Plant Molecular Biology **51:** 9–19, 2003.

G-box binding coincides with increased *Solanum melongena* cysteine proteinase expression in senescent fruits and circadian-regulated leaves

Zeng-Fu Xu[1,3], Mee-Len Chye[1,*], Hong-Ye Li[1], Fang-Xiu Xu[1] and Kwok-Ming Yao[2]
[1]*Department of Botany, University of Hong Kong, Pokfulam Road, Hong Kong, China (*author for correspondence; e-mail: mlchye@hkucc.hku.hk);* [2]*Department of Biochemistry, University of Hong Kong, Hong Kong, China; present address: Key Laboratory of Gene Engineering of MOE, Zhongshan University, Guangzhou, SI0275, China*

Received 20 November 2001; accepted in revised form 30 April 2002

Key words: cis-elements, clock, DNase I footprinting, eggplant, ethylene

Abstract

We have previously shown that *SmCP*, the gene encoding *Solanum melongena* cysteine proteinase, is expressed during developmental events associated with programmed cell death (PCD) suggesting its involvement in protein degradation during these events (Xu and Chye, Plant Journal 17 (1999) 321–327). Here, we investigated the regulation of *SmCP* expression and showed that it is ethylene-inducible and is under circadian control. This circadian rhythm is entrained by light/dark (LD) cycling with peak expression in the late light period, as opposed to that in early light for *rbcS*, suggesting that protein degradation and photosynthesis are temporally separated by circadian control. Northern blot analysis shows that the pattern of ethylene induction of *SmCP* is consistent with our previous observation of its significantly increased expression at leaf senescence and fruit ripening when endogenous ethylene is abundant. To further understand *SmCP* regulation, we have cloned the *SmCP* promoter and identified a G-box (CACGTG) at −85/−80 by DNase I footprinting analysis of the −221/+17 region. Its specific interaction with nuclear proteins in *S. melongena* leaves and fruits was confirmed by competitive electrophoretic mobility shift assays using oligonucleotides containing the G-box and mutant derivatives. G-box binding activity was stronger in senescent than young fruits. In circadian-regulated leaves, stronger binding activity coincided with peak circadian expression of *SmCP*. This correlation between binding activity and expression suggests that G-box binding factors enhance *SmCP* transcription and that the G-box likely plays a role in circadian regulation of genes affected by LD cycling.

Introduction

PCD in plants occurs when unwanted or damaged cells are removed during the hypersensitive response to pathogen invasion and during developmental events including organ senescence, fruit ripening, xylogenesis, aleurone degeneration and valve dehiscence in seed pods (Pennell and Lamb, 1997; Buckner *et al.*, 1998). Proteinases participate in PCD by degrading proteins within the affected cells (Callis, 1995). Cysteine proteinases (thiol proteinases) are one class of such proteinases involved in this process; the others being aspartic acid proteinases, metalloproteinases and serine proteinases (Callis, 1995). We have previously shown that the expression of a single gene encoding cysteine proteinase in *Solanum melongena* (brinjal or eggplant), designated *SmCP*, coincides with developmental events associated with PCD (Xu and Chye, 1999). Localization of *SmCP* mRNA by *in situ* hybridization studies to the xylem, the epidermis and the endothecium of the anther, and the nucellar cells of the ovule, suggests a role for SmCP in protein degradation during xylogenesis, anther senescence and ovule development, respectively (Xu and Chye,

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number AF101032 (promoter of *S. melongena* cysteine proteinase gene, *SmCP*)

1999). Northern blot analysis showed that *SmCP* expression increases significantly at senescence in leaf and fruit (Xu and Chye, 1999).

The expression of *SmCP* is not limited to a specific organ or cell type (Xu and Chye, 1999) in contrast to cysteine proteinases associated with leaf senescence (Hensel *et al.*, 1993; Lohman *et al.*, 1994; Drake *et al.*, 1996; Buchanan-Wollaston and Ainsworth, 1997), petal senescence (Jones *et al.*, 1995; Guerrero *et al.*, 1998) or xylogenesis (Ye and Varner, 1996) which share low (34–41%) identity to SmCP. Unlike the expression of actinidin, which does not peak in ripened fruits (Praekelt *et al.*, 1998) and that of tomato cysteine proteinase which is highly expressed at leaf but not fruit senescence (Drake *et al.*, 1996), *SmCP* expression rises significantly in both senescent leaves and ripened fruits (Xu and Chye, 1999). Since the expression pattern of *SmCP* differs from those of other cysteine proteinases, we have initiated investigations towards understanding *SmCP* regulation. Here, we show that *SmCP* is ethylene-inducible and is under circadian control. We further cloned the *SmCP* promoter region and identified a G-box that binds nuclear factors in leaves and fruits by DNase I footprinting and electrophoretic mobility shift assays (EMSA). We first report the presence of a G-box in a plant proteinase promoter and of circadian regulation of a PCD-associated gene.

Materials and methods

Plant material, growth conditions and treatments

Unless otherwise stated, *S. melongena* L. plants were grown in a growth chamber at 24 °C with a light/dark (LD) regime of 12 h light (08:00–20:00) and 12 h dark (20:00–08:00). Ethephon treatment was carried out on 5-week old seedlings as described by Greenberg and Ausubel (1993) and young leaves were harvested at 06:00 (time of treatment), 09:00 and 12:00 for RNA extraction in northern blot analysis. Young leaves from the untreated control, similarly grown, were also harvested for RNA extraction. In experiments on circadian control, 5-week old seedlings germinated at 24 °C in 12 h light (08:00–20:00)/12 h dark (20:00–08:00) were shifted to either continuous light or continuous darkness at 24 °C. In another experiment, 5-week old seedlings, germinated in continuous light at 24 °C, were subjected to either thermocycling (12 h 21 °C/12 h 24 °C) or LD and thermocycling (12 h light 08:00–20:00 24 °C/12 h dark 20:00–08:00 21 °C).

Plant DNA isolation and Southern blot analysis

DNA was isolated from *S. melongena* leaves as previously described (Dellaporta *et al.*, 1983). For Southern blot analysis, 25 μg genomic DNA was digested, separated on a 0.8% agarose gel and transferred onto Hybond-N membrane (Amersham) following the procedure of Sambrook *et al.* (1989). The blot was pre-hybridized in a solution containing 30% deionized formamide, 6× SSC, 5× Denhardt's solution, 1% SDS, 50 μg/ml denatured, sonicated salmon sperm DNA at 42 °C for 4 h. Probes labeled with [^{32}P]dCTP were prepared by random-primed labeling either the full-length *SmCP* cDNA (Xu and Chye, 1999) or the 5′-end *Eco*RI fragment of *SmCP* cDNA (nucleotides 1–993 of the cDNA; nucleotide 1 of the cDNA corresponds to position +29 of the gene in Figure 4a). After overnight hybridization with the probe, the blot was washed in 0.1× SSC, 0.1% SDS at room temperature.

