The EMBO Journal vol.6 no.5 pp.1189–1195, 1987

# Structure and elicitor or u.v.-light-stimulated expression of two 4-coumarate:CoA ligase genes in parsley

**Carl Douglas, Heidi Hoffmann, Wolfgang Schulz and Klaus Hahlbrock**

Max-Planck-Institut für Züchtungsforschung, D-5000 Köln 30, FRG

Communicated by H.Saedler

**We have isolated genomic clones encoding 4-coumarate:CoA ligase (4CL), a key enzyme of general phenylpropanoid metabolism, and have analysed the structure and regulation of the genes contained on these clones. Restriction enzyme and sequence analysis indicated that two distinct 4CL genes, Pc4CL-1 and Pc4CL-2, are represented on the clones and that additional 4CL genes are not present in parsley. Two lines of evidence suggest that each gene is transcriptionally activated by both elicitor and u.v. irradiation: cDNA clones corresponding to each gene were found in cDNA libraries made with RNA from both elicitor-treated and u.v-irradiated cells, and run-off transcripts homologous to a Pc4CL-2-specific intron probe were induced by both treatments. This induction was about half of the induction measured using probes homologous to both genes. The transcription initiation sites of both genes were determined. Comparison of the nucleotide sequences of the two genes 5′ to these sites showed that they are highly homologous for several hundred base pairs and that they contain features potentially involved in regulation by elicitor and u.v. irradiation.**

*Key words:* gene activation/promoter region/phenylpropanoid metabolism/fungal elicitor

## Introduction

Inducible defense responses are important mechanisms by which plants defend themselves against a variety of environmental stresses. Among the best studied of these responses is the induced accumulation of secondary metabolites, which may provide protection against the applied stress. For example, u.v. irradiation leads to the accumulation of u.v. protective flavonoids in plants (Hahlbrock *et al.*, 1982), while pathogen infection often induces the accumulation of phytoalexins (plant antimicrobial compounds) and the deposition of lignin and other phenolic compounds in plant cell walls (Hahlbrock *et al.*, 1982; Ride, 1983). The latter responses play well-documented roles in plant defense against pathogens (Hahlbrock and Scheel, 1987) and can be induced in the absence of pathogens by elicitors, low mol. wt compounds derived from fungal cell walls or culture filtrates (Dixon, 1986).

Many plant secondary metabolites (e.g. isoflavonoid and other phenolic phytoalexins, flavonoid glycosides, lignin and related phenolic compounds), whose accumulation is induced by stress, are phenylpropanoid derivatives (Hahlbrock *et al.*, 1982). Thus, several recent studies of the molecular mechanisms underlying induced defense responses have focussed on the stimulation of phenylpropanoid metabolism in plant cells. The coenzyme A esters of 4-coumarate and other cinnamate derivatives are central intermediates in the synthesis of many phenylpropanoid compounds in higher plants (Hahlbrock and Grisebach, 1979). These thiol esters are synthesized by 4-coumarate:CoA ligase (4CL), which has a pivotal role at the branch point between general phenylpropanoid metabolism and pathways leading to the synthesis of various specific products. In cultured parsley (*Petroselinum crispum*) cells, the differential induction of flavonoids by u.v. light and several coumarin derivatives [most notably furanocoumarins, the phytoalexins of parsley (Knogge *et al.*, 1987)] by fungal elicitor, is associated in each case with rapid, transient activation of 4CL gene transcription (Chappell and Hahlbrock, 1984) and subsequent transient increases in steady-state 4CL mRNA levels (Kreuzaler *et al.*, 1983; Kuhn *et al.*, 1984) and enzyme activities (Hahlbrock *et al.*, 1981; Ragg *et al.*, 1981). The transcriptional activation of 4CL following either elicitor treatment or u.v. irradiation occurs coordinately with the activation of phenylalanine ammonia-lyase (PAL), the first enzyme of general phenylpropanoid metabolism (Chappell and Hahlbrock, 1984). Chalcone synthase (CHS), the first specific enzyme in the branch pathway leading to flavonoid synthesis, is transcriptionally induced by u.v. light but not elicitor (Chappel and Hahlbrock, 1984). Similar findings have been reported in cultured bean cells, where PAL and CHS are both involved in phytoalexin synthesis and are transcriptionally induced following elicitor treatment (Cramer *et al.*, 1985; Edwards *et al.*, 1985).

Because of its key position in phenylpropanoid metabolism and its inducibility by both fungal elicitor and u.v. irradiation, we have chosen 4CL for further studies on the mechanism of stress-induced transcriptional activation of plant genes. We report here studies on the structure and regulation of individual 4CL genes in parsley. We demonstrate that parsley contains two genes encoding 4CL and present evidence suggesting that each gene is transcriptionally activated by both fungal elicitor and u.v. irradation. Comparison of the sequences 5′ to the transcription start sites of the genes reveals a high degree of homology between the genes for several hundred base pairs, suggesting that sequences involved in the transcriptional activation of the genes are located here.

## Results

### *Genomic clones and copy number of 4CL genes*

The number and organization of 4CL genes in parsley was established using molecular cloning and filter hybridization techniques. Genomic clones were isolated from a library constructed in λ-EMBL 4 using a 450-bp 4CL cDNA (Kuhn *et al.*, 1984) as a hybridization probe. The structures of three representative 4CL clones analysed in detail are shown in Figure 1. All three clones contain complete copies of 4CL genes, as determined by DNA hybridization of various restriction digests probed with a near full-length 4CL cDNA of 1700 bp. The approximate 5′ and 3′ ends of the transcribed areas were located using subcloned probes from the two ends of the cDNA. Two clones (λ-Pc23 and λ-Pc32) were very similar, having only minor restriction site differences 3 kb or more distal to the 3′ transcription termini. Portions of these two clones sequenced were identical (see below),

