

Analysis of cis-active sequences involved in the leaf-specific expression of a potato gene in transgenic plants

(*Agrobacterium*/chimeric genes/regulatory sequences/organ-specific expression/photosynthetic tissue)

JÖRG STOCKHAUS, PETER ECKES*, MARIO ROCHA-SOSA, JEFF SCHELL, AND LOTHAR WILLMITZER

Institut für Genbiologische Forschung Berlin, Ihnestr. 63, 1000 Berlin 33, and Max-Planck-Institut für Züchtungsforschung, 5000 Köln 30, Federal Republic of Germany

Contributed by Jeff Schell, July 24, 1987

ABSTRACT *ST-LS1* is a light-inducible, single-copy gene from potato that is expressed only in photosynthetic tissues. Various sequences derived from the 5'-upstream region of this gene were fused to the coding region of the chloramphenicol acetyltransferase (CAT) gene and to the gene 7 termination region of the transfer DNA (T-DNA) from the *Agrobacterium* Ti plasmid pTiACH5 and transferred to tobacco using Ti-plasmid vectors. After regeneration of whole plants, tissues were assayed for the expression of the CAT gene. Sequences derived from the 5'-upstream region of the *ST-LS1* gene comprising positions -334 to +11 were sufficient to confer a leaf/stem-specific as well as a light-inducible expression of the CAT gene. Destruction of chloroplasts by treatment with the herbicide norflurazon and subsequent exposure to light drastically reduced the expression of the CAT gene indicating that this upstream sequence most likely interacts with a chloroplast-dependent signal. When sequences from position -98 to position +675 were fused to a truncated inactive fragment of the cauliflower mosaic virus 35S promoter in a head-to-head manner, the corresponding chimeric genes were again expressed in photosynthetic tissues only, indicating that these sequences have enhancer-like properties.

Differential gene expression in higher plants is controlled by environmental as well as by developmental factors. Polysomal RNA sequences of various vegetative organs such as leaf, root, or stem contain at least 6000–11,000 mRNA species that cannot be detected in polysomes of other organs. The organ-specific expression of a large number of these genes is most likely controlled at the level of transcription (1, 2).

We are studying factors involved in developmentally controlled gene expression. To this end we have isolated several genes from potato that are expressed specifically during development. The expression of one of these genes (named *ST-LS1*) is restricted to photosynthetic tissues (leaf/stem) and is light inducible (3, 4). A single copy of this gene was detected in the haploid genome of potato (4). This is in contrast to otherwise similar genes encoding either the small subunit of the ribulose-1,5-bisphosphate carboxylase (*rbcS*) or the chlorophyll a/b binding protein (*lhcp*) that are members of fairly large gene families (5, 6). Genes encoding the *rbcS* or the *lhcp* proteins isolated from pea and wheat as well as the chalcone synthase from *Antirrhinum majus* have been used in gene-transfer experiments to define cis-active regulatory elements important for light-dependent gene expression (7–10).

To approach as much as possible the normal conditions that prevail in photosynthetic tissues such as leaves and stems, our analysis was based primarily on gene-expression studies in fully developed transgenic plants. This is in

contrast to the studies mentioned above that were mainly performed on callus or teratoma-like tissue.

We have shown before that the expression of the *ST-LS1* gene is qualitatively and quantitatively very similar in transgenic potato or tobacco plants (11). Therefore, for reasons of experimental convenience, we chose to use tobacco as a host for the gene-transfer experiments described below.

MATERIALS AND METHODS

Recombinant DNA Techniques. Standard procedures were used for recombinant DNA work (12).

Isolation and Analysis of Nucleic Acids. Isolation of RNA and DNA from plants and from *Agrobacteria* and their subsequent analysis by blot hybridizations were performed as described (4, 13).

