The EMBO Journal vol.8 no.13 pp.4197–4204, 1989

The ocs-element is a component of the promoters of several T-DNA and plant viral genes

D.Bouchez[1], J.G.Tokuhisa, D.J.Llewellyn, E.S.Dennis and J.G.Ellis

CSIRO Division of Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia

[1]Present address: Laboratoire de Biologie Cellulaire INRA, F78026 Versailles Cedex, France

Communicated by K.Marcker

The ocs-element is an enhancer element first identified in the promoter of the octopine synthase gene (*OCS*) where it occurs as a 16 bp palindromic sequence. The transcriptional enhancing activity of the ocs-element correlated with *in vitro* binding of a transcription factor. We have now identified ocs-elements in the promoter regions of six other T-DNA genes involved in opine synthesis and three plant viral promoters including the 35S promoter of cauliflower mosaic virus. These elements bind the ocs transcription factor *in vitro* and enhance transcription in plant cells. Comparison of the sequences of these 10 elements has defined a 20 bp consensus sequence, TGACG(T/C)AAG(C/G)(G/A)(A/C)T(G/T) ACG(T/C)(A/C)(A/C), which includes the 16 bp palindrome in its central region. We propose the name ocs-element for this class of promoter elements of similar sequence and function.

Key words: caulimovirus promoters/DNA–protein interaction/octopine synthase enhancer/T-DNA genes/transcription signals

Introduction

The development of routine methods for introducing DNA into plant cells has allowed rapid advances in the understanding of the control of plant gene expression. Several environmentally induced and tissue-specific plant gene promoter elements have been identified quite precisely by mutagenesis (Walker *et al.*, 1987; Green *et al.*, 1988; Gidoni *et al.*, 1989) and by the study of interactions between mutant and wild-type promoter elements and *trans*-acting nuclear factors (Green *et al.*, 1988; Schulze-Lefert *et al.*, 1989; Singh *et al.*, 1989).

One group of promoters, expressed in plants and the subject of considerable analysis, are the apparently constitutive promoters of 'parasitic' genes, e.g. the nopaline synthase promoter from the T-DNA of *Agrobacterium tumefaciens* (Ebert *et al.*, 1987; Mitra and An, 1989), or the 35S promoter of the cauliflower mosaic virus (Odell *et al.*, 1985; Ow *et al.*, 1987; Fang *et al.*, 1989). These promoters have been widely used for the construction of chimeric genes (Bevan *et al.*, 1983; Rogers *et al.*, 1987) and an understanding of their regulation is of considerable practical importance. Promoters of T-DNA genes such as the nopaline synthase (*NOS*) promoter, the octopine synthase (*OCS*) promoter and the mannopine synthase (*MAS*) promoters have been studied in transgenic plants using sensitive reporter genes. Although these promoters are active in undifferentiated cells and in various plant organs, a tissue or organ preference has been reported, with the highest activity often in roots or wounded leaf tissue (An *et al.*, 1988; Teeri *et al.*, 1989).

We have studied the octopine synthase promoter and have defined a specific *cis*-regulatory sequence, the 16 bp palindromic ocs-element (Ellis *et al.*, 1987a). The insertion of synthetic oligonucleotides containing this ocs-element upstream of the maize *Adh1* promoter, inactivated by deletion, reactivated the expression of this promoter in maize and tobacco protoplasts. This assay system has been used to study the effect of mutations introduced into the ocs-element on transcriptional enhancement and on interaction with nuclear protein extracts (Ellis *et al.*, 1987a; Singh *et al.*, 1989).

The ocs-element is a binding site for a transcription factor, OCSTF (Singh *et al.*, 1989; J.G.Tokuhisa *et al.*, in preparation). Gel-retardation assays using labelled DNA fragments from the *OCS* promoter or from cloned synthetic oligonucleotides containing the 16 bp palindrome reveal two specific retarded bands of different mobility resulting from the interaction between the DNA probe and OCSTF (Singh *et al.*, 1989; J.G.Tokuhisa *et al.*, in preparation). The same protein is involved in the two complexes, with a monomeric protein bound to the DNA probe in the faster migrating ('lower band') complex and a dimeric protein in the slower migrating ('upper band') complex (J.G.Tokuhisa *et al.*, in preparation). A set of mutant ocs-elements containing one or more base changes introduced into the 16 bp palindrome by oligonucleotide synthesis was analysed for protein binding activity in gel-retardation assays and for enhancer activity in transient expression assays (Singh *et al.*, 1989). A strong correlation was observed between the presence of the upper band DNA–protein complex in gels and activation of transcription *in vivo*. In other words, association of two protein units with the ocs-element is necessary for maximal transcriptional activation.

The ocs-element functions as an enhancer in dicot (tobacco) and monocot cells (maize) (Ellis *et al.*, 1987a). OCSTF is also present in both species and the properties of this factor in DNA binding and transcriptional activation are essentially identical (Singh *et al.*, 1989; J.G.Tokuhisa *et al.*, in preparation). OCSTF therefore is a transcription factor that is conserved, at least at a functional level, across the plant kingdom.

In this paper, we show that DNA sequences with homology to the 16 bp palindrome of the *OCS* promoter, all of which bind OCSTF, occur in 10 promoters, seven from T-DNA opine synthase genes from Ri and Ti plasmids and three from the promoters of the large transcript of three

© IRL Press

gene	5'	ocs element region	3'	distance ocs-tata	homology with ocs	reference
Ti OCS	AA	**ACGTAAGCGCTTACGT**	AC	141	16/16	De Greve *et al.* (1982)
CaMV 35S	TG	**ACGTAAGGGATGACGC**	AC	27	12/16	Franck *et al.* (1980)
FMV	TG	**ACGAACGCAGTGACGA**	CC	14	10/16	Richens *et al.* (1987)
CERV	AG	**ACGTCATGCATGACGT**	TT	27	10/16	Hull *et al.* (1986)
Ri MAS1'	TT	**ACGTCATCGCTTACGC**	AT	146	13/16	this work
Ri MAS2'	AT	**GCGTAAGCGATGACGT**	AA	128	13/16	this work
Ti MAS1'	TG	**ACTCGGATACTTACGT**	CA	26	9/16	Velten *et al.* (1984)
Ti MAS2'	TG	**ACGCTCGCGGTGACGC**	CA	39	10/16	Barker *et al.* (1983)
Ti NOS	TG	**AGCTAAGCACATACGT**	CA	85	12/16	Depicker *et al.* (1982)
Ti AGS	TG	**ACGCATCGAATGACGC**	GA	97	8/16	Barker *et al.* (1983)
CONSENSUS	**TG**	**ACG(T/C)AAG(CGA/GAC)(T)(G/T)ACG(T/C)**	**AA/CC**			

Fig. 1. Sequences homologous to the OCS 16 bp palindrome (top line) found in viral genes and in opine synthase genes of T-DNAs from Ti and Ri plasmids. The Ri *MAS1'* and Ri *MAS2'* are the same sequence in reverse orientations. A consensus sequence (bottom) has been derived from this alignment according to Cavener (1987). The distances in bp of these sequence elements from their respective TATA boxes are indicated and their homologies with the OCS palindrome tabulated.

related caulimoviruses, cauliflower mosaic virus (CaMV), figwort mosaic virus (FMV), carnation etched ring virus (CERV). We propose the name ocs-element to include all these promoter sequences which bind OCSTF and have sequence homology to a 20 bp consensus derived by comparing protein binding sites of these 10 promoters.