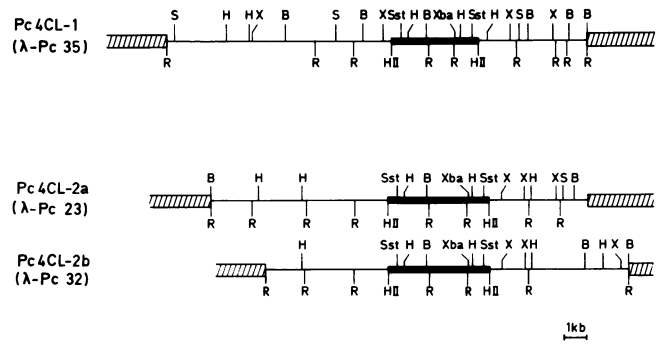


**Fig. 1.** Genomic clones of 4CL in λ-EMBL 4. The restriction maps of three representative clones, λ-Pc35, λ-Pc23 and λ-Pc32, are shown. The large (left) and small (right) arms of the cloning vector are indicated by hatched boxes. Approximate locations of transcribed areas within the cloned genes (Pc4CL-1, Pc4CL-2a and Pc4CL-2b), as determined by hybridization of a 4CL cDNA to genomic restriction fragments, are shown by solid boxes. The cDNA used for these hybridization is 1700 bp in length, ~300 bp shorter than the expected mRNA length (Kuhn *et al.*, 1984). Hybridization of 5′ and 3′ probes derived from this cDNA to genomic fragments indicated that the genes are transcribed from left to right as they are represented in the figure. Symbols: B, *Bam*HI; H, *Hin*dIII; HII, *Hin*cII; R, *Eco*RI; S, *Sal*I; Sst, *Sst*I; X, *Xho*I; Xba, *Xba*I. Not all *Hin*cII, *Sst*I or *Xba*I sites are shown.

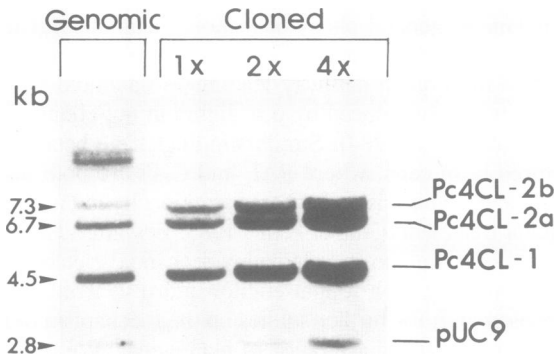


**Fig. 2;** Genomic DNA blot with copy number reconstructions of Pc4CL-1, Pc4CL-2a and Pc4CL-2b. *Bam*HI-digested genomic DNA was fractionated on an agrarose gel adjacent to mixtures of *Bam*HI-digested plasmid DNA containing subcloned *Bam*HI fragments 6.7 kb in size from λ-Pc23 (Pc4CL-2a), 7.3 kb in size from λ-Pc32 (Pc4CL-2b), and 4.5 kb in size from λPc35 (Pc4CL-1). The mixtures contained the *Bam*HI fragments in amounts approximately equivalent to 1,2 or 4 copies (1×, 2× or 4×) per haploid parsley genome. The DNA blot was hybridized with the $^{32}$P-labelled 1700-bp 4CL cDNA. The 2.7-kb fragment seen in the reconstruction is linearized pUC9 DNA from the plasmid subclones, which hybridized with small amounts of plasmid DNA present in the cDNA hybridization probe.

indicating that the clones may contain two alleles of the same gene. A third clone, λ-Pc35, had no restriction sites in common with λ-Pc23 or λ-Pc32 outside of the region hybridizing with the cDNA, suggesting that a separate 4CL gene is contained on this clone.

The authenticity of the genomic clones was verified by comparing DNA blots of various restriction digests of genomic DNA with blots of cloned DNA. Fragments found in the clones were also represented in genomic DNA and the lack of additional hybridizing genomic fragments, even at reduced stringency, suggested that no additional 4CL genes are present in parsley (data not shown). The gene located on λ-Pc35 was designated as Pc4CL-1 and the putative alleles of the second gene located on λ-Pc23 and λ-Pc32 were designated as Pc4CL-2a and Pc4CL-2b, respectively. To estimate the copy number of these 4CL genes, DNA blots containing *Bam*HI-digested genomic DNA and copy number reconstructions of subcloned *Bam*HI fragments from Pc4CL-1, Pc4CL-2a and Pc4CL-2b were hybridized with the 4CL cDNA (Figure 2). The intensities of hybridization suggest that Pc4CL-1 is present at ~1 copy, and Pc4CL-2a and Pc4CL-2b at ~1/2 to 1 copy per haploid parsley genome.

### *3′-Untranslated regions of Pc4CL-1 and Pc4CL-2*

Individual genes can often be distinguished from each other by sequence divergence in regions corresponding to the 3′-untranslated portions of mature mRNAs (for example, Fluhr *et al.*, 1986b). Thus Pc4CL-1, Pc4CL-2a and Pc4CL-2b were sequenced near the expected 3′ transcriptional termini to identify possible heterogeneity which would facilitate studies on the expression of the individual genes. The exact locations of these termini were established by comparing the 3′ sequences of several 4CL cDNA clones with the genomic sequences. Multiple polyadenylation sites were found within a 100-bp region so that the 3′-untranslated portions varied from ~140 to 240 bp in length (Figure 3a). The positions of the translation stop codons, which occurred in identical positions in both genes, were deduced from the nucleotide sequences of two near full-length 4CL cDNA clones (H.Hoffmann *et al.*, in preparation). No consensus AATAAA polyadenylation signal was found, but possible variants are present. The sequences of Pc4CL-2a and Pc4CL-2b were identical within the entire 3′-untranslated region and for > 100 bp downstream of the most distal 3′ polyadenylation site, supporting the assumption that they are two alleles of the same gene. The sequence of Pc4CL-1, although very similar to Pc4CL-2, differed in four nucleotide positions in the 3′-untranslated region (Figure 3a), showing that the two genes can be distinguished at the sequence level. Further differences between Pc4CL-1 and Pc4CL-2 occurred distal to the polyadenylation sites.

### *Expression of Pc4CL-1 and Pc4CL-2*

Since the nucleotide differences between Pc4CL-1 and Pc4CL-2 were insufficient for the construction of gene-specific hybridization probes, we assessed the expression of the two 4CL genes by determining the presence of cDNAs specific to each gene in cDNA libraries prepared using RNA from either elicitor-treated or u.v.-irradiated cells. The 3′-untranslated portions of 17 4CL cDNAs from both types of libraries were sequenced and compared with the genomic sequences. All 4CL cDNAs analysed matched exactly the sequences of either Pc4CL-1 or Pc4CL-2 (Figure 3b). Furthermore, both cDNA types were present in libraries prepared with RNA from cells treated with either stimulus. These data demonstrate that both genes are expressed and suggest that both may be activated by either elicitor treatment or u.v. irradiation.