Transformation of Tobacco Plants with Various Chimeric Genes. Chimeric genes consisting of various 5'-upstream segments of the *ST-LS1* gene, the chloramphenicol acetyltransferase (CAT)-coding region and the termination region of the gene 7 of the transfer DNA (T-DNA) of the Ti plasmid ACH5 (14) were cloned into the intermediate vector pMPK110 (4) and mobilized into the *Agrobacterium* receptor strain C58C1 (pGV3850kan) as described (13). Transformation and regeneration of tobacco plants followed established procedures (13). All manipulations were performed according to the "Richtlinien für den Umgang mit neu-kombinierten Nukleinsäuren" of the Bundesministerium für Forschung und Technologie.

CAT Assays. Plant extracts were assayed for CAT activity essentially as described (15) using 0.25 M Tris-HCl (pH 7.5) as buffer throughout.

Induction of Callus Tissue. For induction of callus tissue leaf discs of transformed plants were put onto basic MS medium (16) supplemented with naphthylacetic acid at 1 mg/liter and benzylaminopurine at 0.2 mg/liter. Growing callus tissue was subcultured every 3 weeks on the same medium and kept under a dark/light regime of 8 hr of dark/16 hr of light.

RESULTS

Transgenic Tobacco Plants with Genes Containing the CAT Gene Fused to Different Deletions of the 5'-Upstream Region of the *ST-LS1* Gene. Chimeric genes were constructed by cloning three fragments of the 5'-upstream region of the *ST-LS1* gene into *EcoRI*-*Sma* I-digested pMPK110 plasmid DNA. At the 3' end the fragments terminated at position +11 of the *ST-LS1* gene, whereas they extended to positions

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: CAT, chloramphenicol acetyltransferase; T-DNA, transfer DNA.

*Present address: Hoechst AG, 6000 Frankfurt/Main, Federal Republic of Germany.

-130, -334, and -1600 at their respective 5' ends. A *Bam*HI-*Eco*RI fragment containing the coding region of the Tn9-CAT gene as well as the termination region of gene 7 of the pTiACH5 T-DNA (14) was subsequently inserted using the *Sal*I site of the polylinker region behind the various 5' fragments. The resulting constructions contain an untranslated leader of 78 nucleotides in front of the ATG of the CAT gene composed of 11 nucleotides derived from the 5'-untranslated leader of the *ST-LS1* gene and of 67 nucleotides originating from both the polylinker region and the original leader of the CAT gene.

A schematic view of the various constructions is shown in Fig. 1. The construction O-CAT consists of the CAT coding region and of the termination fragment without any 5' sequences of the *ST-LS1* gene and serves as a control.

The various constructs were mobilized into "disarmed" (cf. ref. 24) *Agrobacterium* strains, and transgenic plants were obtained by using a leaf-disc infection procedure. Regenerated plants were shown to contain the correct structure of the transferred gene by Southern blotting (data not shown), and only the plants that contained at least one copy of the intact nonrearranged gene were used for further analysis.

A Segment of 130 Nucleotides of the 5'-Upstream Region Is Not Sufficient for Expression of the CAT Gene in Transgenic Plants. For each construct, leaves of five independently derived transgenic plants were assayed for the CAT enzyme. As expected no activity was observed with the promoterless O-CAT construction (data not shown). More interestingly, however, we were also unable to detect any activity with the construction carrying 130 nucleotides of the 5'-upstream region fused to the CAT gene (130-CAT) (data not shown), although the TATAAA and the CAAT boxes are contained in this fragment. As this result does not exclude the possibility that this construct could be active in other organs, we decided to test the activity in stem and root tissue of several transgenic plants. No activity was detected in five independent transgenic plants tested (data not shown).

As shown in Fig. 2, the constructs carrying 334 and 1600 nucleotides of the 5'-upstream region fused to the CAT gene (334-CAT and 1600-CAT, respectively) are active in tobacco leaves, although the relative activity in individual transformants varied considerably. By assaying pooled extracts of 30 independently transformed plants for each construct, evidence was obtained that the 1600-CAT construct was expressed at ≈ 3 -fold higher level than the 334-CAT construct in leaves of transgenic tobacco plants (data not shown).

