Characterisation of LTR sequences involved in the protoplast specific expression of the tobacco Tnt1 retrotransposon

Josep M.Casacuberta and Marie-Angèle Grandbastien

Laboratoire de Biologie Cellulaire, INRA, Centre de Versailles, 78026 Versailles Cedex, France

Received February 22, 1993; Revised and Accepted April 5, 1993

ABSTRACT

The tobacco Tnt1 retrotransposon is the only plant retrotransposon that has been shown to be transcriptionally active, and its transcription is strongly induced when preparing leaf-derived protoplasts. We have analysed in this paper the LTR sequences important for Tnt1 expression in tobacco protoplasts. We show that LTR sequences upstream of the TATA box are sufficient to confer protoplast-dependent induction to a heterologous promoter. We also show that this region contains two short activator elements, and that one of these sequences, BII, interacts with protoplast-specific nuclear factors.

INTRODUCTION

Retrotransposons are a family of mobile genetic elements, structurally and functionally related to retroviruses, that transpose via an RNA intermediate (1). As in retroviruses, the coding region of retrotransposons is flanked by long direct repeated sequences (LTRs) that are necessary for both transcription and transposition. Transcription of retroviral-like elements begins at the transcription start site in the 5' LTR and ends at the 3' LTR polyadenylation site, dividing the LTR into three zones: the U3 region, containing LTR sequences up to the transcription start, the R region, containing LTR sequences from the transcription start up to the polyadenylation site, and the U5 region, containing LTR sequences downstream of the polyadenylation site.

Retrotransposons represent the major class of transposable elements in animal genomes, and have recently been shown to be also ubiquitous components of plant genomes (2, 3). Several plant retrotransposons have been described to date (reviewed in 4) but most of them seem to be inactive. Direct evidence of activity has been provided for only two plant retrotransposons, Bsl from maize (5) and Tnt1 from tobacco (6), which were isolated after insertional mutagenesis. The tobacco Tnt1 retrotransposon, a member of the Ty1-*copia* family of retroelements, remains, however, the only plant retrotransposon that has been shown to be transcriptionally active, and thus the only one for which transcriptional regulation can be studied.

Transcription of retrotransposons is usually regulated by different stimuli, and this is also the case within the Ty1-*copia* family of retroelements. *Drosophila copia* retrotransposon expression is under developmental control (7), and the steady state level of Ty1 RNA depends on the mating type (8). Nevertheless, both Ty1 and *copia* are strongly expressed in most of the tissues or conditions tested, their RNA being, in most cases, the major polyadenylated RNA in the cell. In contrast, Tnt1 transcription is strongly regulated, which may indicate that control of transcription plays a primary role in the control of transposition of this element.

Tnt1 expression is low in adult roots and almost absent in leaves, but is strongly induced in leaf-derived protoplasts (9). This induction was shown to be associated with the presence of cell wall hydrolases (9). Tnt1 transcription is initiated in the 5' LTR and a major transcriptional start site, as well as a minor transcriptional start site, have been determined (9). Furthermore, transgenic plants transformed with translational fusions between the 5' LTR and the β -glucuronidase reporter gene (10) (LTR-GUS), showed a pattern of GUS activity similar to the pattern found for Tnt1 expression (9), demonstrating that LTR sequences are sufficient to mediate protoplast-specific expression.

In this paper we describe the analysis of the transcriptional control regions present in the LTR of Tnt1. We show that the sequences upstream of the TATA box, within the U3 region of the element, are sufficient to mediate induction during protoplast preparation. This region, composed of different activator elements, contains a short repeated sequence that interacts with a protoplast-specific nuclear factor(s), and is a good candidate to explain the induction of Tnt1 expression associated with the presence of cell wall hydrolases.

MATERIALS AND METHODS

Plant material and protoplast isolation

Plants of *Nicotiana tabacum* (cv Xanthi) wild type line XHFD8 (11) were grown in a greenhouse under low light conditions. Protoplasts were isolated from fully expanded young leaves according to Chupeau *et al.* (12). Cell wall enzymatic digestions

* To whom correspondence should be addressed

were done overnight (16h) at 22°C in To medium (13). The standard concentrations of cell wall hydrolases were: Macerozyme R10 (Yakult Biochemicals Co., Ltd) 200 mg/l; Onozuka R10 (same origin), 1000 mg/l, and Driselase (Kyowa Hakko Kogyo Co., Ltd) 500 mg/l.

