Molecular Studies of Transcriptional Regulation of the *rolC* Gene in Higher Plants

Rikyu Matsuki

1994

(1)

Molecular Studies of Transcriptional Regulation of the *rolC* Gene in Higher Plants

Rikyu Matsuki

1994

Contents

Summary

Transgenic rice plants were obtained from protoplasts treated with two plasmids by electroporation. Primary transformants were selected on the basis of resistance to hygromycin, which was conferred by one of the co-transferred plasmids. Two plants out of 26 hygromycin resistant plants showed the reporter gene activity due to another plasmid possessing a chimeric gene consisting of a promoter region (about 900 bp upstream non-coding strand) of ORF12 gene (*rolC*) of Ri plasmid and the coding region for β-glucuronidase (GUS). GUS activity was strong in leaves and roots, but weak in callus. Colorimetric reaction using the GUS enzyme always showed the localization of the gene expression to vascular tissues only. Furthermore, transgenic rice plants possessing the *rolC* gene were obtained. Their progeny plants showed the reduction of plant height.

Gel shift analysis revealed that sequence specific binding factors interacting with the 5'-upstream region of the *rolC* gene exist in nuclear extracts from both leaves and calli. The binding site visualized by DNase I footprinting experiments possesses AT-1 box-like-sequence "ATATTTTTAT", located from -76 bp to -67 bp. In a further upstream region, an AT-rich region from -203 bp to -164 bp that was protected by DNase I digestion was seen using leaf nuclear proteins, but not with callus nuclear proteins.

A novel DNA binding protein RCS2 (*rolC* single stranded DNA binding protein 2) was identified in the nuclear extract of tobacco seedlings that interacted with the region from -136 to -111 nt of the *rolC* gene promoter. DNA-protein gel shift and competition

assays demonstrated that RCS2 bound to single-stranded DNA in a sequence-specific manner. A five-base direct repeat (GCATC) was shown to be important for the DNA binding of RCS2. South-Western blot analysis suggested that the size of RCS2 is approximately 43 kDa.

In the *rolC* promoter, three G-box like elements are found at -136, -219, and -364 positions. Parsley CPRF2 and CPRF3, which belong to G-box binding factor (GBF) group, interacted with these G-box like elements. Rice cDNA clone (RGBF1: rice GBF1) was isolated from root cDNA library by DNA hybridization using CPRF3 cDNA as a probe. RGBF1 contains a proline-rich region at its N-terminus and a bZIP (basic region-leucine zipper) motif at its C-terminus. A high degree of conservation was detected between the rice cDNA and GBF group proteins in both regions of the bZIP domain and N-terminal acidic domain.

General Introduction

Most dicotyledonous plants could be infected with the soil bacterium *Agrobacterium rhizogenes*, resulting in the production of adventitious roots (Moore et al., 1979; White and Nester, 1980). The phenomenon is a consequence of the expression of genes contained in the T-DNA (transfer-DNA) of Ri (root-inducing) plasmid that had been transfered to host cells and become stably integrated into host nuclear genome. Plants regenerated from root cultures possessing the T-DNA of the Ri plasmid showed abnormal morphological characteristics such as wrinkled leaves, short internodes and reduced apical dominance (Tepfer 1984; David et al., 1984; White et al., 1985; Cardarelli et al., 1987; Pasqua et al., 1987; Nakamura et al., 1988). Furthermore, insertion of the *rolC* gene, one of the genes in the TL-DNA of Ri plasmid, causes a reduction in plant growth, loss of apical dominance, and early flowering (Oono et al., 1987, 1990) Schmülling et al., 1988). Bacteria expressing RolC were shown to have a β-glucosidase activity that hydrolyses cytokinin-glucoside and releases free cytokinins (Estruch et al., 1991). Quantifications of several plant hormones in *rolC* containing plants showed that expression of the RolC protein clearly influences the metabolism of cytokinins and gibberellins, and in both cases RolC seemed to reduce the rate of synthesis of these plant hormones (Nilsson et al., 1993).

The transcriptional level of the *rolC* gene was greatly different among organs, the order being roots > stems > leaves (Nakamura et al., 1988). To investigate the specific site of expression of the *rolC* gene in plant tissues, the β-glucuronidase reporter gene (GUS)

under the control of 5'-flanking sequences of the *rolC* gene was introduced into plants and its expression was histochemically analyzed. Expression of the *rolC*-GUS chimeric gene was detected mainly in phloem cells throughout plant body (Sugaya et al., 1989; Schmülling et al., 1989). Variations of the *rolC* gene expression in different organs seems to be due to the proportion of phloem cells in each organ rather than organ specific expression.

In somatic embryogenesis in carrot, the expression of the *rolC* gene shows an interesting pattern. Although carrot cells maintained in a medium that contained 2,4-D did not generate embryos, removal of 2,4-D stimulated the activity of the product of the *rolC*-GUS chimeric gene with the development of somatic embryos. The expression of the GUS gene was detected as early as the globular stage. However, suppression of the *rolC* promoter was apparent when embryos became disorganized cells in the presence of 2,4-D. Addition of GA_3 to the basal medium did not affect the formation of embryos nor GUS activity, but ABA suppressed both phenomena to some extent. The results suggest that 2,4-D may be indirectly involved in the suppression of the *rolC*-GUS gene expression in carrot cells.

Recently, the effects of high concentrations of sucrose on activation and expression of the *rolC*-GUS fusion gene were investigated (Yokoyama et al., submitted for publication). In the cotyledons and hypocotyls of young tobacco seedlings, GUS activity increased almost 20-fold with the addition of exogenous sucrose (400 mM). Since the *rolC* promoter was not activated by sorbitol, its activation may be correlated with sucrose metabolism rather than osmotic pressure. The correlation between sucrose-dependent

activation and phloem-cell specificity of the *rolC* expression is still under investigation.

The structural characteristics of *rolB* and *rolC* promoters have been investigated. *rolB* and *rolC* share a 1.85 kbp bi-directional 5'-upstream region. Thus, *cis*- regulatory DNA elements involving the cell- or tissue-specific or signal-dependent expression described above may be located in this region. To characterize the *cis* elements, serial deletions of the 5'-noncoding regions of *rolB* or *rolC* were produced and analyzed histochemically using GUS as the reporter gene. Analysis of the *rolB* 5'-deletion mutants revealed that the region from -623 bp to -471 bp (from the translational start codon) controls the level of *rolB* expression but not its tissue specificity (Capone et al., 1991). When the region from -341 bp to -306 bp was deleted, the tissue specificity was dramatically altered. Furthermore, Leach and Aoyagi (1991) pointed out that high-level activity of the *rolC* promoter required the presence of a sequence 350 bp upstream the site of initiation of transcription. In a related study, two regions, from -860 bp to -501 bp and from -289 bp to -230 bp (from the site of the initiation of transcription of the *rolC* gene)(Sugaya and Uchimiya 1992), were found to have enhancer-like activity, whereas a region from -94 bp to +23 bp regulated phloem-specific expression (Yokoyama et al., submitted for publication).

The expression of many genes is regulated by transfactors that bind to *cis*-regulatory elements. Detection and characterization of protein(s) that bind(s) to *cis*-DNA sequences is an important step towards an understanding of the mechanism of gene expression *per se*. To detect such binding protein, a gel-shift assay was performed

using nuclear proteins extracted from tobacco hairy roots. The results indicated the presence of a protein that binds to a DNA fragment containing the region from -95 bp to +23 bp of the *rolC* gene (Kanaya et al., 1990). This protein could not be detected in extracts of tobacco leaves, in which *rolC* was expressed at low levels. DNase I footprinting analysis revealed the apparent protection of regions from -83 bp to -60 bp and from -55 bp to -50 bp (Kanaya et al., 1990). With the wheat germ extract, a different mobility pattern was observed in the gel-shift assay. Two regions of the top strand and three regions of the bottom strand were protected during the footprinting analysis (Kanaya et al., 1991). One of the two regions in the top strand was a region that was protected by both tobacco and wheat germ extracts. The DNA sequence of this region included 5'-AATATTTTTATT-3' sequence that is homologous to the AT-box in the 5'-upstream region of the light-regulated *rbcS*-3A gene from tomato (Datta and Cashmore, 1989), and the AT-rich motif in the upstream regions of the genes for leghemoglobin (Bustos et al., 1989), phaseolin (Jensen et al., 1988) and helianthinin (Jordano et al., 1989).

In addition to the DNA-binding proteins described above, a single-stranded DNA-binding protein that interacts with a DNA fragment from -94 bp to +23 bp of the *rolC* was detected in a nuclear extract of cultured cells of carrot. Methylation interference experiments showed that the G residue at position -41 on the bottom strand between the CAAT and TATA boxes was important for binding to the DNA (Suzuki et al., 1992).

In the 5'-upstream region of *rolC*, there are some sequences that have been identified as binding sites for nuclear proteins (for

example, tobacco 3AF-1, wheat EmBP-1, and maize Rc and Bc; Katagiri and Chua, 1992). The relationship between these elements and the regulation of transcription is presently being investigated.

In this thesis, (Part I) the expression of the *rolC* gene promoter in rice plants (graminaceous monocots), (Part II) the evidence demonstrating the existence of rice nuclear proteins capable of binding to specific sequences of the *rolC* upstream region, (Part III) the presence of sequence specific single-stranded DNA binding protein, and (Part IV) the isolation of rice DNA binding protein are presented.

Part I

Analysis of the *rolC* gene expression in transgenic rice plants

Introduction

Recent advances in gene transfer methods have expanded our understanding of the mechanism underlining gene expression in higher plants. In dicots, numerous studies concerning gene function and regulation have been reported from transgenic plants using the *Agrobacterium* vector system (reviews by Schell, 1987; Uchimiya et al., 1989). In contrast, such studies have not been successful in most monocots, due to a limitation of host range by *Agrobacterium*. Recently, the production of transgenic plants has been successful by direct gene transfer methods in monocots (Toriyama et al., 1988; Gordon-Kamm et al., 1990). Consequently, the analysis of gene expression in monocotyledonous plants such as rice has become feasible (Terada et al., 1990; Battraw et al., 1990).

Furthermore, the incorporation of an appropriate reporter gene such as β-glucuronidase (GUS) provided reliable means for the analysis of gene expression in higher plants (Jefferson et al., 1987). The GUS gene is now the most commonly used reporter gene in higher plants, because it is highly sensitive and can be used for histochemical detection.

In this thesis, the chimeric gene consisting of the *rolC* promoter and GUS coding region was transfered into rice protoplasts by electroporation. Molecular analysis were conducted using transgenic rice plants regenerated from the protoplasts.

Materials and Methods

plant material

Suspension cells derived from callus of the mature embryo of rice (*Oryza sativa* L. var. Yamahoushi) were maintained in a liquid AA medium (Toriyama and Hinata, 1985). The culture was kept on a gyratory shaker (100 rpm) at 25°C and subcultured every two weeks.

Plasmids

A DNA fragment (*Eco*RI/*Hin*dIII; 3.1kb) containing a promoter region (5'-upstream non-coding DNA) from the ORF12 gene (*rolC*) of Ri plasmid, the coding region of the β-glucuronidase (GUS) gene (Jefferson et al., 1987) and a terminator from the nopaline synthase gene (Bevan et al., 1983), was inserted into the *Eco*RI/*Hin*dIII sites of pUC12. The resulting plasmid was named pUC12-O12-GUS (Fig. 1).

The plasmid pCH was constructed from pUC12 and a chimeric gene comprised of the cauliflower mosaic virus (CaMV) 35S promoter, the coding region of hygromaycin phosphotransferase (hph: Gritz and Davies, 1983) and a terminator of the nopaline synthase gene (Fig. 1).

The *Eco*RI/*Hin*dIII fragment of T-DNA of Ri plasmid (ORF12; 1.9 kbp) (Oono et al., 1987) was inserted into the *Eco*RI/*Hin*dIII sites of pBR328 (Fig. 7A). The resulting plasmid TM1 was used for transformation.

A DNA fragment (*Hin*dIII/*Bam*HI; 2.0 kb) containing the maize *Ubi-1* promoter, first exon, and first intron (Christensen et al., 1992)

was inserted into pUC12-BN possessing a structural gene of *bar* (Murakami et al., 1986) and a terminator from nopaline synthase gene, resulting in the production of pUBA (Fig. 7A).

Electroporation and selection of transformants

Suspension cultured cells (4-5 day-old) were incubated in a enzyme solution containing 2.0 % Cellulase R10, 0.2 % Macerozyme R10, 0.1 % $CaCl_2 \cdot 2H_2O$ and 6.5 % mannitol, pH 5.8 for 2 hr (Toriyama and Hinata, 1985). Protoplasts were filtered through nylon mesh (30 μm), and centrifuged at 100 xg for 5 min. Pellets were suspended in buffer A (0.5 mM MES (pH 5.8), 7 mM KCl, 4 mM $CaCl_2 \cdot 2H_2O$ and 0.35 M mannitol), and re-centrifuged. Protoplasts ($1x10^6$ cells/ml) and two plasmid DNA solutions (1 μg/ml each) were mixed in the buffer A. *Hin*dIII-digested DNA (linear form) was employed for these experiments. A mixture of DNA and protoplasts in a plastic cuvette (inter-electrode distance: 0.4 cm) was subjected to electrical pulse from a 220 μF capacitor charged at 375 V/cm (hand-made product)(Toriyama et al., 1988). After electroporation, the same cuvette was kept on ice for 10 min. Protoplasts were then embedded in 1.2 % agarose medium containing B5 salts (Gamborg et al., 1968), 2 % sucrose, 5 % glucose and 2 mg/l 2,4-D. After 2 weeks, 50 μg/ml hygromycin B or 10 μg/ml bialaphos were added to the culture medium. After a further month, growing calli were placed on a medium containing N6 salts (Chu et al., 1975), 3 % sucrose, 1 mg/l kinetin and 1.2 % agarose to regenerate whole plants.

Preparation of rice DNA

Leaf tissue of rice was ground to a fine powder in liquid nitrogen using a mortar and pestle. The powdered tissue was mixed with extraction buffer containing 5 M urea, 0.35 M NaCl, 50 mM Tris-HCl (pH 7.5), 20 mM EDTA, 2 % sodium lauryl sarcosine, 0.5 % SDS, and phenol-saturated TE buffer, and an equal volume of phenol-chloroform-isoamylalcohol (75:24:1, Shure et al., 1983) was added. The solution was incubated at room temperature, and stirred gently for 10 minutes. The mixture was phase-separated by centrifugation at 2,500 xg for 5 minutes at room temperature. The aqueous phase was re-extracted with the phenol-chloroform-isoamylalcohol mixture, and the extraction procedure was repeated. Ether was added to the aqueous phase and mixed gently. The mixture was centrifuged at 2,500 xg for 10 minutes. The upper phase, containing the DNA, was removed and 10 M LiCl was added to a final concentration of 10 %. The DNA solution was centrifuged at 8,000 xg for 10 minutes at room temperature. The supernatant was mixed with 2 volumes of 99.5 % ethanol, and DNA was precipitated, rinsed in 70 % ethanol, and dissolved in dDW.

Preparation of total RNA

Total RNA was prepared from leaves with guanidium thiocyanate using the method of Chirgwin et al. (1979). Frozen tissues were ground in liquid nitrogen with a mortar and pestle. The powdered tissue was homogenized in 4 volumes of GTC buffer, (4M guanidine thiocyanate, 0.5 N-lauryl sarcosine sodium salt, 25 mM sodium citrate (pH 7.0), 0.1 M 2-mercaptoethanol, and 0.2 % Antiform A). The homogenate was layered onto a 3.4-ml pad of

5.7 M CsCl in 0.1 M EDTA and 25 mM sodium citrate (pH 7.0) in a sealed tube (Beckman 70.1Ti). The sample was centrifuged at 33,000 xg for 4.5 hr at 10°C. The RNA pellet was dissolved in GHCl solution [7.5 M guanidium hydrochloride, 25 mM sodium citrate (pH 7.0), and 5 mM DTT], and the RNA was precipitated with ethanol, then dissolved in DEP-treated water.

Analysis of DNA and RNA

DNA was prepared from leaves of rice plants according to the method of Shure et al. (1983). DNA (5 μg) was digested with restriction enzymes and run in a 1.0 % agarose gel. After treatment of DNA with 0.4 N NaOH, denatured DNA was blotted onto a nylon membrane, followed by hybridization with probe DNA which had been labeled with ^{32}P-dCTP using the Multiprime DNA Labeling System (Amersham). Probe DNA was prepared as follows: the *Pvu*II/*Hpa*I (420 bp: 5'-upstream non-coding region of *rolC* gene)(Slightom et al. 1986) fragment of pUC-O12-GUS (Fig. 1) was cloned into the *Sma*I site of pUC12, resulted in the production of pUC-O12-p420. The *Eco*RI/*Hin*dIII fragment (460 bp) of pUC12-O12-p420 was used as a probe for hybridization.

Total RNA (20 μg) was also fractionated by electrophoresis using the 1.0 % agarose gel, and was blotted onto a nylon membrane. The membrane was hybridized with the radio-labeled DNA of the GUS gene coding region.

Fluorometric GUS assay

GUS activity was determined by a fluorometric assay(Jefferson et al., 1987). For quantitative determination of GUS activity, leaf,

root, and callus tissue of rice were homogenized in an extraction buffer [50 mM NaH_2PO_4 (pH 7.0), 1 mM EDTA, 0.1 % Triton X-100, 0.1 % sodium lauryl sarcosine, and 10 mM dithiothreitol] using Eppendorf tubes and a glass pestle with quartz sand. The homogenate was centrifuged at 7,000 xg for ten minutes at 10°C. The supernatant, containing soluble protein, was used for assays of GUS activity. The assays was performed in 500 ml of an extraction buffer containing 1 mM 4-methylumbelliferyl glucuronide (4-MUG). After 60 minutes, 100 ml of the assay mixture was diluted into 2 ml of 0.2 M Na_2CO_3. The fluorescence at 455 nm was determined on a spectrofluorometer using excitation wavelength of 365 nm. The protein concentration of extracts was measured using the Bradford assay (Bradford, 1976).

GUS histochemical staining

Histochemical assays were performed using X-gluc (5-bromo-4-chloro-3-indolyl-β–D-glucuronide) as the substrate. Plant tissues were cut using a razor blade. After fixation in a solution consisting of 0.3 % formaldehyde, 10 mM MES (pH 5.6), and 0.3 M mannitol, samples were incubated in X-gluc solution [1 mM X-gluc and 50 mM NaPO4, pH 7.0] for 3 to 16 hours at 37°C.

PCR analysis

Leaf extracts were prepared according to the method of Edwards et al.(1991). The 2.5 μl aliquots of extracts (A_{260}=1.0) were added to 47.5 μl of the reaction mixture for PCR [400 μM each dNTP, 0.2 μM each primer, 2 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 50 mM KCl and 0.5 μl^{-1} of *AmpliTaq* DNA polymerase (Perkin Elmer

Cetus, U.S.A.); final concentrations]. Synthesized primers corresponding to the 5' and 3' termini of the *rolC* gene were used in the reaction (5'-ATGGCTGAAGACGACCTGTGTT-3' and 5'-TTAGCCGATTGCAAACTTGCAC-3', respectively). The temperature cycle for amplification of *rolC* was 94 °C for 2 min, 60 °C for 2 min, and 72 °C for 3 min, followed by 39 cycles of 94 °C for 1 min, 60 °C 2 min, and 72 °C for 3 min. The final extension was carried out at 72 °C for 15 min. The reaction mixture was loaded onto agarose gel.

Analysis of T_1 generation of *rolC* rice

Seeds (T_1 generation) obtained from self-pollinated T_0 transgenic plants were imbibed with water for 2 days at 30 °C. Then germinated seeds were transfered to soil in pods. Plants were grown under glasshouse conditions of 20 °C — 25 °C and a day length of 14 h. After three weeks from germination, the heights of plants were measured.

Results and Discussion

Establishment of transformants

We attempted to transform rice protoplasts with two plasmids, pCH and pUC12-O12-GUS. We were able to obtain several hundred hygromycin resistant calli, from which 26 plants were regenerated (Table 1). All plants thus obtained appeared to be albino. Every plants were able to grow in the medium containing 50 μg/ml hygromycin sulfate for successive selections.

Fluorometric analysis of GUS activity in the transgenic plants revealed that two plants expressed foreign gene products. Shimamoto et al. (1989) also reported co-transformation of rice protoplasts with the hygromycin B resistant gene and a CaMV promoter-GUS fusion gene.

Presence of foreign DNA

Analysis of DNA was made using leaf tissues of two transformants. Total DNA was digested with *Eco*RI/*Hin*dIII. As seen in Fig. 2, several bands (2.2, 3.8, 7.0, and 8.2 kb), which hybridized to a probe DNA containing 420 bp 5'-upstream non-coding region of *rolC* gene, were obvious in one of the transgenic rice plants, suggesting random integration of a foreign gene. Similar DNA blot patterns were observed in other transformants .

Expression of a foreign gene

GUS activities directed by the 5' non-coding signal of *rolC* gene in leaves and roots of rice transformants were significantly higher than in the control (Table 2). Unlike transgenic tobacco

plants possessing the same chimeric gene, organ-related-activation of GUS genes (i.e. roots > leaves) (Sugaya et al., 1989) was of vascular development in monocots and dicots. On the other hand, GUS activity was weaker in the callus induced from a root of a transgenic plant, than in leaves or roots of that plant (Table 2). Thus the expression of the *rolC* gene promoter was activated in leaves and roots, but suppressed in dedifferentiated callus. In carrot, GUS activity in callus was at a low level, but the activity in somatic embryo differentiated from callus was higher (Fujii and Uchimiya, 1991). The results suggest that the regulation of the *rolC* gene promoter is related to differentiation.

To verify that the observed GUS activity is due to specific expression of the *rolC* gene promoter but not caused by some spurious activation due to other factors, RNA was isolated from leaves and analyzed by a Northern blotting experiment (Fig. 3). In transgenic rice leaves, a band was observed, but no signal was seen in the non-transformant. The size of a band of transgenic rice was shorter than the band detected in tobacco. However, Southern blotting analysis showed that the GUS 3' coding region and nos terminator were integrated intactly. The result presented here may be due to post-transcriptional regulation in rice.

Histochemical analysis

In order to see sites of gene expression directed by the *rolC* promoter, 5-bromo-4-chloro-3-indolyl-β–D-glucuronic acid was given to tissues of transgenic rice. Results observed in leaves indicate that GUS was located only in vascular tissues (Fig. 4b). Such expression was not seen in control plants (Fig. 4a). The leaf

sheath was then dissected to evaluate GUS expression at the cellular level (Fig. 5, a-d). Blue coloration was limited only to the region where phloem cells were organized. Similar analysis was also made using roots (Fig. 6). Cellular localization of GUS expression was the same as observed in the leaf sheath. Using transgenic tobacco plants possessing same promoter-GUS fusion gene, Sugaya et al. (1989) reported phloem cell specific expression of the *rolC* gene promoter of Ri plasmid. Therefore tissue specific expression of the *rolC* gene promoter was proven to be the same both in monocotyledonous and dicotyledonous plants. It would be worthwhile to investigate common factors (both *cis* - and *trans* -) regulating the *rolC* gene promoter in higher plants.

Expression of the *rolC* gene in transgenic rice

To analyze the expression and function of RolC in rice, we introduced two plasmid, pUBA and TM1 (Fig. 7A), into rice protoplast by electroporation. Transformed cells were selected by bialaphos and regenerated into whole plants.

To investigate the presence of co-transferred *rolC* genes in the bialaphos resistant rices, PCR analysis was performed using the primers corresponding to the 5' start region and 3' stop region of the *rolC* coding sequence. The agarose gel electrophoresis indicated that the 540 bp band representing *rolC* fragment were detected in four independent transgenic plants (#1, #2, #3, and #5) (Fig. 7B).

To investigate the expression of *rolC* gene in these regenerants, Northern hybridization was performed using RNAs extracted from leaves. Fig. 7C shows that *rolC* transcript was detected in three transformants (#1, #2, #3), but not in the #5

transformant in spite of the presence of the *rolC* coding region. It may be due to integration of the *rolC* gene with inactive form for transcription. The expression of the *rolC* gene in mature leaves of the *rolC*-transformed tobacco was very weak and almost undetectable by Northern analysis. However, three transformants (#1, #2, and #3) showed high expression of the *rolC* gene in the leaves. This result may be due to differences in the ratio of vascular tissues to the other tissues in rice and tobacco leaves.

