Overexpression of *virD1* and *virD2* Genes in *Agrobacterium tumefaciens* Enhances T-Complex Formation and Plant Transformation

K. WANG,[†] A. HERRERA-ESTRELLA, AND M. VAN MONTAGU*

Laboratorium voor Genetica, Rijksuniversiteit Gent, B-9000 Ghent, Belgium

Received 25 January 1990/Accepted 14 May 1990

The VirD1 and VirD2 proteins encoded by an inducible locus of the virulence (*vir*) region of the *Agrobacterium tumefaciens* Ti plasmid are required for site-specific nicking at T-DNA border sites. We have determined the nucleotide sequence of a 3.6-kilobase-pair fragment carrying the *virD* locus from nopaline Ti plasmid pTiC58. In contrast to the previous report (Hagiya et al., Proc. Natl. Acad. Sci. USA 82:2669-2673, 1985), we found that the first three open reading frames were capable of encoding polypeptides of 16.1, 49.7, and 21.4 kilodaltons. Deletion analysis demonstrated that the N-terminal conserved domain of VirD2 was absolutely essential for its endonuclease activity. When extra copies of the *virD1* and *virD2* genes were present in an *A. tumefaciens* strain carrying a Ti plasmid, increased amounts of T-strand and nicked molecules could be detected at early stages of *vir* induction. Such strains possessed the ability to transform plants with higher efficiency.

Agrobacterium tumefaciens is the etiological agent of plant crown gall disease. It was demonstrated that the crown galls are produced as the result of transfer of a segment of DNA, the transferred DNA (T-DNA), from the tumor-inducing (Ti) plasmid of the bacterium to the plant nuclear genome (13, 31). The T-DNA is delimited by imperfect 25-base-pair (bp) direct repeats (termed border sequences). The right 25-bp border sequence is crucial in *cis* for directing efficient T-DNA transfer into the plant nuclear genome (23, 40). Another set of elements responsible for the transfer process reside outside of the T-DNA, in the 35-kbp virulence (*vir*) region of the Ti plasmid. The *vir* region is composed of seven operons (*virA*, *virB*, *virC*, *virD*, *virE*, *virF*, and *virG*) (30). Except for the *virA* gene, which is constitutively expressed in *A. tumefaciens*, the other six operons are induced in response to small phenolic compounds, such as acetosyringone (AS), excreted by wounded plant cells (3, 33). Activation of *vir* gene expression results in the generation of site-specific nicks within the bottom strand of all T-DNA border repeats and a linear single-stranded DNA molecule (the T-strand) corresponding to the bottom strand of the T-DNA region (1, 34, 41, 44). The polarity of T-strands suggests that they are produced in a right-to-left direction, utilizing border nicks as initiation and termination sites. Both nicking and T-strand synthesis require two polypeptides encoded by the 5' half of the *virD* locus (*virD1* and *virD2*) (35, 44).

The T-strand is considered the intermediate that is transferred from *Agrobacterium* spp. to the plant cell. Recent work has provided more evidence to support this hypothesis. First, it has been shown that the *virE2* gene encodes a single-stranded DNA-binding protein which interacts with the T-strand nonspecifically (6-8, 14). As the transfer intermediate, the DNA molecule must be protected from exo- and endonucleolytic degradation during the whole transfer process, until it integrates into the plant nuclear genome.

The fact that *Agrobacterium* has evolved a single-stranded DNA-binding protein specifically under *vir* control supports the argument that the T-strand may be the intermediate molecule. Second, in addition to the endonuclease activity of the VirD1 and VirD2 proteins, the VirD2 protein was found to be firmly associated with the 5' end of the T-strand (the T-complex) (16, 18, 42, 45). The asymmetrical binding of the VirD2 protein to the T-strand (only at the 5' end) may therefore confer its property of polarity as an intermediate in the transfer process.

If the T-complex were the true transfer intermediate, one could expect a direct correlation between the expression of the *virD1* and *virD2* genes, the production of the T-strand, and tumor formation on plants. In this work, we show that when extra copies of *virD1* and *virD2* are present in *Agrobacterium* strains containing a wild-type Ti plasmid, increased amounts of T-strand can be detected soon after AS induction. In addition, increased expression of *virD1* and *virD2* leads to elevated tumorigenicity on plants.

Due to the inconsistencies between the sequences reported for the *virD* genes from different *Agrobacterium* species (10, 15, 17, 20) and our own results, we present our nucleotide sequence for the *virD* gene of the nopaline plasmid pTiC58. The discrepancies between our sequence and the previously reported one (15) are discussed. The amino acid comparison of the VirD2 proteins from different *Agrobacterium* species shows that its N-terminus but not the C-terminus is highly conserved. The C-terminal 50% of VirD2 can be deleted or replaced without affecting its endonuclease activity in *Escherichia coli*. Although the C-terminal domains of VirD2 proteins are highly divergent, they display similar hydropathy profiles. This suggests that this domain is involved in the steps subsequent to border cleavage and T-strand synthesis.

MATERIALS AND METHODS

Bacterial strains and plasmids. *E. coli* MC1061 (4) was used as the host for all cloning and T-strand assays. Plasmid pUC18 (43) was used as the vector for the various subclones. pGA472' (2) is a plasmid with a broad-host-range origin of

* Corresponding author.

