Journal of Plant Biology, December 2003, 46(4) : 231-244

Molecular Cloning and Structural Analysis of the Cucumber (*Cucumis sativus* L.) Phosphoe*nol*pyruvate Carboxykinase (*CsPCK*) Gene

Dae-Jae Kim*
School of Science Education, College of Education, Chungbuk National University, Chungbuk 361-763, Korea

Genomic sequence of the ATP-dependent phosphoe*nol*pyruvate carboxykinase (*CsPCK*) gene has been determined first from cucumber. Several putative clones were isolated in three rounds of genomic library screening with designated cDNA probes. These clones were analyzed via restriction digests, Southern hybridization, and nucleotide sequencing to ascertain the structure of the *CsPCK* gene. Analysis of a selected positive clone (λ*cscpk*-4A) demonstrated that this gene consists of 13 exons and 12 introns, spanning 9 kb in the cucumber genome. Exon 1 contains only 23 nucleotides of the 5'-noncoding region of cucumber *PCK* cDNA, whereas Exon 2 comprises 12 nucleotides of the 5'-noncoding region with an N-terminal PEPCK coding sequence. All the exon-intron junction sequences agree with the GT/AG consensus, except for the 5 donor site of Intron 7, where GC replaces the GT consensus. As with rice (*Oryza sativa*), cucumber contains only one copy of the *CsPCK* gene in its haploid genome. The overall number of exons and the structure of this gene are similar to those for both *Arabidopsis* Chromosome 4 (Atg4) *PCK* and the rice *PCK* genes, which contain 13 and 12 exons, respectively. Two additional *Arabidopsis PCK* genes can be found in the fifth chromosome (Atg5), which contains 9 exons and 8 introns (with 628 and 670 amino acids, respectively) of the PEPCK peptide. The *CsPCK* gene promoter has conserved plant-specific *cis*-acting elements within 2 kb of the 5' flanking region. Several common *cis*-acting elements of the isocitrate lyase (*icl*) and malate synthase (*ms*) gene promoters, identified in the *CsPCK* gene, are responsible for the sugar response during plant development, especially at germination. These conserved elements are discussed here.

Keywords: C3 plant, exon, genomic library, gluconeogenesis, intron, promoter

The ATP-dependent enzyme, phosphoe*nol*pyruvate carboxykinase (PEPCK or PCK; EC 4.1.1.49), catalyzes reversibly (PEP ↔ oxaloacetate) in the gluconeogenic pathway during the early phase of germination, proceeding from lipids to sugars in C3 oil-seed species (e.g., cucumber, pumpkin, watermelon, and *Arabidopsis*) (Thomas and Rees, 1972; Trevanion et al., 1995). Together with glyoxysome-specific isocitrate lyase (ICL) and malate synthase (MS), PEPCK is synthesized and accumulated rapidly in the cotyledons for a few days immediately after imbibition (Kim and Smith, 1994). Afterward, gene expression is strongly repressed when photosynthesis begins in those tissues. Nevertheless, small amounts of PEPCK proteins still can be detected via western blot analysis in green leaves and developing roots. Furthermore, both transcripts and the proteins of cucumber PEPCK reappear in senescing cotyledons and leaves as nutrients are recycled (Graham et al., 1992; Kim and Smith, 1994). During this second phase of activity, lipids can be converted to carbohydrates via the glyoxylate cycle and gluconeogenesis in order to supply energy and source molecules.

The role of PEPCK appears to be more complex in C4 species than in C3 plants because enzymes in the former are involved in carbon, sulfur, and nitrogen metabolisms (Edwards et al., 1971; Gerwick and Black, 1979; Hatch, 1997). PEPCK is a component of the C4 carbon-concentrating cycle in one subgroup of C4 grasses as well as a wide range of CAM plants. Although this C4 PEPCK has been extensively investigated per its physiological and molecular aspects, no complete genomic analysis has been conducted for its related genes (Finnegan and Burnell, 1995; Finnegan et al., 1999; Furumoto et al., 1999).

It is not surprising that amino acid sequences are dissimilar between C3 and C4 plants because different developmental processes and cell types are involved in each (Tsuyoshi et al., 1999). For example, in the C4 grass *Urochloa panicoides*, four *PCK* genes are differentially expressed-- *PCK1* and *PCK2* in the photosynthesizing leaves versus *PCK3* and *PCK4* predominantly in the developing roots (Finnegan et al., 1999). Nevertheless, high levels of activity for C4 enzymes have also been detected in subtype C3 plants (Hibberd and Quick, 2002). In particular, PEPCK activity has been

*Corresponding author; fax : +82-43-271-0526
e-mail djkim@chungbuk.ac.kr

shown to be 9-fold higher in the cells surrounding the xylem and phloem than in the leaves of tobacco. This clearly indicates that some C4-like biochemical pathways are associated with photosynthesis in C3 species. Therefore, the role of PEPCK may be more complex in those plants, requiring further physiological and biochemical studies to elucidate the function of decarboxylation enzymes and corresponding genes.

The primary structure of ATP enzymes is homologous to the ATP-dependent PEPCK found in bacteria, yeast, trypanosomes, and plants (Kim and Smith, 1994). However, those enzymes share little similarity in their primary structure with the GTP-dependent PEPCK enzymes from animals, which catalyze the first committed step in gluconeogenesis (Utter and Kolenbrander, 1972; Kim and Smith, 1994). C4 plant genes for PEPCK have been cloned and partially sequenced from *U. panicoides* (Finnegan et al., 1999). Genome-sequencing projects have demonstrated three complete *PCK* gene sequences from *Arabidopsis thaliana* (Accession nos. AL021684, NP_680468, and AL161592) as well as a single *PCK* gene sequence (AC099325) from rice (*Oryza sativa*).

We have previously used Southern blot analysis to theorize that the cucumber genome contains a single copy of the *PCK* gene, and have also reported a cDNA clone for the 74-kD subunit of PEPCK (Kim and Smith, 1994). A mature form of this subunit has now been purified from cucumber (Kim and Smith, 1994; Walker and Leegood, 1995). Although the plant PEPCK enzyme plays a role in central metabolic processes for lipid mobilization and carbon metabolism in C3 species (Walker et al., 1999), little information is available on how the gene is organized and regulated in plant development. Though the *Arabidopsis* and rice *PCK* gene sequences have been opened to the public, their molecular structures have not been examined yet. Therefore, as a first step in evaluating the structural organization of this gene, we report here the isolation of the *CsPCK* gene from the cucumber leaf genomic library to determine its entire genomic structure and analyze its putative regulatory sequences. Our objective was to identify common *cis*-acting elements as they coordinate plant development in this species.

MATERIALS AND METHODS

Plant Material

Seeds of cucumber (*Cucumis sativus* L.), were obtained from Heunong Jong Myo (Korea), then imbibed in sterile water at 4°C for 12 h and sown in wet vermiculite. The resultant plants were maintained in a growth chamber (LAB-LINE Instrument Inc. USA) at 25°C and 70% humidity, with a 12-h photoperiod. The first emerging leaves were collected for genomic DNA isolation approximately 10~14 d after seed imbibition, and were immediately frozen in liquid nitrogen.

Bacterial Strains, Plasmids, and Media

Bacterial plasmids were propagated (via pBlueScriptII) in an *Escherichia coli* XL-1 Blue MRF strain (Stratagene, USA). The phage was manipulated on *E. coli* strain XL-1/Blue MRA P2, following the manufacturers protocol (Stratagene). A pBlueScriptII vector was used for subcloning and nucleotide sequencing of the cloned genomic DNA fragments. Bacteria were cultured with either LB broth or LB-containing agar. Competent cells were prepared with a SOB medium (Sambrook et al., 1989).

Nucleic Acid Isolation

We used a DNAeasy Plant Maxi Kit (Qiagen GmBH, USA) to extract genomic DNA from cucumber leaves. Phage DNA was purified with a Qiagen Lambda Mini Kit (Qiagen GmBH) after being propagated on an LB agar plate (Sambrook et al., 1989). The plasmid DNA isolation was carried out with a Plasmid Isolation Kit (Bioneer, Korea). A QIAEXII Gel Extraction Kit (Qiagen GmBH) was used to purify the DNA fragments from the agarose gel. We estimated the concentration of purified DNA at 260 nm with a UV spectrophotometer (Sun-II Science, Korea).