A Segment of 334 Nucleotides of the 5'-Upstream Region of the *ST-LS1* Gene Is Sufficient to Confer Organ-Specific Expression to Chimeric CAT Genes. As outlined above the *ST-LS1* gene is expressed organ specifically in photosynthetic (leaf/stem) tissues of potato plants. Leaf, stem, and root tissue of transgenic tobacco plants were, therefore, analyzed for expression of the various chimeric genes. The CAT assays in Fig. 3 show that the 1600-CAT and the 334-CAT constructs are highly expressed in leaves and moderately

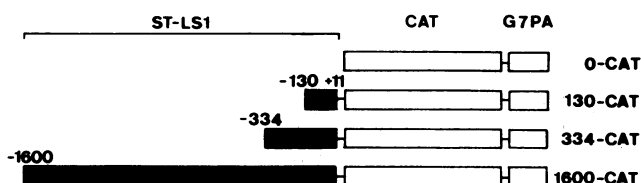


FIG. 1. Schematic structure of CAT chimeric genes under the control of various fragments of the 5'-upstream region of the *ST-LS1* gene. The various 5'-upstream fragments of the *ST-LS1* gene that have been fused to the CAT gene are shown. The poly(A) addition site of gene 7 of the T-DNA of pTiACH5 (14) was used in all construction.

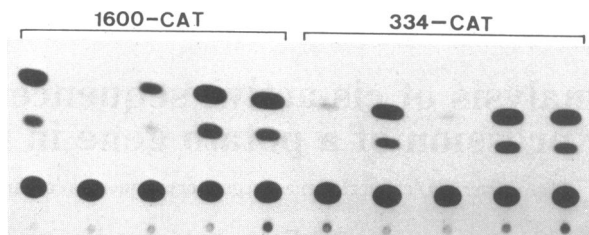


FIG. 2. CAT assays of leaf extracts of each of five independently obtained transgenic tobacco plants transformed with either the 334-CAT or 1600-CAT. Extracts assayed for CAT activity were adjusted to same amounts of protein.

expressed in stems, but no activity above background was detected in roots. This organ-specific expression was consistently observed in 9 out of 10 transformants tested with one exception, a 334-CAT transformant where stem tissue displayed a higher CAT activity than leaf tissue (data not shown). In all cases the extracts assayed for CAT activity were adjusted for similar protein content. The same result in terms of organ-specific expression was obtained when analysis was performed by RNA gel blot experiments (data not shown). The expression of the *ST-LS1* gene has been shown to be low in undifferentiated white callus tissue (3). To determine how the chimeric genes would be expressed under these circumstances callus containing either the 334-CAT or the 1600-CAT construct was initiated from leaf tissue of transgenic plants that had been shown to express the CAT gene. The results in Fig. 3 show that in both cases the CAT activity in callus tissue is low compared to leaf tissue. Again no expression was detected in calli containing the 130-CAT construct (data not shown).

A Segment of 334 Nucleotides of the 5'-Upstream Region of the *ST-LS1* Gene Is Sufficient to Confer Light-Inducible Expression to Chimeric Genes. Fig. 4 shows the activity of the 1600-CAT and of the 334-CAT constructs in leaves of plantlets that were either kept in the dark for 10 days or put back into light for 2 days after the dark treatment. In all transformants analyzed, the CAT activity increases considerably after light treatment although the extent of the relative induction rate varied from 4-fold to >20 -fold (data not shown).