[†] Present address: Research Department, Garst Seed Company, Slater, IA 50244.

replication which contains the right and left border fragments of nopaline T-DNA. pGV910 (Van den Eede et al., unpublished data) is a plasmid containing the origin of replication from the broad-host-range plasmid pVS1 (19). pND1, pND2, pND3, and pND4 (16) are pUC18 derivatives containing various parts of the nopaline *virD* genes. pGV3850 (46) is a nopaline Ti plasmid mutant in which all the oncogenic functions of the T-DNA were substituted by pBR322. pGV2301 (J.-P. Hernalsteens, unpublished data) is a derivative of the nopaline plasmid pTiC58 carrying the kanamycin resistance gene from transposon Tn903.

Plasmid constructions. Figure 3A illustrates the subclones of the *virD* gene in the pUC18 vector. Plasmid pND5 was constructed by introduction of a 1,556-bp *EcoRV* fragment from pND1 into the *SmaI* site of pUC18. Plasmid pND6 was constructed by deletion of a 108-bp *XmaIII* (coordinate 1696) to *EcoRI* (coordinate 1804) fragment from pND3.

A 3.8-kilobase (kb) *SacII-BamHI* fragment containing the origin of replication from pVS1 was obtained from pGV910 and ligated to the corresponding sites of pND1 to generate pVND1. The same fragment was introduced into the *NdeI* site of pND3 to generate pVND3. Plasmid pVND3-Km was constructed as follows. A 1.5-kb *EcoRI-NdeI* fragment carrying the neomycin phosphotransferase II (*nptII*) gene was isolated from plasmid pLKM91 (gift from J. Botterman), which was derived from pKm9 (27). This fragment was ligated into the *EcoRI* site of pVND3 to generate a VirD2-NPTII fusion protein.

Nucleotide sequence analysis. Nucleotide sequence was determined by the chain-termination method (28) after subcloning of the *virD* operon into plasmid pUC18. Reagents for DNA sequencing were from P-L Biochemicals. The computer program Genalign (Software; Intelligenetics, Inc.) was used for homology searching of proteins. The program Pep was used for hydrophobicity searching (prediction based on the Hopp and Wood index).

T-strand and border nick assays. The assays for the T-strand and the border nick in *E. coli* and *A. tumefaciens* were done as described previously (16).

NPTII assays. The neomycin phosphotransferase II (NPTII) activity of AS-induced *Agrobacterium* cells was assayed by the method of Platt and Yang (25). The NPTII activity of the VirD2-NPTII fusion protein in *E. coli* carrying pVND3-Km was examined by native polyacrylamide gel electrophoresis (PAGE) as described by Van den Broeck et al. (38).

Virulence assays on plants. Bacterial inoculations were performed on various plants as described by Wang et al. (39) for tobacco (*Nicotiana tabacum* W38) and for potato (*Solanum tuberosum* cv. Bintje), by Joos et al. (21) for kalanchoe (*Kalanchoe daigremontiana*), and by Peralta et al. (24) for carrot (*Daucus carota*).

RESULTS

Nucleotide sequence of the *virD* locus. Plasmid pND1 (16) contains a 3.7-kb *Sall-KpnI* fragment overlapping the *virD* locus of nopaline Ti plasmid pTiC58 (9). Figure 1 shows the nucleotide sequencing strategy with pND1 and the resulting sequence. The first three open reading frames (ORFs) encoded three polypeptides of 16.13, 49.65, and 21.35 kilodaltons (kDa). The fourth open reading frame was only partially sequenced. This sequence was reported previously by Hagiya et al. (15). However, a comparison of our sequence with their data showed that there were differences at 28 positions. For example, Hagiya et al. (15) suggested that

ORF1 encoded a polypeptide comprised of 109 amino acids, while we deduced that it should encode a polypeptide of 147 amino acids. This difference is due to two missing guanidine residues (coordinates 963 and 966, Fig. 1B) at corresponding positions 470 and 473 in the sequence of Hagiya et al. (15). The 147-amino-acid VirD1 deduced from our data was 93% homologous with the VirD1 proteins of *Agrobacterium rhizogenes* pRiA4b and the octopine Ti plasmid pTiA6NC (both 149 amino acids in length) (17, 20, 44).

Similar nucleotide discrepancies occurred in ORF3. This region was 597 bp long and encoded a polypeptide of 199 amino acids (Fig. 1B). The corresponding region determined by Hagiya et al. (15) was designated ORF4 and found to be capable of encoding a polypeptide of 273 amino acids. Sequence comparison revealed that this difference was due to a guanidine residue missing (coordinate 3008, Fig. 1B) at corresponding position 2503 in the sequence of Hagiya et al. (15). The VirD3 proteins had low homology between *Agrobacterium* species (17, 26). Neither molecular size nor amino acid composition was similar between octopine pTiA6NC and *A. rhizogenes* pRiA4b (17). Nevertheless, the VirD3 protein of 199 amino acids deduced by our data had a mass similar to that of the VirD3 protein of octopine pTiA6NC (201 amino acids [26]) and had 31% homology overall (data not shown).