Construction and screening of S. melongena *genomic library*

Genomic DNA was digested to completion with *Eco*RI and fragments of ca. 5 kb were obtained by NaCl gradient ultra-centrifugation and cloned into predigested λZAP II/*Eco*RI/CIAP (Stratagene). This genomic library was screened with a [^{32}P]dCTP-labeled probe prepared from the 5′-end *Eco*RI fragment of *SmCP* cDNA. Plaque lifts with Hybond-N membranes (Amersham) and DNA hybridization were according to the manufacturer's recommendations. The pBluescript phagemid containing the insert DNA was rescued from the purified λ phage by co-infection with helper phage ExAssist (Stratagene).

DNA sequencing and analysis

DNA fragments containing the sequence of interest were sequenced using the DNA Sequencing Kit with Sequenase Version 2.0 according to the instructions of the manufacturer (UBS, Amersham Life Science). The DNA sequence data was analyzed with the GCG sequence analysis software package (Genetics Computer Group).

Northern blot analysis

Total RNA was isolated from *S. melongena* by the method of Nagy *et al.* (1988b). For northern blot analysis, 20 μg RNA was denatured at 50 °C in the presence of glyoxal, separated by electrophoresis in 1.5% agarose gel and blotted onto Hybond-N (Amersham) membranes. To ensure that equal amounts of RNA were used, RNA was stained with ethidium bromide after gel electrophoresis, before blotting. The RNA blot was hybridized at 42 °C with [^{32}P]dCTP-labeled *SmCP* cDNA in a solution containing 50% deionized formamide, 1× Denhardt's solution, 6× SSPE, 0.1% SDS, 100 μg/ml denatured sonicated salmon sperm DNA and 10% dextran sulfate. The blot was washed at 65 °C in 0.1× SSC, 0.1% SDS. With the use of the *Hevea rbcS* cDNA (Chye *et al.*, 1991) probe, the blot was washed at room temperature in 0.1× SSC, 0.1% SDS. Bands were detected by autoradiography. Autoradiographs were scanned with a densitometer (BioRad Gel Doc) with Quantity One v4.2 software.

Primer extension analysis

To map the 5′ end of the *SmCP* mRNA, a [^{32}P]ATP-end-labeled oligomer (5′-AGAAATGATGCTCGGCG-3′) complementary to positions +181 to +197 of *SmCP* was hybridized to 30 μg of total RNA from *S. melongena* fruit. After extension with AMV reverse transcriptase (Ausubel *et al.*, 1995), the product of the primer extension reaction was electrophoresed in a 6% polyacrylamide gel against sequencing reactions with the same oligomer and sequencing template prepared from the genomic clone containing the 5′ end of *SmCP*.

Preparation of nuclear protein extracts

Nuclear protein extracts were prepared according to Martino-Catt and Kay (1994) from young fruits at 20 days after pollination (DAP), mature fruits at 50 DAP and senescent fruits at 70 DAP. Leaf extracts were prepared from 5-week old seedlings, grown in 12 h light (08:00–20:00)/12 h dark (20:00–08:00), at 17:00 (peak circadian expression) and 06:00 (low circadian expression). Protein concentration of extracts was determined by the method of Bradford (1976) with BioRad Protein Assay Kit I.

In vitro *DNase I footprinting*

Coding strand and non-coding strand probes were prepared by linearizing plasmid pSm108, a pGEM-T Easy (Promega) derivative containing the *SmCP* promoter region (−221/+17), with *Nco*I or *Spe*I, respectively, and end-labeling with [$\alpha-^{32}$P]dCTP by filling-in 3′ recessed ends with Klenow. After removal of unincorporated [$\alpha-^{32}$P]dCTP in Microspin G-25 columns (Amersham Pharmacia Biotech), the ^{32}P-labeled probes were released from the labeled linearized plasmids by digestion with *Spe*I or *Nco*I, respectively. Probes were purified with a preparative non-denaturing 5% polyacrylamide gel. DNase I (0.75 units, Amersham Pharmacia Biotech, FPLCpure) was added to 25 μl of reaction mixtures containing ca. 20 000 cpm of ^{32}P-labeled probe (quantified by scintillation counting), 2 μg of poly (dI-dC), 150–250 μg of crude nuclear protein or bovine serum albumin, 25 mM Hepes pH 7.6, 45 mM KCl, 1 mM EDTA and 1 mM DTT (Sierra, 1990), and incubated for 5 min on ice. The digestion was stopped by adding 120 μl of stop solution (200 mM NaCl, 30 mM EDTA, 1% SDS) and the products were phenol-extracted and precipitated with ethanol. Maxam-Gilbert sequencing (G+A) reactions of the labeled promoter fragments were performed according to Sambrook *et al.* (1989). Samples were analyzed by electrophoresis on an 8% polyacrylamide sequencing gel. The dried gels were examined by autoradiography.

Electrophoretic mobility shift assay (EMSA)

Double-stranded oligonucleotides comprising the G-box region (G-wt) and those comprising a mutated G-box (G-mut) (Figure 7a) were used in EMSA as radiolabeled probes or as unlabeled competitors. The oligonucleotides were end-labeled with [$\alpha-^{32}$P]dCTP by filling-in 3′ recessed ends with Klenow. Crude nuclear proteins (10 μg) were incubated with or without a ten-fold molar excess of unlabeled competitor oligonucleotide (250 fmol) in 15 μl reaction mixtures containing 3 μg of poly(dI-dC), 25 mM Hepes pH 7.6, 45 mM KCl, 1 mM EDTA, 1 mM DTT and 10% glycerol. After 10 min on ice, 25 fmol of labeled oligonucleotide was added and further incubated for 15 min at room temperature. Reaction products were analyzed on a 6% native polyacrylamide gel. The dried gels were examined by autoradiography.

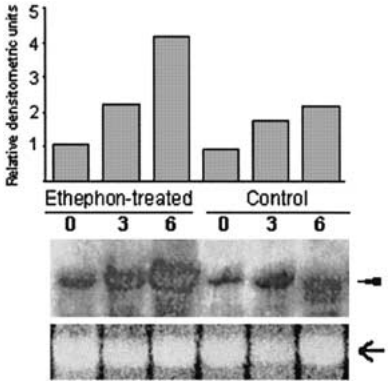

Figure 1. Northern blot analysis on ethylene induction of *SmCP* expression with a [^{32}P]dCTP-labeled cDNA probe. Densitometry analysis, depicted graphically, is shown above the blot. The number above each lane indicates time (h) after ethephon treatment. Zero hour corresponds to 06:00. Arrowhead indicates the 1.5 kb *SmCP* mRNA. Arrow denotes 25S rRNA band in the ethidium bromide-stained gel, photographed before blotting (bottom panel).

