TGTCACA Motif Is a Novel cis-Regulatory Enhancer Element Involved in Fruit-specific Expression of the cucumisin Gene*

Received for publication, October 15, 2001, and in revised form, December 26, 2001
Published, JBC Papers in Press, January 8, 2002, DOI 10.1074/jbc.M109946200

Hiroshi Yamagata‡, Kiyoki Yonesu, Ayako Hirata, and Yasuo Aizono
From the Laboratory of Biochemistry, Faculty of Agriculture, Kobe University, Nada, Kobe, 657-8501 Japan

Cucumisin, a subtilisin-like serine protease, is expressed at high levels in the fruit of melon (Cucumis melo L.) and accumulates in the juice. We investigated roles of the promoter regions and DNA-protein interactions in fruit-specific expression of the cucumisin gene. In transient expression analysis, a chimeric gene construct containing a 1.2-kb cucumisin promoter fused to a β-glucuronidase (GUS) reporter gene was expressed in fruit tissues at high levels, but the promoter activities in leaves and stems were very low. Deletion analysis indicated that a positive regulatory region is located between nucleotides −234 and −214 relative to the transcriptional initiation site. Gain-of-function experiments revealed that this 20-bp sequence conferred fruit specificity and contained a regulatory enhancer. Gel mobility shift experiments demonstrated the presence of fruit nuclear factors that interact with the cucumisin promoter. A typical G-box (GACACGTGTC) present in the 20-bp sequence did not bind fruit protein, but two possible cis-elements, an I-box-like sequence (AGATATAAAAA) and an odd base palindromic TGTCACA motif, were identified in the promoter region between positions −254 and −215. The I-box-like sequence bound more tightly to fruit nuclear protein than the TGTCACA motif. The I-box-like sequence functions as a negative regulatory element, and the TGTCACA motif is a novel enhancer element necessary for fruit-specific expression of the cucumisin gene. Specific nucleotides responsible for the binding of fruit nuclear protein in these two elements were also determined.

Timing and levels of gene expression are critical to the proper development of eukaryotic organisms. Regulation of the expression pattern of a particular gene can involve the specific binding of trans-acting factors to the cognate cis-elements, constituting a crucial step in transcriptional initiation and, in turn, on the spatial and temporal expression of genes. Plant genes that show tissue specificity, developmental specificity, and a wide range of expression levels have been characterized, whereas their expression patterns are also influenced by environmental stimuli. A family of genes for fruit proteins provide a model system for the study of the regulatory mechanisms of plant genes, since their expression is restricted to a specific tissue and stage during fruit development (1). A number of fruit-specific genes that are activated during ripening have been isolated from tomato and other fruits, and genes responding to ethylene and nonethylene signals have been identified (1, 2). The promoters of fruit-specific genes would also be of great interest for use in strategies to manipulate fruit metabolism and produce valuable proteins such as antibody, biopharmaceuticals, and edible vaccines through methods of genetic engineering (3–5). However, the detailed mechanisms by which the expression of fruit protein genes are regulated are poorly understood, as many of the essential cis-elements have not been identified.

Melon cucumisin, an extracellular subtilisin-like serine protease, is expressed at high levels in fruit, where it comprises more than 10% of the total protein (6). No expression is detected in other parts of the melon plant. The accumulation of a large amount of the cucumisin protein in juice in the central parts of the fruit at a specific stage during fruit development and the single copy character of the corresponding gene suggest that the cucumisin gene contains a strong promoter, which also controls the fruit-specific expression of cucumisin (7). The cucumisin gene is also characterized by its conspicuous expression pattern in specific tissues such as the placenta, locule, and peripheral tissues around seeds in the fruit and early during fruit development. These specific temporal and spatial expression patterns of the cucumisin promoter may be explained as the result of regulatory assemblies of several transcriptional activators that recognize the cis-elements implicated in fruit-specific expression. The cucumisin promoter may also be useful for genetic engineering of fruit. To understand the mechanisms of fruit-specific gene expression and for the design and application of fruit-specific promoters for the improvement of fruit quality by genetic engineering, detailed analyses of the cucumisin promoter and the transcription factors are required.

Here, we describe cloning of the cucumisin promoter and experiments designed to define cis-acting elements involved in proper expression of the gene. A stretch of only 20 bp of the promoter sequence is required for the promoter activity in young fruit.

EXPERIMENTAL PROCEDURES

Plant Material—Musk melons (Cucumis melo L. var. reticulatus cv. Teresa) were cultivated in a greenhouse at the experimental farm attached to the Faculty of Agriculture, Kobe University, from March to August. Fruits were tagged upon self-pollination, and developing fruits were harvested on the 10th day after pollination (DAP). For the analysis of transient gene expression, developing fruits, leaves, and stems were used immediately after harvest. The central parts of the fruit were separated from the sarcocarp and testae and used for isolation of nuclear proteins, or frozen in liquid nitrogen and stored at −80 °C until use for poly(A)* RNA isolation.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF055805.

† To whom correspondence should be addressed: Laboratory of Biochemistry, Faculty of Agriculture, Kobe University, Rokkodai-cho 1, Nada-ku, Kobe, 657-8501, Japan. Tel/Fax: 81-78-803-5875; E-mail: yamagata@kobe-u.ac.jp.

‡ To whom correspondence should be addressed: Laboratory of Biochemistry, Faculty of Agriculture, Kobe University, Rokkodai-cho 1, Nada-ku, Kobe, 657-8501, Japan. Tel./Fax: 81-78-803-5875; E-mail: yamagata@kobe-u.ac.jp.

The abbreviations used are: DAP, day after pollination; GUS, β-glucuronidase; LUC, luciferase; PIPES, 1,4-piperazinediethanesulfonic acid; GBF, G-box binding factor.
Reagents—Restriction and modification enzymes were obtained from New England Biolabs Inc., Roche Molecular Biochemicals, and Promega. Megaprime DNA labeling kits, Hybond-N nylon membranes, [32P]dATP, [α-32P]dCTP, [α-32P]dGTP, and [α-32P]dTTP were obtained from Amersham Biosciences, Inc. Specific activities of all radioactive nucleotide species was 150 Ci/mmol. All other commonly available reagents were of analytical grade.

