The Pennsylvania State University
The Graduate School
Intercollege Program in Plant Physiology

MOLECULAR ANALYSIS OF STARCH BRANCHING ENZYME GENES IN MAIZE (*ZEA MAYS* L.)

A Thesis in
Plant Physiology
by
Kyung-Nam Kim

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 1997

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

We approve the thesis of Kyung-Nam Kim

Date of Signature

Mark J. Guiltinan

May 28, 1997

Mark J. Guiltinan
Assistant Professor of
Plant Molecular Biology
Thesis Advisor
Chair of Committee

Jack C. Shannon

May 7 1997

Jack C. Shannon
Professor of Plant Physiology

Jill Deikman

May 7, 1997

Jill Deikman
Assistant Professor of Biology

David S. Gilmour

May 7, 1997

David S. Gilmour
Associate Professor of
Molecular and Cell Biology

Eva J. Pell

May 7, 1997

Eva J. Pell
Steimer Professor of
Agricultural Sciences
Chair of the Intercollege Graduate Program
in Plant Physiology

ABSTRACT

A full-length maize genomic clone was isolated and sequenced (-2190 to +5929) which contains the entire coding region of starch branching enzyme I (*SbeI*) as well as 5'-and 3'-flanking sequences. A consensus TATA-box and a G-box containing a perfect palindromic sequence, CCACGTGG, were found in the 5'-flanking region. The transcription initiation site of the *SbeI* gene was determined by primer extension analysis to be at an adenine nucleotide located 24 bp downstream from the TATA box. Studies using 3'-rapid amplification of cDNA ends (3'RACE) indicated that polyadenylation occurs 29 bp downstream from a putative polyadenylation signal (AATAAA) which is located 264 bp downstream from the translation stop codon. Sequence alignment with an *SbeI* cDNA revealed that the transcribed region consists of 14 exons and 13 introns, distributed over 5.7 kb. Southern blot analysis suggested that two *SbeI* genes with divergent 5'-ends of the coding region exist in the maize genome.

A transcriptional chimeric construct of the 5'-flanking region (-2190 to +27) fused to the luciferase gene (pKL101) showed promoter activity after it was introduced into maize endosperm suspension cells by particle bombardment. Addition of the first exon and intron of the *SbeI* gene (pKLN101) into the chimeric construct dramatically increased gene expression (14 fold). Transient expression assays of 5' deletion constructs revealed that a 60-bp region between -315 and -255 relative to the transcription initiation site is critical for promoter activity. Further analysis of the 60-bp region with linker-scan constructs identified two positive *cis*-regulatory elements located in regions, -315 to -295 and -284 to -256, respectively. Nuclear proteins binding to the

60-bp fragment containing these two elements were detected in nuclear extracts prepared from maize kernels by electrophoretic mobility shift assays.

Northern blot analysis demonstrated that the endogenous *Sbe1* mRNA level in cultured maize endosperm suspension cells is modulated by sugar concentrations in the media. Therefore, a transient expression system was used to determine whether the *Sbe1* promoter activity is regulated by sugar availability. The chimeric construct pKLN101 showed approximately two-fold greater LUC activity in sugar containing media than in sugar-free media. Further analysis showed that the promoter sequences between -314 and -145 are necessary for sugar regulation. Within the 170-bp sequence, the -314/-295 region displayed a striking similarity with the sugar-response elements of the potato patatin-1 and sporamin promoters, suggesting a common *cis*-acting sugar response element.

A maize genomic DNA fragment containing the entire coding region of starch branching enzyme IIb (*Sbe2b*) as well as 5'- and 3'-flanking sequences was isolated and sequenced (-2,964 to +15,556). A consensus TATA-box sequence was found 28 bp upstream of the transcription initiation site as determined by primer extension analysis. The complete genomic structure of the *Sbe2b* gene was established by alignment with the *Sbe2b* cDNA. The transcribed region of the gene is composed of 22 exons and 21 introns distributed over 15,347 bp. The introns vary in length from 76 bp to 3,051 bp, all of which have the conserved junction sequences (GT--AG). Although the canonical polyadenylation signal (AATAAA) was not found in the 3'-end of the gene, a similar sequence, AATTAAA, was observed 29 bp upstream of the polyadenylation site. Genomic Southern blot analyses suggested that a single *Sbe2b* gene is present in the maize genome. Promoter activity was confirmed with a transcriptional fusion of the *Sbe2b* 5'-flanking region between -2,964 and +100 to the luciferase gene, which was tested by transient expression assays in maize endosperm suspension cells. Unlike the first exon and intron of the *Sbe1* gene, the corresponding region of the *Sbe2b* gene did not increase the level of gene expression. 5' deletion analysis revealed that the 122-bp region from -160 to -49 is essential for promoter activity.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my thesis advisor, Dr. Mark J. Guiltinan, for his guidance, encouragement, and support during my graduate study and research. I would also like to express my sincere appreciation to Drs. Jack C. Shannon, Jill Deikman and David S. Gilmour for their invaluable advice and discussions throughout this work. I am also grateful to Dr. Charles D. Boyer. He gave me the opportunity to study at The Pennsylvania State University where I have enjoyed unforgettable academic and cultural experiences. My appreciation is also extended to fellow graduate students and staff in the laboratory for their friendship, help, humor and encouragement.

Special thanks to my family. I could never thank enough for their constant love, support and encouragement. Thank God for everything!

Chapter 1

REVIEW OF PERTINENT LITERATURE

1.1. Starch Granule

Starch is the major storage carbohydrate of most higher plants and can be found in many organs such as seeds, roots, tubers, stems, leaves, fruits and pollen (reviewed in Shannon and Garwood, 1984). Depending on its source, the granule can have a variety of sizes ranging from 0.5 to 175 μm in diameter, as well as shapes such as spheres, ellipsoids, polygons, platelets and irregular tubules (Zobel, 1988). Starch is found as a water-insoluble granule consisting of crystalline and amorphous regions and is mainly composed of two different polysaccharides, amylose and amylopectin (Takeda et al., 1988). Amylopectin is present in both areas but is considered to be solely responsible for the crystallinity of the starch granule. In contrast, amylose is associated only with the amorphous regions.

Amylose is considered to be an essentially linear α-1,4-linked glucose chain of about 1,000 residues long. In fact, however, it is branched to a very small extent (approximately 1 branch per 1,000 glucose residues) by α-1,6-glucosidic bonds (Takeda et al., 1990). Amylose has the ability to form complexes with many hydrophobic molecules, such as iodine, fatty acids or a number of hydrocarbon chains, since it has a helical conformation with six glucose residues per turn. It can also crystallize out of aqueous solution and shrink depending on conditions such as temperature, degree of polymerization, and concentration (retrogradation) (Zobel, 1988). Amylopectin is a more highly branched macromolecule consisting of linear α-1,4-glucose chains with α-1,6-glucosidic bonds at branch points. Branches occur about every 21 glucose residues (Kainuma, 1988). Amylopectin is stable in aqueous solution and generally does not form complexes, remaining fluid and giving high viscosity and elasticity to pastes and thickeners.

Usually, amylose and amylopectin constitute about 23-31% and 69-77% of starch, respectively. However, the proportion can vary considerably depending on starch source (Shannon and Garwood, 1984). Since the two polysaccharides, amylose and amylopectin, have distinctive properties, it is apparent that the ratio of amylose and amylopectin is important in determining many of the chemical and physical properties of starch.

1.2. Accumulation and Uses of Starch

Starch granules are synthesized in both photosynthetic and non-photosynthetic cells and can be divided into two types, transitory and reserve, depending on how they are utilized by plants. A typical example of transitory starch can be found in photosynthetic organs such as leaves, in which starch is accumulated in chloroplasts during the day and is mobilized for translocation at night. Transitory starch can also be synthesized in the non-photosynthetic tissues such as meristems, pollen grains and root cap cells. The major site of reserve starch accumulation is in the amyloplasts of storage organs including seeds, fruits and tubers (reviewed in Shannon and Garwood, 1984; Martin and Smith, 1995).

In plants, starch provides carbon and energy for vegetative and reproductive development and plays a very important role in buffering the level of soluble sugar via a diurnal starch-sucrose interconversion cycle (Stitt et al., 1987). According to the starch-statolith theory, starch is also involved in sensing gravity (reviewed in Salisbury, 1993). Although there is some controversy regarding a statolith-based model for gravity perception, many plant biologists strongly support the essence of the starch-statolith theory, i.e., dense particles functioning in the early stage of gravitropism.

Starch is not only a major source of food and feed but also is an important industrial commodity throughout the world. It is utilized in the production of paper, adhesives, textiles and biodegradable plastics, not to mention its uses in the food industry. Starch is also used in the production of ethanol by fermentation. Other uses include building materials, packaging and pharmaceuticals (reviewed in White, 1994). Depending on the industrial applications of starch, however, its desired properties vary. For example, starch with high amylose content is required for adhesives and films, while high amylopectin starch is used for salad dressings.

1.3. Starch Biosynthetic Pathway

Although the pathway of starch biosynthesis is not completely understood, there is no doubt that it involves at least four groups of committed enzymes: ADP-glucose pyrophosphorylase (EC 2.7.7.23), starch synthase (EC 2.4.1.21), starch branching enzyme (EC 2.4.1.28) and starch debranching enzyme (EC 2.4.1.41) (reviewed in Preiss, 1991; Martin and Smith, 1995).

ADPG pyrophosphorylase is a heterotetrameric enzyme catalyzing a rate-limiting reaction in starch biosynthesis . It is capable of producing ADP-glucose and pyrophosphate from glucose-1-phosphate and ATP. Starch synthases elongate amylose and amylopectin by transferring glucose from ADP-glucose to the nonreducing end of the polymers. They can be classified into two groups, soluble starch synthases and granule-bound starch synthase (GBSS). Starch branching enzymes (SBEs) catalyze the formation of amylopectin by introducing α-1,6 branch points into the linear α-1,4 linked glucose chains. Introduction of branches facilitates starch synthesis by increasing the number of nonreducing ends, the sites of glucose addition by starch synthases. Thus, SBEs are of crucial importance for the quantity and quality of starch synthesized in the plant (Edwards et al., 1988). In fact, mutation in *Sbe* genes of pea, maize and rice severely decreased total starch content and changed the ratio of amylose and amylopectin (Shannon and Garwood, 1984; Smith, 1988; Mizuno et al., 1993).

Recent evidence strongly supports the possibility that an additional enzyme, a starch debranching enzyme which hydrolyzes α-1,6 glycosidic bonds, also plays an important role in producing the final branching pattern of amylopectin (James et al., 1995). Apparently, the balanced actions of SBEs and debranching enzymes are critical for determining the final degree of branching in amylopectin.

Apart from the enzymes mentioned above, there is a possibility that starch phosphorlylase (EC 2.4.1.1), which can extend a preexisting glucan chain using glucose-1-phosphate as a substrate *in vitro*, may be involved in starch biosynthesis (reviewed in Nelson and Pan, 1995). However, to date no mutants have been identified that definitely demonstrate phosphorylase involvement in starch synthesis.

Starch synthases require a preexisting oligosaccharide as a substrate to carry out their function in starch biosynthesis. At present, however, we do not know how the initial primer is produced *in vivo*. Two possible candidates generating the primer have been found in potato tubers and maize endosperm. One of them is a UDPG-protein transglucosylase (EC 2.4.1.112) which is capable of transferring a single glucose residue from UDP-glucose to itself (autoglucosylation) (Ardita and Tandecarz, 1992). The other is an isoform of starch phosphorylase (phosphorylase II) which can synthesize long α-1,4 glucose chains in the absence of preexisting oligosaccharides (reviewed in Preiss, 1991).

1.4. Starch Biosynthesis in Maize Endosperm Cells

The study of maize mutants showing altered starch content and composition has provided vital information necessary for elucidating the starch biosynthetic pathway in higher plants. Since sucrose translocated from leaves is the primary source of carbon skeletons for starch synthesis in maize endosperm, the biochemical processes involved in sucrose entry into the endosperm will be reviewed first.