Results

SmCP *expression is induced by ethephon treatment*

Since *SmCP* expression is induced during leaf senescence and fruit ripening (Xu and Chye, 1999), and both these processes are brought about by ethylene; it was pertinent to examine the effect of ethephon treatment on *SmCP* expression. Treatment with ethephon led to an increase in *SmCP* mRNA in leaves after 3 h and, more significantly, 6 h as compared to the control (Figure 1). In the untreated control, we noticed an increase in *SmCP* mRNA over this 6 h period (Figure 1), albeit considerably less than the ethephon-treated plants, suggesting that *SmCP* may be under circadian regulation.

Circadian control of SmCP *expression*

To investigate circadian regulation of *SmCP*, northern blot analysis was carried out with RNA extracted, every 3 h for 24 h, from 5-week old *S. melongena* seedlings germinated in 12 h light/12 h dark cycles at 24 °C. Figure 2a confirms that *SmCP* expression is under circadian control with lowest expression in early light (ca. 06:00–09:00) and highest expression in late light (ca. 15:00–18:00). When these seedlings were transferred to continuous light, circadian regulation of *SmCP* was maintained during the first 24 h, after which a shift in the peak was observed some 3–6 h later (between 21:00–24:00) the following day (Figure 2b). Alternatively, when seedlings were

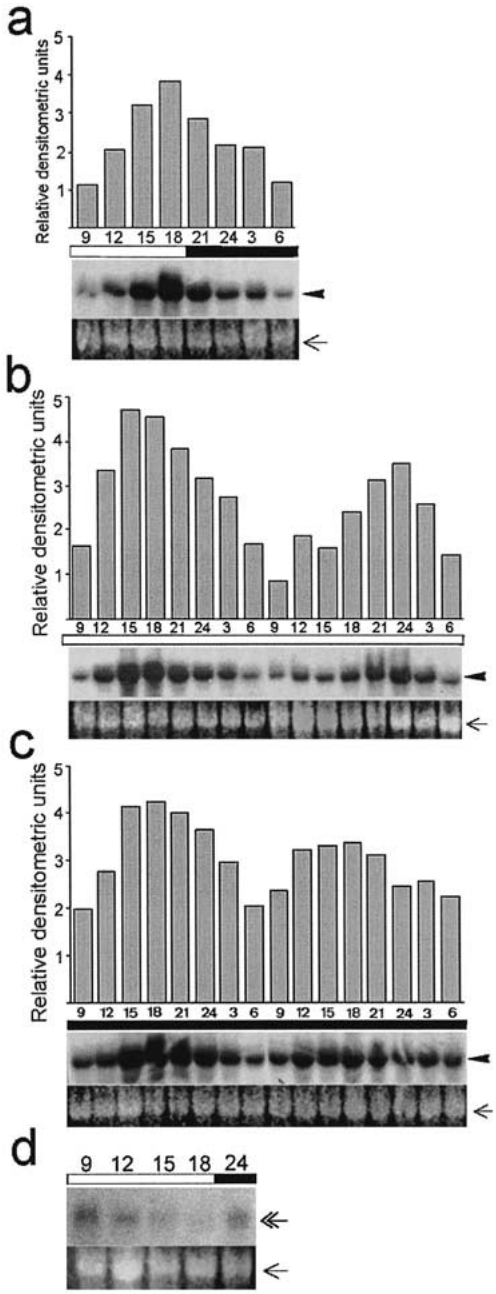

Figure 2. Northern blot analysis on circadian regulation of *SmCP* and *rbcS* in *S. melongena*. A [^{32}P]dCTP-labeled *SmCP* probe (a–c) or a [^{32}P]dCTP-labeled *Hevea rbcS* probe (d) was hybridized to RNA from aerial parts of 5-week old *S. melongena* seedlings germinated in 12 h light [08:00–20:00]/12 h dark [20:00–08:00] and subsequently grown in one of the following conditions: (a) 12 h light [08:00–20:00]/12 h dark [20:00–08:00]; (b) continuous light from 08:00; (c) continuous darkness from 08:00; (d) 12 h light [08:00–20:00]/12 h dark [20:00–08:00]. Densitometric analysis is shown above the blot. Numbering above each lane indicates time of day (24 h clock). White and black bars indicate light and dark periods, respectively. Arrowhead indicates the 1.5 kb *SmCP* mRNA. Arrow denotes 25S rRNA band (bottom panel). The *rbcS* mRNA is denoted with a double arrow in (d).

shifted to continuous darkness, the circadian regulation dampened-off (Figure 2c). Our results suggest that LD cycling is essential in maintaining circadian control of *SmCP*. We also demonstrated, using a heterologous *Hevea rbcS* (Chye *et al.*, 1991) probe, that in contrast to *SmCP*, circadian expression of *S. melongena rbcS* is maximum in early light (Figure 2d).

Circadian control of SmCP *mRNA is entrained by LD cycles*

To establish whether *SmCP* mRNA experiences circadian control under continuous light, we harvested leaves for northern blot analysis, every 3 h from 5-week old *S. melongena* seedlings germinated in continuous light at constant temperature (24 °C) and found that *SmCP* expression was uniform (Figure 3a). When these seedlings, maintained in continuous light, were subsequently exposed to temperature cycling (12 h 21 °C/12 h 24 °C) *per se*, *SmCP* expression still lacked circadian rhythm (Figure 3b). However, circadian control was initiated when these seedlings also experienced periods of 12 h light/12 h dark (Figure 3c), suggesting that LD cycling is essential not only for maintaining, but also for entraining the rhythm.

The molecular basis of SmCP *regulation: cloning the* SmCP *promoter*

The *SmCP* promoter was cloned to elucidate the molecular basis of *SmCP* regulation. To this end, Southern blot analysis using a probe prepared from the 5′-end *Eco*RI *SmCP* cDNA (nucleotides 1–993 of the cDNA) was carried out. This probe cross-hybridized to one *Eco*RI fragment, ca. 5 kb, while the full-length 1.5 kb cDNA cross-hybridized to two *Eco*RI fragments, ca. 5 kb and 10.5 kb. These results suggest that the 5 kb fragment corresponds to a more upstream fragment that likely contains the *SmCP* promoter.