Cloning of the Cucumis Species—Genomic DNA was prepared from 10 g of young leaves using cetyltrimethylammonium bromide according to the method of Rogers and Bendich (8) as described previously (7). To determine the DNA sequence of the 5’ portion of the exon/intron region of the uidA gene, the cDNA was used as a template for a PCR using the forward primer (5'-GATCGAGCTCCAATAA) and reverse primer (5'-TGGTTCTGACTATATCCTT-3') that were synthesized on an Applied Biosystems 380A oligonucleotide synthesizer (Foster City, CA) based on the sequence of cumin cDNA (7). The PCR conditions were 94°C for 2 min, followed by 30 cycles of 1 min at 94°C, 1.5 min at 55°C, and 2 min at 72°C. The amplified fragment (1.8 kb) was digested with ScaI and HindIII, inserted into pbBlueScript II SK (subclone name pUK2) and sequenced. Sequence analyses were carried out using the Genetyx program (Software Development Co., Ltd., Tokyo).

Primer Extension Analysis—Primer extension analysis was carried out as described by Sambrook et al. (10). The synthetic oligonucleotide Ex (5’-TAAGCGAGTCTGACTATATCCTT-3’) was 5’-end-labeled with [γ-32P]ATP and T4 polynucleotide kinase (New England Biolabs) and then used as a primer for a reverse transcription reaction with poly(A)+ RNA prepared from immature melon fruit (10 DAP) as described by Foster et al. (15) using avian myeloblastosis virus reverse transcriptase (Invitrogen). The extended product was compared with the sequence ladder obtained from the pUK2 primer with the same synthetic oligonucleotide (Ex).

Particle Bombardment—Immature melon fruits (10 DAP) cut horizontally into slices (2-mm thickness), young leaves, and stems were put into Petri dishes. Each sample was bombarded twice with gold particles (1.6 μm in diameter) coated with plasmids. In all experiments, we bombarded the samples with two kinds of plasmid, cucumin promoter fused to β-glucuronidase (uidA; Ref. 11) (denoted as GUS) and the firefly luciferase (LUC) reporter gene (pBI11FF) constructed as described below as an internal standard. An equimolar amount of each plasmid (GUS construct 2.5 μg and 2.5 μg of the pBI11FF plasmid mixed in 3 mg of gold particles and suspended in 100 μl of ethanol. Each plant material was bombarded with a 12-μl aliquot of the suspension (360 μg of gold particles) per shot using a helium-driven Biolistic PDS1000 system (Bio-Rad) with a 28-mm Hg vacuum. The distance between the rupture disc (1,100 psi) and macrocarrier and that between the macrocarrier and sample were 3.0 and 6.0 cm, respectively. After bombardment, the tissues were cultured with aluminum foil and incubated in darkness for 24 h at 25°C.

β-Glucuronidase (GUS) and Luciferase (LUC) Assays and Measurement of Protein Concentration—The central parts of fruit sections, leaves, and stems that were bombarded as described above were ground in liquid N2 using a chilled mortar and pestle. The powder was dispersed into microcentrifuge tubes and mixed with 500 ml of Pica Gene lysis buffer (LC/PGC-51; Wako, Osaka, Japan) containing 2 μl dioxylpropylfluorophosphate and then centrifuged at 13,000 × g for 5 min at 4°C. The supernatant was frozen at −80°C until the enzyme assay was conducted. Proteins were measured using a Bio-Rad protein assay kit with bovine serum albumin as a standard. GUS activity was measured by the method of Jefferson et al. (11), using 4-methylumbelliferyl β-D-glucuronide (Nacalai Tesque, Kyoto, Japan) as substrate. The fluorometer was calibrated with 4-methylumbelliferone standards in the same solution used for sample assays and blanks against the zero time point. LUC assays were performed as described by Miller et al. (12) using a Pica Gene LUC assay kit (Wako). Photon emission derived from LUC activity was counted with a luminometer (model TD-2020; Turner Designs, Sunnyvale, CA) using an equimolar amount of each construct (12) in place of pBI221. Both pBI11FF and pBI221 have the same cauliflower mosaic virus 35S constitutive promoter and no-palinate synthase terminator. For the gain-of-function experiments, the 4 × 20-bp tandem repeat sequence was prepared as follows. Two synthesized oligonucleotides, pHisPeX (5’-GCGGGGATCCGAACATGCTTATACATTGAAGC-3’) and pHiSimX (5’-GGGTTCTAGAATTAGGTGTCACACCTCTTAACTGTTAAGCTTCC-3’) were annealed and digested with EcoRI and XbaI to yield a 90-bp fragment denoted as pHis. Then pHis was ligated with EcoRI-NolI and XbaI-NotI and sequenced with the corresponding restriction enzymes. The NotI-XbaI fragment of this sequence in the normal orientation and NotI-NotI fragment in the reverse orientation were placed into NolIXbaI and SacNI-digested pBluescript plasmids, respectively, to yield pN-X (normal orientation) and pS-N (reverse orientation), and these plasmids were used for particle bombardment. All inserted sequences in each construct were confirmed by DNA sequence analysis.

