The EMBO Journal vol.14 no.14 pp.3585–3595, 1995

The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome

Bruno Tinland[1], Fabrice Schoumacher[2], Véronique Gloeckler, Ana María Bravo-Angel and Barbara Hohn

Friedrich Miescher-Institut, PO Box 2543, CH-4002 Basel and [2]Zentrum für Lehre und Forschung, Hebelstrasse 20, CH-4031 Basel, Switzerland

[1]Corresponding author

The VirD2 protein of *Agrobacterium tumefaciens* was shown to pilot T-DNA during its transfer to the plant cell nucleus. We analyze here its participation in the integration of T-DNA by using a *vir*D2 mutant. This mutation reduces the efficiency of T-DNA transfer, but the efficiency of integration of T-DNA *per se* is unaffected. Southern and sequence analyses of integration events obtained with the mutated VirD2 protein revealed an aberrant pattern of integration. These results indicate that the wild-type VirD2 protein participates in ligation of the 5′-end of the T-strand to plant DNA and that this ligation step is not rate limiting for T-DNA integration.

Key words: *Agrobacterium*/plant/recombination/T-DNA/virulence protein

Introduction

Among all bacterial conjugation systems, the one evolved by *Agrobacterium* to stably transform plant cells is especially remarkable because proteins encoded by prokaryotic genes have to achieve functions specific to eukaryotic cells. In *Agrobacterium* the T-DNA (transferred DNA) is flanked by 25 bp repeats, termed borders, and is part of a plasmid called the tumor-inducing (Ti) plasmid. During the process of plant cell transformation by *Agrobacterium*, a single-stranded form of T-DNA, the T-strand, is transferred to the plant cell nucleus (Stachel *et al.*, 1986; Tinland *et al.*, 1994; Yusibov *et al.*, 1994). The production and transfer of the T-strand are mediated by bacterial virulence proteins (Vir proteins; for reviews see Ream, 1989; Zambryski, 1992).

VirD2 protein has an endonuclease activity which, in association with VirD1 protein, cleaves the lower strand of the right border of the T-DNA and, simultaneously, attaches covalently to the 5′-end of the T-strand via Tyr 29 (see Figure 1; Ward and Barnes, 1988; Young and Nester, 1988; Dürrenberger *et al.*, 1989; Howard *et al.*, 1989; Vogel and Das, 1992; Pansegrau *et al.*, 1993). A second cleavage at the left border leads to the liberation of a T-strand–VirD2 complex from the parental Ti plasmid. In addition, the T-strand–VirD2 complex may be coated by the single-stranded DNA binding protein VirE2, leading to the transferable structure defined as the T-complex (Howard and Citovsky, 1990).

The T-strand processing events that take place in *Agrobacterium* resemble those occuring in bacterial conjugation of plasmids belonging to the IncP family (Waters and Guiney, 1993; Lessl and Lanka, 1994). A site-specific cleavage/attachment reaction has been measured *in vitro* using purified VirD2 protein and single-stranded oligonucleotides or small DNA fragments containing the right border sequence (Pansegrau *et al.*, 1993; Jasper *et al.*, 1994). However, purified VirD2 alone does not cleave double-stranded DNA *in vivo* in the bacteria; association of VirD2 with other virulence proteins (i.e. VirD1) may relax or unwind the T-DNA borders, allowing site-specific cleavage by VirD2 (Howard *et al.*, 1989; Filichkin and Gelvin, 1993; Pansegrau *et al.*, 1993; Scheiffele *et al.*, 1995).

Once in the plant cell, the T-DNA–protein complex presumably interacts with eukaryotic components as it enters the nucleus and integrates into the plant cell genome. VirD2 protein has been shown to contain a nuclear localization signal (NLS) responsible for efficient transfer of the bacterial DNA to the plant cell nucleus (Herrera-Estrella *et al.*, 1990; Howard *et al.*, 1992; Shurvinton *et al.*, 1992; Tinland *et al.*, 1992; Rossi *et al.*, 1993b). So far, the bacterial function involved in the integration of T-DNA has not been identified. Nevertheless, it is because of its efficient and precise integration properties that *Agrobacterium*-mediated T-DNA transfer is used so widely in plant transformation (Gasser and Fraley, 1989; Hooykaas and Schilperoort, 1992).

Sequence analysis of several T-DNA insertions and of their respective pre-insertion sites (insertional target sites) in the plant DNA revealed some remarkable features (Matsumoto *et al.*, 1990; Gheysen *et al.*, 1991; Mayerhofer *et al.*, 1991): (i) T-DNA integration is not site-specific and occurs by some type of illegitimate recombination; (ii) T-DNA integration leads to small deletions of plant DNA at the insertion sites (13–73 nt); (iii) the 3′-end of the integrated T-strand is poorly conserved, i.e. in the vast majority of the cases studied it is truncated by from 3 to 100 nt and, in addition, one can find homologies with the pre-insertion site of at least five mostly contiguous nucleotides; (iv) the 5′-end of the integrated T-strand, on the other hand, is very well conserved and shows less homology to the pre-insertion site (for reviews see Tinland and Hohn, 1993; Koncz *et al.*, 1994). This latter observation suggests that the two T-DNA ends are joined to plant DNA by different mechanisms. Here, therefore, we directly address the question whether VirD2 protein, which is attached to the 5′-end of the T-strand, participates in ligation of the 5′-terminus of the T-strand with a 3′-end in the plant DNA. We chose to mutate an amino acid common to motifs which have been found in the sequence of VirD2 (Figure 1): domain III of the TraI protein of the conjugative plasmid RP4 (Lessl and Lanka, 1994;

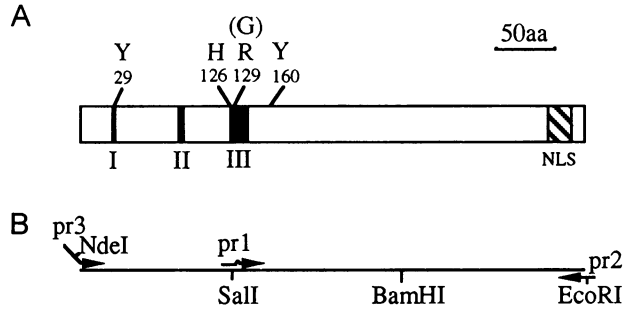


Fig. 1. (**A**) VirD2 domains. NLS, nuclear localization signal. I, II and III are the three TraI motifs. Amino acids H126, R129 and Y160 constitute the putative site-specific integrase triad. (G) stands for the R129G mutation. Upon T-DNA processing, VirD2 attaches to the 5′-end of the T-strand at its amino acid Y29. (**B**) Position of the primers used for constructions (see Materials and methods).

Pansegrau *et al.*, 1994) and the motif H-(2)-R-(32+/–2)-Y found in recombinases (Argos *et al.*, 1986).

Results

Effect of an R→G mutation of VirD2 on T-DNA processing

To disrupt a function of VirD2 which might be responsible for integration of the T-DNA, we mutated the Arg129 (R129) residue to glycine (G) (see Materials and methods; Figure 1). An *Agrobacterium* strain carrying this mutation exhibited a strongly reduced T-DNA transfer efficiency (see below). Since this could, in principle, be due to an impairment of the processing events in the bacterium and/or to the efficiency of the transfer *per se*, we first analyzed different activities of the mutant VirD2 protein within the bacterial cell.

VirD2R129G processes the T-DNA in Agrobacterium inefficiently. The plasmid pAVD63 carrying the gene *vir*D2R129G was introduced into the *vir*D2-defective strain ATvirD2$^-$, resulting in ATvirD2R129G (see Materials and methods). Plasmid pTd33 carrying a T-DNA with a *uidA* and an *nptII* gene (see Figure 2; Materials and methods) was introduced into the *Agrobacterium* strains containing Ti plasmids with either the mutated *vir*D2R129G, the wild-type *vir*D2 gene or the deleted *vir*D2 gene. The resulting strains, ATvirD2R129G(pTd33), ATvirD2(pTd33) and ATvirD2$^-$(pTd33), were tested for their ability to produce T-strands from pTd33. DNA from *Agrobacterium* cells induced by acetosyringone (acetosyringone is a plant phenolic compound shown to be an inducer of the *vir* genes of the Ti plasmid; for a review see Ream, 1989) was analyzed for its content of processed T-DNAs by hybridizing a blot performed under non-denaturing conditions with a radioactively labeled 2.2 kb *Bam*HI fragment containing the *uidA* structural gene (see Figure 2; Materials and methods).

