC2 enhances T-strand production is unknown at present; potentially, it associates with proteins, such as the VirD1-VirD2 endonuclease, or with the DNA substrate at or surrounding the T-DNA border region to promote T-strand production. That VirC1 plays an ancillary role in T-strand synthesis fits with earlier genetic studies suggesting this locus is not absolutely necessary for T-DNA transfer; mutants in the *virC* locus are only attenuated in plant cell transformation (Stachel and Nester 1986).

The *tac* promoter constructs direct VirD1, VirD2, and consequent T-strand synthesis at levels higher than the constructs carrying the native *Agrobacterium* promoter, and at least equivalent to the T7 constructs. Several *tac*-derived constructs carrying deletions across the *virD* region were used to determine that VirD1 and only the N-terminal 50% of VirD2 are necessary for T-strand synthesis. Potentially, the C-terminal half of the protein plays a role in

steps subsequent to T-strand synthesis. For example, VirD2 remains tightly associated with the T-strand and can be purified as a T-strand-protein complex from *vir*-induced *Agrobacterium* (Howard *et al.*, unpublished). This association may be necessary for transport of the T-strand out of the bacterial cell, and the C-terminal half of VirD2 may play a role in these later steps.

The *tac* constructs were used also to compare border nicking and T-strand production by using two different T-DNA containing plasmids as substrate. Whereas border nicking occurred at approximately the same levels with T-DNA plasmid, T-strands were effectively produced only from the lower copy number plasmid. Thus, the role of VirD1 and VirD2 in border nicking is significantly different than their role in T-strand production. For example, if the results are explained only as a reflection of the differences in plasmid copy number, then catalytic amounts of VirD1 and VirD2 may be sufficient for nicking, but more stoichiometric amounts may be necessary for T-strand production. The data may also be explained if the lower copy number, and larger pLAFR plasmid specifies additional functions that can enhance T-strand production. Whatever this pLAFR function might be, it is not absolutely necessary as T-strands can be produced, albeit at a low level, in cells containing the pACYC plasmid. T-strand production could require several helicases and polymerases, and these activities can be provided by essential bacterial genes. The pLAFR plasmid may synthesize other such activities that are somehow more compatible with the generation of the T-strand molecule.

In summary, T-strand production in *E. coli* requires VirD1 and the N-terminal 50% of VirD2, and VirC1 when VirD1 and VirD2 are limiting. The data on T-strand production directed by *vir* expression from different promoters, as well as the difference in T-strand synthesis with low (pLAFR-derived) and high (pACYC-derived) copy number T-DNA substrates, suggest that the amounts of VirD1 and VirD2 produced are directly related to the amounts of T-strand synthesized. However, that synthesis of T-strands was consistently higher in strains carrying the pLAFR plasmid compared with the pACYC plasmid independent of whether the native, *tac*, or T7 promoters were used as a source of *virD* proteins also suggests that other as yet unknown factors (some of which may be encoded by the pLAFR vector) play a role in T-strand generation.

This work also describes several results relating to our attempts to overproduce VirD1 and VirD2 by using T7 expression vectors. First, polypeptides carrying 11 amino acids of the T7 gene 10 protein plus nine amino acids encoded by VirD1 upstream sequences fused to the native VirD1 polypeptide (the ATC frame) have no activity in T-strand synthesis compared with the native VirD1 polypeptide. Thus, the N-terminus of VirD1 must be essential to its optimal activity. Second, the native VirD1 translational start is strong enough to override translational termination in its vicinity; a T7-fusion construct that was not expected to yield a T7-VirD1 fusion (GAT frame) instead was able to produce genuine VirD1 polypeptide. Third, the strong T7 promoter and its cognate Shine-Dalgarno ribosome binding site fused to native *virD* upstream sequences did not yield the expected high levels of VirD1 polypeptide. Besides VirD1, a larger fusion polypeptide was synthesized in relatively high amounts presumably in initiation of translation at a rarely used codon. If the above factors are taken into account, it may be possible to design a vector for optimal VirD1 (and VirD2) expression, for example, by fusing the T7 promoter plus T7 Shine-Dalgarno sequences to within 10 nucleotides of the native VirD1 ATG translational start codon.

We have shown that *E. coli* can be used as a host cell to study the molecular reactions involved in the synthesis of the *Agrobacterium* T-strand transfer intermediate. This system is very amenable to manipulation, and future studies will aim to identify proteins in addition to VirD1, VirD2, and VirC1 that serve in this process. The data comparing nicking versus T-strand production with pLAFR and pACYC T-DNA plasmids suggest that VirD1 and VirD2 are primarily used in the border-nicking reaction. It will be interesting to determine whether VirC1 or other host cell or pLAFR-encoded functions play roles in the unwinding, processing, or replacement strand synthesis that must be coupled to T-strand production. Ultimately, these types of analyses will be useful to the design of methods to increase the efficiency of T-strand production and its effective transfer to plant cells.