Isolation of Nuclei and Protein—The nuclear isolation procedure for the plant material that was described was performed with some modifications. All procedures were performed at 4°C. The central parts of the immature melon fruit (about 10 DAP) were cut out from the pericarp, and the seeds were removed with a spatula. The tissue including the placenta, columella, and fibrous tissue (500 g) was ground using a commercial blender in 500 ml of nuclear grinding buffer (1 M hexylene glycol, 10 mM PIPES/KOH, pH 7.0, 10 mM MgCl2, 0.2% (w/v) diisopropylfluorophosphate). After centrifugation of the mixture at 10,000 g for 15 min, the supernatant was ultracentrifuged at 100,000 g for 90 min using a Beckman SW/28 rotor. Then solid ammonium sulfate was added to the supernatant to 0.25 g/ml, and the protein precipitate was collected by centrifugation at 10,000 × g for 15 min. The pellet was resuspended in 20 ml of nuclear lysis buffer (110 mM KCl, 15 mM HEPES/KOH, pH 7.5, 5 mM MgCl2, 1 mM dithiothreitol, 5 mg/ml antipain, 5 μg/ml leupeptin, and 5 μg/ml chymostatin). Then 2 ml of 4 M ammonium sulfate was added to the mixture in several small aliquots with gentle mixing, and the mixture was ultracentrifuged at 100,000 × g for 90 min using a Beckman SW41 rotor. Then solid ammonium sulfate was added to the supernatant to 0.25 g/ml, and the protein precipitate was collected by centrifugation at 10,000 × g for 15 min. The pellet was resuspended with 0.5 ml of nuclear extract buffer (70 mM KCl, 25 mM HEPES/KOH, pH 7.5, 0.1 mM EDTA, 20% (v/v) glycerol, 1 mM dithiothreitol, 5 μg/ml antipain, 5 μg/ml leupeptin, and 5 μg/ml chymostatin). The gel was boiled in 10 mM MgCl2, 1 M dithiothreitol, 5 mM 1,5-diaminonaphthalene-8-sulfonate fluoride, and 2 mM diisopropylfluorophosphate). The mixture was filtered through two layers of gauze and 50-μm nylon mesh. The filtrate was centrifuged at 2,000 × g for 10 min, and the pellet (crude nuclei) was gently resuspended with a soft brush in 80 ml of nuclear wash buffer (0.5% hexylene glycol, 10 mM PIPES/KOH, pH 7.0, 10 mM MgCl2, 0.2% (v/v) Triton X-100, 5 mM 2-mercaptoethanol, and 1.0 mM diisopropylfluorophosphate). After centrifugation of the mixture at 3,000 × g for 5 min, the pellet was resuspended in 20 ml of nuclear lysis buffer (110 mM KCl, 15 mM HEPES/KOH, pH 7.5, 5 mM MgCl2, 1 mM dithiothreitol, 5 mg/ml antipain, 5 μg/ml leupeptin, and 5 μg/ml chymostatin). Then 2 ml of 4 M ammonium sulfate was added to the mixture in several small aliquots with gentle mixing, and the mixture was ultracentrifuged at 100,000 × g for 90 min using a Beckman SW/Ti60 rotor. Then solid ammonium sulfate was added to the supernatant to 0.25 g/ml, and the protein precipitate was collected by centrifugation at 10,000 × g for 15 min. The pellet was resuspended with 0.5 ml of nuclear extract buffer (70 mM KCl, 25 mM HEPES/KOH, pH 7.5, 0.1 mM EDTA, 20% (v/v) glycerol, 1 mM dithiothreitol, 5 μg/ml antipain, 5 μg/ml leupeptin, and 5 μg/ml chymostatin). The gel was boiled in 10 mM MgCl2, 1 M dithiothreitol, 5 mM 1,5-diaminonaphthalene-8-sulfonate fluoride, and 2 mM diisopropylfluorophosphate). After centrifugation at 12,000 × g for 10 min, the extract was frozen in liquid N2 and stored at −80°C until use.
gesized and used as probes (only the upper strands are shown for Ga, Gb, and Gc): Gp, 5'-AGATATGATAAATATAGCCACGTGTC-CAACCCTA-3'; Gm, 5'-ATTACCGGACATGTCTGCTGCTACTTATTACATCACACG-3'; Hp, 5'-ACCTAATAGATATATCGATCTGCTACTTATTACATCACACG-3'; Hm, 5'-AAGACGTGAGATATATCGATCTGCTACTTATTACATCACACG-3'.

Complementary single-stranded oligonucleotides, Gp and Gm or Hp and Hm containing a protruding AA or AT at the 5'-end were annealed and labeled with 0.925 Mbq of [α-32P]dATP and [α-32P]dGTP by fill-in reaction with Klenow fragment (New England Biolabs). For Ga, Gb, and Gc oligonucleotides, each complementary single-stranded oligonucleotide was annealed, resulting in blunt ends and end-labeled with 1.5 Mbq [α-32P]dATP and T4 polynucleotide kinase (New England Biolabs). The labeled oligonucleotides were fractionated on G-25 microspin columns (Amersham Biosciences), and the radioactivity was measured on a liquid scintillation counter. Gel mobility shifts were performed in a volume of 20 μl containing labeled oligonucleotide probe (200,000 cpm) and 6 μg of nuclear protein in binding buffer (25 μl HEPES-KOH, pH 7.5, 100 mM KCl, 100 mM Na2, 10% (v/v) glycerol, and 3 mg of poly(dI-dC)-poly(dI-dC)). For competition assays, the following oligonucleotides were synthesized and used as unlabeled competitors (only the upper strands are shown): GH, 5'-Deletion Analyses and Transient Assays Showed That the Cucumisin Upstream Region Had cis-Regulatory Regions for cis-Element Involved in Fruit-specific Gene Expression.