1.4.1. Entry of Sucrose into the Endosperm and Its Utilization

It has been known that sucrose is the major sugar released from the pedicel tissue of maize kernels (Porter et al., 1985). Most of the sucrose is then hydrolyzed by invertases (EC 3.2.1.26) into glucose and fructose before it enters the endosperm, where sucrose is resynthesized by the combined actions of sucrose phosphate synthase and sucrose phosphate phosphatase (Shannon, 1968; Chourey et al., 1993). Sucrose phosphate synthase catalyzes the formation of sucrose-6-phosphate and UDP from fructose-6-phosphate and UDP-glucose. The sucrose-6-phosphate is then hydrolyzed by sucrose phosphate phosphatase to produce sucrose and phosphate. Contrary to the implication of the name, however, sucrose synthase (EC 2.4.1.13) does not seem to be involved in sucrose synthesis in the maize endosperm although it can synthesize sucrose *in vitro*. In fact, sucrose synthase encoded by the *Shrunken1 (Sh1)* locus on chromosome 9 is involved in converting sucrose into UDP-glucose and fructose, which are then converted to glucose-1-phosphate by UDP-glucose pyrophosphorylase and enzymes of the glycolytic pathway, respectively (Chourey and Nelson, 1976; Chourey and Nelson, 1979). Homozygous *sh1* mutant kernels produce 55-60 % less starch than nonmutant kernels and show a cavity within the endosperm and a shrunken phenotype. The finding that not all sucrose synthase activity in the endosperm is under control of the *Sh1* locus led to cloning of the second gene for sucrose synthase, designated *Sus1* (McCarty et al., 1986). Chourey and Taliercio (1994) suggested that the *Sus1* gene should be dispensable, because loss of its function did not have any discernible phenotypic effect on the maize plant or kernels.

As mentioned earlier, most sucrose is hydrolyzed and resynthesized prior to its utilization in starch biosynthesis in the maize endosperm, suggesting an important role of invertases in starch biosynthesis. It has been known that *miniature* (*mn1*) mutant seed has significantly reduced levels of both soluble and wall-bound invertase activity in the basal portion of the developing seed (Miller and Chourey, 1992), which are encoded by separate structural genes in maize (Xu et al., 1993). Homozygous *mn1* mutant seeds, which are smaller than normal, contain fewer soluble sugars and a higher percentage of sucrose compared to nonmutant seed (reviewed in Nelson and Pan, 1995). Since the reduction in invertase activity in the seed is correlated with the degeneration of maternal cells in the pedicel, at present we do not know whether the *Mn1* gene product is actually an invertase or some protein involved in the development or differentiation of basal endosperm cells.

1.4.2. Metabolites Crossing the Amyloplast Envelope

It is still controversial what metabolite(s) is translocated into amyloplasts in the maize endosperm cells. Several reports indicated that triose phosphates were taken up by amyloplasts via a phosphate translocator, which were subsequently converted into hexose phosphates via gluconeogenesis requiring the activity of fructose-1,6-bisphosphatase in the amyloplasts (reviewed in Nelson and Pan, 1995). However, the significance of triose phosphate uptake by amyloplasts and their use in starch biosynthesis has been seriously questioned by the following findings. First, fructose-1,6-bisphosphatase, a key regulatory enzyme in gluconeogenesis which is necessary for transforming triose phosphates to hexose phosphates, was not found in the amyloplasts of storage tissues (Entwistle and ap Rees, 1990). Second, when specially labeled hexoses were fed to developing kernels of maize, the pattern of label redistribution among the carbon positions of the hexoses in starch did not support a pathway utilizing triose phosphates as intermediates (Hatzfeld and Stitt, 1990).

By measuring metabolite levels in kernels of several starch-deficient mutants of maize, Tobias et al.(1992) have found that *sh1* kernels contained reduced levels of hexose phosphates, which resembled the reduction of starch accumulation in amyloplasts. In contrast, levels of triose phosphates were elevated. These findings led them to suggest that hexose phosphates are the major compounds which are taken up by amyloplasts and are utilized in the starch biosynthetic pathway.

It has been reported that ADP-glucose can be taken up by amylopasts and subsequently incorporated into starch, providing a third possible metabolite crossing the amyloplast for starch biosynthesis (Pozueto-Romero et al., 1991a and 1991b). Importantly, a significant reduction of this uptake was observed in amyloplasts from the endosperm of *bt1* mutant kernels (Liu et al., 1992). The *bt1* locus on chromosome 5 was first identified in 1926 based on the collapsed endosperm phenotype at seed maturity, and it was cloned by transposon tagging (Sullivan et al., 1991). The effect of the *bt1* locus appears to be endosperm-specific, with no apparent defect in embryo or plant development. The *bt1* kernels contained 80% less starch than normal kernels, and sucrose content was greatly increased (Tobias et al., 1992).

Along with the reduced uptake of ADP-glucose in the *bt1* kernels, several reports suggested that the BT1 protein belongs to a family of adenylate translocaters localized in amyloplast membranes in endosperm. First, the level of ADP-glucose in the cytosol of *bt1* immature maize endosperm increased almost 13-fold compared to that in normal endosperm (Pien et al., 1993). Second, the deduced amino acid sequence of the *Bt1* cDNA showed greatest similarity to a yeast ATP/ADP carrier protein and also had considerable similarity to a family of mitochondrial inner membrane adenylate translocator proteins (Sullivan et al., 1991). Third, Li et al. (1992) showed that *in vitro* synthesized BT1 protein can be imported into pea chloroplasts and become part of the inner membrane. Finally, antibody raised against the BT1 protein reacted with a membrane-bound protein surrounding the starch granules (Sullivan and Kaneko, 1995) or membranes isolated from maize endosperm amyloplasts (Cao et al., 1995).

ADP-glucose pyrophosphorylase, the major source of ADP-glucose for starch synthesis, was initially known to be located in the amyloplast of maize endosperm cells (reviewed in Preiss, 1991; Martin and Smith, 1995), which raised a contradiction in considering the BT1 protein as an adenylate translocator in amyloplast membranes. However, the problem was later solved by the following reports. First, Shannon et al. (1996) found that ADP-glucose levels in endosperm extracts of the *bt1* kernels as well as the *sh1 bt1* double mutant kernels were 13-fold elevated compared to those of normal kernels. This finding suggested that the activity of ADPG pyrophosphorylase may be localized predominantly in the cytoplasm rather than in the amyloplast of the maize endosperm cells. Later, Denyer et al. (1996) actually showed that more than 95% of the activity of ADPG pyrophosphorylase in maize endosperm is present in the cytoplasm.

The cytosolic location of ADPG pyrophosphorylase is consistent with earlier evidence that the amino acid sequences, predicted from the cDNAs, of both the small and large subunits of the major isoforms of ADPG pyrophosphorylase in barley and maize seeds apparently lack any cleavable N-terminal transit peptides. These peptides are required for all proteins that are encoded by the nucleus and targeted to the stroma of plastids.

1.4.3. ADP-Glucose Pyrophosphorylase

ADP-glucose pyrophosphorylase (AGPase) catalyzes the synthesis of ADP-glucose and pyrophosphate from glucose-1-phosphate and ATP, which is the first committed reaction in the starch biosynthetic pathway (reviewed in Preiss, 1991). Biochemical and genetic evidence indicated that the maize endosperm AGPase is heterotetrameric and composed of two small and two large subunits which are encoded by *Shrunken2 (Sh2)* and *Brittle2 (Bt2)*, respectively (reviewed in Nelson and Pan, 1995). The cloning of *Sh2* and *Bt2* provided definitive evidence for the existence of two dissimilar subunits (Bae et al., 1990; Bhave et al., 1990). Maize mutants carrying *sh2* or *bt2* showed severely decreased starch accumulation in the kernels, indicating both subunits are required for full AGPase activity (Martin and Smith, 1995).

The catalytic activity of leaf AGPase is allosterically regulated by the small effector molecules, 3-phosphoglyceric acid (3-PGA) and inorganic phosphate (Pi), which activate and inhibit enzyme function, respectively. However, it is still unclear whether 3-PGA and Pi have the similar effect on AGPase activity in non-photosynthetic tissues. Investigations into the biosynthetic pathway of starch in amyloplasts indicate that 3-PGA is not a key intermediate translocated between the stroma and the cytoplasm (review in Martin and Smith, 1995). This suggests a diminished role for allosteric activation of AGP in non-photosynthetic tissues. Also, the maize endosperm AGP is known to be less sensitive to 3-PGA activation compared to that of the chloroplast enzyme, leading Hannah and Nelson (1975) to speculate that this activation may not be physiologically relevant.

If the genes encoding endosperm AGPs were derived through evolution from those encoding the chloroplast activity, then the activation properties may simply represent evolutionary relics. In fact, the enzyme from barley endosperm is not activated by 3-PGA *in vitro*, suggesting that 3-PGA activation of AGP is not a prerequisite for starch synthesis in the endosperm (Kleczkowski et al., 1993).

In *sh2* and *bt2* mutants, AGP activity was not reduced in the embryo, suggesting different genes encode the enzyme in these two organs (Bryce and Nelson, 1979). This was later confirmed by cloning of the embryo transcripts hybridizing to *Sh2* and *Bt2* clones, which were shown to be different from *Sh2* and *Bt2* and represent a separate set of genes (Giroux and Hannah, 1994; Giroux et al., 1995). In addition, the *Bt2* counterpart in leaf tissue has been isolated and it has been shown from sequence analysis that it is not *Bt2* nor the embryo small subunit encoded by the *AGP2* gene (Prioul et al., 1994). Analysis of *sh2* and *bt2* mutants also revealed the presence of a second set of AGP structural genes, *AGP1* and *AGP2*, which are expressed at low levels in the endosperm (Giroux and Hannah, 1994). This appears very similar to the case for *Sh1* and *Sus1* in controlling endosperm and embryo sucrose synthase.

1.4.4. Starch Synthase

Starch synthases catalyze the elongation of amylose and amylopectin by adding ADP-glucose to the nonreducing end of the polymers by α-1.4 glycosidic bonds and can be divided into soluble starch synthases (SSS) and granule-bound starch synthase (GBSS) (reviewed in Nelson and Pan, 1995). The *waxy* (*wx*) mutants of maize, which are deficient in GBSS, lacked the amylose component of starch, but starch content was not affected, implying that the ADP-glucose not used by GBSS to make amylose is used by other SSSs to make amylopectin. This also suggested that amylopectin is not synthesized from amylose. This finding led Martin and Smith (1995) to hypothesize that the ability of GBSS to make amylose is not an intrinsic property of the protein but rather a function of its location in the starch granule.

According to the hypothesis, GBSS can make amylose only when it is located within the granule. On the other hand, amylopectin is synthesized by combined actions of GBSS, which is closer to the outer edge of the granule, and SSSs as well as starch branching enzymes which are active predominantly in the soluble phase. This might explain in part why amylose production appears to lag behind that of amylopectin during development of storage organs (Shannon and Garwood, 1984).

Two SSSs have been identified in maize endosperm, SSSI and SSSII (Ozbun et al., 1971). SSSI is capable of unprimed synthesis if the reaction medium contains citrate and bovine serum albumin, whereas SSSII requires a primer. So far, no mutation affecting the activity of either enzyme has been identified in a higher plant. Three possible reasons for this observation were suggested by Nelson and Pan (1995). First, such mutations may be lethal. Second, loss of either enzyme may not reduce the amount of starch so that a mutation would not be detectable. Third, each enzyme may be encoded by duplicate loci and mutations in each locus would be required before starch synthesis is affected.

1.4.5. Starch Branching Enzyme

Starch branching enzymes (SBEs) catalyze the formation of amylopectin by introducing α-1.6 branch points into the linear α-1,4 linked glucose chains. Introduction of branches facilitates starch synthesis by increasing the number of nonreducing ends, the sites of glucose addition by starch synthases. Thus, SBEs are of crucial importance for the quantity and quality of starch synthesized in the plant (Edwards et al., 1988). In fact, mutation in *Sbe* genes of pea, maize and rice severely decreased total starch content and changes the ratio of amylose and amylopectin (Shannon and Garwood, 1984; Smith, 1988; Mizuno et al., 1993).

Multiple forms of starch branching enzymes (SBE) differing in enzymatic and biochemical properties have been identified and characterized in various plant tissues, such as spinach (Hawker 1974) , pea (Matters and Boyer, 1981; Smith, 1988), potato (Griffin and Wu, 1968; Khoshnoodi et al., 1993), teosinte (Boyer and Fisher, 1984), rice (Mizuno et al., 1992; Yamanouchi and Nakamura, 1992; Nakamura et al., 1992a), and maize (Hodges et al., 1969; Boyer and Preiss, 1978; Dang and Boyer, 1989). They can be grouped into two distinct families based on their structural relatedness and are named according to the prototypic family member from maize. The SBEI family consists of maize SBEI, rice SBEI, and pea SBEII, and the SBEII family encompasses maize SBEII, rice SBEIII, and pea SBEI.