Results

Sequences homologous to the ocs-element in other genes

Sequences homologous to the 16 bp palindromic element of the *OCS* promoter have been identified in the promoter regions of 10 promoters from T-DNA opine synthase genes or plant viral genes (Figure 1). These sequences were found from 27 to 146 bp upstream of their respective TATA boxes (Figure 1). In the 35S promoter of CaMV, a sequence identical at 12 out of 16 bases of the 16 bp ocs-element occurs at positions −79 to −64. An internal deletion from −89 to −73 that removed approximately half of this sequence was shown to decrease the activity of the promoter 50-fold but the specific *cis*-acting sequence in this region was not identified (Ow *et al.*, 1987). The sequence homologous to the ocs-element overlaps the pentanucleotide direct repeat TGACG observed by Fang *et al.* (1989). Sequences homologous to the ocs-element also occur in the genome of two other caulimoviruses, FMV and CERV. These two viruses have a genomic organization similar to CaMV and the potential ocs-elements occur in the intergenic promoter region which is highly conserved between the three viruses (Richins *et al.*, 1987). The TGACG direct repeat occurs in the FMV genome and with a single mismatch in the CERV genome. In both cases, as in the 35S promoter, the pentanucleotide repeats are separated by 7 bp.

Sequences homologous to the ocs-element occur in six other T-DNA genes. As previously reported, a 12 out of 16 match is found in the promoter of the nopaline synthase gene (Ellis *et al.*, 1987a). A 13 out of 16 match occurs in the T-DNA of the Ri-plasmid of *Agrobacterium rhizogenes* strain A4 approximately equidistant from the transcription starts of the two divergent genes, *MAS1'* and *MAS2'*, involved in the biosynthesis of mannopine. Two homologous

Fig. 2. The CaMV 35S ocs-element (**a**) sequences of the 35S ocs-element and the three mutant sequences obtained by oligonucleotide directed mutagenesis showing the homology to the *OCS* palindrome, the location of the restriction sites used for isolating the gel retardation probes and the sequence alterations (underlined) in the mutants. The activity of a *GUS* reporter gene linked to the wild-type and mutant promoters is indicated and is the average of three experiments using *N.plumbaginifolia* protoplasts. (**b**) Gel retardation obtained with the 31 bp *Eco*RV−*Fok*I fragment of the wild-type 35S promoter. A strong upper band and a faint lower band were observed under the binding conditions used. **Lane 1**, no competitor DNA; **lane 2**, 48 bp competitor fragment from the *OCS* promoter; **lane 3**, competitor fragment from a cloned synthetic 16 bp *OCS* palindrome; **lane 4**, pUC polylinker as competitor. (**c**) A labelled DNA fragment containing a cloned *OCS* palindrome was used as probe. No competitor DNA (**lane 1**), pUC polylinker (**lane 2**) or the 31 bp *Eco*RV−*Fok*I fragment (**lane 3**) were added to the binding reaction. The competition of binding to the *OCS* promoter probe in this and subsequent figures was done with one extract preparation and the formation of mainly the lower band (lb) complex was due to a lower concentration of binding activity in the extract compared with that used to generate (b). (**d**) Gel-retardation pattern using the *Eco*RV−*Fok*I fragment from wild-type (**lane 1**) and mutant (**lanes 2−4**) 35S promoters. fp, free probe; lb, lower band; ub, upper band.

sequences occur close to the TATA boxes in the promoters of the equivalent mannopine synthase genes *MAS1'* (9/16 match) and *MAS2'* (10/16 match) that occur in the T-DNA

of octopine Ti-plasmids (Velten *et al.*, 1984). In the *MAS2′* promoter, direct repeats of TGACG are present with 7 bp spacing. The published sequences (Barker *et al.*, 1983; Velten *et al.*, 1984) differed in this region. Resequencing this region of pTiAch5 confirmed the sequences shown in Figure 1. Another similar sequence was detected in the promoter of the agropine synthase gene in the T-DNA of an octopine Ti plasmid (Barker *et al.*, 1983). The homology with the *OCS* palindrome is only 8 out of 16 but the direct repeat of the pentanucleotide TGACG occurs with a spacing of 7 bp between the repeated units. No obvious homology was detected in other T-DNA promoters, including the agropine synthase promoter of the Ri-plasmid, pRiA4. In order to test whether these sequences are functional ocs-elements, we analysed the ability of these sites to bind to the plant transcription factor, OCSTF (Singh *et al.*, 1989; J.G.Tokuhisa *et al.*, in preparation), and to act as transcriptional activators *in vivo*.

An active ocs-element in the 35S promoter of the CaMV

A 31 bp fragment (−89 to −59) of the 35S promoter containing the sequence homologous to the ocs-element was analysed in a gel-retardation assay with maize nuclear protein extract. This fragment bound protein from the nuclear extract to produce two retarded bands (Figure 2b, lane 1). The same pattern of retarded complexes is seen in gel-retardation experiments using fragments containing the 16 bp palindrome of the *OCS* promoter (Singh *et al.*, 1989). The binding to the 35S promoter fragment was competed by addition of DNA fragments containing the 16 bp palindrome of the *OCS* promoter (Figure 2b, lanes 2 and 3). In controls, no competition was observed when the same amount of pUC19 polylinker DNA was used as a competitor (Figure 2b, lane 4). Conversely, the binding to *OCS* promoter fragments is competed by the 35S promoter fragment (Figure 2c, lane 3). These experiments indicate that the 31 bp fragment from the 35S promoter binds the transcription factor, OCSTF. The same pattern of binding to this region of the 35S promoter to nuclear extracts from tobacco leaves has been recently reported by Prat *et al.* (1989), who have shown the importance of the pentanucleotide repeats by footprint and methylation interference assays.

Mutations of the 35S promoter that affected binding of OCSTF also affected transcription. Three different sequence alterations were introduced by oligonucleotide-directed mutagenesis (Figure 2a). The first mutation changed the sequence TAAG between the two repeats of the pentanucleotide elements to CCTC, while the second one changed the ACG trinucleotide of the 5′ repeat to TAT. The third mutation changed the sequence TG of the 5′ pentanucleotide element to GA. The effect of these changes on promoter activity was tested by transient expression assays in protoplasts. Mutant 1 reduced the expression of a linked GUS reporter gene by 76% compared to the level seen with the wild-type 35S promoter. Mutant 2 had the most dramatic effect on GUS activity, reducing it to background levels. Mutant 3 had wild-type activity (Figure 2a). The effect of the mutations on protein binding was assayed by gel retardation. The −89 to −59 fragments of the wild-type promoter and the three mutants were analysed. Binding to mutant probes 1 and 3 was reduced (Figure 2d, lanes 2 and 4). No upper band was detected with the probe from mutant 2 (Figure 2, lane 3) which is the mutant with lowest promoter activity.

