# Purification of tobacco nuclear proteins binding to a CACGTG motif of the chalcone synthase promoter by DNA affinity chromatography

Dorothee STAIGER, Frank BECKER, Jeff SCHELL, Csaba KONCZ and Klaus PALME

Max-Planck-Institut für Züchtungsforschung, Köln, Federal Republic of Germany

(Received January 4/March 8, 1991) – EJB 91 0016

The activity of various light-regulated and developmentally regulated plant gene promoters critically depends upon the presence of a conserved sequence with a central CACGTG motif. Using band-shift assays, we have identified nuclear factor(s) from *Nicotiana tabacum*, termed CG-1, that specifically recognize(s) this transcriptional element in the ultraviolet-light-regulated *Antirrhinum majus* chalcone synthase promoter. CG-1 activity is constitutively expressed in tobacco seedlings grown in the absence of ultraviolet light as well as in seedlings induced for chalcone synthase gene expression by ultraviolet light irradiation. CG-1 activity has also been detected in flower tissue. DNA-protein cross-linking studies identified three polypeptides with apparent molecular masses of 20, 32 and 42 kDa binding to the CACGTG motif. Proteins interacting with the CACGTG motif were purified from *N. tabacum* seedlings using differential sequence-specific DNA affinity chromatography employing wild-type and mutated CG-1-binding sites. Denaturing polyacrylamide gel electrophoresis revealed major polypeptides of approximately 20, 30 and 40 kDa which are highly enriched in the affinity-purified fractions binding specifically to the CACGTG motif.

Plant flavonoids are synthesized in response to diverse environmental stimuli. They have important functions in protection against ultraviolet light (Schmelzer et al., 1988), defense against pathogens (Lamb et al., 1989), and induction of nodulation (Peters et al., 1986). In addition, flavonoids serve as pigments for flowers and seed coats (Brouillard, 1988). External stimuli induce the transcription of genes involved in synthesis of flavonoids (Hahlbrock and Scheel, 1989). To study signal transduction mechanisms regulating flavonoid gene expression the induction of the chalcone synthase gene was characterized in detail (for review see Dangl et al., 1989). Chalcone synthase catalyzes the first step in the branch of phenylpropanoid metabolism specific for the synthesis of flavonoids (Heller and Hahlbrock, 1980). Its expression is induced both by ultraviolet light and developmental signals in various plant species such as parsley, snapdragon, tobacco and petunia (Kaulen et al., 1986; Lipphardt et al., 1988; Schulze-Lefert et al., 1989a; van der Meer et al., 1990).

Promoter elements regulating ultraviolet light induction of the *Antirrhinum majus* chalcone synthase gene (Sommer and Saedler, 1986) were identified by deletion analysis using a parsley protoplast transient expression system or transgenic tobacco seedlings (Kaulen et al., 1986; Lipphardt et al., 1988). This resulted in identification of a 150-bp promoter element immediately upstream of the TATA box (positions −39 to −197), containing *cis*-acting signals for ultraviolet-light-regulated expression. Of specific interest is a hexameric motif within this region having internal dyad symmetry. This element, CACGTG, located at positions −129 to −134 within the ultraviolet-light-responsive promoter region, contains the binding site for the nuclear factor(s) CG-1 (Staiger et al., 1989). Since multiple protein-DNA complexes were observed in band-shift assays, CG-1 might be composed of more than one polypeptide. CG-1 has been found in a variety of plant species including tobacco and *A. majus* (Staiger et al., 1989).

*In vivo* footprinting of the parsley chalcone synthase promoter has identified a related sequence motif, CCACGTGG, representing a target site for a protein-DNA interaction following ultraviolet irradiation (Schulze-Lefert et al., 1989a). Mutation of this motif results in loss of ultraviolet-light-responsive chalcone synthase gene expression in parsley protoplasts. It is intriguing that a related sequence element, referred to as G box, is highly conserved within 5′-flanking sequences of the light-regulated ribulose-1,5-bisphosphate carboxylase small subunit and chlorophyll-*a*/*b*-binding-protein genes. The G box has been shown to bind a nuclear factor from *Arabidopsis* and tomato (Giuliano et al., 1988; Schindler and Cashmore, 1990). Site-specific mutations within the G box of the *Arabidopsis* ribulose-1,5-bisphosphate carboxylase 1A promoter drastically reduce the expression of linked reporter genes demonstrating its *cis*-regulatory function *in vivo* (Donald and Cashmore, 1990). In addition, competition studies revealed that nuclear factors binding *in vitro* to the G box (G-box-binding factor, GBF) and to the CACGTG motif (CG-1) are closely related (Staiger et al., 1989). G-box-like motifs are also present within the promoter region of non-photoregulated plant genes. Recently, *in vivo* and *in vitro* binding studies have identified a protein factor interacting with the G box of the *Arabidopsis* alcohol dehydrogenase gene (McKendree et al., 1990; DeLisle and Ferl, 1990). Furthermore, CACGTG sequence elements are apparently present also in mammalian, yeast and *Xenopus* promoters (Sawadogo and Roeder, 1985; Moncollin et al., 1990; Scotto et al., 1989).

*Correspondence to* D. Staiger, Institute for Plant Sciences, Swiss Federal Institute of Technology, ETH Zentrum, Universitätsstrasse 2, CH-8092 Zurich, Switzerland

Interestingly, a mammalian promoter element containing this motif competes with the *A. majus* chalcone synthase CACGTG motif for binding to plant nuclear factor(s) CG-1 (Staiger et al., 1989). This suggests that a protein with an evolutionarily conserved DNA-binding domain recognizes this motif.

The ubiquity and importance of the CACGTG motif for expression of many plant genes suggests that CG-1 is a general transcription factor involved in the regulation of many different plant genes. Study of the developmental and molecular characteristics of this factor therefore is an important step towards unravelling the function of the cognate *cis*-regulatory element.

In this report we describe the analysis of CG-1 in *Nicotiana tabacum* using *in vitro* DNA-binding assays and photoaffinity labelling. We present data on the purification of proteins binding to the CACGTG motif by differential sequence-specific DNA affinity chromatography employing wild-type and mutated binding sites.

## MATERIALS AND METHODS

### *Plant material*

For preparative scale nuclei isolation seedlings of *N. tabacum* (Var. W38) were grown under normal greenhouse conditions. For ultraviolet-light induction experiments, seedlings were protected during growth against ultraviolet light (290 – 350 nm) by plexiglass shields. Induction was subsequently peformed for 18 h by continuous illumination with ultraviolet light using the following lamps: Osram L40 W15, Osram L40 W73 and Philipps TL40 W21. During illumination seedlings were protected against low-wavelength ultraviolet light by glass filters (below 310 nm; optisches Farbglas WG305, Schott, Mainz, FRG).

### *Northern blot hybridization*

RNA from tobacco seedlings was isolated according to Chirgwin et al. (1979). Tobacco chalcone synthase clone 2 was the kind gift of Drs G. Drews and R. Goldberg. A *Hin*dIII fragment of 1 kb carrying the coding sequence was labelled by multiprime kit (Boehringer). Northern blotting and hybridization were performed using established procedures (Maniatis et al., 1982). RNA filters were washed at 56° C with 0.3 M NaCl and 0.03 M sodium citrate.