Subsequently, a *S. melongena* genomic library enriched for *Eco*RI fragments ca. 5 kb was constructed and screened with the *SmCP* cDNA probe. Three putative positive clones obtained were identical, each with a 5 kb *Eco*RI insert carrying the *SmCP* promoter. This 5 kb fragment was cloned and partially sequenced (Figure 4a). Primer extension analysis was carried out to locate the site of transcription initiation of *SmCP* mRNA (Figure 5). This site was mapped to an 'A' (position + 1, Figure 5). Consistent with this being the transcription initiation site, CAAT and TATA boxes could be identified at nucleotide positions

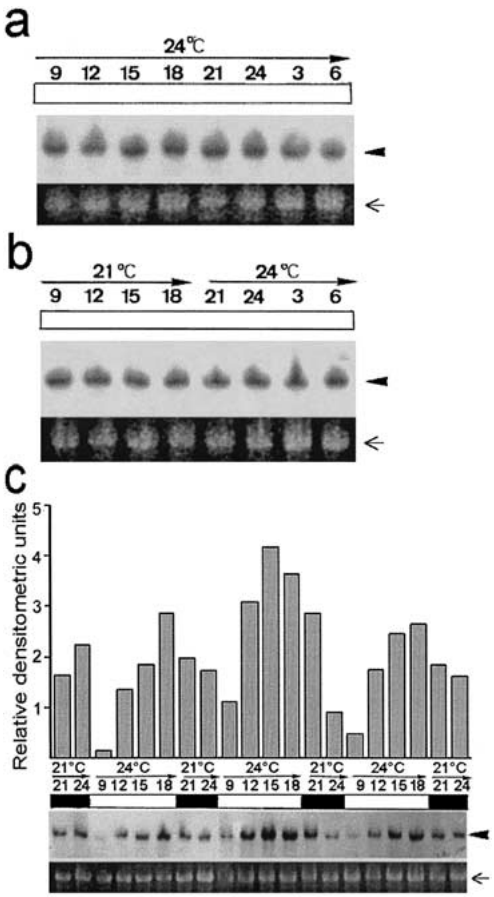

Figure 3. Northern blot analysis on the effect of LD cycling in initiating circadian regulation of *SmCP*. RNA was extracted at 3 h intervals from aerial parts of 5-week old *S. melongena* seedlings germinated in (a) continuous light and constant temperature (24 °C); (b) continuous light and a thermoperiod of 21 °C [08:00–20:00]/24 °C [20:00–08:00]; (c) continuous light and constant temperature (24 °C) and subsequently shifted (at 20:00) to 12 h (20:00–08:00) dark 21 °C/12 h light (08:00–20:00) 24 °C. The first sample was collected 1 h after the shift to 21 °C, i.e. at 21:00. Arrowhead indicates the 1.5 kb *SmCP* mRNA. Arrow denotes 25S rRNA band (bottom panel).

−91/−88 and −29/−26, respectively (Figure 4a). In comparison, the 5′ end of the longest cDNA obtained (Xu and Chye, 1999) was identified at position +29 (Figure 4a).

Since *SmCP* expression is ethylene-inducible, we attempted to locate putative ethylene-responsive elements (EREs) in its promoter. A putative 8 bp ERE (ATTTCAAA) was observed at positions −141/−134 (Figure 4b) and it is identical to the ERE of carnation *GST1* (Itzhaki *et al.*, 1994) and differs by one nucleotide from the sequence AATTCAAA required for ethylene-responsiveness in the tomato fruit-ripening

a

```
      TAGCACTTAATTTTTTCAAGATAAGGCAAGAATATTATTATTA -584
AAATAAAATGTGGGTCCCAGAAAATGGTCTCCATGAAAAAGATATAAATA -534
GATATTTCTATACCCTATCTTAACTTATTAGCTACCATGGAAGCAACCTC -484
TCTTTTGTTATTTATTTGCAAAAAATAGTCCTATCATCAACTTGGAGCAT -434
AAATAAATATAAAATAGTCCTCGATTCACATCAAATGAATTATTATATTT -384
TTTAATAAAAAATAAGTGAATTTGTTAAAATTCAAGAGCATTGTTCATTT -334
TTTTTATATTTATGTTTTTAATGAATTTCTTAATTACTCACACTTTCAAG -284
AAAAATAACTTGAAAAAAAAGATTAAAAGTGAGTCAAACAATATCATACT -234
TTTTAATTTACTTAACATGAACTATGGAGAGAAAAATGAATAGATTATAA -184
TAAAATAATAAAAATACATTTATTAAATGCATATTTAATCTGATTTCAAA -134
AAAAAATTGTGAAATATAGCTACCACACGTATCATGAGGCGACAATTACA -84
CGTGTCAGCATCACAACGGCTAACTCACATACCTAACCATGTTCTCGGAG -34
CCAATATATACTCGAGTAGCTAAAGAGGAGAGGAGAAAATATCCGCATTC +17
ATTTCTTCTTCATTCTTCTTTTTATTATTATTTCTCCCATGGATCGTCTT +67
TTTCTCTTATCTCTTCTCGCTTTCGCGCTTTTCTCGTCGGCGATTGCTTT +117
CTCCGACGACGATCCGTTGATCCGGCAAGTCGTATCGGAAACCGATGACA +167
ACCATATGTTAAACGCCGAGCATCATTTCTCACTTTTCAAGTCAAAGTAT +217
GGAAAGATCTATGCTTCTCAGGAGGAACATGACCATAGATTGAAGGTGTT +267
CAAGGCTAATCTCCGCCGTGCAAGGCGTCACCAGCTCCTTGACCCTACCG +317
CTGAGCACGGTATTACGCAATTCTCCGATCTGACTCCGTCGGAGTTTCGC +367
```

b

```
SmCP  -166 ATTTATTAAATGCATATTTAATCTGATTTCAAAAAAA -130
GST1  -521 GATATATACAGTGAATTTACCACCTATTTCAAAGCCT -485
E4    -142 GTTTTTGTTTTTACTACCAACAAGAAATTCAAATGGC -106
```

Figure 4. Nucleotide sequence analysis of the *SmCP* promoter region. a. Nucleotide sequence of *SmCP* promoter showing the transcription start site as defined by primer extension analysis (+1; in bold italics). TATA and CAAT boxes are underlined. The ATG start codon is overlined. The putative ERE (ATTTCAAA) and the G-box (CACGTG) are in bold. The sequence (+181/+197) complementary to the primer used in primer-extension is underlined. b. Comparison of the putative ERE of *SmCP* with EREs of *GST1* (Itzhaki *et al.*, 1994) and of *E4* (Montgomery *et al.*, 1993). Regions defined by DNase I footprinting are underlined. DNA sequences are numbered relative to the transcription start. The ATTTCAAA motif is in bold. Vertical lines denote homology.

E4 promoter (Montgomery *et al.*, 1993). Analysis of the *SmCP* promoter also revealed the presence of a G box (CACGTG at −85/−80, Figure 4a), a ubiquitous regulatory element in plants (Menkens *et al.*, 1995).