Constructs

The 2120 plasmid (pBMCV 102120) (Fig 1) containing a translational LTR-GUS fusion (containing 110 bp of tobacco nitrate reductase genomic sequences, the entire Tnt1 5' LTR, and the first 160 bp of the internal Tnt1 sequences) in a pBluescript KS plasmid has been previously described (9). The 2146 plasmid (pBMCV 102146), in which the LTR upstream sequences have been deleted, was obtained by deletion of the 300 bp HindIII fragment of the 2120 construct (from the HindIII site found in the polylinker to the HindIII site located at nucleotide 302 of the LTR sequence) (Pauls et al, in preparation). The 7011 construct (pBMCV 107011), that lacks the central LTR region, was made by cloning the 300 bp HindIII fragment of the 2120 plasmid into a 2120-derived plasmid from which the 550 bp EcoRV fragment (from the polylinker to the EcoRV site located at nucleotide 432 of the LTR sequence) had been deleted. Orientations of reinserted fragments were checked by sequencing. The 7031 plasmid (pBMCV 107031), lacking the 3' final region of the LTR, was made by cloning the 550 bp EcoRV fragment of the 2120 plasmid into a 2120-derived plasmid from which the 700 bp EcoRV-SspI fragment (from the polylinker to the SspI site located at nucleotide 581 of the LTR sequence) had been removed. Orientations of reinserted fragments were checked by sequencing.

The EcoRI-BglII fragment of the EMBL-GUS46 plasmid (14), containing the minimal (-46, +8) CaMV 35S promoter fused to the reporter gene GUS and the 3C terminator from the pea *rbcs* SE9 gene, was subcloned in the EcoRI-BamHI sites of a pUC18 plasmid, to give the construct pG46. The 300 bp HindIII fragment of the 2120 plasmid was cloned into the XbaI site of pG46, after filling-in with Klenow, to give the 4691 plasmid (pBMCV 104691) (Fig 2), in which the LTR sequence is in the same orientation as the 35S promoter, and the 4692 plasmid (pBMCV 104692) in which this sequence is in reverse orientation. The constructs 4651 and 4652 (pBMCV 104651 and pBMCV 104652) (Fig 3) were obtained by cloning the 5' HindIII-HincII fragment (from the polylinker to the HincII site located at nucleotide 79 of the LTR sequence) in both direct and reverse orientation in the XbaI site of pG46. The 4661 and 4662 plasmids (pBMCV 104661 and pBMCV 104662) were obtained by cloning the HincII-HindIII fragment of the 2120 construct (from the HincII site located at nucleotide 79 to the HindIII site located at nucleotide 302 of the LTR sequence) in both direct and reverse orientations, in the XbaI site of pG46.

The same LTR fragments where cloned in the SphI site of a p249 plasmid (Vaucheret, unpublished) containing a truncated (-90, +4) CaMV 35S promoter fused to the gene GUS and the nos terminator, giving the 9091 and 9092 plasmids (pBMCV 109091 and pBMCV 109092) when containing the 300 bp HindIII fragment of the 2120 plasmid in both direct or reverse orientation, the 9051 and 9052 plasmids (pBMCV 109051 and pBMCV 109052) when containing the 5' HindIII-HincII fragment of the 2120 plasmid, and the 9061 and 9062 plasmids (pBMCV 109061 and pBMCV 109062) when containing the 3' HincII-HindIII fragment of the 2120 plasmid.

Synthetic double-stranded oligonucleotides corresponding to boxes BI (5'-GAAGAAGTATTAGGCATGTGCCTAATA-3') or BII (5'-TTTGGTTTGGTAGCCAACCTTGTTGAAT-3') were cloned in tandem, after phosphorylation of only one strand, in SmaI-HincII cleaved pUC18 (dephosphorylated only on one strand). Clones containing two copies of the BII sequence and five copies of the BI sequence, were digested by EcoRI-HindIII, and the internal fragments were cloned in the SphI site of the p249 plasmid, giving the B151 plasmid (pBMCV 100151) containing a tandem of 5 BI sequences in front of the truncated promoter, and the B221 plasmid (pBMCV 100221) containing a tandem of 2 BII sequences in front of the truncated 35S promoter.