A fragment corresponding to a full-length, correctly processed T-DNA was found for the strain ATvirD2-(pTd33), whereas no signal was found for the mutant ATvirD2R129G(pTd33) nor, as expected, for ATvirD2$^-$ (pTd33) (data not shown). However, a VirD2 mutation resulting in inefficient processing of the T-DNA would render the simultaneous nicking of both borders improbable; this could explain why we did not detect any free T-strands by Southern DNA hybridization analysis. Right border processing was therefore revealed by creating a precise second nick in the T-DNA thereby generating a smaller T-strand. This was done by digestion with *Hin*dIII, which has been shown to cleave single-stranded T-DNA (Stachel *et al.*, 1986). Following this digestion, the signal corresponding to the full-length T-DNA disappeared for the strain containing wild-type *vir*D2 and gave rise to a smaller one corresponding to the expected *Hin*dIII site right border fragment. However, this fragment was still not detectable for the strain containing *vir*D2R129G. The sensitivity of the T-strand assay was tested by loading different quantities of the *Hin*dIII-digested DNA of induced wild-type bacteria. This test proved that we could have detected T-strand formation corresponding to 1% of that produced by ATvirD2(pTd33) (see Figure 3A). Thus T-strand formation in ATvirD2R129G(pTd33) occurs at an efficiency of $<1\%$ of that of the wild-type *vir*D2-containing strain. Western blot analysis showed that this low efficiency of T-strand processing could not be due to a reduced amount of VirD2R129-G, since similar amounts of VirD2 and VirD2R129G were found in the respective strains (see Materials and methods; data not shown).

Two possible reasons for the absence of detectable nicking activity in *Agrobacterium* were envisioned: (i) due to a conformational defect, VirD2R129G does not easily have access to the single-stranded border motif which is the substrate for the cleavage reaction; (ii) the site-specific cleaving/attachment reaction of VirD2 on the single-stranded border is affected. Accordingly, we tested VirD2R129G for enzymatic cleavage and joining activity *in vitro*.

VirD2R129G cleaves the right border efficiently and correctly in vitro. The proteins VirD2 and VirD2R129G were produced in *Escherichia coli* and tested *in vitro* for cleavage of an oligonucleotide containing the border sequence (see Materials and methods). The cleavage products obtained after incubation of VirD2 or VirD2-R129G with a 35 base oligonucleotide containing the border sequence, ^{32}P-labeled at its 5′-end, were separated in a polyacrylamide gel. A product with the electrophoretic mobility of a 27mer, resulting from correct cleavage of the oligonucleotide, was found in similar amounts for both the VirD2 and VirD2R129G proteins used at a concentration limiting for the reaction (see Figure 3Ba). Thus the absence of detectable T-strands in induced cells of ATvirD2R129G(pTd33) is not due to a decrease in cleavage activity on the single-stranded border sequence.

The VirD2R129G protein is tightly attached to the 5′-end of the cleaved oligonucleotide. The mutated and wild-type VirD2 proteins were incubated with a 35mer oligonucleotide containing the 25 bp border sequence described above, but in this case labeled at its 3′-end. The proteins, separated by 10% SDS–PAGE, retained the radioactive label at similar levels, indicating covalent attachment of the cleaved oligonucleotide to VirD2 as well as to VirD2-R129G (Figure 3Bb).

The VirD2R129G protein performs the hemi-site ligation of oligodeoxyribonucleotides similarly to the wild-type VirD2 protein. The VirD2 protein was shown to perform the ligation of an attached oligonucleotide moiety to a

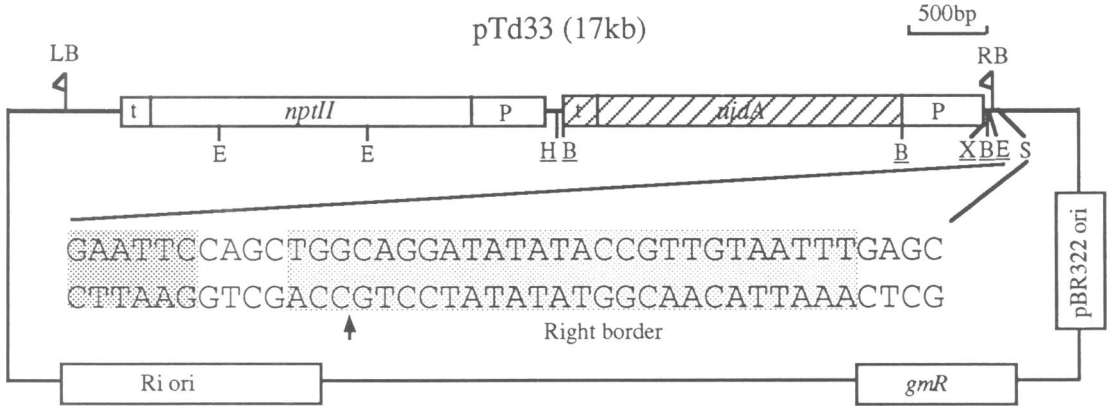


Fig. 2. Binary plasmid pTd33. Only the T-DNA is drawn to scale. The sequence replacing the *Eco*RI–*Sac*I fragment of pCGN1559 is given with the right T-DNA border indicated by the light grey area and the arrow pointing to the nick site. The *Eco*RI site used to analyze the T-DNA integration is shadowed dark grey. B, H, E, X and S are restriction sites for *Bam*HI, *Hin*dIII, *Eco*RI, *Xba*I and *Sac*I respectively. The sites important for the study of the integration process are underlined. There are two other *Eco*RI sites outside the T-DNA which are not indicated. P and T are the 35S promoter from cauliflower mosaic virus and the terminator of the nopaline synthase gene respectively. *nptII* codes for neomycin phosphotransferase (kanamycin resistance) and *uidA* for β-glucuronidase. The hatched area corresponds to the region used to probe the T-DNA. pBR322 ori and Ri ori stand for origin of replication in *E.coli* and *Agrobacterium* respectively. The gentamycin resistance gene is indicated by *gmR*.

hemi-site corresponding to a preformed 3′-OH-terminus of the cleavage site (Pansegrau *et al.*, 1993; Jasper *et al.*, 1994; Figure 3Bc). Since this ligation may be analogous to a ligation event occuring during T-DNA integration into the plant genome (although this latter ligation is not site-specific), we compared the ability of VirD2R129-G and VirD2 to perform this reaction. This test was performed using a 35mer oligonucleotide containing the entire border sequence and a 17mer, radiolabeled at the 5′-end, corresponding to the 3′-end of a cleaved border (see Materials and methods). The reaction involved processing of the 35mer oligonucleotide, resulting in attachment of VirD2 or VirD2R129G to an 8mer, and ligation of this attached fragment to the 3′-end of the 17mer. The ligation activity was monitored by the appearance of the 25mer ligation product. Both proteins performed the reaction with similar efficiency (Figure 3Bc).

Effect of an R→G mutation of VirD2 on T-DNA transfer to the plant cell nucleus and integration into the genome

The R129G mutation of VirD2 does not affect the site-specific cleavage/attachment activity on single-stranded DNA *in vitro*. Thus the step limiting the processing in the bacteria may be inefficient melting of the border DNA. Although T-strand formation is reduced to below the detection limit, a small amount of transferable T-strand molecules must nevertheless be produced in a strain carrying this mutation, because T-strands are detected in the plant cells (the assay is more sensitive; see below). We next compared the transfer and integration properties of these T-strands with those produced by wild-type VirD2. T-DNA transfer and integration was tested after co-cultivation of 1–2-week-old tobacco seedlings with the different *Agrobacterium* strains containing the plasmid pTd33. The presence of *uidA* and *nptII* genes on the T-DNA of pTd33 allowed us to measure transient expression and stable integration of T-DNA-encoded genes within the same batch of inoculated seedlings (see Materials and methods).

The R129G mutation drastically affects the transfer of the T-DNA to the plant cell nucleus. We define the relative transfer efficiency as the number of T-DNA molecules transferred to the plant cell nucleus by the bacterial strain to be tested compared with this number for a wild-type bacterium. This was measured by the activity of β-glucuronidase (number of blue spots/tested seedling) resulting from the transient expression of the transferred *uidA* gene (see Materials and methods; Jansen and Gardner, 1989; Machida *et al.*, 1993, Chaudhury *et al.*, 1994). For the VirD2-defective strain ATvirD2$^-$(pTD33) the transfer efficiency was zero, whereas for ATvirD2R129G(pTD33) it corresponded to that of a dilution of ATvirD2(pTd33) to 0.4–1% (see Table I). These experiments demonstrate a low but measurable transfer efficiency of the strain containing *vir*D2R129G which correlates with the observed reduction to an undetectable level (i.e <1%) of T-strand production.

The T-DNA is co-transferred with VirD2R129G to the plant cell nucleus. The VirD2 protein of ATvirD2 was shown previously to contain an NLS necessary for efficient transfer of the T-DNA to the plant cell nucleus (Rossi *et al.*, 1993b; see Introduction). To check whether VirD2R129G pilots the T-DNA to the plant cell nucleus, the strain ATvirD2R129GNLS$^-$(pTd33), containing the *vir*D2R129G gene lacking its NLS sequence, was compared for its transfer efficiency with dilutions of the *vir*D2NLS$^-$ strain ATvirD2NLS$^-$(pTD33) [T-DNA is processed at wild-type levels in ATvirD2NLS$^-$(pTd33); A.M.Bravo Angel, unpublished results]. Transfer efficiency of both NLS$^-$ strains, again measured as transient expression of β-glucuronidase, dropped to ~5% of that of the NLS-containing strains. This indicates that inefficient T-DNA transfer from a strain containing the mutated VirD2 still is dependent on the NLS of VirD2 and that VirD2R129G therefore accompanies the T-DNA during its transfer to the plant cell nucleus. This result is consistent with the finding that VirD2R129G, just as its wild-type relative, remains attached to the 5′-terminus of the 3′-end-labeled oligonucleotide upon *in vitro* cleavage (see above).