Results

Characterization of the Genomic Clone of the Cucumisin Gene and Determination of the Transcriptional Start Site—The 5' part of the cucumisin gene containing the 1.2-kb promoter was obtained by inverse PCR using melon genomic DNA as a template. The combined nucleotide sequence, which was confirmed by several independent rounds of ordinary PCR to correct for possible PCR errors, was 3,049 bp in length and contained five exons encoding Met1-Ala183 of the cucumisin precursor and four introns. Fig. 1 shows the nucleotide sequence of the 1.2-kb promoter and short stretch of transcribed region. The entire sequence is available from the GenBank data base under accession number AY055805. The nucleotide sequence of the coding region was identical to that of the cDNA (7). Introns 1–3 were mapped in the prosequence of the cucumisin precursor, and intron 4 was located in that of the mature protease domain. Primer extension analysis revealed that transcription begins at an adenine residue located 48 bp upstream of the translational start site (Fig. 2). The putative TATA box, TATAAA, and CAAT box, CAAAT, were located 34 and 89 bp upstream, respectively, from this transcriptional start site. In the upstream region, a perfect palindromic G-box (5'-GACACGTGTC-3'), the target-binding site of plant basic leucine zipper proteins (16), was found at -234 to -225 bp from the transcriptional start site. An I-box-like sequence containing a tandem GATA repeat (5'-AGATATGATAAATATAGCCACGTGTC-3'), in that 9 bases were identical to the reported 13-bp I-box element (5'-GGATGAGATAAGA-3') (17), was also found at positions -253 to -241 (Fig. 1).

5'-Deletion Analyses and Transient Assays Showed That the Cucumisin Upstream Region Had cis-Regulatory Regions for cis-Expression in Fruit—We used horizontal sections (2–3 mm thick) of the immature melon fruit to examine reporter gene expression in transient assays. We fused the DNA sequence extending 1,181 bp upstream from the transcriptional start site (P1181 construct; Fig. 1) to the GUS reporter gene and introduced the construct into the central part of the fruit sections by particle bombardment with the 35S-LUC construct (pB10PF) as the internal standard (Fig. 3). When the relative GUS activity of the control plasmid (pBI221) containing a constitutive cauliflower mosaic virus 35S promoter was taken as 100, the activity of the p1181 was 35, indicating that the 1181-bp 5'-upstream region was sufficient for the high level expression of the cucumisin gene in fruit.
expression of the **GUS** gene in melon fruit (Fig. 3A).

To broadly determine the positive enhancer region, 5′ deletion analysis of the 1,181-bp *cucumisin* promoter was conducted, and the activities of the deletion constructs were measured in young fruits, leaves, and stems. The p234 construct showed comparable levels of activity to the full-length p1181 construct in fruit, indicating that a region sufficient to confer higher expression in fruit is present within the −234 bp region from the transcriptional start site (Fig. 3A). However, the activity of p228 constructs lacking 6 nucleotides (GACACG) in the G-box element (5′-GACACGTGTC-3′) was decreased to about half of that of p1181. In the p214 construct, the activity decreased to about one-sixth of that of p1181. No further drop in activity was observed by deleting the promoter to position −89, and these activities were thought to be the basal level of the *cucumisin* minimal promoter (p89) containing the TATA box and CAAT box. These results indicated that critical cis-acting element(s) responsible for high level *cucumisin* promoter activity in fruit were localized in the 20-bp fragment from −234 to −215, the sequence of which is 5′-GACACGTGTCACAC-GATAAAT-3′. This sequence contains a perfect palindrome of a typical G-box element (5′-GACACGTGTC-3′) in the first half. The decrease in activity of the p254 construct suggested the presence of another positive element(s) in the region from −310 to −255 and negative element(s) in the region from −254 to −235, which contains an I-box-like sequence (5′-AGATATGATAAAA-3′). Since p234 had strong activity comparable with the longest p1181, we did not examine further the possible positive element(s) upstream of −234.

### The *Cucumisin* 5′ Upstream Region Can Direct Fruit-specific Expression

All constructs from P1181 to P214 introduced in melon leaves and stems showed very low levels of promoter activity in contrast to the control pB221 that showed a high level of activity in these organs (Fig. 3, B and C). The activities of *cucumisin* promoter constructs in stems were slightly higher than those in leaves. These results indicated that the cis-regulatory element(s) in the upstream region of the *cucumisin* promoter is responsible for not only high level expression in fruits but also for fruit-specific expression.

The 20-bp Sequence Is Sufficient to Confer Fruit-specific Expression on a Cucumisin Minimal Promoter—We conducted gain-of-function analysis to determine whether the 20-bp region from −234 to −215 is able to confer fruit-specific high level expression on the *cucumisin* minimal promoter. Four tandem repeats of the 20-bp sequence were fused in the normal or reverse orientation to the truncated (−89 to −1 from the transcriptional start site) *cucumisin* promoter containing the putative TATA and CAAT boxes to yield pN-X and pS-N, respectively (Fig. 4). In the absence of the 20-bp sequence, the construct (p89) showed very low levels of GUS activity in fruit tissues as in leaves and stems. However, fruit bombarded with construct containing four copies of the 20-bp upstream sequence (pN-X) showed an 8-fold induction of GUS activity, and the activity was comparable with that of fruit bombarded with the full-length *cucumisin* promoter (p1181) (Fig. 4A). It is noteworthy that the tandem repeat of the 20-bp sequence promoted GUS expression in both orientations. However, both pN-X and pS-N did not lead to elevation of basal level of GUS activity derived from p89 in leaf or stem tissues (Fig. 4, B and C). These results indicated that the 20-bp sequence has two functions: to enhance transcription in fruit tissues and to act as a fruit-specific enhancer.

Detection of Nuclear Factors That Interact with the *Cucumisin* Promoter 5′ Upstream Region—Transient expression assays suggested that the 20-bp sequence from −234 to −215 containing a G-box element (5′-GACACGTGTC-3′) contains cis-element(s) responsible for high level and fruit-specific expression and suggested that the region from −254 to −235 containing the I-box-like sequence (5′-AGATATGATAAAA-3′) that has a tandem repeat of the GATA sequence has a negative influence on gene expression. Another I-box-related sequence, a GATA element, was also found between −212 and −207 (5′-AGATAT-3′) and between −195 and −189 on the opposite DNA strand (5′-GATAGAAA-3′) (Fig. 5). To evaluate whether
A.