Transient GUS expression in electroporated protoplasts

Protoplasts were electroporated with the different plasmids essentially as described by Guerche et al. (15). Ten micrograms of the analysed plasmids were mixed with one microgram of the plasmid pD043 containing the Luciferase reporter gene (LUC) placed under the control of the CaMV 35S promoter (16), in order to compare the GUS expression of each construct to the internal standard supplied by this reporter gene. 50 µg of sonicated salmon sperm DNA were used as carrier DNA. After electroporation, the protoplasts were cultured overnight in To medium, washed twice in washing buffer and lysed in a pH 7 Na₂HPO₄/NaH₂PO₄ (50 mM) buffer, containing 10 mM dithiothreitol (DTT), 1 mM EDTA, 10% (v/v) glycerol (GUS buffer). The fluorometric GUS assays were performed on 50 µg of total protein, using a Fluoroskan apparatus (Labsystems), based on the assay described by Jefferson (10). Luciferase activity, assayed in 50 µl extract made in GUS buffer, was initiated by injecting 150 µl of 25 mM tricine buffer (pH 8.1), 10 mM MgCl₂, 0.2 mM ATP, 0.2 mM luciferin and 0.1 mM Coenzyme A according to Wood (17). Total light emission was measured during the initial 10 sec of reaction using a Luminoskan apparatus (Labsystems).

Transient GUS expression after particle bombardment

1 µg of the plasmid was mixed with 0.2 µg of 35S-Luc control plasmid and coated on 0.2 µm tungsten particles as described by Godon et al. (18). Leaves were bombarded on both sides of the central axis and on both sides of the leaf for each bombarded zone. Leaf discs of the two bombarded zones (2.5 cm) were incubated overnight in To medium, one disc with cell wall degrading enzymes, and the other one without cell wall degrading enzymes. Resulting protoplasts were treated as described above for the electroporated protoplasts. Control leaf discs were ground in liquid N₂ and the powder was extracted for 10 min with GUS buffer on ice. Fluorimetric and luminometric assays were performed as described above.

Gel mobility shift analysis

Gel shift assays were performed according to Grange et al. (19). Blunt-end double-stranded oligonucleotides were generated by the annealing of complementary single-stranded synthetic oligonucleotides, labelled on each strand using T4 polynucleotide kinase, and purified on an acrylamide gel. 10,000 cpm of ds oligonucleotide were incubated 20 min on ice with 20 µg of nuclear extracts in 10 mM Tris pH7.9, 1 mM EDTA, 30 mM KCl, 15% glycerol, 0.1 µg/µl poly dI.dC-dI.dC. When indicated, non-labelled ds oligonucleotides were added as competitors to

the incubation mixture (at the concentrations described) before adding the probe. The 8% acrylamide gel was run for 30 min at 4°C in 0.25×TBE buffer.

RESULTS

Upstream regions of the Tnt1 LTR contain strong regulatory elements

To determine the sequences important for the transcription of Tnt1 in leaf-derived protoplasts, a series of deletions of the LTR of the element fused to the GUS reporter gene was made. In each case, transcription and translation beginning in Tnt1 sequences resulted in a Tnt1:GUS fusion protein. Fig. 1 shows a comparison between the intact LTR-GUS fusion (construct 2120) and a series of LTR deletion constructs. Deletion of an internal HindIII-EcoRV region containing the major transcriptional start site and its putative TATA-box (9), giving the 7011 construct, results in a substantial decrease of the activity. However, the activity of this construct, that contains most of the LTR U3 region and the minor transcriptional start site (9), is still high, being similar to that of the control plasmid pBI221, which contains the GUS reporter gene under the control of the complete CaMV 35S promoter, a strong constitutive promoter in plant systems. The pBI221 was used as a standard control in this work, in order to compare the activities found for different plasmids by different transient expression techniques. Deletion of an internal EcoRV-SspI region, that contains the minor transcriptional start site, giving the 7031 construct, resulted in a more severe loss of activity. These results indicate that both deleted regions are necessary to obtain a maximum level of expression. Nevertheless, deletion of the sequences 5' to the first putative TATA-box (referred as 'LTR upstream sequences' in this work) (construct 2146) containing most of the U3 region of the LTR, results in