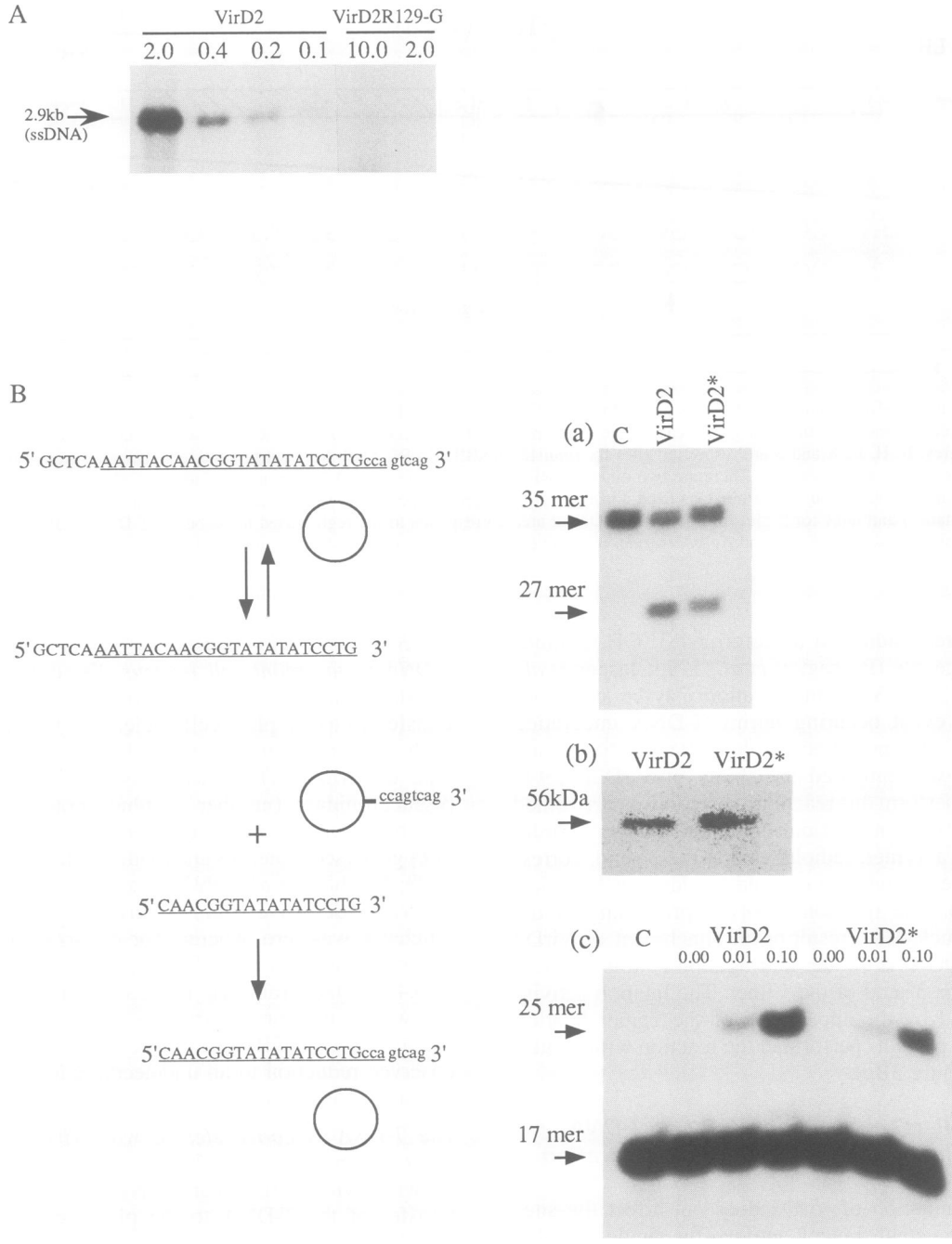


Fig. 3. Analysis of T-DNA processing mediated by VirD2R129G. (**A**) T-Strand production in *Agrobacterium*. DNA was extracted from induced ATvirD2(pTd33) and ATvirD2R129G(pTd33), treated with the restriction enzyme *Hin*dIII and compared by Southern analysis after non-denaturing transfer (see text). The quantities of DNA loaded (μg) are indicated on top of each lane. A 2.9 kb single-stranded DNA fragment containing the *uidA* gene was used as a size marker (not shown). No signal is detected in 10 and 2 μg DNA from ATvirD2R129G(pTd33), while a very weak signal can still be detected on the original autoradiograph in 0.1 μg DNA from ATvirD2(pTd33). (**B**) *In vitro* test of VirD2R129G activities. The oligonucleotides are given in the 5′→3′ orientation. VirD2* stands for VirD2R129G. Lanes C correspond to control reactions performed in the absence of proteins. (**a**) Site-specific cleavage. Either VirD2 or VirD2R129G protein (10 pmol) were incubated with 1 pmol 35mer oligonucleotide containing the octopine border sequence and radiolabeled at the 5′-end. The site-specific cleavage is monitored by the appearence of a 27mer oligonucleotide on a 20% polyacrylamide gel. (**b**) Attachment. The reaction is performed as for (a) except that the 35mer oligonucleotide is radiolabeled at the 3′-end. VirD2 and VirD2R129G retain the labeled oligonucleotide to a similar extent after migration on SDS–PAGE gels. (**c**) Site-specific joining. The reaction was performed with either 10 pmol VirD2 or VirD2R129-G, 1 pmol 17mer oligonucleotide radiolabeled at the 5′-end, containing the non-transferred part of the border, and either 0.00, 0.01 or 0.1 pmol unlabeled 35mer oligonucleotide. The reaction is monitored by the signal corresponding to the original (17mer) and chimeric ligation products (25mer) after separation on a 20% polyacrylamide gel. In (a) and (c) the signals were revealed using an autoradiographic film. In (b) a PhosphorImager device was chosen to reveal the 3′-OH labeling.

The R129G mutation does not affect the integration efficiency. By exploiting the *nptII* gene located adjacent to the *uidA* gene on the T-DNA of pTd33 (Figure 2), we estimated the efficiency of T-DNA integration into the plant chromosomal DNA after selection for kanamycin-resistant calli. The relative transformation efficiency of a

Table I. *Agrobacterium* strains

Name	Relevant characteristics	Reference
ATvirD2$^-$	GV3101(pPM6000K). pPM6000K is a disarmed Ti plasmid with a deletion inactivating the *vir*D2 gene; rifampicin resistant	Rossi *et al.*, 1993a
ATvirD2	GV3101(pPM6000K::pAVD43). Plasmid pAVD63 carrying a functional *vir*D2 gene is co-integrated into pPM6000K; rifampicin and kanamycin resistant	Rossi *et al.*, 1993a
ATvirD2NLS$^-$	GV3101(pPM6000K::pAVD44). The plasmid pAVD44 carrying *vir*D2 deleted in the sequence encoding the C-terminal nuclear localization signal (NLS) is co-integrated into pPM6000K; rifampicin and anamycin resistant	Rossi *et al.*, 1993a
ATvirD2R129G	GV3101(pPM6000K::pAVD63). Plasmid pAVD63 carrying a *vir*D2 gene mutated in the traI/ integrase region, R129G, is co-integrated into pPM6000K; rifampicin and kanamycin resistant	This work
ATvirD2R129GNLS$^-$	GV3101(pPM6000K::pAVD66). Plasmid pAVD66 carrying the *vir*D2 R129G mutation combined with deletion of the sequence encoding the C-terminal NLS is co-integrated into pPM6000K; rifampicin and kanamycin resistant.	This work

Table II. Efficiency of transfer, transformation and integration mediated by VirD2 and VirD2R129-G

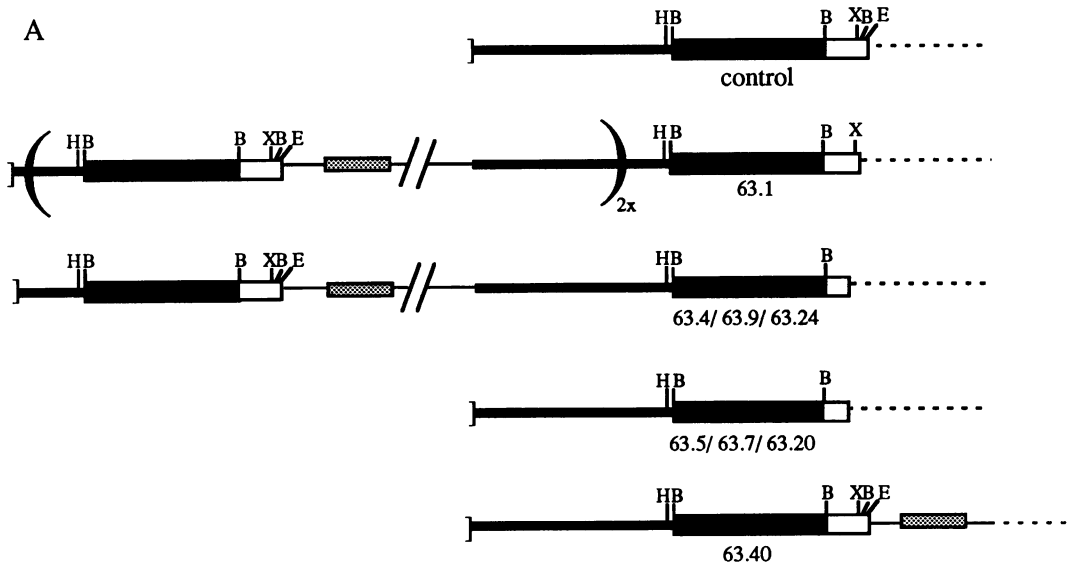
	Transfer efficiency (spots/seedling)			Transformation efficiency (calli/seedling)			Integration efficiency (transformation/transfer)		
	1	2	3	1	2	3	1	2	3
1/100	5.5	9.0	5.9	0.6	0.37	0.4	0.108	0.041	0.068
1/250	3.0	5.0	3.4	0.39	0.13	0.21	0.130	0.026	0.062
1/500	1.6	3.0	1.8	0.28	0.09	0.15	0.175	0.030	0.081
R129-G	3.7	6.0	3.9	0.40	0.08	0.34	0.108	0.013	0.087

