The EMBO Journal vol.4 no.4 pp.891-898, 1985

A Tn*3* *lacZ* transposon for the random generation of β-galactosidase gene fusions: application to the analysis of gene expression in *Agrobacterium*

Scott E.Stachel[1,2], Gynheung An[1,3], Carlos Flores[1,4] and Eugene W.Nester[1]

[1]Department of Microbiology and Immunology, SC-42, University of Washington, Seattle, WA 98195, and [2]Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, CA 94143, USA

[3]Present address: Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340, USA
[4]Present address: Department of Biotechnology, Imperial College of Science and Technology, London SW7 2AZ, UK

Communicated by M.van Montagu

The construction and use of a Tn*3*-*lac* transposon, Tn*3*-HoHo1, is described. Tn*3*-HoHo1 can serve as a transposon mutagen and provides a new and useful system for the random generation of both transcriptional and translational *lacZ* gene fusions. In these fusions the production of β-galactosidase, the *lacZ* gene product, is placed under the control of the gene into which Tn*3*-HoHo1 has inserted. The expression of the gene can thus be analyzed by monitoring β-galactosidase activity. Tn*3*-HoHo1 carries a non-functional transposase gene; consequently, it can transpose only if transposase activity is supplied in *trans*, and is stable in the absence of this activity. A system for the insertion of Tn*3*-HoHo1 into sequences specifically contained within plasmids is described. The applicability of Tn*3*-HoHo1 was demonstrated studying three functional regions of the *Agrobacterium tumefaciens* A6 Ti plasmid. These regions code for octopine catabolism, virulence and plant tumor phenotype. The regulated expression of genes contained within each of these regions was analyzed in *Agrobacterium* employing Tn*3*-HoHo1 generated *lac* fusions.

Key words: transposon Tn*3*/*lacZ* gene/gene fusions/β-galactosidase/*Agrobacterium*

Introduction

Transposon mutagenesis and the generation and analysis of gene fusions have been widely used in the study of gene organization and expression of a variety of organisms. Transposon mutagenesis allows the mutation, identification, and isolation of genes contained within large genetic targets, including prokaryotic (Ruvkun and Ausubel, 1981) and eukaryotic genomes (Bingham *et al.*, 1981; Copeland *et al.*, 1983), while gene fusions allow the analysis and manipulation of the expression and regulation of genes whose functions are unknown or difficult to assay (Bassford *et al.*, 1978; Franklin, 1978).

In gene fusions, the control sequences of a gene of interest are placed in front of the coding sequences of a 'reporter' gene whose product can be readily assayed. Thus, expression of the gene can be monitored by measuring the reporter gene product, and genetic and environmental factors that affect the expression can be determined and manipulated. The *Escherichia coli* β-galactosidase structural gene, *lacZ*, has often been employed as the reporter gene in the contruction of gene fusions. β-Galactosidase is easily and quantitatively assayed, is amenable to genetic manipulations and is active in a variety of organisms and when contained within hybrid proteins (Bassford *et al.*, 1978). Employing *lacZ*, two types of gene fusions can be made. Transcriptional fusions are generated using sequences that contain the *lacZ* gene including its translational initiation signal; expression of a transcription fusion results in the production of wild-type β-galactosidase. In a translational fusion, the translational initiation signal and the amino-terminal coding sequences of a gene are linked 5′ directly to the *lacZ* coding sequences; expression of a translational fusion results in the production of a chimeric protein with β-galactosidase activity. Both transcriptional and translational fusions are useful for studying prokaryotic gene expression. Translational fusions also have a number of additional applications. They can be used to study gene expression in eukaryotic systems, because translation of the prokaryotic reporter gene is placed under the control of eukaryotic translational control sequences (Lis *et al.*, 1983). They also can be used to determine the site of translational initiation of a gene and to analyze its control (Lee *et al.*, 1981; Casadaban *et al.*, 1982). Furthermore, a hybrid protein produced from a translational fusion can be useful for studying properties of the original protein encoded by the target gene (Schuman *et al.*, 1980; Hall and Silhavy, 1981).

Transposable elements that contain *lacZ* coding sequences combine the techniques of transposon mutagenesis and the generation of *lac* gene fusions; such elements allow the random insertion of *lacZ* into a wide variety of DNA sequences. *LacZ* transposons that generate transcriptional gene fusions (Casadaban and Cohen, 1979; Kroos and Kaiser, 1984), or translational gene fusions (Casadaban and Chou, 1984) have been constructed previously, and used to determine the location (Kenyon and Walker, 1980) and the transcriptional orientation of genes (Dixon *et al.*, 1980; Wanner *et al.*, 1981) and to analyze their expression (Kenyon *et al.*, 1982; Stern *et al.*, 1984). In this paper we describe the construction and use of a new Tn*3*-*lacZ* element, Tn*3*-HoHo1. This transposon can be used in the generation of both transcriptional and translational *lac* fusions with plasmid DNA sequences, and these fusions are stable, as Tn*3*-HoHo1 cannot self-transpose.

To test our Tn*3*-*lac* system and to provide novel information about an organism in which questions of gene organization and gene expression were relatively unstudied, we chose to study different genetic regions of the *Agrobacterium tumefaciens* Ti plasmid. *A. tumefaciens* is a phytopathogen that is able to transfer a specific DNA segment, the T-DNA, from its large (200 kb) Ti plasmid to plant cells (Caplan *et al.*, 1983; Nester *et al.*, 1984). The genes contained within the T-DNA do not have a known function within the bacterial cell, but their expression within the transformed plant cell results in altered growth of the plant cell and synthesis of novel compounds, called opines. The Ti-plasmid genes involved in T-DNA transfer and integration are not contained within T-DNA sequences but are located on another region of the plasmid designated the virulence (*vir*) region. The pTiA6 plasmid also contains a genetic region (*occ*) that codes for the

catabolism of octopine, a specific opine. Tn*3*-HoHo1 was used to study the expression of these three genetic regions (T-DNA, *vir*, *occ*) of the Ti plasmid within *Agrobacterium*.

Results and Discussion

Tn3::lac constructions

We set out to construct a molecule that can be used to generate fusions between cloned DNA sequences of interest and the β-galactosidase *lac*Z gene. Such a molecule should be able to transpose at high frequency into plasmid DNA, generate both transcriptional and translational hybrid *lacZ* gene fusions and be stable after transposition. Stability could be achieved by requiring that transposition can occur only if transposase activity is provided in *trans*. The bacterial transposon Tn*3* is well suited for such a construct. Tn*3* preferentially transposes into plasmid *versus* chromosomal DNA sequences (F.Heffron, unpublished results) at a frequency of $\sim 10^{-3}$/cell (Heffron *et al.*, 1977), and its DNA sequence and genetic organization have been determined (Heffron *et al.*, 1979).