To extend these findings, a probe from an intron region present in Pc4CL-2 but not Pc4CL-1 was used to measure the level of transcription from Pc4CL-2. Pc4CL-1 and Pc4CL-2 both consist of five exons separated by introns of varying sizes (Figure 4), as deduced by comparing genomic sequences with cDNA sequences (C.Douglas and H. Hoffmann, unpublished). The structure of both genes is highly conserved in the locations as well as the sequences of introns and exons; the only major difference is in the size of the second intron (Figure 4). In Pc4CL-2, this intron is ~700 bp larger than the corresponding intron in Pc4CL-1. Sequencing of *Bam*HI−*Xba*I fragments 1.05 and 1.8 kb in size from Pc4CL-1 and Pc4CL-2, respectively, which contain most of these introns, showed that both introns contain highly

a

```
Pc4CL-1   TAAGTAATCTCTACAAACAGAAATGGCATAAAGCTGAAGCTGTATGTGTA  50
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Pc4CL-2a  TAAGTAATCTCTACAAACAGAAATGGCATAAAGCTGAAGCTGTATGTGTA
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Pc4CL-2b  TAAGTAATCTCTACAAACAGAAATGGCATAAAGCTGAAGCTGTATGTGTA  50
          End

Pc4CL-1   TCTTTACAAAGTAACTTCTACCTAAAAGAGCTCCGAGTTGTAACTTGTTT  100
          |||||||||||||| ||||||||||||||||||||||||||||||||||||
Pc4CL-2a  TCTTTACAAAGTAAATTCTACCTAAAAGAGCTCCGAGTTGTAACTTGTTT
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Pc4CL-2b  TCTTTACAAAGTAAATTCTACCTAAAAGAGCTCCGAGTTGTAACTTGTTT  100

Pc4CL-1   GTAAATTTTATTTTTTGTATGAAGGAAGATTTATAAGATCATGTAATCAC  150
          ||| |||||||||||| ||||||||||||||||||||||||||||||||||
Pc4CL-2a  GTATATTTTATTTTTTGAATGAAGGAAGATTTATAAGATCATGTAATCAC
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Pc4CL-2b  GTATATTTTATTTTTTGAATGAAGGAAGATTTATAAGATCATGTAATCAC  150

Pc4CL-1   TCATCAAAGTTTAAATATCATCATTTGTATCACTACATTCGGTTTTTCCG  200
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Pc4CL-2a  TCATCAAAGTTTAAATATCATCATTTGTATCACTACATTCGGTTTTTCCG
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Pc4CL-2b  TCATCAAAGTTTAAATATCATCATTTGTATCACTACATTCGGTTTTTCCG  200

Pc4CL-1   ATCATAAACAATGATTTTTTCATGTTAAAAGTACAACATCTGTTTAATTT  250
          ||||||||| ||||||||||||||||||||||||||||||| ||||||
Pc4CL-2a  ATCATAAACATTGATTTTTTCATGTTAAAAGTACAACATCTGTGTAATTT
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Pc4CL-2b  ATCATAAACATTGATTTTTTCATGTTAAAAGTACAACATCTGTGTAATTT  250

Pc4CL-1   TTTTTTT.....AATCTCGAAAATACAAGGAATTAGAGACTCGACGGCAC  295
          |||||||     |||||||||||||||||||||||||||| ||||
Pc4CL-2a  TTTTTTTTTTTTTTATCTCGAAAATACAAGGAATTAGAGACTCGACCGCAC
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Pc4CL-2b  TTTTTTTTTTTTTTATCTCGAAAATACAAGGAATTAGAGACTCGACCGCAC  300
```

b

| | cDNA SOURCE | |
|---|---|---|
| | Elicitor-treated cells | UV-irradiated cells |
| Pc4CL-1 | B11 | UV1 |
| | B14 | UV2 |
| | B15 | |
| | P8 | |
| Pc4CL-2 | E6 | UV2 |
| | E30 | UV4 |
| | E56 | UV6 |
| | P5 | UV7 |
| | | UV9 |
| | | UV14 |
| | | UV15 |

**Fig. 3.** Comparison of the 3'untranslated sequences of Pc4CL-1, Pc4CL-2a and Pc4CL-2b. (**a**) Genomic nucleotide sequences. The regions sequenced are shown by small bidirectional arrows in Figure 4. Vertical arrows show the poly(A) addition sites found in 17 4CL cDNA clones whose sequences were compared with the genomic sequences. The translation stop codon is indicated and is preceded by an open reading frame of 544 amino acids in a near full-length cDNA (H. Hoffmann, unpublished). Stars show positions where base pair differences between Pc4CL-1 and Pc4CL-2 occur (Pc4CL-2a and Pc4CL-2b sequences are identical). **b)** Summary of sequence data from 4CL cDNA clones isolated from cDNA libraries prepared with RNA from elicitor-treated or u.v.-irradiated parsley cells. The 3' sequences of the cDNA clones listed were compared with the sequences of Pc4CL-1 and Pc4CL-2. All cDNAs matched the sequence of either Pc4CL-1 or Pc4CL-2 as shown.

homologous sequences outside of a central region unique to Pc4CL-2 (data not shown). A 0.54-kb fragment isolated from this unique region was used as a Pc4CL-2 specific probe in run-off transcription experiments.

Preliminary experiments had shown that *in vitro* labelled nuclear run-off transcripts hybridize equally well to exon- and intron-containing 4CL fragments. $^{32}$P-Labelled run-off transcripts of nuclei from unstimulated, elicitor-treated and u.v.-irradiated cells were hybridized under DNA excess conditions

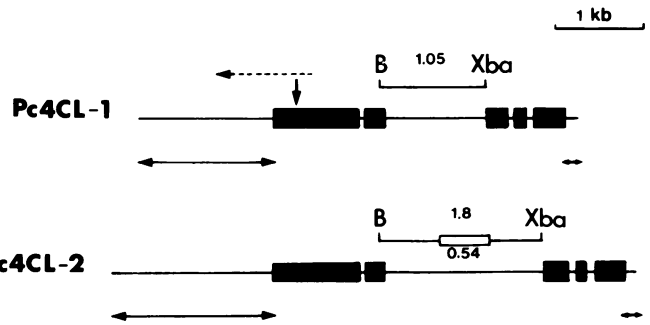


**Fig. 4.** Structure of the two parsley 4CL genes, Pc4CL-1 and Pc4CL-2. The 5' ends of the genes are to the left as the figure is drawn. Filled-in boxes represent the five exons of the two genes; these are separated by four introns ranging in size from 118 to 1734 bp. The locations of the introns and exons were deduced by comparing genomic sequences with cDNA sequences (C.Douglas and H. Hoffman, unpublished). *Bam*HI−*Xba*I restriction fragments from each gene (1.05 and 1.8 kb in size from Pc4CL-1 and Pc4CL-2, respectively), which contain most of the second large intron, are shown above the genes. An open box within the 1.8-kb Pc4CL-2 fragment shows the location of a 0.54-kb fragment used as a probe specific to Pc4CL-2. Also shown above the Pc4CL-1 gene are the locations of a uniformly labelled RNA probe used for S1 protection experiments (dashed horizontal arrows) and of a synthetic 15-bp oligonucleotide used for primer extension experiments (vertical arrow). The 3' and 5' sequences shown in Figure 3 and Figure 7 are indicated by horizontal bidirectional arrows below the genes.