Construction of Genomic DNA Library and Screening

A cucumber genomic library was constructed using the Lambda Fix II Vector Cloning System (Stratagene). The prepared genomic DNA was partially digested with *Sau*3AI enzyme (Roche, Switzerland). After agarose-gel electrophoresis, the fraction having the maximum density of partially digested DNA (range of 10 to 30 kb) was chosen for a fill-in reaction on both ends of the partially digested genomic DNA fragments. This reaction was performed with a Klenow Fill-In Kit (Stratagene). The filled-in genomic DNA fragments were then inserted into the Lambda Fix II vector, using T4 DNA ligase (Stratagene). After overnight ligation at 4°C, the products were *in vitro*-packaged with the Gigapack III Gold packaging extract (Stratagene). The phages were then used to infect the host *E. coli* cells. Titer of the initial

library was estimated to be 0.8×10^5 pfu ml^{-1}. The initial library was amplified immediately and stored at –80°C, as suggested by the manufacturer (Stratagene).

We used standard procedures, according to the manufacturers instructions (Stratagene), to screen the genomic library. For the first screening of the *CsPCK* gene, approximately 5×10^5 plaques were plated out on a pair of 24.5-cm-square Bio-assay dishes (Nunc, USA). The plaque lifts were prepared using a Hybond-N^+ membrane (Amersham Pharmacia Biotech, UK). Pre-hybridization and hybridization reactions were carried out in ECL gold buffer (Amersham Pharmacia Biotech). The ECL direct-labeled cDNA probe was added to the hybridization reaction buffer after 1 h of pre-hybridization at 42°C; hybridization lasted for 16 h at 42°C. The probes used to screen the genomic library were three different fragments of cucumber *PCK* cDNA (Fig. 2; Kim and Smith, 1994) that were prepared from pBRPCK-7A by restriction-enzyme treatments a-0.23-kb *Pst*I fragment, a 1.4-kb *Eco*RI/*Xho*I fragment, and a 0.7-kb *Bam*HI fragment (Fig. 2; Kim and Smith, 1994). The cDNA probes were applied to detect the relevant portion of the *CsPCK* gene. Washing was done first in 5X SSC for 5 min at 55°C, then three times in 0.5X SSC and 0.4% SDS for 15 min each at 55°C, and, finally, twice in 2X SSC for 10 min at room temperature. A hybridization signal was detected with ECL direct detection reagents (Amersham Pharmacia Biotech). After the third round of genomic library screening, two putative clones were selected and named λ*cspck*-4A and λ*cspck*-11B.

Southern Blot Hybridization

Purified cucumber genomic DNA and recombinant phage DNA were digested with restriction enzymes (Takara), separated on a 0.8% (w/v) agarose gel, and transferred onto a Hybond-H^+ membrane by capillary blotting (Sambrook et al., 1989). The blot was washed in 5X SSC for 1 min and dried briefly on 3 MM paper (Whatman, USA). Afterward, the transferred DNA was fixed on the membrane by illumination for 40 sec at 1.2×10^5 μJ/CM^2, using a UV cross-linker (UVP, USA). Labeling of the DNA probes was performed with the ECL Direct Nucleic Labeling System (Amersham Pharmacia Biotech). The labeled DNA was then added to the reaction buffer, and hybridization was carried out for 12 to 15 h at 42°C. The membrane was washed three times at 55°C, for 10 min each, using 5X SSC containing 0.1% (w/v) SDS, and twice (10 min each) at room temperature with 0.5X SSC before detection. Hybridized signals were generated using ECL direct detection reagents (Amersham Pharmacia Biotech). They were then visualized by pre-flashed X-ray film (AGFA), with exposure ranging from 1 min to overnight, depending on the strength of the signal from the first exposure.

Subcloning and Sequencing of Phage DNA Carrying CsPCK

λ*cspck*-4A and λ*cspck*-11B were propagated on Petri dishes (Φ 15 cm) to near-confluence. The phages were then suspended in 5 ml of SM buffer, precipitated, and extracted (Sambrook et al., 1989). Purification was carried out with a Qiagen Lambda Mini Kit (Qiagen GmBH). The phage DNA was subcloned into the plasmid vector pBlueScriptII SK, according to standard procedures (Sambrook et al., 1989). Purified λ*cspck*-4A and λ*cspck*-11B DNA was digested with *Sal*I, *Sac*I, and *Not*I, then fractionated on a 0.8% agarose gel and transferred to a Hybond-N^+ membrane for Southern blot hybridization. The DNA fragment that showed a positive signal was isolated from the gel and again subcloned into pBlueScriptII SK. Nucleotide sequencing of the DNA was conducted by ABI prism (Applied Bioscience, USA) at the Macrogen Co. (Korea). Plasmid subclones that carried a partial *CsPCK* gene were sequenced via primer-walking, using a *CsPCK* gene-specific internal primer and T7 and T3 primers (see the following section). This sequence was given the accession number AF481231.

Oligonucleotide Primers

All primers were synthesized by the Bioneer Co. (Korea), except for T7 and T3, which were supplied by the Macrogen Co. The *CsPCK* gene-specific primers included: 5′-GGGTGAGTAATGGGTATTTATAAG-3′ (5R-1); 5′-CCTGGACCATGGTTAAGTAAAACT-3′ (5R-2); 5′-AGATGGATGAAAGTTGGTTTTAGC-3′ (5R-3); 5′-AAGTTACCTCCAATATTGTAACCT-3′ (5R-4) for the 9-kb *Sac*I fragment subclone; 5′-GTAATTCTTCTTG-CTTGTGATGCA-3′ (3R-2); 5′-GCATTATCTCATGCCG-ATGCGCCA-3′ (3R-1); 5′-AGATTGTGAATAGCAGA-TTCTGCT-3′ (3F-2) for the 3.4-kb *Sac*I fragment subclone; 5′-ATCCACAAATGTCTTCTTTTAAATGGT-3′ (2F-1); and 5′-ACGAGAAAGGGTCATTCATAACGT-3′ (2R-1) for the 3.0-kb *Sac*I fragment subclone.

Computer Analysis of the DNA Sequences

General analysis and an homology search of the sequenced DNA was carried out by BLAST (Altschul et

al., 1990). A comparative analysis was performed at the following web sites: NCBI (http://www.ncbi.nlm.nih.gov/BLAST/), CMSMBR (http://restools.sdsc.edu/cmshp.html), and UK HGMP Resource Centre (http://www.hgmp.mrc.ac.uk). In addition, the TRES (http://bioportal.bic.nus.edu.sg/tres/) comparative promoter analysis program, TRANSFAC (http://transfac.gbf.de/TRANSFAC/) eukaryotic promoter database, PlantCARE (http://oberon.rug.ac.be:8080/PlantCARE/index.html) plant promoter database, CMSMBR, and Cister (http://zalb.bu.edu) were used to locate *cis*-acting elements in the promoter sequences.

RESULTS

Isolation of Cucumber CsPCK Genomic Clones

In the first round of screening, the 5′ end of the 0.23-kb *Pst*I fragment (containing a 5′ UTR with an ORF for the 64 amino-acid portion of the N-terminal) was hybridized to membrane-bound phage DNA. The first probe produced 14 positive signals from about 5 × 10^5 pfu. After de-probing from the membrane, a second probe (1.4 kb of the *Eco*RI/*Xho*I fragment; Fig. 2), which spanned between the 5′ end and the middle of the *PCK* cDNA, was hybridized to the same membrane. This second probe produced 36 positive signals, for a total of eight positive signals that overlapped from the two hybridization reactions. These double-hybridized positive plaques were used for second- and third-round screenings of the library. After the third round, two putative clones (i.e., λ*cspck*-4A and λ*cspck*-11B) were chosen for further analyses.

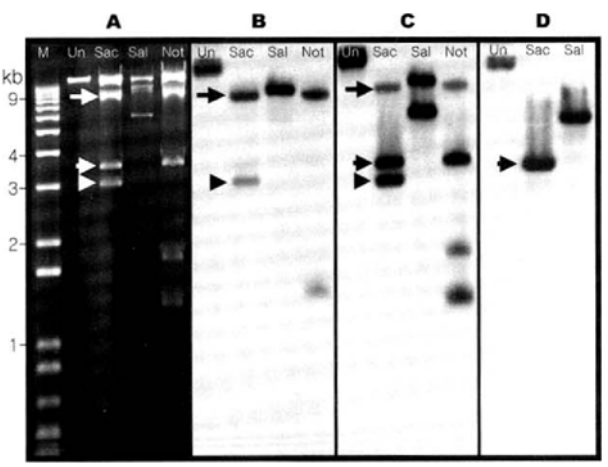

Figure 1. Restriction enzyme digestion and Southern blot analysis for ëc*spck*-4A clone to examine restriction pattern and determine appropriate subcloning procedure. Phage DNA was purified and digested by Lambda Fix II vector unique restriction enzymes (Sac; *Sac*I, Sal; *Sal*I and Not; *Not*I). DNA was size-fractionated on a 0.8% agarose gel (**A**), then Southern-blotted onto a Hybond-N^+ membrane for hybridization by three different ranges of *PCK* cDNA probes, as in genomic library screening. Three sets of blots were hybridized in separate reactions, using 0.23 kb of labeled *Pst*I-digested cDNA fragments (**B**), 1.4 kb of *Eco*RI/*Xho*I-digested cDNA fragments (**C**), and 0.7 kb of *Bam*HI-digested cDNA fragments (**D**) as probes. No *Not*I-digested DNA was included in Blot (**D**) since a well was not available on the agarose gel. M, 1-kb DNA ladder (Gibco-BRL); UN, undigested DNA. Three sized arrows represent relevant DNA fragments with corresponding signals, which were hybridized by designated cDNA probes.