The Tn*3*-*lac* transposon, Tn*3*-HoHo1, which is carried by the plasmid pHoHoI, was constructed as shown in Figure 1A. pFH33 served as the starting point in this construction. pFH33 is a pMB8::Tn*3* derivative that contains an *Eco*RI linker inserted 84 bp inside of the left inverted repeat (IR_L) of Tn*3* (Heffron *et al.*, 1979). This linker mutation results in the abolition of transposase activity and provides a convenient restriction site closely adjacent to IR_L for the insertion of *lac* operon sequences. The *lac* sequences within Tn*3*-HoHo1 are oriented such that gene expression, which initiates in sequences into which the element has transposed and proceeds through IR_L, will result in β-galactosidase production. Transcriptional control sequences do not precede the *lac* sequences within Tn*3*-HoHo1; therefore, β-galactosidase production is dependent on expression that initiates outside of and reads into this element.

Tn*3*-HoHo1 was constructed to serve in the generation of both transcriptional and translational *lac* fusions. The *lac*Z portion of the pGA300-derived *lac* fragment in Tn*3*-HoHo1 encodes a hybrid β-galactosidase protein of mol. wt. ~146 000 daltons, whose amino terminus is composed of sequences derived from *tuf*B and *rpo*B fused in-frame to the eighth amino acid of *lac*Z (Lee *et al.*, 1981; An *et al.*, 1982). This fragment has been inserted into Tn*3* such that translation into and through IR_L is open and in-frame; therefore, such translation can result in the production of a hybrid β-galactosidase protein whose amino terminus is encoded by the gene into which Tn*3*-HoHo1 has inserted. Generation of active gene::*lac* fusions with this element, however, does not depend on in-frame translation. At least six methionine codons occur in-frame 5′ of the *lac*Z structural gene in the construct, and translation can apparently initiate at one or more of these codons if transcription occurs across this region. Because the pGA300-derived *lac* fragment used in the construction of Tn*3*-HoHo1 contains sequences encoding the SV40 early transcript acceptor splice site and polyadenylation signal (An *et al.*, 1982; Figure 1A), these sequences occur within Tn*3*-HoHo1 downstream of and in the same orientation as the *lac*ZYA sequences (Figure 1C). Thus, *lac* fusions functional in both prokaryotic and eukaryotic systems can potentially be generated. Tn*3*-HoHo1 is 14.25 kb in size.

Tn*3*-HoHo1 is defective for transposase activity. To effect Tn*3*-HoHo1 transposition, pSShe was constructed (Figure 1B) to supply *tnp*A activity in *trans*. pSShe contains Tn*3* deleted for sequences rightward of the Tn*3* *Bam*HI site; thus, it encodes transposase activity, but cannot be transposed due to the absence of IR_L. The structures of Tn*3*-HoHo1 and pSShe, and the nucleotide sequence of the left terminus of Tn*3*-HoHo1 are shown in Figure 1.

Tn3-lac transposition

Tn*3*-HoHo1 insertions into target sequences that were cloned into the mobilizable wide-host range plasmid, pVCK102, were selected in *E. coli* using the transconjugant procedure described in detail in Materials and methods. In brief, the transposon donor strain which contained pHoHo1, pSShe and the target plasmid was mated with the recipient strain, SF800, and transconjugants containing Tn*3*-HoHo1::target plasmids were obtained at a frequency of $\sim 10^{-4}$. The mating efficiency of the target plasmid into SF800 is $\sim 10^{-1}$, so Tn*3*-HoHo1 transposes at a frequency of $\sim 10^{-3}$. Transconjugants were obtained at a frequency of $\sim 10^{-7}$ when pSShe was omitted from the donor transposon strain. These frequencies are similar to those obtained by Heffron *et al.* (1977) for complementation of Tn*3* *tnp*A mutations in *trans*. Also, >99% of the transconjugants contained Tn*3*-HoHo1 inserted into the target plasmid, and deletions were associated with ~1% of the Tn*3*-HoHo1::target plasmids.

An alternative method that we have employed to select for transposition of Tn*3*-HoHo1 into target cosmid plasmids is the *in vivo* packaging procedure of White *et al.* (1983). Using this method, Tn*3*-HoHo1 transductants were obtained at a frequency of $\sim 5 \times 10^{-4}$. When the initial target cosmid was >40 kb in size, a majority of the resultant Tn*3*-HoHo1::target transductants contained deletions. Thus, this method can be used to generate random deletions in a target plasmid.

We tested Tn*3*-HoHo1 by using the element to study the organization and regulated expression of three separate genetic regions of the A6 Ti plasmid within *A. tumefaciens*.

Analysis of occ::lac gene fusion

The *occ* genes confer upon *Agrobacterium* the ability to utilize both octopine and arginine as sole carbon and nitrogen sources, and expression of these genes is induced by octopine (Montoya *et al.*, 1977; Klapwijk and Schilperoort, 1979); this induction appears to be mediated through a negative regulatory system (Klapwijk and Schilperoort, 1979). Knauf and Nester (1982) previously constructed pVCK261, which contains the complete *occ* region of pTiA6 cloned into the wide-host range cosmid pVCK102. Their studies suggested that *occ* is encoded at minimum by the 6.2-kb *Eco*RI fragment 8 (De Vos *et al.*, 1981), and is probably expressed from the rightward to leftward direction with respect to the map of pVCK261 shown in Figure 2.

pVCK261 was mutagenized with Tn*3*-HoHo1 and seven pVCK261::Tn*3*-HoHo1 derivatives, pSM100 to pSM106, were introduced by conjugation into *Agrobacterium* strains A348 and A136, that contain or lack pTiA6, respectively. pSM100 carries Tn*3*-HoHo1 within the vector portion of the plasmid, while pSM101 to pSM106 carry Tn*3*-HoHo1 within the cloned *occ* region (Figure 2). These *Agrobacterium* strains were grown to logarithmic phase in liquid minimal media and in liquid minimal media supplemented with 100 μg/ml octopine for 6 h. Bacteria were then harvested and their β-galactosidase activities were determined (Table I). The parental A136 and A348 strains have negligible levels of β-galactosidase activity (<1 unit), and all pSM strains exhibited similar low activities when grown in minimal media. The pSM102, pSM105 and pSM106 strains displayed increased β-galactosidase activities when grown in the presence of octopine (Table I), demonstrating that expression of octopine catabolism genes occurs in a rightward to leftward orien-