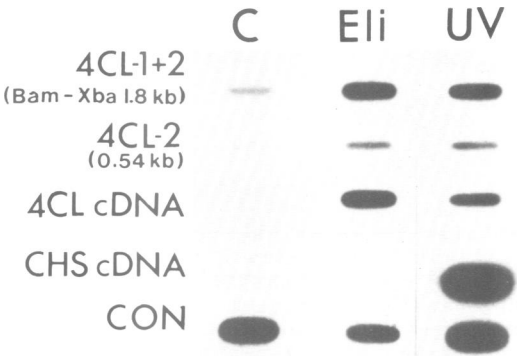


**Fig. 5.** Hybridization of *in vitro* labelled run-off transcripts from parsley cell culture nuclei to various DNA fragments. $^{32}$P-labelled transcripts from control cells (C), or cells treated for 3 h with elicitor (Eli) or u.v. light (UV) were hybridized to slot blots of plasmids containing the following DNA fragments: the *Bam*−*Xba* 1.8-kb fragment containing intron sequences common to Pc4CL-1 and Pc4CL-2 (4CL1+2; see Figure 4), a 0.54-kb Pc4CL-2 specific fragment derived from *Bam*−*Xba* 1.8 (4CL2; see Figure 4), the 1700-bp 4CL cDNA (4CL cDNA), a parsley chalcone synthase cDNA (CHS cDNA) (Kreuzaler *et al.*, 1983) and a cDNA of a constitutively expressed parsley gene (CON) (I.Somssich and W.Schulz, unpublished). Hybridization was quantified by scanning autoradiograms and values were normalized to hybridization to the constitutive clone.

to 4CL cDNA, to the 1.8-kb *Bam*−*Xba* and the 0.54-kb Pc4CL-2 specific fragments, and to cDNAs specific for chalcone synthase mRNA (which is induced by u.v light but not elicitor) and for the mRNA of a constitutively expressed parsley gene to which hybridization was normalized (Figure 5). In addition to the expected behaviour of the latter two controls, the results clearly demonstrate that both 4CL genes are transcriptionally regulated by the two stimuli. Increased hybridization was observed to the Pc4CL-2 probe after both elicitor treatment and u.v. irradiation. Transcription specific to this probe was activated ~6-fold above background by elicitor treatment and ~2.5-fold above background by u.v. irradiation. The increase in signal strength

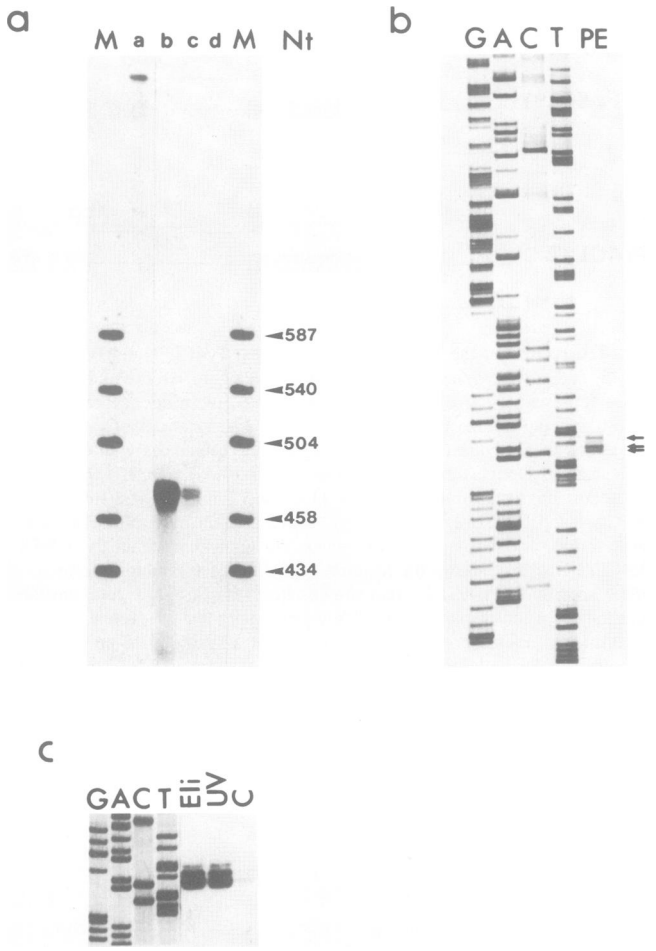


**Fig. 6.** S1 nuclease protection and primer extension analysis of parsley 4CL genes. (**a**) S1 nuclease protection of the 5′ end of Pc4CL-1 and Pc4CL-2. Uniformly labelled *in vitro* synthesized RNA (see Figure 4) was hybridized to total RNA from untreated cells or cells treated 5 h with elicitor. Fragments resistant to S1 nuclease digestion were separated on a 5% sequencing gel. **Lane a**, RNA probe without treatment; **lanes b–d**, samples treated with S1 nuclease: **b**, RNA probe hybridized to 10 μg elicitor-induced RNA; **c**, probe hybridized to 10 μg RNA from untreated cells; **d**, RNA probe alone. M, markers from *Hae*III-digested pBR322 with sizes in nucleotides (Nt). (**b**) Primer extension. A synthetic oligonucleotide primer homologous to both 4CL genes (see Figure 4) was used with 10 μg poly(A)$^+$ RNA isolated from cells treated for 5 h with elicitor. Primer extension products (PE) were separated on a 5% sequencing gel. Arrows indicate the three major products. The sequencing ladder (left lanes) corresponds to M13 template DNA from Pc4CL-1 sequenced using the same oligonucleotide primer. (**c**) Comparison of the primer extension products obtained using 10 μg poly(A)$^+$ RNA from cells treated 5 h with elicitor (Eli) or u.v. light (UV) or from untreated control cells (C). The primer and sequencing ladder are as in (**b**).

after either treatment was two to three times stronger with probes measuring transcription from both 4CL genes, which detected a 15- to 20-fold stimulation of 4CL transcription after elicitor treatment and a 4.5- to 6-fold stimulation after u.v. irradiation. This suggests that the transcriptional induction of 4CL by elicitor or u.v. light is the sum of the activation of both genes and that both genes are activated approximately equally following elicitor treatment or u.v. irradiation.