**Fig. 4.** Gain-of-function analysis of the 20-bp sequence. Chimeric genes in which four tandem repeats of the 20-bp sequence were fused 5′ to the truncated (-89) *cucumisin* promoter-GUS construct (p89) in normal or reverse orientation, and designated as pN-X and pS-X, respectively, are shown. Schematic representations of the promoter region are shown on the left in A. The arrows indicate the polarity of the promoter fragment. Equivalent amounts of each construct were introduced into horizontal slices of immature melon fruit (A), leaves (B), and stems (C) by particle bombardment with the 35S-LUC construct as the internal standard. Average GUS activities in each melon tissue were calculated as described under “Experimental Procedures.” Values are the means of those from at least six (A) or three (B and C) independently bombarded samples, with error bars representing S.E. (n ≥ 6 in A, and n ≥ 3 in B and C).

B.

**Fig. 5.** DNA probes and competitors used for gel mobility shift assays. The nucleotide sequence of the region from -254 to -183 in the *cucumisin* promoter is shown. Typical cis-elements found in the promoters of other plant genes and a TGTACA motif (see text) are boxed. The arrows indicate the synthetic oligonucleotides used as probes and competitors in subsequent gel mobility shift assays.

C.

**Fig. 6.** Gel mobility shift assay of nuclear proteins from melon fruits with the synthetic oligonucleotides. Binding of nuclear proteins from young melon fruit to the labeled oligonucleotide probes G and H was assayed. The arrowhead indicates the bound complex. F, free probe.

these sequences interact specifically with proteins in nuclei from melon fruit, we conducted gel electrophoresis mobility shift assays. Nuclear protein extracts were prepared from central parts of immature fruit. First, we divided the -254 to -183 region that was thought to contain both positive and negative regulatory elements as described above into two regions, G (-254 to -215) and H (-222 to -183) as shown in Fig. 5. When these sequences were used as probes in the gel mobility shift assays, binding of nuclear proteins from central parts of young fruit to these sequences was detected as a broad retarded band (Fig. 6, lanes 2 and 4). The signal of the complex of the fruit protein with the G probe was far stronger than that with the H probe, indicating that the nuclear protein(s) in the fruit binds more tightly to the G region than to the H region.

**Determination of Protein-binding Region within the Cucumber Promoter.** To localize the cis-element responsible for binding with fruit nuclear proteins more precisely within the G region, we further divided the G and H regions into Ga (-254 to -235) and Gb (-234 to -215), and Ha (-222 to -203) and Hb (-202 to -183), respectively (Fig. 5). In addition to these fragments, GH (-240 to -201) spanning the G and H regions was synthesized. These oligonucleotides were used as competitors in the gel mobility shift assays with the G, H, and GH fragments as labeled probes (Fig. 7, A–C). Surprisingly, not only the G competitor but also the H competitor completely abolished the retarded band of the G probe, when these competitors were present in a 100-fold molar excess relative to the labeled probe (Fig. 7A, lanes 2–4). Conversely, the binding of the H probe to the protein competed with not only the H fragment but also with the G fragment (Fig. 7B, lanes 2–4). The addition of the GH and Ga competitors to the binding mixture with the G probe significantly diminished the retarded band, whereas the Gb competitor showed weak competition (Fig. 7A, lanes 5–7). The Hb, but not the Ha, could compete strongly with the G probe (Fig. 7A, lanes 8 and 9). These results confirmed that there are at least two binding regions to the nuclear proteins within the G region (i.e., Ga and Gb), and Ga binds more strongly to the nuclear proteins than Gb. The I-box-like sequence (5′-AGATATGATAAAA-3′) in the Ga region is a possible cis-element for the observed tight binding of the G region to the nuclear proteins, since a similar I-box-related GATA sequence (5′-GATAGAA-3′) was present on the opposite DNA strand within the Hb region that also competed well with the G probe. The weak competition of the Ha with the G probe (Fig.
tors were assayed for their ability to interfere with complexes formed 
excess. Sequences of the probes and competitors are given in Fig. 5.
the estimation of the 
7A, lane 8) can be explained by the presence of another GATA 
sequence (AGATAT) in the Ha fragment, the same sequence 
with the first 6 bases in the I-box-like sequence present in the 
Ga region (Fig. 5). Thus, the I-box-like sequence (5'-AGATAT-GATAAAA-3') in the Ga region and 20 bp of Gb sequence are 
likely to be the target binding sites of the fruit nuclear proteins.

When the H fragment was used as a probe, besides H and G 
fragments, GH, Hb, and Gb fragments were strong competi-
tors, and Ha and Ga fragments were weak competitors (Fig. 7B).
These results agreed well with the above suggestion for the 
estimation of the cis-element in the G region, and I-box-
related GATA sequence in the opposite strand (5'-GATAGAA-
3') in the Hb region and another GATA sequence (5'-AGATAT-
3') in the Ha region seem to be strong and weak binding 
regions, respectively. When the labeled GH fragment was used as 
a probe, the signal of the retarded band was much weaker 
than those seen with G and H probes, and longer exposure of 
the film was needed, indicating that the nuclear protein binds 
only weakly to the GH region (Fig. 7C). The G and H fragments 
as well as the GH fragment itself strongly competed with the 
GH probe. Also, the competition with the Gb and Ha fragments 
was stronger than with the Ga and Hb fragments, indicating 
that the GH region had also at least two elements (Gb and Ha 
regions) involved in protein binding. Since the DNA sequence 
downstream of -214 is thought not to be involved in the high 
level and fruit-specific expression of the cucumisin promoter 
(Fig. 3), the H region was not studied further in the following 
experiments.

To test the possibility that the nuclear protein binds to the 
central parts in the G region spanning the Ga and Gb, we 
synthesized a Gc oligonucleotide (-244 to -225; Fig. 5) and 
used it as a labeled probe for the gel mobility shift assay. As 
shown in Fig. 8A, the Gc probe did not show any retarded band 
in contrast to the Ga and Gb probes that showed strong and 
weak retarded bands, respectively. Therefore, the G-box (5'-
GACACGTGTC-3') in the Gc region could be excluded from the 
candidates of cis-elements responsible for binding to nuclear 
proteins.