G-box specifically binds nuclear proteins in fruits and leaves

To investigate if these putative *cis*-elements in the *SmCP* promoter bind nuclear proteins, we performed DNase I footprinting experiments on both strands of

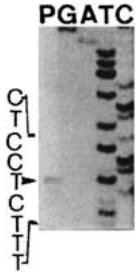

Figure 5. Primer extension analysis to map the 5′ end of the *SmCP* mRNA. A ^{32}P-labeled oligomer was hybridized to 30 μg of total fruit RNA from *S. melongena*. The primer was also used to generate a dideoxy sequencing ladder (GATC) which was electrophoresed next to the extended product (P), denoted by an arrowhead. The template used in this sequencing reaction was derived from the genomic clone containing the 5′ end of *SmCP*.

the *SmCP* promoter. Incubation of the ^{32}P-end-labeled −221/+17 *SmCP* promoter strands with nuclear extracts from mature fruits (50 DAP) revealed a protected area on the top strand, from −93 to −73, and on the bottom strand, −92 to −68 respectively (Figure 6a). Figure 6c shows that the G-box lies within this protected region. Figure 6b shows that protection of the G-box-containing sequence was stronger in senescent fruits (70 DAP) than young fruits (20 DAP).

We further tested the binding activity and specificity of nuclear proteins to this G-box by competitive EMSA using the wild-type G-box containing oligonucleotides (G-wt) and mutant oligomers G-mut (Figure 7a). Results on fruit nuclear extracts show that the binding activity of G-wt was stronger in senescent fruits (Figure 6b, lane 3) than young fruits (Figure 6b, lane 2). This observation is consistent with our results of DNase I footprinting analysis (Figure 7b). The complete abolition of this binding activity with G-mut (Figure 7b, lanes 8 and 9) demonstrates that the observed binding activities in fruit extracts are specific to the G-box. This specificity was further verified by competitive EMSA. Addition of a ten-fold molar excess of unlabeled G-wt completely eliminated the formation of the low-mobility protein-DNA complex (Figure 7b, lanes 4 and 5). In contrast, G-mut could not compete out labeled G-wt in binding fruit nuclear proteins (Figure 7b, lanes 6 and 7).

When we compared G-box binding of fruit extracts (Figure 7c, lanes 1 and 2) with leaf extracts (Figure 7c, lanes 4 and 5), we observed strongest binding in senescent fruits (Figure 7c, lane 2). With extracts from circadian-regulated leaves, stronger binding occurred at peak (Figure 7c, lane 5) rather than low (Figure 7c, lane 4) *SmCP* expression. G-box-binding specificity

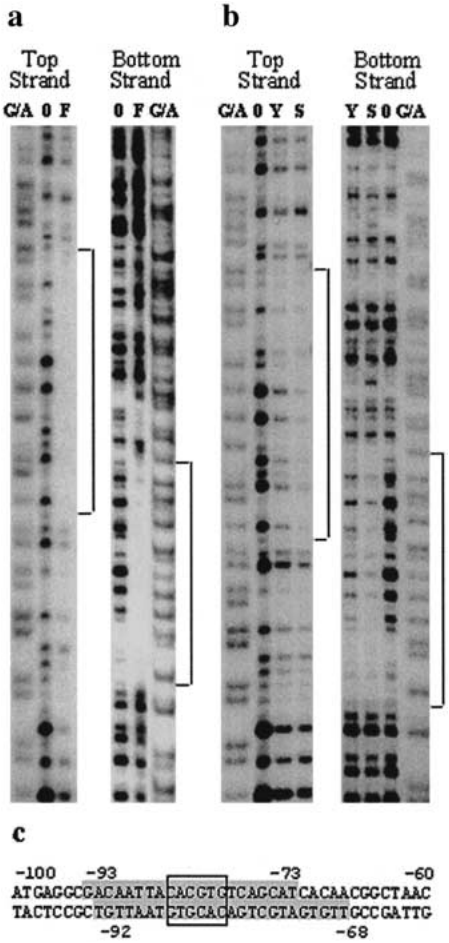

Figure 6. In vitro DNase I footprinting analysis of *SmCP* promoter with [$\alpha-^{32}$P]dCTP-labeled probes (−221/+17) containing the *SmCP* promoter region. a. DNase I protection using nuclear extract from mature fruits at 50 DAP. Lanes: G/A, Maxam-Gilbert sequencing (G+A) reactions of the labeled promoter fragment used in localization of the protected region. 0, DNase I digestion pattern without nuclear extracts; F, crude nuclear protein (250 μg) from mature fruits at 50 DAP. The brackets show the footprinted regions. b. Differential DNase I protection using nuclear extracts from young fruits (20 DAP) and senescent fruits (70 DAP). Lanes: G/A, Maxam-Gilbert sequencing (G+A) reactions of the labeled promoter fragment used for localization of the protected region. 0, DNase I digestion pattern without nuclear extracts; Y, crude nuclear protein (150 μg) from young fruits at 20 DAP; S, crude nuclear protein (150 μg) from senescent fruits at 70 DAP. The brackets show the regions footprinted. c. DNA sequence of protected regions detected in a and b. The *SmCP* promoter sequence from −100 to −60 is shown. The protected regions detected are highlighted in gray. The G-box sequence is boxed.

in circadian-regulated leaves was demonstrated when G-wt (Figure 7c, lanes 6 and 7), but not G-mut (Figure 7c, lanes 8 and 9), competed out binding. That G-mut does not bind nuclear extracts from circadian-regulated leaves is clearly demonstrated in Figure 7c (lanes 10 and 11).

Discussion

We have examined *SmCP* expression by northern blot analysis and observed that *SmCP* is ethylene-inducible and is under circadian regulation. Subsequently, investigations on the molecular basis of *SmCP* regulation were initiated by cloning the *SmCP* promoter for DNase I footprinting and EMSA. Our results revealed the presence of a G-box that specifically interacts with nuclear proteins in fruits and leaves and that strong binding activity at the G-box coincided with high *SmCP* expression.

Ethylene induction of SmCP *expression*

Ethylene induction of *SmCP* expression in northern blot analysis reported here is consistent with our previous observations in northern analysis of significant increases in *SmCP* expression in ripened yellow *S. melongena* fruits and in senescent leaves (Xu and Chye, 1999). Our results strongly suggest that ethylene, which is at particularly high levels in fruit at the onset of fruit ripening (Gillaspy *et al.*, 1993) and which regulates leaf senescence (Grbic and Bleecker, 1995), triggers *SmCP* expression at the level of transcription. Other ethylene-inducible genes encoding cysteine proteinases associated with senescence and ripening are *Arabidopsis SAG2* and *SAG12* (Grbic and Bleecker, 1995), *Citvac* encoding *Citrus* vacuolar processing thiol proteinase (Alonso and Granell, 1995) and *pD-CCP1* encoding carnation cysteine proteinase (Jones *et al.*, 1995) but their corresponding promoters have not been characterized.

