Volume 15 Number 20 1987 Nucleic Acids Research

Plant expression signals of the *Agrobacterium* T-cyt gene

B.S.de Pater*, M.P.Klinkhamer, P.A.Amesz, R.J.de Kam, J.Memelink, J.H.C.Hoge and R.A.Schilperoort

Molbas Research Group, Department of Plant Molecular Biology, Leiden University, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands

Received July 17, 1987; Revised and Accepted September 25, 1987

ABSTRACT

Within the 5' and 3' non-coding regions of the T-cyt gene from the octopine T-DNA of Agrobacterium tumefaciens sequences required for expression of this gene in plant cells were identified by deletion mutagenesis. The results show that 184 bp of the 5' non-coding region and 270 bp of the 3' non-coding region are sufficient for wild-type expression. Within the 5' non-coding region two essential expression signals were identified:(1.) an activator element located between -185 and -129 with respect to the ATG start codon and (2.) one out of two TATA boxes. Deletions of the activator element or the two TATA boxes resulted in nonfunctional genes. Deletion of the upstream TATA box and both putative CAAT boxes did not significantly affect expression. Within the 3' non-coding region, the polyadenylation box most distal to the stop codon was not essential for expression, but sequences more upstream, including a second polyadenylation box were found to be required for wild-type expression.

INTRODUCTION

In the 5' non-coding region of animal genes several consensus sequences have been identified, which have a function in transcription: the TATA box (1, 2) , the CAAT box (3) and a variety of other upstream elements (4). In the 3' non-coding region the polyadenylation box (3) and another sequence element (PyGTGTTPyPy) located downstream from this box (5) are required for 3' processing of the mRNA.

Most plant genes that have been transcript mapped contain a sequence, similar to the TATA box at the expected position, about 30 bp upstream from the cap site. It has been demonstrated that the TATA box of the nopaline synthase gene (nos gene) (6) and the "780" gene (7) of Agrobacterium tumefaciens T-DNA are required for expression of these genes in plant cells. All plant promoters examined sofar (6-12) contain additional elements required for maximum expression levels.

Here we present data concerning plant expression signals in the 5' and 3' non-coding regions of the octopine-type cytokinin gene (T-cyt gene), one of the T-DNA onc-genes of Agrobacterium tumefaciens (13). After infection of

plants by this soil microbe the T-DNA is transferred to the plant genome and its genes, including the onc-genes, are expressed leading to the development of a tumor (crown gall). The T-DNA genes, although of bacterial origin have putative plant expression signals (14). The T-cyt gene codes for an enzyme, isopentenyltransferase, catalyzing the first step in cytokinin biosynthesis (15). The effects of deletions in the non-coding regions of the T-cyt gene on its expression in plant cells were determined by studying their influence on both the tumorous response and the transcript level via Northern blot hybridization.

MATERIALS AND METHODS

Bacterial strains and plasmids

Bacterial strains are listed in Table 1 and plasmids in Table 2.

Microbiological methods

Bacteria were grown in LC-medium (26) at 29°C (*Agrobacterium*) or 37°C (*E. coli*). Selection was done on LC-medium plates solidified with 1.8 % Difco agar containing antibiotics at the following concentrations (mg/l): rifampicin 20, carbenicillin 100 (5 for *Agrobacterium* Ach5-chromosomal background), kanamycin 50, streptomycin 50, tetracyclin 10 (2 for *Agrobacterium*). Conjugations were performed on membrane filters as described by Hooykaas *et al*. (26).

Recombinant DNA techniques

Plasmid DNA was isolated by the procedure of Birnboim and Doly (27) for small scale isolations and by the procedure of Clewell and Helinski (28) for large scale isolations. Restriction enzyme digestions, gel electropho-

Table 1. *Bacterial strains*

E.coli	Characteristics	References
JM 101	Δ(*lac*,*pro*), Sup E, *thi*, {F' *traD*36, *proAB*, *lacI*^{q}, zΔM15}	16
MH-1	*lacX*74, *araD* 139, *gal* U, *galK*, *strA*, *hsr*$^-$, *hsm*$^+$.	17
KMBL 1224	*str*^{r} Δ(*proXIII*, *lacXIII*), *thi*209, *Su* II$^+$, *Sup* E,	P. van der Putte, dept. of Molecular Genetics, Leiden.
A. tumefaciens		
LBA 288	*rif*^{r}	18
LBA 4210	pAL 228, pTiAch5 with Tn 904 in T-cyt gene.	19
LBA 1900	pAL 1900, pAL 969 with substitution of T-cyt gene by part of pBR 322.	this work

Table 2. Plasmids

Plasmids	Characteristics	References
pUC 8/9	Ap^r	20
pBR 322	Ap^r Tc^r	21
pAGS 113	Tc^r, artificial T-DNA	22
pRK 2013	Km^r	23
pAL 969	Km^r, R772::pTiAch5 cointegrate	24
pRAL 3910	Ap^r, Em^r, ocs	25
MOGEN 8	Ap^r, NPT II	B. Dekker (MOGEN BV, Leiden)

resis, ligations and transformations were done essentially according to Maniatis et al. (29). Enzymes were purchased from Anglian Biotechnology and Boehringer Mannheim.