Figure 1. Deletion analysis of the Tnt1 LTR. Diagrams show translational fusions between Tnt1 LTR deletions and the GUS reporter gene. The positions of the TATA-box and the transcription start sites are shown, as well as the sites used to produce the different deletions. Tnt1 sequences are shown by open boxes, GUS sequences are shown by black boxes and CaMV 35S promoter sequences are shown by a bold line. GUS activity in transient expression assays in electroporated tobacco protoplasts is expressed relative to the internal 35S-LUC standard (arbitrary units). Error bars are the standard error for five independent assays

a total loss of the GUS activity although this construct contains both LTR transcriptional start sites and putative TATA-boxes. This result is in agreement with preliminary results obtained in our group pointing to a determinant role of this region in expression of Tnt1 LTR in protoplasts (Pauls et al. in preparation).

U3 sequences upstream from the TATA-box are sufficient to confer protoplast-specific expression to a minimal heterologous promoter

In order to determine if the LTR upstream sequences were sufficient to confer protoplast-specific expression, we have developed a transient expression assay based on particle bombardment. Tobacco leaf discs were bombarded with different constructs containing LTR sequences fused to the GUS reporter gene. In each experiment, two discs of the same leaf were bombarded with the same construct, and one of these two discs was subsequently incubated in the presence of cell wall hydrolases. GUS expression was measured in the resulting protoplasts and compared to that obtained from the untreated control leaf disc. As shown in Fig 2, we have been able, using this experimental approach, to detect protoplast-associated induction of the LTR-GUS translational fusion (construct 2120) bombarded on leaf discs. This induction is also found for the 4691 construct, in which the LTR upstream sequences are placed in front of the 35S minimal (-46, +8) promoter (containing the sequences from the TATA-box to the transcription start site) fused to the GUS gene. As shown in Fig 2, in this case the GUS activity found in protoplasts is 4- to 6-fold higher than the activity found in untreated leaf discs. This induction is clearly greater than the slight induction shown by the control 35S-GUS construct. The GUS expression from the 2146 construct is very low in both leaf discs and protoplasts, and is therefore difficult to judge if the slight induction observed is significant or not.

Different subfragments of U3 region can stimulate activity of a minimal heterologous promoter in tobacco protoplasts

In order to further analyse the inducing effect of the LTR upstream sequences, different constructs containing either the

Figure 2. Analysis of LTR-mediated specific expression. Diagrams show the different constructs tested. GUS activity in leaf discs (shaded bars) and protoplasts (hatched bars) in transient expression assays by particle bombardment is expressed relative to the internal 35S-LUC standard. Error bars are the standard error for three independent assays.

Figure 3. Analysis of U3 subfragments inducing transcriptional activity on heterologous promoters. **A.** Diagram showing a magnification of the LTR upstream sequences (up to the HindIII site) and the position of the HindIII restriction site used to generate the 5' distal and the 3' proximal subfragments tested. Diagrams of the different constructs tested in front of the minimal CaMV 35S promoter (−46, +4) in direct (shaded bars) or reverse (hatched bars) orientation, are shown. GUS activity in transient expression assays in electroporated tobacco protoplasts is expressed relative to the internal 35S-LUC standard. Error bars are the standard error for five independent assays. **B.** Diagram showing the different LTR fragments tested in front of the truncated CaMV 35S promoter (−90, +4) in direct (shaded bars) or reverse (hatched bars) orientation. GUS activity in transient expression assays in electroporated tobacco protoplasts is expressed relative to the internal 35S-LUC standard. Error bars are the standard error for five independent assays.

entire region or different subfragments (a 5' distal and a 3' proximal region), were cloned in either orientation upstream of the CaMV 35S minimal (−46, +8) promoter fused to the gene GUS, and analysed by transient expression in tobacco protoplasts. Fig 3A. shows that each of the two regions can efficiently induce activity of the 35S minimal promoter when placed in direct orientation. The 5' distal region (construct 4651) gives a 6-fold induction, and the 3' proximal region (construct 4661) gives a 10-fold induction. When these regions are separately placed in reverse orientation (constructs 4652 and 4662) their inducing

Figure 4. Analysis of the ability of BI and BII to induce transcriptional activity on heterologous promoters. **A.** Nucleotide sequence of the Tnt1 LTR U3 region. The putative TATA box is underlined and the major transcriptional start site is shown by asterisks. Sequences corresponding to the BI and BII elements are boxed and the dyad symmetry of the BI sequence is shown by arrows. Restriction sites (HincII and HindIII) defining the proximal and the distal regions of the upstream LTR sequences, are shown in italics. **B.** Diagram showing the different constructs analysed, and their relative GUS/LUC activity by transient expression in tobacco protoplasts. Error bars are the standard error for five assays.

effect is clearly weaker. Nevertheless, the proximal region (construct 4662) still confers a 5-fold induction.