Expression of the Chimeric *ST-LS1*-CAT Constructs Requires Intact Chloroplasts. The expression of the *ST-LS1* gene is restricted to photosynthetic green tissue (3). The herbicide norflurazon, an inhibitor of carotenoid biosynthesis, leads to a selective damage of chloroplast structure as a result of photooxidation when the tissue is exposed to strong white light (17). Expression of nuclear-encoded cytosolic and glyoxisomal genes and mitochondrial genes is largely unaffected by the norflurazon treatment (18, 19). It can, therefore, be used to analyze the importance of chloroplast

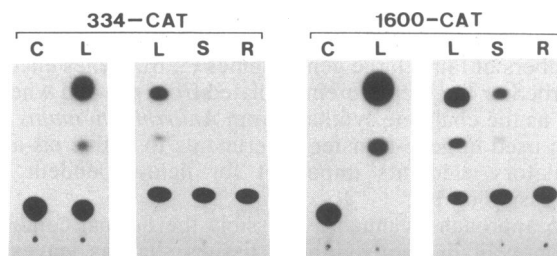


FIG. 3. Developmentally specific expression of 334-CAT and 1600-CAT constructs in leaf, stem, root, and leaf-derived callus tissue of transgenic tobacco plants. CAT assays were performed using extracts from leaf (L), stem (S), root (R), and leaf-derived callus (C) tissue. The activity of callus (C) tissue should be compared to the activity of leaf tissue shown in the accompanying lane (L).

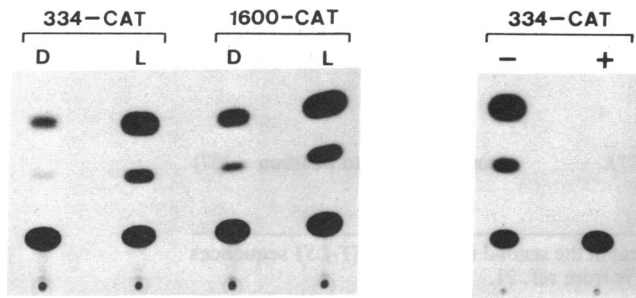


FIG. 4. Effect of light and norflurazon treatment on 334-CAT and 1600-CAT expression. CAT assays were performed on leaf tissue of plants that were kept in dark for 10 days (lanes D) and then put back into light for 2 days (lanes L) or grown in a light/dark rhythm of 16 hr of light and 8 hr of dark in either the presence (+) or absence (-) of $1 \mu\text{M}$ norflurazon. Extracts were adjusted to equal amounts of protein.

structure on the expression of nuclear-encoded genes (20). The data of Fig. 4 show that treatment of transgenic tobacco plants with norflurazon in white light leads to a drastic reduction in expression of the 334-CAT construct. To further elucidate the nature of this effect, green and white tissue segments of the same leaf obtained after norflurazon treatment were analyzed for expression of the 334-CAT construct. A high CAT activity was detected in the green areas of the leaf whereas a much lower level was observed in the white areas (data not shown). This result indicates that irrespective of the exact nature of the signal(s) for the chloroplast-controlled expression of the *ST-LS1* gene, its action seems to be restricted to the cells where it is produced.

5'-Upstream Sequences Controlling the Differential Expression of the *ST-LS1* Gene Have Enhancer-Like Properties. Many regulatory elements of yeast and animal genes can function in a bidirectional manner (21). Such elements were also detected in *rbcS* genes (9, 22) and in the gene encoding a chlorophyll *a/b* binding protein from pea (23).

We, therefore, decided to test whether or not the regulatory region important for the differential expression of the *ST-LS1* gene as defined by deletion analysis would also work when fused to a heterologous promoter. A deletion of the 35S promoter of cauliflower mosaic virus down to position -90 that was fused to the CAT coding region and did not produce detectable CAT activity in leaves of transgenic tobacco plants (cf. Fig. 5, lane -) was used. To this truncated and inactive promoter, we fused a *HincII* fragment containing the 5'-upstream region of the *ST-LS1* gene from position -98 to position -675. To be able to directly test whether or not this *HincII* fragment contains enhancer-like elements, it was fused in a head-to-head fashion to the truncated 35S promoter fragment, this construct was subsequently transferred into tobacco cells, and whole plants were regenerated. The results shown in Fig. 5 demonstrate that the addition of the *HincII* fragment from the *ST-LS1* gene to the truncated 35S promoter results in an organ-specific expression of the chimeric CAT gene. Furthermore, both light inducibility and dependence on chloroplast structure were displayed (data not shown). As the *HincII* fragment was fused in a head-to-head manner to the 35S promoter fragment, these results indicate that the elements of the *ST-LS1* gene controlling its differential expression display enhancer-like activities.