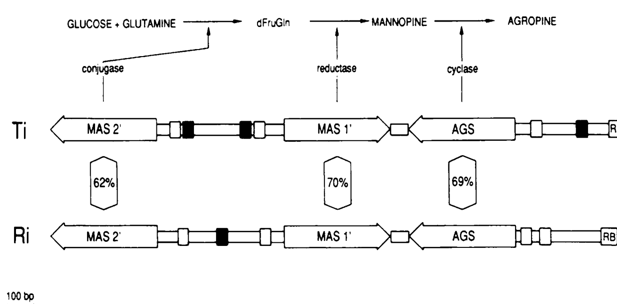

Fig. 3. The mannopine and agropine synthetic pathway is encoded by genes located in the TR-DNA of octopine Ti plasmids (Ti) and agropine Ri plasmids (Ri). Only the promoter regions are to scale. TATA boxes are indicated by open boxes and ocs-elements by black boxes. dFru-Gln is deoxy-fructosyl-glutamine. Homology between sequences of the coding regions are indicated.

From these experiments, it is evident that the 35S promoter contains a functional ocs-element and that this element is essential for the activity of the promoter in plant protoplasts. We have not analysed the function of the homologous sequences in FMV and CERV, but several facts point to their probable role as promoter elements. First, there is a high degree of sequence similarity between the elements of these viruses and the one from CaMV and, secondly, these three elements occur in a region of extensive sequence homology between the three viruses that has been proposed as the promoter region for the major transcripts of FMV and CERV (Richins *et al.*, 1987).

Ocs-elements in the genes for mannopine and agropine synthesis from the T-DNA of Ri and Ti plasmids

The agropine family of opines is synthesized by enzymes encoded by octopine Ti plasmids and agropine Ri plasmids. Three genes are involved in agropine biosynthesis on the T-DNA of octopine Ti plasmids (Ellis *et al.*, 1984; Salomon *et al.*, 1984) (Figure 3). The first two enzymes in the pathway, leading to mannopine formation, are encoded by the 2′ and 1′ genes which are divergently transcribed; here we refer to them as *MAS2′* and *MAS1′*. The third enzyme in the pathway is encoded by the 0′ transcript and converts mannopine to its cyclic form, agropine; we refer to this gene as the agropine synthase or *AGS* gene. The same biosynthetic pathway is encoded by agropine type Ri plasmids (Petit *et al.*, 1983), and the same genetic arrangement occurs (Figure 3). However, the homology between the genes is restricted to the coding regions (D.Bouchez, unpublished results).

Interaction between OCSTF and the sequences homologous to the ocs-element identified in the mannopine and agropine synthase genes was studied by gel-retardation assays (Figure 4). Two retarded bands were observed in experiments involving a 125 bp fragment containing the ocs-element homologous sequences from the promoters of the *MAS1′* and *MAS2′* genes of the Ri plasmid (Figure 4b, lane 1). The specificity of this interaction was demonstrated using competitor DNA fragments containing the ocs-element of the *OCS* promoter (Figure 4b, lanes 3 and 4). These fragments competed with the formation of both retarded complexes. Conversely, the 125 bp fragments of the *MAS*