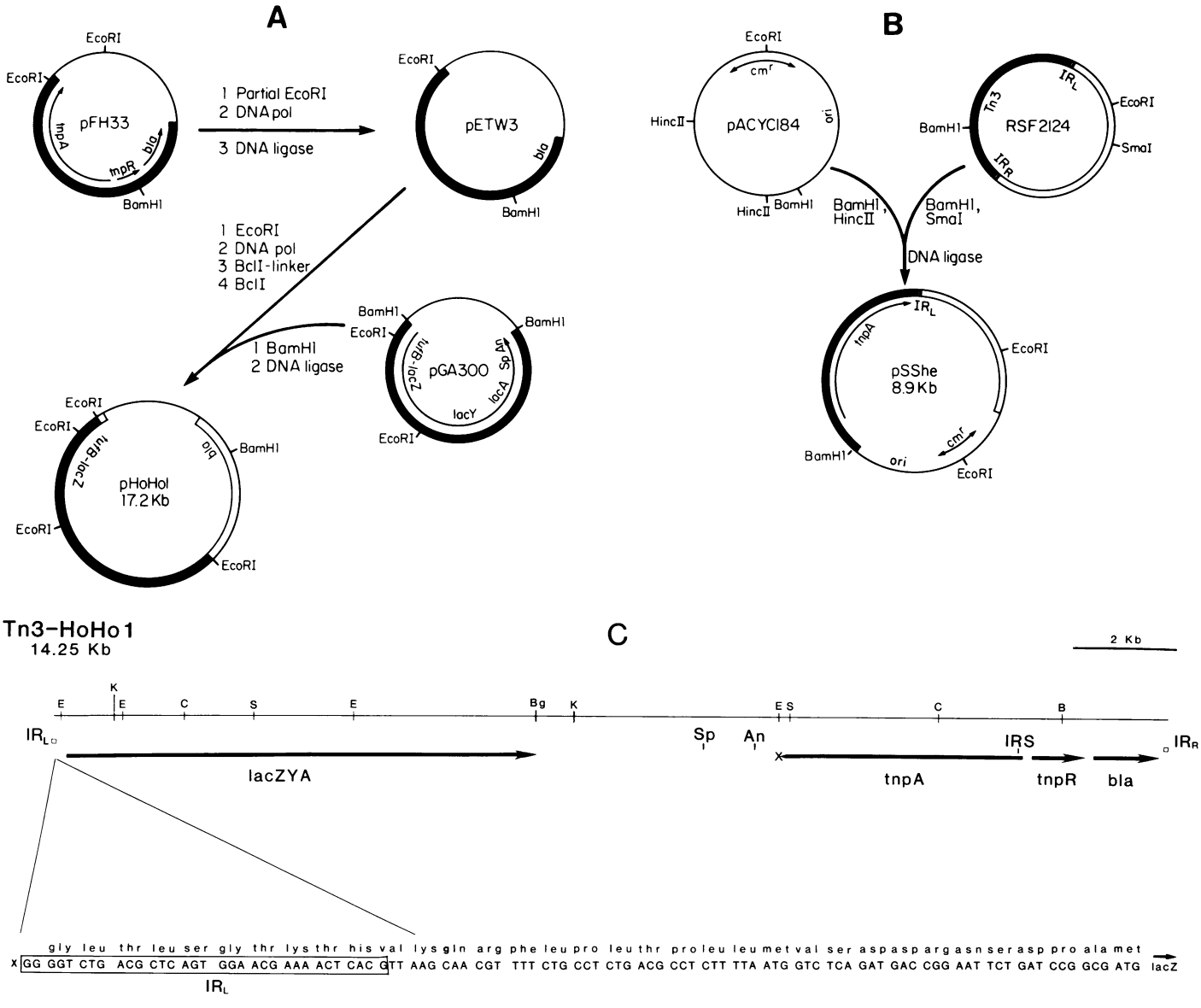


Fig. 1. Tn*3-lac* constructions. **(A)** Construction of pHoHo1. The pMB8 *Eco*RI site of pFH33 (Heffron *et al.*, 1978) was removed to yield pETW3 by partial digestion with *Eco*RI, incubation with DNA polymerase (Klenow fragment) in the presence of dATP and dTTP, and circularization with T4 DNA ligase. The *Eco*RI site of pETW3 was converted to a *Bcl*I site by digestion with *Eco*RI, filling this site by Klenow treatment, followed by the addition of 8-mer *Bcl*I linkers and T4 DNA ligase. This molecule was subsequently digested with *Bcl*I and ligated to the 9.3-kb *Bam*HI fragment of pGA300 (An *et al.*, 1982) to give pHoHo1. The above treatments generated new *Eco*RI sites that flank the *lac* sequences inserted into Tn*3* on pHoHo1. pHoHo1 is 17.2 kb in size. In the plasmid representations, single lines designate pMB8 sequences. Black areas represent Tn*3* sequences in pFH33 and pETW3, and *lac* sequences in pGA300. In the pHoHo1 diagram, the Tn*3-lac* transposon, Tn*3*-HoHo1, is represented by the open and black areas together. The transcriptional orientations of the Tn*3* *tnp*A, *tnp*R and *bla* genes, and the *tuf*B-*lac*Z gene, are indicated by arrows. **(B)** Construction of pSShe. The *Bam*HI-*Sma*I digest of the Tn*3*::ColE1 plasmid RSF2124 (So *et al.*, 1975) was ligated to the *Bam*HI-*Hind*III digest of pACYC184 (Chang and Cohen, 1978) to yield pSShe. pSShe is 8.9 kb in size, and carries the *tnp*A coding region of Tn*3*. In the plasmid representations single lines designate pACYC184 sequences; open areas, ColE1 sequences; black areas, Tn*3* sequences; ori, the pACYC184 origin of replication; and Cm^R, the chloramphenicol resistance gene. **(C)** Structure and organization of Tn*3*-HoHo1. The vertical lines indicate the positions of various restriction sites in Tn*3*-HoHo1. *Hind*III and *Sal*I sites do not occur within the elements. The genetic organization of Tn*3*-HoHo1 is shown below the restriction map. The respective coding region of each gene within the element is indicated by a bold line, and the transcriptional orientation of each gene is indicated by an arrow. *tnp*R and *bla* are wild-type, while *tnp*A is non-functional owing to the *lac* sequences inserted into its 3′ end. The *lac* operon sequences are intact but lack a functional promoter. The DNA sequence of the left end of Tn*3*-HoHo1 was determined by the method of Maxam and Gilbert (1980). The DNA sequence, along with the predicted amino acid sequence, up to the native initiation codon of the *tuf*B-*lac*Z-encoded protein, is shown in the bottom of the figure; the Shine-Dalgarno sequence of *tuf*B has been eliminated in Tn*3*-HoHo1 and translation can potentially initiate at any ATG that occurs upstream of the *lac*Z structural gene. The IR_L sequence is boxed, and X represents sequences leftward of IR_L into which Tn*3*-HoHo1 has inserted. Symbols: B, *Bam*HI; Bg, *Bgl*II; C, *Cla*I; E, *Eco*RI; K, *Kpn*I; S, *Sst*I; IR_L, left-inverted repeat; IR_R, right-inverted repeat; IRS, internal resolution site; Sp, eukaryotic splice site; An, eukaroyotic polyadenylation signal; *lac*ZYA, *EW. coli lac* operon; *tnp*A, transposase; *tnp*R, resolvase; *bla*, β-lactamase.

tation (Figure 2). The β-galactosidase activity measured for each of the *occ*::*lac* strains grown in octopine medium was independent of the presence or absence of wild-type *occ* carried in *trans* on the Ti plasmid (data not shown), suggesting that none of the Tn*3*-HoHo1 insertions affected genes involved in the regulation of *occ* expression. The time course of octopine induction of β-galactosidase activity for the strain A348(pSM102) was determined and is shown in Figure 3. These kinetics closely follow