Based on restriction mapping and Southern blot analysis (Fig. 1), we had expected a 15.4-kb insert DNA, containing λ*cspck*-4A, to encompass the entire *PCK* gene. Instead, our restriction enzyme *Sac*I produced three fragments 9.0-kb, 3.4-kb, and 3.0-kb -- of insert DNA from λ*cspck*-4A (Fig. 1A). Moreover, applying the *Sac*I enzyme to λ*cspck*-11B produced 6.0-kb, 3.4-kb, and 2.8-kb inserts, without the 9-kb fragment found with λ*cspck*-4A. Furthermore, the *Sac*I-digested λ*cspck*-11B DNA did not show any hybridization signal from the first probe (0.23 kb *Pst*I cDNA). We later determined that λ*cspck*-11B contained only a truncated, incomplete *CsPCK* gene (Fig. 2). Nevertheless, its DNA fragments were used in subcloning and partial sequencing, and the plasmid subclones showed complete identity with the λ*cspck*-4A internal sequences from the 5′ end of the second intron. Therefore, we believe that λ*cspck*-4A carries the complete cucumber *CsPCK* DNA sequence.

Copy Number of the CsPCK Gene in Cucumber

Southern blot analysis revealed simple patterns of hybridization signals, with a few minor bands in the Korean cucumber cultivar (Fig. 3). Selected λ*cspck* clones also manifested unique restriction patterns (data not shown). Partial sequencing of the subclones also showed complete homology with λ*cspck*-4A DNA. The number of restriction enzyme recognition sequences of the *CsPCK* gene entirely agreed with the number of bands found with our Southern blots. The full *CsPCK* gene sequences were entered into MBS SeqCUTTER (http://www.mbshortcuts.com) to identify known restriction enzyme recognition positions. Here, we found five *Eco*RI sites (at 2145, 2542, 7030, 7266, and 7405 bp from the 5′ end), two each for *Bam*HI (5523 and 7236 bp) and *Hind*III (4404 and 7574 bp), but none for *Xba*I. We had previously suggested that the cucumber

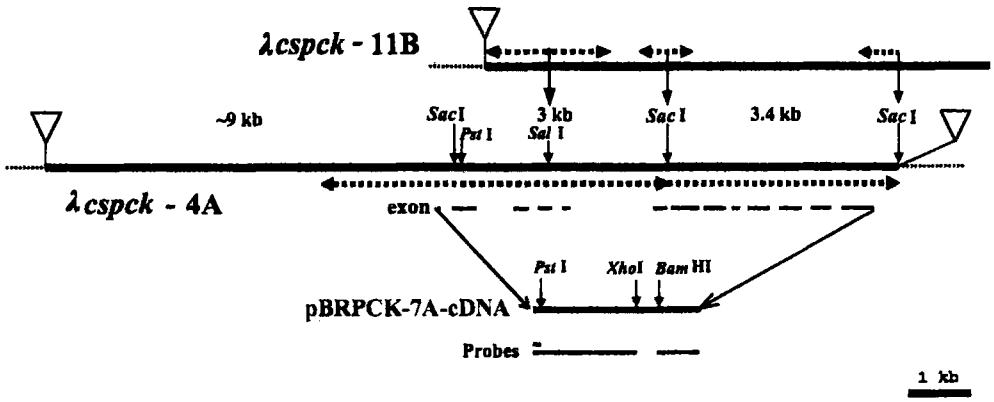

Figure 2. Genomic organization of the *CsPCK* gene in cucumber. Schematic representation of physical map from overlapping genomic lambda clones *cspck*-11B and *cspck*-4A with *Sac*I and *Sal*I recognition sites, according to the Southern hybridization in Fig. 1. Thick, straight lines indicate extension of the cucumber genomic DNA around *CsPCK* gene. Arrows with thick dashed lines show direction of DNA sequencing for subcloned genomic DNA fragments. Exons are indicated by thin broken lines. The pBRPCK-7A is a cucumber *PCK* cDNA (Kim and Smith, 1994) used for probe preparation. Reverse triangles represent cloning site in the Lambda Fix II vector (Stratagene), which contained unique *Not*I, *Sac*I and *Sal*I sites.

genome has a single copy of the *CsPCK* gene (Kim and Smith, 1994), a statement that was confirmed in our analyses in the current study. The rice genome-sequencing project has also demonstrated that a single copy of the *PCK* gene exists on Chromosome 10 of that species. This phenomenon may reflect a similar molecular structure between the two plant systems. Moreover, regulation of their *CsPCK* gene expression may be alike during development.

Sequence and Organization of the CsPCK Gene in Cucumber

We selected the λ*cspck*-4A clone for further analysis of the complete genomic sequence of the *CsPCK* gene. Unique restriction enzymes (*Sac*I, *Sal*I and *Not*I) first were applied to purified λ*cspck*-4A DNA to cut out the inserts and measure the insert genomic DNA. Size fractionation of the enzyme digests on the agarose gel showed relatively simple band patterns, with the total length of the insert DNA estimated at about 15.4 kb (Fig. 1A). The fractionated DNA was then Southern-blotted onto a Hybond-N^+ membrane for hybridization with a range of *PCK* cDNA probes (Fig. 1B, C, and D). Strong signals were revealed at 9 kb + 3 kb from *Sac*I, and at 10 kb from *Sal*I, using the 5′-end cDNA probe (0.23 kb of the *Pst*I fragment; Fig. 1B). The *Not*I-digested DNA produced a strong signal at 9 kb; a weak one, at 1 kb. This indicated that λ*cspck*-4A contains the 5′ end of the coding region for PEPCK and, most likely, a 5′ flanking region for the *CsPCK* gene (Fig. 1B).

A second hybridization with a longer probe (1.4 kb of the *Eco*RI/*Xho*I fragment) resulted in stronger signals, as well as additional signals from each restriction enzyme. Therefore, we believe that the supplementary 3.4 kb on *Sac*I, 6.0 kb on *Sal*I, and the 3.5 kb + 1.8 kb on the *Not*I fragments should contain the middle section of the *CsPCK* gene (Fig. 1C). Here, two *Sac*I fragments (9 kb and 3 kb) corresponded to the upstream region of the promoter. In addition, 3.4 kb of the *Sac*I fragment expanded the internal section of the *CsPCK* gene for the PEPCK peptide, from the middle to the C-terminal.

The third hybridization, using the 3′-end cDNA probe (0.7 kb of the *Bam*HI fragment) produced signals at 3.4 kb of *Sac*I and at 6.0 kb of *Sal*I. We concluded that these fragments should contain a part of the 3′ end of the *CsPCK* gene for the C-terminal coding region of PEPCK (Fig. 1D). Based on this, we subcloned three hybridized *Sac*I fragments (9.0 kb, 3.4 kb, and 3.0 kb) into bacterial plasmid vector pBluescript II SK. After sequence alignment, we obtained 8646 bp of the genomic DNA sequence and directly compared it with the cucumber *PCK* cDNA sequence (Accession no. L31899). The genomic clone λ*cspck*-4A contained the complete cucumber *CsPCK* gene, including the 5′ end of the flanking region plus a 3′ UTR (Fig. 2).

Exon-intron junctions were also determined so that a precise molecular structure for the *CsPCK* gene could be confirmed by GENSCANW, http://genes.mit.edu (Fig. 4 and Table 1). This gene comprised 13 exons and 12 introns, although the first exon contained only 23

Table 1. Exon-intron sturctures and boundaries of the *CsPCK* gene. Capital letters are intron sequences and lower case letters are exon with coding amino acids under the DNA sequences.