### *DNA-binding substrates*

The sequences of oligonucleotides used are shown in Fig. 1. DNA probes for band-shift assays were prepared by *Eco*RI/*Hin*dIII digestion of pUC19 subclones containing five tandemly repeated copies of oligonucleotide 1 comprising the CG-1-binding site (Staiger et al., 1989). After a 'fill-in' reaction with [$^{32}$P]dCTP and [$^{32}$P]dATP using the Klenow fragment of DNA polymerase (Maniatis et al., 1982) labelled fragments were purified from polyacrylamide gels by elution into 10 mM Tris/HCl, pH 7.9, 1 mM EDTA and 100 mM NaCl.

### *Gel retardation assay*

DNA – protein-binding assays were carried out in 25 mM Hepes/KOH, pH 7.9, 100 mM KCl and 10% (by vol.) glycerol, containing 0.5 ng $^{32}$P-labelled DNA probe, nuclear protein extract and *Hpa*II-cut pUC19 competitor DNA. Concentrations are indicated in the figure legends. Specific competitor was linearized pUC19 DNA containing five copies of oligonucleotide 1 or 3. After incubation at room temperature for 10 min DNA-binding mixtures were separated on native 5% polyacrylamide gels in 20 mM Tris base, 8 mM sodium acetate, 15 mM glacial acetic acid and 1 mM EDTA at 4° C (Garner and Revzin, 1981; Fried and Crothers, 1981). For quantitative measurements bands of free and bound fragments were cut out from dried gels and quantified by liquid scintillation counting.

### *Preparation of nuclear extracts*

All procedures were at 4° C. Nuclei from four-week-old seedlings of *N. tabacum* were prepared by homogenization with a Waring blender in 3 vol. ice-cold 25 mM Mes/KOH, pH 6, 0.25 M sucrose, 5 mM EDTA, 10 mM KCl, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride and 1 mM NaF (buffer A) supplemented with 0.5% Triton X-100 and 0.5 mM spermidine. The suspension was treated for $3 \times 10$ s with an Ultra-Turrax (Janke and Kunkel, Staufen, FRG) and subsequently pressed through a 80-μm nylon mesh followed by a passage through a 20-μm nylon mesh. Nuclei were pelleted at $2000 \times g$ and washed with buffer A. The nuclear pellet was resuspended in 3 vol. 25 mM Hepes/KOH, pH 7.9, 5% (by vol.) glycerol, 0.1 mM EGTA, 420 mM KCl, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 1 μM each of leupeptin, pepstatin, antipain and chymostatin, 0.1 mM sodium pyrophosphate and 0.1 mM NaF (buffer B). Proteins were extracted by stirring the suspension of nuclei on ice for 30 min and nuclear debris was pelleted at $100000 \times g$ for 30 min. The supernatant was dialyzed against 500 vol. 25 mM Hepes/KOH, pH 7.9, 20% (by vol.) glycerol, 0.1 mM EDTA, 50 mM KCl, 14 mM 2-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, 0.1 mM NaF, quick-frozen in liquid nitrogen and stored at −70° C.

For preparation of chloroquine extracts, nuclei were resuspended in buffer B, pH 7.0, without KCl. Subsequently an equal volume of 20 mM chloroquine in buffer B, without KCl, was added (final concentration 10 mM; Schröter et al., 1987). After extraction and high speed centrifugation as above, chloroquine was removed by gel filtration on Sephadex G-50 (NAP column, Pharmacia).

For preparation of nuclei from flower tissue, petals from *N. tabacum* were dissected into the pink pigmented limb and the unpigmented tube. After quick-freezing in liquid nitrogen, the tissue was ground with mortar and pestle and stirred into 5 vol. buffer A. After thawing, processing was as described above. To evaluate the quality of extracts from different tissues, ratios between the amount of extracted nuclear protein and DNA content of isolated nuclei were determined.

DNA determination was performed according to Burton (1968). An aliquot of nuclei was incubated with 2 vol. homogenization buffer (2 M NaCl, 10 mM EDTA, 10 mM Tris/HCl, pH 7.0, 1% lauryl sarcosinate, 5 mM $MgCl_2$) for 1 h at room temperature. After repeated extraction with phenol and $CHCl_3$ the DNA was precipitated and dissolved in 50% diphenylamine (100 mg) in glacial acetic acid (10 ml) and sulphuric acid (280 μl). After incubation at 95° C for 15 min the absorption of samples was measured at 595 nm. As standard, calf thymus DNA treated identically was used.

Protein concentrations were determined as described by Bradford (1976) with diluted samples to prevent interference from Triton X-100.

*Cross-linking of photoactivated DNA probes with nuclear proteins*

A photoaffinity probe containing 5-bromodeoxyuridine in the place of thymine was synthesized after annealing of a nonameric primer to the noncoding strand (lower strand) of an oligonucleotide spanning positions −121 to −143 of the *A. majus* chalcone synthase promoter (Fig. 1) by Klenow polymerase in the presence of [$^{32}$P]dCTP and 0.1 mM dATP, dGTP and 5-bromodeoxyuridine triphosphate. The resulting double-stranded 24-mer oligonucleotide was purified by nondenaturing polyacrylamide gel electrophoresis. For isolation of DNA-protein complexes binding reactions were scaled up approximately 20-fold and separated on preparative retardation gels. Cross-linking was performed by placing the wet gel under an inverted transilluminator (312 nm wavelength) at a distance of 10 cm for 1 h at 4° C (Cereghini et al., 1988). The DNA-protein complexes were visualized by autoradiography at 4° C and excised. DNA-protein complexes were eluted in 0.5 × SDS electrophoresis buffer (Laemmli, 1970) using an elucon chamber (Biometra). Eluted fractions were heat-denatured and separated on 10–20% SDS/polyacrylamide gradient gels (Laemmli, 1970) using prestained molecular-mass markers (BioRad).

*Purification of CG-1*

Nuclei were extracted with 0.33 M KCl. After pelleting chromatin and nuclear debris, the supernatant was equilibrated with 25 mM Hepes/KOH, pH 7.9, 20% (by vol.) glycerol, 0.1 mM EDTA, 50 mM KCl, 1 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride, 0.1 mM NaF by passage through Sephadex G-50 columns (NAP column, Pharmacia). Subsequently, the extracts were loaded onto a DEAE-Sepharose Fast Flow column (bed volume 20 ml; Pharmacia) and the column was developed by a linear 50–300 mM KCl gradient generated by a Pharmacia GP-250 gradient programmer. For large scale preparations, CG-1-binding activity was recovered by a 0.2 M KCl step.

For differential sequence-specific DNA affinity chromatography two columns were prepared. The specific DNA affinity column 1 contained double-stranded oligonucleotide 1 (Fig. 1) representing the wild-type binding site. The mutant DNA affinity column contained double-stranded oligonucleotide 3 (Fig. 1) carrying three base changes as compared to wild-type CG-1-binding site.

