Biological Rhythm Research, 1999, Vol. 30, No. 3, pp. 264–271 0929-1016/99/3003-0264$15.00

Transcriptional Regulation of Oscillating Steady-State *Lhc* mRNA Levels: Characterization of two *Lhca* Promoter Fragments in Transgenic Tobacco Plants

Jan-Wolfhard Kellmann, Raimund Hoffrogge and Birgit Piechulla

University of Rostock, Germany

ABSTRACT

Lhc genes from angiosperms display circadian steady-state mRNA oscillations as a result of differential transcriptional activation. We studied promoter regions of tomato *Lhca* genes encoding chlorophyll *a/b* binding proteins, which are associated with photosystem I. Promoter deletion analysis showed that a 278 basepair (bp) DNA fragment from *Lhca3 (cab 8)* and a 119 bp fragment from *Lhca4*1 (cab 11)* was sufficient to mediate circadian gene expression in transgenic plants. Further experiments using the GUS-reporter gene assay showed that an internal 110 bp DNA fragment cloned from the 278 bp region of *Lhca3* (position -278 to -169), however, was not able to drive a circadian reporter gene expression, indicating that additional *cis*-elements located downstream or within the exons and introns of *Lhca3* are necessary.

In contrast, the 119 bp fragment of *Lhca4*1* bears all elements sufficient for a circadian mRNA accumulation pattern. The 119 bp promoter sequence of *Lhca4*1* contains imperfect repeats of the motif CAA(N_{2-4})ATC, which is present in 81% of all published *Lhc* promoter regions. This sequence overlaps with a recently determined binding motif ($AA^{A}/_{C}AATCT$) of the circadianly regulated transcription factor CCA1 isolated from *Arabidopsis*, suggesting that similar transcription factors may be involved in the circadian *Lhc* gene expression of angiosperms.

KEYWORDS: Circadian rhythm, promoter deletion analysis, transgenic tobacco plants, GUS reporter gene, tomato, *Lycopersicon esculentum*.

INTRODUCTION

Studies of *Lhc* genes and *Lhc* gene expression in angiosperms in the past decade lead to the following observations: (i) higher plants accomplish a large gene family encoding light-harvesting proteins for photosystem I (LHCa) and photosy-

Address correspondence to: Birgit Piechulla, University of Rostock, Department of Molecular Physiology of Plants and Microorganisms, Gertrudenstrasse 11a, D-18051 Rostock, Germany. Fax: +49-381-494-2243. E-mail: bpiechu@imppm.bio4.uni-rostock.de

stem II (LHCb) proteins (Jansson and Gustafsson, 1991; Jansson, 1994; McGrath et al., 1992) and (ii) steady-state mRNA levels and transcriptional activation of the *Lhc* genes are under the control of a circadian clock in angiosperms (reviewed in: Piechulla, 1999).

Recent findings regarding the cellular origin of circadian rhythms and their entrainment revealed that autoregulation of gene expression, combined with protein-protein interactions (PER/TIM, Gekakis et al., 1995) may maintain or actually create circadian cycles (for review: Kay and Millar, 1995; Kay, 1997). Feedback cycles leading to autoregulated-circadian gene expression have been shown, for example, for the glycine rich protein *At*GRP7 from Arabidopsis, which was found to bind RNA and could therefore regulate its mRNA steady-state level post-transcriptionally by a negative autoregulatory circuit (Heintzen et al., 1997). Furthermore, at the level of transcription many clock controlled genes as well as transcription factors have been found. The Arabidopsis CCA-1 Myb related transcription factor is transiently induced by phytochrome and its steady-state mRNA levels oscillate with a circadian rhythm (Wang and Tobin, 1998). Phosphorylation and interaction of the CCA1 protein with a regulatory subunit of the CK2 protein kinase stimulates binding of CCA1 to the respective binding site of the *Lhcb1*3* promoter (Sugano et al., 1998). The circadian appearance of the CCA1 protein, successive phosphorylation and binding to the corresponding DNA motif leads to the circadian mRNA level oscillation of the target, the *Lhcb1*3* gene.