Significant differences in the enzymatic properties between the SBE families are well documented (reviewed in Martin and Smith, 1995). SBEs belonging to the SBEII family have lower affinity for amylose than SBEI isoforms and prefer to use shorter glucan chains for further branch formation. Another noteworthy difference between the SBEI and II families is that they are differentially regulated during seed development. The SBEII family genes are expressed earlier than the SBEI family members in developing seeds (Gao et al., 1996 ; Smith, 1988; Burton et al., 1995), which may result in changes in the SBEI/SBEII ratio. Since SBEI and IIb have significantly different *in vitro* catalytic properties as mentioned above, such changes in the SBEI/SBEIIb ratio may cause differences in the starch synthesized during kernel development. During pea embryo development, in fact, changes in the SBE isoform ratio was accompanied by transition in branch lengths of amylopectin (Burton et al., 1995).

1.4.6. Starch Debranching Enzyme

It was generally thought that starch is synthesized through sequential reactions catalyzed by ADP-glucose pyrophosphorylase, starch synthase and starch branching enzyme. However, analysis of the *sugary1* (*su1*) mutants of maize suggested a possible involvement of starch debranching enzyme in starch biosynthesis in higher plants. Starch debranching enzyme (EC 3.2.1.41), pullulanase or R-enzyme, hydrolyzes the α-1,6 glucan branches of amylopectin but does not utilize glycogen as a substrate efficiently. Pan and Nelson (1984) showed that starch debranching enzyme activity is significantly reduced in *su1* mutants which produce reduced amounts of starch with a higher percentage of amylose and a substantial amount of phytoglycogen, the highly branched, high molecular weight and water-soluble polysaccharide (Pan and Nelson, 1984; Shannon and Garwood, 1984). Later, James et al. (1995) isolated the *sugary1* gene in maize by transposon tagging, and demonstrated that the gene has significant sequence similarity to a family of genes involved in the cleavage of α-1,6 glycosidic bonds. Taken together, these data strongly support the hypothesis proposed by Pan and Nelson (1984) that the final structure of amylopectin might be determined by a balance between the activities of debranching and branching enzymes.

Recently, Ball et al. (1996) proposed a model for the biogenesis of the plant starch granule. According to the model, glucan trimming by starch debranching enzyme is required to generate order in the amorphous lamella for subsequent synthesis of the crystalline lattice.

Chapter 2

MOLECULAR ANALYSIS OF A GENOMIC FRAGMENT CONTAINING THE MAIZE (*ZEA MAYS* L.) STARCH BRANCHING ENZYME I GENE

2.1. Introduction

In maize endosperm, three isoforms of starch branching enzymes, SBEI, IIa and IIb, have been identified and their genes have been cloned in our laboratory (Boyer and Preiss, 1978; Fisher et al., 1993; Fisher et al., 1995; Gao et al. 1997). Northern blot analysis showed that the SBEs are expressed in a coordinate fashion with the granule-bound starch synthase (GBSS) and ADP-glucose pyrophosphorylase during maize endosperm development (Gao et al., 1996). In addition, ADP-glucose pyrophosphorylase gene (AGPase S) from potato and the genes encoding GBSS and SBE in cassava plants were induced by exogenous supply of sugars (Muller-Rober et al., 1990; Salehuzzaman et al., 1994). These results indicate that these genes probably share common regulatory mechanisms controlling their expression. However, no promoter elements or transcription factors controlling starch biosynthetic genes have been identified to date.

In order to better understand the regulation of starch biosynthesis in higher plants, our laboratory has been working on molecular characterization of the maize genes encoding starch branching enzymes. It is our ultimate goal to understand the relative roles of each SBE isoform in starch biosynthesis and to define the regulatory sequences and transcription factors involved in their expression. Such knowledge will not only allow us to generate transgenic plants that might enhance total starch production or produce unique starches for special industrial purposes, but will also establish a basis for unraveling the regulatory mechanisms of starch biosynthesis. A prerequisite to accomplishing these goals is to identify all the genes encoding SBE isoforms and analyze their promoter sequences involved regulating gene expression.

In this study, a full-length genomic clone, containing the entire coding region of starch branching enzyme I (*Sbe1*) as well as 5'-and 3'-flanking sequences, was isolated from a maize genomic library and its DNA sequence was completely determined. Using transient expression assays, *cis*-acting elements important for the *Sbe1* gene expression in maize endosperm cells were investigated.

2.2. Materials and Methods

2.2.1. PCR Amplification and Cloning of Maize Genomic DNA Containing an *Sbe1* Gene

Maize genomic DNA prepared from 22 DAP (days after pollination) kernels (W64A) using the method of Rogers and Bendich (1985) was amplified in a 50 µl reaction mixture containing 1 µg of target DNA, 1 µM of each primer, 200 µM of each dNTP, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin and 2.5 U of *Taq* DNA polymerase (Boehringer Mannheim, Indianapolis, IN). The mixture without dNTP was overlaid with 50 µl of mineral oil (Sigma, St.Louis, MO) and incubated for 5 min at 94°C. The dNTPs were added to the mixture at the end of the incubation and then the mixture was cycled 35 times in a thermal cycler (ERICOMP, San Diego, CA) as follows: 94°C for 30 sec; 55°C for 1 min; and 72°C for 2 min; with a final 72°C extension of 7 min. The primers were designed according to the published sequence data of maize *Sbe1* cDNA (Baba et al., 1991). The 5' primer, 5'-GACTGAATTCCTGCGCAGGAGGCAGAGCTT-3', and the 3' primer, 5'-GATCGAATTCCATAGATACGTGGAGCAGCA-3', are homologous and complementary to DNA sequences of the maize *Sbe1* cDNA from 438 bp to 457 bp and 745 bp to 764 bp, respectively. Each primer contains an EcoRI restriction enzyme site and 4 extra nucleotides (underlined) at their 5' ends for convenience of subsequent cloning of the PCR product. After amplification, 15 µl of the reaction sample was run on an agarose gel (1.5 %) in 1X TAE buffer containing 0.04 M Tris-acetate and 1 mM EDTA. A single PCR product of about 0.5 kb was detected on an ethidium bromide-stained agarose gel, digested with EcoRI restriction enzyme and cloned into the corresponding site of pBluescript SK$^-$ (Stratagene, La Jolla, CA) creating plasmid pB1.

2.2.2. Maize Genomic Library Screening and DNA Sequencing

An EMBL-3 genomic library (Clonetech, Palo Alto, CA) prepared from maize seedlings (2 leaf stage, B73) was screened according to Sambrook et al. (1989). Approximately 3 x 10^5 plaques were transferred onto nylon membranes (Hybond-N+, Amersham, UK) and hybridized with the ^{32}P-labeled *SbeI* genomic PCR product corresponding to the region from 438 to 764 of the maize *SbeI* cDNA. Hybridization was performed at 55°C for 20 hours in 0.5 M Na_2HPO_4, pH 7.2, and 7% SDS (Church and Gilbert, 1984) with gentle agitation at 40 cycles per minute on a rotary shaker. Following the hybridization, filters were washed twice in 5% SEN (5% SDS, 1 mM EDTA, 0.04 M Na_2HPO_4, pH 7.2) and once in 1% SEN (1% SDS, 1 mM EDTA, 0.04 M Na2HPO4, pH 7.2) for 15 minutes at 65°C. Plaques strongly hybridizing to the probe were selected and purified through three rounds of screening. Phage DNAs were isolated from the positive plaques according to Chisholm's method (1988) and were digested with PstI to release the inserts from the EMBL-3 vector. The restriction fragments were separated on an 0.8% agarose gel and blotted onto nylon membranes. The blots were probed with ^{32}P-labeled full-length *SbeI* cDNA (Fisher et al., 1995) as described above and hybridizing DNA fragments were identified and subcloned into pBluescript SK$^-$. DNA sequences were determined by the

dideoxynucleotide chain termination method (Sanger et al., 1977) with Sequenase Version 2.0 (United States Biochemical Co., Cleveland, OH). Sequence analyses were performed using programs from DNASTAR Inc. (Madison, WI).

2.2.3. Genomic Southern Blot Analysis

Maize genomic DNA was prepared from 7-day-old etiolated seedlings (B73) according to the method described by Junghans and Metzlaff (1990). 10 µg of genomic DNA was digested with restriction enzymes such as BamHI, EcoRI, BglII, and HindIII, separated on 0.8% agarose gels, and transferred onto nylon membranes (Hybond-N, Amersham, UK) in 20 X SSC containing 3 M NaCl and 0.3 M sodium citrate, pH 7.0 according to Sambrook et al. (1989). The DNA was crosslinked to the membrane by 3.5 minutes of UV irradiation on a transilluminator (312 nm). The genomic blots were prehybridized at 65°C for 1 hour in 0.5 M Na_2HPO_4, pH 7.2, 7% SDS, and 100 µg/ml denatured salmon sperm DNA. 25 ng of a full-length *SbeI* cDNA (Fisher et al., 1995) was labeled with [α-^{32}P]-dCTP using the random primed DNA labeling kit (Boehringer Mannheim, Indianapolis, IN). This labeled probe was added to the prehybridization solution and incubated at 65°C for 18 hours. Blots were washed twice in 5% SEN and once in 1% SEN for 15 minutes at 65°C and were exposed to Kodak X-AR film at -80°C for 48 hours using two intensifying screens.

2.2.4. Primer Extension

To locate the transcription initiation site of the *SbeI* gene, an oligonucleotide, 5'-GGCGACACGAGGCACAGCAT-3', which is complementary to the sense strand sequence of the *SbeI* cDNA from +1 to +20 relative to the translation start site (ATG) was radiolabeled at its 5' terminus with T4 polynucleotide kinase and γ-^{32}P-ATP (Sambrook et al., 1989). Approximately 10^5 cpm of the labeled primer was hybridized at 35°C with 10 µg of total RNA, which was isolated from 30 DAP maize kernels (B73) according to the protocol of Vries et al. (1988). After 8 hours of hybridization, complementary DNA was synthesized from the annealed primer by the addition of reverse transcriptase and dNTP. Following the addition of EDTA and RNAase A into the reaction, the nucleic acid was precipitated with ethanol. The reaction products were resuspended in sequencing gel loading buffer, denatured at 95°C, electrophoresed through a 5% polyacrylamide sequencing gel, and visualized by autoradiography. In order to provide size markers, part of the *SbeI* gene was sequenced with the same primer used in the primer extension experiment.

2.2.5. 3' Rapid Amplification of cDNA Ends (3' RACE)

To isolate the 3' end of the *Sbe1* transcript, the 3' RACE method was used. The first strand cDNA synthesis reaction was performed as follows: 14 µl of a mixture containing 5 µg of total RNA from 30 DAP maize kernels (B73) and 50 pmol of a 39-bp oligonucleotide with 17 dT residues and an adaptor sequence, 5'-GGTCGACTCGAGTCGACATCGATTTTTTTTTTTTTTTTT-3', was heated at 70°C for 10 min and quickly chilled on ice. To the chilled sample, 2 µl of 10X synthesis buffer (200 mM Tris-HCl, pH 8.4, 500 mM KCl, 25 mM $MgCl_2$, and 1 mg/ml BSA), 1 µl of 10 mM dNTP mix, 2 µl of 0.1 M DTT and 1 µl (200U) of SUPERSCRIPT reverse transcriptase (GIBCO BRL, Grand Island, NY) were added and incubated at room temperature for 10 min. The reaction mixture was then placed in a 42°C water bath for 50 min and transferred to a 90°C water bath. After 5 min incubation, 1 µl of RNaseH (2U/µl) was added and incubation at 37°C was continued for 20 min. Next, the first strand cDNA obtained was amplified directly by the PCR method using a gene-specific primer (5'-GACTGAGCTCATACCAAATGAAGCCAGGAG-3'), which is a homologous to sequence of the *Sbe1* gene from +5382 to +5401, and an adaptor primer (5'-GGTCGACTCGAGTCGACATCGA-3'). After amplification, a single DNA band detected on an agarose gel (1.5%, w/v) was isolated and digested with SacI and XhoI (underlined within the primers). The resulting fragment, approximately 350 bp in length, was cloned into a pBlusScript SK$^-$ and sequenced (plasmid pB1-3').