Morphological analysis in progenies of *rolC* rice

Transgenic tobacco plants possessing the *rolC* gene showed dwarfness in plant height (Oono et al., 1987; Schmülling et al., 1988). To analyze the effect of RolC protein on the morphological characteristics of rice plant, plant heights of progenies of *rolC* rice were measured. Seeds (T_1 generation) from self-pollinated transformants were germinated in the tapped water at 30 °C, then transferred to soil in pods. After three weeks, plant height of progenies was measured (Fig. 8). The mean values of plant height in transgenic rice, which contained bialaphos resistant gene but not *rolC* gene (#4), was 25.64 cm. On the other hand, the mean values of plant height in *rolC* transformant #2 and #3 were 20.66 and 22.03 cm, respectively. The distributions of plant height in *rolC* rice plants were lower than that of #4 plant. The plant height of seedlings of rice dwarf race Tanginboze is apparently lower than wild race (Suge et al., 1978). Most rice dwarf genes have function to inhibit gibberellin metabolisms, little endogenous gibberellins were detected in Tanginboze. Recently, it was shown that RolC protein influenced the metabolism of cytokinins and gibberellins

and reduced the rate of synthesis of these compounds (Nilsson et al., 1993). Thus, the reduction of plant height in *rolC* rice seedlings might be caused by alteration of phytohormone balance by effect of RolC protein.

The *rolC* gene product affects other morphological alteration, such as reduced apical dominance and miniaturized flowers (Oono et al., 1990). However, no drastic alterations in phenotypes were observed in T_0 generation of the transgenic rice plant expressing the *rolC* gene. Further analysis of RolC activity in the progenies of such transgenic rice plants is needed.

Table 1. Summary of DNA transformation resulted in the production of transgenic rice plants.(1)

Experiments	No. of Hm^R clones	No. of Hm^R clones producing plants	No. of Hm^R plants expressing GUS
1	495	22	1
2	310	22	1

(1) In each experiment, 2.5×10^6 protoplasts were employed.

Hm^R: hygromycin (30 μg/ml) resistance.

GUS: β-glucuronidase.

Table 2. GUS activity in transgenic rice plants.(1)

Plants	Activity [4-MU p mol/min/mg protein](2)		
	Leaves	Roots	Calli
Control	3.4	7.9	12.4
Transformant 1	458.8	387.7	29.8
Transformant 2	1855.0	188.0	—

(1) 4-methyl-umbelliferyl-β-D-glucuronide was used as a substrate.

(2) 4-MU: 4-methyl-umbelliferon.

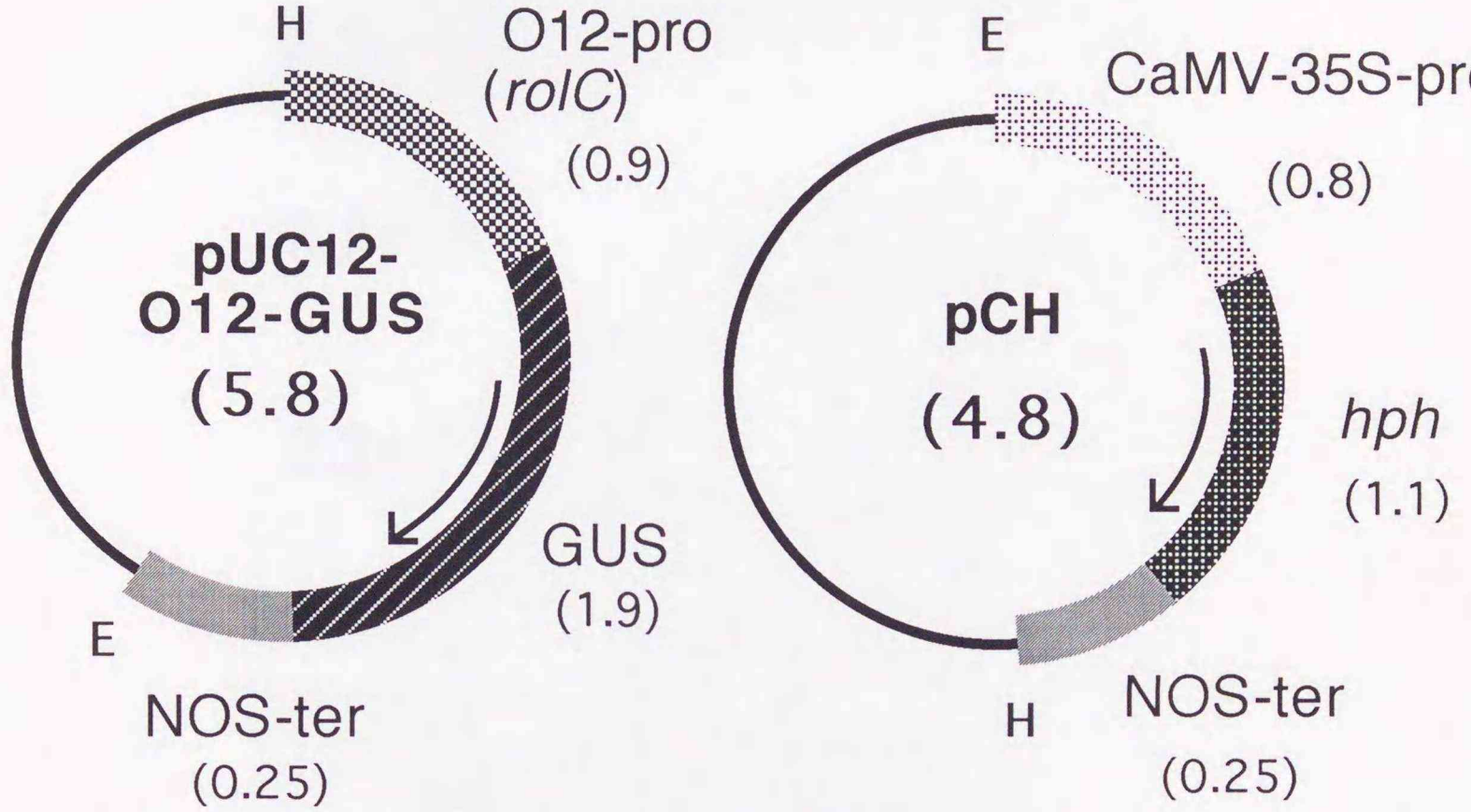


Fig. 1. Physical map of pUC12-O12-GUS and pCH. O12-pro: upstream-non-coding region of ORF12 gene (*rolC*)(Slightom et al. 1986). GUS: a structural gene of bacterial β-glucuronidase (Jefferson et al. 1987). NOS-ter: terminator of nopaline synthase gene (Bevan 1984). CaMV-35S-pro: 35S promoter of cauliflower mosaic virus (Gardner et al. 1981). *hph*: a coding region of hygromycin B phosphotransferase (Gritz and Davies 1983).
E: *Eco*RI. H: *Hin*dIII. Numerals in parenthesis indicate kb.

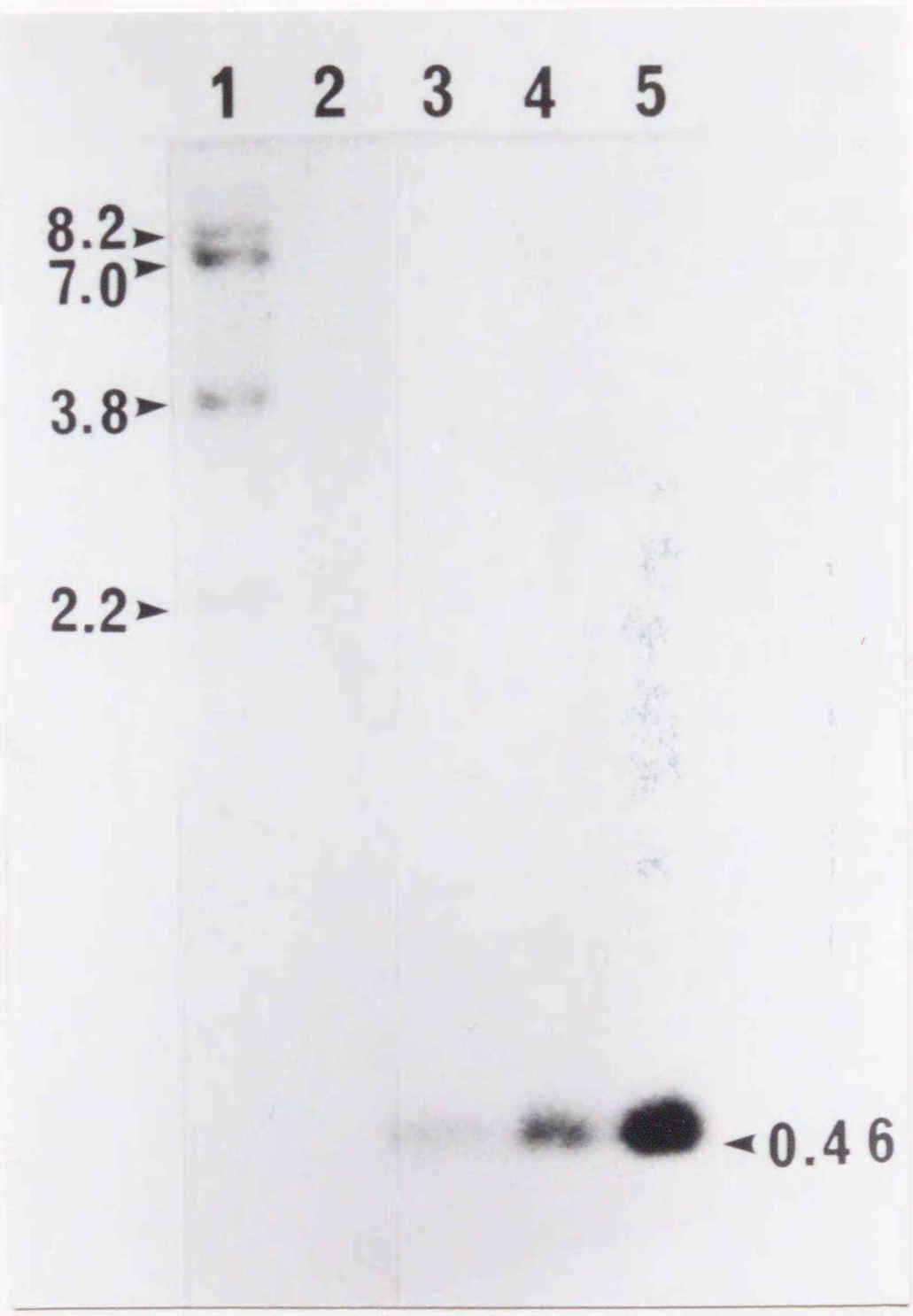


Fig. 2. An autoradiogram of DNA prepared from leaves of normal and transgenic rice plants. DNA fragments (460 bp) containing 5'-upstream non coding region of the *rolC* gene was used as a probe. 1: a transformant, 2: a control plant, 3-5: x1, x2, x5 copy reconstruction of *Eco*RI/*Hin*dIII fragment of pUC-O12-p420 (Details in the Materials and Methods).

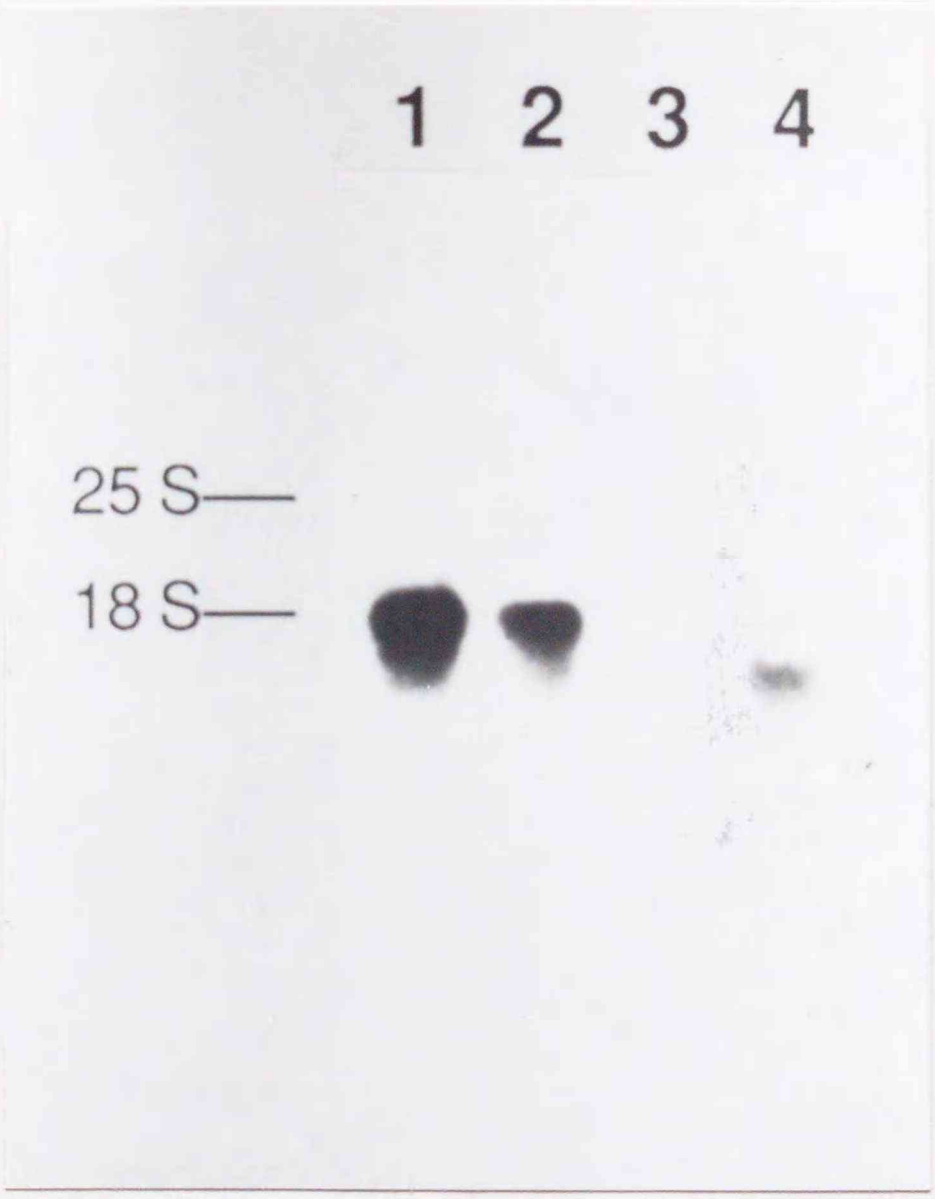


Fig. 3. Northern blot analysis of the *rolC*-GUS gene in transgenic rice plants. Lane 1: leaf midribs from tobacco plants containing CaMV35S-GUS gene, and *rolC*-GUS gene (lane 2). Lane 3: untransformed rice leaves, Lane 4: rice leaves of transformed rice. Total RNA (20 μg) was applied to each lane.

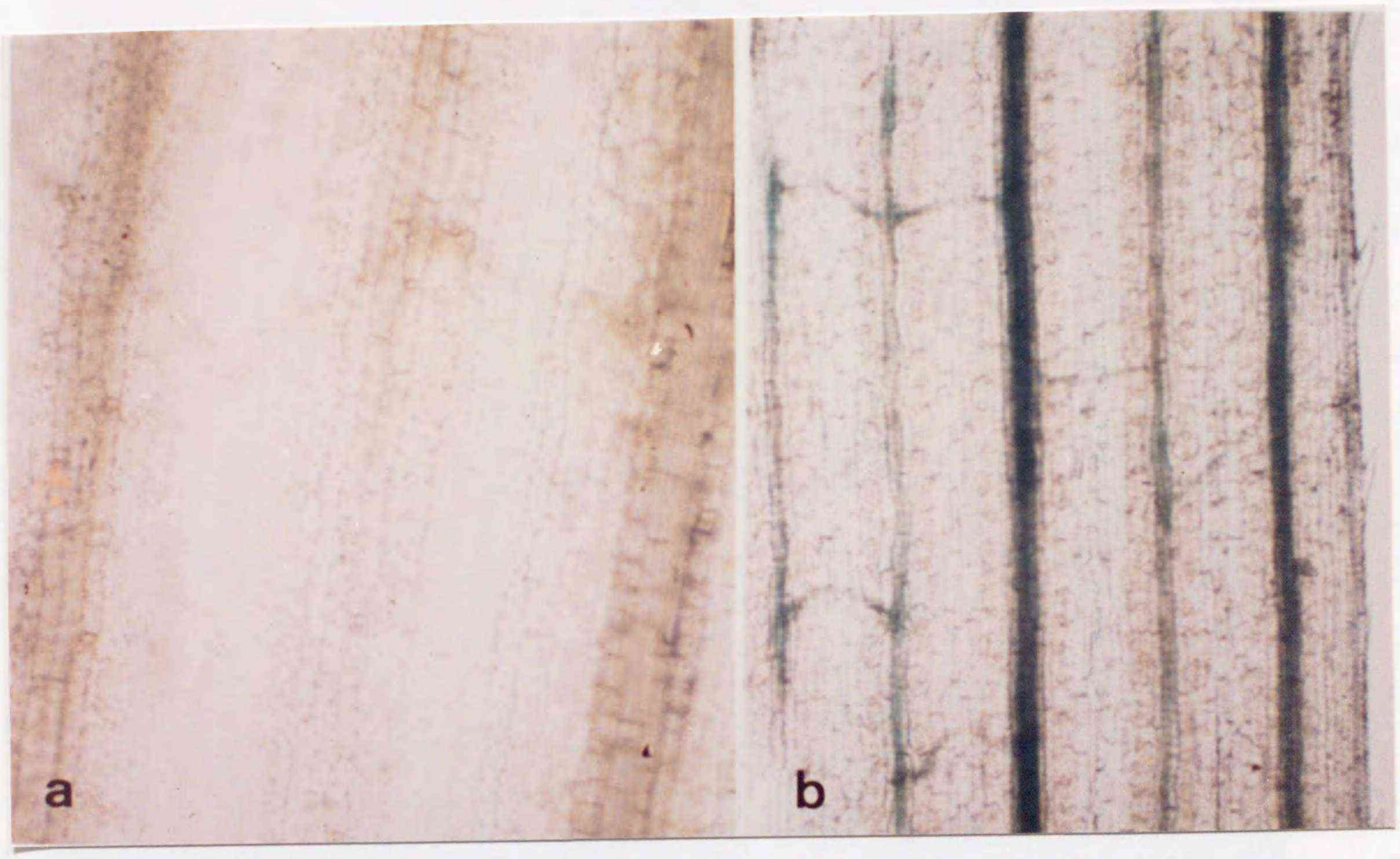


Fig. 4. Histochemical observation of GUS gene expression in leaf of rice plants. (a): a leaf of an untransformed rice, (b): a leaf of a transgenic rice plant. Note: Blue coloration indicates the sites phloem cells where GUS enzyme is present. 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid was used as a substrate.

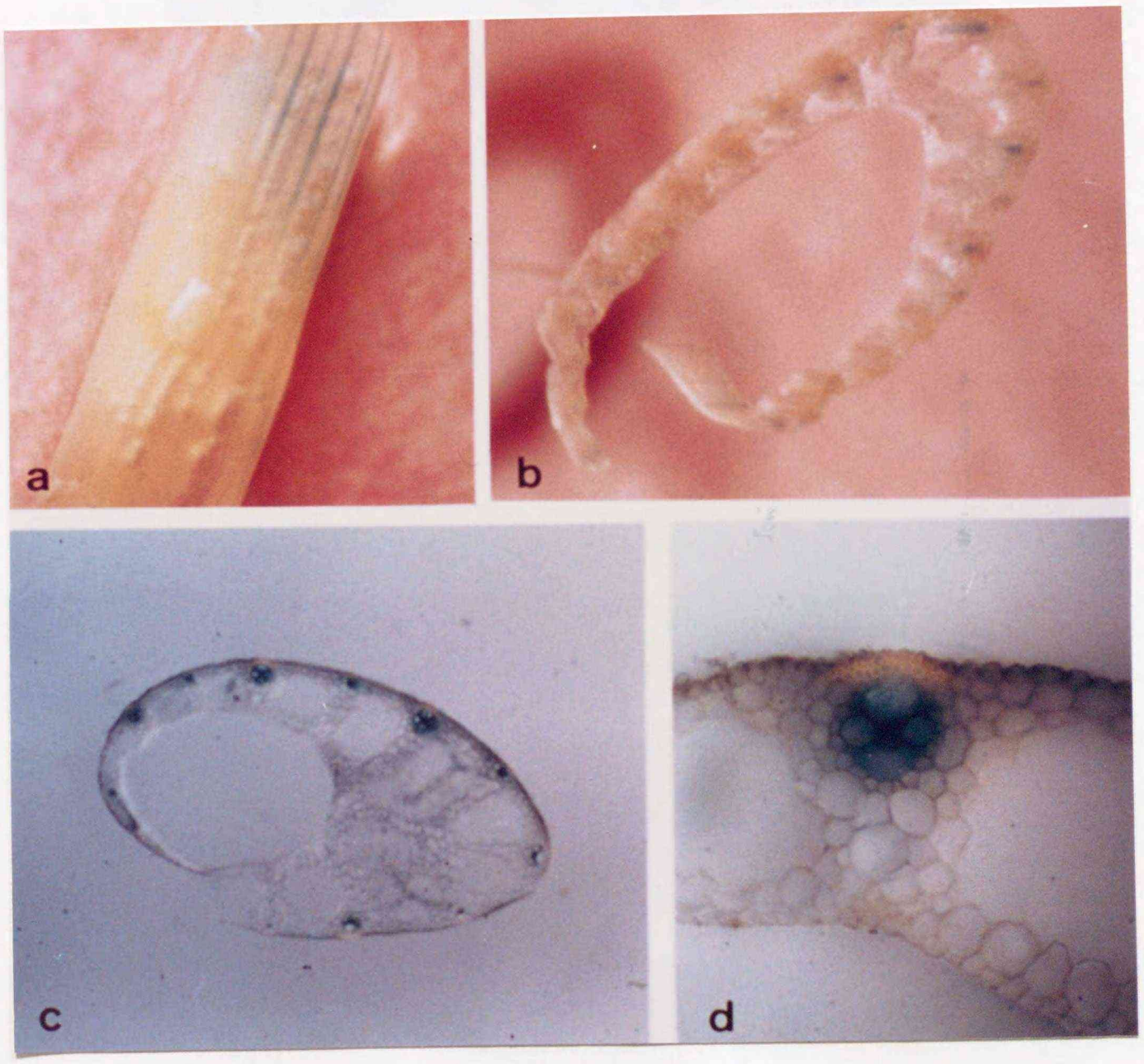


Fig. 5. Histochemical localization of GUS enzyme in leaf sheath of transgenic rice. (a): dissected leaf sheath of transformed rice, (b), (c), and (d): cross-sections of leaf sheath of a transformant.