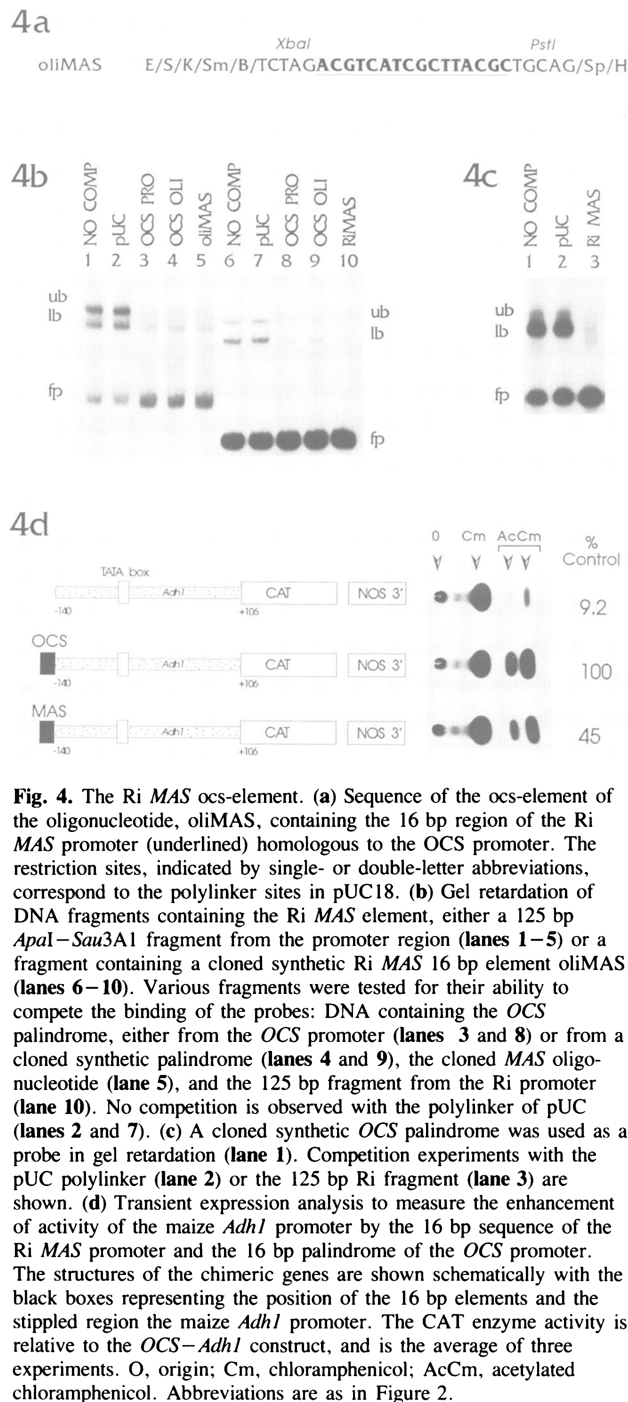

Fig. 4. The Ri *MAS* ocs-element. (**a**) Sequence of the ocs-element of the oligonucleotide, oliMAS, containing the 16 bp region of the Ri *MAS* promoter (underlined) homologous to the OCS promoter. The restriction sites, indicated by single- or double-letter abbreviations, correspond to the polylinker sites in pUC18. (**b**) Gel retardation of DNA fragments containing the Ri *MAS* element, either a 125 bp *Apa*I−*Sau*3A1 fragment from the promoter region (**lanes 1−5**) or a fragment containing a cloned synthetic Ri *MAS* 16 bp element oliMAS (**lanes 6−10**). Various fragments were tested for their ability to compete the binding of the probes: DNA containing the *OCS* palindrome, either from the *OCS* promoter (**lanes 3** and **8**) or from a cloned synthetic palindrome (**lanes 4** and **9**), the cloned *MAS* oligonucleotide (**lane 5**), and the 125 bp fragment from the Ri promoter (**lane 10**). No competition is observed with the polylinker of pUC (**lanes 2** and **7**). (**c**) A cloned synthetic *OCS* palindrome was used as a probe in gel retardation (**lane 1**). Competition experiments with the pUC polylinker (**lane 2**) or the 125 bp Ri fragment (**lane 3**) are shown. (**d**) Transient expression analysis to measure the enhancement of activity of the maize *Adh1* promoter by the 16 bp sequence of the Ri *MAS* promoter and the 16 bp palindrome of the *OCS* promoter. The structures of the chimeric genes are shown schematically with the black boxes representing the position of the 16 bp elements and the stippled region the maize *Adh1* promoter. The CAT enzyme activity is relative to the *OCS−Adh1* construct, and is the average of three experiments. O, origin; Cm, chloramphenicol; AcCm, acetylated chloramphenicol. Abbreviations are as in Figure 2.

promoters competed with the binding of OCSTF to the ocs-element of the *OCS* gene (Figure 4c, lane 3).

The OCSTF binding site in the *MAS* promoters of the Ri plasmid was precisely identified as a 16 bp region homologous to the *OCS* promoter using a cloned oligonucleotide (Figure 4a). A DNA fragment containing the cloned oligonucleotide had the same binding properties as the larger 125 bp fragment (Figure 4b, lanes 6−10).

The enhancer activity of the 16 bp element from the Ri-plasmid was tested by transient expression analysis in *Nicotiana plumbaginifolia* protoplasts. This sequence was able to enhance the activity of the maize *Adhl* promoter to ~50% of the control construct containing the 16 bp palindrome of the *OCS* promoter (Figure 4d). The import-

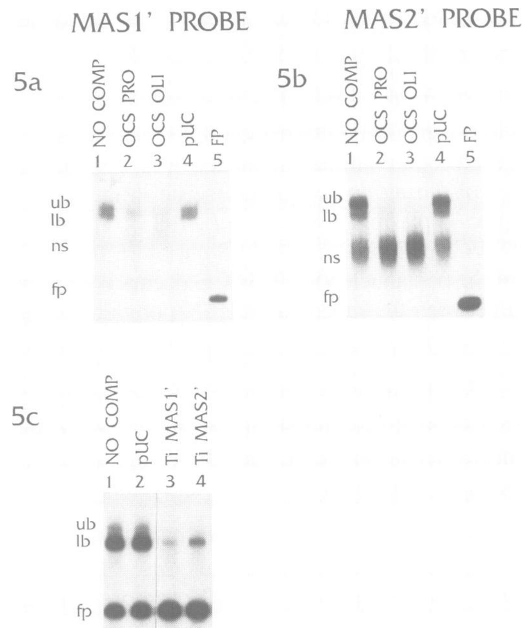

Fig. 5. Gel-retardation assays using the ocs-element of the *MAS1'* and *MAS2'* genes of Ti plasmids. (**a**) *MAS1'* probe with no DNA competitor fragment (**lane 1**), *OCS* promoter competitor (**lane 2**), *OCS* 16 bp palindrome competitor (**lane 3**), pUC polylinker competitor (**lane 4**). (**b**) *MAS2'* probe. In (a) and (b) all free probe was complexed by either specific or non-specific binding activities. (**c**) *OCS* palindrome probe with no competitor (**lane 1**), pUC polylinker competitor (**lane 2**), Ti *MAS1'* fragment (**lane 3**) and Ti *MAS2'* fragment (**lane 4**). ns, a binding activity detected with the Ti *MAS1'* and 2' probes that was not associated with the ocs-element. Other abbreviations are as in Figure 2.

ance of this region of the Ri-plasmid for activity of the divergent *MAS1'* and *MAS2'* promoters was indicated by expression studies in plant tumour cells. Deletion analysis of the *MAS2'* promoter showed that a 70 bp region containing this element is essential for activity, giving a 10-fold decrease in promoter expression when deleted, and similar results were obtained with the *MAS1'* promoter (D.Bouchez, in preparation). These promoters therefore share a common element that binds OCSTF. This element has enhancer-like properties and occurs in a region that is important for the function of the promoters.

The arrangement of the sequences homologous to the ocs-element in the divergent mannopine synthase genes on the T-DNA of the octopine type Ti plasmid pTiAch5 is different from the Ri plasmid; two elements occur in the promoter region of these genes (Figure 3). The *MAS1'* and *MAS2'* sequences are identical to the 16 bp palindrome of the *OCS* promoter at 9 out of 16 bases and 10 out of 16 bases respectively and occur within 40 bases of the TATA boxes (Figure 1). DNA fragments containing each of the two regions of homology were used as probes or binding competitors in gel-retardation experiments. Both fragments competed with the binding of OCSTF to the *OCS* 16 bp palindrome (Figure 5c, lanes 3 and 4) and gave two retarded bands (Figure 5a and b, lane 1). Binding was competed by fragments containing the *OCS* palindrome (Figure 5a and b, lanes 2 and 3). Although a precise mutational analysis of these promoters remains to be done, it is very likely that these binding sites play a major role in the promoters of these genes.