18-mer double-stranded oligonucleotides were multimerized to an average length of 300 bp (Kadonaga and Tjian, 1986) and coupled to CNBr-activated Sepharose 4B following the protocol of Wu et al. (1987). Briefly, CNBr-activated Sepharose was hydrated and washed with 1 mM HCl. Coupling was performed in 10 mM potassium phosphate, pH 8.2, overnight at room temperature. Residual active groups were inactivated by incubating the resin with 0.1 M Tris/HCl, pH 8, for 5 h at room temperature. The matrix was then washed successively with 0.1 M Tris/HCl, pH 8, then 0.1 M potassium phosphate, pH 8.2, followed by 10 mM Tris/HCl/1.5 M NaCl and finally 10 mM Tris/HCl, 1 mM EDTA and 0.1 M NaCl. Affinity resins (bed volume 1 ml) were equilibrated in an Econo-column (Bio-Rad) with 25 mM Hepes/KOH, pH 7.9, 20% (by vol.) glycerol, 5 mM EDTA, 0.05% Nonidet P-40, 1 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride, 0.1 mM NaF and 0.1 M KCl (buffer C). CG-1-containing fractions eluted from the DEAE-Sepharose column were diluted until a conductivity equivalent to the equilibrating buffer

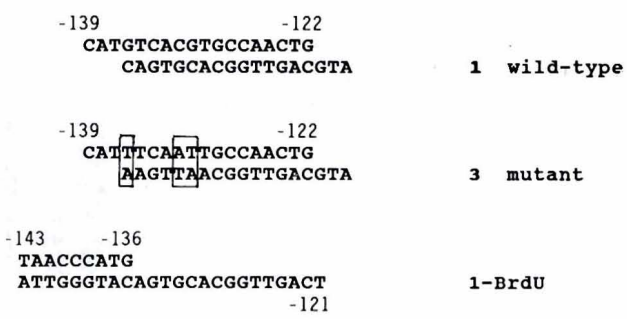


Fig. 1. *Sequence of oligonucleotides used for gel-shift assays and photoactivated protein-DNA crosslinking.* (1) 'Specific' (wild-type) CG-1-binding site (−122 to −139) of the *A. majus* chalcone synthase promoter); (3) mutated CG-1-binding site (Staiger et al., 1989); G-T transversions are boxed; (1-BrdU) wild-type CG-1-binding site (lower strand) containing an 5′ extension to position −143 to allow annealing of a complementary nonameric primer for filling in the upper strand

C was achieved and subsequently passed through the mutant oligonucleotide-3–Sepharose column by gravity flow. The flowthrough was combined with poly[d(U-T)] (5 μg/ml) and poly[d(A-T)] (5 μg/ml) and directly loaded onto the specific affinity matrix 1 (first cycle). This column was washed extensively with buffer C followed by buffer C containing 0.2 M KCl. CG-1-binding activity was eluted with a 0.5 M KCl step. Peak activity fractions were diluted to 0.1 M KCl and mixed with competitor DNA at the same concentration for reapplication to the specific affinity column (second cycle). For the third cycle, 30 μg/ml calf thymus DNA was used as competitor. For the purification described here the material from several DEAE-Sepharose columns was applied in parallel to four mutant oligonucleotide columns, four specific affinity columns (first and second cycle) and two columns of specific affinity chromatography (third cycle).

Purification was monitored by gel-retardation assay of the fractions, using five tandem copies of oligonucleotide 1 as probe. Enrichment of CG-1-binding activity was estimated by liquid scintillation counting of retarded complexes excised from dried retardation gels. Fractions from each purification step were concentrated by precipitation with 10% trichloroacetic acid/0.15% sodium deoxycholate and analyzed on 12% or 15% SDS miniature polyacrylamide gels (Laemmli, 1970) and silver-stained (Morissey, 1981).

## RESULTS

*Analysis of CG-1 DNA binding in tobacco tissue*

Tobacco nuclear factor(s) CG-1 recognize(s) a CACGTG sequence motif located within the ultraviolet-light-responsive region of the *A. majus* chalcone synthase promoter. To test whether *in vitro* binding of CG-1 is regulated by ultraviolet light, nuclear extracts were prepared from seedlings grown in the absence of ultraviolet light and from seedlings which had subsequently been exposed to continuous ultraviolet light for 18 h. Induction of chalcone synthase gene expression by ultraviolet light was confirmed by Northern blot hybridization using the tobacco chalcone synthase 2 gene as a probe (Fig. 2A). Equal amounts of nuclear proteins from untreated seedlings and from seedlings exposed to ultraviolet light were analyzed for CG-1 binding in band-shift assays with the oligonucleotides depicted in Fig. 1. In addition to the three specific complexes observed earlier with five tandem copies