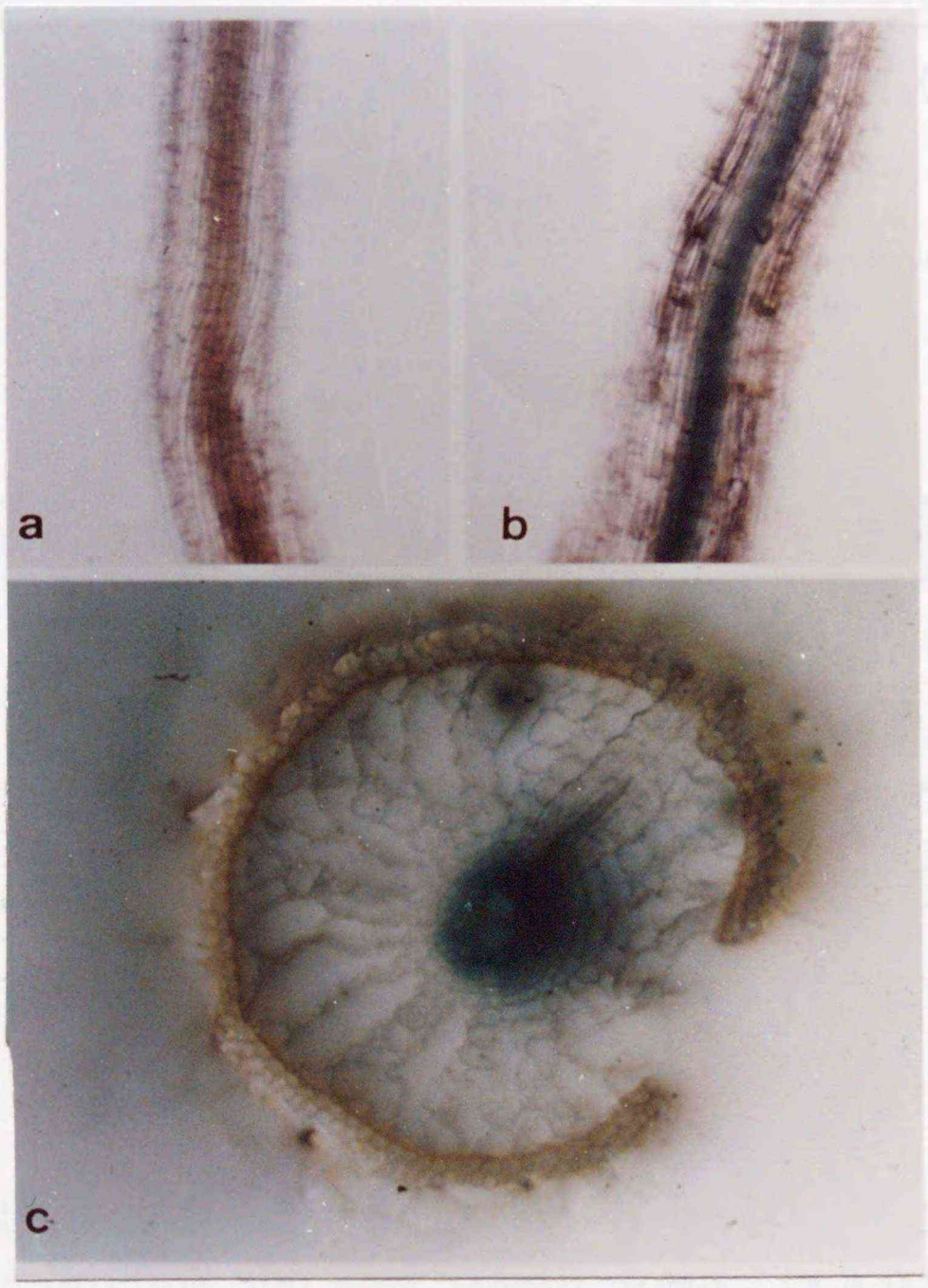


Fig. 6. Expression of GUS gene in roots of rice plants. (a): a control plant, (b): a transformant, (c): a cross section of a transformant.

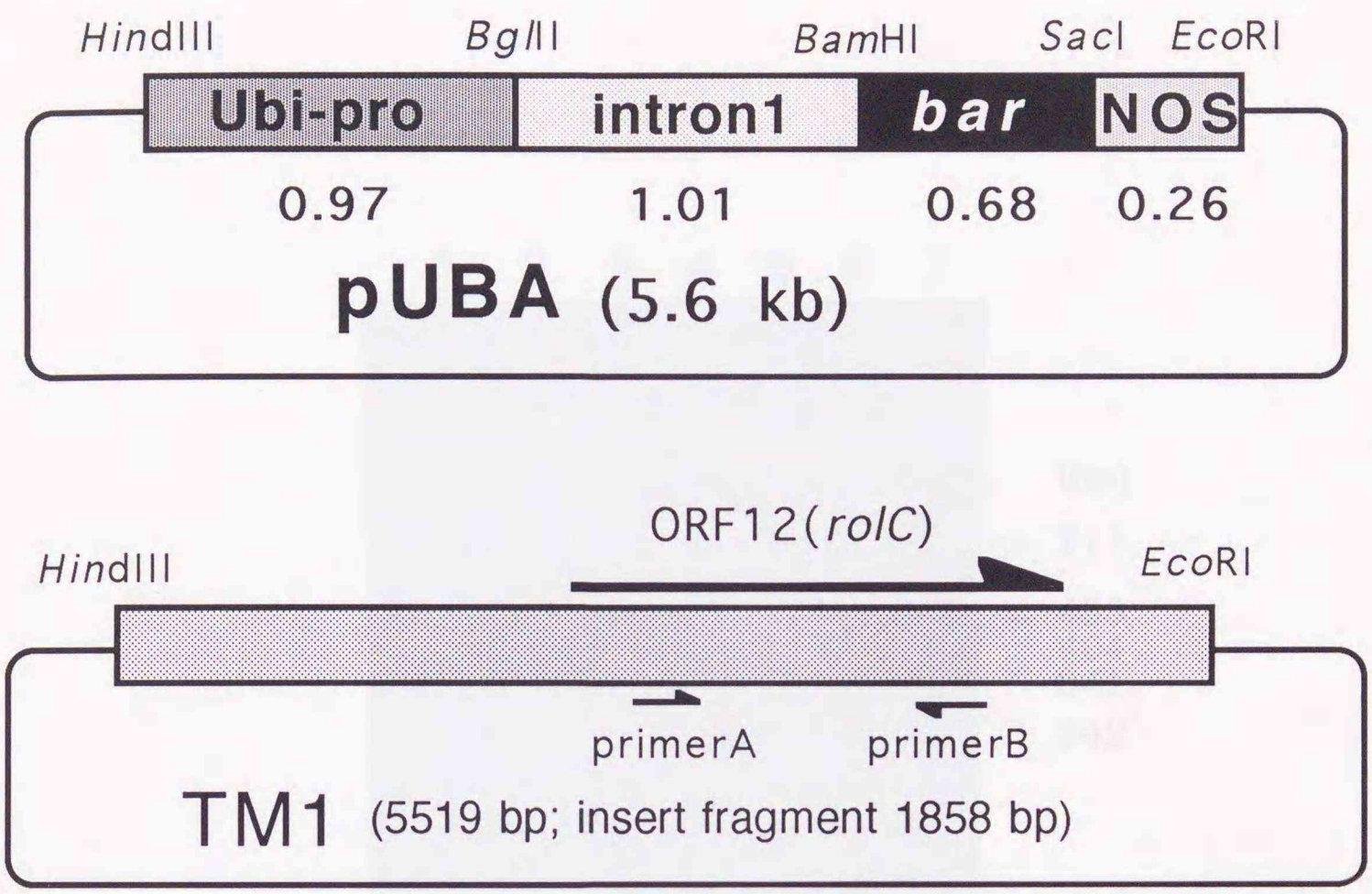
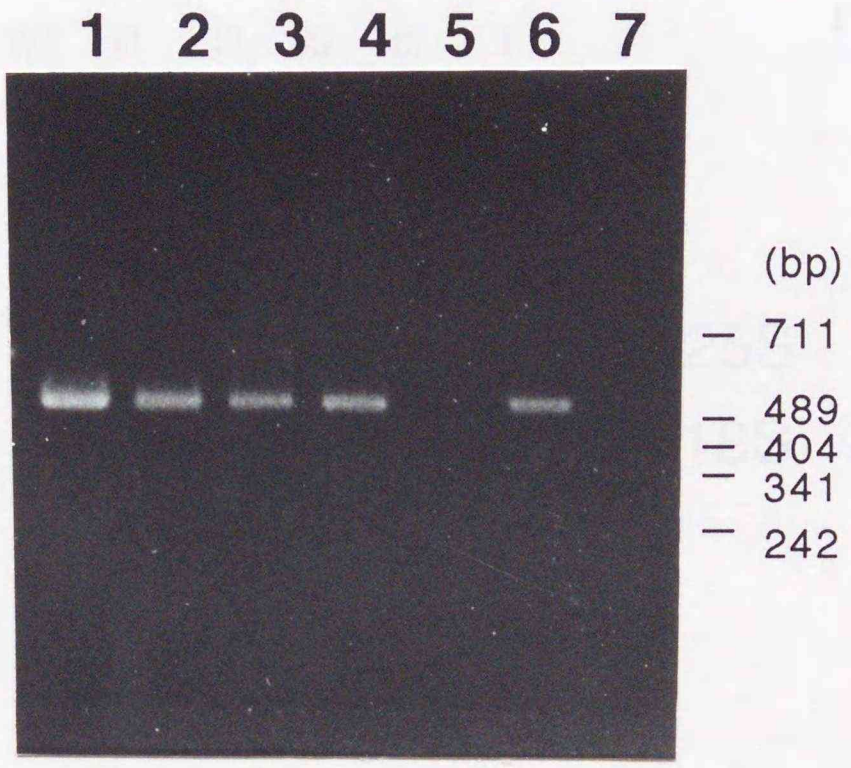
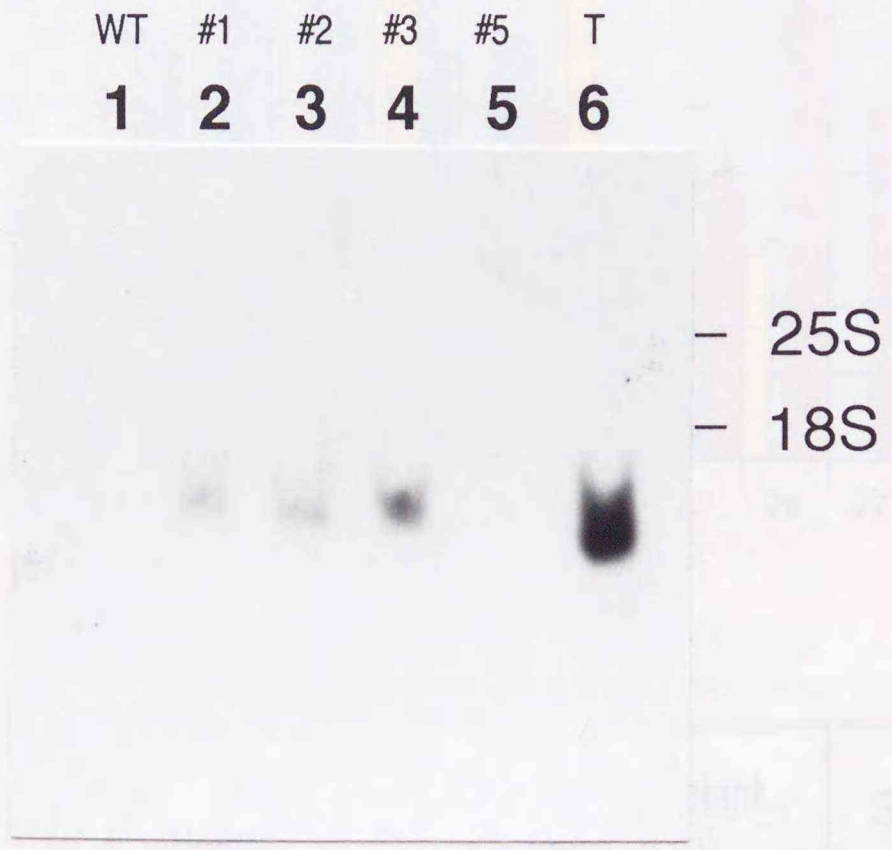


Fig. 7. PCR analysis of rice plants transformed by pUBA and TM1. (A) Diagram of pUBA and TM1. pUBA consists of a 0.97 kb maize *Ubi-1* promoter (Christensen et al., 1992), a 1.01 kb maize *Ubi-1* intron 1, a 0.68 kb *bar* (bialaphos resistance gene) coding region (Murakami et al., 1986), a 0.26 kb nopaline synthase 3' polyA region (NOS) and a 2.6 kb derivative of plasmid pUC12. Numerals indicate kb. TM1 consists of a 1858 bp *Hin*dIII/*Eco*RI fragment containing the *rolC* gene and a 3661 bp derivative of plasmid pBR328. Large arrow indicates the open reading frame 12 (*rolC*). Small arrows indicate PCR primer.

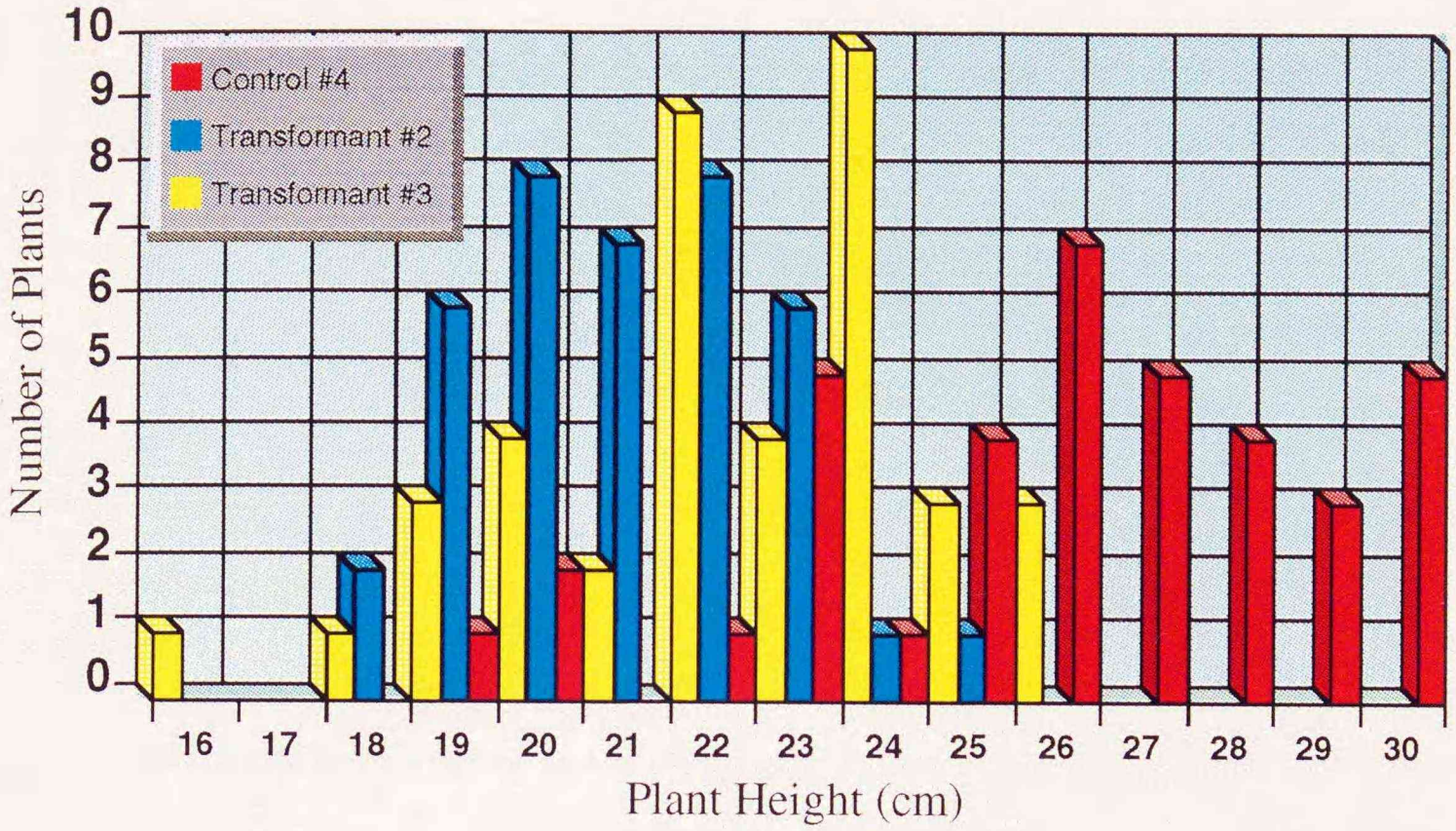


(B) PCR analysis of transgenic rices. PCR was performed with *rolC* primers. Lane 1, *rolC* transformed tobacco, lane 2 to 6, transformed rice #1, #2, #3, #4, and #5 respectively, lane 7, untransformed rice. All transformants (#1-#5) are bialaphos resistant.



(C) Northern hybridization of the *rolC* gene in transgenic plants. Lane 1, non-transformant; lane 2 to 5, transformed rice #1, #2, #3 and #5, respectively; lane 6, transgenic tobacco containing CaMV 35S-*rolC* gene. Total RNA (20 μg) from leaf was applied to each lane.

A



B

	Number of plant	Mean of plant hight (cm)	Standard error
Control #4	39	25.64	±0.50
Transformant #2	39	20.66	±0.26
Transformant #3	40	22.03	±0.36

Fig. 8. Analysis of variance for the plant height of transgenic rice plants.

(A) Distribution of plant height in T_1 population derived from self-crossed transformant. Plant height was measured at 3 weeks after germination. Control #4; transgenic rice possessing bialaphos resistance gene, but not the *rolC* gene. Transformant #2 and #3; transgenic rice expressed the *rolC* gene. Plant designations are the same as those in Fig. 7B. (B) Mean values of plant height.

Part II.

Interaction of rice nuclear proteins with 5'-upstream region of the *rolC* gene

Introduction

Transcriptional regulation of eukaryotic gene expression depends largely on the interaction of *cis*-acting promoter elements with *trans*-acting protein factors that bind to these regulatory DNA sequences (reviewed by Dynan and Tjian, 1985; Maniatis et al., 1987; Ptashne, 1988; Dynan, 1989). In yeast and animal systems, a large number of proteins capable of binding to *cis*-acting elements regulate the transcription of various genes (Jones et al., 1988; Mitchell and Tjian, 1989; Lewin, 1990). Recently, the presence of several regulatory elements in plants have been reported (reviewed by Schell, 1987; Willmitzer, 1988; Benfey and Chua, 1989). Furthermore, *trans*-acting factors have also been identified in higher plants (Tabata et al., 1989; Katagiri et al., 1989; Singh et al., 1990). In the previous chapter (Part I), I presented the evidence that expression of the *rolC* gene promoter in transgenic rice is mainly in vascular tissues of leaves and roots, but not in callus tissues. These results may indicate the presence of tissue specific *trans*-acting factors which might associate with the 5'-upstream region of the *rolC* gene.

To identify DNA-binding proteins related to the tissue specific expression of the *rolC* gene in rice plants, nuclear proteins interacting with 5'-upstream region of the *rolC* gene were analyzed.

Materials and Methods

Preparation of nuclear proteins

Nuclear extracts were prepared from 2-3 week-old rice (*Oryza sativa* L.) leaves and suspension cultured callus. Protein was isolated essentially as described by Dignam et al. (1983). Experiments, otherwise mentioned in the texts, were carried out at 4˚C. Rice leaves and callus were frozen in liquid nitrogen, and homogenized with a mortar and pestle. The fine tissues suspended in 4 volumes of buffer A [0.5 M hexlenglycol, 25 % glycerol, 5 mM $CaCl_2$, 0.25 M sucrose, 10 mM Tris-HCl (pH 8.0), 0.5 mM spermidine, 20 mM β-mercaptoethanol, and 1 mM phenylmethyl-sulfonylfluoride (PMSF)] were homogenized with a dounce type homogenizer. After filtration through a mesh (63 μm), the filtrate was centrifuged (1,000 xg, 5 min.). The crude pellet was dissolved in buffer A containing 0.5 % Triton X-100 and incubated for 5 minutes. After a centrifugation, the precipitate was washed three times in buffer A by centrifuging for 5 minutes at 1,000 xg each time. The pellet fraction containing nuclei was suspended in buffer B [10 mM Tris-HCl (pH 8.0), 5 mM EDTA, 5 mM $MgCl_2$, 25 % glycerol, 1 mM PMSF, and 1 mM dithiothreitol (DTT)] and filtered through a mesh (37 μm). To this mixture was added 0.11 volumes of 4 M NaCl with gentle stirring for 1 hour, followed by a centrifugation (100,000 xg, 30 min). The supernatant was then dialyzed in buffer C [10 mM Tris-HCl (pH 8.0), 100 mM KCl, 1 mM EDTA, 1mM DTT, and 10 % glycerol]. After dialysis, precipitated materials were removed by a centrifugation (25,000 xg, 15 min). The sample was quickly frozen and stored at -80 ˚C. Protein concentration was

estimated by the method of Bradford (1976) using a kit from Bio-Rad Laboratories.

DNA-protein gel shift assay

DNA-protein gel shift assays were performed essentially as described by Kanaya et al.(1990). Binding reaction mixture (10 µl to 20 µl final volume) contained 1 ng DNA probe fragment, 0.5 µg poly[dIdC]^2 or salmon sperm DNA treated 0.5 N NaOH for 1 hour at 100 °C, and 3-5 µg nuclear extracts in a buffer consisting of 20 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 0.5 mM $CaCl_2$, 0.1 mM EDTA, 10 % glycerol, 1.2 mM DTT, and 20 mM KCl. The reaction mixture was incubated for 5 min on ice before addition of the labeled DNA, and the incubation was continued for 25 min. In competition experiments, the unlabeled competitor DNA was added before addition of the labeled DNA. Reaction mixtures were separated on 8 % nondenaturing polyacrylamide gels in 0.5 x TBE buffer at 4 °C. Probe DNA fragments were labeled either with [α-^{32}P]-dATP for filling in the cohesive ends by Klenow large fragment or with [γ-^{32}P]-ATP for kination of the DNA fragment by polynucleotide kinase.

DNase I footprinting

DNA-protein binding reactions were similar to those described for the gelshift assay, except that the amount of probe DNA and of nuclear extracts were increased 5-fold. For DNase I cleavage reactions, DNase I was added to 5 x10^{-4} unit/ml and incubated for 30 seconds at 25°C. The reactions were terminated by the addition of EDTA to 30 mM and immediately applied to a 6 % polyacrylamide

nondenaturing gel for electrophoresis. The DNA was then electrotransferred to DE81 paper (Whatman) at 4°C for 6 hour at 60 V. The DE81 paper was exposed to X-ray film. DNAs that were recovered from the DE81 paper were resuspended in 80 % formamide 1 mM EDTA and applied to 8 % polyacrylamide sequencing gels (Maxam and Gilbert, 1980).

Results

Gel shift analysis of RC94+23 fragment

In order to understand DNA-protein interactions in the upstream region of the *rolC* gene, the gel shift assay was conducted with RC94+23 (-94 to +23) fragment. The numbering of the fragment indicates the position of the two terminal nucleotides in relation to the transcription initiation site at +1.

Using proteins extracted from nuclear fractions of rice leaves, two retarded bands (complexes of DNA-protein) were observed (Fig. 9A, lane 2). These bands were strongly attenuated by the addition of unlabeled fragment to reaction mixtures (lane 3 and 4), but were not affected by the addition of pUC12 plasmid DNA. These results suggested that specific factors interacting with the RC94+23 exist in the rice leaf nuclear proteins, and these DNA-protein interactions were specific to the RC94+23 fragment. When non specific competitor poly[dIdC] was titrated in the binding reaction, a new strong band was detected at a low concentration of poly[dIdC] (Fig. 9B, lane 1 and 2).

Similarly, nuclear proteins prepared from suspension cultured cells were analyzed for binding ability to RC94+23 fragment (Fig. 9C). Five retarded bands were observed (lane 2). The sequence specificities of such DNA-protein interactions were confirmed by competition experiments. Four bands (C1, C3, C4, and C5) disappeared by the addition of 30-fold non-radiolabeled RC94+23 fragment, but they were resistant to the pUC12 DNA. Thus, proteins from rice callus also exhibited the specific binding abilities to the RC94+23 fragment. The migration patterns of C2 and C3 were same

as that of L1 and L2, respectively. These bands may indicate the occurrence of the same DNA-protein complexes. Therefore, the presence of proteins in both leaves and calli, which bind to the same DNA fragment was confirmed.

DNase I footprinting of RC94+23 fragment

RC94+23 fragment was analyzed by DNase I footprinting to determine the precise sites of DNA-protein interactions. After binding reaction with RC94+23 and leaf nuclear proteins, the mixture was treated with DNase I, and immediately electrophoresed on a non-denaturing gel. DNAs recovered from protein-bound DNAs (Fig. 9B, L3) and unbound DNAs (Fig. 9B, F) were separated on a sequencing gel with Maxam and Gilbert ladders of the same fragment. On the top strand, the region from -76 to -67 was protected against DNase I digestion by the binding of the nuclear protein (Fig. 10A). Although the DNase I footprinting experiments on the bottom strand were not performed, the nuclear protein may interact with this region and form a L3 complex.