DISCUSSION

Various parts of the 5'-upstream region of the *ST-LS1* gene, a developmentally controlled gene from potato that is limited in its expression to photosynthetic tissues, have been fused to the coding region of the CAT gene of Tn9 and ligated with the terminator region of gene 7 of the T-DNA. Chimeric genes

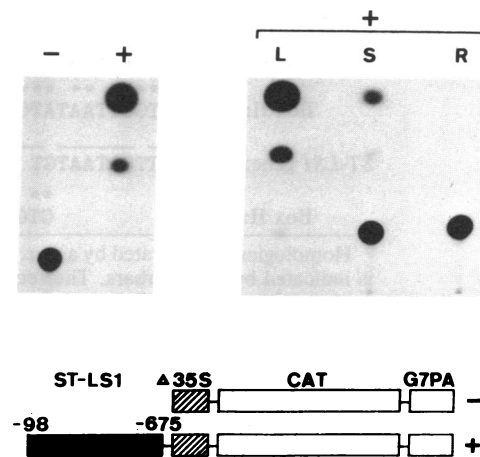


FIG. 5. Schematic structure and developmental-specific expression of a chimeric CAT gene containing regulatory sequences of the 5'-upstream elements of the *ST-LS1* gene fused to a truncated fragment of the 35S promoter of the cauliflower mosaic virus. A segment from position -98 to position -675 was fused to the truncated promoter fragment of the 35S promoter of cauliflower mosaic virus (containing position -90 to position +2) and connected by way of a transcriptional fusion to the CAT coding region and the terminator region of gene 7 of the T-DNA. Leaf (L), stem (S), and root (R) tissue of tobacco plants transformed with this construct (+) as well as leaf tissue (lane -) of plants transformed with the construct devoid of the upstream element of the *ST-LS1* gene were assayed for CAT activity.

have been transferred into tobacco cells using *Agrobacterium* vectors, and whole plants have been regenerated.

Although sequences up to position -130 contain several elements considered to be of prime importance to the expression of plant genes, such as the CAAT box and the TATA box, chimeric 130-CAT genes were apparently inactive in leaves, stems, or roots of transgenic plants. The only tissue where we were able to detect activity was a teratoma-like tissue obtained by transformation using a partly disarmed Ti plasmid, pGV3851 (24) (data not shown). Thus the hormonal content of the plant cells in this case appears to change the control of expression exerted by the upstream elements. This difference in gene expression between completely undifferentiated or only partly differentiated plant tissue and whole plants has also been observed for two other plant genes. Thus, for the chalcone synthase gene from *Antirrhinum majus* (10), expression in whole plants required >600 nucleotides of upstream sequences, whereas much shorter sequences were enough to drive expression in transformed calli or teratoma tissue (H. Kaulen, personal communication). Along the same lines for the *rbcS* gene *E9* from pea, deletion of 5'-upstream sequences from position -352 to position -35, which had little or no effect upon already low expression level in transformed calli (8), led to a drastic decrease in expression level of the same gene in transgenic plants (22).

Transgenic tobacco plants carrying 334-CAT constructions expressed CAT activity in an organ-specific (leaf/stem-specific) as well as light-inducible manner. Expression was, however, drastically reduced as a result of damage to the chloroplast structure by exposing the plant to the herbicide norflurazon. Thus the sequence up to position -334 must contain elements that are important for basic as well as differential expression. The most important elements seem to be located on the 5' upstream side of position -98, since fusing the sequence from position -98 to position -675 to the truncated, inactive 35S promoter led to a leaf/stem-specific, light-inducible, and organ-specific expression of the resulting chimeric genes in transgenic tobacco plants (cf. Fig. 5 and