Putative binding sites for transcription factors (*cis*-elements) can be elucidated by deletion analysis of the 5' untranscribed regions of circadianly expressed genes. Promoter deletion analysis of two tomato *Lhca* and two tomato *Lhcb* genes was performed in transgenic tobacco plants. Regulatory *cis*-elements were found to be present at different positions relative to the transcription start sites of the respective *Lhc* gene (Piechulla et al., 1998). The shortest DNA fragments sufficient for circadian expression of a tomato *Lhc* transgene in tobacco were constricted for *Lhca3* (-278 to -231) and *Lhca4*1* (-119 to +3).

Promoter analysis of *Lhca3* and *Lhca4*1* was carried out by fusing DNA promoter fragments to the GUS reporter gene and subsequent determination of the steady-state levels of GUS mRNA in transgenic plants to further characterize *cis*-elements involved in circadian regulation within tomato *Lhc* promoters.

RESULTS AND DISCUSSION

Expression studies of a reporter gene in transgenic plants are a practicable *in vivo* method to corroborate that a defined regulatory DNA segment would be able to support circadian mRNA fluctuations. When the regulatory DNA fragment to be characterized is placed upstream of the coding region of a reporter gene, it can be

assumed that it can act as an transcriptional enhancer or silencer. As a result, different transcription rates lead to steady-state mRNA fluctuations of the reporter gene.

In our experiments we linked short DNA fragments derived from the tomato *Lhca3* and *Lhca4*1* gene, shown to be sufficient for circadian expression of the respective genes in transgenic plants (Piechulla et al., 1998), to the GUS reporter gene and subsequently integrated the chimeric genes into tobacco plants by *Agrobacterium* mediated gene transfer. The position of the DNA segments and the constructs within the binary plant transforming vectors are depicted in Fig. 1. Twenty-nine individual transgenic plant lines bearing the *Lhca3*-promoter-GUS construct and 64 plants bearing the *Lhca4*1*-promoter-GUS construct were selected and GUS activity was determined in leaf extracts by the fluorimetric assay described by Jefferson (1987). Although the *Lhca3* derived promoter fragment contained the sequence located between –278 and –231 of the 5' upstream region and additional 62 basepairs (bp) including a putative DNA motif involved in circadian control of tomato *Lhc* genes (Fig. 1A; Piechulla et al., 1998), none of the 15 transgenic tobacco plants tested for GUS expression showed GUS enzyme activity. Subsequent northern analysis of total RNA isolated from leaves harvested at different time points did not allow to detect GUS reporter gene transcripts. The lack of GUS expression in these experiments was unexpected because as a precaution a plant transforming vector (pETVgus; Martini et al., 1993) carrying the GUS gene and additional 55 bp of the Cauliflower mosaic virus 35S promoter was chosen. This 55 bp region contains the TATA motif and the transcription start site for the viral 35S-RNA and therefore should have been sufficient to bear the binding and activation elements for eukaryotic RNA polymerase II. PCR analysis revealed that the chimeric promoter-GUS construct was correctly incorporated into the genome of the transgenic tobacco plants demonstrating that the absent GUS enzyme activity was not a result of an artifact generated during T-DNA transfer (data not shown). The fact that (i) the –278 to –231 region is sufficient to drive circadian expression of the tomato *Lhca3* gene in transgenic tobacco plants (Piechulla et al., 1998) and (ii) this region taken out of the nucleotide sequence context of the *Lhca3* promoter and solely fused to a reporter gene is not able to regulate or activate gene expression indicates that additional DNA sequences located within the complete *Lhca3* gene are needed to reveal circadian mRNA accumulation.

Leaf extracts of 21 transgenic tobacco plants bearing the chimeric gene construct composed of the promoter segment –119 to +3 of the *Lhca4*1* gene from tomato and the GUS reporter gene (Fig. 1B) showed sufficient GUS enzyme activity. Two individual transgenic plant lines (d119/37 and d119/74) were chosen for analysis of rhythmic expression of the GUS reporter mRNA. The respective plants were self-pollinated and seeds were grown on medium containing 100