Fig. 10B. showed the results of a similar experiment using callus nuclear protein. When DNA recovered from the C1 complex (Fig. 9C) were electrophoresed, the protected region was found in -76 to -67. The same sequence "ATATTTTTAT" as recognized with leaf and callus nuclear proteins, in spite of the different migration of L3 and C1.

DNA-protein interactions in the upstream region of RC203-92 fragment

To investigate the interaction of further upstream region, RC203-92 (-203 to -92) fragment was used for a gel shift assay. Using leaf nuclear proteins, a DNA-protein complex was observed (Fig. 11A, lane 2). However, no DNA-protein complex was detected in callus nuclear proteins. Thus, factors interacting with RC203-92 fragment may exist only in leaf nuclear proteins.

To verify the specific sequences of RC203-92 fragment interacting with proteins, DNase I footprinting was carried out. DNAs recovered from DNA-protein complexes was scarcely digested by DNase I at the AT-rich region from -203 to -164 (Fig. 11B). The result of these footprinting experiments is summarized in Fig. 12.

Discussion

Objective of this investigation was to verify the specificity of DNA-protein interactions using the *rolC* gene promoter in monocotyledonous plants. Because the *rolC* gene promoter activation was demonstrated in rice plants, the experiments using leaves and calli of rice were performed to study the DNA-protein interactions. DNase I footprinting analysis of RC94+23 fragment showed the specific DNA region from -76 to -67 which interacts with nuclear proteins from either leaf or callus tissues of rice plants. Of special interest is that the same region also interacts with the nuclear proteins extracted from tobacco hairy roots (Kanaya et al., 1990), and wheat germ (Kanaya et al., 1991). Therefore, proteins interacting with this region may exist in various plant species and tissues.

The sequence (-76 to -67) contains a common sequence, namely a AT-1 box "ATATTTTTATT", which was reported in the upstream region of the pea *rbcS* and *cab*-E genes (Datta and Cashmore, 1989). Other homologous sequences with AT-1 box were noted in the upstream regions of the soy bean leghemoglobin gene (Jensen et al., 1988), the French bean β-phaseolin gene (Bustos et al., 1989), and the carrot lipid body membrane protein gene (Hatzopoulos et al., 1990). These sequences are also known to interact with nuclear proteins. Thus, the AT-1 box and similar sequences were widely spread into various gene promoters, because the sequence interacted with nuclear proteins from various sources. Although the role of the AT-1 box-protein complex is

unknown, the AT-1 box-like sequence in the *rolC* gene promoter may be a basic element involved in the regulation of transcription.

The nuclear extracts from calli and leaves of rice plants recognized the same region of the RC94+23 fragment, as the AT-1 box-like-sequences. However, from the gel shift analysis of RC94+23 fragment, differences were observed in the DNA-protein complexes between calli and leaves. Such ambiguity can be explained as follows. One possibility is that the proteins recognizing the AT-1 box-like-sequences may differ in nuclear extracts of calli and of leaves. It is a known fact that various DNA-binding proteins recognize a unique site, such as the octomer element (Scholer et al., 1989; Araki et al., 1988). Alternatively, in the case of leaf nuclear extracts, protein-protein interactions may be crucial for the formation of DNA-protein complexes. In fact, many nuclear proteins interact with each other in the regulation of the transcription (Lewin, 1990). Thus, the latter possibility may be appropriate to speculate the DNA-protein interaction in the transcriptional regulation of the *rolC* gene. That is because the *rolC* promoter is highly activated in leaf tissues, but less in callus tissues of rice plant. However, further analyses remain to be done.

DNase I footprinting analysis of RC203-92 fragment shows that the specific sequences (-203 to -164) interact with nuclear proteins. This region is also an AT-rich sequence as the sequence observed in RC94+23 fragment. The analysis of the transcription activity using the internal deletion mutants of the *rolC* promoter in transgenic tobacco suggested that the region from -196 to -135 was one of the positive regulatory elements (Yokoyama et al., submitted for publication). Therefore, the nuclear proteins which interact with

the sequence from -203 to -164 must be one of the positive transfactors for the *rolC* gene expression. This notion may be supported by the following findings. Gel shift assay of RC203-92 fragment indicated that the proteins interacting with this fragment existed in leaf tissues, in which the *rolC* gene promoter was active. On the contrary, any proteins interacting with this fragment were not detected in the nuclear extracts of calli where the *rolC* gene promoter was not active. As a conclusion, the factors interacting with this DNA region may play a crucial role in the regulation of transcription of the *rolC* gene. Further analysis may be needed in order to understand the biological significance of the *rolC* gene promoter in monocotyledonous plants.

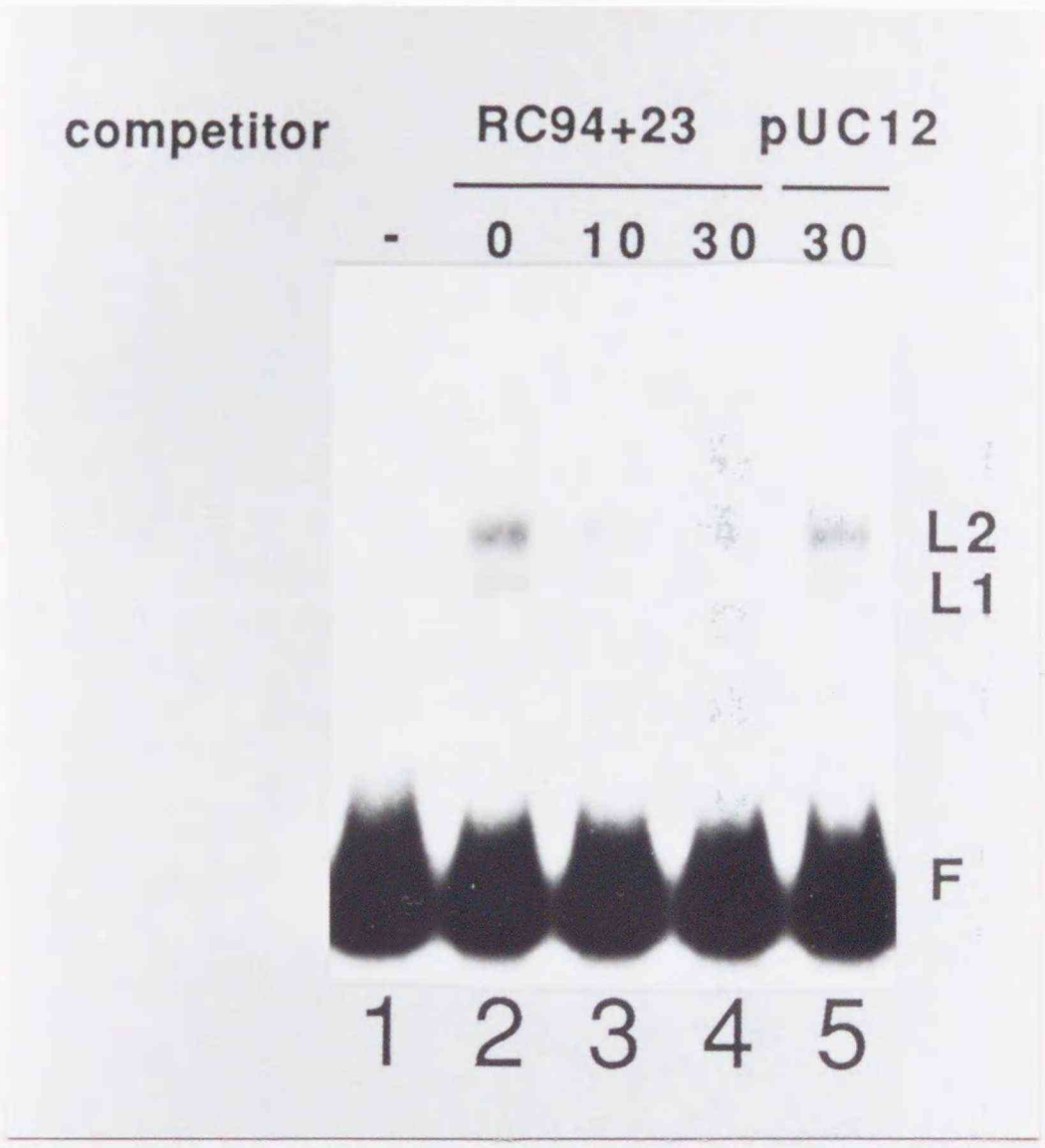
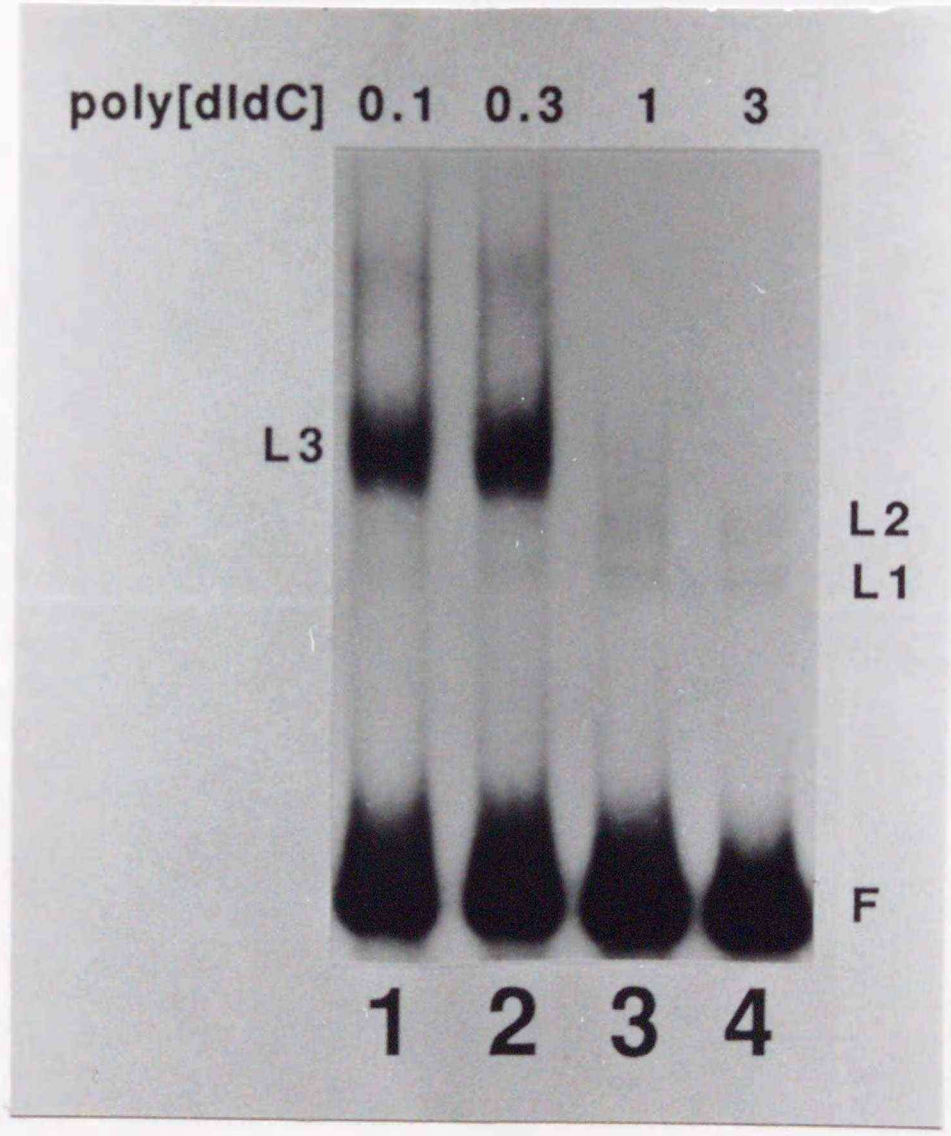
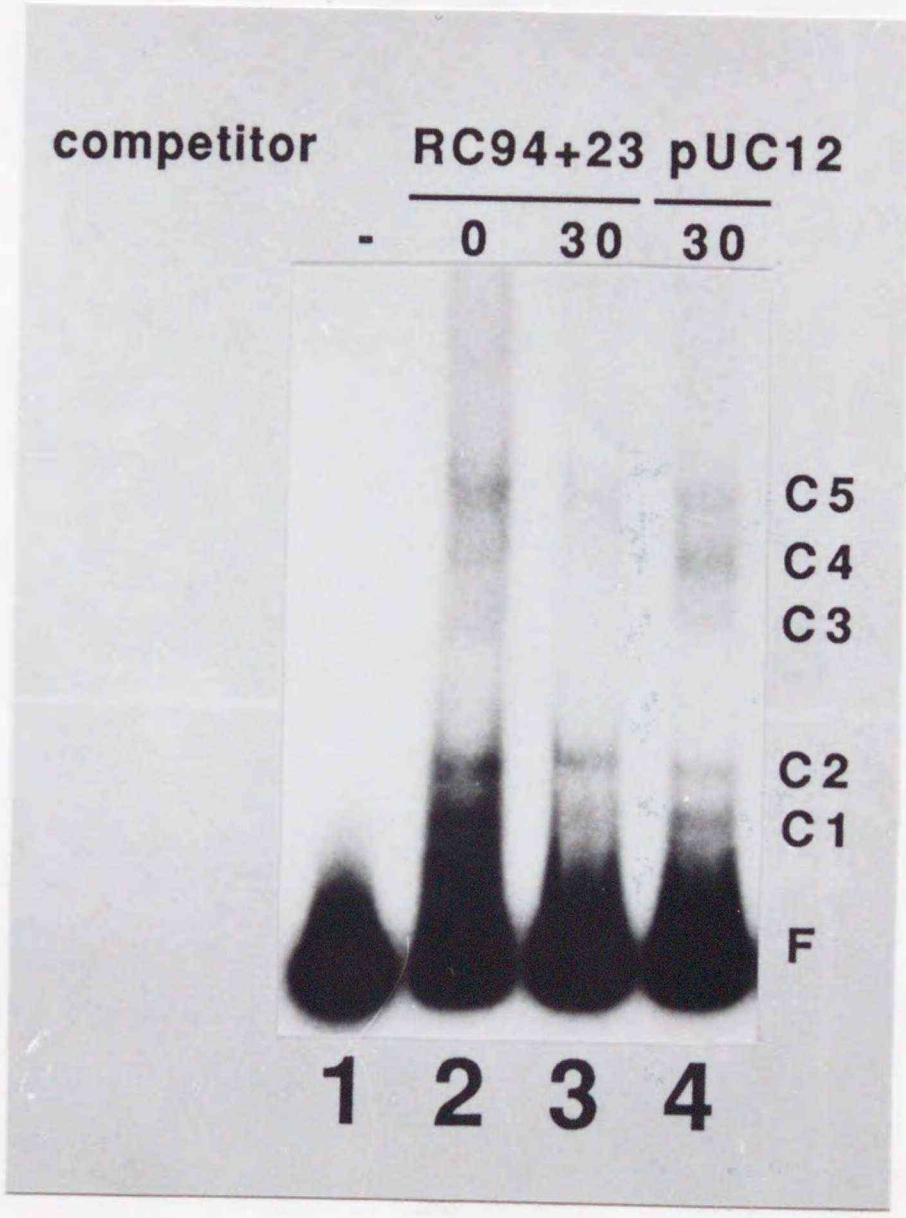


Fig. 9. Gel shift assay of RC94+23 fragment.
(A) A labeled RC94+23 fragment was incubated with 3 μg of nuclear extract from rice leaves (lane 2 to 5). Lane 1: no added protein extract. Unlabeled competitors were added in lane 3 (10-fold RC94+23), lane 4 (30-fold RC94+23), and lane 5 (30-fold pUC12 plasmid DNA). L1 and L2 refer to complexes formed with leaf nuclear protein bound to RC94+23 fragment. F is the unbound DNA fragment.



(B) The gel shift assay of RC94+23 fragment using leaf extract in the presence of increasing concentrations (lane 1: 0.1 μg, lane 2: 0.3 μg, lane 3: 1 μg, lane 4: 3 μg) of poly[dIdC] as the nonspecific competitor.



(C) Gel shift assay of RC94+23 fragment. A labeled RC94+23 fragment was incubated with 3 μg of nuclear extract from rice cultured cells (lane 2 to 4). Lane: no added protein extract. Unlabeled competitors were added in lane 3 (30-fold RC94+23), and lane 4 (30-fold pUC12 plasmid DNA). C1, C2, C3, C4, and C5 refer to complexes formed with callus nuclear protein bound to RC94+23 fragment. F is the unbound DNA fragment.

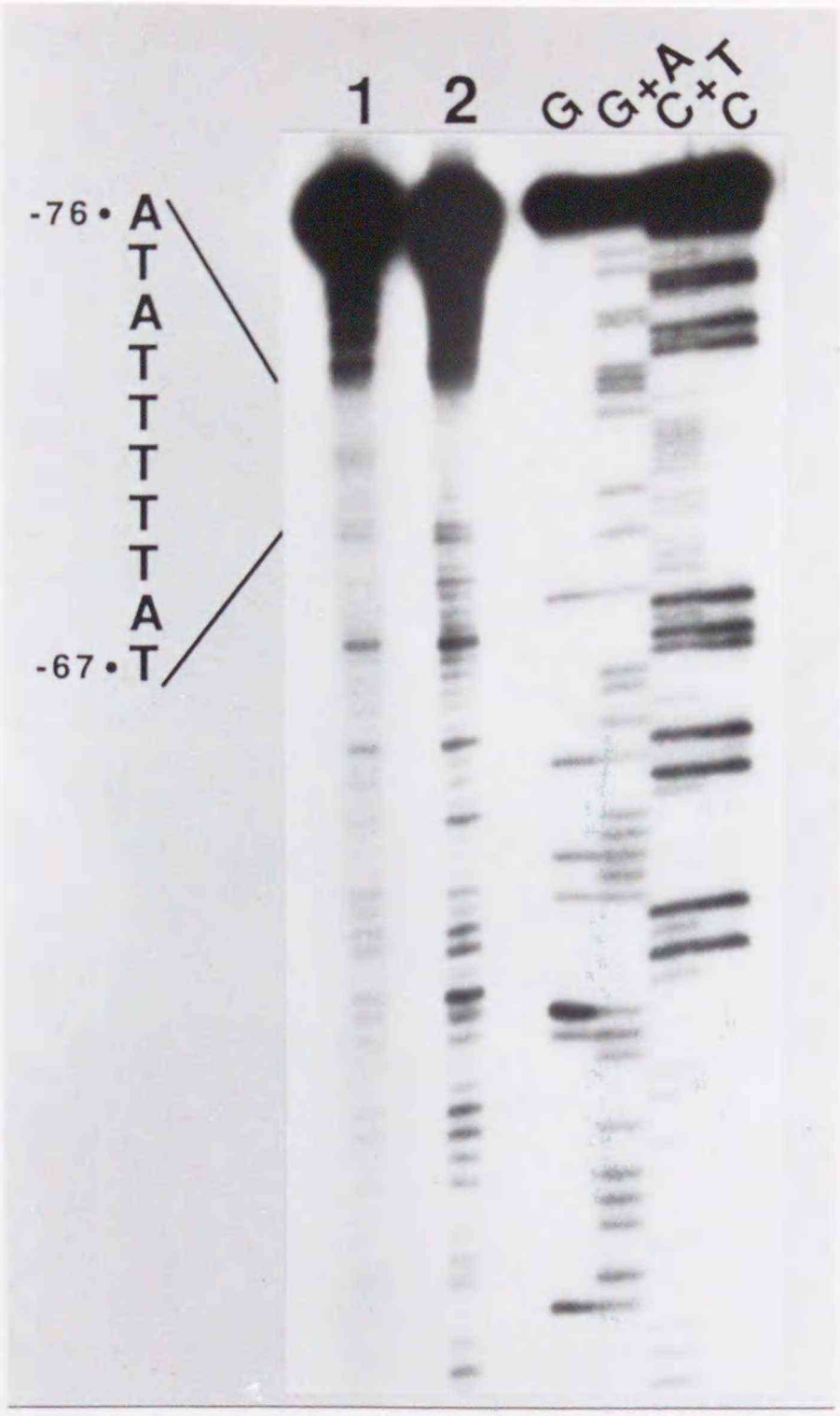
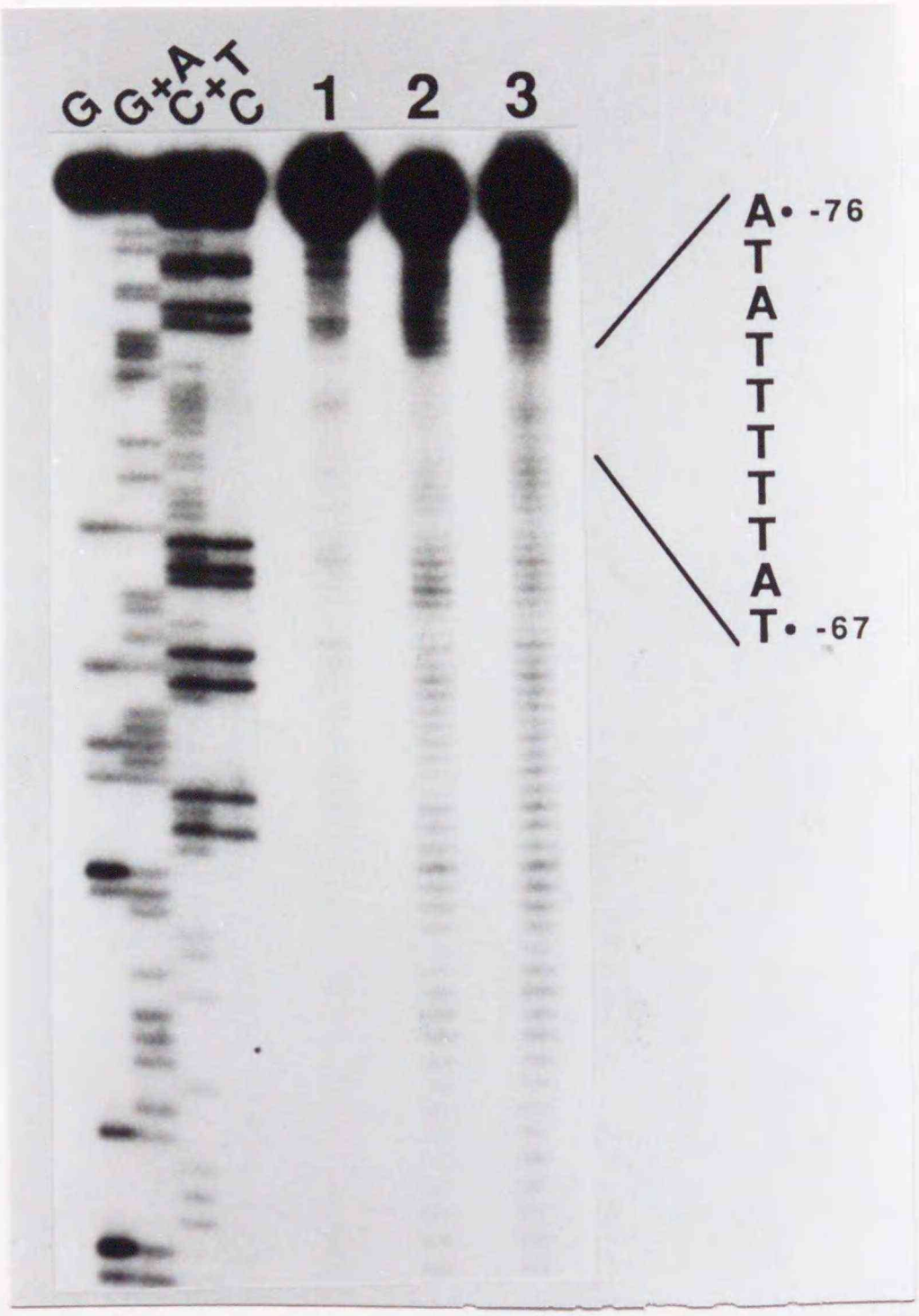


Fig. 10. DNase I footprinting of RC94+23 fragment.
(A) RC94+23 fragment was incubated with leaf nuclear protein and subsequently treated with DNase I. After electrophoretic separation of the DNA-protein complexed, DNAs from the complex (Fig. 9B, L3) and free-probe (Fig. 9B, F) bands were recovered and applied to a sequencing gel. Lane 1: DNA recovered from F, and L3 complex (lane 2). G, G+A, C+T, and C are Maxam and Gilbert sequencing reactions. free-probe (Fig. 9B, F) bands were recovered and applied to a sequencing gel. Lane 1: DNA recovered from F, and L3 complex (lane 2). G, G+A, C+T, and C are Maxam and Gilbert sequencing reactions. Letters indicate the sequences protected from DNase I digestion. Numbers indicate distances from the initiation point of transcription.