2.2.6. Construction of Chimeric Plasmids

For a transcriptional fusion of the *Sbe1* promoter to a luciferase (LUC) reporter gene, a BamHI restriction enzyme site was created just before the translation initiation site of the *Sbe1* gene as follows: the DNA sequence between -253 and +27 of the *Sbe1* gene was amplified via polymerase chain reaction (Figure 2.1). *Pfu* DNA polymerase (Stratagene, La Jolla, CA), which has proofreading activity, was used to enhance the fidelity of PCR amplification. (*Pfu* DNA polymerase was used for all the following PCRs). The 5' primer (PI-1), 5'-CCAGCTCCACGGTTGTTCGTGT-3', is homologous to sequence of the *Sbe1* gene from -253 to -232. An ApaI restriction enzyme site (GGGCCC) is located immediately downstream of the 5' primer binding region of the *Sbe1* promoter, -203 to -198. The 3' primer (PI-2), 5'-CGATGGATCCTGTGACGGCGTGTGAGT CCC-3', consists of a DNA sequence complementary to that of the *Sbe1* gene from +8 to +27 and a BamHI restriction enzyme site (GGATCC) flanked with four random nucleotides (underlined). The PCR product was digested with ApaI and BamHI, and the resulting 236-bp fragment was cloned into pBluescript SK$^-$ and sequenced in order to verify that no misincorporation had occurred in the DNA sequence during the PCR amplification (all the following PCR products were sequenced). Then, the 236-bp fragment was ligated to the 2-kb SalI-ApaI *Sbe1* promoter fragment and cloned into plasmid pLN cut with SalI and BamHI (promoterless LUC-NOS gene in pUC119) (Montgomery et al., 1993), thereby creating plasmid pKL101.

To construct a translational fusion of the *Sbe1* promoter containing the first exon and intron to a LUC reporter plasmid, the DNA sequence between -253 and +228 was amplified with the PI-1 primer and a 3' primer (PI-3), 5'-AGTCGGATCCTCAGGC

GCACATTGCCGCCA-3' designed to anneal to the region just downstream of the first intron of the *Sbe1* gene. The 493-bp PCR product was digested with ApaI and BamHI, and the resulting 436-bp fragment was used to replace the ApaI-BamHI fragment in pKL101. This construct was called pKLN101. To make pKLM101 which contains the *Sbe1* promoter with four exons and introns, the 236-bp ApaI and BamHI fragment in pKL101 was replaced with the 1.8-kb *Sbe1* genomic DNA fragment.

The plasmid pKLNS101 was derived from pKLN101 by replacing the nopaline synthase (NOS) 3' sequence with the native *Sbe1* 3'-flanking sequence. To accomplish this, two primers were designed to amplify *Sbe1* DNA sequences containing the transcription stop signal, and the putative polyadeylation signal and site (from + 5382 to + 5780). The upper primer (PI-4), 5'-GACTGAGCTCATACCAAATGAAGCCAGGAG-3', and the lower primer (PI-5), 5'-ACTGGAATTCGGAACAAGGAACGAAGAAAC-3', contain SacI and EcoRI restriction site, respectively, along with 4 random bases. A 419-bp PCR product was digested with SacI and EcoRI, and the resulting fragment was then used for substituting a 255-bp SacI-EcoRI NOS 3' sequence in the pKLN101.

To create a series of 5' deletions in the *Sbe1* promoter, pKLN101 was first modified as follows: pKLN101 DNA was digested with HindIII and the resulting 7.2 kb fragment lacking the 452-bp HindIII fragment was gel purified. The fragment was blunt ended by Klenow fill-in DNA synthesis and ligated with SalI linkers. After complete digestion with SalI, the DNA fragment was partially digested with BamHI to isolate the approximately 1.9-kb SalI-BamHI fragment, which was then gel purified and cloned into pLN cut with SalI and BamHI to produce pKLN101-1 (Figure 2.1).

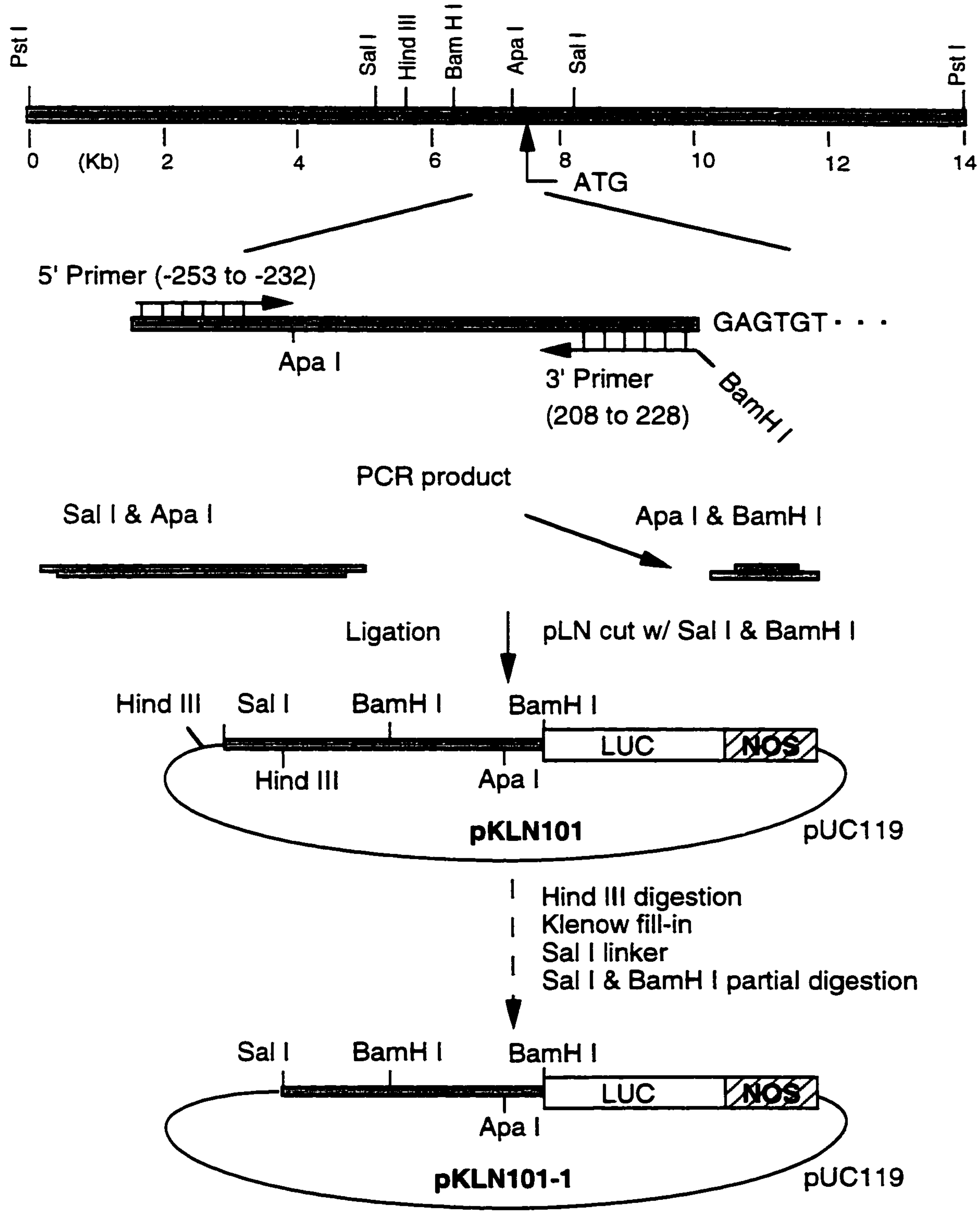


Figure 2.1. Construction of the chimeric construct pKLN101-1

A series of 5' deletion mutants were made from the plasmid pKLN101-1 using an Erase-a-Base system (Promega, Madison, WI) according to the manufacture's directions. Briefly, the plasmid, prepared by two rounds of CsCl/ethidium bromide centrifugation to remove nicked and linear DNA molecules, was digested with HindIII. Next, the recessed 3' ends were filled in with α-phosphorothioate deoxynucleotides and Klenow fragment to protect from exonuclease III digestion. Then, it was digested with SalI to produce a 5' overhanging end which allows exonuclease III to initiate digestion. The double-digested plasmid DNA was incubated with exonuclease III at 37°C and aliquots of the reaction were removed at 30 sec intervals to tubes containing S1 nuclease, which removes the single-stranded tails resulting from exonuclease III digestion. After heat inactivation of the S1 nuclease, Klenow fragment and dNTP were added to generate blunt ends, which were then self-ligated to produce the 5' deletion series plasmids, pKLN102 to pKLN107. All constructs were sequenced with pUC/M13 reverse primer to verify deletion end points.

For the -254 and -196 deletion constructs, two regions of the *SbeI* promoter, -254 to -146 and -196 to -146, were PCR-amplified by primer PI-6 and PI-8, PI-7 and PI-8, respectively. Primers used in PCR to create the *SbeI*-LUC constructs are shown in Table 2.1. Since each 5' primer, PI-7 and PI-8, contain HindIII restriction enzyme sites and a BstXI restriction enzyme site is located between -173 and -162, the PCR products were digested with HindIII and BstXI and the resulting fragments were used to replace the 2,047-bp HindIII-BstXI fragment of pKLN101.

Table 2.1. Oligonucleotides used in PCR to create *SbeI* -LUC chimeric constructs.

Primer[a]	Sequence[b]	Annealing Region[c]
PI-1 U	CCAGCTCCACGGTTGTTCGTGT	-253 to -232
PI-2 L	cgatggatccTGTGACGGCGTGTGAGTCCC	+8 to +27
PI-3 L	agtcggatccTCAGGCGCACATTGCCGCCA	+209 to +228
PI-4 U	gactgagctcATACCAAATGAAGCCAGGAG	+5382 to +5401
PI-5 L	actggaattcGGAACAAGGAACGAAGAAAC	+5761 to +5780
PI-6 U	gatcaagcttACCAGCTCCACGGTTGTTCG	-254 to -235
PI-7 U	attcaagcttCAGATCCGGCTCAGGGTCAT	-196 to -177
PI-8 L	TGCGACAAGGAGGGGGCCAT	-165 to -146

[a] U and L indicate upper (sense) and lower (antisense) primers relative to the *SbeI*, respectively.

[b] The lowercase letters designate restriction sites with four arbitrary bases for cloning.

[c] Numbers represent distance relative to the transcription start site (+1) of the *SbeI*.

2.2.7. Linker-Scanning Mutagenesis

A series of linker-scan mutations were introduced into the 60-bp DNA region from -314 to -255 as described by Kunkel et al (1987). Briefly, the HindIII-BamHI (-314 to +235) fragment from pKLN105, containing the DNA region to be altered, was subcloned into the corresponding sites of a M13mp19 vector to produce a single-stranded template. In order to increase mutant recovery efficiencies, the template was prepared from an *E. coli dut*$^-$ *ung*$^-$ strain (CJ236) which allows the incorporation of uracil into the newly synthesized DNA. Next, a set of oligonucleotides with 10-bp mismatches shown in Table 2.2 were annealed to the template and extended with T7 DNA polymerase. After T4 DNA ligase was added, the resulting heteroduplexes were introduced into a wild-type E.coli strain (MV1190) to generate mutated double-stranded DNAs. DNA sequencing was performed to verify that the desired mutations were correctly introduced and no unintended mutations had occurred.

To create the mutated *Sbe 1* promoter-LUC constructs (pLS1-1 to pLS1-6), the HindIII-BamHI fragment in pKLN105 was replaced by each mutated DNA sequence.

Table 2.2. Oligonucleotides used in linker-scanning mutagenesis.

Constructs	Oligonucleotides
pLS1-1	CCCGGTTTGCCTTTTTTTg cTgcag gacAAGCTTGGCGTAATCAT
pLS1-2	TTGCACGCTTCCCGGTTga Ctgcag gcTATTTTATGTAAGCTTG
pLS1-3	GCCTTTGGGCTTGCACG tcTaga tagcTGCCTTTTTTTATTTTA
pLS1-4	GGGCCGATTGGCCTTTGa ctgcag gtaCTTCCCGGTTTGCCTTT
pLS1-5	AGCTGGTTCTGGGCCGAcga ctgcag aGGCTTGCACGCTTCCCG
pLS1-6	ACAACCGTGGAGCTGGT gtcGac tatcTTGGCCTTTGGGCTTGC

* The mutated bases are shown in lowercase letters, and restriction sites used for convenience of screening are underlined.

2.2.8. Transient Expression Assays

2.2.8.1. Particle Bombardment

Suspension culture cells of maize (*Zea mays*) endosperm (inbred A636) were kindly provided by J. L. Anthony (DEKALB Plant Genetics, Groton, CT). The suspension cells were grown in 250-ml large-mouth Erlenmeyer flasks containing 80 ml of Murashige and Skoog basal salt medium (Murashige and Skoog, 1962) supplemented with 0.4 mg/l thiamine 2 g/l asparagine and 30 g/l sucrose (Shannon and Liu, 1977). The culture was maintained in the dark at 29°C on a rotary shaker (120 rpm) and was subcultured every 7 days by transferring a portion of the cell suspension into fresh medium.