### *5′ Transcription start sites*

As a first step in comparing potential 5′ regulatory sequences of the two 4CL genes, we determined the transcription start sites of the genes by S1 nuclease protection and primer extension assays. Pc4CL-1 and Pc4CL-2 are identical in length from the 3′ end of the first exon to > 200 bp upstream of the translation start point (C.Douglas and H. Hoffmann, unpublished), differing only by individual base pair exchanges within this region. Since the S1 probe and the oligonucleotide used for primer extension were derived from sequences in this region, either method was expected to define the transcription start points of both genes. Uniformly labelled RNA was transcribed *in vitro* in the antisense orientation from a 1.1-kb Pc4CL-1 restriction fragment (Figure 4) and was used as a probe for S1 nuclease protection. When hybridized to RNA from elicitor-treated cells, fragments centred at ~475 nucleotides in length were protected from S1 digestion (Figure 6a, lane b), placing the 5′ transcript termini of both genes at ~75 nucleotides upstream of the translation initiation sites. No other protected fragments were observed, consistent with the absence of additional 4CL genes in parsley. About 20-fold more probe was protected when hybridized to RNA from elicitor-treated cells than when hybridized to RNA from untreated control cells (Figure 6a, lanes b and c), indicating an increase in 4CL mRNA abundance similar to the level of transcriptional activation observed in nuclear run-off experiments.

The transcription start points were confirmed and defined more precisely by primer extension analysis using an end-labelled oligonucleotide primer homologous to sequences in the first exon of both 4CL genes (Figure 4). Three major primer extension products were obtained using poly(A)$^+$ RNA from elicitor-treated cells (Figure 6b). The largest of these places the 5′ transcript termini 78 nucleotides upstream of the translation start sites, in close agreement with the S1 protection data. Two additional extension products, two to three nucleotides shorter, could be due to incomplete extension of RNA templates or alternative transcription start sites used by one or both genes.

Since the two 4CL genes appear to be regulated by both elicitor and u.v. light, it was of interest to determine if transcription is initiated at the same site following either treatment. To this end, primer extension products obtained using RNA from elicitor-treated, u.v.-irradiated or untreated control cells were compared (Figure 6c). In all cases, products of identical length were synthesized, showing that the same transcription start site was used after treatment with either stimulus and during background transcription of 4CL. The relative amounts of products were similar to those observed in run-off transcription and S1 protection assays.

### *5′ Nucleotide sequence of Pc4CL-1 and Pc4CL-2*

In order to compare possible 5′ regulatory elements common to the two 4CL genes, the nucleotide sequences of Pc4CL-1 and Pc4CL-2 upstream of the translation start points were determined. Approximately 1.5 kb of Pc4CL-1 and 1.8 kb of Pc4CL-2 were sequenced as shown in Figure 4. The two genes share large regions of homology up to positions −1088 of Pc4CL-1 and −858 of Pc4CL-2 relative to the transcription start points (Figure 7). No significant stretches of homology were found beyond this. Up to position −440, the two genes are nearly identical. Found within this region are sequences homologous to TATA (position −22) and CAT (position −122) boxes, common features of eukaryotic promoters (Breathnach and Chambon, 1981). Between the TATA and CAT boxes, a 3-fold repetition of the sequence TA(A)CCAA occurs. While the significance of this hexanucleotide for 4CL expression is not known, it is interesting that a similar sequence (TACCA) is repreated three times in similar positions within the promoter of the u.v.light-regulated chalcone synthase gene of *Antirrhinum majus* (Sommer and

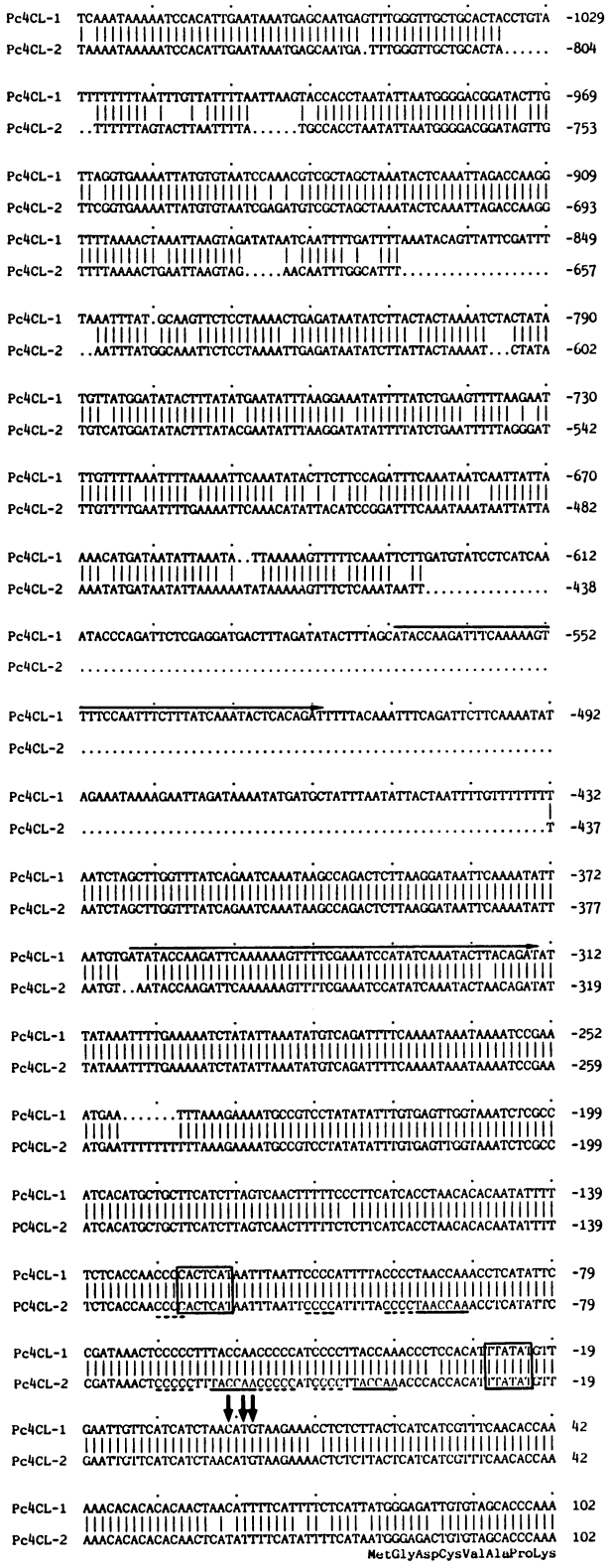


**Fig. 7.** Nucleotide sequences of the 5′ ends of Pc4CL-1 and Pc4CL-2. The sequences are compared with each other by match lines; gaps were introduced to maximize homology between the sequences. The transcription start sites are shown by vertical arrows, and the first eight amino acids of the coding sequences in the first exons are shown. TATA and CAT homologies are boxed in and possible regulatory sequences located between the two are underlined with solid or dashed lines. A 50-kb direct repeat found in Pc4CL-1 is shown by horizontal arrows above the sequence.