(B) RC94+23 fragment was incubated with callus nuclear protein and subsequently treated with DNase I. After electrophoretic separation of the DNA-protein complexed, DNAs from the complex (Fig. 9C, C1) and free-probe (Fig. 9C, F) bands were recovered and applied to a sequencing gel. Lane 1: DNA recovered from F, C1 complex (lane 2), and nonspecific complex (lane 3). G, G+A, C+T, and C are Maxam and Gilbert sequencing reactions. Letters indicate the sequences protected from DNase I digestion. Numbers indicate distances from the initiation point of transcription.

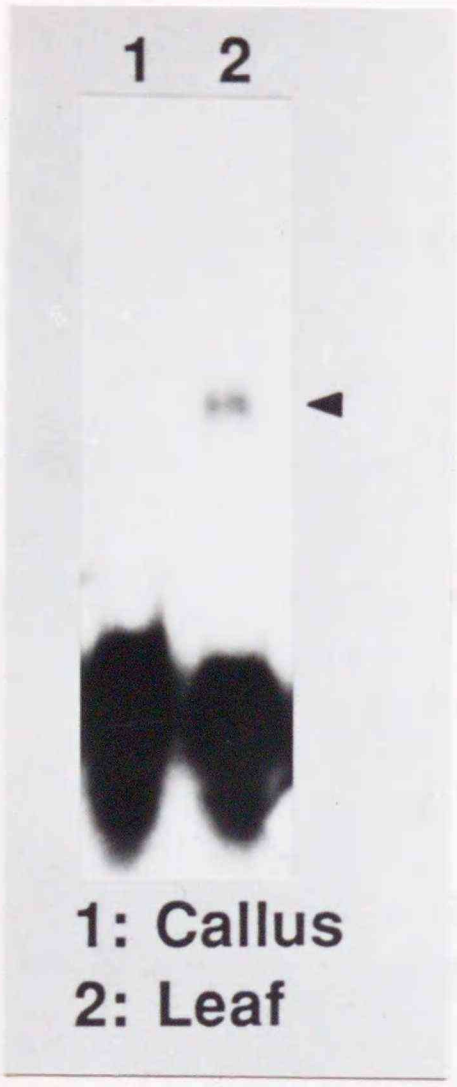


Fig. 11. Gel shift assay and DNase I footprinting of RC203-92 fragment.
(A) A labeled RC203-92 fragment was incubated with callus extract (lane 1) and leaf extract (lane 2).
(B) DNase I footprinting of RC203-92 fragment was performed using leaf nuclear extract. The complex (lane 1) and free probe (lane 2) were isolated from a gel similar to that of (A) lane 2.

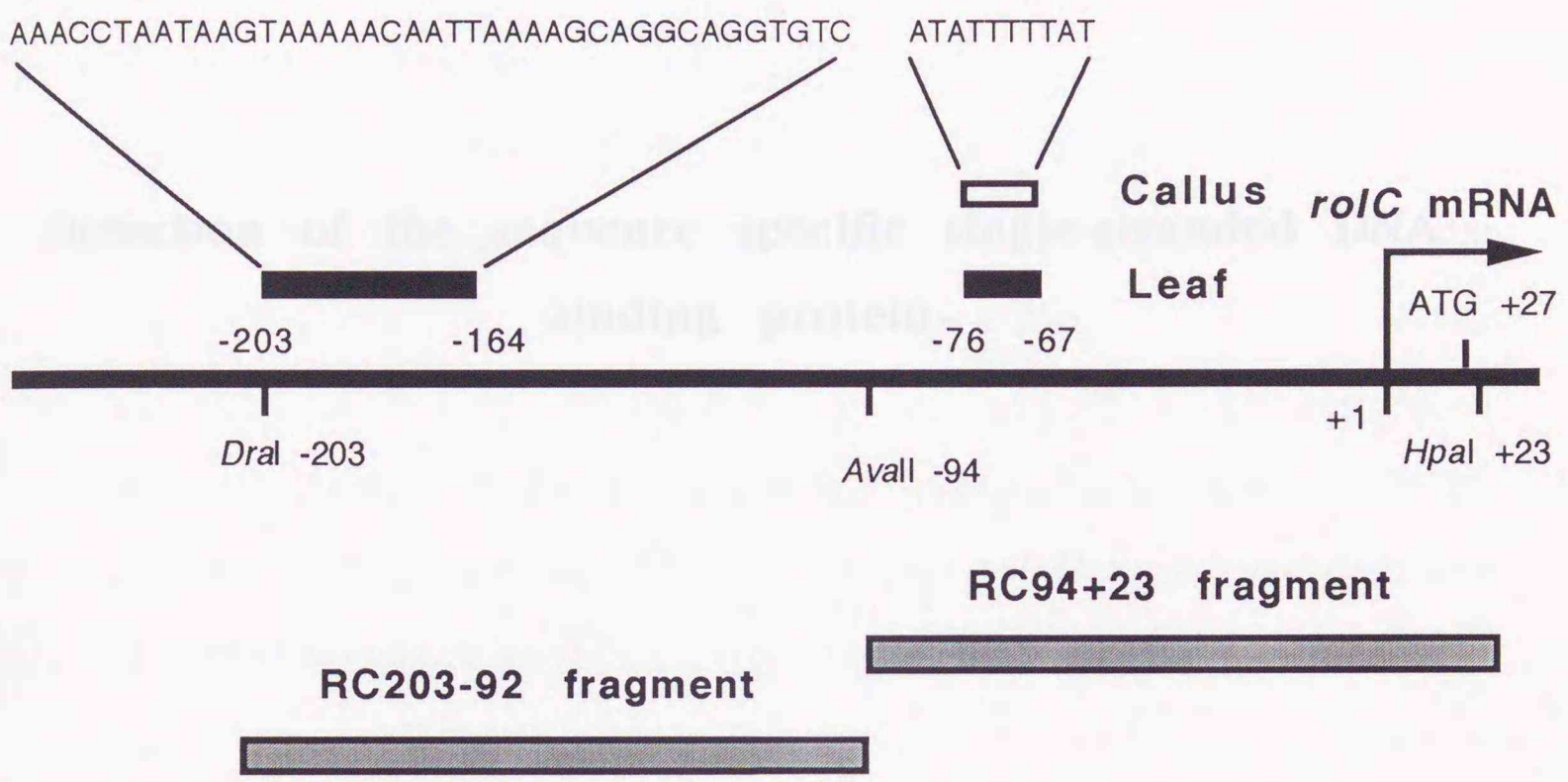


Fig. 12. A schematic illustration of the *rolC* gene upstream region. Numbers indicate the distance from the transcription start point. Bars represent the protected regions defined by DNase I footprinting using leaf nuclear extract (closed) and callus nuclear extract (open).

Part III.

Detection of the sequence specific single-stranded DNA binding protein

Introduction

In our earlier studies, we identified the sequence-specific DNA-binding proteins, RC1 (*rolC* DNA binding protein 1) and RCS1 (*rolC* single-stranded DNA binding protein 1), in nuclear extracts that interact with the proximal sequences of the *rolC* gene promoter regions (Kanaya et al., 1990, 1991; Suzuki et al., 1992). RC1 interacts with the region from -76 bp to -67 bp, namely AT-1 box, in a double-stranded sequence-specific manner. Another DNA-binding protein, RCS1, was capable of binding to single-stranded DNA in the vicinity of the TATAA box. In this Part, I describe the novel DNA-binding protein RCS2 that binds to the top strand of the proximal regulatory sequence of the *rolC* gene promoter.

Materials and Methods

Preparation of nuclear proteins

Nuclear extract was prepared from 10-day-old tobacco (*Nicotiana tabacum* L. var. Petit Havana SR1) seedlings grown on MS medium (Murashige and Skoog, 1962) lacking phytohormones at 25 °C. Protein was isolated essentially as described by Dignam et al.(1983).

DNA-protein gel shift assays

DNA-protein gel shift assays were performed essentially as described by Kanaya et al.(1990). Probe oligonucleotides were labeled with [γ-^{32}P]-ATP by T4 polynucleotide kinase.

South-Western Blotting

South-Western blot analysis was performed essentially as described (Vinson et al., 1988). Nuclear extract was incubated in a sample buffer containing 15 % glycerol, 2 % SDS, 5 % β-mercaptoethanol, 60 mM Tris-HCl (pH 6.8), and 10 μg/ml bromophenol blue for 15 min at room temperature. The samples were loaded onto 12 % SDS-polyacrylamide gel. Electrophoresis was carried out using the buffer system of Laemmli (1970). The separated proteins were then electroblotted onto PVDF membrane (Millipore, Japan). The membrane was pretreated in blocking solution (5 % non-fat milk powder, 10 mM Tris-HCl (pH 8.0), 50 mM KCl, 1 mM EDTA, and 1 mM DTT) for 1 h at 4°C. The membranes were then incubated separately with ^{32}P-labeled oligonucleotides in the absence or the presence of a molar excess (25-fold) of the

competitor oligonucleotides. Incubation was performed for 1 h at 4 °C in binding buffer (0.25 % non-fat milk powder, 10 mM Tris-HCl (pH 8.0), 50 mM KCl, 1 mM EDTA, 1 mM DTT, 10 mg/ml denatured salmon sperm DNA, and $1x10^6$ cpm/ml probe DNA). The membrane was washed three times in the binding buffer without salmon sperm DNA and probe DNA for 15 min at room temperature.

RESULTS

Detection of single-stranded DNA binding activity

We carried out 5' deletion analysis using the *rolC* gene promoter-GUS chimeric gene in transgenic tobacco plants. Our results indicated that the region from -153 bp to downstream is essential for the phloem cell specific expression of *rolC*-GUS in tobacco seedlings (Sugaya and Uchimiya, 1992). We also detected a nuclear protein which binds to the region from -76 bp to -67 bp of the *rolC* gene (Kanaya et al., 1990). In the experiment described in this part, I analyzed nuclear proteins from tobacco seedling for *in vitro* binding ability to -136 bp to -111 bp region of the *rolC* gene promoter, which was demonstrated to be a critical *cis*-acting DNA element.

Figure 13 shows oligonucleotide probes and competitors that were used in gel mobility shift assays. When Top 1 fragment annealed with Bottom 1 was used as the probe, no complex was detected (Figure 14, lane 2). In the case of heat denatured oligonucleotide probe, three bands of DNA-protein complexes were observed (lane 4). Single-stranded oligonucleotides designated as ssDNA was increased by heat treatment. Thus, three bands may represent complexes formed between the single-stranded DNA and nuclear proteins.

To test this possibility, I conducted gel mobility shift assays with single-stranded oligonucleotides probe. Figure 15A shows that two complexes were formed when the Top I oligomer was incubated with the nuclear extract. The fast-migrating complex T1 was competed out by excess amount of unlabeled Top I oligomer

(lane 3-5). But, the formation of the T1 complex was not prevented either by the Bottom 1 oligonucleotides or by the SB 1 oligonucleotides (lane 6-11) and by the annealed Top 1-Bottom 1 oligonucleotides (Fig. 15C lane 4). When the excess complementary oligonucleotides was added as a competitor, the intensity of the T1 complex was slightly diminished (lane 8). This reduced activity may be caused by the circumstance where the labeled Top 1 oligomer and its anti-strand oligonucleotide Bottom 1 formed the double strand DNA, designated as dsDNA, before the formation of the protein-DNA complex. Evidently, the amount of the complex was decreased. On the other hand, the slower migrating complex NC was competed out by the presence of unlabeled Top 1, Bottom 1 and SB 1 oligonucleotide, respectively. This retarded band disappeared when denatured salmon sperm DNA was added as a nonspecific competitor (Fig. 15C lane 2,3). These data suggest that the band (NC) was non-specific complex. With the labeled Bottom 1 probe, only the complex NC' was observed (Fig. 15B). This complex was not also formed in sequence-specific manner.

A direct repeat within Top 1 is important for binding

The Top 1 DNA contains a direct repeat sequence of GCATC motif with a 7 nt spacer. Thus, to investigate whether nuclear proteins interact with this direct repeat sequence, competition assays were carried out by using synthetic oligonucleotides containing mutations within the repeat sequence (GCATC) of the Top 1 (Fig. 16). Mutations in either repeat sequence alone inhibited competitor activity (Fig. 16). Using mutated oligonucleotides as probes, no retarded bands were detected. These data indicate that

the direct repeat sequence within Top 1 DNA is important for the binding.

South-Western blot analysis

To estimate the molecular weight of the protein forming a complex with the Top 1 DNA, we performed South-Western blot analysis (Fig. 17). Using an extract from tobacco seedlings and the labeled Top 1 probe, we detected a DNA-binding signal at approximately 43 kDa. This binding was competed by excess Top 1 DNA, but not by its complementary oligonucleotide or non-specific single-stranded DNA sequences (lane 3, 4).

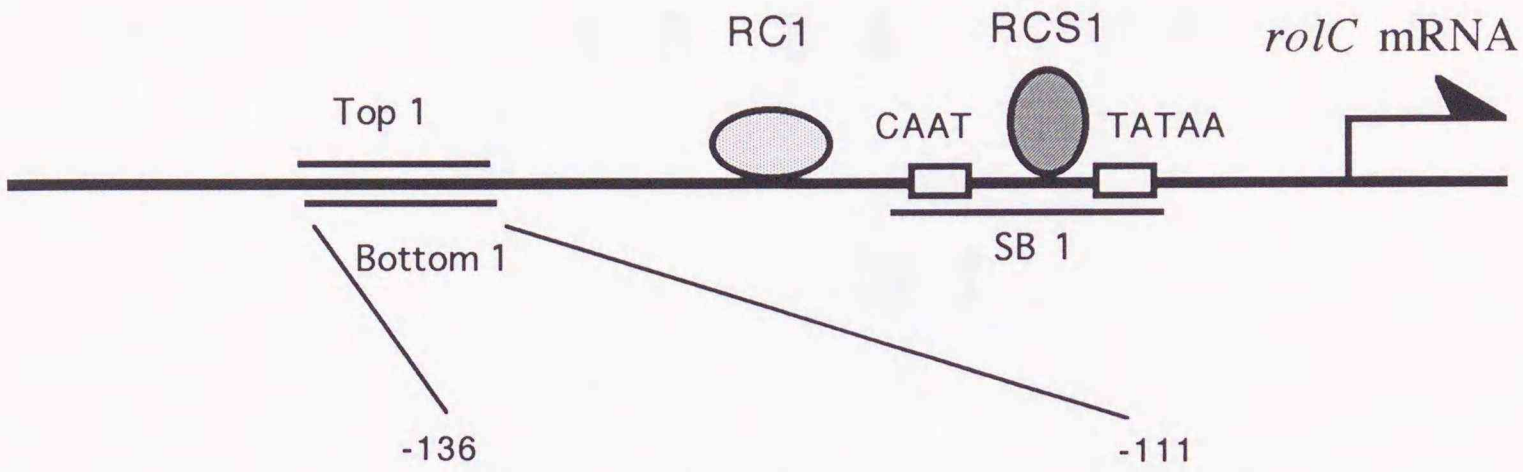
Discussion

I detected novel DNA binding protein RCS2 (*rolC* single-stranded DNA binding protein 2) in a nuclear extract of tobacco seedlings. The binding of RCS2 to single-stranded DNA is sequence specific. Previously we have identified a single-stranded DNA binding protein (RCS1) in the vicinity of TATAA box (Suzuki et al., 1992), but RCS1 and RCS2 have different sequence specificities. For binding to DNA, RCS2 (43 kDa) requires the presence of two direct repeat sequences (GCATC) within Top 1 DNA. Takase et al. (1991) also reported on single stranded DNA binding proteins interacting with the specific promoter region of the wheat histone H3 and H4 encoding gene. As suggested by Suzuki et al. (1992), short stretches of single stranded DNA are supposed to be transiently present in the genome.

Single-stranded DNA binding proteins have been considered to play an important role in DNA replication(Kuno et al., 1990), recombination (Edelmann et al., 1989). Although the majority of reported eukaryotic sequence-specific DNA binding proteins recognize native double-stranded sequences, other transcriptional factors possessing binding activity with single-stranded DNA (Rajavashisth et al., 1989; Kamada and Miwa, 1992). For instance, SRE-BF recognizes both single-stranded and double-stranded SRE-1 (*cis*-acting element necessary for sterol regulation), where the binding affinity of SRE-BF to single-stranded SRE-1 was higher than that to double-stranded SRE-1(Osborne et al., 1988). Such binding specificity of SRE-BF to single-stranded SRE-1 closely correlated

with in vivo ability of SRE-1 to direct sterol responsiveness of transcription.

Recently, we found that the rolC gene promoter was activated by sucrose in phloem tissues of young tobacco seedlings (Yokoyama et al., submitted for publication). Internal deletion studies revealed that the region between -134 and -94 contained a element necessary for the activation of the rolC gene promoter by sucrose. The Top 1 oligonucleotide, which interacted with RCS2, is contained in the sucrose responsive *cis*-DNA element of the *rolC* gene promoter, however, it is unclear whether these sequences play a crucial role for the expression of the *rolC* gene. Therefore, biological significance underlying the RCS2 and sucrose mediated *cis*-acting DNA element, remains to be investigated.



Top 1 5'-ctaGTGGAGCATCGACAGCCGCATCCATT-3'

Bottom 1 3'-ACCTCGTAGCTGTCGGCGTAGGTAAgatc-5'

-61 -29

SB 1 3'-CGAGGTTAAGCGATTCTCGAGACCTAATATTT-5'

Fig. 13. Schematic representation of the *rolC* gene promoter. Top 1 (-136 to -111, top strand), Bottom 1 (-135 to -111, bottom strand), and SB1 (-61 to -29, bottom strand) oligonucleotides were used for protein binding assay. The bent arrow represents the transcription start site (Kanaya et al., 1990). Numerals indicate the nucleotide distance from the transcription start site. TATA; the TATA box, CAAT; the CAAT box, RC1; (*rolC* gene DNA binding protein 1), RCS1; (*rolC* gene ssDNA binding protein 1).

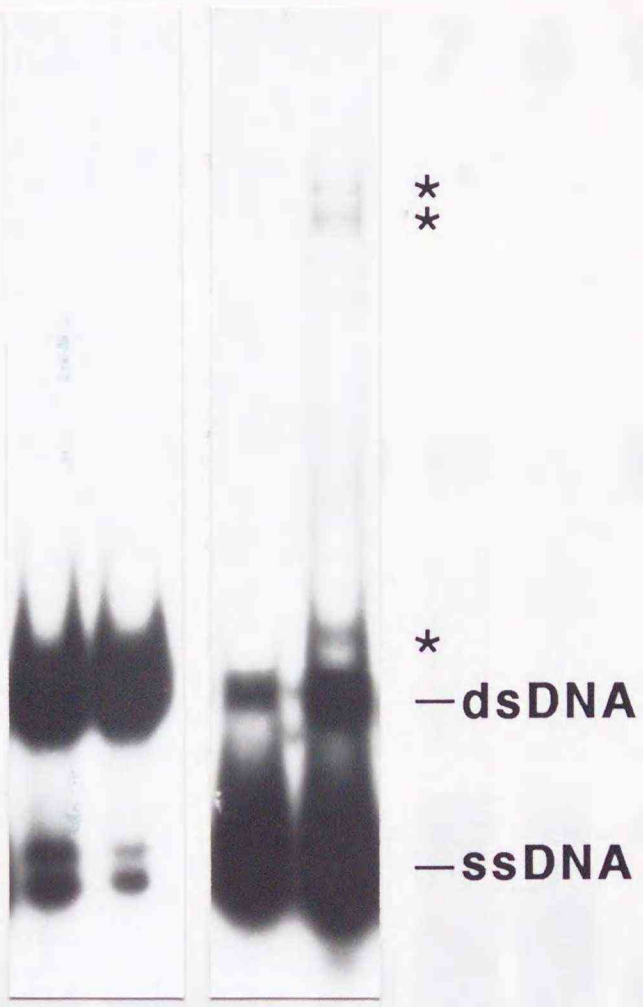


Fig. 14. Nuclear extracts from 10-day-old tobacco seedlings were used for binding analysis using native or heat denatured oligonucleotides. A probe DNA was labeled by filling in the 5' overhang after annealing Top I with Bottom 1. For heat denaturation, the ^{32}P-labeled oligonucleotides mixture was boiling for 2 minutes and then placed on ice. Lane 1, 3; no nuclear extract, lane 2, 4; nuclear extract was added. The complexes formed with DNA probe are indicated by asterisks.