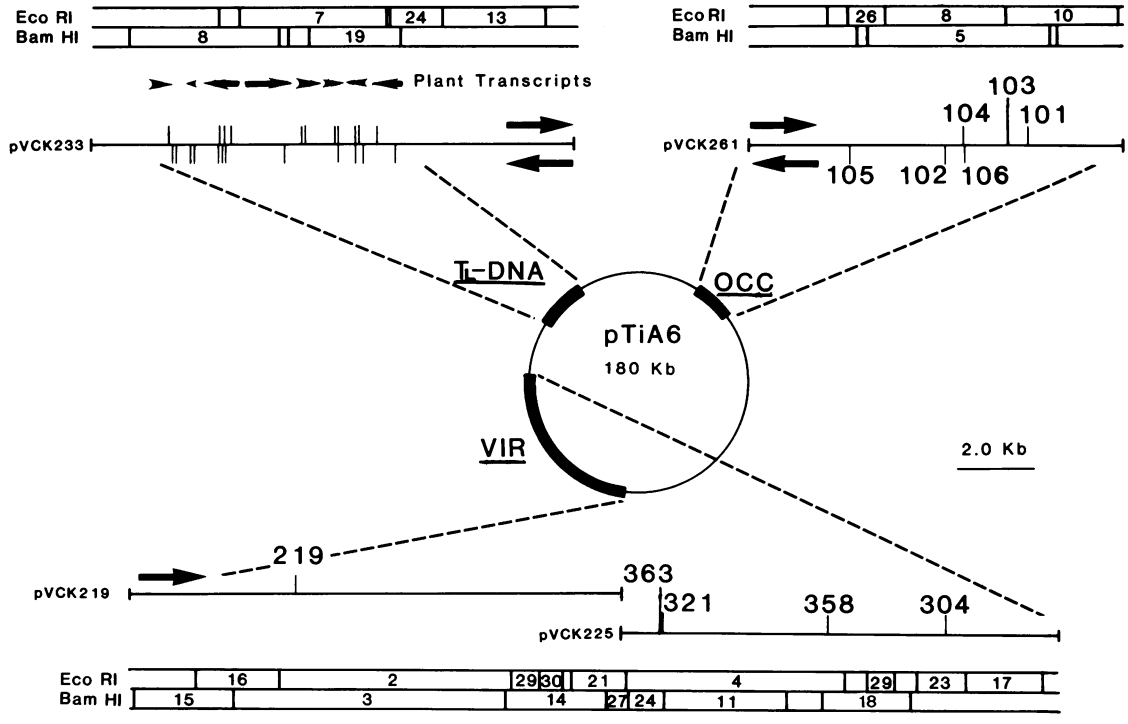


Fig. 2. Tn*3*-HoHo1 insertions within TL-DNA, *occ* and *vir* sequence of pTiA6. The *Bam*HI and *Eco*RI restriction maps (De Vos *et al.*, 1981) of TL-DNA, *occ* and *vir*, and the respective location of each of these regions within pTiA6 (Garfinkel and Nester, 1980; De Greve *et al.*, 1981) are shown. The genetic location, size and transcriptional orientation of the TL-DNA genes known to be expressed in transformed plants (Garfinkel *et al.*, 1981; Willmitzer *et al.*, 1982) are indicated as 'plant transcripts' under the restriction map of the TL-DNA region. TL-DNA, *occ* and *vir* sequences have been separately cloned into the cosmid vector pVCK102 (Knauf and Nester, 1982), and the complete TL-DNA and *occ* regions are carried by pVCK233 and pVCK261, respectively, and *vir* sequences are carried by pVCK219 and pVCK225. These four cosmid clones are represented in the figure by horizontal lines. Vertical lines above and below the horizontal lines represent separate Tn*3*-HoHo1 insertions into the cloned sequences. Tn*3*-HoHo1 can insert into sequences in either of two orientations with regard to its *lacZ* coding sequences. In the figure, the *lacZ* orientation of an insertion represented by an upper line is leftward to rightward while that by a lower line is rightward to leftward as indicated by the arrows.

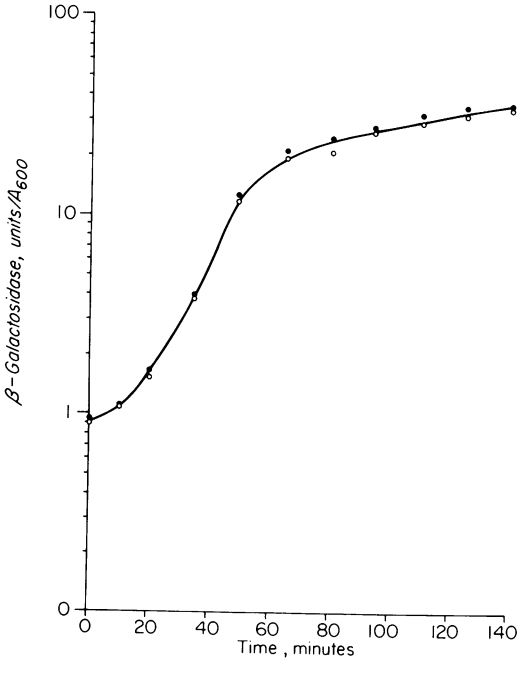


Fig. 3. Kinetics of octopine induction of β-galactosidase in A348(pSM102). Cells were grown in AB/glucose liquid medium. 100 μg/ml octopine was added to two exponential phase cultures, represented by ● and ○. Aliquots were periodically removed and the β-galactosidase activity in each culture as a function of time following octopine addition was determined.

Table I. Characteristics of A136 (*occ*::*lac*) strains

Strain	β-Galactosidase units		Utilization	
	− Octopine	+ Octopine	Octopine	Arginine
A136/100	2.54	2.47	+	+
101	1.35	2.37	+	+
102	0.992	102.6	−	−
103	3.24	1.56	+	+
104	1.07	1.14	−	−
105	0.943	18.93	−	−
106	1.34	82.4	−	−

Seven A136 (pVCK261::Tn*3*-HoHo1) strains were characterized for octopine-inducible β-galactosidase activity and carbon source utilization. The β-galactosidase activity of cells grown 6 h in AB/glucose liquid medium and AB/glucose liquid medium supplemented with 100 μg/ml octopine was determined. Carbon source utilization was assessed by scoring growth on AB agar containing 100 μg/ml octopine and AB agar containing 600 μg/ml arginine and 10 μg/ml octopine.

those seen for octopine-induced octopine uptake by A348 (Montoya *et al.*, 1977; Klapwijk *et al.*, 1977). Interestingly, although pVCK261 does not confer octopine utilization upon *E. coli*, pSM102, pSM105 and pSM106 produced low levels of octopine inducible β-galactosidase activity in the *E. coli lac* deletion strain MC1061. Octopine induction of *occ* in *Agrobacterium* is repressed in the presence of 0.5% casamino acids (A.Montoya, personal communication; S.Stachel, unpublished results); *occ* induction in *E. coli* was similarly seen to be amino acid repressible, suggesting that some mechanisms for gene control might

be analogous between these two organisms.