I-box-like Sequence and TGTCACA Motif Are cis-Elements 
Responsible for the Binding of Fruit Nuclear Proteins—When 
the 40-bp G region was used as a labeled probe and four 10-bp 
segments derived from the G region (G1 (5'-AGATATAGAT-3'), 
G2 (5'-AAAATAATAGAC-3'), G3 (5'-GACACGTGTC-3'), and G4 
(5'-ACAACCTAAT-3') (Fig. 5) were used as competitors, only 
G4 alone weakly competed for the binding of nuclear protein to 
the G probe, whereas G1, G2, and G3 fragments showed little 
competition (Fig. 8B). The observation that the G3 fragment, a 
10-bp palindrome of the G-box, did not compete with the G probe 
indicated again that the G-box did not significantly con-
tribute to the binding of fruit nuclear protein to the cucumisin 
promoter but that the G4 fragment was important for the 
observed protein binding to the Gb probe (Fig. 8A, lane 4). An 
odd base palindrome, TGTCACA motif, is present in the 
center of the Gb region and overlaps with the G-box. We suspected 
this TGTCACA motif to be the cis-element responsible for 
protein binding. A 13-bp I-box-like sequence (5'-AGATAGAT-
GATAAAA-3') inhibited protein binding to the G probe more 
strongly (Fig. 8B, lane 5), consistent with the above observa-
tions for the tight binding of the I-box-like sequence to the 
nuclear protein. However, since G1 (5'-AATATAGAT-3') and 
G2 fragments (5'-AAAATAATAGAC-3') containing the first 9 
bases and the last 4 bases of the I-box-like sequence, respec-
tively, hardly inhibited protein binding (Fig. 8B, lanes 4 and 6), 
the whole I-box-like sequence is probably necessary for protein 
binding.

These observations from gel mobility shift assays showed 
that two possible cis-elements, an I-box-like sequence and 
TGTCACA motif, are present in the G region. The I-box-like 
sequence can bind the nuclear protein tightly, whereas the 
TGTCACA motif shows weak binding. Together with the re-
results of transient expression assays (Figs. 3 and 4), we con-
cluded that the I-box-like sequence in the Ga region is a neg-
ative regulatory element, and the TGTCACA motif in the Gb 
region is responsible for the high level and fruit-specific expres-
sion of the cucumisin promoter.

Critical Nucleotides Responsible for the Binding of Nuclear 
Proteins in the I-box-like Sequence and TGTCACA Motif—To 
confirm the binding of nuclear protein to the I-box-like 
sequence and TGTCACA motif, detailed analysis of sequence-
specific binding was carried out using mutated derivatives of 
these sequences in gel mobility shift assays (Fig. 9). Successive 
single base pair mutations were introduced into these frag-
ments as shown in Fig. 9C (see also “Experimental Proce-
dures”), and the mutated fragments were used as competitors 
in the competition experiments with the wild-type fragment.
As shown in Fig. 9A, the specificity of binding of nuclear 
protein to the G region was demonstrated by the ability to 
compete this complex effectively with a 500-fold molar excess of 
the I-box-like sequence (Fig. 9A, lane 3). However, competitors 
containing a point mutation located at positions 1, 3, 4, 8, 9, 10, 
or 11 in the I-box-like sequence failed to fully compete binding 
of nuclear protein to the wild-type fragment. Especially, the
cis-Element Involved in Fruit-specific Gene Expression

Fig. 8. Binding of fruit nuclear proteins to the G region. A, Ga, Gb, and Gc fragments of the G region of the cucumisin gene, the sequences of which are shown in Fig. 5, were used as probes in gel mobility shift assays with young fruit nuclear proteins. B, several oligonucleotide competitors were assayed for their ability to interfere with complexes formed between G probe and young fruit nuclear protein. Oligonucleotide competitors, the sequences of which are shown in Fig. 5, were added in 100-fold (lane 3) or 500-fold (lanes 4-8) molar excess. F, I-box-like sequence.

oligonucleotide with a point mutation at position 9 (T) in the I-box-like sequence showed no competition. In contrast, oligonucleotides containing mutations at the middle position (oligonucleotides 5-7) and downstream (oligonucleotides 12-14) of this sequence efficiently competed with the binding of nuclear protein. This indicated that one of the nuclear protein binding sites in the G region is located in a region covering the I-box sequence, and the nucleotides important for the binding are A1, A3, T4, A8, T9, A10, and A11, of which T9 is most important (Fig. 9D).

An analogous experiment was carried out for the TGTCACA motif. A 500-fold molar excess of G5 fragment (5'-TGTCACAACCTAAT-3') efficiently interfered with the protein binding to labeled Gb fragment (Fig. 9B, lanes 3 and 15). However, competitors containing point mutations located at position 3 or 7 in the G5 sequence failed to fully compete with binding of nuclear protein to the wild-type fragment. These observations indicated again that the TGTCACA motif is responsible for the binding of nuclear protein to the Gb region, and the critical nucleotides are T3 and A7 in the G5 fragment (Fig. 9D). Together with transient expression assays (Figs. 3 and 4), these results confirmed the role of the TGTCACA motif as a cis-acting enhancer element in the cucumisin promoter.

DISCUSSION

Melon is a typical climacteric fruit, and its ripening is influenced by ethylene, which regulates expression of several ripening-related genes in a variety of climacteric fruits (1, 2). The biosynthesis of ethylene in melon fruit occurs at a late stage in fruit development, and the level of 1-aminocyclopropane-1-carboxylate oxidase, a key ethylene biosynthetic enzyme, is very low in unripe melon fruit (18, 19). Synthesis of cucumisin, the precursor of melon 1-cysteine proteinase (20), is completely inhibited by ethylene, which regulates expression of several ripening-related genes (21). This result is in agreement with many other plant promoter studies in which promoter regions in the range of several hundred base pairs to about 1 kb have been found to reproduce promoter activity compared with the activity of the endogenous promoter. Therefore, introns, downstream sequences, or extended domains in the chromatin at the chromosomal location of the cucumisin gene are unlikely to contribute significantly, if at all, to the regulation of its transcription.