For particle bombardment, about 600 mg (fresh weight) of actively growing cells 3 days after subculture was evenly distributed over the surface of a piece of filter paper (Whatmann #4, 55 mm in diameter) by vacuum filtration of 8 ml of suspension culture. The filter paper bearing the cells was then placed over three layers of filter paper (Whatmann #4, 70 mm in diameter) moistened with 5 ml of the liquid medium containing 12% sucrose and positioned in the middle of a 10 cm petri dish.

60 mg of gold microcarriers (1.6 µm particle size) were washed three times with 1 ml of 100% ethanol and twice with 1 ml of sterile deionized H_2O, resuspended in 1 ml of sterile deionized H_2O, and dispensed in 50-µl aliquots (3 mg/50 µl). The *SbeI* Promoter-LUC constructs and a GUS reference plasmid (pBI 221; Clontech, Palo Alto, CA) were co-precipitated onto gold particles as follows: under continuous vortexing, the following were added in order to each 50-µl aliquot of gold particles: 5 µl of DNA (8 µg of LUC reporter plasmid, 4 µg of GUS reference plasmid), 50 µl of 2.5 M $CaCl_2$, and 20 µl of 0.1 M spermidine (free base, tissue culture grade). The gold particles coated with DNA were pelleted in an eppendorf centrifuge at 10,000 rpm for 10 sec, rinsed with 250 µl of 100% ethanol, and resuspended in 60 µl of 100% ethanol. Immediately after sonication, 8 µl of the DNA-coated gold particles were pipetted onto the center of macrocarriers (Bio-Rad, Hercules, CA) and dried in a low humidity and vibration-free environment.

A Bio-Rad PDS-1000/He Biolistic Particle Delivery system was used for particle bombardment. Bombardment parameters which were optimized include He pressure, gap distance (distance from power source to macroprojectile), and target distance (distance from microprojectile launch site to sample target). After optimization, all bombardments were performed at 650 psi under a vacuum of 26 inches of Hg with a distance of 10 cm between the cells and the barrel of the particle gun. Following the bombardments, the petri dishes were sealed with parafilm and then incubated in the dark at 25°C for 48 hr.

2.2.8.2. GUS and LUC Assays

The bombarded cells were harvested from the plates by vacuum filtration, frozen in liquid nitrogen, and ground with a pestle and mortal to a fine powder. The powder was then transferred into a microfuge tube and extracted with cell culture lysis buffer containing 300 mM Tris-phosphate, pH 7.8, 2 mM DTT, 2 mM 1,2-diaminocyclohexane-N, N, N', N'-tetraacetic acid, 10% glycerol and 1% Triton X-100 (0.3 ml/g of tissue). Cell debris was pelleted in an eppendorf centrifuge at 14,000 rpm for 10 min at 4°C and the supernatant was split into two aliquots for assays of GUS and LUC activity.

For fluorometric GUS assays (Jefferson et al., 1987), 30 µl of the crude extract was incubated at 37°C with 2 mM 4-methyl umbelliferyl glucuronide in 0.3 ml of GUS assay buffer (50 mM $NaPO_4$, pH 7, 10 mM EDTA, 0.1% Triton X-100, 0.1% Sarkosyl, 10 mM β-mercaptoethanol, 20% methanol). After 0, 1, and 2 hr of incubation, 0.1 ml aliquots were removed and added to 0.9 ml of 0.2 M Na_2CO_3 to terminate the reaction. A TKO 100 fluorometer (Hoeffer, San Francisco, CA) calibrated by setting a 100-nM MU to 1,000 fluorescence units was used to measure fluorescence of the product, 4-methyl umbelliferone (4-MU). For each sample, results of GUS assay were plotted in a graph of OD_{405} (Y-axis) versus time in minutes and the GUS activity was expressed simply as the slope of the line. GUS activity from the maize endosperm suspension cells that had been bombarded with the naked gold particles (no DNA) was used as a control.

Using a luminometer (Monolight 1500; Analytical Luminescence Laboratory, San Diego, CA), luciferase activity was determined by measuring luminescence for 10 sec after mixing 20 µl of cell extract with 100 µl of Promega's luciferase assay reagent containing 20 mM tricine, pH 7.8, 1.07 mM $(MgCO_3)_4Mg(OH)_2 \cdot 5H_20$, 2.67 mM $MgSO_4$, 0.1 mM EDTA, 33.3 mM DTT, 270 µM coenzyme A, 470 µM luciferin and 530 µM ATP. LUC activity from the maize endosperm suspension cells that had been bombarded with the pLN (promoterless LUC plasmid) was used as a control. To correct differences in sample variability and transfection efficiency, the luciferase activity (in light unit) was normalized with GUS activity, yielding the LUC/GUS ratio of each sample.

2.2.9. Nuclear Extract Preparation

Maize kernels (inbred B73) were harvested 30 days after pollination and frozen in liquid nitrogen. Nuclear extract was prepared essentially according to the method described by Jensen et al. (1988). The frozen tissue (10 g) was ground with mortar and pestle in liquid nitrogen and transferred into a polypropylene screw-cap centrifuge tube followed by addition of 20 ml of buffer A (10 mM NaCl, 10 mM Mes, pH 6.0, 5 mM EDTA, 0.6% Triton X-100, 0.25M sucrose, 0.15 mM spermine, 0.5 mM spermidine, 20 mM β-mercaptoethanol and 0.2 mM PMSF). After mixing, the suspension was filtered twice through several layers of miracloth and centrifuged for 10 min at 2000 g to pellet crude nuclei. The crude nuclei were twice washed in buffer A, resuspended in 3 ml of buffer C (20 mM HEPES, pH7.9, 25% glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM PMSF and 0.5 mM DTT), and homogenized in a Dounce homogenizer with a B pestle (Kontes, Vineland, NJ). The resulting extract was stirred gently for 2 hr at 4 °C and then centrifuged for 15 min at 15,000 g. The supernatant was dialyzed for 2 hr against 50 volumes of buffer D (20 mM Hepes, pH 7.9, 0.5 mM PMSF and 0.5 mM DTT). Following dialysis, the extract was aliquoted, frozen in liquid nitrogen, and stored at - 80 °C. Protein concentration was determined using a BCA protein assay kit (Pierce, Rockford, IL), according to the manufacturer's instructions.

2.2.10. DNA Probe Preparation

The *SbeI* promoter region from -314 to -255 was PCR-amplified with a forward primer (5'-GGACTTACATAAAATAAAAAAAGG CA) and a reverse primer (5'-TGCTAAGCTTTCTGGGCCGATTGGCCTTTG) which contain BamHI and HindIII restriction enzyme sites, respectively, at their 5' ends (underlined). The PCR product was digested with BamHI and HindIII, and the resulting fragment was cloned into pBlueScript SK^- cut with BamHI and HindIII to create plasmid pRb4-1. The plasmid construct was verified with DNA sequencing.

For electrophoretic mobility shift assays, the DNA fragment was cut out from the plasmid pRb4-1 with HindIII and BamHI, purified from agarose gels, and end-labeled with $[\alpha\text{-}^{32}P]$-dCTP using the Klenow fragment.

2.2.11. Electrophoretic Mobility Shift Assay

The DNA-protein binding reaction was performed in 20 µl of solution containing 0.5 ng of labeled probe, 10 µg of nuclear protein, 1 µg of poly (dI-dC)·poly (dI-dC), 12% glycerol, 12 mM HEPES-NaOH (pH 7.9), 4 mM Tris-Cl (pH 7.9), 60 mM KCl, 1 mM EDTA and 1 mM DTT. After a 20-min incubation at room temperature, the samples were loaded into a 4% native polyacrylamide gel which was pre-run at 4°C for 1 hr at 150 V and electrophoresed for 2.5 hr at 150 V in Tris-glycine buffer at 4°C. Following electrophoresis, the gel was dried with a gel dryer (Bio-Rad) and exposed to Kodak X-ray film with two intensifying screens for 24 hr.

2.2.12. Northern Blot Analysis

Total RNA was isolated according to Vries et al. (1988) from maize endosperm suspension cells which had been incubated for 24 hr in the MS basal salt media supplemented with 0.4 mg/l thiamine, 2 g/l asparagine and various amounts of sucrose from 0% to 15%. Each RNA (10 µg) was denatured at 65°C for 15 min in 20 µl of sample preparation solution containing 2.2 M formaldehyde, 50% (v/v) formamide and 1X MOPS buffer (20 mM MOPS, 1 mM sodium acetate, 10 mM EDTA, pH7.0). After incubation on ice for 5 min, 1 µl of ethidium bromide (1 mg/ml) and 2 µl of formaldehyde gel-loading buffer (50% glycerol, 1 mM EDTA pH 8.0, 0.25%

bromophenol blue and 0.25% xylene cyanol FF) were added to the denatured RNAs. The samples were loaded into a 2.2 M formaldehyde, 1.2% (w/v) agarose gel and run at 4 V/cm for 3-4 hr until the bromophenol blue neared the bottom of the gel. Fractionated RNA was transferred onto a nylon membrane (Hybond N, Amersham) according to the manufacturer's instructions, and fixed by a UV exposure for 30 sec. The blot was prehybridized at 65°C for 1 hr in 0.5 M Na_2HPO_4, pH 7.2, 7% SDS, and 100 µg/ml denatured salmon sperm DNA. 25 ng of a full-length *SbeI* cDNA was labeled with [α-^{32}P]-dCTP using the random primed DNA labeling kit (Boehringer Mannheim, Indianapolis, IN), and purified with a Sephadex G-50 spun column to remove unincorporated nucleotide. After hybridization at 65°C for 18 hr with the probe, the blot was washed twice in 5% SEN and once in 1% SEN for 15 min at 65°C. Radioactivity was detected with a PhosphorImager and quantified with the ImageQuant software program (Molecular Dynamics). To correct for minor loading errors between the lanes, the blot was wash at 95°C in a 0.1% (w/v) SDS solution to remove the ^{32}P-labeled *SbeI* cDNA probe and rehybridized with a ^{32}P-labeled tomato cDNA for 26S rRNA.

2.3. Results

2.3.1. PCR Amplification of Maize Genomic DNA

In order to obtain a probe for genomic library screening, maize genomic DNA prepared from 22 DAP kernels (W64A) was amplified by polymerase chain reaction (PCR) using upper and lower primers designed to anneal the *SbeI* cDNA (Baba et al., 1991) from 438 to 457 and 745 to 764, respectively. A single amplified DNA band, approximately 500 bp in length, was observed on an ethidium bromide-stained agarose gel. The PCR product containing EcoRI sites at both 5'- and 3'-ends was digested with EcoRI restriction enzyme and the resulting fragment was cloned and sequenced.

Alignment of sequences between the maize PCR genomic fragment and the published *SbeI* cDNA (Baba et al., 1991) shows that the PCR fragment contains the predicted *SbeI* cDNA sequence with only 3 different nucleotides and a 103 bp intron (Figure 2.2). The differences could be explained by cultivar polymorphisms or misincorporation of bases during PCR amplification by T*aq* DNA polymerase which does not have proofreading activity. This PCR fragment was ^{32}P-labeled and used as a hybridization probe for screening a maize genomic library to isolate *SbeI* genomic clones.

```
CTGCGCAGGAGGCAGAGCTTATTGGTGACTTCAATGACTGGAATGGTGCAAACCATAAGATGGAGAAGGA  70
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGCGCAGGAGGCAGAGCTTATTGGTGACTTCAATGACTGGAATGGTGCAAACCATAAGATGGAGAAGGA  507

TAAATTTGGTGTTTGGTCGATCAAAATTGACCATGTCAAAGGGAAACCTGCCATCCCTCACAATTCCAAG  140
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TAAATTTGGTGTTTGGTCGATCAAAATTGACCATGTCAAAGGGAAACCTGCCATCCCTCACAATTCCAAG  577

GTTAAATTTCGCTTTCTACATGGTGGAGTACGGGTTGATCGTATTCCTGCATGGATTCGTTATGCGACTG  210
|||||||||||||||||||||||||||||| |||||||||||||||| |||| |||||||||||||||||
GTTAAATTTCGCTTTCTACATGGTGGAGTATGGGTTGATCGTATTCCAGCATTGATTCGTTATGCGACTG  647

TTGATGCCTCTAAATTTGGAGCTCCCTATGATGGTGTTCATTGGGATCCTCCTGCTTCTGAAAGGTCTCT  280
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTGATGCCTCTAAATTTGGAGCTCCCTATGATGGTGTTCATTGGGATCCTCCTGCTTCTGAAAG------  711

TTCTACTGTCCTGGACCACATTGTAAACTTAAGCAATAGACCTTAGCAGACAAAAATATATGACAAAATT  350

----------------------------------------------------------------------

ATAGTACCTTTTACCTCTGGTTTGCAGGTACACATTTAAGCATCCTCGGCCTTCAAAGCCTGCTGCTCCA  420
                           |||||||||||||||||||||||||||||||||||||||||||
---------------------------GTACACATTTAAGCATCCTCGGCCTTCAAAGCCTGCTGCTCCA  754

CGTATCTATG  430
||||||||||
CGTATCTATG  764
```

Top strand : PCR Product

Bottom strand : *Sbe1* cDNA

Figure 2.2. Sequence alignment of the maize genomic PCR product and the published *Sbe1* cDNA sequence (Baba et al., 1991).