Our sequencing data show that ORF2 of the *virD* locus was 1,338 bp in size and encoded a polypeptide of 49.7 kDa. This result does not agree with the earlier report that two polypeptides of 15 and 29 kDa should be produced from this region (15). Eighteen differences at the nucleotide level, including base insertions or deletions, have been found in this 1,338-bp-long sequence when it was compared with the corresponding sequence of Hagiya et al. (15). However, our data are consistent with the reported *virD2* sequences from the octopine strain (20, 44) and from *A. rhizogenes* (17). In both cases, VirD2 was found to encode a polypeptide of 48 kDa. Moreover, a 56-kDa protein was detected in *E. coli* maxicells when a VirD2 antibody was used (10). Our work provided a similar result. An *XmnI* fragment (corresponding to positions 1119 to 2000, Fig. 1B) carrying part of the *virD2* gene was cloned behind the T7 RNA polymerase promoter (36). This fragment produced a truncated polypeptide of 38 kDa when expressed in *E. coli* (data not shown), which is in agreement with the predicted polypeptide of 33 kDa from this fragment. No polypeptide of 15 kDa was detected from the corresponding region (positions 1125 to 1553). Based on these data, we conclude that the nopaline VirD2 protein consists of 446 amino acids with a molecular mass of 49.7 kDa.

Function analysis of the *virD2* gene. A comparison of amino acid sequences of the VirD2 proteins from pTiC58, pTiA6NC, and pRiA4b revealed high homology throughout the N-terminal half (90% identity at residues 1 to 203), while the C-terminal half was completely divergent (Fig. 2). A deletion of 50% of the VirD2 protein at its C-terminal part (up to the *EcoRI* site in pND3, Fig. 3A) did not affect border-nicking or T-strand-binding capacity in *E. coli* (Fig. 3B, lane 1) (9, 16).

To localize the domain essential for nicking and binding activity of the VirD2 protein, more deletions were made. Two constructs, pND6 and pND5, carrying 3' deletions of *virD2* up to the *XmaIII* and *EcoRV* sites (Fig. 3A), respectively, were introduced into an *E. coli* strain which harbored border-carrying plasmid pGA472' (2). *E. coli* cells containing each of the constructs and pGA472' were lysed by sodium dodecyl sulfate (SDS). DNAs from the phenol-aqueous

FIG. 1. Nucleotide sequence of *virD* operon of pTiC58 (GenBank accession no. M33673). (A) Sequencing strategy. Sequencing was done in both directions (as indicated by the arrows) starting from the chosen restriction enzyme sites. B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hind*III; K, *Kpn*I; N, *Nru*I; S, *Sal*I. ORF, Open reading frame. Numbers correspond to those assigned previously (9) to the fragments obtained upon restriction of plasmid pTiC58 with the corresponding enzymes. (B) Nucleotide sequence and deduced amino acid sequence.

FIG. 2. Comparison of amino acid sequence of VirD2 proteins from octopine pTiA6NC (44), nopaline pTiC58, and *A. rhizogenes* pRiA4b (17). The amino acids in the shaded areas are identical. Double dots indicate functionally conservative amino acid substitutions. Single dots indicate that two of three amino acid substitutions are functionally equivalent. Nonconservative differences are indicated by lowercase letters.

interphase and aqueous phase after phenol extraction were examined for the presence of T-strand molecules (16). Since VirD2 was demonstrated to bind firmly (covalently) at the 5' end of the T-strand molecule, the T-strand-VirD2 protein complex (T-complex) was expected to be present only in the phenol-aqueous interphase (Fig. 3B, lanes 1 and 4) (16). Figure 3B shows that bacteria carrying plasmid pND6 or pND5 no longer produced T-complex (Fig. 3A, lanes 2 and 3), indicating that the 108-bp DNA sequence between the *Xma*III and *Eco*RI sites of *virD2* is essential for generating the T-complex. Since these constructs have lost part of the

VirD2 conserved region, it suggests that most, if not all, of the N-terminal domain is functionally involved in nicking and T-strand formation, while the C-terminal domain is not required for the enzymatic activity.

It has been shown that the β -galactosidase gene fused to *vir* genes can be regulated under the *virA* and *virG* system (31). We have used a similar approach to characterize gene regulation but used the *nptII* gene (27). The construct pVND3-Km carries a chimeric *virD2* gene with the *nptII* gene translationally fused in the middle of the reading frame (see Materials and Methods). This plasmid was tested for its

FIG. 3. Functional analysis of *virD2* in *E. coli*. (A) Various subclones of *virD1* and *virD2* in plasmid pUC18. Boxes represent ORF1, ORF2, ORF3, and part of ORF4 of the *virD* operon. Numbers above restriction enzyme sites indicate their coordinates. B, *Bam*HI; E, *Eco*RI; E5, *Eco*RV; P, *Pst*I; S, *Sal*I; X, *Xma*III. (B) Autoradiography of T-strand assay in *E. coli*. *E. coli* strains containing pGA472' and various *virD* subclones were lysed by SDS and phenol extracted. DNAs from the phenol-aqueous interphase (int) or aqueous phase (aq) were treated with pronase and precipitated. After agarose electrophoresis, DNA samples were blotted without denaturation and hybridized with ³²P-labeled plasmid pGA472'.