A perfect palindromic G-box sequence, 5'-GACACGTGTC-3', was found in the 5' upstream region of the cucumisin promoter. G-box elements are currently the best-characterized plant cis-regulatory element. A family of plant basic leucine zipper proteins has been identified that interacts with G-box elements to confer high promoter activity (16). We therefore presumed initially that the G-box in the cucumisin promoter was a target for a transcription factor controlling the observed high level expression and fruit specificity of the cucumisin gene. Transient expression assays with truncated promoters fused to the GUS reporter gene, however, indicated that the G-box is not a single dominant cis-element responsible for fruit expression of the cucumisin gene but that the downstream sequence is necessary for full activity of the cucumisin promoter (Fig. 3). Gain-of-function experiments revealed that the 20-bp sequence (5'-GACACGTGTCACAACCTAAT-3') containing the G-box in the first half includes an enhancer that controls fruit-specific ex-
cis-Element Involved in Fruit-specific Gene Expression

Plant G-box elements contain a CACGTG core sequence that is necessary for efficient binding of basic leucine zipper proteins, and sequences flanking the CACGTG core affect basic leucine zipper protein binding specificity (25) and gene expression in vivo (26). The G-box in the *cucumisin* promoter, a 10-bp perfect palindrome (5'-GACACGTGCT-3'), can be classified as a high affinity class I G-box element that binds type A G-box binding factor (GBF) (16). Gel mobility shift assays, however, showed that the G-box is not responsible for the binding of fruit nuclear protein to the *cucumisin* promoter (Fig. 8, A and B). Instead, an odd base palindromic sequence, TGTCACA, localized in the center of the 20-bp sequence and overlapping with the G-box (Fig. 5), was thought to be an essential element for the binding of fruit protein to the Gb region of the *cucumisin* promoter. Furthermore, competition experiments in the gel mobility shift assay clearly showed that T3 and A7 in the TGTCACA motif are critical nucleotides for the binding of nuclear protein in fruit (Fig. 9). Since the G-box element in the *cucumisin* promoter contains the complete structure in itself for binding of GBF (i.e. further flanking sequence is not necessary for the protein binding (16)), the TGTCACA motif is not likely to be a part of the G-box. Most in vivo expression studies have indicated that G-box elements cannot act on their own but require the presence of additional cis-acting sequences for their functions (16). However, the TGTCACA motif in the *cucumisin* promoter is also unlikely to be such a cooperative element for the G-box, because it overlaps with the G-box. Taken together, we concluded that the TGTCACA motif is the primary element responsible for the binding of nuclear protein to the *cucumisin* promoter in fruits. To our knowledge, the TGTCACA motif is a novel cis-element in the promoters of plant genes. For the full activity of the TGTCACA motif, however, part of the G-box is thought to be necessary, since promoter activity was decreased distinctly without 6 bp of the G-box (Fig. 3). One explanation for this observation is that the TGTCACA motif is a possible core sequence of a longer cis-element, and the sequence-flanking TGTCACA may affect its activity as a transcriptional activator. To understand the precise function of the TGTCACA motif, identification of a TGTCACA motif-binding protein is required, and we are now cloning TGTCACA motif-binding protein using the yeast one-hybrid system.

Three members of the tomato gene family for Rubisco (RBCS1, RBCS2, and RBCS3A) contain a conserved pair of I-box (5'-GGATGAGATAAGA-3') and G-box (5'-CACGGT-3') elements in an identical spatial arrangement (27). The promoters of these genes are equally active in all green tissues, with the exception of young fruit, where the RBCS3A promoter has strongly reduced activity compared with the RBCS1 and RBCS2 promoters (17). Comparative analysis of the RBCS2 promoter in tomato leaves and young fruits showed that the 37-bp domain containing the conserved I-box and G-box elements is required for high level promoter activity in leaves, while only the G-box is required in young fruits (28). They speculated that these *RBCS* genes are activated by the coordinate binding of an I-box binding factor (IBF) and GBF in leaves, while the promoter is activated by GBF alone in fruits. cDNAs encoding GBFs were cloned from young tomato fruit, and these GBFs were shown to bind the G-box sequence in the RBCS1, RBCS2, and RBCS3A promoters in vitro (29). However, the differences in the expression patterns of RBCS1, RBCS2, and RBCS3A in tomato young fruit cannot be explained by the G-box alone, since all of these genes have a G-box element in their promoters. Although a negative regulator, the F-box binding factor, acting on the RBCS2A promoter in developing fruit was found (17), it is uncertain whether G-box sequences in RBCS1 and RBCS2 function as genuine

expression, since four tandem repeats of the 20-bp sequence in both orientations enhanced the transcription in fruit and were sufficient to confer fruit specificity of the gene expression to the minimal promoter (Fig. 4).
single dominant positive cis-element in young tomato fruit. Despite the presence of the same kinds of cis-elements and their similar arrangements in promoters of tomato RBCS and melon cucumisin, the functions of these cis-elements seem to be quite different between RBCS and cucumisin genes.

The I-box is less well characterized than the G-box and seems to be involved in light-regulated and/or circadian clock-regulated gene expression of photosynthetic genes (30, 31). Related GATA motifs are found in many other promoters, some of which are light-regulated but others of which are not (30, 32). A tobacco protein, ASF-2, binds not only to the I-box (5'-GGATGAGATAAGA-3') of which are light-regulated but others of which are not (30, 32). Related GATA motifs are found in many other promoters, some of which are light-regulated but others of which are not (30, 32).