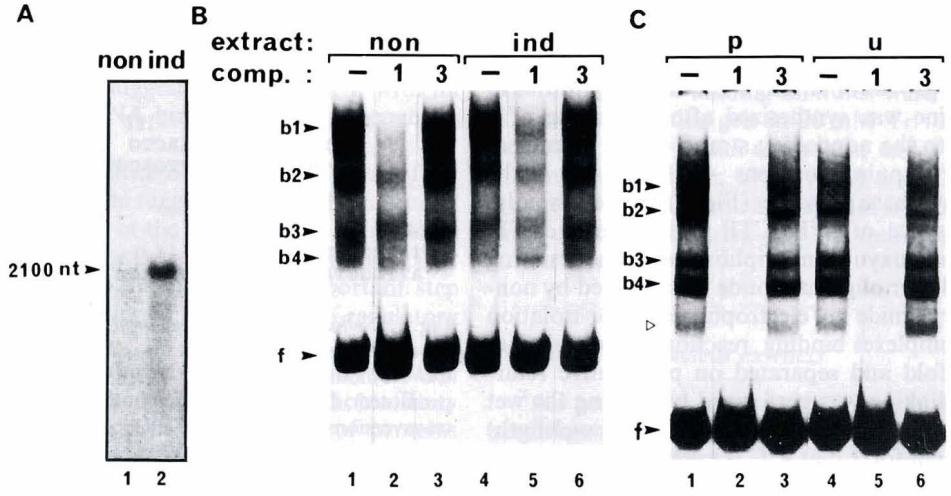


Fig. 2. *Analysis of CG-1 binding activity*. (A) Ultraviolet light induced chalcone synthase gene expression in tobacco seedlings. A Northern blot with 20 µg total RNA isolated from four-week-old tobacco seedlings grown in the absence of ultraviolet light (non, lane 1) and from tobacco seedlings which subsequently had been induced by continuous irradiation with ultraviolet light for 18 h (ind, lane 2) was hybridized with the tobacco chalcone synthase 2 gene (Drews and Goldberg, 1988). (B) Analysis of CG-1 binding activity in ultraviolet-light-induced and noninduced tobacco seedlings. Nuclear extracts prepared from noninduced (lanes 1–3) and ultraviolet-light-induced tobacco seedlings (lanes 4–6) were incubated with labelled five tandem copies of oligonucleotide 1 in the presence of 4 µg *Hpa*II-cut pUC19 DNA (lanes 1 and 4) or 2.5 µg *Hpa*II-cut pUC19 DNA and 1.5 µg linearized plasmid containing five copies of the wild-type oligonucleotide 1 (lanes 2 and 5) or of the mutated oligonucleotide 3 (lanes 3 and 6); comp, competitor DNA. (C) CG-1 binding activity in flower tissue. 1.5 µg nuclear protein isolated from the pink pigmented limb of the petal (p, lanes 1–3) and from the unpigmented tube (u, lanes 4–6) were incubated with 0.5 ng labelled five tandem copies of oligonucleotide 1 in the presence of 1 µg *Hpa*II-cut pUC19 DNA (lanes 1 and 4). For competition, pUC19 DNA was replaced by 1 µg linearized pUC19 plasmid containing five copies of wild-type oligonucleotide 1 (lanes 2 and 5) or of the mutated oligonucleotide 3 (lanes 3 and 6)

of oligonucleotide 1 (Staiger et al., 1989), these experiments revealed a fourth complex, termed b2 (Fig. 2B). Retarded complexes b1 – b4 are formed in equal amounts with either of the extracts.

In tobacco flowers chalcone synthase is expressed predominantly in the pink pigmented limb of petals rather than the unpigmented tube (Drews and Goldberg, 1988). To characterize CG-1-DNA binding in these tissues, nuclear extracts were prepared from the pink pigmented distal petal region and the unpigmented tube. To evaluate the quality of nuclear extracts prepared from these tissues we determined the ratio of nuclear protein in the extract to the DNA content in the nuclei preparation. Both extracts have an identical protein/DNA ratio (data not shown). Therefore we used equal amounts of nuclear protein from either the pink pigmented limb or the unpigmented tube for band-shift analysis. The results shown in Fig. 2C indicated that both flower extracts formed complexes of similar specificity and apparently equal quantity. In addition to b1 – b4 complexes observed in seedlings (cf. Fig. 2B) a low-molecular-mass complex was detected (marked with an open arrow in Fig. 2C) that either arose by proteolytic degradation or binding of an additional polypeptide to the CG-1 recognition sequence.

CG-1-binding activity was also found in stem and leaf tissue from adult tobacco plants as well as in tobacco cell cultures (data not shown).

### *Photochemical cross-linking of polypeptides interacting with the CACGTG motif*

To identify polypeptides binding to the CACGTG motif, a 5-bromodeoxyuridine-containing oligonucleotide was used to selectively label polypeptides by photochemical cross-linking. This photoaffinity probe was synthesized by incorporating 5-bromodeoxyuridine and radioactive deoxycytosine into a partially double-stranded oligonucleotide, containing one copy of the CG-1-binding site (Fig. 1). Incubation of nuclear proteins enriched for CG-1 by DEAE-Sepharose chromatography (see below) with this monomeric CG-1-binding site gave rise to a broad retarded band (Fig. 3A). The complex was specifically competed by wild-type oligonucleotide 1, but not by oligonucleotide 3 which contains three base changes resulting in loss of CG-1 binding (Fig. 1). This complex was subjected to further analysis by cross-linking. In contrast to the pattern seen on five tandem copies of oligonucleotide 1 a high-molecular-mass complex formed on the monomeric binding site was apparent only as a faint band, perhaps due to the lack of stabilizing effects of flanking sequences. This complex was not further analyzed.

Preparative scale binding assays were run on a retardation gel and cross-linking was performed by irradiation of the gel with an ultraviolet-light lamp (312 nm wavelength). The DNA-protein complex was excised from the gel and cross-linked polypeptides were resolved by SDS/polyacrylamide gel electrophoresis (Fig. 3B; + UV). As a control DNA-protein complexes were isolated from a preparative retardation gel omitting ultraviolet-light-activated cross-linking (Fig. 3B; – UV). Only after cross-linking, a major band of approximately 20 kDa and two minor bands of approximately 32 kDa and 42 kDa were selectively labelled by the covalent attachment of the oligonucleotide. These molecular masses are upper estimates because at least a portion of the labelled DNA remained attached on the protein. However, experience, including our own, suggests that oligonucleotides of the size used here do apparently not have a very significant effect on protein mobility in the gel system used (Sorger and Pelham, 1987; Riddihough and Pelham, 1987).

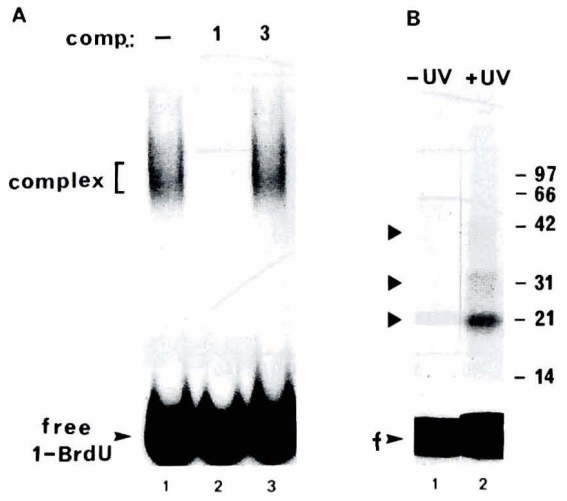


Fig. 3. *Identification of polypeptides binding to the CACGTG motif by photoactivated cross-linking.* (A) Gel-shift assay using $^{32}$P-labelled BrdU-substituted oligonucleotide 1-BrdU as probe. The binding reaction contained 5 µg nuclear protein (0.2 M KCl step from DEAE-Sepharose) and 1.5 µg *Hpa*II-cut pUC19 DNA (lane 1) or 1.5 µg linearized pUC19 DNA containing five copies of either oligonucleotide 1 (lane 2) or oligonucleotide 3 (lane 3); comp, competitor DNA. (B) SDS/polyacrylamide gel electrophoresis of polypeptides involved in complex formation. Protein-DNA complexes were eluted from preparative scale retardation gels after cross-linking (+ UV) and from control gels without cross-linking (– UV) as detailed in Materials and Methods and loaded on a 10–20% gradient gel. The position of the molecular-mass markers (kDa; prestained markers, Bio-Rad) are indicated. f, free probe