FIG. 4. VirD2-NPTII fusion. (A) T-strand assay in *E. coli* carrying various plasmids. All manipulations are described in the legend to Fig. 3B. Lanes: a, pND3 plus pGA472'; b, pVND3-Km plus pGA472'; c, pGA472'. (B) Native PAGE for detecting the VirD2-NPTII (VD2-NPT) fusion protein. Lanes: 1, pVND3-Km; 2, control strain carrying wild-type *nptII*. (C) NPTII assay of *A. tumefaciens*. -AS, Cells incubated without AS; +AS, cells incubated with AS.

ability to generate T-strands in *E. coli* (Fig. 4A). Neomycin phosphotransferase activity was assayed by native PAGE as described previously (38). Figure 4B shows that the fusion product had a different mobility in native polyacrylamide gels. The same plasmid was introduced into an *Agrobacterium* strain carrying a wild-type Ti plasmid, pGV2301 (J.-P. Hernálsteens, unpublished data). As can be seen in Fig. 4C, the enzymatic activity of the *nptII* gene was increased about 10-fold when *A. tumefaciens*(pVND3-Km pGV2301) was incubated with AS (Fig. 4C). These data suggest that the VirD2-NPTII fusion protein is being regulated by the *virA* and *virG* system, as previous studies showed for a *virD-lacZ* fusion (31). The fact that the VirD2-NPTII fusion protein is active in T-strand production provides further evidence that the C-terminal part of the protein is not essential for this function. The C-terminus can be replaced by a totally different peptide without disturbing its activity. On the other hand, introduction of this construct into *A. tumefaciens* allowed us to monitor AS induction with a system different from the widely used *vir-lacZ* fusions. In this case, the level of induction of the different *vir* genes can be studied without interference, due to the well-known stability of the LacZ protein (37).

Overexpression of *virD1* and *virD2* increases T-strand production in *A. tumefaciens* at early times after induction. In previous work we had observed that subclones of the *virD* gene in a pUC18 vector resulted in a higher yield of T-complex molecules than when cloned in a pBR322 vector in *E. coli* (16). Since both *virD* clones contained the *virD* promoter, the difference in yield is probably due to the higher copy number of the pUC18 vector. The T-complex molecule has been hypothesized to be a transfer intermediate; therefore, it was interesting to investigate whether increasing the copy number of the *virD* gene would lead to increased amounts of T-strand in *A. tumefaciens* and, in turn, increased oncogenicity on plants.

A 4-kb fragment containing a broad-host-range origin of replication from pVS1 (19) was introduced into pND2 (see Materials and Methods). This new construct, pVND1, carrying *virD1* and 50% of *virD2*, was mobilized into *A. tumefaciens* pGV3850 (46). *Agrobacterium* strains carrying pVND1 and pGV3850 thus harbored six to seven extra copies of the *virD1* and *virD2* genes in addition to their wild-type *virD* gene. Total DNA from AS-induced *Agrobacterium* was digested with *Hind*III and subjected to alkaline gel electrophoresis (22). Under these conditions, the nicked border fragments generated by the endonuclease activity upon *Hind*III digestion could be strand separated. The separated molecules were blotted onto a nitrocellulose membrane in 20× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and hybridized with the ³²P-labeled fragment *Hind*III-23 containing the right border.

Figure 5B shows the effect on generation of border nicks and T-strands at different time points after AS induction. At 4 h after AS induction, border nick products and T-strand molecules were already detected in the strain carrying pVND1 and pGV3850 (Fig. 5B, lane 2; Fig. 5C, lane 2), while little or no nick products and few T-strands were detected in the strain carrying pGV3850 alone (Fig. 5B, lane 1; Fig. 5C, lane 1). A similar effect was observed after 8 h of AS induction (Fig. 5B, lanes 3 and 4; Fig. 5C, lanes 3 and 4). However, after 16 h of AS induction, this enhancing effect of extra copies of *virD1* and *virD2* on nick and T-strand production in the strain harboring pVND1 and pGV3850 was no longer significant in comparison with that in the strain carrying pGV3850 alone (Fig. 5B and C, lanes 5, 6, 7, and 8). When incubated in the absence of AS, neither strain produced T-strands (Fig. 5C, lanes 9 and 10). This indicates that even though the rate of induction was increased, the final number of nick products and T-strand molecules reached the same level in both *A. tumefaciens* strains.

Overproduction of VirD1 and VirD2 proteins in *A. tumefaciens* elevates transformation efficiency on plants. Plasmid pVND1 was also introduced into *A. tumefaciens* carrying pGV2301. This strain harbors a wild-type nopaline Ti plasmid (J.-P. Hernálsteens, unpublished data). The effect of overproduction of VirD1 and VirD2 on plant transformation was examined by assaying tumorigenicity. Tobacco leaf disks, kalanchoe stems, carrot root slices, and potato tuber disks were chosen as plant material for *Agrobacterium* infection.

The responses of kalanchoe stems to the strain carrying pVND1 and pGV2301 could be observed a few days earlier than those to the strain carrying only pGV2301. The size of tumors caused by the strain carrying pVND1 and pGV2301 was significantly larger than that caused by the strain carrying pGV2301. For example, after 19 days of infection, tumors caused by *A. tumefaciens*(pVND1 pGV2301) were about 8 mm in diameter, while tumors caused by *A. tumefaciens*(pGV2301) were about 3 mm in diameter. Moreover, more roots were generated on the infection sites of *A. tumefaciens*(pVND1 pGV2301) (Fig. 6A). A similar effect occurred on carrot root slices (Fig. 6B). More tumors were observed on slices infected by *A. tumefaciens*(pVND1 pGV2301).