1, 2 and 3 stand for three independent experiments. 1/100, 1/250 and 1/500 correspond to the ratio ATvirD2(pTd33)/ATvirD2$^-$(pTd33) used to inoculate the seedlings. R129-G corresponds to the strain ATvirD2R129-G(pTd33) used undiluted.

strain corresponds to the number of kanamycin-resistant calli/seedling compared with this number for a wild-type bacterium. We thus define the integration efficiency as the ratio of transformation efficiency to transfer efficiency (as defined above). This value represents the fraction of T-DNA molecules which integrate out of those that entered the nucleus. Experimentally we define the integration efficiency as the number of kanamycin-resistant calli divided by the number of blue spots. Whereas the transformation efficiency of ATvirD2 was dependent on the dilution used, the integration efficiency was constant over the range examined, i.e. 0.14 ± 0.03 (mean ± SD), 0.032 ± 0.006 and 0.07 ± 0.008 for experiments 1, 2 and 3 respectively (dilutions from 1:500 to 1:100; see Table II). Although the relative transfer and transformation efficiencies for the mutant were of the order of 0.4–1% of those of the wild-type, the integration efficiencies were very similar for both strains (see Table II). The efficiency of integration *per se* is therefore not affected significantly by the R129G mutation.

The R129G mutation modifies the precision of ligation of the T-DNA right border to the plant DNA. In many cases the 5′-end of the T-strand has been shown to be precisely conserved upon integration into the plant genome (see Introduction). We wanted to test whether the R129G mutation influences the precision of this ligation by analyzing T-DNA–plant DNA junctions of the F1 progeny of tobacco plants transformed with ATVirD2(pTd33) and ATVirD2R129G(pTd33). For this purpose, the plasmid pTd33 was designed with an *Eco*RI restriction site 7 nt to the left of the nick site of the right border (Figure 2). This site allowed us to determine whether the 5′-end of the integrated T-DNA was still present or had been lost during integration.

Southern analysis, using the enzymes *Hin*dIII (which cuts once in the whole plasmid), *Bam*HI and *Xba*I, alone and in combination with *Eco*RI, and with the *Bam*HI fragment spanning the *uidA* gene as probe (Figure 2) revealed the structures at the right T-DNA end–plant DNA junctions, schematically represented in Figure 4A. The following conclusions can be drawn. (i) In six out of eight transformants obtained with a wild-type *Agrobacterium* strain the right T-DNA end was conserved (Figure 4A; control). (ii) Four out of eight transformants obtained with the strain mutated in VirD2 contained multimers of the complete plasmid (plants 63.1, 63.4, 63.9 and 63.24). The presence of plasmid multimers in the plant cell genome can be explained by inefficient processing of the T-DNA by the mutant VirD2 protein. This would result from rare cleavage at the right border, followed by liberation of single-stranded DNA concomitant with its continuous replacement. This multimeric plasmid DNA is transferred to and integrated into the plant cell genome, suggesting that the T-strand is produced by a mechanism related to rolling circle replication. This has in fact been proposed earlier (Miranda *et al.*, 1992; Waters and Guiney, 1993; Wilkins and Lanka, 1993). (iii) Most importantly, seven out of eight transformants obtained with the mutant strain did not contain the *Eco*RI site diagnostic for a precise right T-DNA–plant DNA junction. Only in one case (plant 63.40) has the site been detected, but T-DNA was not found joined to plant DNA but rather was found inserted in conjunction with a truncated plasmid. We conclude that in none of the analyzed cases were T-DNA–plant DNA junctions at the right T-DNA border. However in most cases, junctions occurred very close to the right border: in six out of eight cases the junction of the right end of the T-DNA to plant DNA occurred at a T-DNA position



B

```
          .         .         .         .  -376  .         .         .         .
ATCTTTGGGACCACTGTCGGCAGAGGCATCTTCAACGATGGCCTTTCCTTTATCGCAATGATGGCATTTGTAGGAGCCAC
TAGAAACCCTGGTGACAGCCGTCTCCGTAGAAGTTGCTACCGGAAAGGAAATAGCGTTACTACCGTAAACATCCTCGGTG
                                  ||||||||||||
                                  aacgatggccttggttcatgaagagaa   63.4

          .         .         .         .         .         .         .         .
CTTCCTTTTCCACTATCTTCACAATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGGATATTACCCTT
GAAGGAAAAGGTGATAGAAGTGTTATTTCACTGTCTATCGACCCGTTACCTTAGGCTCCTCCAAAGGCCTATAATGGGAA

          .         .         .      -220       .         .         .         .
TGTTGAAAAGTCTCAATTGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTTTTGGAGTAGACAAGCGTGTCGTGCTC
ACAACTTTTCAGAGTTAACGGGAAACCAGAAGACTCTGACATAGAAACTATAAAAACCTCATCTGTTCGCACAGCACGAG
                            ||||||||||||
                            tcttctgagacttccttgaccctccga   63.7

        -168      .         .         .         .         .         .         .
CACCATGTTGACGAAGATTTTCTTCTTGTCATTGAGTCGTAAGAGACTCTGTATGAACTGTTCGCCAGTCTTTACGGCGA
GTGGTACAACTGCTTCTAAAAGAAGAACAGTAACTCAGCATTCTCTGAGACATACTTGACAAGCGGTCAGAAATGCCGCT
||||||||||||
caccatgttgatataaaattaccaa   63.5

          .         .         .         .  -54   .         .         .         .
GTTCTGTTAGGTCCTCTATTTGAATCTTTGACTCCATGGACGGATCCGGGGAATTGACTCTAGAGGATCCCCGGGTACCG
CAAGACAATCCAGGAGATAAACTTAGAAACTGAGGTACCTGCCTAGGCCCCTTAACTGAGATCTCCTAGGGGCCCATGGC
                                  ||||||||||||
                                  ccatggacggatgacatgtattttgaa   63.20

          .      -1 .         .
AGCTCGAATTCCAGCTGGCAGGATATATACCGTTGTAATT
TCGAGCTTAAGGTCGACCGTCCTATATATGGCAACATTAA
      ||||||||||||
      aattccagctggcacaagaaaaacgaa   33.α3

      aattccagctggacgagaaagtgaatg   33.50

      aattccagctggcacaaaaattcacaa   33.6
```

Fig. 4. (**A**) Schematic representation of the types of insertion found. 'Control' represents a typical transformant obtained after transformation with the wild-type VirD2-containing strain (6/8 of the analyzed cases). The other patterns correspond to the insertions obtained with VirD2R129G. Only restriction sites present on the T-DNA are indicated. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; X, *Xba*I. The left junctions of the T-DNA with the plant DNA have not been investigated and are indicated by the brackets. The parentheses delimits a monomer of pTd33. Dashed line, plant DNA; white rectangle, promoter region of *uidA* gene; black rectangle, coding region of the *uid*A gene (probe used for the Southern analysis); thick line, left part of the T-DNA; thin line, vector part of pTd33; grey rectangle, gentamycin gene. (**B**) Sequence analyses of insertion sites obtained with ATVirD2R129G(pTd33) (63.4, 63.5, 63.7, 63.20) and with ATVirD2(pTd33) (33.α3, 33.50, 33.6). The 410 nt near the right border and inside the T-DNA of pTd33 are indicated by upper case letters. The 24 bp of the right border are in bold. –1 indicates the location of the first nucleotide proposed to be transferred. The sequences of junctions between the T-DNA and the plant DNA are given in lower case letters, plant DNA sequences are underlined. Matches between pTd33 and the junction sequence are highlighted by vertical bars.

located between 34 and 600 nt from the nicking site. In one case the junction was located between 7 and 34 bases (63.1) and in one case inside the vector (63.40).