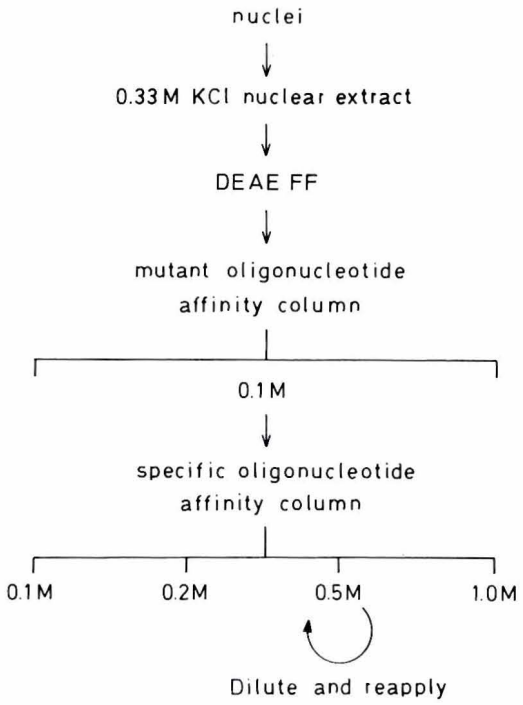


Fig. 4. *Flow diagram of the purification procedure.* Precise details of the column chromatography are given in Materials and Methods. FF, fast flow

*Extraction of CG-1 from tobacco nuclei*

In order to minimize contamination of CG-1-containing nuclear fractions with other proteins, we optimized the extraction procedure. The major CG-1 fraction was extracted with 0.33 M KCl. Upon re-extraction of the nuclei with 0.8 M KCl we recovered a large amount of unspecific DNA-binding activity but little CG-1. It has been reported that certain sequence-specific DNA-binding proteins can be selectively released by DNA intercalating drugs with little contamination of other nuclear proteins (Schröter et al., 1987). To test whether CG-1 could be purified in a similar manner nuclei were extracted with 10 mM chloroquine. However, no substantial release of CG-1 was observed as assayed by band-shift gels (data not shown).

*Chromatographic fractionation of CG-1*

Fractionation steps were monitored by band-shift assay and binding specificity at each purification step was confirmed by competition with wild-type and mutated binding sites. The purification is outlined in the flow diagram shown in Fig. 4.

Crude 0.33 M KCl nuclear extract was desalted on Sephadex G-50 and bound to DEAE-Sepharose at 50 mM KCl. Protein fractions were eluted with a linear 50–300 mM KCl gradient (Fig. 5). CG-1 was eluted as a broad peak at 125–200 mM KCl as determined by conductivity measurement of the active fractions (Fig. 5). Polypeptides involved in the formation of complexes b3 and b4 eluted at slightly lower ionic strength than polypeptides giving rise to the higher-molecular-mass complexes b1 and b2. To rapidly elute concentrated CG-1 fractions in repeated purifications a step elution with 200 mM KCl was used.

The key step of the purification involved differential sequence-specific affinity chromatography using on the one hand the wild-type CG-1-binding site (oligonucleotide 1, Fig. 1) and on the other hand oligonucleotide 3 carrying three base changes that virtually abolish binding. The wild-type affinity matrix was constructed by concatamerization of oligonucleotide 1 and subsequent coupling to CNBr-activated Sepharose. To minimize the contamination of the high-salt eluate from the wild-type affinity column by other proteins binding fortuitously to CACGTG-flanking sequences or to the matrix, a mutant oligonucleotide column was prepared containing concatamerized oligonucleotide 3. The mutant and wild-type columns were run sequentially, as schematically shown in Fig. 4. The flow-through fraction from the mutant oligonucleotide column containing most of the CG-1-binding activity was directly loaded onto the specific column after mixing with competitor DNA-{5 µg/ml poly[d(I-C)] and 5 µg/ml poly[d(A-T)]}. The wild-type column was washed with 0.1 M KCl and eluted stepwise with 0.2 M, 0.5 M and 1.0 M KCl with CG-1 binding confined to the 0.5 M KCl fractions (Fig. 6A). Samples of the chromatographic fractions were analyzed on SDS/PAGE (Fig. 6B). Most of the proteins were not retained by the column. Substantial but incomplete purification of CACGTG-binding factors was obtained. For example, three closely spaced prominent polypeptides in the 20-kda range of the loaded protein fraction were present in the flow-through fraction containing little CG-1-binding activity as well as in the 0.5 M KCl eluate containing the peak binding activity and therefore most likely represent non-specific interaction. The amount of these bands was further reduced by a second cycle of affinity chromatography. Again, CG-1-binding activity was recovered in the 0.5 M KCl eluate.

To achieve a further substantial degree of purification CG-1-containing fractions were diluted to 0.1 M KCl and supplemented with calf thymus DNA (30 µg/ml) as an alterna-

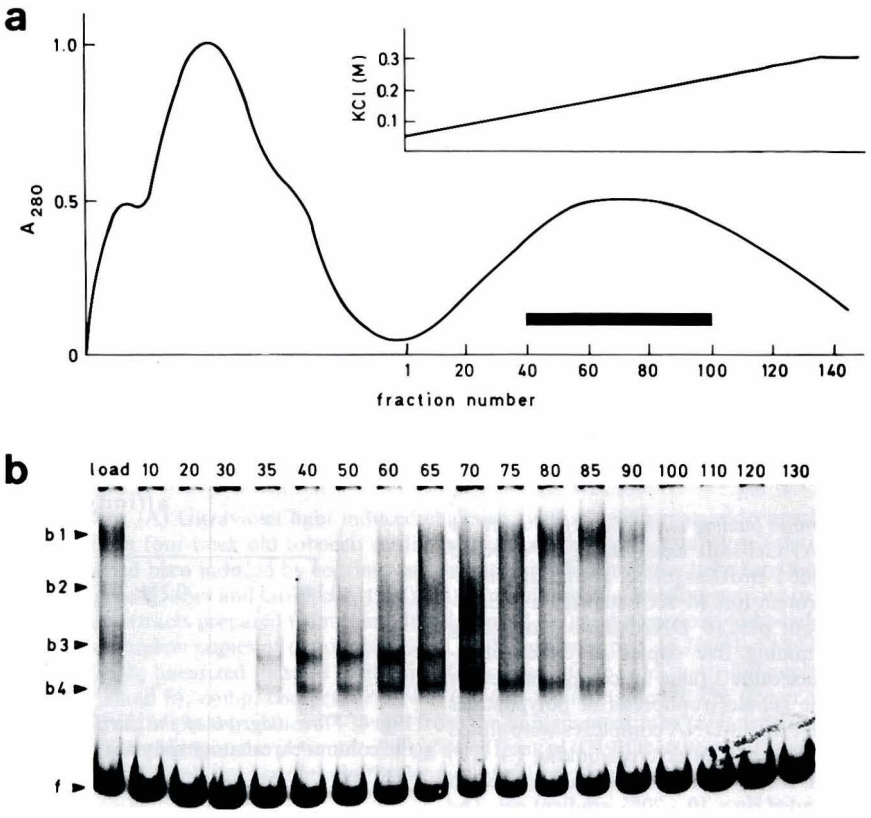


Fig. 5. *DEAE-Sepharose column fractionation of CG-1.* (a) The linear KCl gradient (0.05—0.3 M) was started at fraction 1. Peak fractions are indicated by a bar. (b) Band-shift assay to monitor CG-1 binding activity across the fractions. 8-µl aliquots of DEAE-Sepharose fractions were reacted with 0.5 ng labelled five tandem copies of oligonucleotide 1 in the presence of 2 µg *Hpa*II-cut pUC19 DNA

tive competitor DNA in order to suppress binding of proteins which presumably are not part of the CG-1-specific complexes. In addition, the concentration of Nonidet P-40 detergent was raised to 0.1%. During the third cycle of 'specific' affinity chromatography CG-1 was again recovered in the 0.5 M KCl eluate, giving rise predominantly to two shifted bands corresponding to complexes b3 and b4 in band-shift assays whereas higher-molecular-mass complexes were barely visible (Fig. 6C, left gel). Sequence specificity was verified by competition with unlabelled oligonucleotide 1. Oligonucleotide 3 did not interfere with binding (Fig. 6C, right gel). Analysis by SDS/PAGE revealed a highly enriched polypeptide in the CG-1-containing fraction whose apparent molecular mass was determined to be 20 kDa, and two additional polypeptides of approximately 30 kDa and 40 kDa present in small amounts (Fig. 6D). The molecular masses and the relative abundance of these proteins correlated well with the polypeptide species identified earlier by photoaffinity labelling in the DEAE-Sepharose peak fractions.