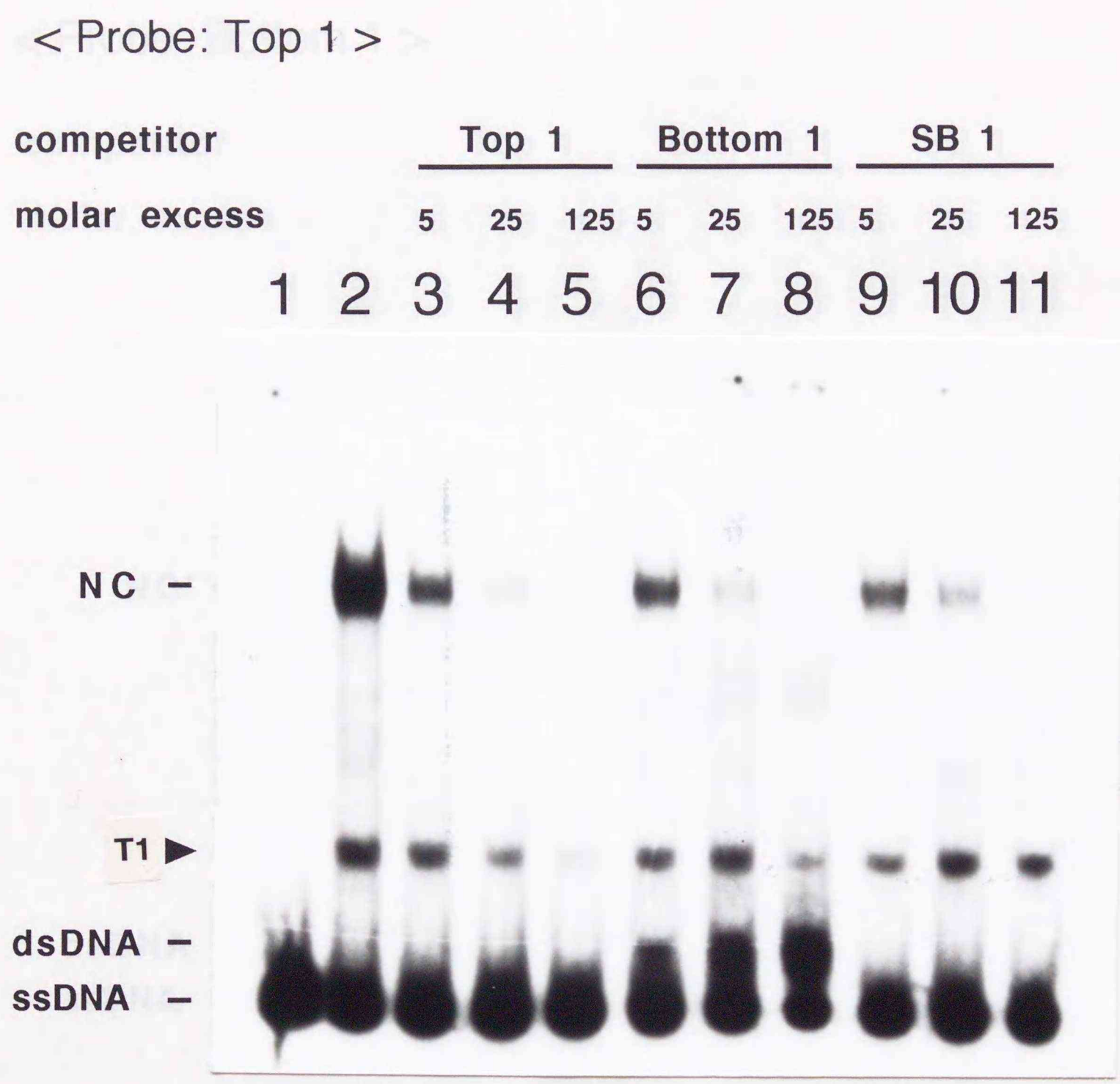
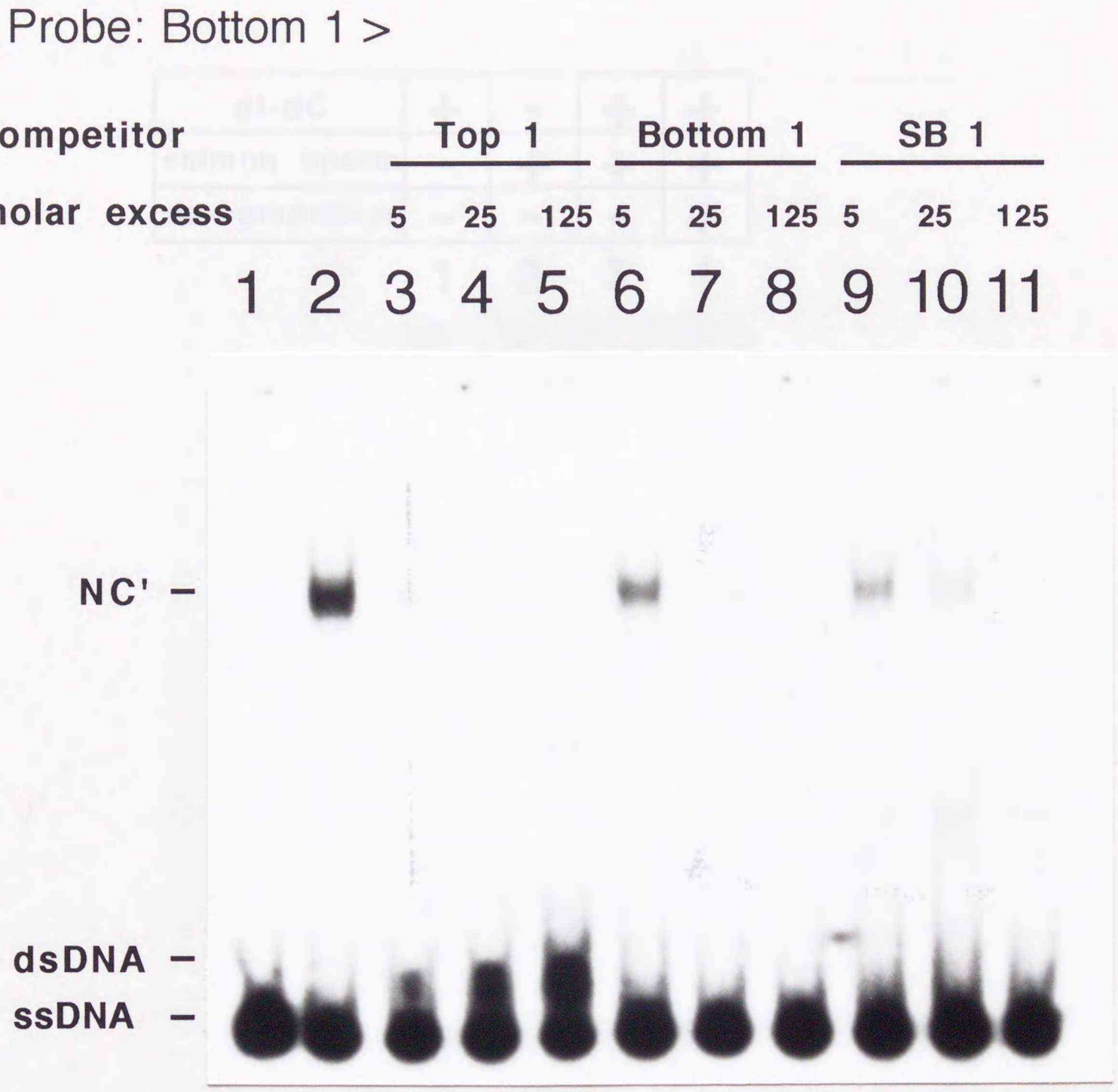


Fig. 15. Gel mobility shift assays showing strand specific interaction between nuclear protein of tobacco seedlings and ssDNA of the *rolC* gene.

(A) Gel mobility shift and competition assay with end-labeled Top 1 fragment.

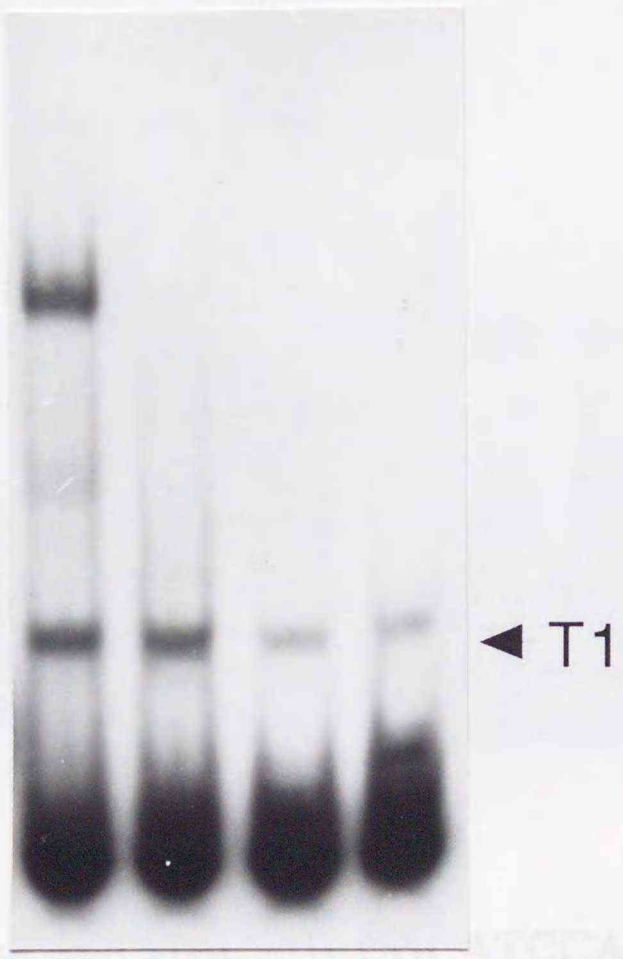
Competitions were performed with 5- to 125-fold molar excess of unlabeled Top 1 (lane 3-5), Bottom 1 (lane 6-8), and non-specific competitor SB1 (lane 9-11). Lane 1; no nuclear extract, lane 2; with nuclear extract (3 μg protein).

< Probe: Bottom 1 >

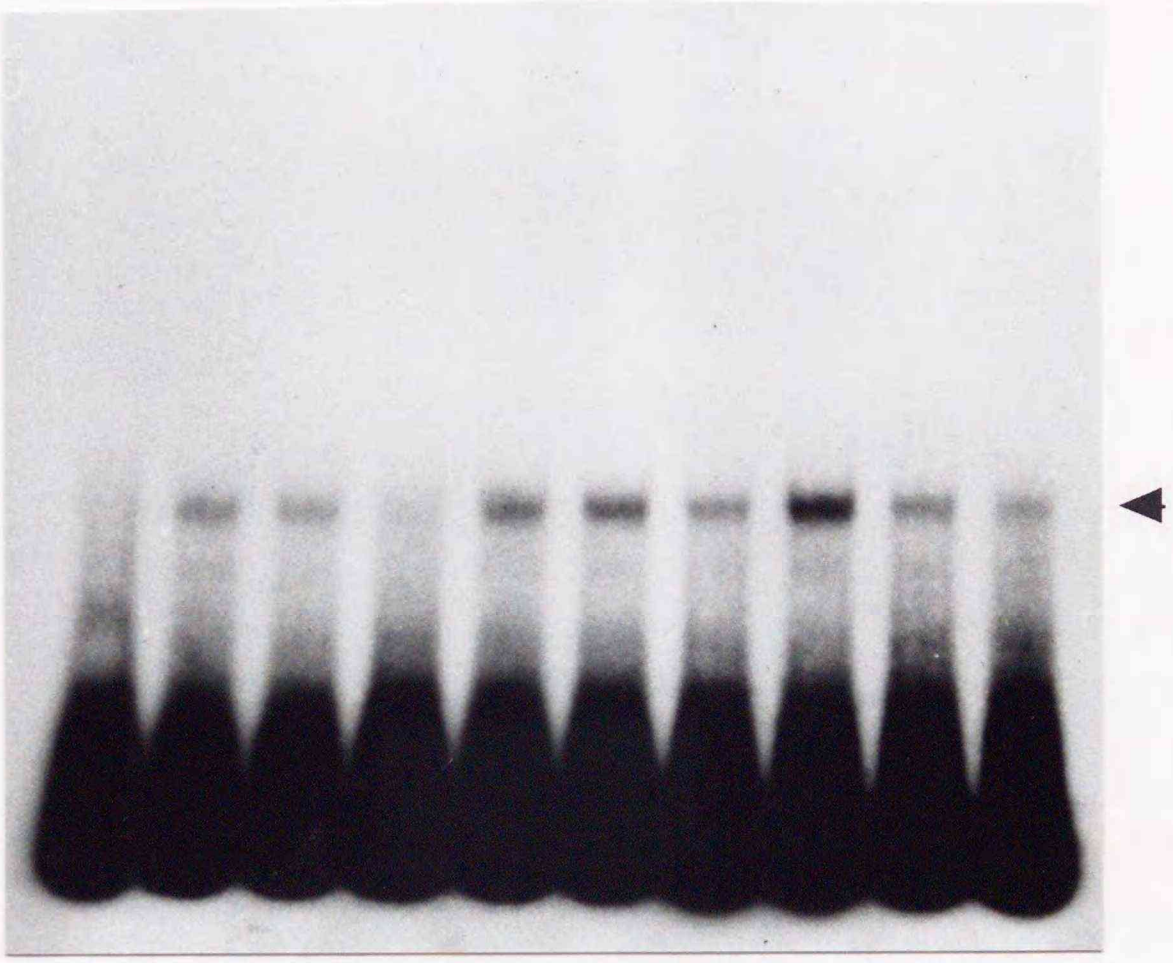


(B) Gel mobility shift and competition assays with end-labeled Bottom 1 fragment. Lane 1; no nuclear extract, lane 2; with nuclear extract (3 μg protein).

dI-dC	+	-	+	+
salmon sperm	-	+	+	+
ds-competitor	-	-	-	+
	1	2	3	4



(C) Binding of nuclear protein to the Top 1. ^{32}P-labeled Top 1 probe was incubated with the nuclear extract in the presence of poly[dI-dC]2 (lane 1), denatured salmon sperm DNA (lane 2), and both non-specific DNA (lane 3). Lane 4 is the binding reaction in the presence of poly[dI-dC], denatured salmon sperm DNA and 25 fold molar excess of unlabeled ds Top1-Bottom1 fragments.



M 1 5'-ctaGTGGAGAAGCGACAGCCGCATCCATT-3'

M 2 5'-ctaGTGGAGCATCGACAGCCGAAGCCATT-3'

M 3 5'-ctaGTGGAGAAGCGACAGCCGAAGCCATT-3'

Fig. 16. Direct repeats within Top1 are required for binding. Three mutated DNA fragments (M1-M3) were assayed for their ability to interact with the nuclear extract by gel mobility shift assay. The nt that were changed in mutant M1, M2 or M3 are indicated by the filled circles and outlined letters. Lane 1; no nuclear protein, lane 2; with nuclear protein, competitions were performed with 25 and 125-fold molar excess of unlabeled Top 1 (lane 3, 4), M1 (lane 5, 6), M2 (lane 7, 8), or M3 (lane 9, 10).

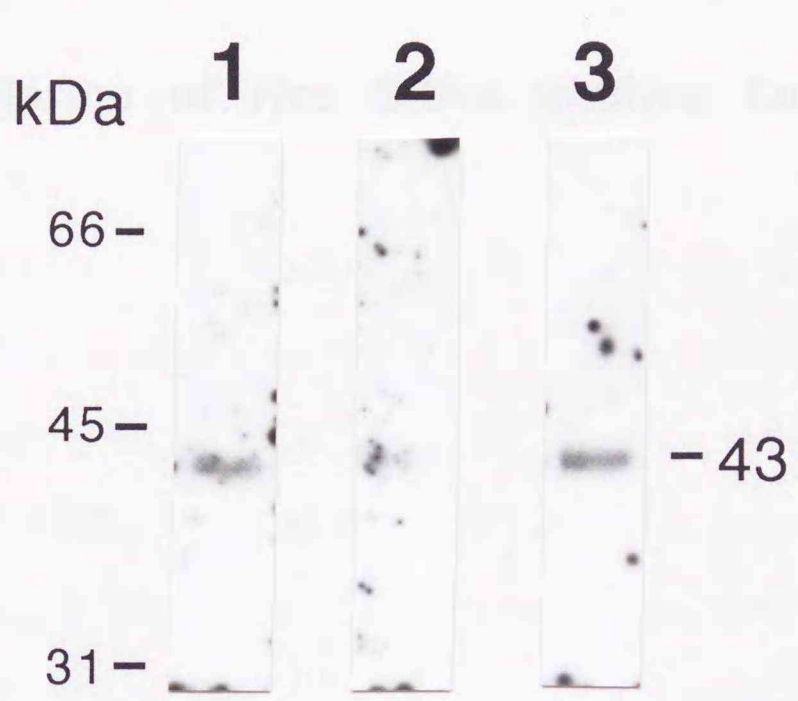


Fig. 17. South-Western blot analysis using ^{32}P-labeled Top 1 as a probe. The membranes were incubated separately with labeled Top I in the absence of competitor (lane 1), in the presence of a molar excess (25-fold) of the unlabeled Top I fragment (lane 2) and SB 1 fragment (lane 3). The molecular weights of the protein responsible for position signal is indicated to the right of the panel.

Part IV.

Isolation of rice G-box binding factor

Introduction

Gene expression is regulated in part by the interaction between sequence-specific transcription factors and *cis*-acting DNA elements. Recently, a large number of DNA binding motifs in many plant genes have been characterized. One of these motifs, the G-box, is defined by the palindromic sequence CCACGTGG. The G-box sequence has been found to be important for activity of several promoters regulating genes involved in photosynthesis (Donald and Cashmore, 1990; Donald et al., 1990; Schindler and Cashmore, 1990), chalcone synthase (Schulze-Lefert et al., 1989), ABA responsiveness (Guiltinan et al., 1990; Oeda et al., 1991), and alcohol dehydrogenase (DeLisle and Ferl, 1990). Recent data indicate that nuclear extracts from many plant species contain factors being able to interact G-box and related sequences containing the ACGT core (Giuliano et al., 1988; Staiger et al., 1989, 1991; McKendree et al., 1990). Indeed, several cDNA clones encoding nuclear proteins that specifically interact with G-box core sequence have been isolated. They are wheat Em binding protein EmBP-1 (Guiltinan et al., 1990), wheat histone binding protein HBP-1a (Tabata et al., 1991), tobacco transcription activator TAF-1 (Oeda et al., 1991), parsley common plant regulatory factors CPRF-1, CPRF-2, and CPRF-3 (Weisshaar et al., 1991), and *Arabidopsis* G-box binding factors GBF-1, GBF-2, and GBF-3 (Schindler et al., 1992). Interestingly all these proteins belong to the bZIP (basic/leucine zipper) family. bZIP motifs are characterized by a region of basic amino acids that required for specific protein-DNA interaction immediately N-terminal to a

leucine zipper that is characterized by several leucine residues regularly spaced at seven-amino acid intervals (Landschulz et al., 1988; Vinson et al., 1989). bZIP proteins interacted with G-box element and G-box like element contain proline rich regions in the N-terminus, which has been demonstrated to activate transcription (Schindler et al., 1992).

Three G-box like elements are found in the *rolC* gene 5' upstream region. To analyze whether these G-box like elements are important for the *rolC* gene transcription, I isolated a rice cDNA clone encoding GBF protein.

Materials and Methods

Bacterial protein extract

A culture of *E. coli* cells was incubated overnight, diluted 1:40 in Super broth medium [3.5 % polypeptone, 2 % yeast extract, 0.5 % NaCl, pH 7.5], and grown for 4 hr to early log phase. IPTG was added to the final concentration to 1 mM, and growth was continued for 2 hr. Cells were harvested by centrifugation, and the pellet was resuspended in Buffer A [40 mM Hepes-KOH pH 7.9, 1 mM EDTA, 1 mM PMSF, and 1 mM DTT]. The suspension was frozen in liquid nitrogen, then thawed rapidly. Lysozyme was added to 500 μg/ml and placed on ice for 15 min. Then, KCl was added to the final concentration to 1 mM, and incubate at 4 °C for 15 min. The lysate was centrifuged at 10,000 *g* for 15 min. The supernatant fraction was dialyzed against Buffer B [10 mM Tris-HCl (pH 8.0), 100 mM KCl, 1 mM EDTA, 1mM DTT, and 10 % glycerol]. After dialysis, the extract was centrifuged in a microfuge for 10 min to remove insoluble materials. The supernatant fraction was divided into aliquots, frozen in liquid nitrogen and stored at -80 °C.

DNA-protein gel shift assays

DNA-protein gel shift assays were performed as described in part II. *Ava*II/*Hin*fI fragment (RC275-92) and *Eco*RI/*Hin*fI fragment (RC392-273) of pUC12-C-P2 (Kanaya et al., 1990) are used as probe DNA.

Construction of cDNA library

Total RNA was prepared from seven days old roots of rice as described in part I. Poly(A)$^+$ RNA was prepared by two rounds of purification by oligotex-dT30 super (Roche) from total RNA. Oligo(dT)-primed cDNA for the phage λgt10 library was constructed using the Amersham cDNA synthesis kit according to the instructions of the manufacturer. cDNA ligated into λgt10 was packaged using Gigapack Gold II (Stratagene) according to the manufacturer's directions.

Screening of cDNA library

Screening of rice cDNA library was performed according to Sambrook et al. (1989). Thirty-seven plates with approximately 20,000 plaques were prepared and immobilized on nylon membrane. Filters were hybridized with CPRF3 (Weisshaar et al., 1991) in 5x SSPE (50 mM NaH_2PO_4 pH 7.4, 750 mM NaCl and 5 mM EDTA), 10x Denhardt's solution, 0.5 % SDS, and 40 μg/ml denatured salmon sperm DNA, and wash for 2 hr at 65 °C in 2x SSPE and 0.1 % SDS. Four plaques hybridizing to CPRF3 were purified and characterized.

Sequencing analysis

cDNA inserts of the λ clones were subclone into plasmid vector pIBI31. For sequencing, unidirectional deletions were obtained using exonuclease III and mung bean nuclease. These clones were sequenced from plasmid DNA by automated dideoxy chain termination on an ABI 373 (Applied Biosystems) according to the

manufacturer's protocol. Nucleotide and amino acid sequence analyze were performed using GeneWorks (IntelliGenetics). The data base search was done using the FASTA program (Pearson and Lipman, 1988).

Results

G-box binding proteins interact with the G-box like motifs of the *rolC* gene promoter

The G-box (CCACGTGG) is crucial for the expression of variety of plant genes. The G-box like elements, which differ from the G-box by 1 or 2 bp substitutions, are also found in other plant promoters. In the *rolC* promoter, three G-box like elements are found at -136 (G1;TAACGTGG), -219 (G2;TCACGTGG), and -364 (G3;CGACGTGG) (Fig. 18A).

Parsley protein CPRF2 and CPRF3 (Common Plant Regulatory Factor) are bZIP transcription factor, and bind specificity to the G-box and the G-box like elements. Whether CPRF2 and CPRF3 interact with the *rolC* G-box like elements, gel mobility shift assays were performed with *E. coli* extracts containing overexpressed CPRF2 and CPRF3 by using fragment containing G-box like sequences as probes.

Using RC392-273 fragment as probe, both CPRF2 and CPRF3 formed a complex (Fig. 18B, lane 2, 5). Binding of CPRF2 is specific to RC392-273 fragment (lane 3, 4), however DNA-protein complex of CPRF3 was competed out by either unlabeled RC392-273 or RC275-92 fragment (lane 6, 7). CPRF3 was also able to bind RC275-92 fragment (Fig. 18C, lane 2), the complex was competitively inhibited by both RC392-273 and RC275-92 but not by RC94+23 (lane 3 to 5). These results suggested that CPRF2 bind specifically to the *rolC* gene G-box like element G1 but CPRF3 are capable of interacting with either G1 or G3 and/or G2. Indeed, CPRF2 and

CPRF3 have different DNA-binding affinity to the G-box like motifs (Armstrong et al., 1992).

Isolation of rice GBF cDNA clone

To isolate rice cDNA encoding proteins that interact with G-box motif, the CPRF3 cDNA was used to screen a rice root cDNA library by DNA hybridization. Four positive clones were obtained from screening of about 750,000 recombinant phages. The cDNA inserts were subcloned into pIBI31 plasmid and sequenced. All of these cDNAs were shown to be derived from the same mRNA species.

The DNA and deduced amino acid sequences of the longest cDNA are shown in Fig. 19. The insert contains a partial cDNA of 1120 bp encoding an open reading frame of 288 amino acids, starting with a glutamine which I have tentatively designated as the first amino acid residue. The deduced protein sequence contains a leucine zipper motif that is composed of five leucine residues and one isoleucine residue regularly spaced at intervals of seven amino acids. In addition, a stretch of basic amino acids is located adjacent to the N-terminus of the leucine repeats. These two structural motifs, the basic domain and the leucine zipper, are characteristic features of the bZIP proteins. Furthermore, the encoded protein contains a small acidic region rich in proline residues at the N-terminus.

The protein encoded by the rice cDNA is 56 % identical to GMGBFA (*Glycine max* G-box binding factor, Hong et al., 1992) and 48 % identical to ATHGBF1 (*Arabidopsis thaliana* GBF1, Schindler et al., 1992). A particularly high degree of conservation is detected within bZIP sequences of other proteins interacted with G-box element (Fig. 20A, B). In addition to the bZIP domain, amino acid

sequence similarities within the N-terminal proline-rich regions were shown in other bZIP proteins (Fig. 20C). Because of the strong similarity in the amino acid sequence, I designated the protein encoded in this partial cDNA as RGBF1 (rice G-box binding factor 1).

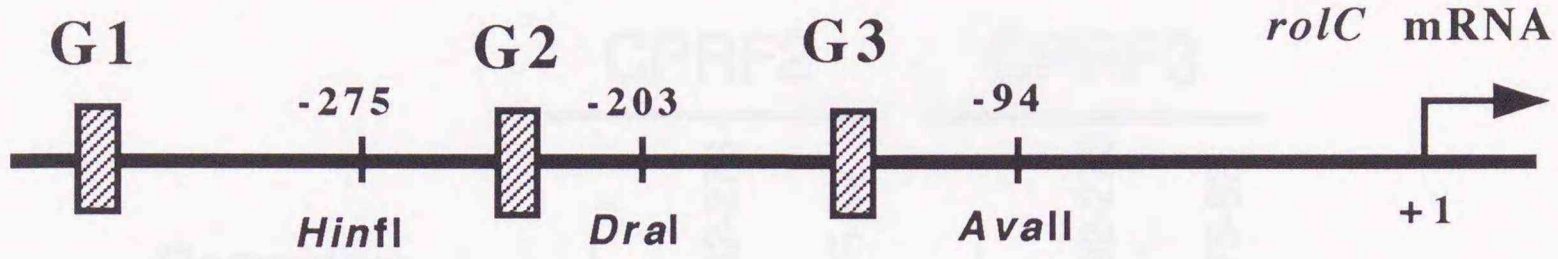
Discussion

Recently, several bZIP proteins interacted with the ACGT core sequence have been isolated from a variety of plant species. These bZIP proteins were classified into three groups by the specificity of DNA recognition sequences (Schindler et al., 1992; Izawa et al., 1993). Proteins interacting with G-box (CCACGTGG) sequence, which include the GBF1 (Schindler et al., 1992), HBP1a (Tabata et al., 1989, 1991), CPRF3 (Weisshaar et al., 1991), and EmBP1 (Guiltinan et al., 1990), belong to group I, and these binding proteins were only observed when the ACGT core was followed by GG or GT. Group III proteins, TGA1a (Katagiri et al., 1989) and HBP1b (Tabata et al., 1991), recognized TGACGT/C sequence. It is important for the binding that 5'-side nucleotides flanking the ACGT core are TG. On the other hand, group II including CPRF2 (Weisshaar et al., 1991) shows comparable binding affinity for both elements. In addition to the differences of recognition sequences, overall protein structures are distinct between group I and group III. Proteins of group I are characterized by a bZIP motif at the C terminus (Fig. 20A). In contrast, group III proteins contain an N-terminal DNA binding domain. Furthermore, the activation domain of group I protein is enriched in proline and located at the N terminus. On the other hand, group III proteins contain a C-terminal acidic or glutamine-rich region. On the basis of protein structure, rice protein RGBF1 belongs to group 1.