Construction of deletions in the 5' and 3' non-coding regions

A 1349 bp RsaI fragment, located between positions 8487 and 9837 on the nucleotide sequence map of the octopine TL-DNA (14) was used for the studies reported here. This fragment contains the 720 bp coding region of the T-cyt gene, 283 bp of the region upstream from the translational start codon and 346 bp of the region downstream from the translational stop codon. The sequence information on this DNA fragment was found to be sufficient for wild-type expression of the T-cyt gene in plant cells.

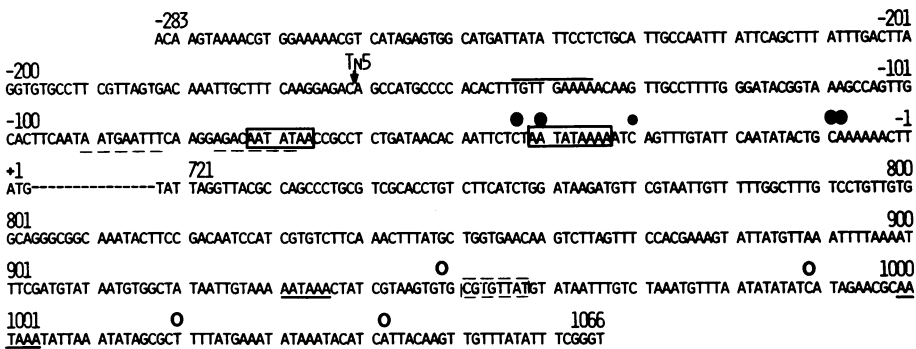

Fig.1. Nucleotide sequence of non-coding regions of the T-cyt gene. The 5'non-coding region (-283bp /-1bp) and 3'non-coding region (721-1066) present on the RsaI fragment in pRAL 3401 are shown. Putative expression signals are indicated: TATA boxes are boxed, CAAT boxes are underlined with broken lines, homology to enhancer core sequence of SV40 is overlined, polyadenylation boxes are underlined with solid lines and homology to the PyGTGTTPyPy consensus sequence is boxed with broken lines. The 5'ends of mRNA, as determined by primer extension (de Pater et al., accompanying paper), are indicated with solid dots; 3'ends of mRNA, as determined by S_1-nuclease mapping (35), are indicated with open dots. The position of Tn5 in a mutant, which shows wild-type expression (35), is shown.

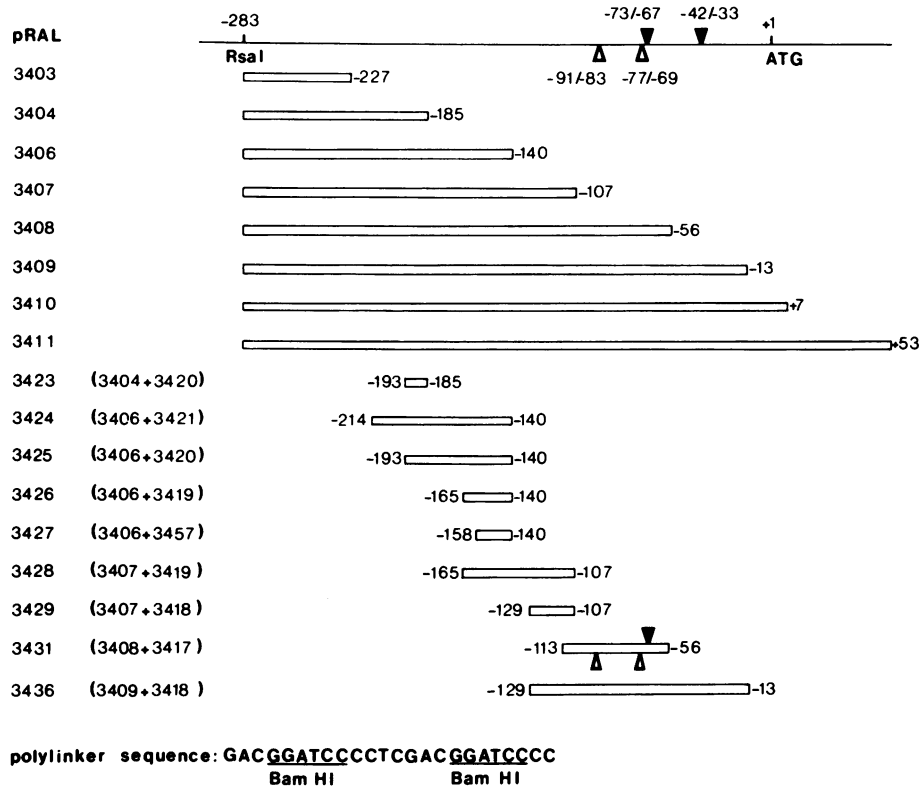

Fig.2. 5' deletion mutants of the T-cyt gene.
The upper line represents the 5' non-coding region of the T-cyt gene. Consensus sequences are indicated (▼ TATA box, Δ CAAT Box). The numbers refer to their positions with respect to the ATG start codon. Open bars represent Bal 31 deletions. The plasmids used for construction of pRAL 3423-3436 are given in parenthesis. Deleted sequences in pRAL 3423-3436 are replaced by a 24 bp sequence, derived from the polylinker of pUC 8/9. This sequence is shown at the bottom of the figure.