Exon size (bp)	3′ Acceptor site sequence	Exon sequence	5′ Donor site sequence
1 (23)		CAACTTCTCTTCctcttatccatcacaaatctccg**GT**ACCCACCTAC	
2 (302)	GTGTTGC**AG**gaggatacaaaaaatg	.....tatctcaatcag**GT**ATGTATATCC M I S I S	
3 (182)	GAAACGTGGC**AG**tgcttcattggc	.....cctctcccgccg**GT**AAGAACTGTC A S L A S P A E	
4 (154)	GGAAAACGAC**AG**agctttacgagc	.....ttggtggggcaa**GT**AAGAACATCC L Y E Q W W G K	
5 (85)	TTTGTGTTGT**AG**gggatcacctaa	.....tcccttgataag**GT**AAAAAAATAT G S P N S L D K	
6 (101)	CTTTTTTTCC**AG**gtatttgtgaat	.....catgcacaacat**GT**ATGCAAGATA V F V N M H N M	
7 (299)	TTTTGATGGC**AG**gtgcattcgacc	.....ttggattatcag**GC**ATGTCCTATT C I R P G L S G	
8 (174)	ATGTTTCTAT**AG**gtactgggaaga	.....agttcgggactg**GT**AACCTCCAAA T G K T F G T V	
9 (66)	TTTTGGCTGC**AG**ttcttgagaatg	.....aatccgttacag**GT**ATTAATTAGA L E N V S V T E	
10 (179)	CTGTCGATAC**AG**agaacactcgag	.....tacactgctttg**GT**ATTATTTACT N T R A Y T A L	
11 (170)	TTGCAAAATC**AG**gtggctggaact	.....gtcaggaggaag**GT**ACAACATAAC V A G T S G G S	
12 (169)	ACTCTTCTGC**AG**ctatggaagtgg	.....ccaataaacacg**GT**TAGTTTGATT Y G S G P I N T	
13 (144)	TTTGAACCAC**AG**tggtcagacaaa	.....cctaccttgtaaTAAAATGGTTAA W S D K P T L *	

bp of 5′ UTR. The UTR was estimated to be approximately at the transcriptional start site from cDNA, most likely being longer than 23 bp of the non-coding first exon. A similar structure for a non-coding Exon 1 was seen in the *Arabidopsis* Atg4 *PCK* gene, which contained 107 bp of the 5′ UTR sequence (Fig. 4). In addition, the *Arabidopsis* Atg5 *PCK* gene possibly had a non-coding Exon 1, as in cucumber, because GENESACNW predicted 40 bp of a non-coding exon in that region (data not shown). The 3′ UTR region was defined as being downstream of the polyadenylation sites from the *PCK* cDNA.

We compared the molecular organization of the *CsPCK* gene with four other *PCK* genes from two other plant species -- one from rice and three from *Arabidopsis* (Fig. 4). The overall exon organization of the *CsPCK* gene was very similar to Atg4 and the rice *PCK* gene, except for the beginning three exons for the N-terminal portion of PEPCK. Although the length of the following exons was almost identical in all three species, two *Arabidopsis* Atg5 *PCK* genes contained two longer exons, due to a reduction in their introns. However, this did not significantly change the length of the PEPCK peptide. The scale of the introns also differed somewhat among all four *PCK* genes. Cucumber has relatively longer intron sequences, the longest being Intron 5, at 1.6 kb (Fig. 4). In addition, an examination revealed that the exon-intron boundaries for the *CsPCK* gene AG (acceptor-site) and GT (donor-site) positions were conserved except for the seventh intron, which replaced the donor-site signal with GC (Table 1). Although we first considered this to be a sequencing error, repeated sequencing and alignment confirmed that the single signal nucleotide was replaced.

The primary structure of the C3 plant PEPCK showed high variability within the 140 amino-acid N-terminal region of the peptide sequences. These variations between cucumber and rice arose from Exons 1 and 2 in their corresponding genes. Cucumber and rice PEPCKs consisted of 670 amino acids (Protein Id: P42066) and 636 amino acids (Protein Id: AAM18765), respectively. The cucumber PEPCK shared 78% sequence homology with that of rice even though the former contained a longer N-terminal portion (i.e., 44 amino acids). The overall primary structure of the cucumber PEPCK peptide also showed the same level of homology

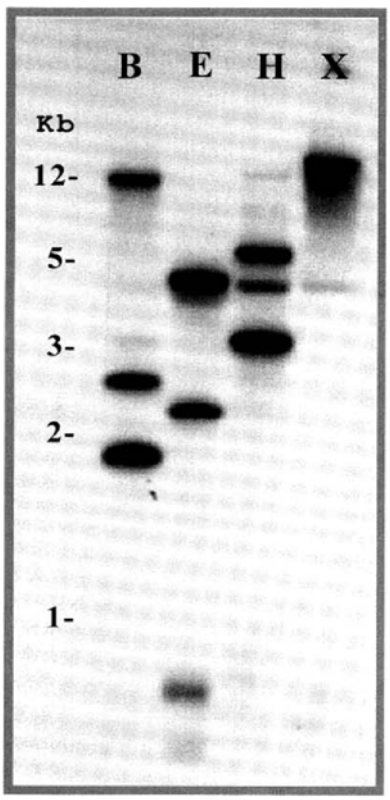

Figure 3. Southern blot hybridization of cucumber genomic DNA by full-length *PCK* cDNA probe. 20 μg of genomic DNA was used in each lane. Cucumber leaf DNA was completely digested with *Bam*HI (B), *Eco*RI (E), *Hind*III (H) and *Xba*I (X), then fractionated on a 0.8% agarose gel and transferred to a Hybond-N^+ membrane using 20X sodium saline citrate (SSC). Hybridization was carried out using ECL 2.4. kb of direct-labeled *PCK* cDNA for16 hr in ECL-gold buffer (Amersham Pharmacia Biotech), then washed at low stringency [0.5 x SSC (pH 7.0), 0.1% SDS without urea, at 55°C]. After a final wash in 2 x SSC, the hybridized signal was generated by ECL direct detection reagents. The membrane was exposed overnight to pre-flashed X-ray film (AGFA).

(78%) with an Atg4 *PCK*-encoded PEPCK of 671 amino acids (Protein Id: CAB80452, http://rarge.gsc.riken.go.jp).

Arabidopsis PEPCK can be separated into two classes according to peptide length and sequence variation in the N-terminal extension. The first class is a larger cucumber type (670 amino acids; Protein Id: At5g65690); the second, a smaller type having 628 amino acids (Protein Id: CAA16690). These differ from those found with C4 *Urochloa panicoides* and *Zea mays* (Kim and Smith, 1994; Finnegan and Burnell, 1995; Finnegan et al., 1999; Furumoto et al., 1999). Interestingly, two *Arabidopsis* PEPCKs were linked in a chromosome (Atg5), and the intron-exon structures of the two *PCK* genes were almost identical, except for three longer exons (numbers 1, 2 and 7) and two shorter introns (numbers 1 and 7) in the 670 amino acid-encoding *PCK* gene (Fig. 4). Moreover, two different classes of the PEPCK peptide from Atg5 *PCK* genes shared 93% homology. Their 7% diversity arose from the missing 38 amino acids (85~122) in the N-terminal region, with four amino acids also being absent in the C-terminal region. Compared with cucumber, two *Arabidopsis* Atg5 PEPCK peptides showed less homology (76 and 73%) to 670 amino-acid Atg5 PEPCK and 628 amino-acid Atg5 PEPCK, respectively. None of the C4 species contained such an extended N-terminal sequence in the PEPCK peptide, possibly a reflection on the regulation of PEPCK enzyme activity between C3 and C4 plants. Therefore, this N-terminal extension may be necessary for the regulation of the PEPCK enzyme in C3 species. However, the genetic and enzymatic control of this enzyme should be further examined in both types of plants.

Conserved Plant-Specific Cis-Acting Elements in the CsPCK Gene

Initially, we compared 2 kb of the *CsPCK* promoter sequences with others present in the eukaryotic promoter databases (EPD) from NCBI (Altschul et al., 1990). In the second stage, we adopted the web-supporting TRES comparative promoter analysis program, TRANSFAC eukaryotic promoter databases, and PlantCARE plant promoter databases. *CsPCK*, *icl*, and *ms* promoter sequences were then examined to identify common conserved *cis*-acting elements. The first TATA signal (5'-TATAAAT) was seen 63 bp upstream (or −216 bp from the translation initiation point) from the relative transcription initiation site (Fig. 5). That site was deduced from cucumber *PCK* cDNA (Kim and Smith, 1994). Two conserved CAAT boxes were found 10 bp (−226 bp) from the 5' TATA box, and 24 bp downstream (−188 bp) from the 5' TATA box. However, a CAAT box usually was located about 40 nucleotides from the TATA box. An AT-rich palindrome sequence showed 90% homology from −1131 bp to −1092 bp in the forward and reverse directions (Figs. 5 and 6A). Moreover, a pair of inverted repeat sequences (92% homology) were positioned between −1200 bp~−1173 bp and -998 bp~−971 bp (Fig. 6B). We believe these palindrome and inverted sequences may function in the control of *CsPCK* gene expression, either as recognition sequences for *trans*-acting factors or in the formation of secondary structures.