Table 1 summarizes the enrichment of the proteins binding to the CACGTG motif from tobacco nuclear extracts resulting in a purification of more than 7000-fold and a yield of approximately 0.6% of the activity present in crude nuclear exctracts.

## DISCUSSION

The nuclear factor(s) CG-1 recognize(s) CACGTG motifs residing within the promoter region of various plant genes (Giuliano et al., 1988; Schulze-Lefert et al., 1989a; Staiger et al., 1989; Schindler and Cashmore,1990). CACGTG motifs have been shown to play a crucial role in the regulation of the chalcone synthase gene in parsley (Schulze-Lefert et al., 1989a, b), a ribulose-1,5-bisphosphate carboxylase small subunit gene in *Arabidopsis* (Donald and Cashmore, 1990) and a dihydroflavonol reductase (*pallida*) gene in *Antirrhinum* (Coen et al., 1986; Almeida et al., 1989). We have initiated the purification of the factor(s) interacting with the CACGTG motif contained within the ultraviolet-light-responsive region of the *A. majus* chalcone synthase gene promoter and have characterized it by photoaffinity labelling.

### *CG-1 factor(s) binding to the chalcone synthase CACGTG motif is constitutively expressed*

The results described above indicate that expression of CG-1 is constitutive in tobacco seedlings, and is not affected by 18 h of ultraviolet-light irradiation which induces the chalcone synthase gene. This suggests that transcriptional activation of the chalcone synthase gene by ultraviolet light does not require *de novo* synthesis of CG-1. Modification of the CG-1 DNA-binding domain upon induction seems also unlikely, unless lost during extract preparation, since identical protein-DNA complexes are formed in both types of extracts (Fig. 2B) as well as in extracts prepared from seedlings irradiation for only 4 h (at this time chalcone synthase mRNA starts to accumulate; Staiger, 1989). Similarly, identical *in vitro* binding of the G-box-binding factor to the G box of the ribulose-1,5-bisphosphate carboxylase gene and chlorophyll-*a*/*b*-binding protein gene promoters has been observed in extracts prepared from light-grown and dark-adapted plants

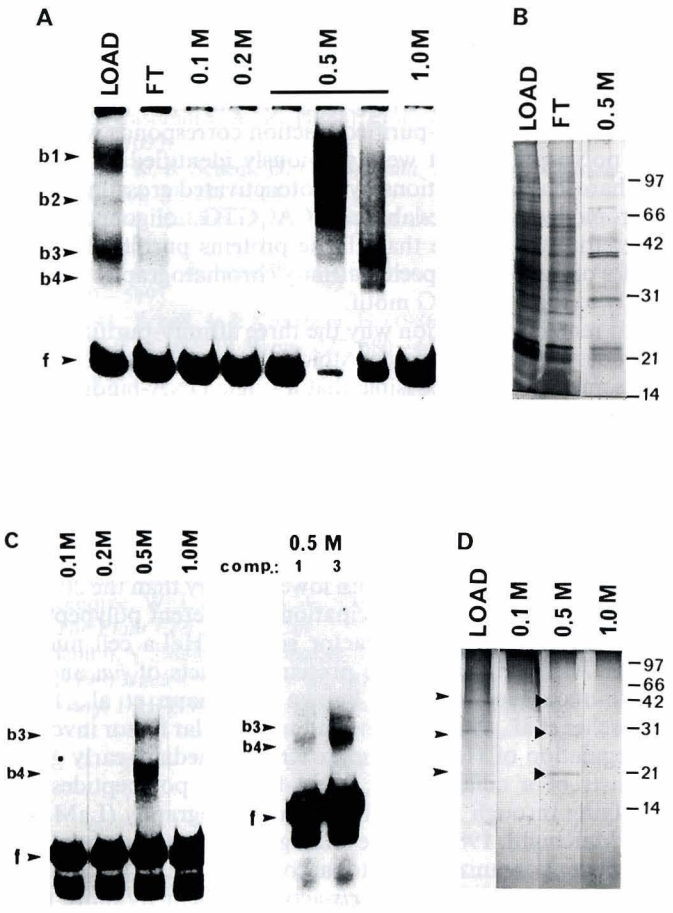


Fig. 6. *Oligonucleotide affinity chromatography*. (A) Band-shift assay of first cycle. 4 µl of the fractions indicated were reacted with five tandem copies of oligonucleotide 1 in the presence of 500 ng *Hpa*II-cut pUC19 DNA. Load, 0.1 M eluant of the mutant oligonucleotide-3—Sepharose column which was loaded directly on the specific affinity column; Ft, flow-through of the specific affinity column. (B) Silver-stained 12% SDS polyacrylamide gel. Samples: 2 µl of the load; 2 µl flow-through (FT); 100 µl of the 0.5 M KCl step which were concentrated by precipitation with 10% trichloroacetic acid/0.15% sodium deoxycholate. (C) Band-shift assay of the affinity-purified fractions of the third cycle. For competition, 500 ng linearized pUC19 DNA containing either five copies of oligonucleotide 1 or of mutant oligonucleotide 3 were included in the binding assay with 4 µl of the 0.5 M KCl peak fraction; comp, competitor DNA. (D) Silver-stained 15% SDS polyacrylamide gel containing 900 µl of the load and 300 µl each of the flow-through fractions, 0.1 M, 0.2 M, 0.5 M and 1.0 M KCl eluate of the third cycle after precipitation with trichloroacetic acid. The positions of molecular-mass markers (kDa) are indicated