Nucleotide sequences of the non-coding regions and putative expression signals are given in Fig. 1.

The RsaI fragment was subcloned in the SmaI site of pUC 8, in the orientation in which the 5' non-coding region faced to the HindIII site (pRAL 3401). Bal 31 deletions in the 5' non-coding region were made after digestion of pRAL 3401 with HindIII. Klenow→blunt/EcoRI fragments, containing the T-cyt gene, were cloned in SmaI-EcoRI digested pUC 8.

Internal 5' deletions were constructed as follows. The BamHI fragment of pRAL 3401, containing the 5' non-coding region and part of the coding

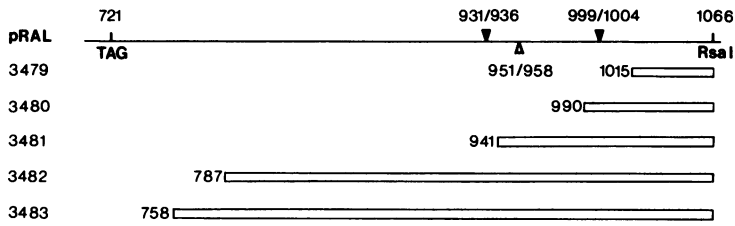

Fig.3. 3' deletion mutants of the T-cyt gene.
The upper line represents the 3' non-coding region of the T-cyt gene. The locations of consensus sequences are shown (▼ polyadenylation box, Δ PyGTGTTPyPy). Open bars represent Bal 31 deletions.

region, was cloned in the BamHI site of pUC 8 (pRAL 3412). After digestion with EcoRI Bal 31 deletions were made, starting in the coding region and ending in the 5' non-coding region. Klenow→blunt/HindIII fragments, containing the 5' end of the T-cyt gene, were cloned in HindII-HindIII digested pUC 9. HindIII-SmaI fragments of the resulting plasmids were cloned in the 5' deletion mutants between the HindIII-SalI(Klenow→blunt) sites. All 5' deletions are shown in Fig. 2. As a result of the cloning strategy the 24 bp polylinker fragment shown in Fig. 2 is present at the site of the 5' internal deletions.

The construction of 3' deletions was performed in a similar way as the construction of the 5'deletions. pRAL 3401 was digested with EcoRI followed by Bal 31 treatment. Klenow→blunt/HindIII fragments were cloned in SmaI-HindIII digested pUC 9. The resulting 3' deletion mutants are shown in Fig. 3.

Bal 31 deletions and double stranded DNA sequencing

Fifty micrograms of linearized plasmid DNA were incubated with 2.5 units Bal 31 (Biolabs) in 200 µl of 20 mM Tris/HCl pH=8.0, 12 mM $MgCl_2$, 12 mM $CaCl_2$, 300 mM NaCl, 1 mM EDTA at 30°C. After 3=12 minutes, samples were taken and the reaction was stopped by adding 1 volume of 50 mM EDTA pH=8.0 of 0°C followed by phenol/chloroform extraction. After alcohol precipitation, blunt ends were generated by incubation with DNA polymerase I (Klenow fragment). The lengths of the deletions were estimated by gel electrophoresis and samples with proper deletions were used for recloning. The precise lengths of the deletions were determined by dideoxy sequencing as follows. Half a microgram of pUC 8/9 derived plasmid was digested with BglI to decrease the length of the DNA (two recognition sites in pUC 8/9). Twenty ng M13-primer (15-mer, Biolabs) or T-cyt-primer (17-mer,

5'-CCTGTGGCAAGTTGGACC-3') were annealed to the plasmid DNA by heating at 100°C for 3 min in 12 µl of 7 mM Tris/HCl pH=7.6, 7 mM $MgCl_2$, 50 mM β-mercaptoethanol, 0.1 mM EDTA (reaction buffer) followed by direct transfer to 0°C. Seven pmoles α^{32}P-dCTP (± 3000 Ci/mmol) and 2 units DNA polymerase I (Klenow fragment) were added to the mixture. Three µl of annealed plasmid/primer were added to 2 µl of A, T, G or C mix. A-mix: 0.01 mM dATP, 0.12 mM ddATP, 0.1 mM dGTP, 0.1 mM dTTP, in reaction buffer; G-mix:0.01 mM dGTP, 0.12 mM ddGTP, 0.1 mM dATP, 0.1 mM dTTP in reaction buffer; T-mix: 0.01 mM dTTP, 0.12 mM ddTTP, 0.1 mM dATP, 0.1 mM dGTP in reaction buffer; C-mix: 2.5 µM dCTP, 0.12 mM ddCTP, 0.1 mM dGTP, 0.1 mM dTTP, 0.1 mM dATP in reaction buffer. After 15 min. at roomtemperature 1 µl of dNTP-chase was added (0.25 mM of each dNTP) and the reaction was continued for 15 minutes. Subsequently the volume was reduced by evaporation. Four µl loading buffer was added (96 % formamide, 0.3 % xylene cyanol, 0.3 % bromophenol blue) and 1-2 µl volumes were analysed on a 6 % sequencing gel.