Sequence analysis of T-DNA insertions. The finding that seven out of the eight analyzed recombination events occurred within a small area to the left of the right border can be explained by two hypotheses: either the mutated VirD2 protein attaches to the T-strand at the correct position, as has been shown *in vitro*, or there are preferred sequences (pseudo-borders) which are recognized more efficiently by the mutated VirD2 protein. To study the latter possibility we sequenced across four T-DNA–plant DNA junctions obtained with the mutated VirD2R129G protein and three junctions obtained with the wild-type VirD2 protein. Inverse polymerase chain reaction (IPCR) technology was used for isolation of junction segments (see Materials and methods). T-DNA breakpoints at the expected site were found in the case of the wild-type VirD2, whereas for the mutant protein alignment with T-DNA internal sequences revealed that all breakpoints were different. Furthermore, we did not detect any specific feature in the T-DNA sequence to the right of the junction which could have served as an alternative border or pseudo-border (Figure 4B). This indicates that VirD2R129G does not have affinity for some kind of degenerated border or for a different DNA sequence. Thus the sequence cleaved by the mutated protein inside the bacterium is most probably the *bona fide* T-DNA border and the observed truncations must have therefore occurred upon integration.

Discussion

As opposed to other integrative mobile DNA elements, such as transposons or retroviruses, T-DNA does not encode proteins enabling it to integrate. Thus integration of T-DNA could be carried out either by the machinery of the recipient plant cell or by proteins co-transferred with it from the donor bacteria, possibly in combination with plant proteins.

We investigated whether the bacterial VirD2 protein, which has been shown to be attached to the 5′-end of the T-strand, participates in the integration process (see Introduction). To analyze the mechanism of T-DNA integration more directly, we studied the consequence of a mutation in VirD2 on processing and integration of T-DNA (see Introduction; Figure 1).

A series of experiments was performed to establish the effect of this mutation, R129G, on activities of VirD2 known to be performed in the bacteria. We found that the R129G mutation does not affect the efficiency of site-specific cleavage by VirD2 *in vitro* nor the precision of joining an oligonucleotide–VirD2 complex to its cognate 3′-end. Further, the amino acid change did not appear to alter the stability of the VirD2–T-strand complex; the VirD2R129G protein was found covalently attached to the processed oligonucleotide after *in vitro* cleavage. However, the efficiency of T-DNA transfer from an *Agrobacterium* strain containing the mutated VirD2 protein was reduced to ~1% of that of a wild-type VirD2-containing strain. This is consistent with the finding that T-strand formation in the bacteria was reduced to <1% and suggests that some aspects of the processing (possibly melting of the T-DNA borders) is the limiting step in transfer mediated by VirD2R129G. Deletion of the NLS sequences of the mutated VirD2 protein caused a further drastic drop in its T-DNA transfer efficiency. This indicates that the small number of T-strands transferred to the plant nucleus by the mutant VirD2 is nevertheless dependent on the nuclear localization signal carried by VirD2R129G. We conclude that VirD2R129G remains attached to the T-strand at least until it reaches the nucleus.

These conclusions are essential to the interpretation of our two most important findings concerning the integration itself: (i) the R129G mutation does not significantly affect the efficiency of integration into the plant cell genome; (ii) the R129G mutation leads to a loss of precision of integration of the 5′-end of the T-strand. We strongly favor the interpretation that it is the integration event during which truncation occurs and that cleavage of the border sequence by the mutated VirD2 protein occurs precisely, but with low efficiency, inside the bacterial cell. This is strengthened by several arguments: (i) VirD2R129G cleaves correctly and efficiently *in vitro* and it therefore has the intrinsic capacity for sequence-specific cleavage; (ii) the sequence truncations at the right end of the T-DNA were found to differ from each other and from the border sequence, which excludes preferential cleavage of T-DNA by the mutated protein at a new site or at pseudo-borders; (iii) in transformants obtained using the agrobacterial strain containing the mutated protein, all but one junction were mapped inside the T-DNA, close to the right border; if randomly cleaved T-strands entered the nucleus, transformants should have been recovered with T-DNA junctions also to the right of the right border.

Model for T-DNA integration

Because the mutant VirD2 protein is transferred with the T-strand to the plant cell nucleus, the loss of nucleotides at the 5′-end of the T-strand presumably occurs upon integration, concomitant with the release of the mutant VirD2 protein. The release of VirD2 (either wild-type or mutant) is required to provide the 5′-end of the T-strand for ligation to the plant DNA. The simplest mechanism to explain release of the wild-type VirD2 protein and conservation of the attached T-strand nucleotide is the use of energy contained in the phosphodiester bond joining Tyr29 of VirD2 and the first nucleotide of the T-strand for the establishment of the new phosphodiester bond of the recombination product. This phosphotyrosine bond is an electrophilic center which can serve as substrate for a nucleophilic 3′-OH provided by nicked plant DNA (see below; Jayaram, 1994). If the high energy of the linkage of VirD2 to the 5′-end of the T-strand is indeed the driving force for precise integration, we can propose that the R129G mutation affects the interaction between the phosphotyrosine bond and the plant cell DNA. As a consequence, the recombination process is mediated by the plant DNA repair system using as substrate a phosphodiester bond between two nucleotides of the T-strand.

Upon detailed analysis of the published right border junctions and their respective pre-insertion sites (Gheysen *et al.*, 1991; Mayerhofer *et al.*, 1991), we noticed that in each case in which VirD2 could be invoked in the integration process, homology existed between at least the nucleotide linked to VirD2 and the last nucleotide of the

```
3'cacaaact 5' nopaline

5'ATGTTT   3' 621-36
  ACACGTGA    621-37
  AGCCCATA    N6H-cs4
  AACGGTGA    N6H-14
  CCTTTTAA    621-2
  AAAAGT      B1

3'gactgacc 5' octopine

5'TGTCATGG 3' B2
  CAAAGGG    B6
```

Fig. 5. Sequence homologies between the 5′-end of the T-strand and the right end of the plant sequences to be deleted upon T-strand integration. The 3 nt derived from the 25 bp repeats are underlined. Octopine and nopaline indicate the origin of the T-DNA. The plant DNA sequences are referred to by the name of the analyzed transformed plant (see Gheysen *et al.*, 1991; Mayerhofer *et al.*, 1991). The exact sites of T-DNA insertion in plants 621-36 and 621-37 have been re-interpreted. Bold characters indicate homologies found between plant DNA pre-insertion sites and T-DNA. Large and small arrows indicate major and minor processing sites respectively (Wang *et al.*, 1987).

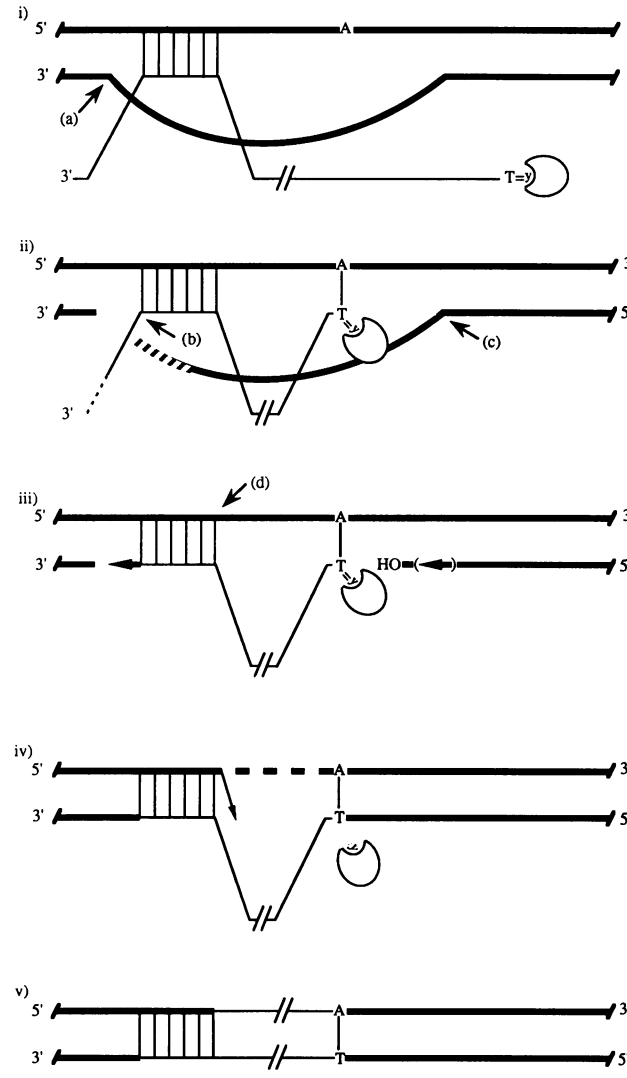


Fig. 6. Model for T-DNA integration (see text). The plant DNA and the T-strand are represented by thick and thin lines respectively. Only base pairing and nucleotides important for integration are indicated (we chose the example of a T-strand processed at the major nick site of a nopaline strain). (a), (b), (c) and (d) indicate the possible sites for endonucleolytic attack. Oval motif, VirD2; =, phosphotyrosine bond. Arrows indicate DNA replication and dashed lines represent DNA strands subjected to nucleolytic degradation. Plant proteins which are not shown in this scheme might stabilize the micro-annealing and/or help in the ligation process.

plant DNA deleted during the integration event (Figure 5). Pairing of only a few bases, termed pairing by 'micro-homologies', has already been reported for illegitimate recombination (Lehman *et al.*, 1994). This observation indicates that pre-annealing of the nucleotide(s) of the T-strand attached to VirD2 may be a prerequisite for integration of an intact 5′ T-strand end. These micro-homologies may provide a minimum 'site specificity' for the recombination reaction by positioning VirD2 for the ligation. Since only one base is sufficient, the homology requirement is very low and should therefore not be rate limiting for integration. This observation, as well as the fact that mutation of VirD2 does not reduce the efficiency of the integration step itself, indicates that recombination of the 5′-end of the T-strand is not a limiting step with respect to the efficiency of integration. Therefore, we propose that the 3′-end of the T-strand realizes the first synapsis with the plant DNA.