The entire LTR upstream sequence (construct 4691), gives a 30-fold induction of the minimal promoter. This induction is greater than the sum of the effects of the two sub-fragments. This fact, that suggests a synergistic effect of the two regions, could also be due to a position effect on the activity of the distal one. In order to investigate this possibility we have analysed a series of constructs in which the same sequences were placed at a different distance from the transcription start site by fusing them to a truncated 35S (−90, +4) promoter-GUS construct. The distance between the sequences analysed and the 35S transcription start is, in this case, 44 bp longer than in the series of constructs done with the 35S minimal promoter. The distance between the distal region and the transcription start site will thus be more similar to that found in the original Tnt1 promoter when this region is analysed alone, but it will be 70 bp longer when the construct contains the entire LTR upstream sequences.

Figure 5. BII sequences interact with protein(s) induced during protoplast preparation. Band shift analysis were performed with 10,000 cpm of 32 P-labelled ds oligonucleotide (0.1 to 0.5ng) corresponding to the BII sequence, and 20 μ g of tobacco nuclear extracts from protoplasts (tracks 1–5) or leaves (tracks 6–10). Competition experiments were performed by adding unlabelled ds BI oligonucleotide (4ng, tracks 2 and 7; 40ng tracks 3 and 8) or unlabelled ds BII oligonucleotide (4ng, tracks 4 and 9; 40ng, tracks 5 and 10). The specific retarded band is shown by an arrow.

Fig 3B shows that the entire LTR upstream sequences gives a 15-fold induction of the 35S truncated promoter, when placed upstream of this promoter (construct 9091). This induction is clearly weaker than the inducing effect of the same sequences placed in front of the 35S minimal promoter (4691, Fig 3A). In contrast, the distal region gives a 12-fold induction when it is placed in front of the 35S truncated promoter (construct 9051), that is two times higher than the inducing effect of the same sequences placed in front of the 35S minimal promoter (4651 Fig 3A). In this case the inducing effect is similar to that of the proximal region or the entire LTR upstream sequences. The inducing effect of the proximal region seems to be less dependent on the distance (9061, Fig 3B vs 4661, Fig 3A), and, in accordance with results obtained when this region was placed in front of the 35S minimal promoter, its inducing effect on the truncated 35S promoter remains independent of the orientation.

Two different short sequences can stimulate transcription of a minimal heterologous promoter in tobacco protoplasts

A detailed analysis of the U3 region of the LTR revealed two particularly interesting sequences. In the distal region of the LTR upstream sequences, a 20-bp sequence (BI, Fig 4A), which displays dyad symmetry, contains in the centre a sequence (CATGTG) similar to the myc binding site (CACGTG) (20) and the consensus sequence (CANNTG) for the factor-binding elements (E boxes) within the immunoglobulin enhancers (21). The sequence found (GGCATGTGC) is also very similar to a sequence in the Bz1 promoter from maize (GGCAGGTGC) that is essential for Bz1 expression in the presence of the maize regulatory genes C1 and R (22). The core sequence also resembles the G-box related sequences found in plant genes regulated by many different stimuli such as light or ABA (23, 24).

A 31-bp sequence (BII, Fig 4A) that does not show any clear similarity with any described transcription factor recognition boxes, is found repeated once in the distal and twice in the proximal region (Fig 4A). Interestingly, the major variation found in the U3 region between different transcriptionally active Tnt1 elements in tobacco is the presence, in most of them, of an extra BII sequence in this region (Casacuberta, unpublished).