Tumor induction

Stems of tomato (cv. Moneymaker), Kalanchoe daigremontiana, Nicotiana glauca and axenic tobacco plants (cv. Petit Havanna, SR1) were wounded and infected with a sterile injection needle that had been dipped in a colony of the strains to be tested. Agrobacterium strains LBA 4210 and LBA 1900 which lack an active T-cyt gene do not induce tumors on tomato and induce small tumors with many roots on Kalanchoe daigremontiana, undifferentiated tumors on Nicotiana glauca and tumors carrying few roots on tobacco. Complementation of these strains in trans or in cis with an active T-cyt gene results in tumors with some roots on Kalanchoe daigremontiana and undifferentiated tumors on tomato, Nicotiana glauca and tobacco. Tomato and Kalanchoe daigremontiana were used for determination of phenotypic T-cyt gene expression (tumorous respons) as the difference between an active and inactive T-cyt gene is most clear here. Nicotiana glauca and tobacco were used for transcript analysis as tumors are obtained independent of T-cyt gene expression.

Isolation of total RNA

Undifferentiated tumors from Nicotiana glauca were directly used for RNA isolation. Tumors from axenic tobacco plants were subcultured for three weeks on LS medium containing 100 µg/ml cefotaxim (Calbiochem, Hoechst) and 100 µg/ml kanamycin (Sigma) before RNA isolation. This resulted in the

disappearence of the roots originally present on the non-complemented tumors. RNA isolations were done according to van Slogteren et al. (30).

Northern Blot Hybridization

Fifty micrograms of total RNA were glyoxylated for 30 min. at 60°C in 1.0 M deionized glyoxal, 10 mM Na-phosphate pH=6.5, 1 mM EDTA, 5 % Fycoll 400, 0.01 % bromophenolblue. Glyoxylated RNA was electrophoresed on 1.5 % agarose gels in 10 mM Na-phosphate pH=6.5, 1 mM EDTA with constant buffer recirculation. Subsequently RNA was transferred from gels to GeneScreen membranes (New England Nuclear) according to Southern (31) with 50 mM Na-phosphate pH=6.5, 5mM EDTA as blotting buffer. After baking for 2-4 hours at 80°C the blots were treated with 0.1 M NaOH for 30 sec. to remove glyoxal from the RNA and neutralized in 50 mM Na-phosphate pH=6.5, 5 mM EDTA for 30 minutes. The blots were impregnated with 50 % deionized formamide, 5xSSPE (1xSSPE=0.18 M NaCl, 10 mM Na-phosphate pH=6,5, 1 mM EDTA), 5 % SDS at 42°C. (50 $\mu l/cm^2$). After 4-16 hours the nick translated, denatured probe (2-6x10^8 cpm/μg) (32) was added to this mixture. After hybridization for 48-65 hours at 42°C, the blots were washed 2-3 times with 0.1xSSPE, 0.1 % SDS for 15 min. at 65°C (1 ml/cm^2). Transcripts were visualized by exposing the blots 2-20 days to Fuji RX films, using Kyokko intensifying screens. Autoradiograms were scanned on a LKB laser densitometer.

RESULTS

Introduction of T-cyt genes into plant cells

Two methods were used for the introduction of wild-type and mutant T-cyt genes into plant cells: (1) cointegration with a R772::Ti plasmid derivative and (2) a binary vector system (33). For application of the first method a R772::Ti plasmid (pAL 969) derivative was made (pAL 1900), in which the RsaI fragment containing the T-cyt gene was replaced in several steps by the PstI-EcoRI fragment of pBR 322 containing the tetracyclin resistance gene, the origin of replication and part of the ampicillin resistance gene. Cointegrates were made by single cross-over between the pUC 8 part of the T-cyt gene plasmids and the pBR 322 fragment present in pAL 1900, via a triparental conjugation with E. coli strain KMBL 1224 (str^r). Subsequently cointegrates were transferred to a Ti plasmid cured Agrobacterium strain (LBA 288) resulting in LBA 1901-1930 (Table 3). In all cases the structures of cointegrates were verified by Southern blot hybridization (results not shown). In the cointegrates 1 to 4 tandemly

Table 3. Complementation with mutated T-cyt genes introduced into LBA 4210 and LBA 1900.