Ethylene elicits a signal transduction pathway leading to the activation of an ethylene-responsive binding protein family of transcriptional activators which induces expression of ethylene-responsive genes (Solano and Ecker, 1998). EREs can be divided into two groups, the AT-rich element and the GCC box. The GCC box, also known as AGCCGCC sequence (Sato *et al.*, 1996) has been located in the promoters of a large number of ethylene-inducible pathogenesis-related (PR) genes. Despite strong conservation to other known EREs, the putative ERE

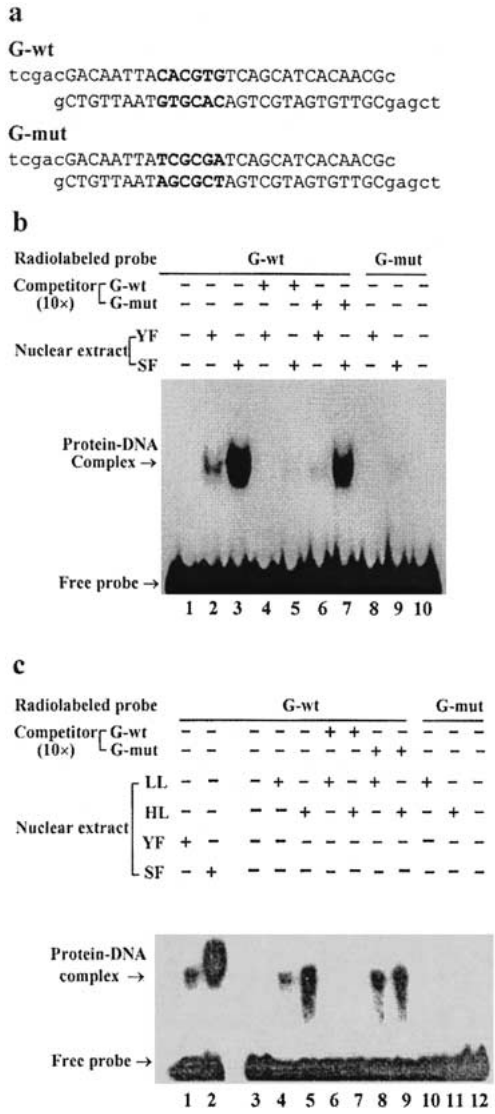

Figure 7. EMSA of the G-box containing footprinted regions of the *SmCP* promoter. a. Nucleotide sequences of double-stranded oligonucleotides used for EMSA, G-box containing G-wt and its mutant, G-mut. The upper-case letters correspond to the footprinted regions of the *SmCP* promoter and the lowercase letters indicate additional sequences designed for end-labeling (Janne and Hammond, 1998). The altered nucleotides in G-mut (Loake *et al.*, 1992) and their corresponding sequences in G-wt are highlighted in bold type. b. Interaction of nuclear proteins of *S. melongena* fruits with the G-box-containing probes. A 10 μg portion of crude nuclear protein from young fruits (YF) at 20 DAP (lanes 2, 4, 6, 8) and senescent fruits (SF) at 70 DAP (lanes 3, 5, 7, 9) was incubated with end-labeled G-wt (lanes 1–7) or G-mut (lanes 8–10), in the absence (lanes 2, 3, 8 and 9) or presence of a ten-fold molar excess of unlabeled competitor, G-wt (lanes 4 and 5) or G-mut (lanes 6 and 7). Lanes 1 and 10 are free probes lacking crude nuclear proteins. c. Interaction of nuclear proteins from *S. melongena* fruits and circadian-regulated leaves with the G-box-containing probes. A 10 μg portion of crude nuclear proteins from young fruits (YF) at 20 DAP (lane 1), senescent fruits (SF) at 70 DAP (lane 2), circadian-regulated leaves at low (lane 4) or peak (lane 5) *SmCP* expression was incubated with end-labeled G-wt (lanes 1–9) and G-mut (lanes 10–12) in the absence (lanes 1, 2, 4, 5, 10, and 11) or presence of a ten-fold molar excess of unlabeled competitor, G-wt (lanes 6 and 7) or G-mut (lanes 8 and 9). Lanes 3 and 12 are free probes lacking crude nuclear proteins.

(ATTTCAAA) at −141/−134 of the *SmCP* promoter did not display protection on DNase I footprinting analysis. The ERE (ATTTCAAA) of *GST1*, a gene associated with petal senescence (Itzhaki *et al.*, 1994), was identified by footprinting and is both necessary and sufficient in conferring ethylene induction to a heterologous promoter (Itzhaki *et al.*, 1994). The 5′-flanking sequences reported necessary for ethylene-induction during tomato fruit ripening in the *E4* promoter (Montgomery *et al.*, 1993) contain the sequence AATTCAAA, which differs in only one nucleotide from the EREs of *SmCP* and *GST1* (Itzhaki *et al.*, 1994). Footprinting of the *E4* promoter has shown that this sequence is protected from DNase I digestion.

Significance in circadian regulation of SmCP

In plants, circadian regulation temporally compartmentalizes many key metabolic processes (Harmer *et al.*, 2000). Genes encoding chlorophyll *a/b*-binding protein (Kloppstech, 1985; Nagy *et al.*, 1988a; Riesselmann and Piechulla, 1990; Kay and Millar, 1993), the small subunit of the ribulose-bisphosphate carboxylase (Kloppstech, 1985; Pilgrim and McClung, 1993) and phosphoenolpyruvate carboxylase (Thomas *et al.*, 1990) show circadian regulation with peak expression in early light and light induction is phytochrome-mediated. In contrast, *SmCP* resembles some non-photosynthetic genes that show peak circadian expression in late light, *e.g.* tobacco cysteine proteinase *CYP-8* (Linthorst *et al.*, 1993), mustard glycine-rich RNA-binding proteins (Heintzen *et al.*, 1994), maize catalase *Cat3* (Redinbaugh *et al.*, 1990), *Arabidopsis CAT3* (Zhong and McClung, 1996) and *Arabidopsis* cold-circadian rhythm-RNA binding protein genes *CCR1* and *CCR2* (Carpenter *et al.*, 1994).

We confirm that *S. melongena rbcS* expression peaks during early light, like *Lhc* genes in other plants and suggest that photosynthesis and protein degradation, anabolic and catabolic events, respectively, are temporally separated by circadian regulation in opposite phases to efficiently maximize their functions. Interestingly, the expression of three *Arabidop-*

sis senescence-associated genes, of which one encodes cysteine proteinase SAG2, increases when the expression of photosynthesis-associated *CAB* and *rbcS* decline during leaf development (Grbic and Bleecker, 1995). This implicates yet another mode of control, over a longer time period, to keep these two metabolic processes apart.