The A136(pSM) strains were grown on octopine minimal media agar (Hooykaas *et al.*, 1979) and on arginine minimal media agar (Knauf and Nester, 1982) containing 10 μg/ml octopine to determine the effect of Tn*3*-HoHo1 insertions on octopine and arginine utilization, respectively (Table I). The pSM100, pSM101 and pSM103 strains grew on both these media, while the pSM102, pSM104, pSM105 and pSM106 strains did not. This suggested that a regulated *occ* promoter occurs within the 1-kb region bounded by the pSM103 and pSM106 insertions. The regulation of this promoter was not affected by the pSM102, pSM105 and pSM106 insertions. These experiments with *occ* indicated that Tn*3*-HoHo1 efficiently generates *lac* fusions that can be used to analyze gene organization and expression in *Agrobacterium*.

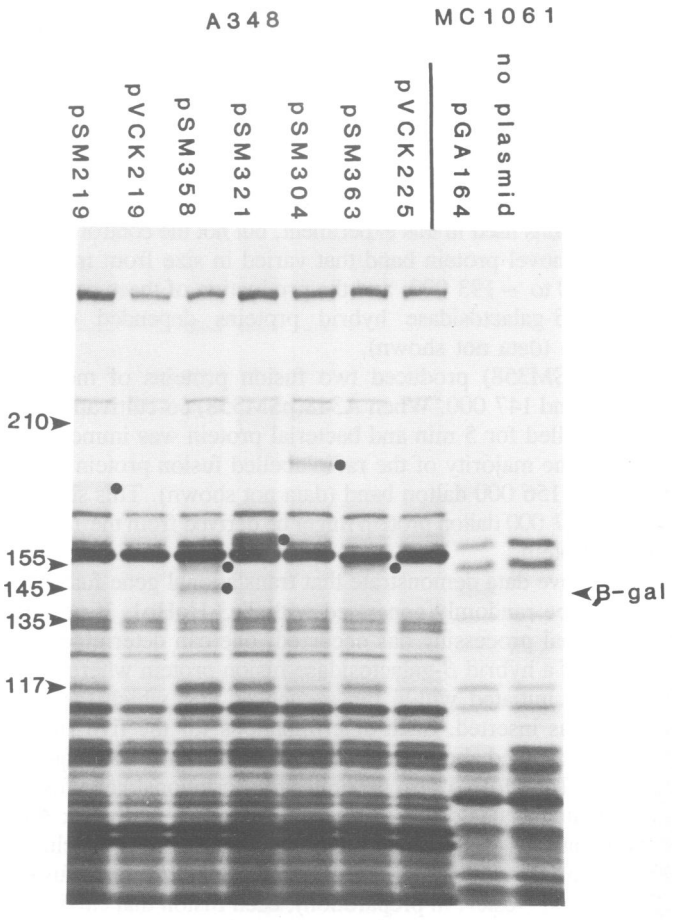
Analysis of translational vir::lacZ gene fusions

The ability of Tn*3*-HoHo1 specifically to generate translational *lac* fusions was demonstrated using the *vir* genes of the *Agrobacterium* pTiA6 plasmid as a model system. The *vir* region of pTiA6 spans ~50 kb of the plasmid and contains at least eight separate transcriptional units (Stachel *et al.*, in preparation). Transposon insertions in these loci result in loss or attenuation of virulence (Garfinkel *et al.*, 1981; Ooms *et al.*, 1980; Hooykaas *et al.*, 1984; Klee *et al.*, 1983; Hille *et al.*, 1984; Stachel *et al.*, in preparation).

Adjacent segments of the pTiA6 *vir* region, plasmid clones pVCK219 and pVCK225 (Knauf and Nester, 1982; Figure 2), were mutagenized with Tn*3*-HoHo1 and introduced into A348. Five A348(*vir*::*lac*) transconjugants, pSM219, pSM304, pSM321, pSM358 and pSM363 (Figure 2) are analyzed here in detail. These strains were grown both in minimal medium and minimal medium supplemented with octopine, and the units of β-galactosidase activity were determined. Three of these strains, pSM321, pSM363 and pSM358, displayed significantly higher units of β-galactosidase activity for growth in minimal medium than other *vir*::*lac* strains or than the *occ*::*lac* strains (Figure 4, Table I). These results suggest that the regions of Tn*3*-HoHo1 insertion in these strains are constitutively expressed in *Agrobacterium*. Indeed, Gelvin *et al.* (1981) have reported that steady-state transcription occurs from the same region of *vir* as that of the Tn*3*-HoHo1 insertions in pSM321 and pSM363. None of the *vir*::*lac* strains exhibited octopine-induced β-galactosidase activity.

Plant cells produce a small mol. wt. factor(s) that specifically induces the expression of *Agrobacterium vir* genes (Stachel *et al.*, 1984 and in preparation). Regenerating meosphyll protoplasts of *Nicotiana tabacum* Xanthi, which are susceptible to high efficiency transformation by *Agrobacterium* (An *et al.*, 1985) produce this factor. *Vir* induction has been seen by direct assay of *vir* RNA and has also been assessed using *vir*::*lac* fusions. In the latter experiments, the levels of β-galactosidase activity in A348(*vir*::*lac*) strains are seen to increase markedly, with respect to other growth conditions, as a result of co-cultivation with Xanthi protoplasts.

The *lacZ* gene used in the construction of Tn*3*-HoHo1 encodes a β-galactosidase protein of mol. wt. ~146 000. To demonstrate that translational fusions can be generated with Tn*3*-HoHo1, it is necessary to visualize hybrid proteins greater than this size synthesized from Tn*3*-HoHo1-generated *lac* fusions. Because the five *vir*::*lac* strains discussed above exhibited comparatively high levels (7- to 30-fold over the levels of octopine-induced β-galactosidase activity exhibited by the *occ*::*lac* strains) of plant-



Strain		β-Galactosidase Units − plant cells	β-Galactosidase Units + plant cells	β-Galactosidase Fusion M.W. ($\times 10^{-3}$)
A348/	pVCK219	2.2	1.9	-
	pVCK225	1.8	2.0	-
	pSM219	10.3	726	182
	pSM304	8.2	859	193
	pSM321	81	1050	164
	pSM358	46	3120	156,147
	pSM363	72	1063	156
MC1000/no plasmid		1.2	-	-
	pGA164	1320	-	146