	pAGS 113 derivatives	LBA 4210 binary			LBA 1900 binary			LBA 1900 cointegrates			
pRAL	pRAL	LBA	K	T	LBA	K	T	LBA	#	K	T
3401	3108	4608	+	+				1901	3	+	+
3403	3109	4609	+	+	1935	+		1903	2	+	+
3404	3111	4611	+	+	1936	+	+	1904	2	+	+
3406	3112	4612	-	-	1937	-	-	1906	2	-	-
3407	3113	4613	-		1938	-		1907	3	-	
3408	3114	4614	-		1939	-		1908	1	-	
3409	3115	4615	-					1909	2	-	
3410	3116	4616	-					1910	2	-	
3411	3117	4617	-								
3423	3119	4619		+				1912	3	+	+
3424	3120	4620			1945	-	-	1913	1	-	-
3425	3121	4621	-	-	1946	-	-	1914	2	-	-
3426	3122	4622	-	-	1947	-	-	1915	2	-	-
3427	3123	4623	-	-				1925	2	-	-
3428	3124	4624	-	-	1949	-	-	1916	3	-	-
3429	3125	4625	+	+				1917	2	+	+
3431	3127	4627	+	+	1952	+	+	1919	1	+	+
3436	3132	4632	-	-	1957	-	-	1924	3	-	-
3479	3133	4633	+					1926	3	+	
3480	3134	4634	+					1927	4	+	+
3481	3135	4635	±					1928	3	+	±
3482								1929	2	-	-
3483	3136	4636	-					1930	2	-	-

Plasmids containing mutated T-cyt genes (pRAL 3401-3483) were cloned in pAGS 113. The resulting plasmids (pRAL 3108-3136) were conjugatively transferred to LBA 4210, resulting in LBA 4608-4636, and to LBA 1900, resulting in LBA 1935-1957. Cointegrates between T-cyt gene plasmids (pRAL 3401-3483) and pAL 1900 were made. The number of T-cyt gene plasmids in the cointegrates is indicated (#). Agrobacteria were used for stem infection of *Kalanchoe daigremontiana* (K) and tomato (T), and expression of T-cyt genes was determined by analysing the phenotypes of the tumors. +: wild-type expression, ±: attenuated expression, -: no expression.

arranged copies of T-cyt gene plasmids were present (Table 3).

For application of the second method, T-cyt gene plasmids were linearized with AccI and cloned in the ClaI site of the wide host range plant vector pAGS 113 (22) with the 5' non-coding region of the T-cyt gene closest to the NPT II gene. Using pRK 2013 as a helper plasmid, pAGS 113 derivatives (pRAL 3108-3136, Table 3) were conjugatively transferred to *Agrobacterium* strains containing a Ti plasmid with an inactivated T-cyt gene, i.e. LBA 4210 (Tn 904 in T-cyt gene) resulting in LBA 4608-4636, or LBA 1900 (T-cyt gene replaced by pBR 322 sequences) resulting in LBA 1935-1957. The plasmids were shown to be present in the agrobacteria by small scale

plasmid isolation, restriction enzyme digestion and Southern blot hydridization (results not shown).

Phenotypic expression of 5' deletion mutants

Kalanchoe daigremontiana and tomato were used for complementation studies with mutated T-cyt genes introduced into Agrobacterium T-cyt gene mutants LBA 4210 and LBA 1900. Introduction of the wild-type T-cyt gene into these strains resulted in wild-type tumors. The results obtained after infection with T-cyt gene deletion mutants (Fig. 2) are summarized in Table 3. A 184 bp fragment of the 5' non-coding region was sufficient for wild-type expression (pRAL 3404). Further reduction of the 5' non-coding region resulted in inactivation of the gene (pRAL 3406-3411). Internal deletions extending into a -184 bp/-130 bp region with respect to the ATG start codon resulted in inactive genes (pRAL 3424-3428). These observations suggest that an activator element is situated between positions -185 and -129. When the two TATA boxes were deleted (pRAL 3436) no phenotypic expression was detected, independent of the presence of the activator element. Deletion of one TATA box and both putative CAAT boxes (pRAL 3431) did not affect expression. This indicated that the CAAT boxes did not have much, if any, influence on expression and that at least one TATA box was required for expression.

Phenotypic expression of 3' deletion mutants

The expression of 3' deletion mutants (Fig. 3) was analysed analogous to the 5' deletion mutants (Table 3). The results show that 270 bp of the 3' non-coding region were sufficient for wild-type expression (pRAL 3479, 3480). This means that the most distal polyadenylation box is not essential. Deletion of a region also including sequences between the two polyadenylation boxes reduced the expression (pRAL 3481). Further deletion of the 3' non-coding region, including the second putative polyadenylation box, abolished phenotypic expression completely (pRAL 3482, 3483).