We identified various transcription factors for binding

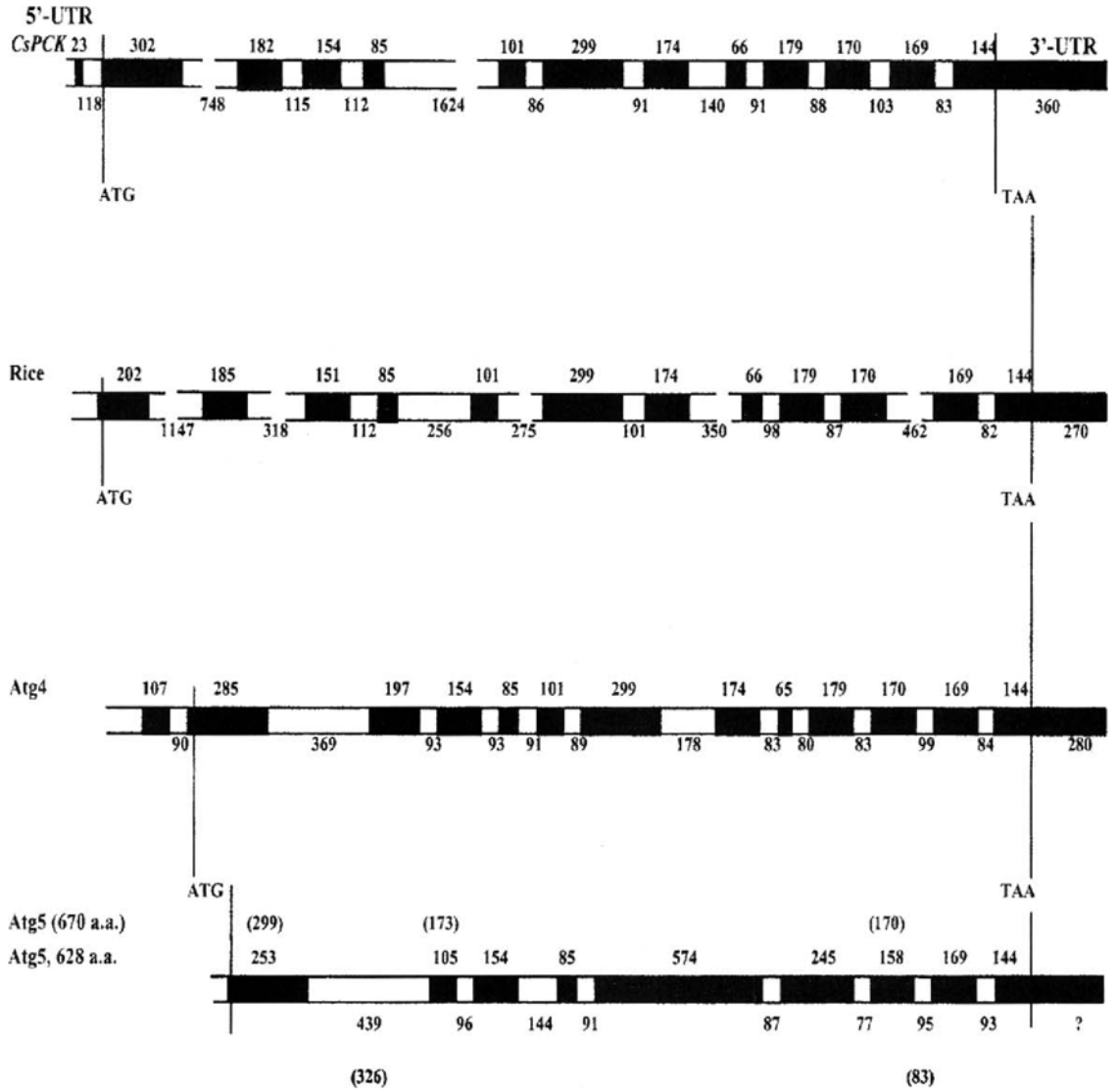

Figure 4. Molecular structure of *PCK* genes from cucumber (*CsPCK*), rice, and *Arabidopsis* (Atg4 and Atg5). Except for cucumber, four plant *PCK* gene sequences were obtained from NCBI. Accession numbers include AL161592 for Atg4 *PCK* gene (Chromosome 4), AL021684 and NP_680468 for linked Atg5 *PCK* genes (Chromosome 5) from *Arabidopsis*, and AC099325 for the rice *PCK* gene on Chromosome 10. Distances shown are in base pairs. Numbers in parentheses with Atg5 represent 670 a.a. PEPCK encoded the change in lengths for exons and introns of *PCK* gene. Position of transcription start site (thin, straight line) was deduced from cDNA sequences, so the size of the 5 UTRs is not precise. The 3 UTR of *Arabidopsis* also is not precise, but the coding sequences in the last exons are equal (144 bp) for all five *PCK* genes. Black boxes indicate exons and numbers refer to length (in nucleotides) for both exons and introns. Lengthy introns were interrupted by broken lines for alignment in cucumber and rice.

plant-unique *cis*-acting elements within 1 kb of the *CsPCK*, *icl*, and *ms* promoters (Table 2). The GATA-motif (as its anti-sense sequence) for the sugar repression-responsible rice alpha-amylase gene (*Ramy3D*)-specific *cis*-acting elements was detected 3 bp downstream from the relative transcription initiation site (–151 bp), which was similarly situated from the relative translation initiation site (Fig. 5). This motif has been located in the first exon with an I-Box (5'-GATAAG) on a complementary sequence (–146 bp~–151 bp), using an overlapping mode (Donald and Cashmore, 1990; Rose et al., 1999). Furthermore, a plant promoter unique to the Dof transcription factor that binds three pyrimidine boxes (5'-CCTTTT) has been reported at –412

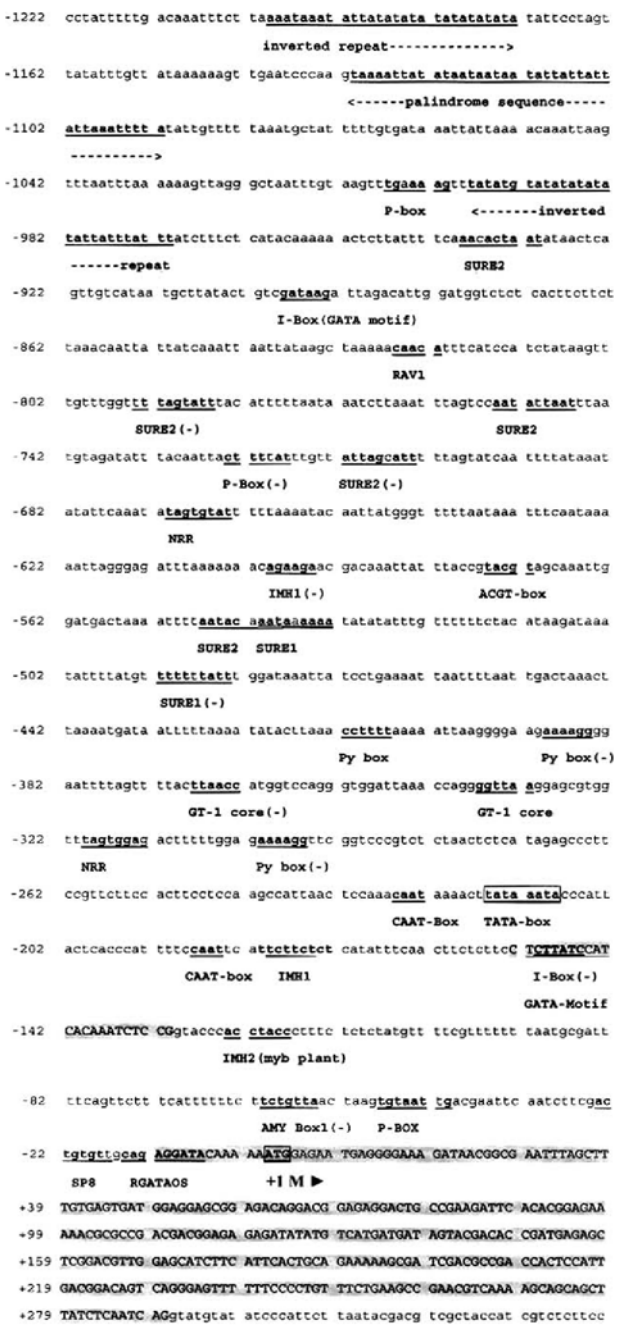

Figure 5. 5′ flanking region DNA sequences from the *CsPCK* gene. Start codon (ATG) is boxed and in bold type. Capital letters are exons; lowercase letters, intron sequences. First and second exons are highlighted with shaded boxes. The first non-coding exon starts from –153 bp to –130 bp, relative to the translation start site. Putative TATA-box is located –216 bp from the relative translation initiation site (63 bp upstream from the relative transcription initiation site). Matching *cis*-acting core sequences are in bold type. Matching *cis*-acting sequences are also summarized in Table 2. +1 indicates relative translation initiation site. M, methionine.