(Schindler and Cashmore, 1990). In addition to being involved in protection against ultraviolet light, flavonoids play a major role in floral pigmentation. In tobacco flowers, chalcone synthase is highly expressed in the pink pigmented limb, but only very weakly expressed in the unpigmented tube (Drews and Goldberg, 1988). The same expression pattern is conferred by the *A. majus* chalcone synthase promoter upon a linked β-glucuronidase reporter gene in transgenic tobacco plants (K. Fritze, D. Staiger, I. Czaja, R. Walden, J. Schell and D. Wing, unpublished results). CG-1 was found in the pink pigmented limb as well as in the unpigmented tube (Fig. 2C), again indicating that its expression is not correlated with chalcone synthase expression. In *A. majus* flowers, CG-1 binding to the CACGTG motif was also found in both tubes and lobes (D. Staiger, unpublished results). This agrees with genetic evidence indicating that in *A. majus* a CACGTG motif contained within the promoter of the *pallida* gene encoding dihydroflavonol reductase, a further flavonoid-biosynthesis enzyme acting after chalcone synthase in the pathway, binds common regulatory factors rather than tissue-specific *trans*-acting factors in flowers (Almeida et al., 1989). Final proof of whether the CG-1 factor(s) identified in nuclear extracts of tobacco and *A. majus* are directly involved in regulating expression of the *Antirrhinum* chalcone synthase gene awaits molecular cloning and functional analysis of the corresponding genes *in vivo*. Ultraviolet-light-dependent changes in protein binding to the parsley chalcone synthase G box *in vivo* show that factors binding to G box elements participate in regulation of ultraviolet-light-induced chalcone synthase gene expression (Schulze-Lefert et al., 1989a).

### *CG-1 is composed of multiple polypeptides*

Photoactivated cross-linking identified three polypeptides of apparent molecular masses of approximately 20, 32 and 42 kDa which specifically bind a photoaffinity probe for the CACGTG motif. This indicates that more than one nuclear factor can bind to the same CACGTG motif. It is apparent that the 20-kDa polypeptide either is more abundant in nuclear extracts or cross-links more efficiently than the 32-kDa and 42-kDa polypeptides. Alternatively, the lower-molecular-mass bands might represent proteolytic degradation products of the 42-kDa protein. Similar observations have been reported for the *Arabidopsis Adh* gene G box which cross-linked to 18-kDa and 31-kDa polypeptides in crude whole-cell extracts of *Arabidopsis* leaves and cell cultures (DeLisle and Ferl, 1990).

To characterize further nuclear factor(s) CG-1, a purification scheme was developed. The low efficiency of cross-

Table 1. *Purification of CG-1*
For specific activity, 1 U binding activity is the amount of CG-1 that, under standard gel-retardation assay conditions, retards 1 fmol end-labelled five tandem copies of oligonucleotide 1

| Fraction | Volume | Total protein | Total activity | Specific activity | Purification | Yield |
|---|---|---|---|---|---|---|
| | ml | mg | U | U/mg | -fold | % |
| Nuclear extract | 400 | 380 | 30000 | 79 | 1 | 100 |
| DEAE-Sepharose | 86 | 95 | 28500 | 300 | 3.8 | 95 |
| Mutant oligo-Sepharose | 172 | 77.2 | 24589 | 318.5 | 4 | 82 |
| Specific oligo-Sepharose 1 | 10 | 0.375 | 4057 | 10820 | 137 | 13.5 |
| 2 | 100 | 0.04 | 2467 | 61687 | 780 | 8.2 |
| 3 | 2 | 0.0003 | 173 | 576666 | 7299 | 0.6 |

linking of proteins to the photoaffinity DNA probe (at least 90% of the DNA fragment contained in the retarded complexes after photoactivated cross-linking migrated as free fragments in denaturing polyacrylamide gels, Fig. 3B) prevented us using this technique to affinity-label CG-1 polypeptides in crude extracts. Therefore, purification under denaturing conditions monitored by liquid scintillation counting of the radioactive affinity label was not feasible. Instead, a binding site affinity procedure was used and purification of CG-1 factor(s) was monitored by *in vitro* DNA-binding assays.

To minimize the loss of nuclear proteins caused by partial disruption of nuclei a modified nuclei isolation protocol was employed that omitted extensive purification of nuclei through Percoll gradients. Attempts to selectively release CG-1 from tobacco nuclei by the use of the intercalating drug chloroquine were not successful. This technique has proven itself to be an efficient first step in the purification of HeLa cell protein p67 which binds to the c-fos serum-response element and could be extracted with 20-fold enrichment in comparison to conventional nuclear salt extracts (Schröter et al., 1987). Therefore, 0.33 M KCl extracts of nuclei were prepared and chromatographed on DEAE-Sepharose to remove residual nucleic acids. CG-1 could be recovered in a concentrated form in a buffer of medium ionic strength (0.2 M KCl) allowing subsequent application to a sequence-specific oligonucleotide column without prior dialysis. The crucial purification step was the use of differential sequence-specific affinity chromatography. Our mutational analysis of binding-site requirements of CG-1 factor(s) demonstrated that mutation of three G residues almost completely eliminated CG-1 binding (Staiger et al., 1989). On this basis we constructed a mutant oligonucleotide Sepharose column. The efficiency of this DNA column in eliminating nonspecific binding proteins from the affinity-purified fractions was previously confirmed by assaying the column fractions for binding of a different tobacco nuclear factor which recognizes a positive regulatory upstream element within the *A. majus* chalcone synthase promoter (Staiger et al., 1990). This factor was found in the flow-through fraction of the tandem columns and in the 0.5 M KCl step of the mutant Sepharose column but not in the high-salt eluant from the specific affinity column (data not shown).

Substantial purification of proteins binding specifically to the CACGTG motif was achieved as judged by the diversity of polypeptides present in the DEAE-Sepharose pool as compared to the peak fraction after three sequential passes over the DNA affinity column. However, the protocol did not yield a single polypeptide species copurifying with sequence-specific binding to the CACGTG motif but rather a collection of polypeptides ranging over 20–40 kDa (see Fig. 6D). Peak fractions of the third round of specific affinity chromatography resulted predominantly in the two faster-migrating complexes, b3 and b4, in band-shift assays, indicating that this step largely separated away proteins involved in b1 and b2 complexes (see Fig. 6C). The observation that proteins from b1 and b2 complexes also eluted in a slightly earlier peak from the DEAE-Sepharose column than those of b3 and b4 complexes similarly indicated that they are chromatographically separable entities (see Fig. 5B). The low yield of purified CG-1-binding polypeptides precluded renaturation of the activity from SDS/polyacrylamide gel slices and to demonstrate unambiguously the identity of the polypeptides with components of CG-1. Two lines of evidence, however, strongly suggest that the purified polypeptides belong to nuclear factor(s) CG-1. The coelution of the three polypeptides with CG-1-DNA-binding activity during the third cycle of specific affinity chromatography suggests that probably all of them are responsible for the sequence-specific binding of the affinity-purified fraction (Fig. 6C). Moreover, the polypeptide profile of the affinity-purified fraction corresponds well with the polypeptides that were previously identified in DEAE-Sepharose active fractions by photoactivated cross-linking to 5-bromodeoxyuridine-labelled CACGTG oligonucleotide. Therefore we assume that all the proteins purified by three cycles of CACGTG-specific affinity chromatography bind directly to the CACGTG motif.