Transcript analysis of 5'deletion mutants

In order to investigate transcriptional activity of 5' deletion mutants, Nicotiana glauca and axenic tobacco plants were infected with LBA 1900, 1935-1938, 1949 and 1957 (Table 3). For each bacterial strain 36 independent tumors were harvested and pooled in order to reduce any variation in promoter activity of the introduced genes due to different integration sites, methylation etc. Total RNA preparations extracted from pooled Nicotiana glauca tumors, coded CNG 1900-1957, according to the Agrobacter-

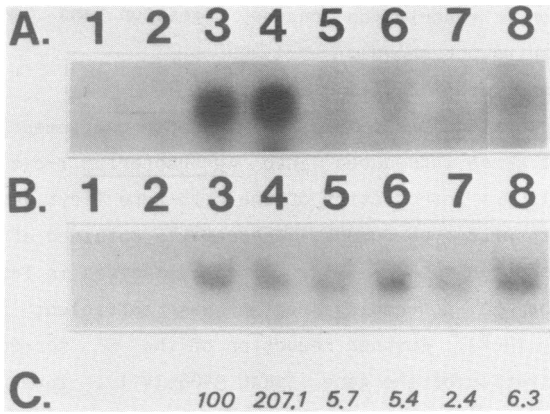

Fig.4. Transcript analysis of 5' deletion mutants.
RNAs from N. glauca tumors, induced by 5' deletion mutants, were hybridized to (A) a T-cyt gene probe (pRAL 3403) and (B) an NPT II gene probe (MOGEN 8). Lane 1: untransformed N.glauca stem tissue, lane 2: CNG 1900, lane 3: CNG 1935 (pRAL 3403), lane 4: CNG 1936 (pRAL 3404), lane 5: CNG 1937 (pRAL 3406), lane 6: CNG 1938 (pRAL 3407), lane 7: CNG 1949 (pRAL 3428), lane 8: CNG 1957 (pRAL 3436). (C) T-cyt transcript levels, relative to NPT II transcript levels, as determined by densitometric scanning are given as a percentage of the relative level found in CNG 1935.

ium strains that were used to induce the tumors, were analysed by Northern blot hybridization. Northern blots were hybridized with a T-cyt gene probe (pRAL 3403) and, after removal of this probe, with an NPT II gene probe (MOGEN 8) (Fig. 4). The NPT II gene in pAGS 113 fused to the nos promoter for constitutive expression in plant cells was used as a reference gene.

No RNA homologous to the T-cyt or NPT II probe was detected in nontransformed stem tissue of N. glauca or in CNG 1900 tissue, which contained only the T-DNA of pAL 1900. (Fig. 4A, B, lanes 1 and 2). Two bands were visible in the other tissues after hybridization with the NPT II probe (Fig. 4B). This is due to the presence of two nos-promoter fragments placed in tandem upstream from the NPT II gene. The level of NPT II transcripts was approximately similar in all RNA preparations, indicating that sufficient tumors had been pooled. In CNG 1935 (pRAL 3403) and 1936 (pRAL 3404) tissues a band, corresponding to transcripts of the T-cyt gene, was seen after hybridization with pRAL 3403 (Fig. 4A, lanes 3 and 4). CNG 1935 and 1936 contained T-cyt genes with small deletions in the promoter region leaving the activator element intact. In CNG 1937 (pRAL 3406), 1938 (pRAL 3407) and 1949 (pRAL 3428), containing T-cyt genes with deletions

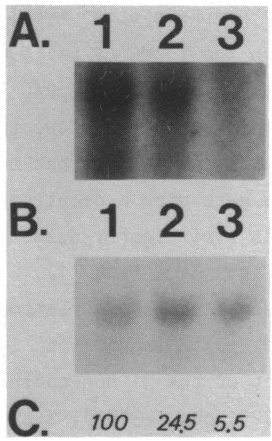

Fig.5. Transcript analysis of 3' deletion mutants.
RNAs from N. glauca tumors, induced by 3' deletion mutants were hybridized to (A) a T-cyt gene probe (pRAL 3403) and (B) an ocs gene probe (pRAL 3910). Lane 1: CNG 1927 (pRAL 3480), lane 2: CNG 1928 (pRAL 3481), lane 3: CNG 1929 (pRAL 3482). (C) T-cyt transcript levels relative to ocs transcript levels, as determined by densitometric scanning are given as a percentage of the relative level found in CNG 1927, corrected for the number of T-cyt genes in the cointegrates.

extending into the activator element, T-cyt transcripts were detected, but only at very low levels (Fig. 4A, lanes 5-7). In the case that the activator region was present but both TATA boxes were deleted, the promoter activity was also strongly reduced (CNG 1957 (pRAL 3436) ; Fig. 4A, lane 8). Similar results were obtained with tobacco tumors (results not shown).

In order to compare T-cyt transcript levels relative to NPT II transcript levels (RTL), the autoradiograms were scanned on a densitometer. The RTL in CNG 1935 was taken as 100 %. The RTL's are listed in Fig. 4C. The RTL in CNG 1936 was 2-fold higher than the RTL in CNG 1935. The RTL's of activator deletion mutants (CNG 1937, 1938 and 1949) and the deletion mutant eliminating both TATA boxes (CNG 1957) were decreased approximately 20-fold.