Saedler, 1986). Also of note in the region between the TATA and CAT boxes of 4CL is a 6-fold repetition of the sequence CCCC(C), since GC-rich elements have been found to be of importance in the expression of some eukaryotic genes (Dynan and Tjian, 1985).

Further upstream, an imperfect direct repeat of 50 bp is found at positions −304 and −521 in the Pc4CL-1 gene (Figure 7). While such direct repeats are also present upsteam of the *A. majus* chalcone synthase (Sommer and Saedler, 1986) and the proteinase inhibitor I gene of tomato (Lee *et al.*, 1986), the relevance of this feature for the expression of 4CL is unclear, since one of the repeats is not present in the Pc4CL-2 gene. Still further upstream (data not shown), sequences which show homology to the SV40 enhancer core sequence (Weiher *et al.*, 1983) are found: in Pc4CL-1, GGTGCAAAGTGT (position −1236), GGGTGCAAGTGT (position −1268) and GGGTGCAAAGTG (position −1355); and in Pc4CL-2, GTGTTCAAT (position −1109) and GTGTTAAAGT (position −1543). These motifs occur within regions of Pc4CL-1 and Pc4CL-2 which otherwise lack any apparent homology with each other.

## Discussion

We have described genomic clones from parsley cells containing complete copies of genes encoding 4CL. These clones appear to contain copies of two separate 4CL genes, and restriction fragments from each clone which hybridize to the 4CL cDNA can be assigned to corresponding genomic restriction fragments. Additional genomic fragments which could contain copies of 4CL were not found. Thus, there appear to be only two genes encoding 4CL in the parsley genome. In other systems, the analysis of the expression of individual members of gene families has been possible by the use of hybridization probes specific to diverged, non-coding portions of individual genes (Fluhr *et al.*, 1986b). Due to the high degree of homology between Pc4CL-1 and Pc4CL-2, this approach was not feasible for the two 4CL genes. Thus, we exploited small sequence differences in the 3′-non-coding portions of the genes to determine the presence of cDNA clones derived from transcripts of each gene in libraries made with RNA from elicitor-treated or u.v.-irradiated cells. We also used a probe from a unique portion of the large second intron of Pc4CL-2 to measure run-off transcripts specific to this gene. The results from these experiments provide evidence that Pc4CL-2 is transcriptionally activated by both fungal elicitor and u.v. irradiation and, taken with the cDNA data, imply that Pc4CL-1 is similarly regulated. These results are being further tested in a separate study using Pc4CL-1 and Pc4CL-2, suitably marked, in gene transfer experiments.

The increase in 4CL mRNA abundance after elicitor treatment or u.v. irradiation was measured by S1 nuclease protection and primer extension assays. The relative increase in abundance was very similar to the relative increase in transcription of 4CL following either stimulus (~6-fold after u.v. irradiation and 15- to 20-fold after elicitor treatment). RNA and nuclei were both isolated at time points near the peak of transiently induced 4CL mRNA accumulation (Schmelzer *et al.*, 1986) or transcriptional activation (I.Somssich *et al.*, in preparation), respectively. It can thus be concluded that the transient accumulation of 4CL transcripts is due primarily to the transcriptional activation of the corresponding genes in both cases, as expected from previous results (Chappell and Hahlbrock, 1984). Primer extension analysis also showed that identical mRNA transcripts are synthesized after elicitor treatment, u.v. irradiation and in untreated

control cells, indicating that the same promoters (which are identical in Pc4CL-1 and Pc4CL-2) are used in all three cases. Thus, the activity of these promoters must be regulated by both fungal elicitor and u.v. light.

A number of plant genes have been shown to be regulated at the transcriptional level (Chappell and Hahlbrock, 1984; Silverthorn and Tobin, 1984; Berry-Lowe and Meagher, 1985; Eckes *et al.*, 1985; Hagen and Guilfoyle, 1985; Walling *et al.*, 1986; Somssich *et al.*, 1986; Lawton and Lamb, 1987) and in some cases the regulation of individual members of gene families has been analysed. In pea, for example, two genes encoding the small subunit of ribulose bisphosphate carboxylase are apparently each regulated by both red and blue light (Fluhr and Chua, 1986). Our results with 4CL demonstrate for the first time the transcriptional activation of the same plant gene by two physically distinct external stimuli. The 4CL genes thus show similarities to stress-induced genes in animal systems, such as mouse (Durnam *et al.*, 1984) and human (Karin *et al.*, 1984; Friedmann and Stark, 1985) metallothionine genes and the human HSP 70 heat shock gene (Wu *et al.*, 1986), which are responsive at the transcriptional level to multiple stimuli.

By analogy to other inducible plant and animal genes, the inducibility of 4CL by elicitor and u.v. light is likely to be conferred by regulatory elements 5′ to the transcription initiation sites of the genes. The 5′ nucleotide sequences of Pc4CL-1 and Pc4CL-2 are nearly identical for > 400 bp and retain significant regions of homology well beyond this, to around −850 to −1100 bp upstream of transcription initiation. The high degree of sequence conservation between the two genes within the first 400 bp suggests that elements responsible for elicitor and u.v. light activation may be present here. Elements involved in inducible expression of animal genes (for example, Karin *et al.*, 1984; Wu *et al.*, 1986; and references therein) and regions required for light inducibility of plant genes (Fluhr *et al.*, 1986a) have been found within similar distances upstream of transcription initiation sites. In several plant genes these regions contain sequences which share homology with the SV40 core enhancer sequence (Eckes *et al.*, 1986; Fluhr *et al.*, 1986a; Lee *et al.*, 1986; Sommer and Saedler, 1986). The role played by such enhancer-like sequences in plant gene expression is not clear, but it is interesting that similar sequences are also found in locations upstream of the 4CL genes. These sequences are found in positions quite remote from the transcription initiation sites (well over 1000 bp upstream).