These new results and analyses, together with the known data about integration of the 3′-end of the T-DNA, allow the postulation of a refined model for T-DNA integration (see Figure 6). (i) The 3′-end or adjacent sequences of the T-strand find(s) some homologies to the plant DNA (possibly in a region in which the DNA is already melted; Koncz *et al.*, 1994) and anneals. As a consequence, the displaced plant DNA is subjected to endonucleolytic digestion at (a). (ii) The 3′ overhanging part of the T-strand and the displaced plant DNA strand are digested away, either by endonucleases, at (b) and (c) respectively, or by 3′→5′ exonucleases. The nucleotide(s) directly attached to VirD2 'look(s)' for micro-homologies to anneal (the annealing between micro-homologies is unlikely to be stable so it can be either transient or stabilized by plant cell proteins). (iii) This annealing by micro-homologies brings the electrophilic phosphotyrosine bond in proximity to the nucleophilic 3′-OH resulting from digestion of plant DNA (or upon repair synthesis of the plant DNA). The 3′-end of the annealed T-strand primes the synthesis of plant DNA (see Koncz *et al.*, 1994). The upper strand of the plant DNA is nicked at (d). (iv) Ligation of the 5′-end of the T-strand to the 3′-OH of plant DNA is achieved, the upper strand of the plant DNA is degraded, and the T-strand is copied by the plant repair machinery. (v) The T-DNA insertion is finalized. As part of this model, we speculate that the 5′-end of the T-strand reacts with a 3′-OH of the plant DNA which is the result of integration

of the 3′-end of the T-strand. We consider this 'chain reaction' as a possible explanation for the short deletions of plant DNA created upon T-DNA integration. We chose to present integration of the single-stranded T-DNA derivatives and not of double-stranded versions thereof, because the two extremities reacting with the plant DNA are parts of one and the same T-strand and thus it is not necessary to invoke a double-stranded integrating T-DNA. Furthermore, it has recently been shown that T-DNA derivatives are mainly single-stranded in the plant cell nucleus (Tinland *et al.*, 1994).

However, an interesting question remains: why do the junctions of the right T-DNA end obtained with the mutated VirD2 protein occur in proximity to the 5′-end of the T-strand? Our model would rather have favored a distribution of the recombination events all along the T-strand, because the phosphodiester bonds within a DNA molecule are equivalent. We can offer two explanations: (i) VirD2 has a general affinity for DNA which would preferentially bring the nucleotides closest to the 5′-end in contact with the plant DNA; (ii) VirE2 protein molecules, which were shown to bind single-stranded DNA cooperatively and to play a role in the plant cell (Citovsky *et al.*, 1992), protect the internal part of the T-strand and allow or stimulate the recombination reaction to start only at T-strand ends.

Conclusion

Our results with the R129G mutation allow us to state that the wild-type VirD2 protein has an important role in preserving the 5′-end of the T-strand during integration. This mechanism allows the T-strand to be liberated from VirD2 and to integrate precisely, i.e. without loss of bases from the T-DNA side. This property is essential, because the genes located at the 5′-end of the T-DNA, the opine synthesis genes, are, in the natural system, an important advantage to the bacterium (Guyon *et al.*, 1980; Petit *et al.*, 1983). The major roles of VirD2 in mediating efficient transformation of plant cells are nuclear targeting and precise integration of T-DNA. VirD2 itself, however, does not significantly influence the efficiency of the integration step. This information is to be kept in mind for plant gene targeting experiments with *Agrobacterium* (Lee *et al.*, 1990; Offringa *et al.*, 1990; reviewed in Ohl *et al.*, 1994). It seems rather that the main requirement for an optimal gene targeting substrate is the availability of homologies at the 3′-end of the T-DNA, possibly including the left border sequence.

Materials and methods

Plasmids

Construction of pTd33. The *Eco*RI fragment containing the chimeric *uidA* gene from plasmid pGUS23 (Schultze *et al.*, 1990; Puchta and Hohn, 1991) was cloned, after filling of the *Eco*RI ends, into the *Hin*cII site of pUC18, with the promoter region facing the resident *Eco*RI site. The *Eco*RI–*Hin*dIII fragment containing the *uidA* gene was then subcloned from this plasmid into the corresponding sites present in the polylinker of the T-DNA of the binary plasmid pCGN1559 (McBride and Summerfelt, 1990), which contains a eukaryotic *nptII* gene (neomycin phosphotransferase II gene) and a bacterial gene for gentamycin resistance. The last step was to exchange the *Sac*I–*Eco*RI fragment overlapping the right border of the T-DNA with a *Sac*I–*Eco*RI oligonucleotide in order to introduce the *Eco*RI recognition site at 7 nt from the right border nicking site (the modified sequence is presented in Figure 2). The resulting plasmid pTd33 was introduced into the *Agrobacterium* strains by electroporation.

Construction of virD2R129G. The CGC codon corresponding to Arg129 (R129) of VirD2 was changed to GGC coding for glycine (G). This mutation was introduced by PCR using the *vir*D2 gene of pAVD43 (Rossi *et al.*, 1993b) as template. Two primers were used: one which spans the *Sal*I site of *vir*D2 and carries as the only mismatch the point mutation (primer pr1); the other contains the *Eco*RI site located on pVD43 after the 3′-end of the *vir*D2 gene (primer pr2; see Figure 1B). The amplified fragment containing the mutation was then used to replace the corresponding fragment of pAVD43, using the enzymes *Eco*RI and *Sal*I, resulting in the plasmid pAVD63.

Construction of virD2R129-GNLS⁻. The *Bam*HI–*Eco*RI fragment of the *vir*D2NLS⁻ gene of pAVD44 (Rossi *et al.*, 1993b) was introduced in place of the corresponding fragment of pAVD63, resulting in pAVD66.

pAVD63 and pAVD66 were mobilized by conjugation into the *Agrobacterium* strain AT*vir*D2⁻ and stabilized by co-integration into the resident Ti plasmid (see Rossi *et al.*, 1993b). Structures of the resulting strains were confirmed by Southern blot analysis.

Introduction of virD2 and virD2R129G into the Escherichia coli expression vector pET3a. The coding sequence of *vir*D2 was amplified by PCR, using a primer containing an *Nde*I site overlapping the start codon of *vir*D2 (primer pr3; see Figure 1B) in combination with pr2. The amplified fragment was then cloned into the *Nde*I–*Eco*RI sites of the vector pET3a (Studier *et al.*, 1990), resulting in plasmid pFSvirD2, which contains a *vir*D2 gene under control of the promoter for T7 RNA polymerase. The plasmid pFSvirD2R129G was obtained by replacing the *Sal*I–*Eco*RI fragment of pFSvirD2 with the corresponding fragment of pVD63 (see Figure 1B).

Cloned oligonucleotides and fragments obtained by PCR described in this section were confirmed by sequencing (sequencing kit from USB, Cleveland, OH).

In vitro tests of VirD2

The VirD2/VirD2R129G proteins were overexpressed in *E.coli* BL21(DE3) (Studier *et al.*, 1990) and purified by HPLC. The cells from 1 l of culture were resuspended in 50 ml buffer A [50 mM Tris–HCl, pH 8.5, 150 mM NaCl, 5 mM EDTA, 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 10 mM β-mercaptoethanol, 0.5 mg/ml lysozyme, 0.1% Tween-20]. After 1 h incubation at 4°C, the lysate was centrifuged and the pellet resuspended in the same volume of buffer A without lysozyme and incubated at 4°C for 30 min. After centrifugation, proteins present in the pellet were solubilized by incubation for 1 h in 50 ml buffer B (25 mM NaOAc, pH 5, 8 M urea, 10 mM β-mercaptoethanol, 0.1 mM PMSF). The supernatant obtained after centrifugation was loaded onto an ion exchange column (EconoPac S; BioRad, Hercules, CA) equilibrated with buffer B. The proteins were eluted with a linear NaCl gradient in buffer B. VirD2 peak fractions were dialyzed against 1 l buffer C (50 mM Tris–HCl, pH 8.5, 150 mM NaCl, 5 mM $MgCl_2$, 10 mM β-mercaptoethanol, 0.05% Tween-20, 0.1 mM PMSF) and then applied to a heparin column (Affi-Gel Heparin; BioRad) equilibrated in buffer C. The heparin-purified fractions (200 µl, 100 µg) were loaded on a C18 reverse phase column (Vydac 218TP54). The column was equilibrated with 0.1% trifluoroacetic acid (TFA). Proteins were eluted with a linear acetonitrile gradient in 0.1% TFA with a flow rate of 0.5 ml/min. The peak of material eluting at ~50% acetonitrile was collected and lyophilized. The VirD2 proteins were resuspended in 100 µl buffer C and stored at –80°C in 10% glycerol.