Fig. 4. Analysis of hybrid β-galactosidase proteins encoded by translational *vir*::*lac* fusions. Radiolabelled total cell extracts were analyzed by one-dimensional 7% SDS-polyacrylamide gel electrophoresis. *Agrobacterium* strains are A348 containing different plasmids. pSM219 is derived from pVCK219, and pSM304, pSM321, pSM358 and pSM363 are derived from pVCK225. The Tn*3*-HoHo1 insertions in pSM304, pSM321, pSM358 and pSM363 are located within *vir* genes and cause an avirulent phenotype when recombined into the Ti plasmid. The pSM219 insertion occurs within a plant-inducible sequence that is not required for virulence (Stachel *et al.*, in preparation). The *tufB*-*lacZ* coding sequences contained within Tn*3*-HoHo1 encode a protein of 146 000 daltons that has 226 amino acids of *tufB* and ~50 amino acids of *rpoB* fused to the eighth amino acid of β-galactosidase. This protein was expressed in the *E. coli* strain MC1061 from the plasmid pGA164 (Lee *et al.*, 1981), and is indicated by the arrow on the right. This plasmid is not expressed in *Agrobacterium*. The units of β-galactosidase activity expressed by the *Agrobacterium* strains in AB/glucose liquid medium (minus plant cells) and in co-cultivation (plus plant cells), along with the estimated size of the induced hybrid β-galactosidase proteins, are shown in the lower section of the figure.

induced β-galactosidase activity (Figure 4; Table I), they were examined for the presence of hybrid β-galactosidase proteins. These strains, along with the parent plasmid strains A348(pVCK219) and A348(pVCK225), were radiolabelled for 16 h with [^{35}S]methionine during co-cultivation with plant cells, and total bacterial protein was subsequently analyzed by SDS-polyacrylamide gel electrophoresis (Figure 4). Each of the *vir*::*lac* strains used in this experiment, but not the control strains, yielded a novel protein band that varied in size from mol. wt. ~147 000 to ~193 000, and the production of these presumptive *vir*::β-galactosidase hybrid proteins depended on co-cultivation (data not shown).

A348(pSM358) produced two fusion proteins of mol. wt. 156 000 and 147 000. When A348(pSM358) co-cultivation was pulse-labelled for 5 min and bacterial protein was immediately isolated, the majority of the radiolabelled fusion protein occurred in the 156 000 dalton band (data not shown). This suggests that the 147 000 dalton protein might be derived from the 156 000 dalton protein.

The above data demonstrate that translational gene fusions to *lac*Z can be randomly generated with Tn*3*-HoHo1. If no post-translational processing has occurred, one can determine from the size of a hybrid β-galactosidase fusion protein where translation has initiated within the sequences into which the Tn*3*-HoHo1 has inserted. Such determinations for the five *vir*::*lac* fusions examined above have been borne out by DNA sequence analysis. Tn*3*-HoHo1 also generates transcriptional gene fusions that initiate β-galactosidase translation from one of the six ATG codons internal to the Tn*3*-HoHo1 *lac*Z gene. This conclusion is based on the analysis of several hundred *vir*::*lac* gene fusions (Stachel and Nester, in preparation); each fusion that carries *lac* orientated in the direction of transcription of the *vir* gene to which it is fused displays *vir*-inducible *lac*Z expression. If Tn*3*-HoHo1 only generates translational fusions, then statistically only 1/3 of the fusions would be in-frame and thus inducible. Since all fusions are active, transcriptional as well as translational fusions are present.

Expression of TL-DNA sequences within Agrobacterium

The expression of T-DNA sequences within *Agrobacterium* has been previously reported. Gelvin *et al.* (1981) observed that the pTiB6-806 T-DNA region is transcribed at a low steady-state level in *Agrobacterium*, and that the pattern of this transcription appears to be less specific than that seen for the same sequences in crown gall cells. Similar results were obtained by Janssens *et al.* (1984) for the T-DNA region of pTiC58. Also, Schröder *et al.* (1984) demonstrated that the transcript 2 gene of the pTiAch5 T-DNA is functionally expressed in *Agrobacterium*. At present, however, no data exist which indicates that an *Agrobacterium* cell which harbors T-DNA genes has a selective advantage over a cell which does not. If T-DNA genes do serve a functional role in *Agrobacterium* their expression might be expected to be specifically regulated within the bacterial cell. We have sought to find evidence for such regulation. T-DNA::*lac*Z fusions generated with Tn*3*-HoHo1 were used to examine the expression of T-DNA sequences within *Agrobacterium* under a variety of conditions.

The complete pTiA6 TL-DNA (Thomashow *et al.*, 1980) has been cloned into pVCK102 to yield pVCK233 (Knauf and Nester, 1982; Figure 2). This plasmid was mutagenized with Tn*3*-HoHo1, and the 23 pVCK233::Tn*3*-HoHo1 derivatives (Figure 2) were introduced into A136 and A348. Twelve of these derivatives carry Tn*3*-HoHo1 within seven of the eight pTiA6 T-DNA genes with the *lac*Z coding sequences oriented in the direction of transcription of these genes (Willmitzer *et al.*, 1982; Barker *et al.*, 1983; Gielen *et al.*, 1984). Each of the transconjugant strains was grown under a variety of different conditions and the β-galactosidase activity units were determined. In minimal media, activities varied between 3 and 15 units for the different strains. These low values suggest that the expression in *Agrobacterium* of the fused T-DNA sequences is negligible or low. Also, the differences in β-galactosidase activity seen between the different strains appeared to be unrelated to the position or orientation of the Tn*3*-HoHo1 insertions with regard to the T-DNA genes. A136(pVCK233::Tn*3*-HoHo1) strains and their respective A348 strains exhibited identical activities, indicating that pTiA6 probably does not carry a repressor of T-DNA transcription.

No growth conditions were found that caused a consistent and marked (>2-fold) increase or decrease in the levels of β-galactosidase activity of each A348 (pVCK233::Tn*3*-HoHo1) strain. Growth conditions included Mg/L medium and minimal medium, and these media supplemented separately with octopine, arginine, indole-acetic acids, trans-zeatin or tobacco cell wall extract. Bacteria from mesophyll protoplast co-cultivations also failed to exhibit increased expression. To date, we have been unable to find any evidence of regulated expression of T-DNA sequences within *Agrobacterium* using T-DNA::*lac* fusions generated with Tn*3*-HoHo1.

Conclusion

We describe the construction of a Tn*3*-*lac* transposon, Tn*3*-HoHo1, and its application in the random generation of *lac* fusions useful in the study of gene organization and expression. In such fusions, the production of β-galactosidase is placed under the control of sequences into which Tn*3*-HoHo1 has transposed. This element, like other transposon-*lac* constructs, is useful in the determination of the location and transcriptional orientation of genes, and in the analysis of factors affecting gene expression. Tn*3*-HoHo1 has a number of properties which make it novel. It is derived from Tn*3*, and thus transposes with high specificity into plasmid DNA sequences. It does not carry a functional transposase and cannot self-transpose. Transposition only occurs if transposase is supplied in *trans* by a plasmid that encodes functional transposase; Tn*3*-HoHo1 insertions are stable in the absence of this plasmid.

The *lac*Z coding sequences within Tn*3*-HoHo1 are open and in-frame to the end of the Tn*3* left terminal repeat, and both transcriptional and translational gene fusions can be generated with this element. When Tn*3*-HoHo1 inserts into a gene so that the coding sequences of the gene are in-frame to the *lac*Z coding sequences, expression of the gene will result in the production of a hybrid β-galactosidase protein. If insertion occurs so that the coding sequences are not in-frame, gene expression will result in the production of a non-hybrid β-galactosidase protein that initiates translation within the transposon *lac*Z sequences. We have also constructed a second Tn*3*-*lac*Z, Tn*3*-HoHo2, that can be used in the generation of transcriptional, but not translational fusions, because several stop codons occur between IR_L and the initial ATG of the *lac*Z gene (S.Stachel, unpublished results).