Light/dark cycling is the key stimulus for circadian regulation of SmCP

SmCP mRNA does not undergo oscillations in the absence of initial entraining by LD cycles. Expression of *S. melongena* cysteine proteinase in 5-week old seedlings germinated in continuous light showed constitutive expression until they experienced LD cycles. The overriding importance of LD cycles over thermocycles in eliciting the circadian rhythm was further confirmed when *SmCP* mRNA did not display circadian regulation in seedlings germinated in continuous light with rhythmic thermocycles. Like *SmCP*, maize *Cat3* is expressed at constant levels in seedlings grown in continuous light and circadian rhythm was induced on exposure to LD cycles (Boldt *et al.*, 1995). Two putative RNA-binding protein transcripts from *Sinapis alba* L., *Sagrp1* and *Sagrp2*, which show similar circadian regulation to *SmCP*, are also expressed constitutively in plants germinated in continuous light (Heintzen *et al.*, 1994). However unlike *SmCP*, thermocycling (18 °C/22 °C) *per se* was sufficient for their entrainment in continuous light, suggesting that temperature can act as an alternative external stimulus to entrain the oscillator regulating *Sagrp* transcription (Heintzen *et al.*, 1994). In contrast, circadian regulation of wheat *Cab-1* (Nagy *et al.*, 1988a) and tomato *cab* (Riesselmann and Piechulla, 1990) is independent of LD cycling in initiating the endogenous rhythm because seedlings germinated in continuous light displayed circadian control. Further understanding of variations in circadian regulation would require dissection of these gene promoters, their interactive clock components and transcription factors.

G-box mediation of SmCP *transcription*

We are the first to report here the occurrence of a G-box in a plant proteinase gene promoter. The G-box has been reported to occur in promoters of plant genes that are affected by environmental cues, including light and stress, and it regulates gene expression in conjunction with another adjacent *cis*-element (Ferl and Paul, 2000). As in the case of other circadian-regulated genes, of which the biological clock serves to coordinate internal events with the external environment (Harmer *et al.*, 2000), the environmental cue for G-box mediation in *SmCP* regulation would be LD cycling. In the *SmCP* promoter, we observed a correlation between high *SmCP* expression and strong G-box binding in fruits and circadian-regulated leaves. Interestingly, a G-box motif has been observed at the promoter of *Arabidopsis Atgrp7*, which encodes a clock-regulated glycine-rich RNA-binding protein that shows similar circadian regulation to *SmCP*, but its significance has not been established by DNase I footprinting analysis or EMSA (Staiger and Apel, 1999). The G-boxes in light-regulated *RBCS-1A*, *CCA1*, *LHY* and *SPA1* have been shown to interact with PIF3, a transcription factor that reversibly binds phytochrome B upon light exposure (Martinez-Garcia *et al.*, 2000). Detection of G-box binding activity at the *SmCP* promoter suggests that G-boxes are not restricted to light-regulated photosynthesis-associated promoters, and they, too, play a role in regulating non-photosynthetic clock genes affected by LD cycling.

In this study, although we localized a putative ERE about 50 bp upstream from the G-box, which could potentially accompany the G-box in regulating *SmCP*, we could not establish its function by DNase I footprinting analysis. Hence, further analysis by DNase I footprinting of sequences 5′ to those analyzed in this study and investigations by using reporter gene fusion constructs in transgenic plants would be conducted to locate other *cis*-elements including evening elements. These have recently been identified (Harmer *et al.*, 2000) by computational analysis of 31 promoters that show peak circadian expression in late light, like *SmCP*.

Acknowledgements

A Postgraduate Studentship from the University of Hong Kong to F.X.X. is gratefully acknowledged. This project was supported by grants from the University of Hong Kong CRCG grant (335/021/0011) and from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project HKU 7181/01M).

References

Alonso, J.M. and Granell, A. 1995. A putative vacuolar processing protease is regulated by ethylene and also during fruit ripening in *Citrus* fruit. Plant Physiol. 109: 541–547.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. 1995. Current Protocols in Molecular Biology. Greene Publishing Associates/Wiley Interscience, New York.

Boldt, R. and Scandalios, J.G. 1995. Circadian regulation of the *Cat3* catalase gene in maize (*Zea mays* L.): entrainment of the circadian rhythm of *Cat3* by different light treatments. Plant J. 7: 989–999.

Bradford, M.M. 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72: 248–254.

Buchanan-Wollaston, V. and Ainsworth, C. 1997. Leaf senescence in *Brassica napus*: cloning of senescence related genes by subtractive hybridization. Plant Mol. Biol. 33: 821–834.

Buckner, B., Janick-Buckner, D., Gray, J. and Johal, G.S. 1998. Cell-death mechanism in maize. Trends Plant Sci. 3: 218–223.

Callis, J. 1995. Regulation of protein degradation. Plant Cell 7: 845–857.

Carpenter, C.D., Kreps, J.A. and Simon, A.E. 1994. Genes encoding glycine-rich *Arabidopsis thaliana* proteins with RNA-binding motifs are influenced by cold treatment and an endogenous circadian rhythm. Plant Physiol. 104: 1015–1025.

Chye, M.-L., Tan, S., Tan, C.-T., Kush, A. and Chua., N.-H. 1991. Nucleotide sequence of a cDNA clone encoding the precursor of ribulose-1,5-bisphosphate carboxylase small subunit from *Hevea brasiliensis* (rubber tree). Plant Mol. Biol. 16: 1077–1078.

Dellaporta, S.L., Wood, J. and Hicks, J.B. 1983. A plant DNA minipreparation: version II. Plant Mol. Biol. Rep. 1: 19–21.

Drake, R., John, I., Farrell, A., Cooper, W., Schuch, W. and Grierson, D. 1996. Isolation and analysis of cDNAs encoding tomato cysteine proteases expressed during leaf senescence. Plant Mol. Biol. 30: 755–767.

Ferl, R. and Paul, A.-L. 2000. Genome organization and expression. In: B.B. Buchanan, W. Gruissem and R.L. Jones (Eds) Biochemistry and Molecular Biology of Plants, American Society of Plant Physiologists, Rockville, MD, pp. 312–357.

Gillaspy, G., Ben-David, H. and Gruissem, W. 1993. Fruits: a developmental perspective. Plant Cell 5: 1439–1451.

Grbic, V. and Bleecker, A.B. 1995. Ethylene regulates the timing of leaf senescence in *Arabidopsis*. Plant J. 8: 595–602.

Greenberg, J.T. and Ausubel, F.M. 1993. *Arabidopsis* mutants compromised for the control of cellular damage during pathogenesis and aging. Plant J. 4: 327–341.

Guerrero, C., de la Calle, M., Reid, M.S. and Valpuesta, V. 1998. Analysis of the expression of two thioprotease genes from daylily (*Hemerocallis* spp.) during flower senescence. Plant Mol. Biol. 36: 565–571.

Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.-S., Han, B., Zhu, T., Wang, X., Kreps, J.A. and Kay, S.A. 2000. Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. Science 290: 2110–2113.

Heintzen, C., Melzer, S., Fischer, R., Kapeler, S., Apel, K. and Staiger, D. 1994. A light- and temperature-entrained circadian clock controls expression of transcripts encoding nuclear proteins with homology to RNA-binding proteins in meristematic tissue. Plant J. 5: 799–813.

Hensel, L.L., Grbic, V., Baumgarten, D.A. and Bleecker, A.B. 1993. Developmental and age-related processes that influence the longevity and senescence of photosynthetic tissues in *Arabidopsis*. Plant Cell 5: 553–564.

Itzhaki, H., Maxson, J.M. and Woodson, W.R. 1994. An ethylene-responsive enhancer element is involved in the senescence-related expression of the carnation glutathione *S*-transferase (*GST1*) gene. Proc. Natl. Acad. Sci. USA 91: 8925–8929.

Janne, M. and Hammond, G.L. 1998. Hepatocyte nuclear factor-4 controls transcription from a TATA-less human sex hormone-binding globulin gene promoter. J. Biol. Chem. 273: 34105–34114.

Jones, M.L., Larsen, P.B. and Woodson, W.R. 1995. Ethylene-regulated expression of a carnation cysteine proteinase during flower petal senescence. Plant Mol. Biol. 28: 505–512.

Kay, S.A. and Millar, A.J. 1993. Circadian-regulated *cab* gene transcription in higher plants. In: M.W. Young (Ed.) Molecular Genetics of Biological Rhythms, Marcel Dekker, New York, pp. 73–89.

Kloppstech, K. 1985. Diurnal and circadian rhythmicity in the expression of light-induced plant nuclear messenger RNAs. Planta 165: 502–506.

Linthorst, H.J.M., van der Does, C., Brederode, F.T.M.W. and Bol, J.F. 1993. Circadian expression and induction by wounding of tobacco genes for cysteine proteinase. Plant Mol. Biol. 21: 685–694.

Loake, G.J., Faktor, O., Lamb, C.J. and Dixon, R.A. 1992. Combination of H-box [CCTACC(N)7CT] and G-box (CACGTG) *cis* elements is necessary for feed-forward stimulation of a chalcone synthase promoter by the phenylpropanoid-pathway intermediate *p*-coumaric acid. Proc. Natl. Acad. Sci. USA 89: 9230–9234.

Lohman, K., Gan, S., John, M. and Amasino, R.M. 1994. Molecular analysis of natural leaf senescence in *Arabidopsis thaliana*. Physiol. Plant. 92: 322–328.

Martinez-Garcia. J.F., Huq, E. and Quail, P.H. 2000. Direct targeting of light signals to a promoter element-bound transcription factor. Science 288: 859–863.

Martino-Catt, S.J. and Kay, S.A. 1994. Optimization of DNase I footprinting experiments. In: S.B. Gelvin and R.A. Schilperoort (Eds) Plant Molecular Biology Manual, 2nd ed., Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. I2: 1–13.

Menkens, A.E., Schindler, U. and Cashmore, A.R. 1995. The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. Trends Biochem. Sci. 20: 506–510.

Montgomery, J., Goldman, S., Deikman, J., Margossian, L. and Fischer, R.L. 1993. Identification of an ethylene-responsive region in the promoter of a fruit ripening gene. Proc. Natl. Acad. Sci. USA 90: 5939–5943.

Nagy, F., Kay, S.A. and Chua, N.-H. 1988a. A circadian clock regulates transcription of the wheat *Cab-1* gene. Genes Dev. 2: 376–382.

Nagy, F., Kay, S.A. and Chua, N.-H. 1988b. Analysis of gene expression in transgenic plants. In: S.B. Gelvin and R.A. Schilperoort (Eds) Plant Molecular Biology Manual, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. B4: 1–29.

Pennell, R.I. and Lamb, C. 1997. Programmed cell death in plants. Plant Cell 9: 1157–1168.

Pilgrim, M.L. and McClung, C.R. 1993. Differential involvement of the circadian clock in the expression of genes required for ribulose-1,5-bisphosphate carboxylase /oxygenase synthesis, assembly, and activation in *Arabidopsis thaliana*. Plant Physiol. 103: 553–564.

Praekelt, U.M., McKee, R.A. and Smith, H. 1988. Molecular analysis of actinidin, the cysteine proteinase of *Actinidia chinensis*. Plant Mol. Biol. 10: 193–202.

Redinbaugh, M.G., Sabre, M. and Scandalios, J.G. 1990. Expression of the maize *Cat3* catalase gene is under the influence of a circadian rhythm. Proc. Natl. Acad. Sci. USA 87: 6853–6857.

Riesselmann, S. and Piechulla, B. 1990. Effect of dark phases and temperature on the chlorophyll *a/b* binding protein mRNA level oscillations in tomato seedlings. Plant Mol. Biol. 14: 605–616.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Plainview, NY.

Sato, F., Kitajima, S., Koyama, T. and Yamada, Y. 1996. Ethylene-induced gene expression of osmotin-like protein, a neutral isoform of tobacco PR-5, is mediated by the AGCCGCC *cis*-sequence. Plant Cell Physiol. 37: 249–255.

Sierra, F. 1990. BioMethods: A Laboratory Guide to *In Vitro* Transcription. Birkhäuser Verlag, Berlin.

Solano, R. and Ecker, J.R. 1998. Ethylene gas: perception, signalling and response. Curr. Opin. Plant Biol. 1: 393–398.

Staiger, D. and Apel, K. 1999. Circadian clock-regulated expression of an RNA-binding protein in *Arabidopsi*s: characterisation of a minimal promoter element. Mol. Gen Genet. 261: 811–819.

Thomas, M., Cretin, C., Vidal, J., Keryer, E., Gadal, P. and Monsinger, E. 1990. Light regulation of phosphoenol/pyruvate carboxylase mRNA in leaves of C4 plants: evidence for phytochrome control on transcription during greening and for rhythmicity. Plant Sci. 69: 65–78.

Xu, F.-X. and Chye, M.-L. 1999. Expression of cysteine proteinase during developmental events associated with programmed cell death in brinjal. Plant J. 17: 321–327.

Ye, Z.H. and Varner, J.E. 1996. Induction of cysteine and serine proteases during xylogenesis in *Zinnia elegans*. Plant Mol. Biol. 30: 1233–1246.

Zhong, H.H. and McClung, C.R. 1996. The circadian clock gates expression of two *Arabidopsis* catalase genes to distinct and opposite circadian phases. Mol. Gen. Genet. 251: 196–203.