G-box and G-box like elements mediate transcriptional regulation by several stimuli, such as UV (Weisshaar et al., 1991), visible light (Schulze-Lefert et al., 1989; Block et al., 1990; Donald

and Cashmore, 1990), and ABA (Guiltinan et al., 1990; Oeda et al., 1991). GBFs can *trans*-activate the expression of the promoter containing G-box motif (Oeda et al., 1991; Schindler et al., 1992). The expression of the *rolC* gene may be regulate by GBFs, such as CPRF2, CPRF3, and RGBF1. Further analysis of the interaction between RGBF1 and G-box like sequence is needed.

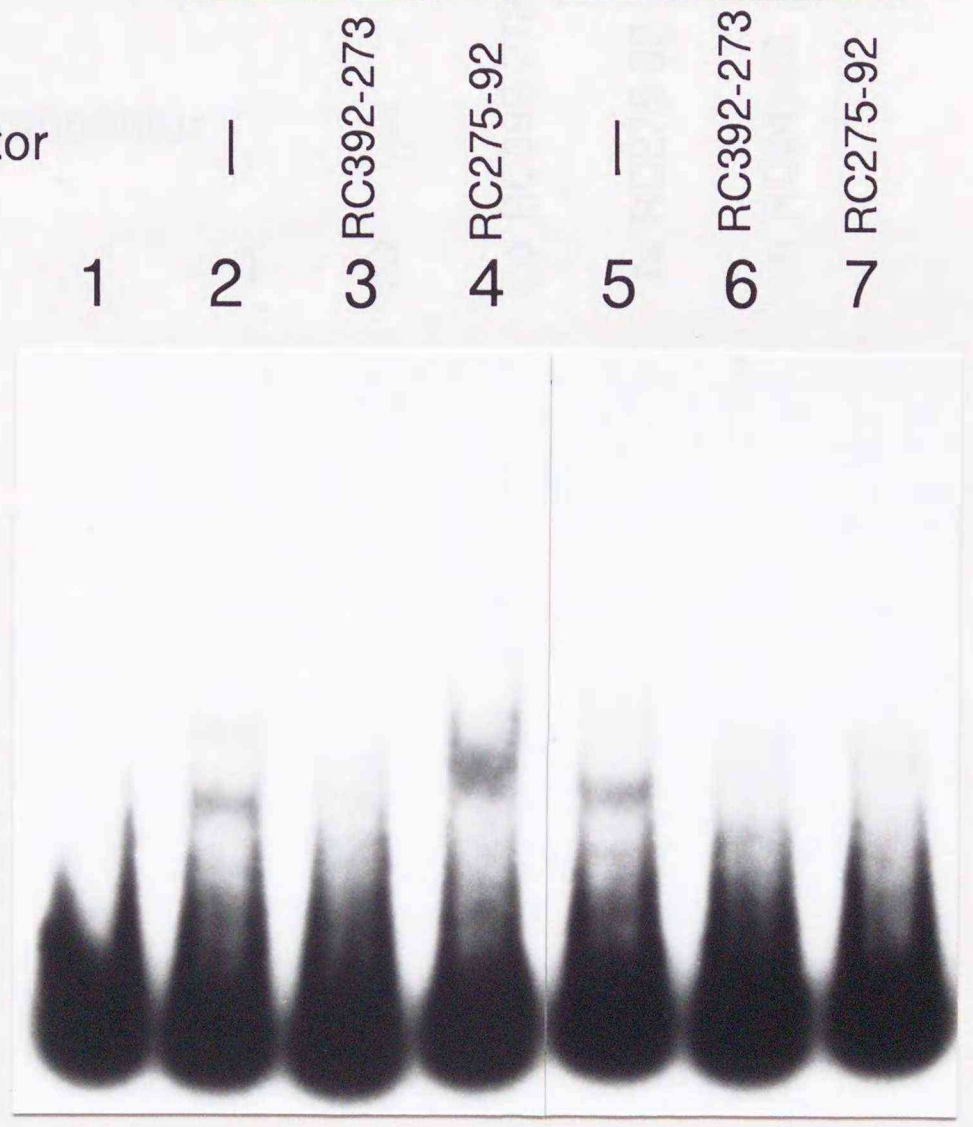
A



G-box		CCACGTGG
G1	-367	TAACGTGG -360
G2	-222	TCACGTGG -215
G3	-139	CGACGTGG -132

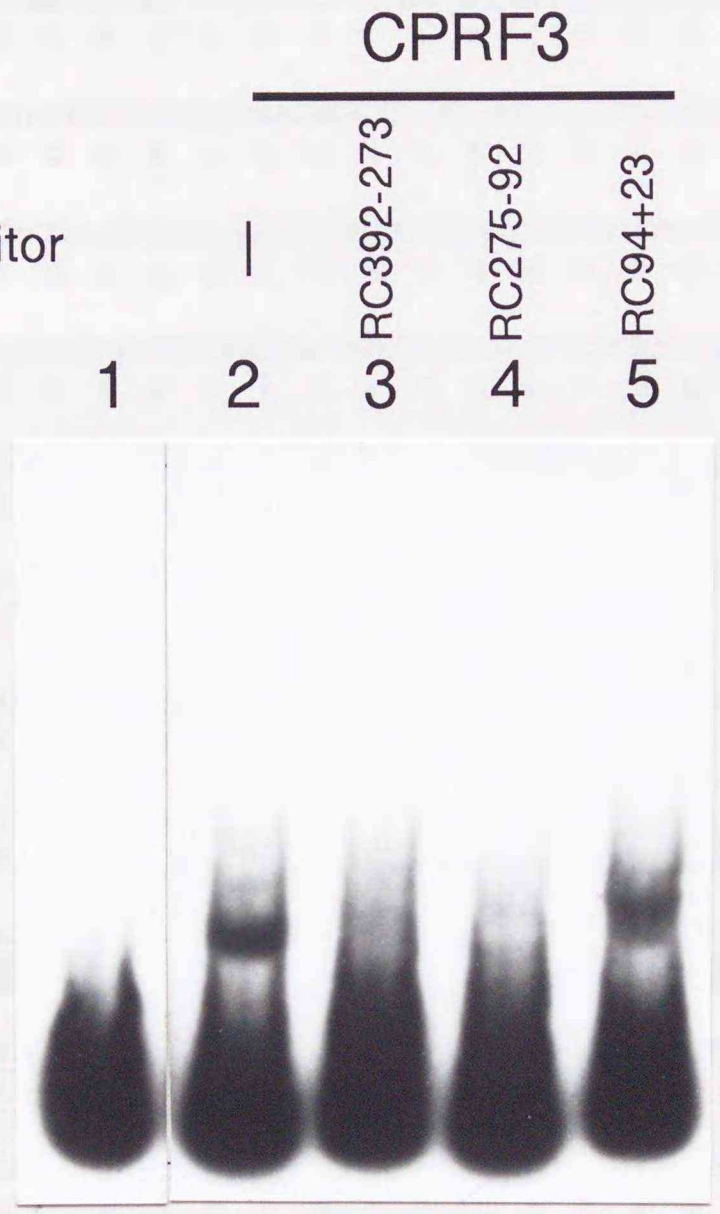
Fig. 18. The parsley CPRFs bind to DNA fragment containing the G-box like element of the *rolC* gene promoter.
(A) Schematic diagram of the 5'-upstream region of the *rolC* gene. The bent arrow represents the transcription start point. Numerals indicate the nucleotide distance from the transcription start point. The hatched boxes represent G-box like element.

Probe: RC392-273



(B) DNA mobility shift assay of RC392-273 fragment. *E. coli* extract containing overexpressed CPRF2 (lane 2 to 4) and CPRF3 (lane 5 to 7) was used in binding assays. Competitions were performed with 50 fold molar excess of unlabeled RC392-273 (lane 3, 6) and RC275-92 (lane 4, 7).

Probe: RC275-92

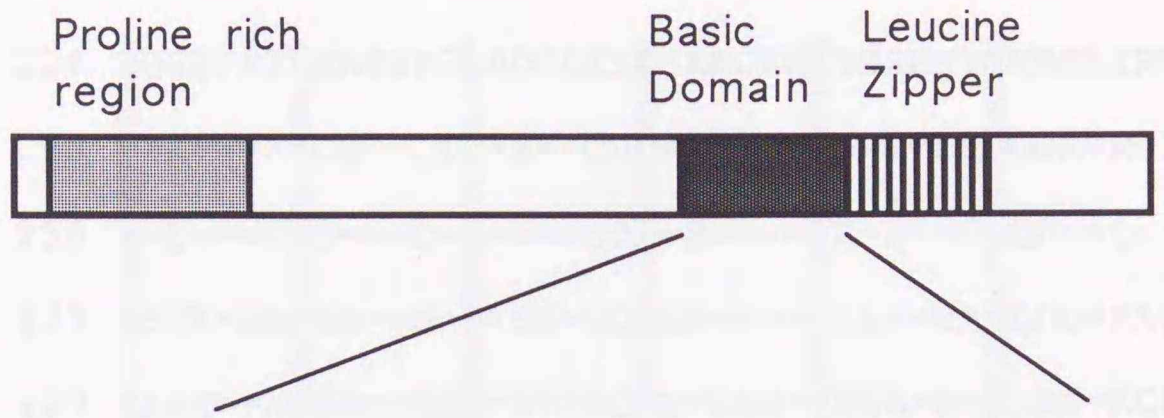


(C) Bacterial produced CPRF3 was used in a gel mobility assay with RC275-92 fragment. Lane 1; no extract, lane 2 *E. coli* extract was added. Unlabeled competitors were added in lane 3 (50-fold RC275-92), lane 4 (50-fold RC392-273), and lane 5 (50-fold RC94+23).

```
   1 CAGCATCCTTTAATGCCACCGTATGGTACTCCTATTCCATATCCAGTATATCCTCCAGGG     60
   1  Q  H  P  L  M  P  P  Y  G  T  P  I  P  Y  P  V  Y  P  P  G     20

  61 GGGATGTATGCTCATCCTAGTATTGCTACGAATCCTAGCATGGTACCAACCGCGGAATCA    120
  21  G  M  Y  A  H  P  S  I  A  T  N  P  S  M  V  P  T  A  E  S     40

 121 GAAGGAAAAGCAGTTGATGGGAAGGACAGAAATCCAACCAAAAAGTCGAAGGGAGCTTCA    180
  41  E  G  K  A  V  D  G  K  D  R  N  P  T  K  K  S  K  G  A  S     60

 181 GGAAATGCTAGTTCTGGTGGTGGTAAAGCTGGAGACAGTGGAAAGGCCNCCTCAAGTTCA    240
  61  G  N  A  S  S  G  G  G  K  A  G  D  S  G  K  A  X  S  S  S     80

 241 GGAAATGATGGCGGCACACAAAGTGCTGAAAGCGGAAGCGATGGATCATCGGATGGAGGA    300
  81  G  N  D  G  G  T  Q  S  A  E  S  G  S  D  G  S  S  D  G  G    100

 301 AGTGATGAGAATACTAACCATGAATTTTCAACTGGGAAGAAAGGAAGTTTTCATCAGATG    360
 101  S  D  E  N  T  N  H  E  F  S  T  G  K  K  G  S  F  H  Q  M    120

 361 CTCGCAGATGGAGCCAGTGCACAGAATACTGTAGCTGGTTCAGTCCCTGGGAATGCATTG    420
 121  L  A  D  G  A  S  A  Q  N  T  V  A  G  S  V  P  G  N  A  L    140

 421 GTTTCTGTACCCGCAGCTAATCTTAATATTGGAATGGACTTGTGGAATGCTTCTCCTGCT    480
 141  V  S  V  P  A  A  N  L  N  I  G  M  D  L  W  N  A  S  P  A    160

 481 GGAAATGGATCTCTAAAAGTGCGTCAAAATCCCTCCGCTGCTGTTGTACCAGGAACAGTG    540
 161  G  N  G  S  L  K  V  R  Q  N  P  S  A  A  V  V  P  G  T  V    180

 541 ATGGGTCGCGATGCGATGATGCCTGATCAGTGGGTCAATCAGGATGAACGTGAGCTGAAG    600
 181  M  G  R  D  A  M  M  P  D  Q  W  V  N  Q  D  E  R  E  L  K    200

 601 AGACAAAAGAGGAAGCAATCTAACAGAGAGTCTGCTCGGAGATCAAGATTACGGAAGCAG    660
 201  R  Q  K  R  K  Q  S  N  R  E  S  A  R  R  S  R  L  R  K  Q    220

 661 GCTGAGTGTGAAGAGCTACAGGGAAGGGTAGAGACGCTTAATAATGAAAACCGCAGTCTC    720
 221  A  E  C  E  E  L  Q  G  R  V  E  T  L  N  N  E  N  R  S  L    240

 721 AGAGATGAGCTGAAGAGNCTGTCTGAGGAGTGTGAGAAACTTACATCANNAAATAATACA    780
 241  R  D  E  L  K  X  L  S  E  E  C  E  K  L  T  S  X  N  N  T    260

 781 ATAAAGGAAGAGTTAATCAGAGTTTACGGTGCAGATGAGGTTTCTAAACTTGGTATTCAA    840
 261  I  K  E  E  L  I  R  V  Y  G  A  D  E  V  S  K  L  G  I  Q    280

 841 TCTTTCGATGACGGAGGTGACAGTTAACACATAAAATTTACTTGTTTTACTTCGTCTGGT    900
 281  S  F  D  D  G  G  D  S  *                                     290

 901 GAAGCAGNAACTTCTTTAAAGTTGATTGCAGANCCTATGGCTTAGATTTAGTTAGTAGGC    960
 961 AATGCTTGTCCTCACTAATACATGAACTTTTCAGTGTAAATTTGCCTTTACTACCTATGA   1020
1021 GATACTAACATGAGAGAAAGCTCTCAACTTATGCTTCATGTTGTNATAGCAATTAGACGA   1080
1081 GANTGAGAAGACGTCTCATTTTATAACCAAAAAAAAAAAA                       1120
```

Fig. 19. Sequence of the RGBF1 cDNA clone and the predicted amino acid sequence. Numbers on the both sides correspond to the base pair and amino acid positions. The proline residues within the N-terminal domains are underlined. The amino acids representing the basic region are highlighted, the leucine and isoleucine residues within leucine zipper domains are boxed.

A




```
RGBF1;      195 DERELKRQKRKQSNRESARRSRLRKQAECEE
SOYGBFA;    267 ======K=========================
ATGBF1;     219 ============================Q
ATGBF2;     246 N=K=V==E====================T=Q
ATGBF3;     198 N======ER====================T==
HBP-1a;     249 ======K====L====================
EmBP-1;     100 =======ER==================Q====
PCCPRF1;    269 ND=D===ER====================A==
PCCPRF2;    187 =PSDA==VR=ML==========R====HMT=
PCCPRF3;    193 ========R==================KSD=
TGA1a;       69 SKPVE=VLR=LAQ===A==K=====K=YVQQ
HBP-1B;      41 KNGDQ=TMR=LAQ===A==K=====K=YVQQ
```

Fig. 20. Similarity between RGBF1 and other plant bZIP proteins.
(A) Amino acid comparison of the basic region of RGBF1 and other plant bZIP proteins (Hong et al., 1992; Schindler et al., 1992; Tabata et al., 1989; Guiltinan et al., 1990; Weisshaar et al., 1991; Katagiri et al., 1989; Tabata et al., 1991). Numbers on the left refer to the amino acid positions within the individual proteins. Amino acids identical to RGBF1 are indicated by equal signs.

B

```
RGBF1;      226 LQGRVETLNNENRSLRDELKXLSEECEKLTSXNNTIKEELIRV
SOYGBFA;    298 ==K===S=GS==QT==E==QRV=========E=DS=QGRVRTI
ATGBF1;     250 ==Q===S=S===Q=======QR==S=====K=E==S=QD==Q==
ATGBF2;     277 =SVK=DA=VA==M===SK=GQ=NN=S===RLE=EA=LDQ=KAQ
ATGBF3;     229 =ARK==A=TA==MA==S==NQ=N=KSD==RGA=A=LLDK=KCS
HBP1A;      280 =GQ=AEA=KS==S===I==DRIKK=Y=E=L=K=TSL=AK=GES
EmBP-1;     131 =AQK=SE=TAA=GT==S==DQ=KKD=KTMETE=KKLMGKILSH
PCCPRF1;    300 =AIK=DS=TA==MA=KA=INR=TLTA=====ND=SRLL=VMKNA
PCCPRF2;    218 =ETQ=SQ=RV==S==LKR=TDI=QRYNDAAVD=RVL=ADIETM
PCCPRF3;    224 ==E=LDN=SK===I==KN=QRI==A=AEV==E=HS=====L=N
TGA1a;      100 =ENSKLK=IQLEQE=ERAR=QGMCVGGGVDASQLSYSGTASSG
HBP-1b;      72 =ENSRLK=TQLEQE=QRARQQGIFISSSADQSHSMSGNGALAF
```

C

```
RGBF1;      1 ------QHPLMPPYGTPIPYP-VYPPGGMYAHPSIA-TNPSMVPTAESEGKAVDGKDRNPTKKSK
                  |||| ||| || |||  |||| ||||| | |       |  ||  |||||   | ||
SOYGBF1;   66 SYMWGSQHPLIPPYSTPVPYPAIYPPGNVYAHPSMAMTLSTTQNGTEFVGKGSDGKDRVSAKSSK
                |||| ||   ||| ||||||| |||| |||||||                    ||  | | ||
ATGBF1;    59 PYMWGAQHHMMPPYGTPVPYPAMYPPGAVYAHPSMP-MPPNSGPT-----NKEPAKDQASGKKSK
                |||| ||    ||| |  |||   ||    ||  |   ||  ||       |        | |
CPRF3;     60 PYMWRNQHRFILPYGIPMQYPALFLPGGIFTHPIVP-TDPNLAPT-----SGEVGRKISDEKGRT
                |||| ||    ||| |       ||    ||  |   |   |       |
HBP1a;     62 PYMWGPQH-MVPPYGTPPPPYMMYPPGTVYAHPTAPGVHPFHYPM---QTNGNLEPAGAQGAAPG
```

(B) Sequence comparison of the proposed leucine zipper region of RGBF1 with other plant bZIP proteins. The leucine and isoleucine residues within the leucine zipper are highlighted.
(C) Amino acid similarities within the N-terminal proline-rich regions of RGBF1, SOYGBF1, ATHGBF1, CPRF3, and HBP1a. Vertical bars indicate identical amino acids. Numbers refer to the amino acid positions within the individual proteins. Gaps (-) were introduced to optimize the alignment.

General Conclusion

The present study was conducted to elucidate the transcriptional regulation of the *rolC* gene of *Agrobacterium rhizogenes* in the higher plants. Though the *rolC* gene is originated from a bacterial plasmid, the promoter structure of the *rolC* gene was the class II gene type of eucaryotes. In this thesis, it is described that the *rolC* gene promoter was expressed in monocotyledonous rice plants, which is not infected with *Agrobacterium rhizogenes* in the natural world. The expression in transgenic rice was located in the phloem cells. The phloem cell-specific expression of the *rolC* gene was also detected in tobacco, carrot and *Arabidopsis*. These results suggest that the *rolC* gene expression is regulated by common factors that are present in higher plants.

Transcriptional regulation of gene expression in eukaryotic cells is mediated by the concerted action of transcriptional factors that interact with regulatory elements residing in the promoter regions. Promoter analysis using GUS reporter gene showed that the *rolC* gene 5'-upstream region contained several transcriptional regulatory elements (Fig. 21). Analysis of the *rolC* 5'-deletion mutants and internal mutants revealed that at least two positive *cis*-acting elements were located in the region from -848 bp to -255 bp and from -196 bp to -135 bp (Sugaya and Uchimiya, 1992; Fujii et al., in press; Yokoyama et al., submitted for publication). In a related study, an element necessary for the activation by sucrose was found in the region between -134 bp and -94 bp. Furthermore,

analysis using heterologous fusion promoters consisting of the CaMV 35S core promoter region and several DNA fragment of the *rolC* promoter revealed that the region from -94 bp to +27 bp was important for phloem cell specific expression. In cell culture of carrot, the expression of the *rolC* gene was activated during somatic embryogenesis (Fujii and Uchimiya, 1990). It was suggested that the combination of three regulatory regions (element I; -848 bp to -255 bp, element II; -255 bp to -94 bp, and element III; -94 bp to +23 bp) is required for the somatic embryogenesis-related activation (Fujii et al., in press).

In this study, I detected three DNA-binding factors interacted with the *rolC* 5'-upstream region. RC1 bound to AT-1 box (ATATTTTTAT) located from -76 bp to -67 bp. RC1 exists not only in rice leaf and callus but also in tobacco hairy root, carrot culture cells, and wheat germ (Kanaya et al., 1990, 1991). RC2 is interacted with an AT-rich -203 bp to -164 bp region. The binding to this AT-rich region is seen in leaf nuclear proteins, but not in callus nuclear proteins. RCS2 (43 kDa protein) is single stranded DNA binding protein which bound specifically to the top strand of -136 bp to -111 bp region. The DNA binding activity of RCS2 required the presence of a direct repeat sequence (GCATC). In addition, Suzuki et al. (1992) have identified another single stranded DNA binding protein RCS1 showing specific binding to the region between the CAAT and TATAA box. These factors were shown to recognize part of regulatory regions identified by *cis*-analysis, but no information is available whether these factors play a crucial role for the expression of the *rolC* gene. Further investigations are needed to characterize the function of these DNA binding proteins.

In the *rolC* gene promoter, several sequences conserved in other plant genes were found. G-box or G-box like motifs have been found in several plant gene promoters and shown to be crucial for their expression. The *rolC* promoter contains three G-box like elements, which may control transcriptional activity. In this study, I isolated rice G-box binding factor (RGBF1) by DNA hybridization. Structure of RGBF1 is similar to the bZIP class of DNA-binding protein. RGBF1, like other GBFs, may regulate the activity of the promoter containing G-box motif. Further analysis of GBFs may facilitate an understanding at molecular level of the regulatory mechanism of the *rolC* gene.