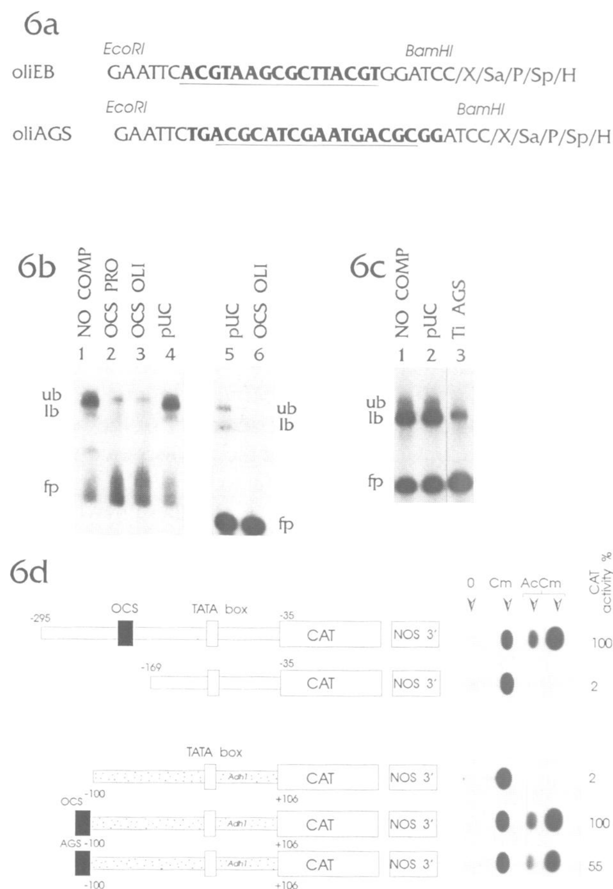

Fig. 6. The ocs-element of the Ti plasmid *AGS* gene. (**a**) The cloned synthetic AGS 19mer, oliAGS, and the cloned *OCS* 16 bp palindrome, oliEB. The homologous 16 bp regions are underlined. The restriction sites, indicated by single or double letters, correspond to the polylinker sites of pUC18. (**b**) Gel retardation: 64 bp *Hae*III–*Alu*I probe from the *AGS* promoter (**lanes 1–4**), *Eco*RI–*Hin*dIII probe of oliAGS (**lanes 5** and **6**). Competitors were *OCS* promoter fragment (**lane 2**), *OCS* 16 bp palindrome (**lanes 3** and **6**) and pUC polylinker (**lanes 4** and **5**). (**c**) Competition of binding to *OCS* 16 bp palindrome probe—no competition (**lane 1**), pUC polylinker (**lane 2**) and 64 bp *Hae*III–*Alu*III fragment of *AGS* promoter. (**d**) Transient expression analyses of deletions of the *AGS* promoter and *Adh1* promoter reactivation using oliEB (OCS 16 bp palindrome) and oliAGS (*AGS* 19mer). The CAT enzyme activity relative to the −205 *AGS* construct or the *OCS*–*Adh1* construct are indicated and are the average of three experiments. O, origin; Cm, chloramphenicol; AcCm, acetylated chloramphenicol. The structure of the chimeric genes and deletion end-points are indicated schematically. The *AGS* promoter is indicated with open boxes (numbering relative to the translation start) and the *Adh1* promoter is stippled (numbering relative to the transcription start).

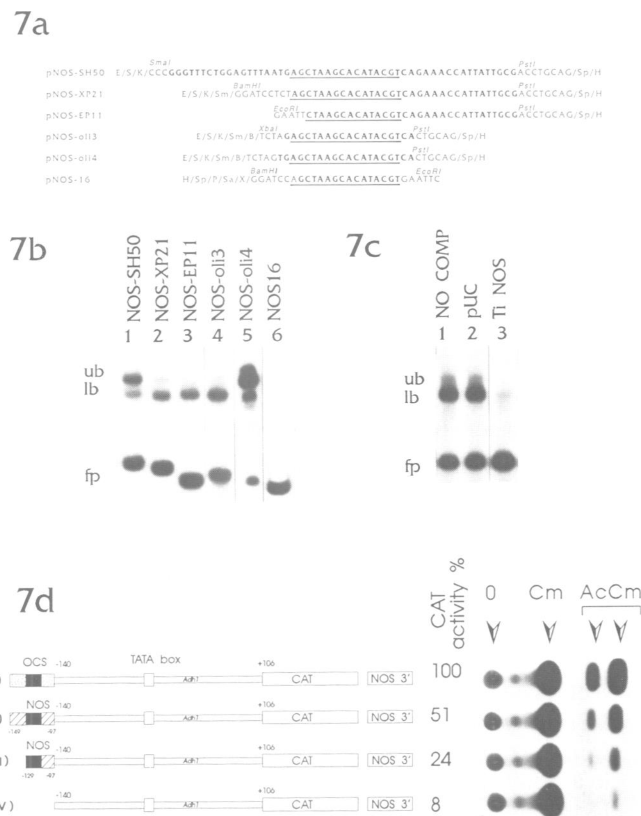

Fig. 7. The ocs-element in the *NOS* promoter. (**a**) Promoter fragments used for gel-retardation and expression studies. Restriction sites, indicated by single or double letters, correspond to the polylinker sites of pUC18. (**b**) Gel retardation of probes made from the fragments shown in (a). (**c**) Competition study using an *OCS* 16 bp palindrome probe with no competitor (**lane 1**), pUC polylinker (**lane 2**) and pNOS-SH50 (**lane 3**). (**d**) Transient expression analysis showing activation of a maize *Adh1* promoter by the pNOS-SH50 fragment (construct ii), the pNOS-XP21 fragment (construct iii) and a 48 bp *OCS* promoter fragment containing the 16 bp palindrome (construct i). Construct (iv) is a negative control. Black boxes are the ocs-elements, striped boxes are flanking *NOS* promoter sequences and stippled boxes are the flanking *OCS* promoter sequences. Numbers above the constructs are the co-ordinates of the *Adh1* promoter and numbers below the line are the co-ordinates of the *OCS* or *NOS* promoter fragments. Numbering is relative to start of transcription of the *OCS*, *NOS* and maize *Adh1* genes. CAT activity, expressed relative to the *OCS* enhancer construct, is the average of four experiments. O, origin; Cm, chloramphenicol; AcCm, acetylated chloramphenicol.

The ocs-element homology in the *AGS* gene of the Ti plasmid was studied in more detail for both binding to OCSTF and *in vivo* transcriptional activity to test whether the low sequence similarity (50%) with the palindrome of the *OCS* promoter could have a functional significance. A 64 bp fragment containing this region of the *AGS* promoter and a synthetic 19mer containing this site competed the binding of OCSTF to the 16 bp palindrome of the *OCS* promoter (Figure 6c). Both *AGS* promoter probes bound to OCSTF and produced double band gel-retardation patterns (Figure 6b).

Two 5′-deletion mutants of the *AGS* promoter, linked to a CAT-NOS3′ cassette, were tested for transient expression in plant protoplasts. The 5′ end-points of the deletions were −295 and −169. While the −295 deletion gave high expression, the −169 promoter gave background levels, showing the presence of an essential promoter element in the −295 to −169 region (Figure 6d); the ocs-element is located at positions −209 to −194. Furthermore, when the *AGS* synthetic oligonucleotide was inserted in front of an inactive mutant of the maize *Adh*1 promoter linked to a *CAT*–*NOS3′* reporter cassette, transient expression analysis revealed that this sequence enhances the expression of the reporter gene (Figure 6d). It is therefore clear that the promoter of the agropine synthase gene of pTiAch5 contains a functional ocs-element that interacts specifically with OCSTF *in vitro* and activates transcription *in vivo*.