ACKNOWLEDGMENTS

We thank all the members of our laboratory for helpful and stimulating discussion of the data herein.

P. Z. was supported by NSF grant DMB-8617772.

LITERATURE CITED

Albright, L. M., Yanofsky, M. F., Leroux, B., Ma, D., and Nester, E. W. 1987. Processing of the T-DNA of *A. tumefaciens* generates border nicks and linear, single stranded T-DNA. J. Bacteriol. 169:1046-1055.

Alt-Moerbe, J., Rak, B., and Schroeder, J. 1986. A 3.6 kbp segment from the *vir* region of Ti plasmids contains genes responsible for border sequence directed production of T-region circles in *E. coli*. EMBO J. 5:1129-1135.

Bagdasarian, M. M., Amann, E., Lurz, R., Ruckert, B., and Bagdasarian, M. 1983. Activity of the hybrid *trp-lac* (*tac*) promoter of *Escherichia coli* in *Pseudomonas putida*. Construction of broad-host range controlled expression vectors. Gene 26:273-282.

Chang, A. C. Y., and Cohen, S. N. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. 134:1141-1156.

Christie, P. J., Ward, J. E., Winans, S., and Nester, E. W. 1988. The *Agrobacterium tumefaciens virE*2 gene product is a single stranded DNA binding protein that associates with T-DNA. J. Bacteriol. 170:2659-2667.

Citovsky, V., DeVos, G., and Zambryski, P. 1988. A novel single stranded DNA binding protein encoded by the *virE* locus is produced following activation of the *A. tumefaciens* T-DNA transfer process. Science 240:501-504.

Citovsky, V., Wong, M. L., and Zambryski, P. 1989. Cooperative interaction of *Agrobacterium virE*2 protein with single stranded DNA: Implications for the T-DNA transfer process. Proc. Natl. Acad. Sci. USA. (In press.)

Close, T. J., Tait, R. C., Rempel, H. C., Hirooka, T., Kim, L., and Kado, C.I. 1987. Molecular characterization of the *virC* genes of the Ti plasmid. J. Bacteriol. 169:2336-2344.

Das, A. 1988. The *A. tumefaciens virE* locus encodes a single stranded DNA binding protein. Proc. Natl. Acad. Sci. USA 85:2909-2913.

de Boer, H. A., Comstock, L. J., and Vasser M. 1983. The *tac* promoter: A functional hybrid derived from the *trp* and *lac* promoters. Proc. Natl. Acad. Sci. USA 80:21-25.

Depicker, A., DeWilde, M., DeVos, G., DeVos, R., Van Montagu, M., and Schell, J. 1980. Molecular cloning of overlapping segments of the nopaline Ti plasmid of pTiC58 as a means to restriction endonuclease mapping. Plasmid 3:193-211.

Friedman, A. M., Long, S. R., Brown, S. E. Buikema, W. J., and Ausubel, F. M. 1982. Construction of a broad host range cloning vector and its use in the genetic analysis of *Rhizobium* mutants. Gene 11:291-298.

Gietl, C., Koukolikova-Nicola, Z., and Hohn, B. 1987. Mobilization of T-DNA from *Agrobacterium* to plant cells involves a protein that binds single stranded DNA. Proc. Natl. Acad. Sci. USA. 84:9006-9010.

Gold, L. 1988. Post-transcriptional regulatory mechanisms in *E. coli*. Annu. Rev. Biochem. 57:199-233.

Hagiya, M., Close, T. J., Tait, R. C., and Kado, C. I. 1985. Identification of pTiC58 plasmid encoded proteins for virulence in *A. tumefaciens*. Proc. Natl. Acad. Sci. USA 82:2669-2673.

Hallewell, R. A., Masiarz, F. R., Najarian, R. C., Puma, J. P., Quiroga, M. R., Randolph, A., Sanchez-Pescador, R., Scandella, C. J., Smith, B., Steimer, K. S., and Mullenbach, G. T. 1985. Human Cu/Zn superoxide dismutase cDNA: Isolation of clones synthesizing high levels of active or inactive enzyme from an expression library. Nucleic Acids Res. 13:2017-2034.

Jayaswal, R. K., Veluthambi, K., Gelvin, S. B., and Slightom, J. L. 1987. Double stranded T-DNA cleavage and the generation of single stranded T-DNA molecules in *E. coli* by a *virD* encoded border specific endonuclease from *A. tumefaciens*. J. Bacteriol. 169:5035-5045.

Leroux, B., Yanofsky, M. F., Winans, S. C., Ward, J. E. Ziegler, S. F., and Nester, E. W. 1987. Characterization of the *virA* locus of *Agrobacterium tumefaciens*. EMBO J. 6:849-856.