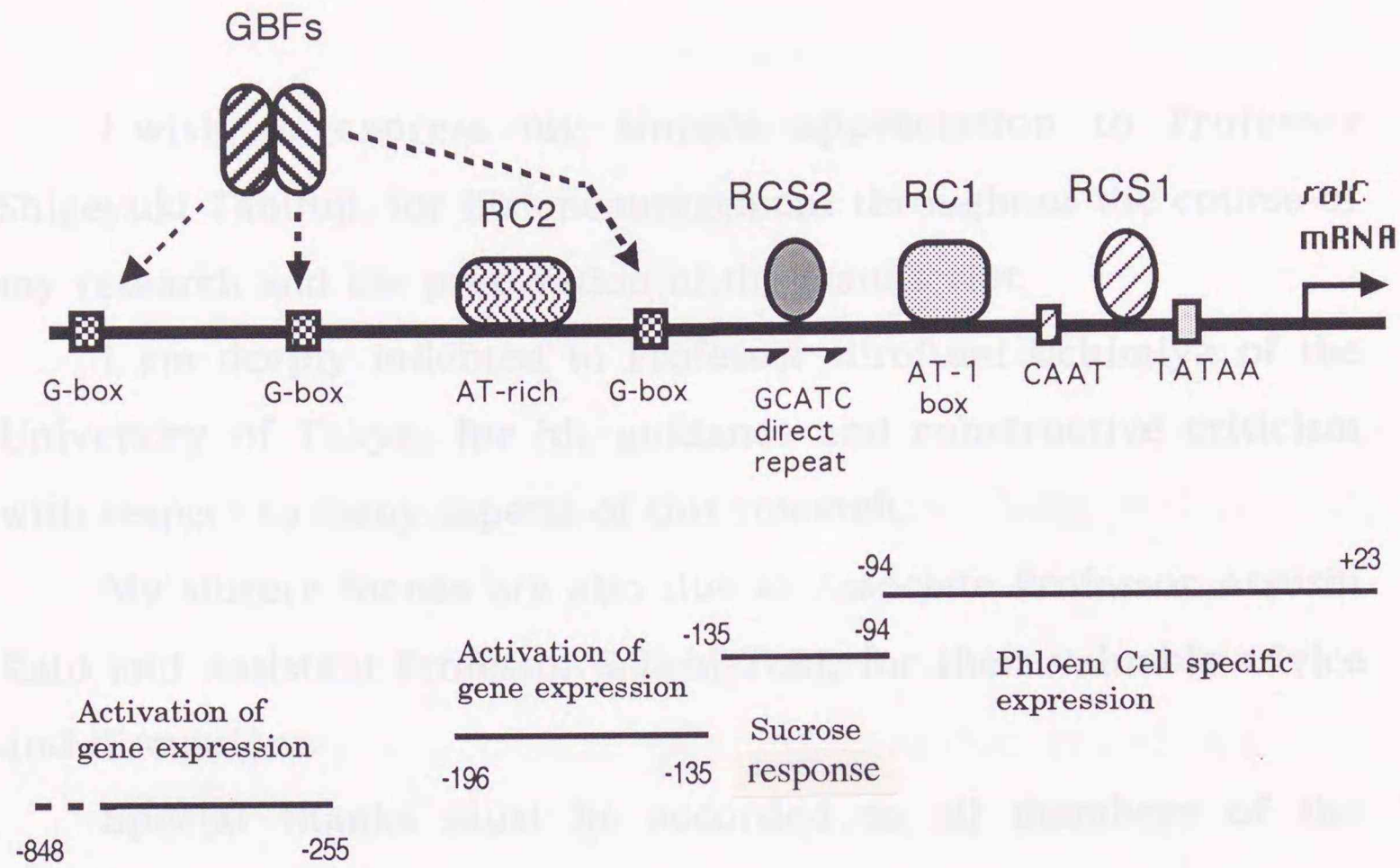


Fig. 21. Schematic representation of the *rolC* gene promoter.
The bent arrow represents the transcription start site. Numerals indicate the distance from the transcription start site. RC1 and RC2; *rolC* gene DNA binding protein, RCS1 and RCS2; *rolC* single strand DNA binding protein), GBFs; G-box binding factors. TATAA; TATAA box, CAAT; CAAT box.

Acknowledgments

I wish to express my sincere appreciation to Professor Shigeyuki Tanifuji, for his encouragement throughout the course of my research and the preparation of the manuscript.

I am deeply indebted to Professor Hirofumi Uchimiya of the University of Tokyo, for his guidance and constructive criticism with respect to many aspects of this research.

My sincere thanks are also due to Associate Professor Atsushi Kato and Assistant Professor Seiichi Toki, for their valuable advice and discussions.

Special thanks must be accorded to all members of the laboratory for useful discussions on the results.

References

Araki, K., Maeda, H., Wang, J., Kitamura, D. and Watanabe, T.: Purification of a nuclear *trans*-acting factor involved in the regulated transcription of a human immunoglobulin heavy chain gene. Cell 53 (1988) 723-730.

Armstrong, G.A., Weisshaar, B. and Hahlbrock, K.: Homodimeric and heterodimeric leucine zipper proteins and nuclear factors from parsley recognize diverse promoter elements with ACGT cores. Plant Cell 4 (1992) 525-537.

Battraw, M. and Hall, T.: Histochemical analysis of CaMV 35S promoter-b-glucuronidase gene expression in transgenic rice plants. Plant Mol. Biol. 15 (1990) 527-538

Benfey, P.N. and Chua, N.H.: Regulated genes in transgenic plants. Science 244 (1989) 174-181.

Bevan, M.: Binary *Agrobacterium* vectors for plant transformation. Nucleic Acids Res. 12 (1984) 8711-8721

Bevan, M., Barners, W.M., and Chilton, M.: Structure and transcription of the nopaline synthase gene region of T-DNA. Nucl. Acids Res. 1 (1983) 369-385.

Block, A., Dangle, J.L., Hahlbrock, K. and Schulze-Lefert, P.: Functional borders, genetic fine structure, and distance requirements of cis elements mediating light responsiveness of the parsley chalcone synthase promoter. Proc. Natl. Acad. Sci. USA 87 (1990) 5387-5391.

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

Bustos, M.M., Guiltinan, M.J., Jordano, J., Begum, D., Kalman, F.A. and Hall, T.C.: Regulation of b-glucuronidase expression in transgenic tobacco plants by an A/T rich, *cis*-acting sequence found upstream of a French bean b-phaseolin gene. Plant Cell 1 (1989) 839-853

Capone, I., Cardarelli, M., Trovato, M. and Costatino, P.: Upstream non-coding region which confers polar expression to Ri plasmid root inducing gene *rolB*. Mol. Gen. Genet. 216 (1989) 239-244

Cardarelli, M., Mariotti, D., Pomponi, M., Spanò, L., Capone, I. and Costantino, P.: *Agrobacterium rhizogenes* T-DNA genes capable of inducing hairy root phenotype. Mol. Gen. Genet. 209 (1987) 475-480.

Chirgwin, J., Przybyla, A., MacDonald, J. and Rutter, W.: Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18 (1979) 5294-5299

Christensen, A.H., Sharrock, R.A., and Quail, P.H.: Maize polyubiquitin genes; structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Mol. Biol. 18 (1992) 675-689.

Chu, C.C., Wang, C.C., Sun, C.S., Hsu, C., Yin, K.C., Chu, C.Y. and Bi, F.Y.: Establishment of an efficient medium for anther culture of rice through comparative experiments on the nitrogen sources. Sci. Sin. 18 (1975) 659-668

Datta, N. and Cashmore, A.R.: Binding of a pea nuclear protein to promoters of certain photoregulated genes is modulated by phosphorylation. Plant Cell 1 (1989) 1069-1077

David, C., Chilton, M.D. and Tempe, J.: Conservation of T-DNA in plants regenerated from hairy root cultures. Bio/Technology 2 (1984) 73-76.

DeLisle, A.J. and Ferl, R.J.: Characterization of the *Arabidopsis Adh* G-box binding factor. Plant Cell 2 (1990) 547-557.

De Paolis, A., Mauro, M.L., Pomponi, M., Cardarelli, M., Spano, L. and Costantino, P.: Localization of agropine synthesizing function in the TR region of the root-inducing plasmid of *Agrobacterium rhizogenes* 1855. Plasmid 13 (1985) 1-7

Dignam, J, D., Lebovitz, R.M. and Roeder, R.G.: Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucl. Acid Res. 11 (1983) 1475-1489.

Donald, R.G.K. and Cashmore, A.R.: Mutation of either G box or I box sequences profoundly affects expression from the *Arabidopsis rbcS-1A* promoter. EMBO J. 9 (1990) 1717-1726.

Donald, R.G.K., Schindler, U., Batschauer, A. and Cashmore, A.R.: The plant G-box promoter sequence activates transcription in *Saccharomyces cerevisiae* and is bound in vitro by a yeast activity similar to GBF, the plant G box binding factor. EMBO J. 9 (1990) 1127-1735.

Dynan, W.S. and Tjian, R.: Control of eukaryotic messenger RNA synthesis by sequence-specific DNA binding proteins. Nature 316 (1985) 774-778

Dynan, W.S.: Modularity in promoters and enhancers. Cell 58 (1989) 1-4

Edelmann, W., Kröger, B., Goller, M. and Horak, I.: A recombination hotspot in the LTR of a mouse retrotransposon identified in an in vitro system. Cell 57 (1989) 937-946.

Edwards, K., Johnstone, C. and Thompson, C.: A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucl. Acids Res. 19 (1991) 1349.

Estruch, J.J., Chiriqui, D., Grossmann, K., Schell, J. and Spena, A.: The plant oncogene *rolC* is responsible for the release of cytokinins from glucoside conjugates. EMBO J. 10 (1991) 2889-2895.

Fujii, N. and Uchimiya, H.: Condition favorable for the somatic embryogenesis in carrot cell culture enhance expression in the *rolC* promoter-GUS fusion gene. Plant Physiol. 95 (1991) 238-241.

Fujii, N., Yokoyama, R. and Uchimiya, H.: Analysis of the *rolC* promoter region involved in somatic embryogenesis-related activation in carrot cell cultures. Plant Physiol. (1994) in press.

Gamborg, O.L., Miller, R.A. and Ojima, K.: Nutrient requirements of suspension cultures of soybean root cells. Exp. Cell. Res. 50 (1968) 151-158

Gardner, R.C., Howarth, A.J., Hahn, P., Brown-Luedi, M., Shepherd, R.J. and Messing, J.: The complete nucleotide sequence of an infectious clone of cauliflower mosaic virus by M13mp7 shotgun sequencing. Nucl. Acid Res. 9 (1981) 2871-2888.

Giuliano, G., Pichersky, E., Malik, V.S., Timko, M.P., Scolnik, P.A. and Cashmore, A.R.: An evolutionary conserved protein binding sequence upstream of a plant light-regulated gene. Proc. Natl. Acad. Sci. USA 85 (1988) 7089-7093.

Gordon-Kamm, W., Spencer, M., Mangano, M., Adams, T., Daines, R., Start, W., O'Brien, J., Chambers, S., Adams, W., Willetts, N., Rice, T., Mackey, C., Krueger, R., Kausch, A. and Lemaux, P.: Transformation of maize cells and regeneration of fertile transgenic plants. Plant Cell 2 (1990) 603-618

Gritz, L. and Davies, J.: Plasmid-encoded hygromycin-B resistance - the sequence of hygromycin-B phosphotransferase gene and its expression in *Escherichia coli* and *Saccharomyces cerevisiae*. Gene 25 (1983) 179-188.

Guiltinan, M.J., Marcotte, W.R. and Quatrano, R.S.: A plant leucine zipper protein that recognizes an abscisic acid response element. Science 250 (1990) 267-271.

Hatzopoulos, P., Franz, G., Choy, L. and Sung, R.Z.: Interaction of nuclear factors with upstream sequences of a lipid body membrane protein gene from carrot. Plant Cell 2 (1990) 457-467

Hong, J.C., Cheong, Y.H., Nagao, R.T., Bank, J.D., Key, J.L. and Cho, M.J.: Soybean G-box binding factor: The identification of a negative regulator of transcription activation by GBF proteins. EMBL/GenBank/DDBJ data base. Accession No. L01447 (1993).

Huffman, G.A., White, F.F., Gordon, M.P. and Nester, E.W.: Hairy-root-inducing plasmid: physical map and homology to tumor inducing plasmids. J. Bacteriol. 157 (1984) 269-276.

Izawa, T., Foster, R. and Chua, N.H.: Plant bZIP protein DNA binding specificity. J. Mol. Biol. 230 (1993) 1131-1144.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.: GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. 6 (1987) 3901-3907

Jensen, E.O., Marcker, K.A., Schell, J. and Bruijn, F.J.: Interaction of a nodule specific, *trans*-acting factor with distinct DNA elements in the soybean leghaemoglobin *lbc3* 5' upstream region. EMBO J. 7 (1988) 1265-1271

Jones, N.C., Rigby, P.W.J. and Ziff, E.B.: *Trans*-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. Genes Dev. 2 (1988) 267-281.

Jordano, J., Almoguera, C. and Thomas, T.L.: A sunflower helianthinin gene upstream sequence ensemble contains an enhancer and sites of nuclear protein interaction. Plant Cell 1 (1989) 855-866.

Jouanin, L: Restriction map of an agropine-type Ri plasmid and its homologies with Ti plasmids. Plasmid 12 (1984) 91-102

Kamada, S. and Miwa, T.: A protein binding to CArG box motifs and to single-stranded DNA functions as a transcriptional repressor. Gene 119 (1992) 229-236.

Kanaya, K., Tabata, T., Iwabuchi, M. and Uchimiya, H.: Specific binding of nuclear protein from tobacco hairy roots cultured in vitro to a 5'-upstream region of the *rolC* gene of the Ri plasmid. Plant Cell Physiol. 31 (1990) 941-946.

Kanaya, K., Hayakawa, K. and Uchimiya, H.: In vitro binding of wheat-germ proteins to the 5'-upstream region of the *rolC* gene of Ri plasmid. Plant Cell Physiol. 32 (1991) 295-297.

Katagiri, F., Lam, E. and Chua, N.H.: Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. Nature 340 (1989) 727-730

Keith, B. and Chua, N.H.: Monocot and dicot pre-mRNAs are processed with different efficiencies in transgenic tobacco. EMBO J. 5 (1989) 2419-2425

Keller, J., Shanklin, J., Vierstra, R. and Hershey, H.: Expression of a functional monocotyledonous phytochrome in transgenic tobacco. EMBO J. 8 (1989) 1005-1012

Kuno, K., Murakami, S. and Kuno, S.: Single-strand-binding factor(s) which interact with ARS1 of *Saccharomyces cerevisiae*. Gene 95 (1990) 73-77.

Landschulz, W.H., Johnson, P.F. and McKnight, S.L.: The leucine zipper: A hypothetical structure common to a new class of DNA-binding proteins. Science 240 (1988) 1759-1764.

Leach, F. and Aoyagi, K.: Promoter analysis of the highly expressed *rolC* and *rolD* root-inducing genes of *Agrobacterium rhizogenes*: enhancer and tissue-specific DNA determinants are dissociated. Plant Sci. 79 (1991) 69-46.

Lewin, B.: Commitment and activation at Pol II promoters: a tail of protein-protein interactions. Cell 61 (1990) 1161-1164

Laemmli, K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227 (1970) 680-685.

Maniatis, T., Goodbourn, S. and Fischer, J.A.: Regulation of inducible and tissue-specific gene expression. Science 236 (1987) 1235-1245

Matsuki, R., Onodera, H., Yamauchi, T. and Uchimiya, H.: Tissue-specific expression of the *rolC* promoter of the Ri plasmid in transgenic rice plants. Mol. Gen. Genet. 220 (1989) 12-16.

Maxam, A. and Gilbert, W.: Sequencing end-labeled DNA with base-specific chemical cleavages. Methods in Enzymol. 65 (1980) 499-560.

McKendree, W., Paul, A-L., DeLisle, A, and Ferl, R.: In vivo and in vitro characterization of protein interactions with the dyad G-box of the *Arabidopsis Adh* gene. Plant Cell 2 (1990) 207-214.

Mitchell, P.J. and Tjian, R.: Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science 245 (1989) 371-378

Moore, L., Warner, G. and Strobel, G: Involvement of plasmid in the hairy root disease of plants caused by *Agrobacterium rhizogenes*. Plasmid 2 (1979) 617-626.

Murakami, T., Anzai, H., Imai, S., Satoh, A., Nagaoka, K. and Thompson, C.J.: The bialaphos biosynthetic genes of *Streptomyces hygroscopicus*: molecular cloning and characterization of the gene cluster. Mol. Gen. Genet. 205 (1986) 42-50.

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

Nakamura, T., Handa, T., Oono, Y. Kanaya, K., Michikawa, M. and Uchimiya, H.: Organ-specific mRNA in transgenic tobacco plants possessing T-DNA of Ri plasmids. Plant Sci. 15 (1988) 213-218.

Nilsson, O., Moritz, T., Imbault, N., Sandberg, G. and Olsson, O.: Hormonal characterization of transgenic tobacco plants expressing the *rolC* gene of *Agrobacterium rhizogenes* T_L-DNA. Plant Physiol. 102 (1993) 363-371.

Oeda, K., Salinas, J. and Chua, N.H.: A tobacco bZIP transcription activator (TAF-1) binds to a G-box-like motif conserved in plant genes. EMBO J. 10 (1991) 1793-1802.

Oono, Y., Handa, T., Kanaya, K. and Uchimiya, H.: The TL-DNA gene of Ri plasmids responsible for dwarfness of tobacco plants. Jpn. J. Genet. 62 (1987) 501-505.

Oono, Y., Kanaya, K. and Uchimiya, H.: Early flowering in transgenic tobacco plants possessing the *rolC* gene of *Agrobacterium rhizogenes* Ri plasmid. Jpn. J. Genet. 65 (1990) 7-16.

Osborne, T.F., Gil, G., Goldstein, J.L. and Brown, M.S.: Operator constitutive mutation of 3-hydroxy-3-methylglutaryl coenzyme A reductase promoter abolishes protein binding to sterol regulatory element. J. Biol. Chem. 263 (1988) 3380-3387.

Pasqua, G., Monacelli, B., Altamura, M.M. and Calzecchi-Onesti, B.: Transformed phenotype and *in vitro* flower neoformation in tobacco hairy root regenerants. J. Plant Physiol. 130 (1987) 221-231.

Pearson, W.R. and Lipman, D.J. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85 (1988) 2444-2448.

Ptashne, M.: How eukaryotic transcriptional activators work. Nature 335 (1988) 683-689

Rajavashisth, T.B., Taylor, A.K., Andalibi, A., Svenson, K.L. and Lusis, A.J.: Identification of a zinc finger protein that binds to the sterol regulatory element. Science 245 (1989) 640-643.

Sambrook, J., Fritsch, E.F. and Maniatis, T.: Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory) (1989).

Schell, J.: Transgenic plants as tools to study the molecular organization of plant genes. Science 237 (1987) 1176-1183

Schernthaner, J., Matzke, M. and Matzke, A.: Endosperm-specific activity of a zein gene promoter in transgenic tobacco plants. EMBO J. 7 (1988) 1249-1255.

Schindler, U. and Cashmore, A.R.: Photoregulated gene expression may involve ubiquitous DNA binding proteins. EMBO J. 9 (1990) 3415-3427.

Schindler, U., Menkens, A.E., Beckmann, H., Ecker, J.R. and Cashmore, A.R.: Heterodimerization between light-regulated and ubiquitously expressed *Arabidopsis* GBF bZIP proteins. EMBO J. 11 (1992) 1261-1273.

Schindler, U., Terzaghi, W., Beckmann, H., Kadesch, T. and Cashmore, A.R.: DNA binding site preferences and transcriptional activation properties of the *Arabidopsis* transcription factor GBF1. EMBO J. 11 (1992) 1275-1289.

Schmülling, T., Schell, J. and Spena, A.: Single genes from *Agrobacterium rhizogenes* influence plant development. EMBO J. 7 (1988) 2621-2629.

Schmülling, T., Schell, J. and Spena, A.: Promoters of the *rolA, B,* and *C* genes of *Agrobacterium rhizogenes* are differentially regulated in transgenic plants. Plant Cell 1 (1989) 665-670.

Scholer, H.R., Balling, R., Hatzopoulos, A.K., Suzuki, N. and Gruss, P.: Octomer binding proteins confer transcriptional activity in early mouse embryogenesis. EMBO J. 8 (1989) 2551-2557.

Schulze-Lefert, P., Dangle, J.L., Becker-André, M., Hahlbrock, K. and Schulze, W.: Inducible *in vitro* DNA footprints define sequences

necessary for UV light activation of the parsley chalcone synthase gene. EMBO J. 8 (1989) 651-656.

Shimamoto, K., Terada, R., Izawa, T. and Fujimoto, H.: Fertile transgenic rice plants regenerated from transformed protoplasts. Nature 338 (1989) 274-276

Shure, M., Wessler, S. and Fedoroff, N.: Molecular identification and isolation of the *waxy* locus in maize. Cell 35 (1983) 225-233

Singh, K., Dennis, E.S., Ellis, J.G., Llewellyn, D.J., Tokuhisa, J.G., Wahleithner, J.A. and Peacock, W.J.: OCSBF-1, a maize ocs enhancer binding factor: isolation and expression during development. Plant Cell 2 (1990) 891-903

Slightom, J.L., Durand-Tardif, M., Jouanin, L. and Tepfer, D.: Nucleotide sequence analysis of TL-DNA of *Agrobacterium rhizogenes* agropine type plasmid: identification of open-reading frames. J. Biol. Chem. 261 (1986) 108-121

Spena, A., Schmülling, T., Koncz, C. and Schell, J.: Independent and synergistic activity of *rolA*, *rolB*, *rolC* loci in stimulating abnormal growth in plants. EMBO J. 6 (1987) 3891-3899.

Staiger, D., Kaulen, H. and Schell, J.: A CACGTG motif of the *Antirrhinum majus* chalcone synthase promoter is recognized by an evolutionary conserved nuclear protein. Proc. Natl. Acad. Sci. USA. 86 (1989) 6930-6934.

Staiger, D., Becker, F., Schell, J., Koncz, C. and Palme, K.: Purification of tobacco nuclear proteins binding to a CACGTG motif of the chalcone synthase promoter by DNA affinity chromatography. Eur. J. Biochem. 199 (1991) 519-527.

Sugaya, S., Hayakawa, K., Handa, T. and Uchimiya, H.: Cell-specific expression of the *rolC* gene of the TL-DNA of Ri plasmid in

transgenic tobacco plants. Plant Cell Physiol. 30 (1989) 649-653.

Sugaya, S. and Uchimiya, H.: Deletion analysis of the 5'-upstream region of the *Agrobacterium rhizogenes* Ri plasmid *rolC* gene required for tissue specific expression. Plant Physiol. 99 (1992) 464-467.

Suzuki, A., Kato, A. and Uchimiya, H.: Single-stranded DNA of 5'-upstream region of the *rolC* gene interacts with nuclear proteins of carrot cell cultures. Biochem. Biophys. Res. Commun. 188 (1992) 727-733

Tabata, T., Takase, H., Takayama, S., Mikami, K., Nakatsuka, A., Kawata, T., Nakayama, T. and Iwabuchi, M.: A protein that binds to a *cis*-acting element of wheat histone genes has a leucine zipper motif. Science 245 (1989) 965-967.

Tabata, T., Nakayama, T., Mikami, K. and Iwabuchi, M.: HBP-1a and HBP-1b: Leucine zipper-type transcription factors of wheat. EMBO J. 10 (1991) 1459-1467.

Takase, H., Minami, M. and Iwabuchi, M.: Sequence-specific single-strand DNA-binding proteins that interact with the regulatory regions of wheat histone H3 and H4 genes. Biochem. Biophys. Res. Commun. 176 (1991) 1593-1600

Tepfer, D.: Transformation of several species of higher plants by *Agrobacterium rhizogenes*: Sexual transformation of the transformed genotype and phenotype. Cell 37 (1984) 959-967

Terada, R. and Shimamoto, K.: Expression of CaMV 35S-GUS gene in transgenic rice plants. Mol. Gen. Genet. 220 (1989) 389-392

Toriyama, K. and Hinata, K.: Cell suspension and protoplast culture in rice. Plant Sci. 41 (1985) 179-183

Toriyama, K., Arimoto, Y., Uchimiya, H. and Hinata, K.: Transgenic rice plants after direct gene transfer into protoplasts. Bio/Technology 6 (1988) 1072-1074

Uchimiya, H., Handa, T. and Brar, D.: Transgenic plants. J Biotechnology 12 (1989) 1-20

Vinson, C.R., LaMarco, K.L., Johnson, P.F., Landschulz, W.H. and McKnight, S.L.: In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Dev. 2 (1988) 801-806.

Vinson, C.R., Sigler, P.B. and McKnight, S.L.: Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science 246 (1989) 911-916.

Weisshaar, B., Armstrong, G.A., Block, A., da Costs e Silva, O. and Hahlbrock, K.: Light-inducible and constitutively expressed DNA-binding proteins recognize a plant promoter element with functional relevance in light responsiveness. EMBO J. 10 (1991) 1777-1786.

White, F.F. and Nester, E.W.: Hairy root: plasmid encodes virulence traits in *Agrobacterium rhizogenes*. J. Bacteriol. 141 (1980) 1134-1141.

White, F.F., Taylor, B.H., Huffman, G.A., Gordon, M.P. and Nester, E.W.: Molecular and genetic analysis of the transferred DNA regions of the root-inducing plasmid of *Agrobacterium rhizogenes*. J. Bacteriol. 164 (1985) 33-44.

Willmitzer, L.: The use of transgenic plants to study plant gene expression. Trends. Genet. 4 (1988) 13-18.

inches 1 2 3 4 5 6 7 8

cm 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Kodak Color Control Patches

Blue Cyan Green Yellow Red Magenta White 3/Color Black

Kodak Gray Scale

C Y M

A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19