An ocs-element in the nopaline synthase promoter

The *NOS* promoter contains a sequence that matches the *OCS* 16 bp palindrome at 12 out of 16 sites. However, a synthetic copy of this sequence had no enhancer activity when linked

to an inactive *Adh1* promoter (Ellis *et al.*, 1987a). When this sequence was excised from pNOS16, the DNA fragment neither bound OCSTF (Figure 7b, lane 6) nor competed with the binding of OCSTF to the 16 bp palindrome of the *OCS* promoter (data not shown). However, when a larger 50 bp fragment (−149 to −97) of the *NOS* promoter that contains this region was used as a probe, a two-band retardation pattern resulted (Figure 7b, lane 1). The same 50 bp fragment competed the binding of OCSTF to the 16 bp palindrome of the *OCS* promoter (Figure 7c, lane 3). In addition, binding of the 50 bp *NOS* promoter fragment was competed by DNA containing the 16 bp palindrome (data not shown). These experiments clearly demonstrate that the *NOS* promoter and the *OCS* promoter interact with the same nuclear factor. In the case of the *NOS* promoter, however, the 16 bp region that is homologous to the *OCS* promoter is not sufficient for binding. Therefore to define precisely the extent of the *NOS* promoter sequence needed to give a double-band binding pattern, several restriction fragments and synthetic oligonucleotides containing smaller fragments of the *NOS* promoter were used as probes and competitors of binding in gel-retardation experiments. The fragment from pNOS-XP21 (Figure 7a) contained the *NOS* promoter sequence from −129 to −97 and the fragment from pNOS-EP11 (Figure 7a) contained sequences from −127 to −97. Deletion of the sequences between −149 and −129 resulted in a major reduction in the upper band (Figure 7b, lane 2). Deletion of the sequences between −149 and −127 eliminated the upper band but the lower band was still formed (Figure 7b, lane 3). Both fragments competed with binding of an OCS promoter probe to OCSTF (data not shown).

The smallest fragment of the *NOS* promoter that we examined and that competed with the interaction between OCSTF and the *OCS* promoter (data not shown), and that gave lower band binding in gel retardation (Figure 7b, lane 4), was derived from pNOS-oli3 (Figure 7a) which contained 19 bp of NOS promoter (−130 to −112). This extended the 16 bp in pNOS16 by 2 bp at the 3′ end and by 1 bp on the 5′ side. When a fragment from pNOS-oli4 (Figure 7a) which contained one more base on the 5′ side and contained 20 bp of *NOS* promoter sequence (−131 to −112) was used as a probe, there was a double-band binding pattern equivalent to that observed with the 50 bp probe of pNOS-SH50 (Figure 7b, lane 5).

The importance of the DNA sequences in this region for the *NOS* promoter activity has been demonstrated by deletion analysis of the *NOS* promoter (An *et al.*, 1986; Ebert *et al.*, 1987; Mitra and An, 1989). We compared sequences from the *NOS* promoter (−149 to −97 and −129 to −97) to the sequence from the *OCS* promoter (−206 to −171), which contains the 16 bp palindrome, for the ability to enhance the inactive *Adh1* promoter used previously in this paper and by Ellis *et al.* (1987a). The *NOS* sequences between −149 and −97 and between −129 and −97 had 51 and 24% respectively of the enhancing activity of the ocs-enhancer (Figure 7d). The difference in enhancement by the *OCS* promoter fragment and the 50 bp *NOS* promoter fragment may reflect differences in *in vivo* affinity of the two sequences for OCSTF, spacing differences between the ocs-element and the TATA box or both.

Discussion

The ocs-element, which was first identified in the promoter of the T-DNA gene encoding octopine synthase, has now been identified in the promoters of six other opine synthase genes and of three caulimoviruses. The basis of the identification of these new ocs-element locations was sequence similarity and, in most cases (the CERV and FMV sequences being the exception), the protein-binding properties of the sequences. The six elements tested by gel retardation all showed the two DNA−protein complexes characteristic of the functionally active ocs-element identified in the promoter of the *OCS* gene (Singh *et al.*, 1989). Protein binding to all of the new sequences was specifically competed by an excess of DNA containing the ocs-element of the octopine synthase gene. Conversely, each of the new ocs-element sites competed with the binding of protein to the ocs-element from the octopine synthase promoter gene. All these newly identified promoter sequences and the *OCS* promoter therefore bind a common protein complex that has been called OCSTF (J.G.Tokuhisa *et al.*, submitted) and that functions as a transcription factor (Singh *et al.*, 1989).

Singh *et al.* (1989) demonstrated that for the ocs-element found in the *OCS* gene, a strong correlation existed between the presence of the higher mol. wt protein−DNA complex ('upper band') in gel-retardation assays and the activity of the promoter in both maize and tobacco protoplasts. In this paper, the importance of the upper band for transcription has been extended to the ocs-element that occurs in the 35S promoter. Mutant 2 (Figure 2) inactivated the promoter in tobacco protoplasts and eliminated the upper band complex in gel-retardation assays. Mutant 3 decreased the amount of upper band complex (Figure 2), but it is not clear why this mutation did not affect transcription. One possibility is that the *in vitro* binding conditions are more stringent than occur *in vivo*.

A comparison of the ocs-element regions of the 10 promoters (Figure 1) shows that sequence similarity extends over a region of 20 bp and that the 16 bp palindrome of the *OCS* promoter forms the core of the 20 bp element. The conclusion that the ocs-element is comprised of a 20 bp sequence is supported by results from DNA footprinting experiments on the *OCS* promoter which reveal a protected region of 20 to 22 bases (J.G.Tokuhisa *et al.*, in preparation) and also by the fact that a 20 bp sequence was the minimal sequence from the *NOS* promoter that could form the slower migrating upper band complex in gel-retardation experiments.

A 20 bp consensus sequence for the ocs-element was derived from the alignment of the 10 sequences (Figure 1). The consensus can be viewed in several ways: as a 20 bp palindrome with two 10 bp half sites or as two direct 8 bp direct repeats separated by 4 bases. These 8 bp repeats themselves have a palindromic structure (Figure 1) and are very similar to the cyclic AMP receptor element consensus TGACGTCA of mammalian genes (reviewed by Roesler *et al.*, 1988). The 20 bp palindrome is an extension of the 16 bp palindrome in the promoter of the *OCS* gene. The 8 bp direct repeats overlap the pentanucleotide direct repeats of the 35S promoter noted by Fang *et al.* (1989). The dis-

tinction between recognition of direct repeats versus inverted repeats could be of critical importance in trying to understand the way OCSTF binds to the ocs-element and how the monomer units interact with each other to produce the higher mol. wt complex that is associated with transcriptional activity.

A considerable degree of sequence divergence occurs between the 10 ocs-element sequences in Figure 1. The tolerance of sequence variability is consistent with the saturation mutagenesis study of Singh *et al.* (1989) where it was demonstrated that the majority of single- and double-base substitutions within the 16 bp palindrome of the *OCS* promoter had only a small effect on enhancing activity of the sequence. Although divergence occurs between the different 20 bp sequences, a high degree of conservation is observed at positions 3, 4, 5 and 13, 15, 16, 17. Mutation of ACG to TAT at positions 3–5 had a drastic effect on activity of the 35S promoter, indicating that this motif is of critical importance in this context. In the *NOS* promoter, however, this region is AGC. Here the divergence from the conserved ACG motif is possibly compensated for by the consensus TG dinucleotide at positions 1 and 2 of the ocs-element. This explanation is supported by the effect of mutations or deletions of the 5′ TG dinucleotide at positions 1–2 of the *NOS* sequence that profoundly affect the stability of the DNA–protein complex and almost completely eliminate the higher mol. wt complex in gel-retardation assays (Figure 7).