DNA sequences of the Sbe1 PCR product and the published cDNA are displayed on top and bottom, respectively. Short vertical lines indicate conserved nucleotides and dashes represent the predicted introns.

2.3.2. Isolation and Analysis of a Maize *Sbe1* Genomic Clone

In a screen of approximately 3 x 10^5 plaque-forming units from a genomic library prepared from maize seedlings (inbred B73), 8 positive lambda clones were isolated which strongly hybridized to the probe, a PCR-amplified genomic fragment corresponding to the region from positions 438 to 764 of the maize *Sbe1* cDNA (Baba et al., 1991). Restriction mapping and partial DNA sequencing of these clones clearly indicated they all originated from the same genetic locus. A full-length clone (λ5-1-1), containing the entire coding region of the *Sbe1* gene as well as 5'- and 3'- flanking sequences, was selected for further analyses. Figure 2.3(A) shows a restriction map of the genomic clone. The 3.0-kb SalI fragment and the 6.2-kb PstI fragment from the clone were subcloned into pBluescript SK$^-$ producing plasmids pBI5-1 and pBI5-2, and their nucleotide sequences were completely determined and shown in Figure 2.4.

Consensus sequences of a TATA-box as well as a G-box containing a perfect palindromic sequence, CCACGTGG, were found in the 5' flanking region of the gene. Primer extension analysis was performed to identify the transcription initiation site of the *Sbe1* gene. As shown in Figure 2.5, a single extended product of 44 nucleotides was observed. The band co-migrated with an A residue in the sequencing ladder generated using the *Sbe1* 5'-flanking region as a template, indicating that the transcription initiates at a position which is located 24 bp downstream from the putative TATA box. This suggests that the TATA sequence may be a functional element of the *Sbe1* promoter. The position of the transcription start site is indicated in Figure 2.4.

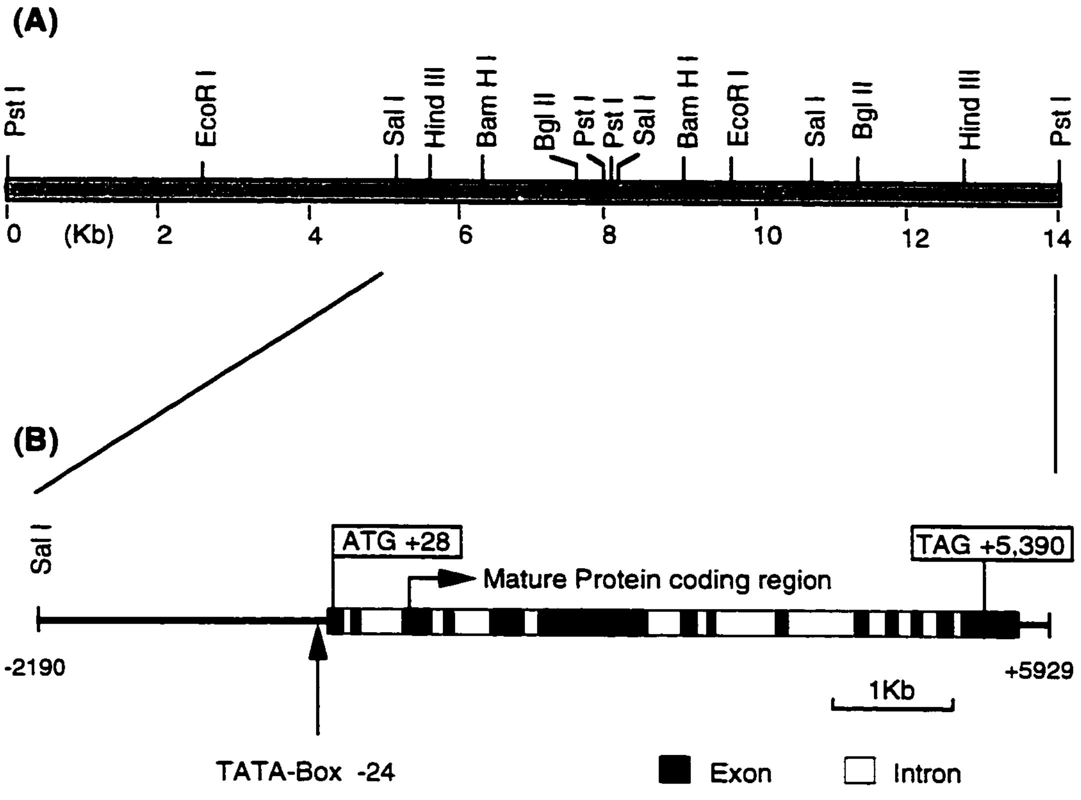


Figure 2.3. Structure of the lambda clone 5-1-1 containing the *SbeI* gene.

(A) Restriction map of the 5-1-1 clone.
(B) Genomic structure of the *SbeI* gene. The thin black lines indicate the 5'- or 3'-flanking sequences of the *SbeI* gene. The solid black boxes indicate exons and the open boxes denote introns. The numbers represent positions relative to the transcription initiation site (+1).

Figure 2.4. Nucleotide sequence of the *SbeI* gene and 5'- and 3'-flanking regions.

Flanking regions and introns are shown in lowercase letters, while exons are presented in uppercase letters. The deduced amino acid sequences are shown below the string of exon sequences. Numbers indicate the distance relative to the transciption start site (+1) which is indicated by the arrowhead. The consensus TATA and G-box sequences as well as putative polyadenylation signal are underlined. The regions containing at least 82% sequence homology with the rice *SbeI* 5' flanking region is also underlined. The asterisk and dot indicate the stop codon and putative polyadeylation site, respectively.