Table 1. Homology between box II and the 5'-upstream region of the *ST-LSI* gene

<i>ST-LSI</i> gene:	TGGCAAAAATG	(position -286 to position -296)
Box II:	*** ** *** TGGTTAATATG	
<i>ST-LSI</i> gene:	TTGATAATGT TGGTAATATTATCTA	(position -143 to position -157)
Box II:	** **** ** *** GTGTGGT TAATAT	

Homologies are indicated by a star. The inverted repeat in the second motif of the *ST-LSI* sequences is indicated by the overbars. The sequence for box II is from ref. 21.

data not shown). When this result is compared to data obtained for other developmentally regulated plant genes studied at the whole plant level, certain common features emerge. Thus for the pea *rbcS* genes *E9* and *3A* (22) and the chlorophyll a/b binding protein gene *AB80* (23) sequences located between positions -80 and -350 are sufficient to confer an organ-specific expression to chimeric genes consisting of these upstream sequences fused to minimal 35S or nopaline synthase promoters. Similarly, 5'-deletion mutants ending at position -257 of the embryo-specific β -conglycinin gene from soybean retained embryo-specific expression (25).

When the 5'-upstream region fused to the CAT gene was extended to around position -1600, the resulting chimeric gene did not differ from the 334-CAT construction in terms of qualitative expression; however, it was expressed at a higher level (\approx 3-fold) when compared to the 334-CAT construct.

The elements located between positions -80 and -350 of the *ST-LSI* gene, the *rbcS* gene *E9* and *3A* (22), and the *lhcp* gene *AB80* (23), which are necessary for differential expression, have certain enhancer-like properties. Indeed these sequences can be added to an otherwise inactive, truncated promoter in a head-to-head orientation giving rise to differential expression of the resulting chimeric gene. The observation that deletion of this region will yield an inactive gene (such as the 130-CAT gene) would in the simplest model predict that this cis-active sequence interacts with a light- and chloroplast-dependent activator, thus explaining the organ-specific expression of genes carrying these upstream elements. That this model is probably too simple was indicated by the finding that these enhancer-like elements can also have silencer-like activities (23).

Data described in this paper and in earlier reports (3, 4, 11) indicate that the organ-specific expression of the *ST-LSI* gene is (supposedly) the consequence of 5'-upstream cis-regulatory sequences of this gene that require factors produced in photosynthetically active tissues. Indeed the *ST-LSI* gene can be expressed in root tissue provided the parenchymatic cells of the root contain mature chloroplasts, which is a spontaneous event in tissue-cultured potato plants (3). Furthermore, this gene is active not only in leaf tissue but also in stem tissue. Finally the expression in leaf tissue is nearly completely abolished if the chloroplasts are damaged by photooxidative processes as a result of the inhibition of carotenoid biosynthesis by the herbicide norflurazon (cf. Fig. 5 and ref. 11). This gene is, therefore, very similar to the genes encoding the pea *rbcS* or the pea chlorophyll a/b binding protein, which also display a light-induced and developmentally specific expression restricted to green tissues. It is, therefore, reasonable to assume certain homologies in the respective 5' regions of these genes and the *ST-LSI* gene. Fluhr *et al.* (22) have described three motifs that are highly conserved between *rbcS* genes of different varieties. These boxes are located in a region that was important for developmentally specific and light-inducible expression of two *rbcS* genes from pea (22). A search for homology among these boxes and the sequences contained between positions -100 and -350 of the *ST-LSI* gene revealed two sequences

that have a high degree of homology to box II as defined by Fluhr *et al.* (22). Between positions -286 and -296, 8 out of 11 nucleotides are homologous to the corresponding 11 nucleotides of box II. A second region of homology is observed between positions -143 and -157. In this case 12 out of 15 nucleotides correspond (Table 1). This second region of homology is contained within a stretch of 28 nucleotides that allows the formation of a hairpin-loop-like structure composed of nine hydrogen bonds. A functional test will be needed to determine whether or not this physical structure is biologically important.