To obtain quantitative data for comparison of transformation efficiency, an assay system on potato tuber disks was used. Uniformly prepared potato tuber disks made by previously established procedures (39), were inoculated with equal amounts bacterial cells, and the number of individual tumors was then scored after 14 days of infection. The results are presented in Table 1. The strain carrying

FIG. 5. Kinetics of the generation of border nicks and T-strands. (A) Possible T-DNA border structure of pGV3850 (46) T-DNA after AS induction. The vertical bars represent *Hind*III restriction sites. *Hind*III-10 is a 6.5-kb *Hind*III fragment containing the left border sequence, and *Hind*III-23 is a 3.3-kb *Hind*III fragment containing the right border sequence. The T-DNA (dotted line) is from pBR322 in this case. Arrows represent the border nicks at the bottom strand of T-DNA after AS induction. Solid circles represent the binding proteins at the 5' ends of nicked products. (B) Autoradiography of alkaline gel electrophoretic analysis. Bacteria were collected after 4, 8, 16, and 24 h of incubation with AS. Total agrobacterial DNA was extracted by the pronase treatment procedure described before (16). DNAs were digested with *Hind*III, separated by alkaline gel electrophoresis, and blotted in 20× SSC. Fragment *Hind*III-23 was used as the probe. The solid triangle indicates *Hind*III-23; the open triangles indicate nick products (sizes shown in kilobases). The arrowhead indicates T-strand molecules. Lanes: a, strain carrying pGV3850; b, strain carrying pVND1 plus pGV3850. Note that all fragments indicated here are single stranded. Other bands are probably the result of cross-hybridization of contaminating pBR322 sequences in the probe, since the *Hind*III-23 fragment was subcloned in a pBR322 vector. (C) T-strand assay in *Agrobacterium* cells. All sample manipulations and the nondenaturing gel system have been described before (16). ³²P-labeled plasmid pBR322 was used as a probe. -AS, Cells incubated without AS. Lanes: a, strain carrying pGV3850; b, strain carrying pVND1 and pGV3850.

pGV2301 produced on average 71.3 tumors per disk (30 disks were assayed). The strain carrying pVND1 and pGV2301 produced on average 102.7 tumors per disk (32 disks were assayed).

Since the data were derived from a large sample size, the probability that the VirD1- and VirD2-overproducing strain carrying pVND1 and pGV2301 is different from the strain carrying only wild-type pGV2301 in its tumor-forming ability could be assessed statistically. As shown in Table 1, a *t* value of 3.78 indicates that the probability that this difference is due to chance is less than 0.001, suggesting that the two strains have a real discrepancy in their ability to form tumors on potato tuber disks.

FIG. 6. Virulence assays on plants. For each panel, the top shows the result with a strain carrying pGV2301 only, and the bottom shows the result with a strain carrying pVND1 plus pGV2301. (A) Inoculations on stems of kalanchoe. (B) Inoculations on root slices of carrots.

While the elevated transformation efficiency of the strain carrying pVND1 and pGV2301 was observed on kalanchoe, carrot, and potato, this difference was not significant when tobacco leaf disks were used as an assay system (data not shown). This is actually consistent with the previous observation that the potato system is more sensitive in detecting differences in tumorigenicity between strains (39).

DISCUSSION

The nucleotide sequence of the *virD* locus from the nopaline-type Ti plasmid pTiC58 was determined in this work. In contrast to the previous report on the same DNA fragment (15), our analysis predicts that the first three open reading frames of this locus could produce three polypeptides of 16.1, 49.7, and 21.4 kDa.

VirD1 proteins have been found to have high homology (93%) between *Agrobacterium* species (K. Wang, unpublished data). This protein is absolutely needed for catalyzing site-specific and strand-specific cleavage at the T-DNA border sequences with the VirD2 protein (16, 18, 35, 44). Recently, it has been reported that the VirD1 protein possesses topoisomerase activity capable of converting supercoiled DNA (form I) to relaxed DNA (form IV) (12). The function of VirD3 is not yet known. This protein is not conserved, since there is only 31% homology between the equivalent proteins of octopine and nopaline Ti plasmids (K. Wang, unpublished data). Deletion of the *virD3* gene has no effect on T-complex formation in *E. coli* (10, 16). However, the need for the *virD3* gene for tumorigenicity has yet to be determined.

The *virD2* gene encodes a protein which is involved in the AS-induced activity for nicking at the T-DNA border sequences and T-strand generation (35, 44). This protein is also found to bind firmly (covalently?) to the 5' end of the T-strand (16, 18, 42, 45). The striking feature of this protein, in contrast to many other *vir*-encoded peptides, is the evolutionary conservation of the amino acid sequence throughout the N-terminal half of the protein (Fig. 2). Ninety percent homology was found between residues 1 and 203 (45% of the total sequence) of VirD2 from nopaline, octopine, and *A. rhizogenes* Ti plasmids, whereas the remaining 55% of the sequence had only 26% homology. Twenty-seven residues in the first 203 amino acids were different

TABLE 1. Oncogenicity of *A. tumefaciens* carrying various plasmids on potato tuber disks^a

Disk no.	No. of tumors		Disk no.	No. of tumors		Disk no.	No. of tumors	
	pGV2301	pVND1 + pGV2301		pGV2301	pVND1 + pGV2301		pGV2301	pVND1 + pGV2301
1	48	84	11	86	91	21	38	139
2	40	67	12	96	120	22	85	96
3	61	56	13	86	99	23	145	46
4	57	87	14	78	107	24	112	102
5	49	61	15	55	121	25	113	93
6	31	95	16	65	140	26	91	88
7	28	103	17	64	143	27	104	56
8	65	93	18	80	73	28	106	69
9	35	214	19	49	60	29	55	170
10	88	162	20	66	116	30	63	117