The existence of plant transcription factors in both monocot and dicot plants that interact with the ocs-element of genes introduced into plants by parasites implies that there are cellular counterparts of this element involved in the regulation of plant genes. What are these genes? We searched the GenBank and EMBL DNA sequence databases for the 20 bp consensus sequence allowing four mismatches. One new sequence of interest was G**G**ACGCAAGGACG**G**ACGGAC occurring ~350 bp 5′ of the transcription initiation site on the non-coding strand of the maize gene, sucrose synthase (Werr *et al.*, 1985). Three mismatches (indicated by bold type) occur between this and the consensus sequence. There are no data available yet that indicate the involvement of this sequence in promoter activity. The shortage of cellular gene candidates containing the ocs-element may reflect the limited number of plant genes in the databases. Alternatively, the ocs-element may have a more flexible sequence requirement than allowed by our search criteria. It is clear that some alternative method to searching the DNA sequence database may be required to identify the plant cellular genes that are regulated by the ocs-element and the OCSTF transcription factor. Another sequence from *Agrobacterium* was also found in our database search. The sequence, with three mismatches, TGACG**AGT**GCAATGACGTCA, was found on the non-coding strand of the *RolA* gene of *Agrobacterium rhizogenes* strain HR1 (ORF 10, Slightom *et al.*, 1986), ~480 bp upstream of the initiation of translation. We have no data yet as to whether this sequence acts as a promoter element in the *RolA* gene.

The expression of opine synthase genes of *Agrobacterium* sp. and of the 35S transcript of the caulimoviruses are key steps in the successful parasitism of plant cells by these parasites. The opines can be considered as the energy source that has driven the evolution of the crown-gall and hairy root systems. In the case of the caulimoviruses, the 35S transcript is the substrate of the virally encoded reverse transcriptase and is essential for viral replication. It is remarkable that the two types of genetic parasite, on one hand a bacterium and on the other hand a DNA virus, have converged on the same *cis*-acting promoter element and host transcription factor to mediate key transcriptional steps in their life cycles.

Materials and methods

Recombinant DNA techniques and gene constructs

Recombinant plasmids were constructed using standard techniques described in Maniatis *et al.* (1982). Site-directed mutagenesis by oligonucleotides was performed according to Kunkel (1985). All oligonucleotides were synthesized using an Applied Biosystems 380A DNA synthesizer. Single-stranded oligonucleotides were cloned using the method of Derbyshire *et al.* (1986), or both strands were synthesized, annealed and cloned as described in Ellis *et al.* (1987a). For expression analysis, synthetic oligonucleotides and promoter fragments were cloned 5′ of the −100 *Adh-cat-nos* 3′ cassette of pACN-100 (Ellis *et al.*, 1987a) or in the case of the RiMAS oligonucleotide and *NOS* promoter fragments, the −140 *Adh-cat-nos* 3′ cassette of pAdcat2 (Ellis *et al.*, 1987b). The AGS promoter vectors contained DNA from the *Acc*I site at −35 to the *Hae*III site at −169 or to the *Acc*I site at −295. These co-ordinates are with respect to the A residue of the translation initiation codon. The transcript start has not been mapped. The 35S-GUS-nos 3′ expression plasmid used for mutagenesis was kindly provided by Dr John Walker (University of Missouri).

Gel-retardation assays

Nuclear protein extracts from maize cell suspension cultures were prepared according to Green *et al.* (1987). The binding reactions typically contained 5000 c.p.m. of the end-labelled probe (0.05–0.2 ng), 1.0 μg of poly(dI.dC), 5 μg of crude nuclear protein extract in binding buffer (25 mM HEPES, 50 mM KCl, 1 mM EDTA, 10% glycerol, 5 mM 2-mercaptoethanol) and the gel retardation was performed as described by Garner and Revzin (1981) and Fried and Crothers (1981). Competitor DNA was added to the reaction in 100- to 300-fold molar excess. The relative proportions of lower and upper band retarded complexes involving the ocs-element is proportional to the amount of input binding activity. Greater binding activity increases the amount of upper band complex. This is due to the fact that the same protein is involved in both complexes (J.G.Tokuhisa *et al.*, in preparation). This accounts for some of the inter-experiment variability in proportions of upper and lower band complexes when different extract preparations are used.

Fragments used for probes and competitors in gel-retardation assays were eluted from 10% acrylamide gels. Probes were end-labelled using Klenow enzyme and [^{32}P]dCTP. The *OCS* promoter fragment was either excised with *Eco*RI and *Hin*dIII as a 52 bp fragment from pUC19pal16 (Ellis *et al.*, 1987a) which contained a cloned synthetic copy of the 16 bp palindrome or excised as a 48 bp fragment with *Xho*I and *Pst*I from a pUC8 clone. This fragment contained bases −206 to −171 (*Xho*I to *Taq*I of Del5′206, Ellis *et al.*, 1987a). The 31 bp fragment of the 35S promoter was excised with *Eco*RV (−89) and *Fok*I (−59). The 125 bp fragment of the Ri MAS promoter was excised from a larger clone of the promoter region with *Apa*I and *Sau*3A1. The complete sequence of this region will be published elsewhere (D.Bouchez, in preparation). A smaller fragment of this promoter was derived from a synthetic 16 bp oligonucleotide (Figure 4) cloned in pUC118 and excised as a 60 bp fragment with *Eco*RI and *Hin*dIII. The *MAS1′* and *MAS2′* promoter fragments of the Ti-plasmid were excised from a larger fragment cloned from pTiAch5 (de Vos *et al.*, 1981; Velten *et al.*, 1984). The *MAS2′* fragment (137 bp) was excised from the promoter using *Cla*I and *Hin*pI. The *MAS1′*fragment (145 bp) was excised from the promoter with *Hae*III, cloned into the *Sma*I site of pUC118 then reisolated as a *Bam*HI–*Eco*RI fragment. The AGS promoter fragment (64 bp) was excised from a larger fragment (*Eco*RI fragment 20, de Vos *et al.*, 1981; Barker *et al.*, 1987) of the T-region of pTiAch5 with *Hae*III and *Alu*I, cloned into pUC118 and reisolated as a 86 bp *Eco*RI–*Bam*HI fragment. A fragment

(54 bp) containing a 19 bp synthetic copy of the ocs-element of the *AGS* gene was excised from pUC19 with *Eco*RI and *Hin*dIII. The *NOS* promoter fragments were isolated from the T-DNA gene or from pUC19 containing various synthetic copies of the NOS promoter region (pNOS-oli3, pNOS-oli4 and pNOS16). pNOS-SH50 contained the *Sac*II (−149) to *Hin*PI (−97) fragment of the NOS promoter. pNOSXP21 contained the *Alu*I (−127) to *Pst*I fragment of pNOS-SH50 cloned into the *Pst*I and filled *Xba*I sites of pUC19. pNOS-EP11 contained the same fragment cloned into the *Pst*I and filled *Eco*RI site of pUC19.