The following oligodeoxyribonucleotides derived from the right border of the octopine-type plasmid pTiA6 were used to test the *in vitro* activities of VirD2 and VirD2R129-G; 35mer, GCTCAAATTACAACGGTATATATCCTGccagtcag; 17mer (hemi-site), CAACGGTATATATCCTG (the lower strand 25 bp border sequence is underlined, lower case letters represent the first nucleotides of the T-strand). 5′-End-labeling was performed using [γ-^{32}P]ATP (Amersham, Little Chalfont, UK) and T4 polynucleotide kinase (Boehringer Mannheim, Mannheim, Germany). The 3′-end was labeled using [α-^{32}P]ddATP (Amersham, Little Chalfont, UK) and terminal deoxynucleotidyltransferase (Boehringer Mannheim, Mannheim, Germany).

In vitro cleavage and hemi-site ligation were performed as described earlier (Pansegrau *et al.*, 1993); oligonucleotides were incubated at 37°C with VirD2 or VirD2R129G in 20 µl reaction buffer (20 mM Tris–HCl, pH 8.5, 5 mM $MgCl_2$, 50 mM NaCl). The reaction products were separated on 20% acrylamide gels containing 8 M urea.

Quantification of radioactive signals
Autoradiography was quantified using a PhosphorImager device and ImageQuant software version 3.3 (Molecular Dynamics, Sunnyvale, CA).

Agrobacterium strains
For *Agrobacterium* strains see Table I.

T-Strand assay
This assay essentially followed the protocol of Stachel *et al.* (1986). Overnight cultures in YEB medium of the various *Agrobacterium* strains were diluted to OD_{600} 0.2 in 6 ml inducing medium (MS; 20 mM MES, pH 5.4, 100 μM acetosyringone) (Murashige and Skoog, 1962). After 8 h growth, 5 ml bacteria were resuspended and incubated for 10 min in 100 μl 50 mM Tris–HCl, pH 8, 20 mM EDTA, 4 mg/ml lysozyme. DNA was then extracted as described (Dhaese *et al.*, 1979). Aliquots of DNA (0.1–10.0 μg/sample, either untreated or digested with *Hin*dIII) were subjected to Southern analysis under non-denaturing conditions.

Western blot analysis
The bacteria from 1 ml of culture grown for the T-strand assay were collected by centrifugation, resuspended in 20 μl sample buffer and subjected to Western blot analysis (Towbin *et al.*, 1979) using a VirD2-specific anti-serum.

Seedling assays
We essentially followed the described protocol (Rossi *et al.*, 1993a). For each experiment, 150 *Nicotiana tabacum* SR1 seedlings were infiltrated with the *Agrobacterium* strain to be tested, incubated for 3 days and then analyzed for transient expression or integration. Transient expression of the *uidA* gene was determined with 50 seedlings by measurement of β-glucuronidase activity. The assay was performed by a histochemical assay using 5-bromo-4-chloro-3-indolyl-glucuronide (X-glu) as described (Rossi *et al.*, 1993a). The incubation time of the seedlings in the X-glu solution was chosen such that a range of 5–10 spots/seedling for the highest concentration of wild-type bacteria (dilution 1/100) was achieved. In all experiments the transfer-defective strain ATVirD2$^-$(pTd33) was used to monitor any T-DNA-independent β-glucuronidase activity and no background activity was found. The remaining seedlings (100) were analyzed for T-DNA integration. The seedlings were washed with 10 mM $MgSO_4$ and placed on MS medium containing 0.1 μg/ml naphtylacetic acid, 1 μg/ml benzylamino purine, 100 μg/ml kanamycin, 500 μg/ml claforan and 500 μg/ml vancomycin. The seedlings were transferred to fresh plates every week. Calli were counted after 4–6 weeks.

Plant transformation
After 7–8 weeks, shoots were collected from independent calli obtained for the integration assay and transferred to the same medium without hormones. After 3–4 weeks the rooted plants were transferred to soil. For analysis, DNA from selfed F1 progeny was extracted as described (Paszkowski *et al.*, 1984) and 10–20 μg of this DNA was analyzed by Southern blots.

Southern blot analysis
Southern blot analysis was performed as described (Sambrook *et al.*, 1989). Radioactive probes were made with a random priming labeling kit (Boehringer, Mannheim, Germany), using [α-^{32}P]dATP (Amersham, Little Chalfont, UK).

Isolation of plant DNA–T-DNA junctions by IPCR
The plant DNA–T-DNA junctions were isolated using IPCR technology (for the principle and methods see Earp *et al.*, 1990). DNA (2 μg) extracted from the transformed plants was digested with the restriction enzyme *Mse*I, which specifically cuts at the sequence TTAA. This sequence is not present in the first 800 nt inside the right border. After digestion, the DNA was religated under diluted conditions (0.5 μg/100 μl) in order to favor intramolecular ligation events. Aliquots (0.5 μg) of the ligated DNA were subjected to PCR using Ampli-Taq polymerase (Perkin-Elmer Cetus). A first PCR was done using as primers pr4, TCCTCTAGAAAGCTTTACAAAGAAGAGGAAGAAGAAACCTC, and pr5, AAATCTAGAAAGCTTGTCCTCTCCAAATGAAATGAAC-TTCC, which are in opposite orientation (the 3′-end of pr4 faces the *uidA* gene and the 3′-end of pr5 faces the right border) and are localized adjacent to each other at ~600 nt from the right border, inside the T-DNA. To obtain a better resolution of the amplified fragment, aliquots of the first PCR were re-amplified using pr1 in combination with primer pr6, AAATCTAGAAAGCTTAGGGTCTTGCGAAGGATAGTGGG (nested primer). pr6 is in the same orientation as pr5 but localized 3′ to it. The re-amplified PCR products were cloned into the *Hin*dIII site of pUC19 using *Hin*dIII sites present at the 5′-termini of pr4 and pr6. Sequencing of the DNA fragments was performed using an Automated Laser Fluorescent Sequencer (Labor für DNA ANALYTIK, Freiburg, Germany).

Acknowledgements

We thank E.Lanka for providing the VirD2 antiserum and information prior to publication. We especially thank M.-D.Chilton, N.Hernandez, H.Puchta, T.Hohn and J.Paszkowski for critically reading the manuscript, and C.Koncz and T.Bickle for discussions.

References

Argos,P. *et al.* (1986) The integrase family of site specific recombinases: regional similarities and global diversity. *EMBO J.*, **5**, 433–440.

Chaudhury,A.M., Dennis,E.S. and Brettell,R.I.S. (1994) Gene-expression following T-DNA transfer into plant cells is aphidicolin sensitive. *Aust. J. Plant Physiol.*, **21**, 125–131.

Citovsky,V., Zupan,J., Warnick,D. and Zambryski,P. (1992) Nuclear localization of *Agrobacterium* VirE2 protein in plant cells. *Science*, **256**, 1802–1805.

Dhaese,P., De Greve,H., Decraemer,H., Schell,J. and van Montagu,M. (1979) Rapid mapping of transposon insertion and deletion mutations in the large Ti-plasmids of *Agrobacterium tumefaciens*. *Nucleic Acids Res.*, **7**, 1837–1849.

Dürrenberger,F., Crameri,A., Hohn,B. and Koukolíková-Nicola,Z. (1989) Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. *Proc. Natl Acad. Sci. USA*, **86**, 9154–9158.

Earp,D.J., Lowe,B. and Baker,B. (1990) Amplification of genomic sequences flanking transposable elements in host and heterologous plants: a tool for transposon tagging and genome characterization. *Nucleic Acids Res.*, **18**, 3271–3279.

Filichkin,A. and Gelvin,S.B. (1993) Formation of a putative relaxation intermediate during T-DNA processing directed by the *Agrobacterium tumefaciens* VirD1, VirD2 endonuclease. *Mol. Microbiol.*, **8**, 915–926.

Gasser,C.S. and Fraley,R.T. (1989) Genetically engineering plants for crop improvement. *Science*, **244**, 1293–1299.

Gheysen,G., Villarroel,R. and van Montagu,M. (1991) Illigitimate recombination in plants: a model for T-DNA integration. *Genes Dev.*, **5**, 287–297.

Guyon,P., Chilton,M.D., Petit,A. and Tempé,J. (1980) Agropine in 'null-type' crown gall tumors: evidence for generality of the opine concept. *Proc. Natl Acad. Sci. USA*, **77**, 2693–2697.

Herrera-Estrella,A., van Montagu,M. and Wang,K. (1990) A bacterial peptide acting as a plant nuclear targeting signal: the amino-terminal portion of *Agrobacterium* VirD2 protein directs a β-galactosidase fusion protein into tobacco nuclei. *Proc. Natl Acad. Sci. USA*, **87**, 9534–9537.

Hooykaas,P.J.J. and Schilperoort,R.A. (1992) *Agrobacterium* and plant genetic enginering. *Plant Mol. Biol.*, **19**, 15–38.

Howard,E. and Citovsky,V. (1990) The emerging structure of the *Agrobacterium* T-DNA transfer complex. *BioEssays*, **12**, 103–108.