^a Mean ($\pm$ SD) number of tumors per disk: pGV2301, 71.3 $\pm$ 27.9; pVND1 plus pGV2301, 102.7 $\pm$ 36.6; $t_{0.001}$, 3.78.

between bacterial strains. Among these differences, 10 residues were conserved substitutions, e.g., arginine for lysine, valine for leucine, alanine for threonine, and aspartic acid for glutamic acid. The conserved N-terminal 50% of VirD2 is essential for the activity of border-nicking and T-strand binding, since a deletion located at the end of the first 203-amino-acid domain (Fig. 3A, pND6) abolished this activity completely (Fig. 3). However, deletions at the variable region of the C-terminal 50% (residues 227 to 447) do not affect T-complex formation in *E. coli* (10, 16).

The role of the C-terminal half is not quite understood. A computer homology search indicated that while this domain was poorly conserved, all VirD2 proteins displayed a very similar hydropathy profile in this region (Fig. 7). Genetic studies showed that when the C-terminus of VirD2 was replaced by β -galactosidase (30) or NPTII protein up to amino acid 227 (this work), T-strand formation nevertheless occurred (35) (Fig. 4A). Moreover, the N-terminal half of the protein was sufficient to stimulate tumor formation by *A. tumefaciens* (Table 1). On the other hand, an *Agrobacterium* strain harboring a *lacZ* insertion mutation in the *virD2* sequence encoding the C-terminus was avirulent (30).

It has been observed that there are conserved amino acid clusters at the C-terminus which resemble the consensus nuclear localization signal found in different nuclear proteins (5, 32). In fact, in addition to the two potential nucleus-targeting signals at its C-terminus (clusters KRPR at positions 417 to 420 and KRXR at positions 431 to 434, Fig. 2), a third putative signal sequence was found at positions 31 to 34 (RKGR). A truncated VirD2 (the first 292 amino acids) containing only this putative signal from the N-terminal half was able to act as a nucleus-targeting sequence in plants (A. Herrera-Estrella, M. Van Montagu, and K. Wang, unpublished data). Therefore, the remaining two postulated signals at the C-terminal region could act as auxiliary nucleus-targeting signals which enhance the efficiency of transport, as shown in animal systems (11). That a *lacZ* insertion mutation in the C-terminus-encoding part of *virD2* was avirulent may be the result of a change of the tertiary structure around the insertion, interfering with possible protein-protein interactions required at a later stage in the T-DNA transfer process.

Overexpression of the *virD1* and *virD2* genes in *A. tumefaciens* accelerates both border cleavage and T-strand formation. After 4 h of AS induction, more T-strands were produced in a strain harboring pVND1 and pGV3850 than in a strain harboring pGV3850 alone. The amount of border nick products was also increased by the presence of more copies of *virD*. While nick products were only seen after 8 h

of AS induction in the strain carrying pGV3850, border nicks could be detected after 4 h of incubation with AS in the strain carrying pVND1 and pGV3850.

A temporal difference in the generation of T-strand and the border nick products was also observed. A few hours after incubation of *A. tumefaciens* with AS, T-strand formation could already be detected, but not border nick products (Fig. 5B, lanes 1 and 2). This delay in production of border nicks could be explained by a DNA synthesis displacement mechanism at the T-DNA. The transfer process is initiated by cleavage of the bottom strand on the T-DNA borders by VirD proteins. The second step is the displacement of the 5' to the 3' nicked border, producing a single-stranded molecule, T-strand. The nicks at the bottom strand of T-DNA borders are simultaneously filled by the replication system with the T-DNA top strand as the template. Thus, although the border cleavages presumably occur prior to the formation of T-strand, nick products are detected later than T-strand production in an induced bacterial population.

It is interesting that the enhancing effect on the generation of T-complex by overexpression of *virD1* and *virD2* in *A. tumefaciens* can only be observed at early hours of AS induction. This suggests that the generation of T-complex is regulated. When the amount of substrate, the T-DNA border sequences, is limiting, increasing the amount of the VirD1 and VirD2 proteins accelerates T-strand formation, but the total amount of T-strand is not elevated. It is unlikely that the increased amount of T-complex was degraded in bacterial cells, since it is known that the T-strand is linked firmly by VirD2 at its 5' end and covered entirely by VirE2 proteins (6, 29). Such molecules would not be easily attacked by nucleases. Another possibility is that excess T-complex is transported out of the bacteria. However, there is no evidence of T-complex being secreted (34). In addition, successful transfer of T-DNA needs cell interaction between bacterium and plant. The third possibility is that the amount of VirD proteins is regulated in an induced bacterial cell. Excess VirD proteins are eventually degraded when the border sequences as substrates are limiting.

The transformation efficiency was elevated on kalanchoe, carrot, and potato, which were infected by an *Agrobacterium* strain carrying extra copies of the *virD1* and *virD2* genes. This was probably correlated with the increased T-complex production soon after the *virD* genes were induced. When more VirD proteins were available in the cell, the reaction between VirD proteins and the T-DNA border sequences accelerated, and T-complex molecules were generated earlier. As a putative transfer intermediate, T-complex might therefore enhance plant transformation effi-

FIG. 7. Hydropathy profile of VirD2 proteins of octopine (pTiA6NC), nopaline (pTiC58), and *A. rhizogenes* (pRiA46) plasmids. Regions above the horizontal axis indicate hydrophilicity, whereas those below the axis indicate hydrophobicity.

ciency. Further work is needed to determine whether plants transformed by VirD-overproducing strains contain more T-DNA insertions than those transformed by the wild-type *A. tumefaciens*.