Transcript analysis of 3' deletion mutants

N. glauca plants were infected with LBA 1927-1929. These strains contain cointegrates of pAL 1900 and pRAL 3480-3482. The copy number of the T-cyt gene plasmids in the cointegrates is respectively 4, 3 and 2, which has been accounted in the determination of transcript levels. For each strain 20 independent tumors were harvested, pooled and the extracted RNAs (CNG

1927-1929) were analysed by Northern blot hybridization. The octopine synthase gene (ocs) present on the T-DNA of pAL 1900 was used as a reference gene. Hybridization with an ocs probe (pRAL 3910) resulted in a band which varies only slightly in intensity when the three lanes in Fig. 5B are compared, indicating that sufficient tumors had been pooled. In CNG 1927 (pRAL 3480), which harbored a gene with a deletion of the distal polyadenylation box, T-cyt transcripts were clearly detectable after hybridization with a T-cyt gene probe (Fig. 5A, lane 1). Deletion of a fragment including the region between the polyadenylation boxes (CNG 1928, pRAL 3481) caused reduction of the T-cyt transcript level (Fig. 5A, lane 2). After deletion of a large part of the 3' non-coding region, including both polyadenylation boxes (CNG 1929, pRAL 3481), T-cyt transcripts were hardly detectable (Fig. 5A, lane 3). The T-cyt transcript levels relative to the ocs transcript levels (RTL) were determined by scanning of the autoradiograms. The RTL's, corrected for the number of T-cyt gene copies in the cointegrates, are given in Fig. 5C. The RTL in CNG 1927 was taken as 100 %. In CNG 1928 the RTL was reduced to 24.5 %. A very low level of T-cyt transcript (5.5 %) was detected in CNG 1929.

DISCUSSION

Sequence elements required for expression in plant cells were identified in the 5' and 3' non-coding regions of the T-cyt gene by making deletions in these regions and determination of their effects on gene expression. Two methods were used for the introduction of mutated T-cyt genes into plant cells: (1) cointegration with a R772::Ti plasmid derivative and (2) a binary vector system. Similar results were obtained using these two methods for phenotypic complementation studies on plants and the results on transcriptional activity were consistent with the phenotypic data.

We have identified two different essential expression signals in the 5' non-coding region (1) an activator element and (2) one out of two TATA boxes. S_1 mapping experiments have demonstrated that each of the two TATA boxes are functional in the wild-type gene (34, 35). Apparently, one TATA box is sufficient for wild-type phenotypic expression. The requirement of a TATA box for plant gene expression has been demonstrated also for other genes (6, 7). The two putative CAAT boxes in the T-cyt gene were found not to be essential for expression. A region upstream from the CAAT boxes however was absolutely required for expression. This activator is situated

between -129 bp (pRAL 3429) and -185 bp (pRAL 3404) and is probably situated downstream from -161 bp, since a Tn5 insertion at this site does not inactivate the gene (35). A 2-fold increase of the T-cyt transcript level in CNG 1936, relative to the NPT II transcript level, was observed. Since there is as much as 2-fold variation in reference NPT II transcript levels in different RNA preparations, we are hesitant to conclude that a 2-fold variation is significant. However, it can as yet not be excluded that sequences within the -185 bp/-226 bp region have some influence on T-cyt gene expression.

All plants genes examined sofar (6-12) contain activator elements in the 5' non-coding region. Several activators contain the consensus core sequence of animal enhancers TGTGG$^{TT}_{AA}$AG (36). We found homology with this consensus core sequence also in the activator region of the T-cyt gene (Fig. 1). Up to 240 bp of the 5' non-coding regions of both the octopine and nopaline type T-cyt genes show greater than 88 % homology (37). Interestingly, the largest conserved sequence (52 bp) falls within the activator region we have identified.

The polyadenylation box distal to the stop codon is not required for wild-type phenotypic expression. The region downstream from the proximal polyadenylation box however has an essential function in 3' processing of the mRNA, as the four sites for polyadenylation of T-cyt mRNA are all within this region (35) (Fig. 1). Deletions of this region reduced phenotypic T-cyt expression in complementation studies indeed. The size of T-cyt transcripts was not affected but a 4-fold reduction of relative T-cyt transcript level was observed. The region between the polyadenylation boxes shows homology to a sequence (PyGTGTTPyPy) of which the significance in animal gene expression has been demonstrated (5). The polyadenylation box proximal to the stop codon and perhaps other sequences deleted in pRAL 3482 are required for 3' processing or stability of mRNA.

Identification of cis-acting elements involved in transcription of the T-cyt gene further attributes to the knowledge of plant gene expression. The mechanisms by which these cis-acting elements conduct gene expression have yet to be established.

ACKNOWLEDGEMENTS

We thank Dr. A.Ledeboer (Unilever Research, Vlaardingen) for providing the T-cyt primer. This work was supported by a grant from the Rockefeller Foundation and by the Foundation for Fundamental Biological Research (BION)

which is subsidized by the Netherlands Organization for the Advancement of Pure Research (ZWO).