Howard,E.A., Winsor,B., De Vos,G. and Zambryski,P. (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.

Howard,E.A., Zupan,J.R., Citovsky,V. and Zambryski,P.C. (1992) The VirD2 protein of *A.tumefaciens* contains a C-terminal bipartite nuclear localization signal: implication for nuclear uptake of DNA in plant cells. *Cell*, **68**, 109–118.

Janssen,B.J. and Gardner,R.C. (1989) Localized transient expression of GUS in leaf discs following cocultivation with *Agrobacterium*. *Plant Mol. Biol.*, **14**, 61–72.

Jasper,F., Koncz,C., Schell,J. and Steinbiβ,H.H. (1994) *Agrobacterium* T-strand production *in vitro*: sequence specific cleavage and 5′-protection of single stranded templates by purified VirD2 protein. *Proc. Natl Acad. Sci. USA*, **91**, 694–698.

Jayaram,M. (1994) Phosphoryl transfer in FLP recombination: a template for strand transfer mechanisms. *Trends Biotechnol.*, **19**, 78–82.

Koncz,C., Nemeth,K., Redei,G.P. and Schell,J. (1994) Homology recognition during T-DNA integration into plant genome. In

Paszkowski,J. (ed.), *Homologous Recombination and Gene Silencing in Plants*. Kluwer, Dordrecht, The Netherlands, pp. 167–189.

Lee,K.Y., Lund,P., Lowe,K. and Dunsmuir,P. (1990) Homologous recombination in plant cells after *Agrobacterium*-mediated transformation. *Plant Cell*, **2**, 415–425.

Lehman,C.W., Trautman,J.K. and Carroll,D. (1994) Illegitimate recombination in *Xenopus*: characterisation of end-joined junctions. *Nucleic Acids Res.*, **22**, 434–442.

Lessl,M. and Lanka,E. (1994) Common mechanisms in bacterial conjugation and Ti-mediated T-DNA transfer to plant cells. *Cell*, **77**, 321–324.

Machida,Y. *et al.* (1993) Molecular interaction between *Agrobacterium* and plant cells. In Nester,E.W. and Verma,D.P.S. (eds), *Advances in Molecular Genetics of Plant–Microbe Interactions*. Kluwer Academic Publishers, Boston, MA, pp. 85–96.

Matsumoto,S., Ito,Y., Hosoi,T., Takahashi,Y. and Machida,Y. (1990) Integration of *Agrobacterium* T-DNA into a tobacco chromosome: possible involvement of DNA homology between T-DNA and plant DNA. *Mol. Gen. Genet.*, **224**, 309–316.

Mayerhofer,R. *et al.* (1991) T-DNA integration: a mode of illegitimate recombination in plants. *EMBO J.*, **10**, 697–704.

McBride,K.E. and Summerfelt,K.R. (1990) Improved binary vectors for *Agrobacterium*-mediated transformation. *Plant Mol. Biol.*, **14**, 269–276.

Miranda,A., Jansen,G., Hodges,L., Peralta,E.G. and Ream,W. (1992) *Agrobacterium tumefaciens* transfers extremely long T-DNAs by a unidirectional mechanism. *J. Bacteriol.*, **174**, 2288–2297.

Murashige,T. and Skoog,F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue. *Physiol. Plant.*, **15**, 473–493.

Offringa,R., de Groot,M.J.A., Haagsman,H.J., Does,M.P. and Hooykaas,P.J.J. (1990) Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. *EMBO J.*, **9**, 3077–3084.

Ohl,S., Offringa,R., van den Elzen,P.J.M. and Hooykaas,P.J.J. (1994) Gene replacement in plants. In Paszkowski,J. (ed.), *Homologous Recombination and Gene Silencing in Plants*. Kluwer, Dordrecht, The Netherlands, pp. 191–210.

Pansegrau,W., Schoumacher,F., Hohn,B. and Lanka,E. (1993) Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. *Proc. Natl Acad. Sci. USA*, **90**, 11538–11542.

Pansegrau,W., Schröder,W. and Lanka,E. (1994) Concerted action of three distinct domains in the DNA cleaving–joining reaction catalysed by relaxase (TraI) of conjugative plasmid RP4. *J. Biol. Chem.*, **269**, 2782–2789.

Paszkowski,J., Shillito,R.D., Saul,M.W., Mandak,V., Hohn,T., Hohn,B. and Potrykus,I. (1984) Direct gene transfer to plants. *EMBO J.*, **3**, 2717–2722.

Petit,A., David,C., Dahl,G., Ellis,J.G., Guyon,P., Casse-Delbart,F. and Tempé,J. (1983) Further extension of the opine concept: plasmids in *Agrobacterium rhizogenes* cooperate for opine degradation. *Mol. Gen. Genet.*, **190**, 204–214.

Puchta,H. and Hohn,B. (1991) A transient assay in plant cells reveals a positive correlation between extrachromosomal recombination rates and length of homologous overlap. *Nucleic Acids Res.*, **19**, 2693–2700.

Ream,W. (1989) *Agrobacterium tumefaciens* and interkingdom genetic exchange. *Annu. Rev. Phytopathol.*, **27**, 583–618.

Rossi,L., Escudero,J., Hohn,B. and Tinland,B. (1993a) Efficient and sensitive assay for T-DNA dependent transient gene expression. *Plant Mol. Biol. Reporter*, **12**, 220–229.

Rossi,L., Hohn,B. and Tinland,B. (1993b) VirD2 protein carries nuclear localization signals important for transfer of T-DNA to plants. *Mol. Gen. Genet.*, **239**, 345–353.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Scheiffele,P.R., Pansegrau,W. and Lanka,E. (1995) Initiation of *Agrobacterium tumefaciens* T-DNA processing: purified proteins VirD1 and VirD2 catalyze site- and strand-specific cleavage of superhelical T-border DNA *in vitro*. *J. Biol. Chem.*, **270**, 1269–1276.

Schultze,M., Hohn,T. and Jiricny,J. (1990) The reverse transcriptase of CaMV is translated separately from the capsid gene. *EMBO J.*, **9**, 1177–1185.

Shurvinton,C.E., Hodges,L. and Ream,W. (1992) A nuclear localization signal and the C-terminal omega sequence in the *Agrobacterium tumefaciens* VirD2 endonuclease are important for tumor formation. *Proc. Natl Acad. Sci. USA*, **89**, 11837–11841.

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**, 706–711.

Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorf,J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. In Goedel,D.V. (ed.), *Gene Expression Technology*. Academic Press, New York, NY, pp. 60–89.

Tinland,B. and Hohn,B. (1993) Plant genome transformation by *Agrobacterium tumefaciens* is mediated by a nucleoprotein complex. In Doerfler,W. and Böhm,P. (eds), *Virus Strategies*. Verlag Chemie, Weinheim, Germany, pp. 349–359.

Tinland,B., Koukolíková-Nicola,Z., Hall,M.N. and Hohn,B. (1992) The T-DNA-linked VirD2 protein contains two distinct functional nuclear localization signals. *Proc. Natl Acad. Sci. USA*, **89**, 7442–7446.

Tinland,B., Hohn,B. and Puchta,H. (1994) *Agrobacterium tumefaciens* transfers single-stranded T-DNA into the plant cell nucleus. *Proc. Natl Acad. Sci. USA*, **91**, 8000–8004.

Towbin,H., Staehelin,T. and Gordon,J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl Acad. Sci. USA*, **76**, 4350–4354.

Vogel,A.M. and Das,A. (1992) Mutational analysis of *Agrobacterium tumefaciens virD2*: tyrosine-29 is essential for endonuclease activity. *J. Bacteriol.*, **174**, 303–312.

Wang,K., Stachel,S.E., Timmerman,B., van Montagu,M. and Zambryski,P.C. (1987) Site-specific nick in the T-DNA border sequence as a result of *Agrobacterium vir* gene expression. *Science*, **235**, 587–591.

Ward,E.R. and Barnes,W.M. (1988) VirD2 protein of *Agrobacterium tumefaciens* is very tightly linked to the 5′ end of T-strand DNA. *Science*, **242**, 927–930.

Waters,V.L. and Guiney,D.G. (1993) Processes at the nick region link conjugation, T-DNA transfer and rolling circle replication. *Mol. Microbiol.*, **9**, 1123–1130.

Wilkins,B. and Lanka,E. (1993) DNA processing and replication during plasmid transfer between Gram-negative bacteria. In Clewell,D.B. (ed.), *Bacterial Conjugation*. Plenum Press, New York, NY, pp. 105–136.

Young,C. and Nester,E.W. (1988) Association of the VirD2 protein with the 5′ end of T strand in *Agrobacterium tumefaciens*. *J. Bacteriol.*, **8**, 3367–3374.

Yusibov,V.M., Steck,T.R., Gupta,V. and Gelvin,S.B. (1994) Association of single-stranded transferred DNA from *Agrobacterium tumefaciens* with tobacco cells. *Proc. Natl Acad. Sci. USA*, **91**, 2994–2998.

Zambryski,P.C. (1992) Chronicles from the *Agrobacterium*–plant cell DNA transfer story. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **43**, 465–490.

Received on March 20, 1995; revised on April 21, 1995