Tn*3*-HoHo1 was used to study the organization and expression of the octopine catabolism, virulence, and T-DNA genetic regions of the A6 Ti plasmid in *A. tumefaciens*. This element can also be applied to studies of gene expression in a wide variety of other prokaryotic organisms, and has recently been used in *Rhizobium sesbania* (M.Holsters, personal communication). Furthermore,

because Tn*3*-HoHo1 contains a eukaryotic polyadenylation sequence, this element might also be applicable to expression studies in eukaryotic organisms.

Materials and methods

Reagents

Restriction and DNA modification enzymes were purchased from New England Biolabs and Bethesda Research Laboratories. X-Gal was from Bethesda Research Laboratories. ONPC, indole-3-acetic acid, and trans-zeatin were from Sigma. [^{35}S]Methionine (1000 Ci/mmol) was purchased from New England Nuclear. L-Octopine was the generous gift of Robert L.Jensen.

Bacterial strains and media

E. coli strains are HB101 : *rec*A$^-$ *hsr*$^-$ *hsm*$^-$ (Boyer and Roulland-Dussoix, 1969); MC1061 : (*lac*IPOZY) (*ara, leu*) *gal*U$^-$ *gal*K$^-$ *hsr*$^-$ *hsm*$^+$ (Casadaban and Cohen, 1980), and SF800 : Nal^R *pol*A$^-$ (Heffron *et al.*, 1977). *E. coli* were grown in LB liquid medium and on LB agar (Miller, 1972), at 37°C, unless otherwise specified. Carbenicillin, kanamycin, chloramphenicol and nalidixic acid were used at 200, 100, 100 and 60 μg/ml, respectively. *Agrobacterium* strains are A136 : C58C1 Nal^R Rif^R, and A348 : A136 carrying pTiA6 (Garfinkel *et al.*, 1981). Agrobacteria were grown on AB minimal agar, and in AB/glucose and in Mg/L liquid media (Chilton *et al.*, 1974), at 28°C unless otherwise specified. Carbenicillin and kanamycin were used at 100 μg/ml.

Procedures

Plasmid DNA transformation was by the method of Cohen *et al.* (1972), and triparental matings were performed by the method of Ditta *et al.* (1980). Transposon insertions into target DNA sequences were selected essentially by the transconjugant procedure of Heffron *et al.* (1977). In this scheme, sequences to be mutagenized were first cloned into pVCK102 (Knauf and Nester, 1982), to yield the target plasmid. pVCK102 is a wide-host range cosmid vector that carries kanamycin and tetracycline resistance genes, can be mobilized for conjugal transfer by the helper plasmid, pRK2013 (Figurski and Helinski, 1979), and carries a *pol*A-independent origin of replication. The target plasmid was transformed into the transposon donor strain, HB101 (pHoHo1, pSShe) and the resultant strain was mated with HB101(pRK2013) and the (Nal^R)*pol*A recipient strain, SF800. The mating mixture was plated on LB agar containing Nal, Cb and Km to select for SF800 (target plasmid::transposon) transconjugants. Replication of pHoHo1, pSShe and pRK2013 is strictly dependent on DNA polymerase I, the *pol*A gene product. Plasmid DNA was isolated from single transconjugants by the method of Birnboim and Doly (1979), and the position and orientation of the transposon within the target plasmid was mapped. Single target plasmid::transposon isolates were transferred by conjugation into *A. tumefaciens*, and *Agrobacterium* transconjugants were selected on BA agar containing Cb and Km, and 0.1% X-gal. X-Gal was included to identify insertions into actively transcribed regions.

Standard recombinant DNA (Maniatis *et al.*, 1982) and Maxam and Gilbert (1980) sequencing procedures have been described. β-Galactosidase activity was determined essentially as described by Miller (1972). Bacteria were collected by centrifugation, and suspended in 600 μl of Z-buffer. Cell density was determined by measuring the absorbance of 100 μl of cells at 600 nm (1 cm), and the remaining bacteria were vortexed with 20 μl 0.05% SDS and 20 μl $CHCl_3$ for 10 s and then incubated for 10 min at 28°C. 100 μl of ONPG (4 mg/ml) was added to start the assay reaction. Reactions were then incubated at 28°C and terminated by adding 250 μl of 1 M Na_2CO_3. Bacteria were removed by centrifugation and the absorbance of the solution at 420 nm (1 cm) was determined. Specific units of β-galactosidase activity were calculated using the formula:

$$\frac{A_{420} \times 10^3}{\text{t(min)} \times A_{600}}$$

Mesophyll protoplasts used for co-cultivation were prepared from *N. tabacum* Xanthi by the method of Wullems *et al.* (1981). 2 ml of cells at a density of 10^5/ml were regenerated in 60 mm dishes for 72 h and exponentially growing bacteria were then added to a density of 5 x 10^7/ml. The co-cultivation mixture was incubated for 16 h at 28°C. During this period, initial protoplast cell division occurred. Bacteria were subsequently separated from the plant cells by passing the mixture through Miracloth® (Calbiochem), or by differential centrifugation, and the units of β-galactosidase activity were determined as described above.

Radiolabelled Agrobacteria were prepared from co-cultivation grown in the presence of 14 μCi(ml [^{35}S]methionine. Following co-cultivation the bacteria were washed in 10 mM Tris, 1 mM EDTA, pH 7.4, and extracts were made by sonication. *E. coli* were radiolabelled overnight in sulfate-free medium containing 10 μCi/ml [^{35}S]methionine. SDS-polyacrylamide gel electrophoresis was carried out according to Laemmli (1970). Fixed gels were dried onto Whatman 3MM paper before exposure to XRP-1 X-ray film.

Acknowledgements

We thank Alice Montoya, Corty Thiens and Brian Watson for their expert technical assistance; Drs Fred Heffron (Scripps Clinic, La Jolla, CA, USA), Roy Parker (UCSF, San Francisco, CA, USA) and Milton Gordon for their suggestions and advice; Dr Patricia Zambryski and Martine De Cock (University of Gent, Belgium) and Dr Paul Allenza and Sharon Bradley for their invaluable advice and help in preparation of this manuscript. This work was funded in part by Grant R01 GM32618 from the NIH. GA was a recipient of a Centennial Fellowship of the Medical Research Council of Canada.