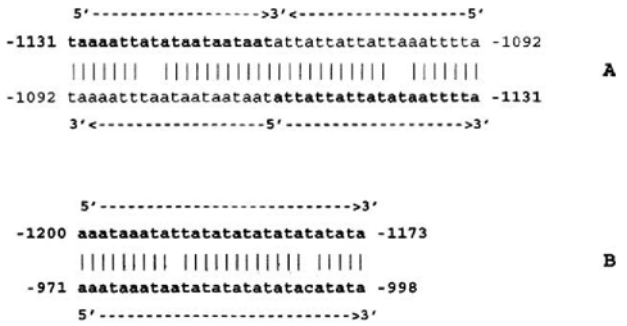

Figure 6. A palindrome (**A**) and the inverted repeat sequences (**B**) of the *CsPCK* promoter region. Numbers begin from the 5′-end of the *CsPCK* gene sequences.

bp, –390 bp, and –301 bp (Mena et al., 1998). Most likely, one of those boxes is partially responsible for sugar repression under the control of the *CsPCK* gene as well as the rice *RAMY* gene (Morita et al., 1998).

General sucrose-inducible SURE elements are well conserved in overlapping regions, at –547 bp for SURE2, and at –541 bp and –492 bp for SURE1 (Grierson et al., 1994). Here, we identified four additional SURE2 elements within the 1-kb 5′ flanking region (i.e., at –939 bp, –794 bp, –755 bp, and –712 bp) of the *CsPCK* promoter (Fig. 5). The SURE1 element was also conserved in both the *icl* and the *ms* promoters, whereas SURE2 sequences were seen only in the *ms* and *CsPCK* promoters (Table 2). Light activation and calcium/calmodulin stimulation can result in two GT-1 binding boxes (5′-GGTTAA) being sandwiched between two pyrimidine boxes, from –337 bp on the coding strand and from -368 bp on the complementary strand (Lam and Chua, 1990; Gourrierec et al., 1999; Zhou, 1999). These GT-1 core elements are possibly involved in the expression of both *CsPCK* and *ms* during cucumber development, as shown by the transcription of multiple genes in various organs (Eyal et al., 1995; Buchel et al., 1996; Lopes Cardoso et al., 1997). Together with a I-box, eight GT-1 boxes were found in thaumatin-like gene (*MdTL1*) promoter in apple (Kim et al., 2003). Therefore, it can be assumed that these elements in many plant promoters are responsible for light- or carbohydrate-responsive expression of the related genes.

The recognition motif (5′-CAACA) for the *Arabidopsis* novel DNA-binding protein RAV1 (homologous to the AP2 DNA-binding domain) can be seen only at –826 bp in the *CsPCK* promoter (Kagaya et al., 1999). Plant bZIP proteins that bind the ACGT *cis*-acting element (A-Box) may be functional in both *CsPCK* (–576 bp) and *icl* (–790 bp), in response to diverse environmen-

Table 2. Summary of common plant specific *cis*-acting elements in *icl*, *ms* and *CsPCK* genes which are involved in lipid mobilization during early germination and senescence. Descriptions of *cis*-acting elements are as follows under the table. * single nucleotide mismatch in the core sequences. ** matching signals are located further upstream before−1 kb region.

Name of family	Sequences	Genes	Position (+/−)	Reference
A-Box	ACGTA	*CsPCK*	−576(+/−)	Toyofuku et al. (1998)
		icl	−790(+/−)	
		ms	−	
AMY Box1	TAACARA	*CsPCK*	−717(−), -61(−)	Huang et al. (1990), Morita et al. (1998)
		icl	−406(+), -282(+)	
		ms	−758(+)	
ERE(E4)	AWTTCAAA	*CsPCK***	−	Itzhaki et al. (1994)
		icl	−238(+)	
		ms	−789(−)	
GT-1 (Box II)	GGTTAA	*CsPCK*	−368(−), -337(+)	Zhou (1999)
		icl	−	
		ms	−510 (−)	
I-Box	GATAAG	*CsPCK*	−899(+), −151(−)	Donald and Cashmore (1990)
		icl	−182(−)	Rose et al. (1999)
		ms	−774 (+)	
NRR	TAGTGGAT*	*CsPCK*	−671(+), −320(+)	Elliott and Shirsat (1998)
		icl	−	
		ms	−976(−), −197(+)	
NtBBF1	ACTTTA	*CsPCK*	−	Baumann et al. (1999)
		icl	−465(+)	
		ms	−421(+)	
Pyrimidine Box	CCTTTT	*CsPCK*	−412(+), −390(−), -301(−)	Morita et al. (1998), Mena et al. (2002)
		icl	−	
		ms	−931(+)	
RGATAOS	CAGAAGATA*	*CsPCK*	−513(+), -15(+)	Yin et al. (1997)
		icl	−513(−), -258(+)	
		ms	−955(+)	
SP8	ACTGTGTA*	*CsPCK*	−24(+)	Ishiguro and Nakamura (1992, 1994)
icl			−576(+)	Kim et al. (1997)
		ms	−832(+), −824(−)	
SURE1	AATAGAAAA*	*CsPCK*	−541(+), 492(−)	Grierson et al. (1994)
		icl	−734(+)	
		ms	−304(−), −155(−)	
SURE2	AATACTAAT*	*CsPCK*	−939(+), 794(−), 755(+), −712(−), 547(+)	Grierson et al. (1994)
		icl	−	
		ms	−485(+)	
TCA1	TCATCTTCTT	*CsPCK*	−	Goldsbrough et al. (1993)
		icl	−361(−)*	
		ms	−253(−)	

A-box, ACGT elements in *RAMY3D* gene for alpha amylase of rice; AMY-box, gibberellin-response *cis*-elements (or GARE) in alpha amylase gene promoter; ERE (E4), ethylene responsive element in the senescence-related gene expression; GT-1, Box II promoter *cis*-acting element binding for many light regulated plant genes; I-box, conserved regulatory core sequences upstream of light-regulated genes (e. g. *rbcsS*); NRR; negative regulatory sequences in promoter of *Brassica napus extA* (extensin) gene; NtBBF1, a Dof protein from tobacco binds to tissue specific and auxin regulated *rolB* oncogene promoter in *Agarbacterium rhizogenes*; Py, pyrimidine box for partially responsible for sugar repression in alpha amylase gene *RAmy1A* (Dof1 binding); RGATAOS, promoter elements for phloem-specific gene expression from the *Tungro bacilliform* virus (RTBV) in rice; SP8, conserved 5'-upstream region regulatory sequences in sweet potato three different genes (sporamin and beta-amylase) which interact with binding factor SP8BF; SURE, sucrose responsive elements first found in potato tuber protein *patatin* gene promoter; TCA1, tobacco nuclear protein 1 binding *cis*-acting element in stress-inducible genes by salicylic acid-inducible expression of many genes.

tal and physiological stimuli (Foster et al., 1994). Five IMH (ICL-MS Homology) sequences have been suggested for the *icl* and *ms* genes in cucumber, but only two potential IMH-related sequences have been observed in the *CsPCK* promoter (Sarah et al., 1996). The IMH2-homologous Myb-binding sequences (ACCA/TACC),

as well as the probable germination-responsive core element (TCTTCT) IMH1, are found at −124 bp and −180 bp, respectively (Sarah et al., 1996). We were unable to identify IMH sequences from the promoter sequence databases because no assignment has been made for them yet.