Maniatis, T., Fritsch, E. F., and Sambrook, J. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Miller, J. H. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Porter, S. G., Yanofsky, M. F., and Nester, E. W. 1987. Molecular characterization of the *virD* operon from *A. tumefaciens*. Nucleic Acids Res. 15:7503-7517.

Ronson, C. W., Nixon, B. T., and Ausubel, F. M. 1987. Conserved domains in bacterial regulatory proteins respond to environmental stimuli. Cell 5:579-581.

Rosenberg, A. H., Lade, B. N., Chiu, D., Dunn, J. J., and Studier, W. 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene 56:125-135.

Sancar, A. Wharton, R. P., Selzer, S., Kacinski, B. M., Clarke, N. D., and Rupp, W. D. 1981. Identification of the *uvrA* gene product. J. Mol. Biol. 148:45-62.

Shinedling, S., Swebelius-Singer, B., Gayle, M., Pribnow, D., Jarvis, E., Edgar, B., and Gold, L. 1987. Sequences and studies of bacteriophage T4 rII mutants. J. Mol. Biol. 195:471-480.

Stachel, S. E. Messens, E., Van Montagu, M., and Zambryski, P. 1985. Identification of the signal molecules produced by wounded plant cells which activate the T-DNA transfer process in *Agrobacterium tumefaciens*. Nature 318:624-629.

Stachel, S. E., and Nester, E. W. 1986. The genetic and transcriptional organization of the *vir* region of the A6 Ti plasmid of *Agrobacterium*. EMBO J. 5:1445-1454.

Stachel, S. E., and Zambryski, P. 1986. *virA* and *virG* control the plant-induced activation of the T-DNA transfer process of *Agrobacterium tumefaciens*. Cell 46:325-333.

Stachel, S. E., Timmerman, B., and Zambryski, P. 1986. Generation of single stranded T-DNA molecules during the initial stages of T-DNA transfer from *Agrobacterium tumefaciens* to plant cells. Nature 322:706-712.

Stachel, S. E. Timmerman, B., and Zambryski, P. 1987. Activation of *Agrobacterium tumefaciens vir* gene expression generates multiple T-strand molecules from the pTiA6 T-region: Requirement for 5′ *virD* products. EMBO J. 6:857-863.

Studier, F. W., and Moffatt, B. A. 1986. Use of bacteriophage T7 RNA polymerase to direct selection of high level expression of cloned genes. J. Mol. Biol. 189:113-130.

Tait, R. C., and Kado, C. I. 1988. Regulation of the *virC* and *virD* promoters of pTiC58 by the *ros* chromosomal mutation of *Agrobacterium tumefaciens*. Mol. Microbiol. 2:385-392.

van den Elzen, P., Lee, K. Y., Townsend, J., and Bedbrook, J. 1985. Simple binary vectors for DNA transfer to plant cells. Plant Mol. Biol. 5:149-154.

Veluthambi, K., Jayaswal, R. K., and Gelvin, S. B. 1987. Virulence genes A, G, and D mediate the double stranded border cleavage of T-DNA from the *Agrobacterium* Ti plasmid. Proc. Natl. Acad. Sci. USA 84:1881-1885.

Veluthambi, K., Ream, W., and Gelvin, S. 1988. Virulence genes, borders, and overdrive generate single stranded T-DNA molecules from the A6 Ti plasmid of *Agrobacterium tumefaciens*. J. Bacteriol. 170:1523-1532.

Vieira, J., and Messing, J. 1987. Production of single stranded plasmid DNA. Methods Enzymol. 153:3-11.

Wang, K., Stachel, S. E., Timmerman, B., Van Montagu, M., and Zambryski, P. 1987. Site-specific nick occurs within the 25 bp transfer promoting border sequence following induction of *vir* gene expression in *Agrobacterium tumefaciens*. Science 235:587-591.

Winans, S. C., Ebert, P. R., Stachel, S. E., Gordon, M. P., and Nester, E. W. 1986. A gene essential for *Agrobacterium* virulence is homologous to a family of positive regulatory loci. Proc. Natl. Acad. Sci. USA 83:8278-8282.

Yanofsky, M. F., and Nester, E. W. 1986. Molecular characterization of a host-range determining locus from *Agrobacterium tumefaciens*. J. Bacteriol. 168:244-250.

Yanofsky, M. F., Porter, S. G., Young, C., Albright, L. M., Gordon, M. P., and Nester, E. W. 1986. The *virD* operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. Cell 47:471-477.

Zambryski, P. 1988. Basic processes underlying *Agrobacterium* mediated DNA transfer to plant cells. Annu. Rev. Genet. 22:1-30.