ACKNOWLEDGMENTS

We thank A. Caplan and M. Holsters for discussions and critical reading of the manuscript, R. Villarroel and J. Gielen for DNA

sequencing, Z.-M. Chen for plant work, J. Coppieters for computer analysis, and M. De Cock, S. Van Gijsegem, and V. Vermaercke for preparation of the manuscript.

K.W. and A.H.E. are indebted to the CEC for a training grant. This work was supported by a grant from the Services of the Prime Minister (UIAP 12.0C01.87).

LITERATURE CITED

1. Albright, L. M., M. F. Yanofsky, B. Leroux, S. Ma, and E. W. Nester. 1987. Processing of the T-DNA of *Agrobacterium tumefaciens* generates border nicks and linear, single-stranded T-DNA. *J. Bacteriol.* **169**:1046-1055.
2. An, G., B. D. Watson, S. Stachel, M. P. Gordon, and E. W. Nester. 1985. New cloning vehicles for transformation of higher plants. *EMBO J.* **4**:277-284.
3. Bolton, G. W., E. W. Nester, and M. P. Gordon. 1986. Plant phenolic compounds induce expression of the *Agrobacterium tumefaciens* loci needed for virulence. *Science* **232**:983-985.
4. Casadaban, M. J., and S. N. Cohen. 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.* **138**:179-207.
5. Chelsky, D., R. Ralph, and G. Jonak. 1989. Sequence requirements for synthetic peptide-mediated translocation to the nucleus. *Mol. Cell. Biol.* **9**:2487-2492.
6. Christie, P. J., J. E. Ward, S. C. Winans, and E. W. Nester. 1988. The *Agrobacterium tumefaciens* virE2 gene product is a single-stranded DNA-binding protein that associates with T-DNA. *J. Bacteriol.* **170**:2659-2667.
7. Citovsky, V., G. De Vos, and P. Zambryski. 1988. Single-stranded DNA-binding protein encoded by the virE locus of *Agrobacterium tumefaciens*. *Science* **240**:501-504.
8. Das, A. 1988. *Agrobacterium tumefaciens* virE operon encodes a single-stranded DNA-binding protein. *Proc. Natl. Acad. Sci. USA* **85**:2909-2913.
9. Depicker, A., M. De Wilde, G. De Vos, R. De Vos, M. Van Montagu, and J. Schell. 1980. Molecular cloning of overlapping segments of the nopaline Ti-plasmid pTiC58 as a means to restriction endonuclease mapping. *Plasmid* **3**:193-211.
10. De Vos, G., and P. Zambryski. 1989. Expression of *Agrobacterium* nopaline-specific VirD1, VirD2, and VirC1 proteins and their requirement for T-strand production in *E. coli*. *Mol. Plant-Microbe Interactions* **2**:43-52.
11. Dworetzky, S. I., R. E. Lanford, and C. M. Feldherr. 1988. The effects of variations in the number and sequence of targeting signals on nuclear uptake. *J. Cell Biol.* **107**:1279-1287.
12. Ghai, H., and A. Das. 1989. The virD operon of *Agrobacterium tumefaciens* Ti plasmid encodes a DNA-relaxing enzyme. *Proc. Natl. Acad. Sci. USA* **86**:3109-3113.
13. Gheysen, G., P. Dhaese, M. Van Montagu, and J. Schell. 1985. DNA flux across genetic barriers: the crown gall phenomenon, p. 11-47. In B. Hohn and E. S. Dennis (ed.), *Genetic flux in plants* (Advances in plant gene research, vol. 2). Springer Verlag, Vienna.
14. Gietl, C., Z. Koukolíková-Nicola, and B. Hohn. 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.
15. Hagiya, M., T. J. Close, R. C. Tait, and C. I. Kado. 1985. Identification of pTiC58 plasmid-encoded proteins for virulence in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. USA* **82**:2669-2673.
16. Herrera-Estrella, A., Z.-M. Chen, M. Van Montagu, and K. Wang. 1988. VirD proteins of *Agrobacterium tumefaciens* are required for the formation of a covalent DNA-protein complex at the 5' terminus of T-strand molecules. *EMBO J.* **7**:4055-4062.
17. Hirayama, T., T. Muranaka, H. Ohkawa, and A. Oka. 1988. Organization and characterization of the virCD genes from *Agrobacterium rhizogenes*. *Mol. Gen. Genet.* **213**:229-237.
18. Howard, E. A., B. A. Winsor, G. De Vos, and P. Zambryski. 1989. Activation of the T-DNA transfer process in *Agrobacterium* results in the generation of a T-strand-protein complex: tight association of VirD2 with the 5' ends of T-strands. *Proc. Natl. Acad. Sci. USA* **86**:4017-4021.