Analysis of the *CsPCK* promoter region, together with *icl* and *ms*, revealed few common sugar response-related *cis*-acting sequences (such as AMY-box I, I-Box, SP8, or SURE1) within 1 kb of the 5 flanking region (Table 2). The SP8 regulatory sequence-binding SPF-1-type protein (a WRKY family), which encodes the cDNA sequence, has been reported from cucumber cotyledons (Kim et al., 1997). Since homologous SP8 *cis*-acting sequences are conserved in three cucumber promoters, the *CsPCK* gene can be controlled during germination through the coordination of the *icl* and *ms* genes. The response to senescence by *CsPCK* may also be associated with the activity of the *icl* and *ms* genes during cotyledon development in that species (Kim and Smith, 1994). An ethylene-responsive element (ERE or E4), involved in senescence-related *GST1* gene expression in carnation (Itzhaki et al., 1994), was found within the previously described 1-kb region in both *icl* and *ms*, but not in *CsPCK*. However, three conserved ERE elements locate further upstream of the promoter region, at −2089, −1872, and −1546 bp.

DISCUSSION

Cucumber PEPCK is encoded by a single gene throughout plant development, from germination to cotyledon senescence. Our genomic DNA Southern analysis detected simple hybridization bands, confirming the presence of that single gene. Although a few additional weak bands were detected in this study, the recognition positions for restriction enzymes agreed with the number of Southern-detected bands. Therefore, we have further demonstrated that a single *CsPCK* gene is active throughout development. In contrast to that species, C4 grasses, such as *Urochloa panicoides*, contain *PCK* multigene families with at least four members (Finnegan et al., 1999). This multigene encoding of the subunits for PEPCK may be ubiquitous in higher plants. For example, the sequencing project for *Arabidopsis* has revealed three *PCK* genes in its genome (NCBI Accession nos. AL021684, NP_680468, and AL161592). Likewise, three species in the genus *Brassica* (*B. napus, B. campestris,* and *B. oleracea*) contain several *PCK* genes (Saez-Vasquez et al., 1995). Nevertheless, genome sequencing of rice, a C3 plant, has shown that a single *PCK* gene (NCBI Accession no. AC099325) exists on Chromosome 10. Therefore, our results add to the information available for the biochemical study of PEPCK, as well as the molecular genetic investigation of *PCK* genes in two C3 species.

We examined the entire genomic structure of the *CsPCK* gene to elucidate its molecular organization in cucumber. Since the first report was made on the single, complete cDNA sequence (Kim and Smith, 1994), several *PCK* cDNA sequences have been identified from the C4 grass *U. panicoides* (Finnegan and Burnell, 1995), as well as *Brassica napus* (Saez-Vasquez et al., 1995), *Lycopersicon esculentum* (Bahrami et al., 2001), *Arabidopsis* EST sequences (T06034), *Zea mays* (AB018744), and partial sequences in *Medicago sativa* (AF212090). Genome sequencing of *Arabidopsis thaliana* has also revealed three complete *PCK* gene sequences and several uncharacterized *PCK*-related sequences. However it is unclear which *PCK* gene or genes are active during phases in the *Arabidopsis* life cycle. Although that C3 species contains the smallest genome, multiple-copy genes are necessary to ensure its survival. One or two of its *PCK* genes may be a pseudogene, with their activity having been halted at one time. However, if three of its *PCK* genes can be detected during plant development, the control of gene expression will be more complex compared with that for species, e.g., cucumber and rice, that contain only a single copy .

Although four *PCK* genes have been identified in C4 *U. panicoides*, no further information is known about its gene structure and regulation. Five complete *PCK* gene sequences are available for *Arabidopsis*, rice, and cucumber. A pair of linked *Arabidopsis PCK* genes (AL021684 and NP_680468) has 8 introns each on Chromosome 5 (Atg5) for 628 amino-acid and 670 amino-acid PEPCK peptides. Another *Arabidopsis PCK* gene (AL161591) consists of 13 exons and 12 introns on Chromosome 4 (Atg4). Similarly, a *PCK* gene containing 11 introns can be found on Chromosome 10 in rice (AC099325). The *CsPCK* gene, which contains 12 introns, is the longest *PCK* gene, spanning >9 kb in the *Cucumis sativus* genome. This gene is closely related to those found in major crops and weeds, from rice to *Arabidopsis*. However, the 5′ sequence of its coding region, which is responsible for the N-terminal peptide extension, is an exceptional configuration; N-terminal amino-acid sequences vary among plant PEPCKs. Such extensions are also found in two *Arabidopsis* PEPCK (Accession no. CAA16690 and CAB80452) peptide sequences, although they differ somewhat (Finnegan et al., 1999). Overall, cucumber and tomato (*Lycoper-*

sicon esculentum) share >81% sequence homology, the highest among plant species (Bahrami et al., 2001). The primary structure of a single PEPCK subunit comprises 670 amino acids in cucumber and 662 amino acids in tomato, with that difference in length mainly resulting from a smaller amount of N-terminal extension in the latter. The tomato *PCK* gene, however, is not yet available.

The control of plant gene expression is a complex biochemical process that is affected by numerous internal and external factors, including nutrient status, hormonal levels, light intensity, and temperature changes. Carbohydrates (i.e., sugars) are a key metabolic constituent of growth and development. Therefore, controlling intracellular sugar levels is a central element for the regulation of metabolic pathways in living cells. Many of the genes associated with carbon assimilation and sugar transport are directly or partly regulated by fluxes in the sugar levels of target cells and tissues (Sheen, 1990; Graham et al., 1994; Koch, 1996; Sherson et al., 2000). Because of this, the metabolic regulation of plant gene expression is now a central focus in studying the plant sugar-transport mechanism. Several sugar-sensing systems and signal pathways have been suggested for this process (Smeekens, 1998; Gibson and Graham, 1999). In fact, the promoters of many metabolism- and transport-related genes contain common and essential regulatory elements for controlling their corresponding genes and regulating expression at the transcription level. Some of these intrinsic critical factors include *cis*-acting elements in individual promoter sequences as well as regulatory elements.

Genes related to the metabolic pathway have been cloned and inspected to elucidate the regulation of expression during plant development (Sarah et al., 1996; Morita et al., 1998). In particular, sequence comparisons and deletion analyses have been conducted for the promoters of two glyoxylate cycle-specific genes, *icl* and *ms* (Sarah et al., 1996). In our cucumber study, two highly conserved germination-response elements (IMH1 and IMH2) were identified in the *CsPCK* gene as being necessary for germination and, possibly, senescence. Although the biological function of the IMH1 sequence is not yet clear, similar sequences have been shown in fungi (*Neurospora crassa*) and yeast (*Saccharomyces cerevisiae*), suggesting that it may be required for either a sugar response (starvation) or germination response (Graham et al., 1994). The IMH2 sequences (with similarity to MYB binding sequences) in the *icl*, *ms*, and *CsPCK* promoters might direct the carbohydrate regulation of expression for many plant genes (Graham et al., 1994; Grotewold et al., 1994). However, the IMH sequences from cucumber genes are apparently uncommon. Therefore, their biological roles should be further examined in other species.

Our investigation of the cucumber *CsPCK* gene revealed a number of possible *cis*-acting sequences within 1 kb of the 5′ flanking region, suggesting that this gene can be regulated in several ways. First, developmental control is a key factor for its expression during germination and senescence. The germination-responsive IMH elements may play some role in controlling the *CsPCK* gene in cucumber, as with the *icl* and *ms* genes. Second, either the palindrome or the inverted repeat sequences may help regulate *CsPCK* gene expression in cucumber. Third, the non-coding first exon may be involved in gene expression for both cucumber and *Arabidopsis*. Based on our sequence analysis, the cucumber *CsPCK* gene contained highly conserved *cis*-acting sequences within the putative promoter region; their functions were postulated based on those for other plant genes. Therefore, we are using deletion-analysis experiments to continue our study of the natural function for these *cis*-acting elements.

ACKNOWLEDGEMENTS

This research was supported by a grant (No. R05-2001-000-00568-0) from the Korea Science and Engineering Foundation to D.-J. Kim.

Received September 25, 2003; accepted November 6, 2003.