To analyse the possible role of these sequences in the regulation of Tnt1 expression, we have cloned tandem repeats of corresponding oligonucleotides upstream of a truncated 35S promoter (–90, +4) fused to the GUS gene. Results of the analysis of these constructs by transient expression in tobacco protoplasts are shown in Fig4B. A tandem of 5 BI oligonucleotides induces a level of GUS expression that is twice that of the promoter alone. A tandem of only 2 BII oligonucleotides is able to induce the GUS gene expression 3 to 4-fold. The inducing effect of BII on the expression of a heterologous promoter is thus clearly higher than that of BI.

BII sequences interact with protoplast specific nuclear protein(s)

Most cis-elements involved in the regulation of gene expression interact with nuclear protein factors. In order to investigate the protein factors that could play a role in Tnt1 activation in protoplasts, we studied by band shift analysis the interaction of BI and BII sequences with tobacco nuclear proteins. No specific interactions were detected with the BI probe (data not shown). Band shift analysis using the BII oligonucleotide probe, and protoplast nuclear extracts, reveal several retarded bands (Fig 5). Most of the retarded bands are competed by increasing concentrations of unlabelled BI oligonucleotide, and so are non-specific interactions. A retarded band, marked with an arrow in Fig 5, is resistant to competition with BI oligonucleotide but is efficiently competed by similar amounts of BII oligonucleotide. This retarded band is thus due to a specific interaction between BII sequence and tobacco protoplast nuclear protein(s). This binding activity is absent in leaf nuclear extracts (tracks 6 to 10, Fig 5) whilst the non-specific binding activities are apparently present at the same level. This binding activity is thus induced during protoplast preparation and could be involved in the induction of Tnt1 transcription associated with the presence of cell wall hydrolases.

DISCUSSION

Previous work done in our group has shown that Tnt1 retrotransposon expression is strongly regulated. Its RNA is not detectable in most of the tissues of the adult plant, being almost absent in adult tobacco leaves but strongly induced when preparing leaf-derived protoplasts (9). We present here the localisation of LTR sequences involved in the control of Tnt1 retrotransposon expression during protoplast preparation.

The results of transient expression analysis of LTR-GUS translational fusions have revealed that all the different LTR regions analysed are necessary for a maximum level of Tnt1 expression. Deletion of each of the three different regions tested results in a severe loss of activity of the reporter gene. This loss of activity when deleting the central or the 3' region of the LTR could be due to the presence, in each of these regions, of one of the two transcriptional start sites, as well as putative TATA boxes. However, transcription regulatory elements may also be present within these regions. Transcriptional regulators downstream of the transcription start site have been described for other retrotransposons. For example, the *drosophila copia* retrotransposon contains a transcriptional enhancer within the leader untranslated region downstream of the 5' LTR (25), and the transcriptional regulators of yeast Ty elements, controlled by the mating-type locus, MAT, are also located downstream of the 5' LTR (26, 27).

The most drastic loss of Tnt1 promoter activity was found when deleting the sequences upstream of the first putative TATA-box (LTR upstream sequences). This deletion gives an almost inactive construct although it contains both LTR transcription starts and putative TATA-boxes. This result points to a major role for the LTR upstream sequences in the control of Tnt1 transcription, and is in agreement with the presence of regulatory elements in the U3 region of the LTR of most of the retroelements previously described (28–30). In particular, upstream activator sequences have been described in the U3 region of Ty2 (31) and *copia* (25).

In order to determine if the LTR upstream sequences are sufficient to mediate protoplast-associated induction, we have analysed their ability to activate a minimal heterologous promoter in transient expression during protoplast preparation, using a particle bombardment approach. The convenience of the experimental approach used was determined by analysing the induction of a LTR-GUS fusion that was known to be induced in protoplasts obtained from transgenic leaves (9). The results obtained show that the LTR-GUS fusion is induced in protoplasts obtained from leaves bombarded with this construct. However, the level of induction is 2 to 5 fold lower than the one obtained with leaves stably transformed with this construct (9), and the GUS activity, when compared to that of the 35S-GUS control construct, is lower than in electroporated protoplasts (fig 2 versus fig1). This lower level of induction is possibly due to the fact that bombarded plasmids are transcribed during the first few hours after bombardment (32), and in our case, part of this time is needed to produce protoplasts and thus to induce the bombarded construct. Nevertheless, as induction is clearly detectable by this approach, we have used this method to test the ability of the LTR upstream sequences to mediate transcriptional induction when preparing leaf-derived protoplasts. Our results show that the LTR upstream sequences are sufficient to confer protoplast-associated induction to a minimal heterologous promoter and thus are probably responsible, at least in part, for the transcriptional regulation of Tnt1 and for its induction in protoplasts associated with the presence of cell wall hydrolases. The possibility remains that other sequences within or outside the LTR could also play a role in reinforcing this control.