Gene transfer experiments will be required to define what role the potential upstream regulatory sequences may play in the expression of 4CL. In view of the apparent inducibility of both 4CL genes by fungal elicitor and u.v. light, it will be of particular interest to determine if common or distinct regulatory sequences are necessary for activation by these two stimuli. This information will be helpful in elucidating the mechanism of stress regulation of plant genes.

## Materials and methods

*Plant cell cultures and induction*

Parsley cell cultures were grown in continuous darkness as described (Ragg *et al.*, 1981) and were harvested either without treatment, after application of Pmg elicitor derived from *Phytophthora megasperma* f. sp. *glycinea* (Ayes *et al.*, 1976; Kuhn *et al.*, 1984) or after u.v. irradiation. U.v. irradiation was performed essentially as described (Ragg *et al.*, 1981) except that cells were transferred to sterile quartz glass flasks prior to irradiation.

*RNA isolation*

RNA was extracted from parsley cells as described (Ragg *et al.*, 1981). Poly(A)$^+$ RNA was isolated by chromatography over oligo(dT)−cellulose as described by Maniatis *et al.* (1982).

*Library construction and screening*

A genomic library was constructed by ligation of partially *Mbo*I-digested, size-fractionated parsley DNA with the large and small arms of *Bam*HI-digested λ-EMBL4 (Frischhauf *et al.*, 1983). Approximately $1 \times 10^6$ recombinant plaques were screened for hybridization to a 450-bp 4CL cDNA (Kuhn *et al.*, 1984). Plaques giving a positive hybridization signal were carried through four rounds of purification. Complementary DNA was synthesized as described (Gubler and Hoffmann, 1983) using poly(A)$^+$ RNA from parsley cells which had been treated with Pmg elicitor or irradiated with u.v. light, and was ligated with λNM1149 (Murray, 1983). The cDNA libraries were screened initially with the 450-bp 4CL cDNA and subsequently with a 1700-bp 4CL cDNA obtained from a cDNA library made with elicitor-induced RNA.

*Filter hybridization, restriction digests and subcloning*

DNA blots and restriction digests were carried out according to standard methods (Maniatis *et al.*, 1982). Filter hybridizations were carried out with nick-translated probes in 6 × SSC at 68°C and blots were washed twice for 30 min in 0.2 × SSC, 0.5% SDS at 68°C. For copy number reconstructions, *Bam*HI-digested genomic DNA was fractionated on an agarose gel adjacent to mixtures of *Bam*HI-digested plasmid DNA containing subcloned *Bam*HI fragments from the 4CL genes. The amount of plasmid DNA needed for copy number reconstructions was calculated using an estimated haploid parsley genome size of $1.9 \times 10^9$ bp (Kiper, 1979) and was mixed with herring sperm DNA before electrophoresis. Restriction maps were constructed by single or multiple restriction enzyme digests of the genomic clones or subclones of these. The 0.54-kb subfragment within the *Bam*−*Xba* 1.8-kb restriction fragment of Pc4CL-2 was isolated by treating the purified 1.8-kb restriction fragment for 2 min with exonuclease III at 37°C, followed by S1 nuclease digestion (300 units/ml for 30 min at 18°C), fill-in with DNA polymerase (Klenow fragment) and ligation with *Sma*I-digested pUC9 DNA.

*DNA sequencing*

The dideoxy chain termination method was used to sequence M13 clones (Sanger *et al.*, 1977; Vieira and Messing, 1982). Overlapping clones were prepared using exonuclease III to create unidirectional deletions as described (Yanisch-Perron *et al.*, 1985).

*Run-off transcription*

Transcriptionally active nuclei were isolated as described (Chappell and Hahlbrock, 1984) from parsley cells which were harvested either without treatment, 3 h after application of Pmg elicitor or 3 h after beginning u.v. irradiation. Run-off transcription was performed in 200-μl reaction mixtures containing 200 μg DNA and 200 μCi [$^{32}$P]UTP(410 Ci/mmol, Amersham) as described (Chappell and Hahlbrock, 1984). The *in vitro* synthesized $^{32}$P-labelled RNA was extracted by first adding 16 μl tRNA (5 mg/ml), 20 μl Hepes buffer (20 mM Hepes, 5 mM $MgCl_2$, 1 mM $CaCl_2$) and 30 units DNase I (RNase-free) to the reaction mixture and incubating for 1 h at 37°C. 22 μl 10 × SET (100 mM Tris−HCl, pH 7.5, 50 mM EDTA, 1% SDS) and 2 μl proteinase K (10 mg/ml) were then added and incubation continued for 25 min at room temperature. The mixture was extracted twice with phenol/chloroform/isoamylalcohol (25:24:1) and twice with chloroform. Volumes of 200 μl were passed twice through 1-ml Sephadex G-50 spin columns (Maniatis *et al.*, 1982) and $^{32}$P-labelled transcripts used immediately for hybridization. For hybridization, 1 μg of recombinant plasmid DNA was linearized, denatured and applied to nitrocellulose filters using a Manifold II slot blot apparatus (Schleicher and Schuell). Hybridization of filters with [$^{32}$P]RNA and washing conditions were as described (Chappell and Hahlbrock, 1984). Hybridization was quantified by scanning autoradiogramms exposed for various lengths of time with an LKB Laser Densitometer. Values were normalized to hydridization to the cDNA of the constitutively expressed gene. In addition, [$^{32}$P]RNA hybridized to the filters was quantified by scintillation counting of radioactivity released from filters following treatment for 5 min in water at 100°C, 30 min in 0.4 M NaOH at room temperature, and neutralization with acetic acid.

*S1 nuclease protection and primer extension*

Radioactively labelled RNA homologous to a *Hin*dIII−*Xho*I fragment from the 5′ end of Pc4CL-1 (see Figure 4) was synthesized in the antisense orientation using a Bluescribe vector (Vector Cloning Systems, San Diego) and T3 polymerase according to the manufacturers instructions. The RNA probe was hybridized to 10 μg total RNA isolated from parsley cells 5 h after elicitor treatment or from untreated cells. Hybridization was for 16 h at 49°C in 30μl of a buffer containing 40 mM Pipes, pH 6.4, 400 mM NaCl, 1 mM EDTA and 80% formamide. Single-stranded RNA was subsequently digested by adding 300 μl of a buffer containing 30 mM sodium acetate, pH 4.6, 200 mM NaCl, 4.5 mM $ZnSO_4$ and 500 units/ml S1 nuclease (Boehringer) and incubating for 30 min at 20°C. RNA was then precipitated and separated on a 5% polyacrylamide sequencing gel. For primer extension studies, a synthetic 15-bp oligonucleotide (located as shown in Figure 4), was end labelled using polynucleotide kinase and used as a primer for reverse transcriptase. About 75 000 c.p.m. (10 ng) primer was annealed to 10 μg poly(A)$^+$ RNA in 50-μl reaction mixtures containing 50 mM Tris−HCl, pH 8.3, 50 mM KCl, 8 mM $MgCl_2$, 10 mM DTT, 100 units RNasin and 0.5