Transient expression analysis

Nicotiana plumbaginifolia protoplasts were isolated from the cell suspension culture NpT5 as previously described (Llewellyn *et al.*, 1987). Protoplasts were electroporated using 10 μg of plasmid DNA. After 24 h incubation at 25°C, they were lysed by sonication and the protein concentration in the extracts was measured.

CAT activity was assayed as previously described (Llewellyn *et al.*, 1987). The conversion of [^{14}C]chloramphenicol to acetylated chloramphenicol was determined by cutting the substrate and acetylated product from TLC plates and counting by liquid scintillation. GUS activity was measured according to Jefferson *et al.* (1987) using the fluorogenic substrate 4-methyl umbelliferyl glucuronide.

Acknowledgements

We wish to acknowledge the help of Professor M.Ronis, First Secretary (Science and Technology), Ambassade de France, Canberra for his help in facilitating the visit of D.B. to Australia. This research has been supported by Lubrizol Corporation.

References

An,G., Ebert,P.R., Yi,B.-Y. and Choi,C.-H. (1986) *Mol. Gen. Genet.*, **203**, 245−250.

An,G.,Costa,M.A., Mitra,A., Ha,S.B. and Marton,L. (1988) *Plant Physiol.*, **81**, 301−305.

Barker,R.F., Idler,K.B., Thompson,D.V. and Kemp,J.D. (1983) *Plant Mol. Biol.*, **2**, 335−350.

Bevan,M.W., Flavell,R.B. and Chilton,M.-D. (1983) *Nature*, **304**, 184−187.

Cavener,D.R. (1987) *Nucleic Acids Res.*, **15**, 1353−1361.

DeGreve,H., Dhaese,P., Seurinck,J., Lemmers,M., Van Montagu,M. and Schell,J. (1982) *J. Mol. Appl. Genet.*, **1**, 499−510.

Depicker,A., Stachel,S., Dhaese,P., Zambryski,P. and Goodman,H.M. (1982) *J. Mol. Appl. Genet.*, **1**, 561−573.

Derbyshire,K.M., Salvo,J.J. and Grindley,N.D.F. (1986) *Gene*, **46**, 145−152.

De Vos,G., De Beuckeker,M., Van Montagu,M. and Schell,J. (1981) *Plasmid*, **6**, 249−253.

Ebert,P.R., Ha,S.B. and An,G. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5745−5749.

Ellis,J.G., Ryder,M.H. and Tate,M.E. (1984) *Mol. Gen. Genet.*, **195**, 466−473.

Ellis,J.G., Llewellyn,D.J., Walker,J.C., Dennis,E.S. and Peacock,W.J. (1987a) *EMBO J.*, **6**, 3202−3208.

Ellis,J.G., Llewellyn,D.J., Dennis,E.S. and Peacock,W.J. (1987b) *EMBO J.*, **6**, 11−16.

Fang,R.-X., Nagy,F., Sivasubramaniam,S. and Chua,N.-H. (1989) *The Plant Cell*, **1**, 141−150.

Franck,A., Guilley,H., Jonard,G., Richards,K. and Hirth,L. (1980) *Cell*, **21**, 285−294.

Fried,M.G. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6506−6525.

Garner,M.M. and Revzin,A. (1981) *Nucleic Acids Res.*, **9**, 3047−3060.

Gidoni,D., Brosio,P., Bond-Nutter,D., Bedbrook,J. and Dunsmuir,P. (1989) *Mol. Gen. Genet.*, **215**, 337−344.

Green,P.J., Kay,S.A. and Chua,N.-H. (1987) *EMBO J.*, **6**, 2543−2549.

Green,P.J., Yong,M.-H., Cuozzo,M., Kano-Murakami,Y., Silverstein,P. and Chua,N.-H. (1988) *EMBO J.*, **7**, 4035−4044.

Hull,R., Sadler,J. and Longstaff,M. (1986) *EMBO J.*, **5**, 3083−3090.

Jefferson,R.A., Kavanagh,T.A. and Bevan,M.W. (1987) *EMBO J.*, **6**, 3901−3907.

Kunkel,T.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 488−492.

Llewellyn,D.J., Finnegan,E.J., Ellis,J.G., Dennis,E.S. and Peacock,W.J. (1987) *J. Mol. Biol.*, **195**, 115−123.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mitra,A. and An,G. (1989) *Mol. Gen. Genet.*, **215**, 294−299.

Odell,J.T., Nagy,F. and Chua,N.-H. (1985) *Nature*, **313**, 310−312.

Ow,D., Jacobs,J.D. and Howell,S.H. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4870−4874.

Petit,A., David,C., Dahl,G., Ellis,J.G., Guyon,P., Casse-Delbart,F. and Tempé,J. (1983) *Mol. Gen. Genet.*, **190**, 204−214.

Prat,S., Willmitzer,L. and Sanchez-Serrano,J.J. (1989) *Mol. Gen. Genet.*, **217**, 209−214.

Richins,R.D., Scholthof,H.B. and Shepherd,R.J. (1987) *Nucleic Acids Res.*, **15**, 8451−8466.

Roesler,W., Vanderbank,G.R. and Hanson,R.W. (1988) *J. Biol. Chem.*, **263**, 9063−9066.

Rogers,S.G., Klee,H.J., Horsch,R.B. and Fraley,R.T. (1987) *Methods in Enzymol.*, **153**, 253−257.

Salomon,F., Deblaere,R., Leemans,J., Hernalsteens,J.P., Van Montagu,M. and Schell,J. (1984) *EMBO J.*, **3**, 141−146.

Schulze-Lefert,P., Dangl,J.L., Becker-Andre,M., Hahlbrock,K. and Schulz,W. (1989) *EMBO J.*, **6**, 651−656.

Singh,K., Tokuhisa,J.G., Dennis,E.S. and Peacock,W.J. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 3733−3737.

Slightom,J., Durand-Tardif,M., Jouanin,L. and Tepfer,D. (1986) *J. Biol. Chem.*, **261**, 108−121.

Teeri,T.H., Lehvaslaiho,H., Franck,M., Uotila,J., Heino,P., Plava,E.T., Van Montagu,M. and Herrera-Estrella,L. (1989) *EMBO J.*, **8**, 343−350.

Velten,J., Hain,R. and Schell,J. (1984) *EMBO J.*, **3**, 2723−2730.

Walker,J.C., Howard,E.A., Dennis,E.S. and Peacock,W.J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6624−6628.

Werr,W., Frommer,W.-B., Maas,C. and Starlinger,P. (1985) *EMBO J.*, **4**, 1373−1380.

Received on August 24, 1989; revised on September 26, 1989