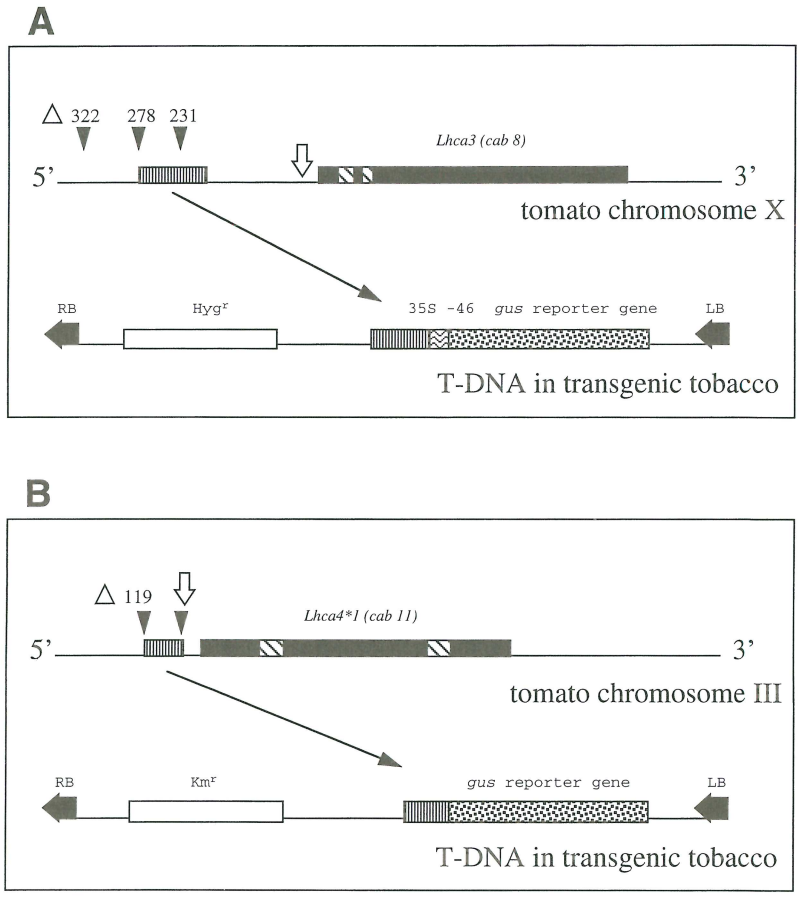

Fig. 1. Construction of *Lhca* promoter – *gus* chimeric genes
A: The *Lhca3* promoter fragment was pcr-amplified with 5' EcoRI and 3' XbaI primers and subsequently cloned into the binary plant transforming vector pETVgus (Martini et al., 1993). Misincorporation of nucleotides was ruled out after sequencing. *Agrobacterium tumefaciens* strain GV3101 (pMP90RK) was transformed and transgenic plants were generated by leaf disc inoculation of *Nicotiana tabacum* plants (Horsch et al., 1985). Transformed plants were grown in tissue culture on MS medium supplemented with 2% sucrose (Murashige and Skoog, 1962) and selected with 15 mg/l hygromycin.
B: *The Lhca4*1* promoter fragment was pcr-amplified with 5' HindIII and 3' BamHI primers and subsequently cloned into the binary plant transforming vector pBI101.2 (Jefferson et al., 1987). Plant transformation was performed as described in (A) using *A. tumefaciens* strain GV 3101 (pGV2260) and transformed plants were selected on 100 mg/l kanamycin.
Black arrow heads: position of *Lhca* promoter deletions; white arrows: transcription start site; black bars: exons; diagonal hatched bars: introns; wavy line hatched bar: Cauliflower mosaic virus 35S gene TATA box region (–46 to +8); Hyg^r: hygromycin phosphotransferase gene; Km^r: neomycin phosphotransferase gene; RB/LB: right and left borders of the Ti plasmid derived T-DNA from *Agrobacterium tumefaciens*; gus: *E. coli uidA* gene encoding β-glucuronidase. The *uidA* gene is 3' terminated by the polyA signals of the *rbcS-3C* gene from pea (A) or the nopaline synthase gene from *A. tumefaciens* (B).

mg/l kanamycin to select for outcrosses of the transgene. GUS enzyme activity was measured in 9 individual plants per line displaying a value between 20 – 400 pkat/mg protein. Subsequently, total RNA was isolated from transgenic leaf tis-