A more detailed study of the LTR upstream sequences has been performed by analysing separately a distal and a proximal subfragment in front of heterologous promoters. Two equivalent series of constructs, containing these subfragments in front of two different promoters, a minimal and a truncated CaMV 35S promoter, fused to the GUS reporter gene, have been analysed. The distance between each one of the different fragments and the transcriptional start site is different in the two series of constructs, being 44 bp longer in the series done with the truncated 35S promoter. In addition, the truncated 35S promoter, a weak promoter in plants (14), contains sequences recognised by plant transcription factors (33–34). The inducing activity of the entire LTR upstream sequences, in transient expression in tobacco protoplasts, is two times lower on the truncated 35S promoter than on the minimal 35S promoter. In contrast, the 5' distal subfragment, gives an activation that is two times higher when placed upstream of the truncated 35S promoter. The activation found for the 3' subfragment is equivalent on both promoters. These results indicate that there is no positive interaction between the Tnt1 sequences analysed and the binding sequences that lie within the 35S truncated promoter, and that two different types of inducing elements are found in the U3 region of the Tnt1 LTR. The elements present in the distal region

seem to induce transcription in a position/orientation-dependent manner, whilst elements present in the proximal region induce transcription in a less position dependent way. These two types of inducing elements could correspond to the two different boxes, BI and BII, present in the U3 LTR region. BI, the palindromic E box-related sequence present in the distal region of the LTR upstream sequences, gives a two fold increase of GUS activity when placed in 5 copies upstream of a truncated 35S promoter in transient expression experiments. Although we have not been able to detect interactions between this sequence and tobacco nuclear extracts by band shift analysis, we think that it could play a role in inducing Tnt1 transcription by interaction with nuclear factors, as other E box-related sequences do (20, 21). The other box analysed here, BII, is found repeated once in the distal and twice in the proximal regions and gives a three to four fold increase on GUS activity, when placed in a tandem of only 2 copies upstream of a truncated 35S promoter in transient expression experiments. In addition, this BII sequence interacts with tobacco nuclear proteins from protoplasts but not from leaves. The factor(s) that interact(s) with the box BII is thus induced during protoplast preparation. This result reinforces the hypothesis that this sequence plays a key role in the activation of Tnt1 transcription. The presence of tandemly-repeated promoter elements has also been described for the drosophila 1731 retrotransposon, in which a hormone-responsive element is located in a 28 bp repeated sequence within the U3 region of the element (35). The repetition in tandem of important promoter sequences could prevent these elements from inactivation by point mutations introduced by error-prone reverse transcription.

Stable transformation experiments with constructs containing these sequences, as well as characterisation of the BII binding factor(s), will allow us to more precisely define the role played by this sequence in the control of transcription, and therefore, of transposition of the Tnt1 tobacco retrotransposon.

ACKNOWLEDGEMENTS

We are grateful to C.Godon for excellent expertise on bombardment experiments and to P.Guerche for precious help on isolating tobacco protoplasts and electroporation. We also wish to thank T.Grange and M.Ll.Espinas for helpful discussions, and M.Caboche, J.Ellis, H.Lucas, C.Meyer, I.Small and H.Vaucheret for critical reading of the manuscript. J.M.C. was supported by an EMBO long-term fellowship.