LITERATURE CITED

Altschul SF, Glash W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410

Bahrami AR, Chen Z-H, Walker RP, Leegood RC, Gray JE (2001) Ripening-related occurrence of phospho*enol* pyruvate carboxykinase in tomato fruit. Plant Mol Biol 47: 499-506

Baumann K, De Paolis A, Costantino P, and Gualberti G (1999) The DNA binding site of the Dof protein NtBBF1 is essential for tissue-specific and auxin-regulated expression of the *rolB* oncogene in plants. Plant Cell 11: 323-333

Buchel AS, Molenkamp R, Bol JF, Linthorst HJM (1996) The PR-1a promoter contains a number of elements that bind GT-1 like nuclear factors with different affinity. Plant Mol Biol 30: 493-504

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

Edwards GE, Kanai R, Black CC (1971) Phosphoenolpyruvate in leaves of certain plants which fix CO_2 by the C4 - dicarboxylic acid cycle of photosynthesis. Biochem Biophys Res Comm 45: 278-285

Elliott KA, Shirsat AH (1998) Promoter regions of the *extA* extension gene from *Brassica napus* control activation in response to wounding and tensile stress. Plant Mol Biol 38: 675-687

Eyal Y, Curie C, McCormick S (1995) Pollen specificity elements reside in 30 bp of the proximal promoters of two pollen-expressed genes. Plant Cell 7: 373-384

Finnegan PM, Burnell JN (1995) Isolation and sequence analysis of cDNAs encoding phospho*enol*pyruvate carboxykinase from the PCK-type C4 grass *Urochloa panicoides*. Plant Mol Biol 27: 365-376

Finnegan PM, Suzuki S, Ludwig M, Burnell JN (1999) Phospho*enol*pyruvate carboxykinase in the C(4) monocot *Urochloa panicoides* is encoded by four differentially expressed genes. Plant Physiol 120: 1033-1042

Foster R, Izawa T, Chua NH (1994) Plant bZIP proteins gather at ACGT elements. FASEB J 8: 192-200

Furumoto T, Hata S, Izui K (1999) cDNA cloning and characterization of maize phospho*enol*pyruvate carboxykinase, a bundle sheath cell-specific enzyme. Plant Mol Biol 41: 301-311

Gibson SI, Graham IA (1999) Another player joins the complex field of sugar-regulated gene expression in plants. Proc Natl Acad Sci USA 96: 4746-4748

Gerwick BC, Black CC Jr (1979) Sulfur assimilation in C4 plants. Plant Physiol 64: 590-593

Goldsbrough AP, Albrecht H, Stratford R (1993) Salicylic acid-inducible binding of a tobacco nuclear protein to a 10 bp sequence which is highly conserved amongst stress-inducible genes. Plant J 3: 563-571

Gourrierec JL, Li Y-F, Zhou D-X (1999) Transcriptional activation by *Arabidopsis* GT-1 may be through interaction with TFIIA-TBP-TATA complex. Plant J 18: 663-668

Graham IA, Leaver CJ, Smith SM (1992) Induction of malate synthase gene expression in senescent and detached organs of cucumber. Plant Cell 4: 349-357

Graham IA, Baker CJ, Leaver CJ (1994) Analysis of the cucumber malate synthase gene promoter by transient expression and gel retardation assays. Plant J 6: 893-902

Grierson C, Du JS, De Torres ZM, Beggs K, Smith C, Holdsworth M, Bevan M (1994) Separate *cis* sequences and *trans* factors direct metabolic and developmental regulation of a potato tuber storage protein gene. Plant J 5: 815-826

Grotewold E, Drummond BJ, Bowen B, Peterson T (1994) The myb-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. Cell 76: 543-553

Hatch MD (1997) Resolving C4 photosynthesis: Trials, tribulations and other unpublished stories. Aust J Plant Physiol 24: 413-422

Hibberd JM, Quick WP (2002) Characteristics of C4 photosynthesis in stems and petioles of C3 flowering plants. Nature 415: 451-454

Huang N, Sutliff TD, Litts JC, Rodriguez RL (1990) Classification and characterization of the rice alpha-amylase multigene family. Plant Mol Biol 14: 655•668

Ishiguro S, Nakamura K (1992) The nuclear factor SP8BF binds to the 5'-upstream regions of three different genes coding for major proteins of sweet potato tuberous roots. Plant Mol Biol 18: 97-108

Ishiguro S, Nakamura K (1994) Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato. Mol Gen Genet 244: 563-571

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

Kagaya Y, Ohmiya K, Hattori T (1999) RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. Nucl Acids Res 27: 470-478

Kim D-J, Smith SM (1994) Molecular cloning of cucumber phospho*enol*pyruvate carboxykinase and developmental regulation of gene expression. Plant Mol Biol 26: 423-434

Kim D-J, Smith SM, Leaver CJ (1997) A cDNA encoding a putative SPF-1 type DNA-binding protein from cucumber. Gene 185: 265-269

Kim S-H, Lee J-R, Kim S-R (2003) Molecular characterization of a fruit-preferential thaumatin-like gene from apple (*Malus domestica* cv. Fuji). J Plant Biol 46: 52-58

Koch KE (1996) Carbohydrate-modulated gene expression in plants. Annu Rev Plant Physiol Plant Mol Biol 47: 509-540

Lam E, Chua N-H (1990) GT-1 binding site confers light-responsive expression in transgenic tobacco. Science 248: 471-474

Lopes Cardoso MI, Meijer AH, Rueb S, Queiroz M, Memlink J, Hoge JHC (1997) A promoter region that controls basal and elicitor-inducible expression levels of the NADPH/cytochrome P450 reductase gene (*Cpr*) from *Catharanthus roseus* binds nuclear factor GT-1. Mol Gen Genet 256: 674-681

Mena M, Vicente-Carbajosa J, Schmidt RJ, Carbonero P (1998) An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. Plant J 16: 53-62

Morita A, Umemura T, Kuroyanagi M, Futsuhara Y, Perata P, Yamaguchi J (1998) Function dissection of a sugar-repressed α-amylase gene (*Ramy1A*) promoter in rice embryos. FEBS Lett 423: 81-85

Rose A, Meier I, Wienand U (1999) The tomato I-box binding factor LeMYBI is a member of a novel class of Myb-like proteins. Plant J 20: 641-652

Saez-Vasquez J, Raynal M, Delseny M (1995) A rapeseed cold-inducible transcript encodes a phospho*enol*pyruvate carboxykinase. Plant Physiol 109: 611-618

Sambrook J, Fritsch E, Maniatis T (1989) Molecular Cloning, A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory, New York

Sarah CJ, Graham IA, Reynolds SJ, Leaver CJ, Smith SM (1996) Distinct *cis*-acting elements direct the germination and sugar responses of the cucumber malate synthase gene. Mol Gen Genet 250: 153-161

Sheen J (1990) Metabolic repression of transcription in higher plants. Plant Cell 2: 1027-1038

Sherson SM, Hemann G, Wallace G, Fobes S, Germain V, Stadlier R, Bechtold N, Sauer N, Smith SM (2000) Monosaccharide/protein symporter AtSTP1 plays a major role in uptake and response of *Arabidopsis* seeds and seedlings to sugars. Plant J 24: 849-858

Smeekens S (1998) Sugar regulation of gene expression in plants. Curr Opin Plant Biol 1: 230-234

Thomas SM, Rees T (1972) Gluconeogenesis during the germination of *Cucurbita pepo*. Phytochem 11: 2177-2185

Toyofuku K, Umemura T, Yamaguchi J (1998) Promoter elements required for sugar-repression of the *Ramy3D* gene for -amylase in rice. FEBS Lett 428: 275-280

Trevanion SJ, Brooks AL, Leegood RC (1995) Control of gluconeogenesis by phospho*enol*pyruvate carboxykinase in cotyledons of *Cucurbita pepo* L. Planta 196: 653-658

Tsuyoshi F, Hata S, Izui K (1999) cDNA cloning and characterization of maize phospho*enol*pyruvate carboxykinase, a bundle sheath cell-specific enzyme. Plant Mol Biol 41: 301-311

Utter MF, Kolenbrander HM (1972) Formation of oxaloacetate by CO_2 fixation on phosphoenolpyruvate. In Boyer PD, ed, The Enzymes, Vol 6. Academic Press, New York, pp 117-168

Walker RP, Leegood RC (1995) Purification and phosphorylation *in vivo* and *in vitro* of phospho*enol*pyruvate carboxykinase from cucumber cotyledons. FEBS Lett 362: 70-74

Walker RP, Chen ZH, Tecsi LI, Famiani F, Lee PJ, Leegood RC (1999) Phospho*enol*pyruvate carboxykinase plays a role in interactions of carbon and nitrogen metabolism during grape seed development. Planta 210: 9-18

Zhou DX (1999) Regulatory mechanism of plant gene transcription by GT-elements and factors. Trends Plant Sci 4: 210-214