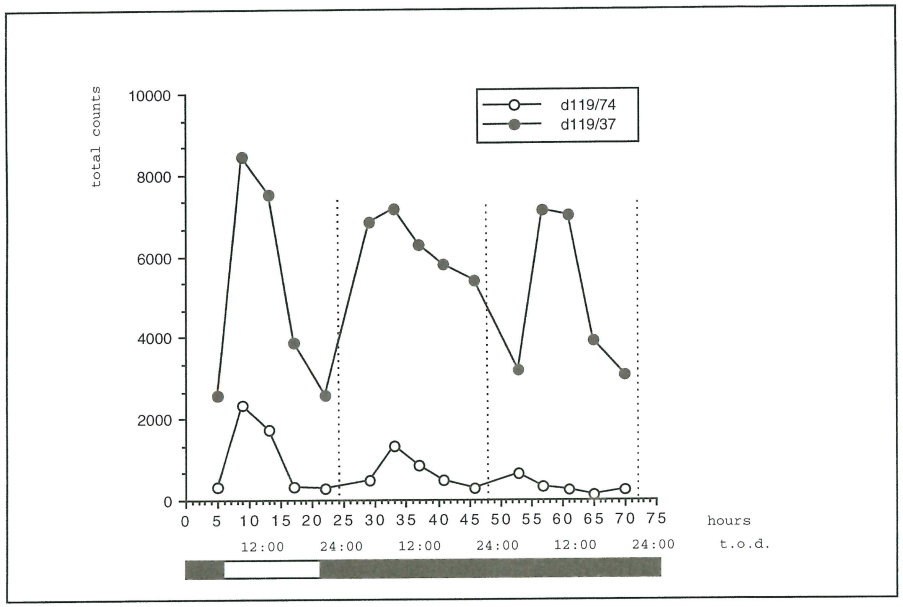

Fig. 2. GUS mRNA accumulation in individual transgenic tobacco plants d119/37 and d119/74 harboring the Lhca4*1-GUS construct.
Plants grown in a phytochamber supplemented with fluorescent white light (OSRAM Lumilux daylight L 12950; 75 µmol m^{-2} s^{-1}) were adapted to a 15 h light/9h dark cycle for several weeks before leaf samples from different plants were taken at the indicated time points (t.o.d.: time of day). Total RNA was isolated by grinding the leaf material with liquid nitrogen and subsequent extraction in a buffer containing 0,35 M sorbitol, 50 mM Tris/HCl, 25 mM EDTA, 15 mM 2-mercaptoethanol, 2 mM DTT, 1 mg/ml polyvinylpyrrolidone, 5 mM aurintricarboxylic acid (pH 8,0). After lysis, 1/10 vol. of a buffer containing 5% N-laurylsarcosine, 50 mM Tris/HCl, 25 mM EDTA (pH 8,0) was added. The suspension was 3× extracted with phenol/chloroform/isoamyl alcohol (24+24+1 vol.) and the supernatant was supplied with 1/10 vol. of 5 M ammonium acetate (pH 5,5). Nucleic acids were precipitated with 1 vol. of isopropanol. The nucleic acids were resolved in 10 mM Tris/HCl, 1 mM EDTA buffer (pH 7,5) and RNA was precipitated by adding LiCl to a final concentration of 2 M. Pellets were resuspended in H_2O and the RNA concentration was determined photometrically at 260 and 280 nm. Fifteen µg of total RNA of each sample were loaded on an denaturing agarose gel and blotted onto nylon membrane (Hybond, AMERSHAM). The blot was probed with random primed DNA fragments bearing the whole coding region of the GUS gene (specific activity: 1,6 * 10^8 cpm per µg of template DNA). The filter was hybridized at 62°C in a solution containing 0,9 M NaCl, 0,09 M Na-citrate (= 6× SSC), 0,1% of Ficoll 44, polyvinylpyrrolidone and bovine serum albumin, 100 µg / ml denatured salmon sperm DNA, 0,5% sodium dodecylsulfate (pH 7,0) over night and washed two times for 15 min. in 2× SSC at 62°C. GUS-mRNA was detected with an Instant Imager System (PACKARD) by direct counting of the radioactivity for four hours. T.o.d.: time of day.