REFERENCES

1. Varmus, H. and Brown, P. (1989) In Berg, DE and Howe MM (eds), *Mobile DNA*. American Society for Microbiology, Washington DC, pp. 53–108.
2. Flavell, A.J., Dunbar, E., Anderson, R., Pearce, S.R., Hartley, R. and Kumar, A. (1992) *Nucleic Acids Res* 20, 3639–3644.
3. Voytas, D.F., Cummings, M.P., Konieczky, A., Ausubel, F.M. and Roderick, S.R. (1992) *Proc Natl Acad Sci USA* 89, 7124–7128.
4. Grandbastien, M.A. (1992) *Trends Genet* 8, 103–108.
5. Johns, M.A., Mottinger, J. and Freeling, M. (1985) *EMBO J* 4, 1093–1102.
6. Grandbastien, M.A., Spielmann, A. and Caboche, M. (1989) *Nature* 337, 376–380.
7. Parkhurst, S.M. and Corces, V.G. (1987) *EMBO J* 6, 419–424.
8. Elder, R.T., St John, T.P., Stinchcomb, D.T., Davis, R.D. (1980) *Cold Spring Harbor Symp Quant Biol* 45, 581–584.
9. Pouteau, S., Huttner, E., Grandbastien, M.A. and Caboche, M. (1991) *EMBO J* 10, 1911–1918.
10. Jefferson, R.A. (1987) *Plant Mol Biol Rep* 5, 387–405.
11. Bourgin, J.P. and Missonier, C. (1973) *Haploid Inform Serv* 8, 7.

12. Chupeau, Y., Bourgin, J.P., Missonier, C., Dorion, N. and Morel, G. (1974) *C R Acad Sci (D) (Paris)* **278**, 1565–1568.
13. Bourgin, J.P., Chupeau, Y. and Missonier, C. (1979) *Physiol Plant* **45**, 288–292.
14. Benfey, P.N., Ren, L. and Chua, N.-H. (1990) *EMBO J* **9**, 1677–1684.
15. Guerche, P., Bellini, C., LeMoullec, J.M. and Caboche, M. (1987) *Biochimie* **69**, 621–628.
16. Ow, D.W., Keith, V., De Luca, M., de Wet, J.R., Helinski, D.R. and Howell, S.H. (1986) *Science* **234**, 856–859.
17. Wood, K.V. (1990) *J. Cell. Biol.* **111**, 380A.
18. Godon, C. Daniel-Vedele, F. and Caboche, M. (1993) *Biochimie* (in press).
19. Grange, T., Roux, J., Rigaud, G. and Pictet, R. (1991) *Nucleic Acids Res* **19**, 131–139.
20. Kerkhoff, E., Bister, K. and Klempnauer, K.-H. (1991) *Proc Natl Acad Sci USA* **88**, 4323–4327.
21. Murre, C., McCaw, S.P. and Baltimore, D. (1989) *Cell* **56**, 777–783.
22. Roth, R.A., Goff, S.A., Klein, T.M. and Fromm, M.E. (1991) *Plant Cell* **3**, 317–325.
23. Giuliano, G., Richersky E., Malik, V.S., Timko, M.P., Scolnik, P.A. and Cashmore, A.R. (1988) *Proc Natl Acad Sci USA* **85**, 7089–7093.
24. Marcotte, W.R.Jr, Russel, S.H. and Quatrano, R.S. (1989) *Plant Cell* **1**, 969–976.
25. Sneddon, A. and Flavell, A.J. (1989) *Nucleic Acids Res* **17**, 4025–4035.
26. Roeder, G.S., Rose, A.B. and Pearlman, R.E. (1985) *Proc Natl Acad Sci USA* **82**, 5428–5432.
27. Company, M. and Errede, B. (1987) *Mol Cell Biol* **7**, 3205–3211.
28. Varmus, H. (1988) *Science* **240**, 1427–1435.
29. Boeke, J.D. and Corces, V.G. (1989) *Annu Rev Microbiol* **43**, 403–434.
30. Arkhipova, I.R. and Ilyin, Y.V. (1992) *BioEssays* **14**, 161–168.
31. Liao, X.-B., Clare, J.J. and Farabaugh, P.J. (1987) *Proc Natl Acad Sci USA*, **84**, 8520–8524.
32. Rolfe, S.A. and Tobin, E.M. (1991) *Proc Natl Acad Sci USA* **88**, 2683–2686.
33. Prat, S., Willmitzer, L. and Sanchez Serrano, J.J. (1989) *Mol Gen Genet* **217**, 209–214.
34. Bouchez, D., Tokuhiya, J.G., Llewellyn, D.J., Dennis, E.S. and Ellis, J.G. (1989) *EMBO J* **8**, 4197–4204.
35. Ziarczyk, P. and Best-Belpomme, M. (1991) *Nucleic Acids Res.* **19**, 5689–5693.