Transcriptional activation activity (33). The binding of LeMYB1 has been shown to be an activating factor but not light responsiveness (32). In tomato, the I-box has been shown to be an activating cis-element of the RBCS14. Gallie, D. R., Sleat, D. E., Watts, J. W., Turner, P. C., and Wilson, T. M. A. (1992) Plant Physiol. 101, 727-739

The I-box-like sequence in the RBCS3A promoter was mapped to the whole sequence of the I-box (5'-GGATGAGATAAGA-3'). This would be consistent with our observation that the binding site of nuclear protein in transcriptional activation activity (33). The binding of LeMYB1 in the I-box-like sequence in the cucumisin promoter extended throughout the sequence, and most of the critical nucleotides for protein binding were mapped within the repeats of the core GATA motif (Fig. 9). Thus, we supposed that the cis-element responsible for protein binding in the Ga region is only this I-box-like sequence, and the Myb-like protein probably binds to it. Our finding that the I-box-like sequence is a down-regulating cis-element in gene expression in melon fruit suggests a new function of the I-box-like element as a negative regulator in fruit. The I-box-like sequence and its binding protein may down-regulate the expression of the cucumisin gene developmentally and/or cell type specifically in fruit. In cells in the central part of developing fruit where the cucumisin gene is highly expressed, the binding protein to the I-box-like sequence may be inactive through interaction with other cis-regulatory elements and/or transcription factors.

In conclusion, we have demonstrated that the 20-bp promoter fragment between positions -234 and -214 from the transcriptional start site of the melon cucumisin gene is sufficient to mediate the fruit-specific expression pattern observed for this gene. This pattern is most likely established via the action of fruit-specific transcriptional regulators interacting with cis-acting DNA sequences within the analyzed promoter fragment, in that both positive and negative regulators were found. Candidates for such fruit-specific positive and negative elements were identified as a TGTCACA motif and an I-box-like sequence, respectively. Further studies of their functions in the regulation of gene expression in melon fruit and their potential interactions with transcription factors might provide new insights into the mechanisms of organ-specific transcription in plants.

Acknowledgment—We are grateful to Dr. T. Iwasaki for encouragement and support.

REFERENCES

1. Giovannoni, J. (2001) Annu. Rev. Plant Physiol. Plant Mol. Biol. 52, 725-749
2. Lelièvre, J.-M., Latché, A., Jones, B., Bouzyazen, M., and Pech, J.-C. (1997) Physiol. Plant. 101, 727-739
3. Gray, J., Petton, S., Shabbeer, J., Schub, W., and Grierson, D. (1992) Plant Mol. Biol. 19, 69-87
4. Ruf, S., Hermann, M., Berger, I. J., Carrer, H., and Bock, R. (2001) Nat. Biotechnol. 19, 870-875
5. Danielli, H., Streafeld, S. J., and Wycoff, K. (2001) Trends Plant Sci. 6, 219-226
6. Yamagata, H., Ueno, S., and Iwasaki, T. (1989) Agric. Biol. Chem. 53, 1009-1017
7. Yamagata, H., Masuzawa, T., Nagoaka, Y, Ohnishi, T., and Iwasaki T. (1994) J. Biol. Chem. 269, 32725-32731
8. Rogers, S. G., and Bendich, A. J. (1986) Plant Mol. Biol. 5, 69-78
9. Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) Genetica 120, 621-625
10. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Vol. 1 pp. 7.73-7.81, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
11. Jefferson, R. A., Ravanau, T. A., and Bevan, M. W. (1987) EMBO J. 6, 3901-3907
12. Miller, A. J., Short, S. R., Hirasaka, K, Chua, N.-H., and Kay, S. A. (1992) Plant Mol. Biol. Rep. 10, 324-337
13. Bowler C., Neuhaus, G, Yamagata, H., and Chua, N.-H. (1994) Plant Cell 7, 1009-1113
14. Yamamoto, M., Miki, T., Ishikii, Y, Fujinami, K., Yanagisawa, Y, Nakagawa, H., Ogura, N., Hirabayashi, T., and Sato, T. (1995) Plant Cell Physiol. 36, 591-596
15. Ayubu, R., Gius, M., Amor, M. B., Gillot, L., Roustan, J.-P., Latché, A., Bouzyazen, M., and Pech, J.-C. (1996) Nature Biotechnol. 14, 862-866
16. Fluhr, R. (1998) Trends Plant Sci. 3, 141-146
17. Ohme-Takagi, M., and Shinshi, H. (1995) Plant Cell 7, 173-182
18. Potrykus, I. (1991) Annu. Rev. Plant Physiol. Plant Mol. Biol. 42, 205-255
19. Goff, S. A., Cone, K. C., and Fromm, A. E. (1991) Genes Dev. 5, 298-309
20. Goff, S. A., Klein, T. M., Roth, B. A., Fromm, M. E., Cone, K. C., Radicella, J. P., and Chandler, V. L. (1996) EMBO J. 9, 2517-2522
21. Williams, M. E., Foster, R., and Chua, N.-H. (1992) Plant Cell 4, 485-496
22. Ishige, T., Takaichi, M., Foster, R., Chua, N.-H., and Oeda, K. (1999) Plant J. 18, 443-448
23. Manzana, T., Carrasco, P., and Gruss, W. (1991) Plant Cell 3, 1305-1316
24. Baum, K., Groth, B., and Meier, I. (1997) Plant Cell, 463-469
25. Meier, I., and Gruss, W. (1994) Nucleic Acids Res. 22, 470-478
26. Terzaghi, W. B., and Cashmore, A. R. (1995) Annu. Rev. Plant Physiol. Plant Mol. Biol. 46, 445-474
27. Anderson, S. L., and Kay, S. A. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 1500-1504
28. Lam, E., and Chua, N.-H. (1989) Plant Cell 1, 1147-1156
29. Rose, A., Meier, I., and Wienand, U. (1999) Plant J. 20, 641-652