*To whom correspondence should be addressed

REFERENCES

1. Cordon, J., Walsylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. and Chambon, P. (1980). Science **209**, 1406-1414.
2. Breathnach, R. and Chambon, P. (1981). Ann. Rev. Biochem. **50**, 349-383.
3. Benoist, C., O'Hara, K.,Breathnach, R. and Chambon, P.(1980) Nucl. Acids Res. **8**, 127-142.
4. Dynan, W.S. and Tjian, R. (1985). Nature **316**, 774-778.
5. Mclauchlan, J., Gaffney, D., Whitton, J.L. and Clements, J.B. (1985). Nucl. Acids Res. **13**, 1347-1368.
6. An, G., Ebert, P.R., Yi, B.-Y. and Choi, C.-H. (1986). Mol. Gen. Genet. **203**, 245-250.
7. Bruce, W.B. and Gurley, W.B. (1987). Mol. Cell. Biol. **7**, 59-67.
8. Shaw, C.H., Carter, G.H., Watson, M.D. and Shaw, C.H. (1984). Nucl. Acids Res. **12**, 7831-7846.
9. Morelli, G., Nagy, F., Fraley, R.T., Rogers, S.G. and Chua, N.-H. (1985). Nature **315**, 200-204.
10. Simpson, J., Timko, M.P., Cashmore, A.R., Schell, J., Van Montagu, M. and Herrera-Estrella, L. (1985). EMBO J. **4**, 2723-2729.
11. Chen, Z.-L., Schuler, M.A. and Beachy, R.N. (1986). Proc. Natl. Acad. Sci. USA **83**, 8560-8564.
12. Kaulen, H., Schell, J. and Kreuzaler, F. (1986). EMBO J. **5**, 1-8.
13. Hooykaas, P.J.J. and Schilperoort, R.A. (1984). In: Scandalios, J.G. (Ed), Molecular genetics of plants. Advances in Genetics 22. Academic Press Inc. Orlando Floria, USA, pp 209-283.
14. Barker, R.F., Idler, K.B., Thompson, D.V. and Kemp, J.D. (1983). Plant Mol. Biol. **2**, 335-350.
15. Buchmann, I., Marner, F.-J., Schröder, G., Waffenschmidt, S. and Schröder, J. (1985). EMBO J. **4**, 853-859.
16. Messing, J. (1979). Recombinant DNA Technical Bulletin, NIH Publication No. 79-99 **2**, 43-48.
17. Goddard, J.M., Caput, D., Williams,S.R. and Martin Jr., D.W.(1983). Proc. Natl. Acad. Sci. USA **80**, 4281-4285.
18. Hooykaas, P.J.J.,den Dulk-Ras, H., Ooms, G. and Schilperoort, R.A. (1980). J. Bacteriol. **143**, 1295-1306.
19. Ooms, G., Klapwijk, P.M., Poulis, J.A. and Schilperoort, R.A. (1980). J. Bacteriol. **144**, 82-91.
20. Vieira, J. and Messing,J. (1982). Gene **19**, 259-268.
21. Sutcliffe, J.G. (1978). Cold Spring Harbor Symp. **43**, 77-90.
22. Van den Elzen, P., Lee, K.Y., Townsend, J. and Bedbrook, J. (1985). Plant Mol. Biol. **5**, 149-154.
23. Figurski, D.H. and Helinski, D.R. (1979). Proc. Natl. Acad. Sci. USA **76**, 1648-1652.
24. Hooykaas, P.J.J., den Dulk-Ras, H. and Schilperoort, R.A. (1980). Plasmid **4**, 64-74.
25. Hoekema, A., van Haaren, M.J.J., Fellinger, A.J., Hooykaas, P.J.J. and Schilperoort, R.A. (1985). Plant Mol. Biol. **5**, 85-89.
26. Hooykaas, P.J.J., Klapwijk, P.M., Nuti, M.P., Schilperoort, R.A. and Rörsch, A. (1977). J. Gen. Microbiol. **98**, 477-484.

27. Birnboim, H.C. and Doly, J. (1979). Nucl. Acids Res. **7**, 1513-1523.
28. Clewell, D.B. and Helinski, D.R. (1969). Proc. Natl. Acad. Sci. USA **62**, 1159-1166.
29. Maniatis, T., Fritsch, E. and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
30. Van Slogteren, G.M.S., Hoge, J.H.C., Hooykaas, P.J.J. and Schilperoort, R.A. (1983). Plant Mol. Biol. **2**, 321-333.
31. Southern, E.M. (1975). J. Mol. Biol. **98**, 503-517.
32. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977). J. Mol. Biol. **113**, 237-251.
33. Hoekema, A., Hirsch, P.R., Hooykaas, P.J.J. and Schilperoort, R.A. (1983). Nature **303**, 179-180.
34. Heidekamp, F., Dirkse, W.G., Hille, J. and van Ormondt, H. (1983) Nucl. Acids Res. **11**, 6211-6223.
35. Lichtenstein, C., Klee, H., Montoya, A., Garfinkel, D., Fuller, S., Flores, C., Nester, E. and Gordon, M. (1984). J. Mol. Appl. Genet. **2**, 354-362.
36. Sassone-Corsi, P. and Borrelli, E. (1986). Trends Genet. **2**, 215-219.
37. Goldberg, S.B., Flick, J.S. and Rogers, S.G., (1984). Nucl. Acids Res. **12**, 4665-4677.