We thank Astrid Blau and Petra Krieger for performing the transformation of the plant, Dirk Scheel for scanning the chromatograms, Jeff Velten for providing pCAP, a plasmid containing the CAT gene and the 3' end of gene 7, Frank Staedtler for help in performing CAT assays, and Mrs. Meyer for editing the manuscript. This work was supported by a grant from the Bundesministerium für Forschung und Technologie (Molekulare Zellund Gentechnologie BCT 035/2).

- Kamalay, J. C. & Goldberg, R. B. (1980) *Cell* **19**, 935-946.
- Kamalay, J. C. & Goldberg, R. B. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2801-2805.
- Eckes, P., Schell, J. & Willmitzer, L. (1985) *Mol. Gen. Genet.* **199**, 216-224.
- Eckes, P., Rosahl, S., Schell, J. & Willmitzer, L. (1986) *Mol. Gen. Genet.* **205**, 14-22.
- Dean, C., van den Elzen, P., Tamaki, S., Dunsmuir, P. & Bedbrook, J. (1985) *EMBO J.* **4**, 3055-3061.
- Dunsmuir, P. (1985) *Nucleic Acids Res.* **13**, 2503-2518.
- Herrera-Estrella, L., Van den Broeck, G., Maenhaut, R., Van Montagu, M., Schell, J., Timko, M. & Cashmore, A. (1984) *Nature (London)* **310**, 115-120.
- Morelli, G., Nagy, F., Fraley, R. T., Rogers, S. & Chua, N. H. (1985) *Nature (London)* **315**, 200-204.
- Timko, M., Kausch, A. P., Castresana, C., Fassler, J., Herrera-Estrella, L., Van den Broeck, G., Van Montagu, M., Schell, J. & Cashmore, A. R. (1985) *Nature (London)* **318**, 579-582.
- Kaulen, H., Schell, J. & Kreuzaler, F. (1986) *EMBO J.* **5**, 1-5.
- Stockhaus, J., Eckes, P., Blau, A., Schell, J. & Willmitzer, L. (1987) *Nucleic Acids Res.* **15**, 3479-3491.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- Sanchez-Serrano, J., Keil, M., O'Connor-Sanchez, A., Schell, J. & Willmitzer, L. (1987) *EMBO J.* **6**, 303-306.
- Velten, J. & Schell, J. (1985) *Nucleic Acids Res.* **13**, 6981-6998.
- Gorman, C. M., Moffat, L. & Howard, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044-1051.
- Murashige, T. & Skoog, F. (1962) *Physiol. Plant.* **15**, 473-497.
- Frosch, S., Jabben, M., Bergfeld, R., Kleinig, H. & Mohr, H. (1979) *Planta* **145**, 497-505.
- Reiß, T., Bergfeld, R., Link, G., Thien, W. & Mohr, H. (1983) *Planta* **159**, 518-528.
- Oelmüller, R. & Mohr, H. (1986) *Planta* **167**, 106-113.
- Simpson, J., Van Montagu, M. & Herrera-Estrella, L. (1986) *Science* **233**, 34-38.
- Gluzman, Y. & Shenk, T., eds. (1984) *Enhancers and Eukaryotic Gene Expression* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

22. Fluhr, R., Kuhlemeier, C., Nagy, F. & Chua, N.-H. (1986) *Science* **233**, 1106–1112.
23. Simpson, J., Schell, J., Van Montagu, M. & Herrera-Estrella, L. (1986) *Nature (London)* **323**, 551–554.
24. Zambryski, P., Herrera-Estrella, L., De Block, M., Van Montagu, M. & Schell, J. (1984) in *Genetic Engineering, Principles and Methods*, eds. Setlow, J. & Hollaender, A. (Plenum, New York), Vol. 6, pp. 253–276.
25. Chen, Z.-L., Schuler, M. A. & Beachy, R. N. (1986) *Proc. Natl. Acad. Sci. USA* **33**, 8560–8564.