It remains a question why the three affinity-purified polypeptides appear at different stoichiometry. If they represent the same factor it is possible that a single DNA-binding protein has been proteolyzed and degradation products of this larger factor remain associated during fractionation or the 40-kDa band may arise by dimerization of the 20-kDa polypeptide under the experimental conditions. Alternatively, it is possible that different CACGTG-binding polypeptides form a multicomponent complex where the 30-kDa and 40-kDa proteins bind to the DNA with lower affinity than the 20-kDa protein. Examples for participation of different polypeptides in a single DNA-binding factor such as HeLa cell nuclear factor AP1, comprising the protein products of *jun* and *fos* proto-oncogenes, are well known (Bohmann et al., 1987; Rauscher et al., 1988). Furthermore, a cellular factor involved in regulation of Herpes simplex virus immediate-early genes consists of a collection of DNA-binding polypeptides copurifying through DNA affinity chromatography (LaMarco and McKnight, 1989). These analogies suggest that the multicomponent organization of tobacco factor(s) CG-1 may allow optimal use of the cognate *cis*-acting element by more than one nuclear factor that may regulate promoter activity independently in response to different external stimuli. The identification of multiple tobacco nuclear polypeptides involved in CACGTG recognition provides basic information for further exploration of this question at the molecular level.

We thank B. Gronenborn for supplying the oligonucleotides, G. Drews and R. Goldberg for the tobacco chalcone synthase 2 clone, A. Bachmair for constructive discussions and I. Moore and M. Hergersberg for reading of the manuscript. D. S. was supported by a fellowship from the Fritz Thyssen Foundation.

## REFERENCES

Almeida, J., Carpenter, R., Robbins, T. P., Martin, C. & Coen, E. S. (1989) *Genes & Dev. 3*, 1758–1767.

Bohmann, D., Bos, T. J., Admon, A., Nishimura, T., Vogt, P. K. & Tjian, R. (1987) *Science 238*, 1386–1392.

Bradford, M. M. (1976) *Anal. Biochem. 72*, 248–254.

Brouillard, R. (1988) in *The flavonoids, advances in research since 1980* (Harborne, J. B., ed.) pp. 525–538, Chapman and Hall, London.

Burton, R. (1968) *Methods Enzymol. 12*, 163–166.

Cereghini, S., Blumenfeld, M. & Yaniv, M. (1988) *Genes & Dev. 2*, 957–974.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Coen, E. S., Carpenter, R. & Martin, C. (1986) *Cell 47*, 285–296.

Dangl, J. L., Hahlbrock, K. & Schell, J. (1989) *Cell Cult. Somatic Cell Genet. Plants 6*, 155–173.

DeLisle, A. J. & Ferl, R. J. (1990) *The Plant Cell 2*, 547–557.

Donald, R. G. K. & Cashmore, A. R. (1990) *EMBO J. 9*, 1717–1726.

Donald, R. G. K., Schindler, U., Batschauer, A. & Cashmore, A. R. (1990) *EMBO J. 9*, 1727–1735.

Drews, G. & Goldberg, R. (1988) *J. Cell. Biol. 12C* Suppl., 145.

Fried, M. G. & Crothers, D. M. (1981) *Nucleic Acids Res. 9*, 6505–6525.
Garner, M. M. & Revzin, A. (1981) *Nucleic Acids Res. 9*, 3047–3060.
Giuliano, G., Pichersky, E., Malik, V. S., Timko, M. P., Scolnik, P. A. & Cashmore, A. R. (1988) *Proc. Natl Acad. Sci. USA 85*, 7089–7093.
Hahlbrock, K. & Scheel, D. (1989) *Annu. Rev. Plant Physiol. Plant Mol. Biol. 40*, 347–364.
Heller, W. & Hahlbrock, K. (1980) *Arch. Biochem. Biophys. 200*, 617–619.
Kadonaga, J. T. & Tjian, R. (1986) *Proc. Natl Acad. Sci. USA 83*, 5889–5893.
Kaulen, H., Schell, J. & Kreuzaler, F. (1986) *EMBO J. 5*, 1–8.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
LaMarco, K. L. & McKnight, S. L. (1989) *Genes & Dev. 3*, 1372–1383.
Lamb, C. J., Lawton, M. A., Dron, M. & Dixon, R. A. (1989) *Cell 56*, 215–224.
Lipphardt, S., Brettschneider, R., Kreuzaler, F., Schell, J. & Dangl, J. L. (1988) *EMBO J. 7*, 4027–4033.
Maniatis, T., Fritsch, E. F. & Sambrock, J. (1982) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
McKendree, W. L., Paul, A.-L., DeLisle, A. J. & Ferl, R. J. (1990) *The Plant Cell 2*, 207–214.
Moncollin, V., Stalder, R., Verdier, J.-M., Sentenac, A. & Egly, J.-M. (1990) *Nucleic Acids Res. 18*, 4817–4823.
Morissey, J. H. (1981) *Anal. Biochem. 117*, 307–310.
Peters, N. K., Frost, J. M. & Long, S. R. (1986) *Science 233*, 977–980.
Rauscher, F. J. III, Voulalas, P. J., Franza, B. R. Jr. & Curran, T. (1988) *Genes & Dev. 2*, 1687–1699.
Riddihough, G. & Pelham, H. R. B. (1987) *EMBO J. 6*, 3729–3734.
Sawadogo, M. & Roeder, R. G. (1985) *Cell 43*, 165–175.
Schindler, U. & Cashmore, A. R. (1990) *EMBO J. 9*, 3415–3427.
Schmelzer, E., Jahnen, W. & Hahlbrock, K. (1988) *Proc. Natl Acad. Sci. USA 85*, 2989–2993.
Schulze-Lefert, P., Dangl, J. L., Becker-Andre, M., Hahlbrock, K. & Schulz, W. (1989a) *EMBO J. 8*, 651–656.
Schulze-Lefert, P., Becker-Andre, M., Schulz, W., Hahlbrock, K. & Dangl, J. L. (1989b) *The Plant Cell 1*, 707–714.
Schröter, H., Shaw, P. E. & Nordheim, A. (1987) *Nucleic Acids Res. 15*, 10145–10158.
Scotto, K. W., Kaulen, H. & Roeder, R. G. (1989) *Genes & Dev. 3*, 651–661.
Sommer, H. & Saedler, H. (1986) *Mol. Gen. Genet. 202*, 429–434.
Sorger, P. K. & Pelham, H. R. B. (1987) *EMBO J. 6*, 3035–3041.
Staiger, D., Kaulen, H. & Schell, J. (1989) *Proc. Natl Acad. Sci. USA 86*, 6390–6394.
Staiger, D., Kaulen, H. & Schell, J. (1990) *Plant Physiol. 93*, 1347–1353.
Staiger, D. (1989) PhD thesis. University of Cologne, FRG.
van der Meer, I. M., Spelt, C. E., Mol, J. N. M. & Stuitje, A. R. (1990) *Plant Mol. Biol. 15*, 95–109.
Wu, C., Wilson, S., Walker, B., Dawid, I., Paisley, T., Zimarino, V. & Ueda, H. (1987) *Science 238*, 1247–1253.