References

An,G., Hidaka,K. and Siminovitch,L. (1982) *Mol. Cell. Biol.*, **2**, 1628-1632.
An,G., Watson,B.D., Stachel,S., Gordon,M.P. and Nester,E.W. (1985) *EMBO J.*, **4**, in press.
Barker,R.F., Idler,K.B., Thompson,D.V. and Kemp,J.D. (1983) *Plant Mol. Biol.*, **2**, 335-350.
Bassford,P., Beckwith,J., Berman,M., Brickman,E., Casadaban,M., Guarente,L., Saint-Giron,I., Sarthy,A., Schwartz,M., Shuman,H. and Silhavy,T. (1978) in Miller,J.H. and Reznikoff,W.S. (eds.), *The Operon*, Cold Spring Harbor Laboratory Press, NY, pp. 245-262.
Bingham,P.M., Levis,R. and Rubin,G.M. (1981) *Cell*, **25**, 693-704.
Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513-1523.
Boyer,H.W. and Roulland-Dussoix,D. (1969) *J. Mol. Biol.*, **4**, 459-472.
Caplan,A., Herrera-Estrella,L., Inzé,D., Van Haute,E., Van Montagu,M., Schell, J. and Zambryski,P. (1983) *Science (Wash.)*, **222**, 815-821.
Casadaban,M.J. and Chou,J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 535-539.
Casadaban,M.J., Chou,J. and Cohen,S.N. (1982) *Cell*, **28**, 345-354.
Chang,A.C.Y. and Cohen,S.N. (1978) *J. Bacteriol.*, **134**, 1141-1156.
Chilton,M.-D., Currier,T.C., Farrand,S.K., Bendich,A.J., Gordon,M.P. and Nester,E.W. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 3672-3676.
Cohen,S., Chang,A.C.Y. and Hsu,L. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 2110-2114.
Copeland,N., Jenkins,N. and Lee,B. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 247-249.
De Greve,H., Decraemer,H., Seurinck,J., Van Montagu,M. and Schell,J. (1981) *Plasmid*, **6**, 235-248.
De Vos,G., De Beuckeleer,M., Van Montagu,M. and Schell,J. (1981) *Plasmid*, **6**, 249-253.
Ditta,G., Stanfield,S., Corbin,D. and Helinski,D.R. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 7347-7351.
Dixon,R., Eady,R., Epin,G., Hill,S., Iaccarino,M., Kahn,D. and Merrick,M. (1980) *Nature*, **286**, 128-132.
Figurski,D.H. and Helinski,D.R. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1648-1652.
Franklin,N.C. (1978) *Annu. Rev. Genet.*, **12**, 193-221.
Garfinkel,D.J. and Nester,E.W. (1980) *J. Bacteriol.*, **144**, 732-743.
Garfinkel,D.J., Simpson,R.B., Ream,L.W., White,F.F., Gordon,M.P. and Nester,E.W. (1981) *Cell*, **27**, 143-153.
Gelvin,S.B., Gordon,M.P., Nester,E.W. and Aronson,A.I. (1981) *Plasmid*, **6**, 17-29.
Gielen,J., De Beuckeleer,M., Seurinck,J., Deboeck,F., De Greve,H., Lemmers, M., Van Montagu,M. and Schell,J. (1984) *EMBO J.*, **3**, 835-846.
Hall,M.N. and Silhavy,T.J. (1981) *J. Mol. Biol.*, **146**, 23-43.
Heffron,F., Bedinger,P., Champoux,J. and Falkow,S. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 702-706.
Heffron,F., McCarthy,B., Ohtsubo,H. and Ohtsubo,E. (1979) *Cell*, **18**, 1153-1163.
Hille,J., Van Kan,J. and Schilperoort,R. (1984) *J. Bacteriol.*, **158**, 754-756.
Hooykaas,P.J.J., Roobol,C. and Schilperoort,R.A. (1979) *J. Gen. Microbiol.*, **110**, 99-109.
Hooykaas,P.J.J., Hofker,M., den Dulk-Ras,H. and Schilperoort,R.A. (1984) *Plasmid*, **11**, 195-205.
Janssens,A., Engler,G., Zambryski,P. and Van Montagu,M. (1984) *Mol. Gen. Genet.*, **195**, 341-350.
Kenyon,C. and Walker,G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 2819-2823.
Kenyon,C., Brent,R., Ptashne,M. and Walker,G.C. (1982) *J. Mol. Biol.*, **160**, 445-457.
Klapwijk,P.M. and Schilperoort,R.A. (1979) *J. Bacteriol.*, **139**, 424-431.
Klapwijk,P.M., Oudshoorn,M. and Schilperoort,R.A. (1977) *J. Gen. Microbiol.*, **102**, 1-11.
Klee,H.J., White,F.F., Iyer,V.N., Gordon,M.P. and Nester,E.W. (1983) *J. Bacteriol.*, **153**, 878-883.
Knauf,V.C. and Nester,E.W. (1982) *Plasmid*, **8**, 45-54.
Kroos,L. and Kaiser,D. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 5816-5820.

Laemmli,U.K. (1970) *Nature,* **227**, 680-685.
Lee,J.S., An,G., Friesen,J.D. and Fiil,N.P. (1981) *Cell,* **25**, 251-258.
Lis,J.T., Simon,J.A. and Sutton,C.A. (1983) *Cell,* **35**, 403-410.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning, A Laboratory Manual,* published by Cold Spring Harbor Laboratory Press, NY.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.,* **65**, 499-559.
Miller,J.H. (1972) *Experiments in Molecular Genetics,* published by Cold Spring Harbor Laboratory Press, NY.
Montoya,A., Chilton,M.D., Gordon,M.P., Sciaky,D. and Nester,E.W. (1977) *J. Bacteriol.,* **129**, 101-107.
Nester,E.W., Gordon,M.P., Amasino,R.M. and Yanofsky,M.F. (1984) *Annu. Rev. Plant Physiol.,* **35**, 387-413.
Ooms,G., Klapwijk,P.M., Poulis,J.A. and Schilperoort,R.A. (1980) *J. Bacteriol.,* **144**, 82-91.
Ruvkun,G.B. and Ausubel,F.M. (1981) *Nature,* **289**, 85-88.
Schröder,G., Waffenschmidt,S., Weiler,E.W. and Schröder,J. (1984) *Eur. J. Biochem.,* **138**, 387-391.
Schuman,H., Silvahy,T. and Beckwith,J. (1980) *J. Biol. Chem.,* **255**, 168-174.
So,M., Gill,R. and Falkow,S. (1975) *Mol. Gen. Genet.,* **142**, 239-249.
Stachel,S., An,G. and Nester,E.W. (1984) *J. Cell. Biochem.,* Suppl. **8B**, 64.
Stern,M.J., Higgins,C.F. and Ames,G.F.-L. (1984) *Mol. Gen. Genet.,* **195**, 219-227.
Thomashow,M.F., Nutter,R., Montoya,A.L., Gordon,M.P. and Nester,E.W. (1980) *Cell,* **19**, 729-739.
Wanner,B., Wider,S. and McSharry,R. (1981) *J. Bacteriol.,* **146**, 93-101.
White,F.F., Klee,H.J. and Nester,E.W. (1983) *J. Bacteriol.,* **153**, 1075-1078.
Willmitzer,L., Simons,G. and Schell,J. (1982) *EMBO J.,* **1**, 139-146.
Wullems,G.J., Molendijk,L., Ooms,G. and Schilperoort,R.A. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 4344-4348.

Received on 7 December 1984; revised on 8 February 1985