gtcgactgccctctagaccccgatgaagtcaaggatagtgagcgcccaaccaagtttaggctcccccagc -2121

cagatgatcatgttaggtgttccactgccataaaagaaagcgtccgtgatgtcgtacttgggggcggtga -2051

atacgtaaggggacagatgctcatataccttcatctgttttactcgaaattactaatgttttccctatgt -1981

cagctaaagtgaaatctcacccgaccaatcttgcagattgtgaatgacttgaagtttaccgtgaaccatg -1911

ctgtggaacccatcaatgaaaagctgcacatgatatccgagaacatcaaaaagcgtgagaagggaaagag -1841

gaaaacgaatgatgatagctgcatcagttcaccaacatccttgaccagggtcatcagtgggattgatgat -1771

ccgctcctgaacggaagcttgagtgacaacagcggtccaaagaaagcacggaggcagcgcaggaaatcag -1701

gctatgcaagcgcagagagcggtggggagagcagcgaccatggcctaggcgggtttgagatccgagggaa -1631

cagatggattactagggagtgacagctaggagcaatctgtggagatactatgtaaatgtcacaggagttg -1561

gcctgtggtgatgtgcgataaccgggccagtgggttctgagcgtaagacgaggtctctagcaatgatatg -1491

ttgtagcgtgactaaatcgagatgagcaatcaggttgttgtctgttgattctttcggctgaccatttcgg -1421

ctgctgaggccttgatcctgtatttgagatgtcaggcctacttagagcaaatgctggaaggaagcaatgt -1351

gcctgttctgaggtccctccagcatcacttgtaatttgtaaagtactcaagtacagactagaactagtca -1281

tgcataccagaagttgaggcctgagatgccgtgccactctatttgctctgtttcatgctttgccgttatt -1211

atatcactattttcactgttcattctaaggccgttgtcacacgggtaccagaagagcagaacaccttttt -1141

acgttccttttgctgctatggatggacttcgagagttccttgagaaagcactctcatctatctttccaat -1071

tccttgtttccttgtggatcctgatcactgaagatggccatttgatttttcacttttcagttatatataa -1001

agtgaagacggttgtctcacgcaatatgtgaagtgcaagatgagctctacactgcaaggaaccccaagaa -931

aaaaccacaggtaagtttgtctgattgacaacggctagagctgtctgacgatgcagatgactccgcatgt -861

tcattcatcactttaatttgttagaaaagctacacaaaaagcacatcagaaaaaaagaagaaaaagctat -791

ttagtggtactactagttgcaacttgcaataatgatgatgataaatctgcacaagccatagctatgctat -721

atgctatagctatgtatatgtacacaaaaaatacatttttgtgctatttttttaccgctagtataata -651

tccatgtcttgctacaacacacaatcatatttaatacctataaaaaataaatttaatattaaaataaaca -581

tatggtccacaccatatattaaactgctaaaaacaaatattataactcatttgatcgttcatcctctttc -511

ggttagtgaggtggacagtgagagcgctgcatcgtgttattgggtttgactggtttctcacggctcatct -441

```
gtgttgtaacgacctatctatggtcaaacaaactattaggattattgttaggcgaaaaatgagggagata -371

aataaacctataagcaggacacatgaaacatatgctttaaacagtagagattaacaacataaaataaaaa -301

aaggcaaaccgggaagcgtgcaagcccaaaggccaatcggcccagaaccagctccacggttgttcgtgtc -231

cgcccacgtggcacgcccggccattccgggcccacagatccggctcagggtcatgtgccactgccatggc -161
   G-box
cccctccttgtcgcagcggcagattgcgacggggaagaaaggtgaggagaccaagcgaaaaaaatcacgc -91

tttcattgcgagggggggaggatagggaggaagacgccaagccagctccagtccggcacccgatataaagc -21
                    ▼                                              TATA-box
ggcaggcacttggattgctg ACGAGATGGGACTCACACGCCGTCACAATGCTGTGCCTCGTGTCGCCCTC   50
                                                M  L  C  L  V  S  P  S

TTCCTCGCCGACTCCGCTTCCGCCGCCGCGGCGCTCTCGCTCGCATGCTGATCGGGCGGCACCGCCGGGG   120
  S  S  P  T  P  L  P  P  P  R  R  S  R  S  H  A  D  R  A  A  P  P  G

ATCGCG gtaagctgcggcgatccgggggctggtggatttactacttttgcttttttttcttctgtactga   190
 I  A

tggttcatagtttcgcag GGTGGCGGCAATGTGCGCCTGAGTGTGTTGTCTGTCCAGTGCAAGGCTCGCC   260
                    G  G  G  N  V  R  L  S  V  L  S  V  Q  C  K  A  R

GGTCAGGGGTGCGGAAG gtagatctttccccagattagaacacgttgttacggacaaaatggaatttggt   330
R  S  G  V  R  K

agtagttttggagagagaaaaaaaactctggaatttgtggattgggtcctgctcgctttggatttggctc   400

tagtgctttgttatcagtagtaatcattttaccaagtctagcggatctcagatttcaactgaagttgttt   470

aaattttggtggttggttttcttcagccatcctaatcatttattacataataattagttgaaataaagtg   540

ctatatttgatatttctgtggttgctgcagcagctaattagcttcactacattttcacgtgttagcttct   610

attgctcgtgtaattctgaatctggctttaacgcactgtgcaag GTCAAGAGCAAATTCGCCACTGCAGC   680
                                              V  K  S  K  F  A  T  A  A

TACTGTGCAAGAAGATAAAACTATGGCAACTGCCAAAGGCGATGTCGACCATCTCCCCATATACGACCTG   750
  T  V  Q  E  D  K  T  M  A  T  A  K  G  D  V  D  H  L  P  I  Y  D  L

GACCCCAAGCTGGAGATATTCAAGGACCATTTCAGGTACCGGATGAAAAGATTCCTAGAGCAGAAAGGAT   820
 D  P  K  L  E  I  F  K  D  H  F  R  Y  R  M  K  R  F  L  E  Q  K  G

CAATTGAAGAAAATGAGGGAAGTCTTGAATCTTTTTCTAAG gttaggcttaatacattcaaatgctaat   890
S  I  E  E  N  E  G  S  L  E  S  F  S  K

acaaggaccatttcacaatatttatagcacctgcccatgccgcaaatattatctgcatcatcaacattt   960
```

```
ccggattttgtatttttttcag GCTATTTGAAATTTGGGATTAATACAAATGAGGATGGAACTGTATATC 1030
                        G  Y  L  K  F  G  I  N  T  N  E  D  G  T  V  Y

GTGAATGGGCACCTGCTGCGCA gtaagttctaatgttgtcatgcaaacatgatgtactggcggggtatcg 1100
R  E  W  A  P  A  A  Q

ttttttcccattttgcttgtgaagagatatcgtctgatagtaactatattaaaaaaaaagaacagctaaa 1170

acacaggagagtaatcaatgtttgggagtagaaaactgttcggatatttcattcttgtgcaagactgtca 1240

tgtcttagctgaggaggctattttttttcttgctgtgattacataatgctgttttcttcttatgacaata 1310

tactttccactattttaatagattgatgcatttgacttgagttttttactcatgggtgtag GGAGGCAGA 1380
                                                                E  A  E

GCTTATTGGTGACTTCAATGACTGGAATGGTGCAAACCATAAGATGGAGAAGGATAAATTTGGTGTTTGG 1450
 L  I  G  D  F  N  D  W  N  G  A  N  H  K  M  E  K  D  K  F  G  V  W

TCGATCAAAATTGACCATGTCAAAGGAAAACCTGCCATCCCTCACAATTCCAAGGTTAAATTTCGCTTTC 1520
 S  I  K  I  D  H  V  K  G  K  P  A  I  P  H  N  S  K  V  K  F  R  F

TACATGGTGGAGTATGGGTTGATCGTATTCCAGCATTGATTCGTTATGCGACTGTTGATGCCTCTAAATT 1590
L  H  G  G  V  W  V  D  R  I  P  A  L  I  R  Y  A  T  V  D  A  S  K  F

TGGAGCTCCCTATGATGGTGTTCATTGGGATCCTCCTGCTTCTGAAAG gtctctttctactgtcctggac 1660
  G  A  P  Y  D  G  V  H  W  D  P  P  A  S  E  R

cacattgtaaacttaagcaatagaccttagcagacaaaaatatatgacaaaattatagtaccttttacct 1730

ctggtttgcag GTACACATTTAAGCATCCTCGGCCTTCAAAGCCTGCTGCTCCACGTATCTATGAAGCCC 1800
             Y  T  F  K  H  P  R  P  S  K  P  A  A  P  R  I  Y  E  A

ATGTAGGTATGAGTGGTGAAAAGCCAGCAGTAAGCACATATAGGGAATTTGCAGACAATGTGTTGCCACG 1870
H  V  G  M  S  G  E  K  P  A  V  S  T  Y  R  E  F  A  D  N  V  L  P  R

CATACGAGCAAATAACTACAACACAGTTCAGTTGATGGCAGTTATGGAGCATTCGTACTATGCTTCTTTC 1940
 I  R  A  N  N  Y  N  T  V  Q  L  M  A  V  M  E  H  S  Y  Y  A  S  F

GGGTACCATGTGACAAATTTCTTTGCGGTTAGCAGCAGATCAGGCACACCAGAGGACCTCAAATATCTTG 2110
G  Y  H  V  T  N  F  F  A  V  S  S  R  S  G  T  P  E  D  L  K  Y  L

TTGATAAGGCACACAGTTTGGGTTTGCGAGTTCTGATGGATGTTGTCCATAGCCATGCAAGTAATAATGT 2180
V  D  K  A  H  S  L  G  L  R  V  L  M  D  V  V  H  S  H  A  S  N  N  V

CACAGATGGTTTAAATGGCTATGATGTTGGACAAAGCACCCAAGAGTCCTATTTTCATGCGGGAGATAGA 2250
 T  D  G  L  N  G  Y  D  V  G  Q  S  T  Q  E  S  Y  F  H  A  G  D  R
```

GGTTATCATAAACTTTGGGATAGTCGGCTGTTCAACTATGCTAACTGGGAGGTATTAAGGTTTCTTCTTT 2320
G Y H K L W D S R L F N Y A N W E V L R F L L

CTAACCTGAGATATTGGTTGGATGAATTCATGTTTGATGGCTTCCGATTTGATGGAGTTACATCAATGCT 2390
S N L R Y W L D E F M F D G F R F D G V T S M L

GTATCATCACCATGGTATCAATGTGGGGTTTACTGGAAACTACCAGGAATATTTCAGTTTGGACACAGCT 2460
Y H H H G I N V G F T G N Y Q E Y F S L D T A

GTGGATGCAGTTGTTTACATGATGCTTGCAAACCATTTAATGCACAAACTCTTGCCAGAAGCAACTGTTG 2530
V D A V V Y M M L A N H L M H K L L P E A T V

TTGCTGAAGATGTTTCAGGCATGCCGGTCCTTTGCCGGCCAGTTGATGAAGGTGGGGTTGGGTTTGACTA 2600
V A E D V S G M P V L C R P V D E G G V G F D Y

TCGCCTGGCAATGGCTATCCCTGATAGATGGATTGACTACCTGAAGAATAAAGATGACTCTGAGTGGTCG 2670
R L A M A I P D R W I D Y L K N K D D S E W S

ATGGGTGAAATAGCGCATACTTTGACTAACAGGAGATATACTGAAAAATGCATCGCATATGCTGAGAGCC 2740
M G E I A H T L T N R R Y T E K C I A Y A E S

ATGATCAG gtaccctgcattatataataactttataaggaagtatagttggatgcagcaaagtattttat 2810
H D Q

ttccttgtggggtagacaacacatggggcaagtatcttttaatcagacagtactatatttcatcacattt 2880

ctatcctgattgctagtaactggattagttattcataaaaaattaggccctcctaagcatgtttaagaat 2950

ttgcaatcacatatgacctgtttatcttatactggaattaaaatcatgcaaattttcag TCTATTGTTG 3020
S I V

GCGACAAAACTATTGCATTTCTCCTGATGGACAAGGAAATGTACACTGGCATGTCAGACTTGCAGCCTGC 3090
G D K T I A F L L M D K E M Y T G M S D L Q P A

TTCACCTACAATTGATCGAGGGATTGCACTCCAAAAG gttccattctcctcaagttctggttgaacctgg 3160
S P T I D R G I A L Q K

ttttcaataattcttgaatgagttccctcaatataccttgcttattgtgttctgtacactccag ATGATT 3230
M I

CACTTCATCACAATGGCCCTTGGAGGTGATGGCTACTTGAATTTTATGGGAAATGAG gtgaaatctcggt 3300
H F I T M A L G G D G Y L N F M G N E

cttttgaaaaatgtttcccattaataagcctgcataaaccttttgtatttttttttctatttgaggtttca 3370 gcatttagagtcgaccctctatttctcagtaactgaagttcctaatggctatttgaatggttaatgtta 3440

attccgagtagcgctcgttgcacatgaaaatgtggtctgtttatggctgctaatccttttgtgacttgtc 3510

caatatactcgtacaacatgggatacatgtaacccttcactatcaagcatcgcttaaatgtcaacatcat 3580

acagagatattatcaagtctttctcctgcactatcctggaatgaaaaaaaagagttgtttcatatcgtat 3650

ctgtacttttacaaatgagtgtgatctttggagaagttagttgacggaatggtccactgtatttgtttg 3720

ttatatgtttctccttattcatagtgtattcaattgatgtgcaattctcttctag TTTGGTCACCCAGAA 3790
F G H P E

TGGATTGACTTTCCAAGAGAAGGGAACAACTGGAGCTATGATAAATGCAGACGACAGTGGAGCCTTGTGG 3860
W I D F P R E G N N W S Y D K C R R Q W S L V

ACACTGATCACTTGCGGTACAAG gttatgtctatgaatgcaatccttataagattttgttctggcacca 3930
D T D H L R Y K

ttctaggctctctttgacctttacctcttccatttatcgctgtctcctatactagctagatctgtctctg 4000

agactagctagatgtcagttccttactactattacactcttgagttgagctgggtccagatatgcttttc 4070

cacactttgttttgctgccccactgccctactatagtaaattgcaccaaccccgtacatgttactcaata 4140

aacaggttggattggcctgtgcaaacctagtattttgtaagttgacctggctttatgtattgccattgtt 4210

gaataaaaataattttgatgcattgattggtatgtcattttcataactttgcagacctaggaatgtactt 4280

ctcaatcattggttgttttgttgatagaaattattctgtttacagcatatggtggactaaatatccaagt 4350

gaacgtttgttttgtctgtttttttttctacccctcaagtttacaatttatcccggctaacatattaatg 4420

ttttctccttggctctaatgctgaacag TACATGAATGCGTTTGACCAAGCGATGAATGCGCTCGATGAG 4490
Y M N A F D Q A M N A L D E

AGATTTTCCTTCCTTTCGTCGTCAAAGCAGATCGTCAGCGACATGAACGATGAGGAAAAG gtaaggattt 4560
R F S F L S S S K Q I V S D M N D E E K

cagaatacttggtatgagagtgttgcgacgtttctttctttcttaggaagagtggtaactctagcttgtg 4630

tgctgtattttgccgattaaggcgagcttcaatttaggagtggatggtttgtcagcattatttttctac 4700

ttcgtttcag GTTATTGTCTTTGAACGTGGAGATTTAGTTTTTGTTTTCAATTTCCATCCCAAGAAAACT 4770
V I V F E R G D L V F V F N F H P K K T

```
TACGAGGG gtaagtcaccagttgtaaaaccctgtcttttcagagtccctttgcctgtggtataatataa 4840
 Y  E  G

tatagtgtgcttacctcccatctctgctctgcaacttatatgcag CTACAAAGTGGGATGCGATTTGCCT 4910
                                                Y  K  V  G  C  D  L  P

GGGAAATACAGAGTAGCCCTGGACTCTGATGCTCTGGTCTTCGGTGGACATGGAAGA gtaagtagtgacg 4980
 G  K  Y  R  V  A  L  D  S  D  A  L  V  F  G  G  H  G  R

gtagacgctgaaagatgtttttttttttttaattttgccctactcccgtagttggggccggccggccgtc 5050

gagaatcgatcatatcaggctgtcgtgtttgcgagactcatggatgctgtgacgcag GTTGGCCACGACG 5120
                                                            V  G  H  D

TGGATCACTTCACGTCGCCTGAAGGGGTGCCAGGGGTGCCCGAAACGAACTTCAACAACCGGCCGAACTC 5190
V  D  H  F  T  S  P  E  G  V  P  G  V  P  E  T  N  F  N  N  R  P  N  S

GTTCAAAGTCCTTTCTCCGCCCCGCACCTGTGTG gtaatgttcacttactctcagctgaactgatgagct 5260
  F  K  V  L  S  P  P  R  T  C  V

agaatgtatccgccctgacaaaccgtcctgctgccag GCTTATTACCGTGTAGACGAAGCAGGGGCTGGA 5330
                                       A  Y  Y  R  V  D  E  A  G  A  G

CGACGTCTTCACGCGAAAGCAGAGACAGGAAAGACGTCTCCAGCAGAGAGCATCGACGTCAAAGCTTCCA 5400
 R  R  L  H  A  K  A  E  T  G  K  T  S  P  A  E  S  I  D  V  K  A  S

GAGCTAGTAGCAAAGAAGACAAGGAGGCAACGGCTGGTGGCAAGAAGGGATGGAAGTTTGCGCGGCAGCC 5470
R  A  S  S  K  E  D  K  E  A  T  A  G  G  K  K  G  W  K  F  A  R  Q  P

ATCCGATCAAGATACCAAATGAAGCCAGGAGTCCTTGGTGAGGACTGGACTGGCTGCCGGCGCCCTGTTA 5540
  S  D  Q  D  T  K  *

GTAGTCCTGCTCTACTGGACTAGCCGCCGCTGGCGCCCTTGGAACGGTCCTTTCCTGTAGCTTGCAGGCG 5610

ACTGGTGTCTCATCACCGAGCAGGCAGGCACTGCTTGTATAGCTTTTCTAGAATAATAATCAGGGATGGA 5680

TGGATGGTGTGTATTGGCTATCTGGCTAGACGTGCATGTGCCCAGTTTGTATGTACAGGAGCAGTTCCCG 5750
                                       ●
TCCAGAATAAAAAAAAAAACTTGTTGGGGGGTTTTTCTTCTA ctctgtagtctgttcttgctaggttgacga 5820
    Poly (A) signal
agatgtttgatattagatgccatagatcatgtactgttaagtttcttcgttccttgttccctgtccagtt 5890

cacatttgattccagctgttcagcaggccggtcagctcagctccacaccgggggccaggccggcntcacg 5960

caggatcttttcatttcatgcgtcataacacaaacacttttgatcttttcagacaaaaaagatggatcg  6029
```