sue at different time points during one day with a light period of 15 hours and two days under continuous darkness and subjected to northern analysis (one example is exemplified in Fig. 2). Both plants exhibited diurnal and circadian GUS – mRNA accumulation patterns with the same period length but different amplitudes, indicating that the tomato *Lhca4*1* gene promoter segment bears all elements needed to express a gene with a circadian mRNA accumulation pattern. Therefore, it was worthwhile to study this short, 123 bp long DNA fragment in more detail (Fig. 3). The transcription start site of *Lhca4*1* is located close to the 3' end of the investigated regulatory segment. Included is a 20 bp upstream located TATA box. Further upstream two imperfect repeats of the motif "CAA (N_{2-4}) ATC" are present, which are thought to play a role in circadian regulation of *Lhc* gene expression. This element was discovered by computer alignment analysis in 81% of all published *Lhc* promoter sequences (Piechulla et al., 1998). Wang et al. (1997) isolated a transcription factor CCA1 (Circadian Clock Associated 1) from *Arabidopsis* and showed by DNA binding and DNA footprint assays that this factor binds to a motif designated "AA$^{A}/_{C}$AATCT". This motif is also present in the *Lhca4*1* promoter fragment and overlaps with the "CAA (N_{2-4}) ATC" motif (Fig. 3). Recent studies by Wang and Tobin (1998) revealed that CCA1 expression is transiently induced by phytochrome action and oscillates itself with a circadian rhythm. Overexpression of CCA1 in transgenic *Arabidopsis* plants lead to delayed flowering and abolished the circadian expression of *Arabidopsis Lhcb*- and other circadianly expressed genes in continuous light and continuous darkness.

The presence of both the "AA$^{A}/_{C}$AATCT" and "CAA (N_{2-4}) ATC" motif in the *Lhca4*1* regulatory promoter fragment allows the suggestion that at least in toma-

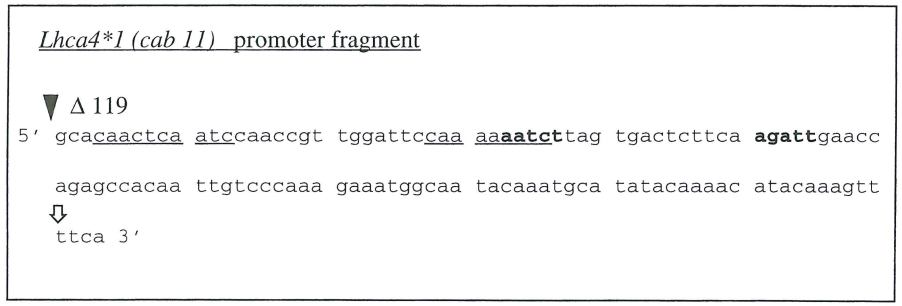

Fig. 3. Sequence of the –119 to +3 DNA fragment of the Lhca4*1 promoter.
The black arrow head indicates the position of the Δ 119 deletion within the *Lhca4*1* gene 5' upstream region. The white arrow marks the transcription start site. Two imperfect repeats of the „CAA (N_{2-4}) ATC" motif are underlined and two inverted repeats of the CCA1 core motif „AATCT" are depicted in bold letters.

to and in tobacco a CCA1 homologue exists which might be able to control *Lhc* genes of photosystem I as well. Our recent investigations using the yeast one-hybrid system (Luo et al., 1996) should answer the question whether the *Lhca4*1* promoter fragment is able to bind the *Arabidopsis* CCA1 transcription factor indicating that the CCA1 - regulated circadian expression of *Lhc* genes may be a common feature in circadian gene regulation in plants.

In summary, the presence of specific *cis*-elements ("CAA (N_{2-4}) ATC" and "$AA^{A}/_{C}AATCT$") in combination with promoter analysis strongly suggests that CCA1 is an important player in the circadian clock regulatory circuit of photosystem I and II *Lhc* genes in *Arabidopsis*, tobacco and tomato. Further investigations with other members of the concerted expressed *Lhc* gene family in tomato (Kellmann et al., 1993) as well as various plant species will inaugurate whether a common mechanism of transcriptional regulation underlies this phenomenon.