mM dNTPs. Primers were extended using reverse transciptase (Life Sciences, 33 units per reaction) at 42°C for 30 min. Nucleic acids were precipitated, resuspended in 40 μl 0.4 M NaOH and incubated 16 h at room temperature to hydrolyse RNA. Following neutralization and precipitation, DNA was separated on a 5% sequencing gel. Single-stranded M13 templates containing Pc4CL-1 and Pc4CL-2 DNA were sequenced using the same oligonucleotide primer and loaded on the gel adjacent to the primer extension products. The relative amounts of primer extension products and S1 nuclease-protected RNA probe were determined by scanning autoradiograms with an LKB Laser Densitometer.

## Acknowledgements

We thank Imre Somssich for help with run-off transcription experiments, Jeff Dangl for advice and reading of this manuscript and Annette Steinmeier for preparation of the manuscript. This work was supported by the Max-Planck Society and Fonds der Chemischen Industrie.

## References

Ayers,A.R., Ebel,J., Finelli,F., Berger,N. and Albersheim,P. (1976) *Plant Physiol.*, **57**, 751–759.
Berry-Lowe,S. and Meagher,R.B. (1985) *Mol. Cell Biol.*, **5**, 1910–1917.
Breathnach,U. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **45**, 349–384.
Chappell,J. and Hahlbrock,K. (1984) *Nature*, **311**, 76–78.
Cramer,C., Ryder,T.B., Bell,J. and Lamb,C. (1985) *Science*, **227**, 1240–1243.
Dixon,R.A. (1986) *Biol. Rev.*, **61**, 239–291.
Durnam,D.M., Hoffmann,J.S., Quaite,C.J., Bennitt,E.P., Chen,H.Y., Brinster,R.L. and Palmiter,R.D. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1053–1056.
Dynan,W.S. and Tjian,R. (1985) *Nature*, **316**, 774–778.
Eckes,P., Schell,J. and Willmitzer,L. (1985) *Mol. Gen. Genet.*, **199**, 216–224.
Eckes,P., Rosahl,S., Schell,J. and Willmitzer,L. (1986) *Mol. Gen. Genet.*, **205**, 14–22.
Edwards,K., Cramer,C., Bolwell,G.P., Dixon,R.A., Schuch,W. and Lamb,C. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 6731–6735.
Fluhr,R. and Chua,N.-H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2358–2362.
Fluhr,R., Kuhlemeier,C., Nagy,F. and Chua,N.-H. (1986a) *Science*, **32**, 1106–1112.
Fluhr,R., Moses,P., Morelli,G., Coruzzi,G. and Chua,N.-H. (1986b) *EMBO J.*, **5**, 2063–2071.
Friedmann,R.L. and Stark,G.R. (1985) *Nature*, **314**, 637–639.
Frischhauf,A.-M., Lehrach,H., Poustka,A. and Murray,N. (1983) *J. Mol. Biol.*, **170**, 827–842.
Gubler,U. and Hoffmann,B. (1983) *Gene*, **25**, 263–269.
Hagen,G. and Guilfoyle,T.J. (1985) *Mol. Cell Biol.*, **5**, 1197–1203.
Hahlbrock,K. and Grisebach,H. (1979) *Annu. Rev. Plant Physiol.*, **30**, 105–130.
Hahlbrock,K. and Scheel,D. (1987) In Chet,I. (ed.), *Non-Conventional Approaches to Plant Disease Control*. John Wiley and Sons, New York, in press.
Hahlbrock,K., Lamb,C.J., Purwin,C., Ebel,J., Fautz,E. and Schäfer,E. (1981) *Plant Physiol.*, **67**, 768–773.
Hahlbrock,K., Kreuzaler,F., Ragg,H., Fautz,E. and Kuhn,D.N. (1982) In Jaenicke,L. (ed.), *Biochemistry of Differentiation and Morphogenesis*. Springer, Berlin, pp. 34–43.
Karin,M., Haslinger,A., Holtgreve,H., Richards,R.I., Krauter,P., Westphal,H.M. and Beato,M. (1984) *Nature*, **308**, 513–519.
Kiper,M. (1979) *Nature*, **278**, 279–280.
Knogge,W., Kombrink,E., Schmelzer,E. and Hahlbrock,K. (1987) *Planta*, in press.
Kreuzaler,F., Ragg,H., Fautz,E., Kuhn,D.N. and Hahlbrock,K. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2591–2593.
Kuhn,D.N., Chappell,J., Boudet,A. and Hahlbrock,K. (1984) *Proc. Natl. Acad. Sci. USA*, **82**, 1102–1106.
Lawton,M.A. and Lamb,C.J. (1987) *Mol. Cell Biol.*, in press.
Lee,J.S., Brown,W.E., Graham,J.S., Pearce,G., Fox,E.A., Dreher,T.W., Ahern,K.G., Pearson,G.D. and Ryan,C.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7277–7281.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
Murray,N.E. (1983) In *Lambda II*. Cold Spring Harbor Laboratory Press, New York, pp. 395–432.
Ragg,H., Kuhn,D.N. and Hahlbrock,K. (1981) *J. Biol. Chem.*, **256**, 10061–10065.
Ride,J.P. (1983) In Callow,J.A. (ed.), *Biochemical Plant Pathology*. John Wiley and Sons, New York, pp. 215–236.
Sanger,F., Nicklens,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schmelzer,E., Somssich,I. and Hahlbrock,K. (1986) *Plant Cell Rep.*, **4**, 293–296.
Silverthorn,J. and Tobin,E.M. (1984) *Mol. Cell Biol.*, **5**, 1112–1116.
Sommer,H. and Saedler,H. (1986) *Mol. Gen. Genet.*, **202**, 429–434.
Somssich,I., Schmelzer,E., Bollmann,J. and Hahlbrock,K. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2427–2430.
Vieira,J. and Messing,J. (1982) *Gene*, **19**, 259–268.
Walling,L., Drews,G.N. and Goldberg,R.B. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2123–2128.
Weiher,H., Konig,M. and Gruss,P. (1983) *Science*, **219**, 626–631.
Wu,B.J., Kingston,R.E. and Morimoto,R.I. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 629–633.
Yannisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.

*Received on February 4, 1987*