19. Itoh, Y., J. M. Watson, D. Haas, and T. Leisinger. 1984. Genetic and molecular characterization of the *Pseudomonas* plasmid pVS1. *Plasmid* 11:206-220.
20. Jayaswal, R. K., K. Veluthambi, S. B. Gelvin, and J. L. Slightom. 1987. Double-stranded cleavage of T-DNA and generation of single-stranded T-DNA molecules in *Escherichia coli* by a *virD*-encoded border-specific endonuclease from *Agrobacterium tumefaciens*. *J. Bacteriol.* 169:5035-5045.
21. Joos, H., B. Timmerman, M. Van Montagu, and J. Schell. 1983. Genetic analysis of transfer and stabilization of *Agrobacterium* DNA in plant cells. *EMBO J.* 2:2151-2160.
22. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. Peralta, E. G., and L. W. Ream. 1985. T-DNA border sequences required for crown gall tumorigenesis. *Proc. Natl. Acad. Sci. USA* 82:5112-5116.
24. Peralta, E. G., R. Hellmiss, and W. Ream. 1986. Overdrive, a T-DNA transmission enhancer on the *A. tumefaciens* tumour-inducing plasmid. *EMBO J.* 5:1137-1142.
25. Platt, S. G., and N.-S. Yang. 1987. Dot assay for neomycin phosphotransferase activity in crude cell extracts. *Anal. Biochem.* 162:529-535.
26. Porter, S., M. F. Yanofsky, and E. W. Nester. 1987. Molecular characterization of the *virD* operon from *Agrobacterium tumefaciens*. *Nucleic Acids Res.* 15:7503-7517.
27. Reiss, B., R. Sprengel, and H. Schaller. 1984. Protein fusions with the kanamycin resistance gene from transposon Tn5. *EMBO J.* 3:3317-3322.
28. Sanger, F. 1981. Determination of nucleotide sequences in DNA. *Science* 214:1205-1210.
29. Sen, P., G. J. Pazour, D. Anderson, and A. Das. 1989. Cooperative binding of *Agrobacterium tumefaciens* VirE2 protein to single-stranded DNA. *J. Bacteriol.* 171:2573-2580.
30. Stachel, S. E., and E. W. Nester. 1986. The genetic and transcriptional organization of the *vir* region of the A4 Ti plasmid of *Agrobacterium tumefaciens*. *EMBO J.* 5:1445-1454.
31. Stachel, S. E., and P. Zambryski. 1986. *Agrobacterium tumefaciens* and the susceptible plant cell: a novel adaptation of extracellular recognition and DNA conjugation. *Cell* 47:155-157.
32. Stachel, S. E., and P. C. Zambryski. 1989. Generic trans-kingdom sex? *Nature (London)* 340:190-191.
33. Stachel, S. E., E. Messens, M. Van Montagu, and P. Zambryski. 1985. Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. *Nature (London)* 318:624-629.
34. Stachel, S. E., B. Timmerman, and P. Zambryski. 1986. Generation of single-stranded T-DNA molecules during the initial stages of T-DNA transfer from *Agrobacterium tumefaciens* to plant cells. *Nature (London)* 322:706-712.
35. Stachel, S. E., B. Timmerman, and P. Zambryski. 1987. Activation of *Agrobacterium tumefaciens vir* gene expression generates multiple single-stranded T-strand molecules from the pTiA6 T-region: requirements for 5' *virD* gene products. *EMBO J.* 6:857-863.
36. Tabor, S., and C. C. Richardson. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* 82:1074-1078.
37. Teeri, T. H., H. Lehvshlaiho, M. Franck, J. Uotila, P. Heino, E. T. Palva, M. Van Montagu, and L. Herrera-Estrella. 1989. Gene fusions to *lacZ* reveal expression patterns of chimeric genes in transgenic plants. *EMBO J.* 8:343-350.
38. Van den Broeck, G., M. P. Timko, A. P. Kausch, A. R. Cashmore, M. Van Montagu, and L. Herrera-Estrella. 1985. Targeting of a foreign protein to chloroplasts by fusion to the transit peptide of ribulose 1,5-bisphosphate carboxylase. *Nature (London)* 313:358-363.
39. Wang, K., C. Genetello, M. Van Montagu, and P. Zambryski. 1987. Sequence context of the T-DNA border repeat element determines its relative activity during T-DNA transfer to plant cells. *Mol. Gen. Genet.* 210:338-346.
40. Wang, K., L. Herrera-Estrella, M. Van Montagu, and P. Zambryski. 1984. Right 25-bp terminus sequences of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. *Cell* 38:455-462.
41. Wang, K., S. Stachel, B. Timmerman, M. Van Montagu, and P. Zambryski. 1987. Site-specific nick in the T-DNA border sequence following *vir* gene expression in *Agrobacterium*. *Science* 235:587-591.
42. Ward, E. R., and W. M. Barnes. 1988. VirD2 protein of *Agrobacterium tumefaciens* is very tightly linked to the 5' end of T-strand DNA. *Science* 242:927-930.
43. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-119.
44. Yanofsky, M. F., S. G. Porter, C. Young, L. M. Albright, M. P. Gordon, and E. W. Nester. 1986. The *virD* operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. *Cell* 47:471-477.
45. Young, C., and E. W. Nester. 1988. Association of the VirD2 protein with the 5' end of T strands in *Agrobacterium tumefaciens*. *J. Bacteriol.* 170:3367-3374.
46. Zambryski, P., H. Joos, C. Genetello, J. Leemans, M. Van Montagu, and J. Schell. 1983. Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. *EMBO J.* 2:2143-2150.