ACKNOWLEDGMENTS

We greatfully acknowledge the assistance of Hannelore Menzel during the RNA preparations. The authors are grateful to Norbert Martini for supplying the pETVgus binary vector and to Bernd Reiss for supplying the *Agrobacterium tumefaciens* strain GV 3101 (pMP90RK). This work was supported by grants of the Deutsche Forschungsgemeinschaft (DFG Pi 8-3 and 12-1) to B.P.

REFERENCES

Gekakis N, Saez L, Delahaye-Brown AM, Myers MP, Sehgal A, Young MW Weitz CJ (1995): Isolation of timeless by PER protein interaction: defective interaction between timeless protein and long-period mutant PER^{L}. *Science 270:* 811–815.

Heintzen C, Nater M, Apel K, Staiger D (1997): ATGRP7, a nuclear RNA-binding protein as a component of a circadian-regulated negative feedback loop in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA 94:* 8515–8520.

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG Fraley RT (1985): A simple and general method for transferring genes into plants. *Science 227:* 1229–1231.

Jansson S (1994): The light-harvesting chlorophyll *a/b* binding proteins. Biochim. *Biophys. Acta 1184:* 1–19.

Jansson S, Gustafsson P (1991): Evolutionary conservation of the chlorophyll *a/b*-binding proteins: cDNAs encoding Typ I, II and III LHC I polypeptides from the gymnosperm Scots pine. *Mol. Gen. Genet. 229:* 67–76.

Jefferson RA (1987): Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol. Biol. Reporter 5:* 387–405.

Jefferson RA, Kavanagh TA Bevan MW (1987): GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J. 6:* 3901–3907.

Kay SA, Millar AJ (1995): New models in vogue for circadian clocks. *Cell 83:* 361–364.

Kay SA (1997): PAS, present and future: Clues for the origins of circadian clocks. *Science 276:* 753–754.

Kellmann JW, Merforth N, Wiese M, Pichersky E, Piechulla B.(1993): Concerted circadian transcript level oscillations of nineteen *Lha/b (cab)* genes in *Lycopersicon esculentum* (tomato). *Mol. Gen. Genet. 237:* 439–448.

Luo Y, Vijaychander S, Stile J, Zhu L (1996): Cloning and analysis of DNA-binding proteins by yeast one-hybrid and two-hybrid systems. *Biotechniques 20:* 564–568.

Martini N, Egen M, Rüntz I, Strittmatter G (1993): Promoter sequences of a potato pathogenesis-related gene mediate transcriptional activation selectively upon fungal infection. *Mol. Gen. Genet. 236:* 179–186.

McGrath JM, Terzaghi WB, Sridhar P, Cashmore AR, Pichersky E (1992): Sequence of the fourth and fifth photosystem II type I chlorophyll *a/b*-binding protein genes of Arabidopsis thaliana and evidence for the presence of a full complement of the extended *cab* gene family. *Plant Mol. Biol. 19:* 725–733.

Murashige T, Skoog F (1962): A revised medium for rapid growth and bio assays with tobacco tissue culture. *Physiol. Plant. 15:* 473–497.

Piechulla B (1998): Circadian expression of the light harvesting complex protein genes in plants. Chronobiology International, in press.

Piechulla B, Merforth N, Rudolph B (1998): Identification of tomato *Lhc* promoter regions necessary for circadian expression. *Plant Mol. Biol. 38:* 655–662.

Sugano S, Andronis C, Green RM, Wang ZY, Tobin EM (1998): Protein kinase CK2 interacts with and phosphorylates the *Arabidopsis* circadian clock-associated 1 protein. *Proc. Natl. Acad. Sci. USA 95:* 11020–11025.

Wang ZY, Kenigsbuch D, Sun L, Harel E, Ong MS, Tobin EM (1997): A myb-related transcription factor is involved in the phytochrome regulation of an *Arabidopsis Lhcb* gene. *The Plant Cell 9:* 491–507.

Wang ZY, Tobin EM (1998): Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell 93:* 1207–1217.