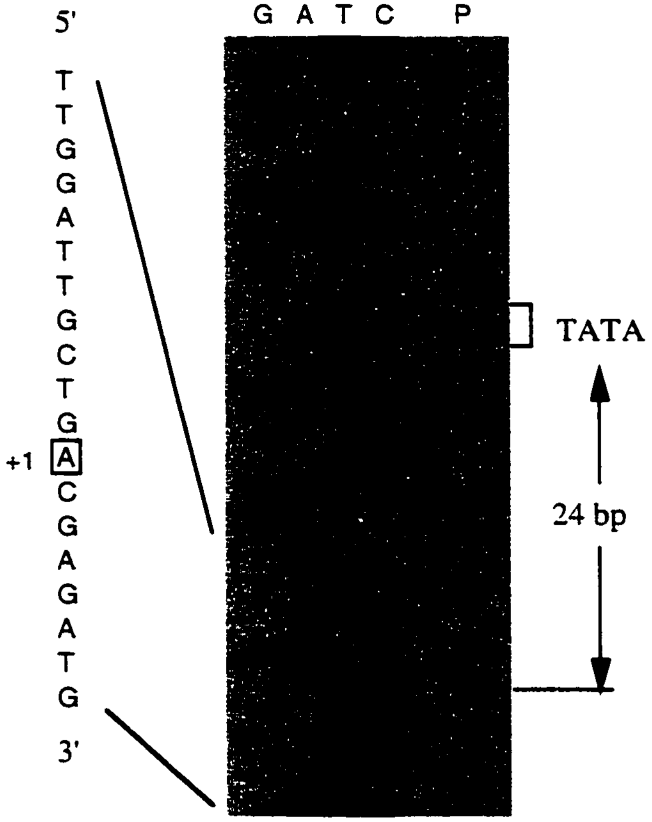


Figure 2.5. Primer extension analysis of the transcription initiation site of the *Sbe1* gene.

A ^{32}P-end labeled antisense primer was annealed to 10 μg of total RNA isolated from maize kernels (30 DAP) and extended using reverse transcriptase. The extended cDNA product was then analyzed on a 5% sequencing gel (lane P) along with a *Sbe1* 5'-flanking genomic DNA sequencing ladder (G, A, T, C) generated by the same primer. The number inside the vertical double-headed arrow on the right refers to distance from TATA-box to the primer extension product indicated by the horizontal arrow. The sense sequence around the product is shown on the left and the transcription initiation site is boxed and assigned +1.

To determine the polyadenylation site of the *SbeI* gene, 3' RACE was conducted (data not shown). It demonstrated that a poly (A) tail occurs at an adenine nucleotide located 29 bp downstream from a putative polyadenlyation signal (AATAAA) in the *SbeI* gene (Figure 2.4). Along with the primer extension result, this indicates that the transcribed region in the *SbeI* gene is 5,690 bp in length.

2.3.3. Structure of the *SbeI* Genomic Clone

Alignment of the genomic sequence with the published maize *SbeI* cDNA sequence (Fisher et al., 1995) revealed that the gene is composed of 14 exons and 13 introns distributed over 5.7 Kb. Figure 2.3(B) summarizes the organization of the maize *SbeI* gene. The cDNA sequence is identical to the corresponding genomic sequence except for a 1 bp mismatch in exon 14. Table 2.3 shows the sequences around the exon/intron junctions and a list of putative branch point consensus sequences, which were derived as described by Brown (1986). The introns, relatively AT-rich (61%) compared to the exons (52%), vary in length from 73 bp to 565 bp, and all of which have the conserved sequences at their 5' and 3' ends, following the "GT··AG" rule of plant introns (Brown, 1986). Exon 1, containing 27 bp of 5' untranslated DNA sequence, and exon 2 occur in the transit peptide region which may be essential for transporting the gene product into the amyloplast. Exon 14 contains the translation stop codon (TGA) and 3' untranslated region as well as the putative polyadenylation signal (AATAAA). The exons vary in length from 63 to 907 bp.

Table 2.3. List of introns and sequences of exon/intron borders in the *Sbe1* gene.

Intron number	Exon /		Putative intron branch point	/Exon	Intron size(bp)	GC content (%)
1	TCGCG	GTAAG ····	CTGAT··17··CGCAG	GGTGG	82	46.3
2	GGAAG	GTAGA ····	CTGAA··24··GGAAG	GTCAA	377	36.9
3	TAAAG	GTTAG ····	ATCAT··30··TTCAG	GCTAT	120	34.2
4	GCGCA	GTAAG ····	TTGAC··25··TGTAG	GGAGG	319	34.2
5	GAAAG	GTCTC ····	ATGAC··33··TGCAG	GTACA	103	35.9
6	ATCAG	GTACC ····	ATGAC··43··TTCAG	TCTAT	262	32.1
7	AAAAG	GTTCC ····	CTCAA··33··TCCAG	ATGAT	97	39.2
8	ATGAG	GTGAA ····	TTGAT··17··TCTAG	TTTGG	488	35.2
9	ACAAG	GTTAT ····	CTAAC··37··AACAG	TACAT	565	37.3
10	AAAAG	GTAAG ····	GTCAG··25··TTCAG	GTTAT	160	39.4
11	GAGGG	GTAAG ····	CTTAC··31··TGCAG	CTACA	107	41.1
12	GAAGA	GTAAG ····	CTCAT··16··CGCAG	GTTGG	140	49.3
13	GTGTG	GTAAT ····	CTGAC··18··GCCAG	GCTTA	73	49.3

Consensus sequences between introns are underlined.

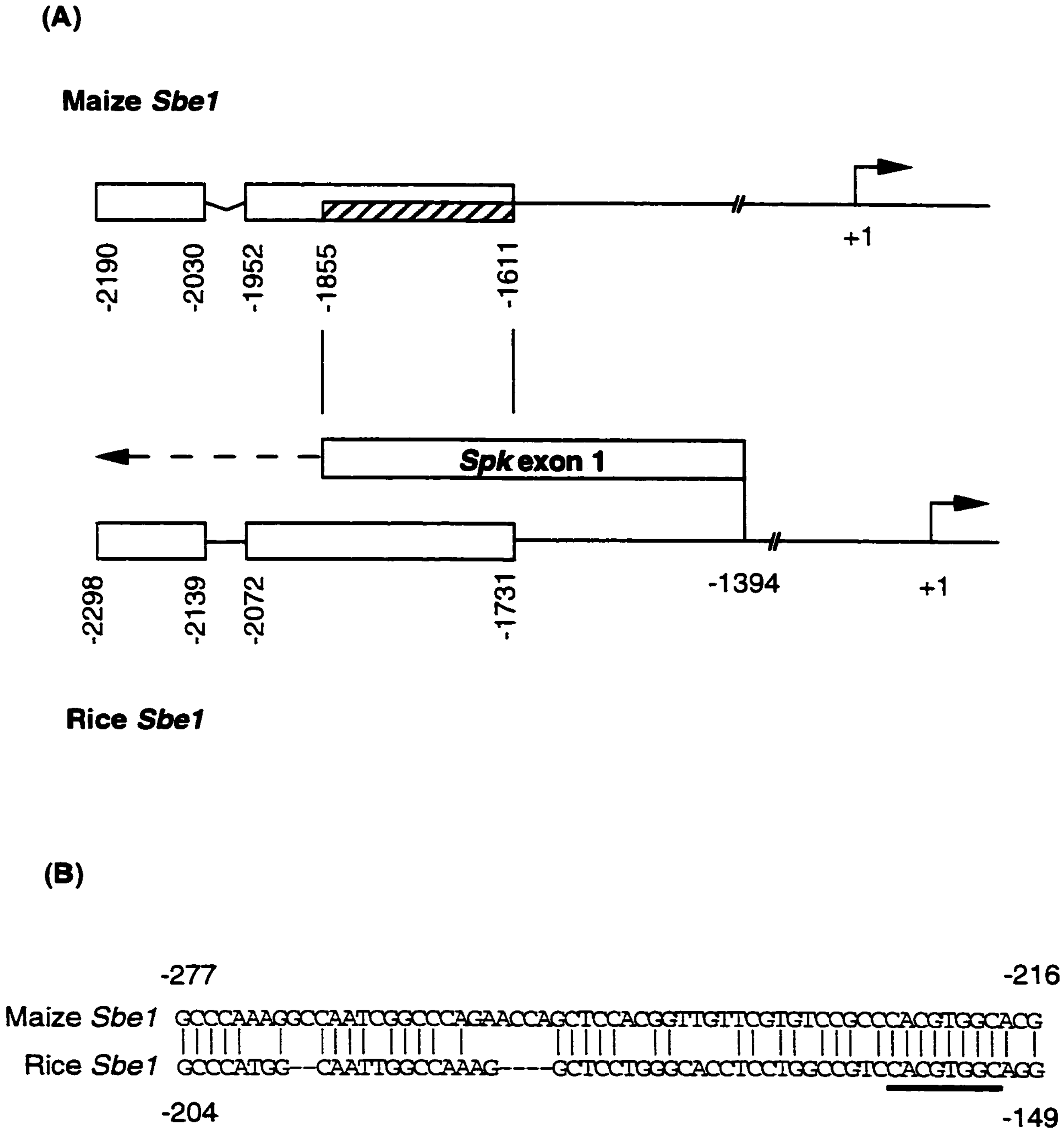
Comparison of the maize and rice *Sbe1* genomic DNA sequences (Kawasaki et al., 1993a) revealed two large highly conserved regions in the 5'-flanking sequences. The results are shown in Figure 2.4 and Figure 2.6(A). One, 161 bp in length, is present between -2190 and -1890 in the maize *Sbe1* 5'- flanking sequence, and has 82% similarity with the corresponding rice region. The other, 342 bp in length, is located from -1804 to -1611, and shows 85% similarity with the corresponding rice region. These sequences conserved between the species suggest that the regions may play a very important role in gene expression.

Interestingly, a part of the latter region from -1804 to -1611 (194 bp in length) shares 83% similarity with the 5' untranslated sequence of a Ca^{2+}-dependent protein kinase gene in rice (Kawasaki et al., 1993b), suggesting that a maize version of protein kinase gene may be located immediately upstream of the *Sbe1* gene. In the rice genome, the two genes are separated by approximately 1.4 kb, and transcribed divergently from each other (Figure 2.6(A)).

5'-flanking sequences downstream from the two conserved regions did not show sequence similarity between the two genes (less than 25%) except for a G-box motif and several small segments adjacent to the G-box as shown in Figure 2.6(B). This may indicate that the G-box plays a role in regulation of *Sbe1* gene expression. The G-box is found in other plant genes which respond to diverse environmental or physiological stimuli and are often associated with additional regions which possibly act as coupling elements determining signal response specificity (Menken et al., 1995).

Figure 2.6. Sequence comparison between the maize and rice *SbeI* promoters.

(A) Highly conserved sequences found in the distal promoter regions. The open boxes indicate the conserved regions containing at least 82% sequence similarity between the two *SbeI* promoters. The striped box denotes 83% similarity between the maize *SbeI* and the rice *Spk* (Ca2+-dependent protein kinase) cDNA sequence which is represented by the stippled box. The arrows indicate transcription directions and the numbers indicate the distance relative to the relevant transciption start sites.
(B) Proximal sequences showing conservation between the two *SbeI* promoters. Short verticle lines indicate conserved nucleotides and dashes represent gaps to maximize aligment. The G-box is underlined.



```
Maize Sbe1  GCCCAAAGGCCAATCGGCCCAGAACCAGCTCCACGGTTGTTCGTGTCCGCCCACGTGGCACG
Rice Sbe1   GCCCATGG--CAATTGGCCAAAG----GCTCCTGGGCACCTCCTGGCCGTCCACGTGGCAGG
```

Other notable feature derived from the sequence comparison is that the maize and rice *SbeI* genomic structures are quite similar. As shown in Figure 2.7, both genes consist of 14 exons and 13 introns, the positions of which are conserved between the two species. The sizes of exons (exon 3 to exon 13) constituting most of the mature proteins are identical except exon 5, in which the rice *SbeI* gene has one more codon for glycine residue compared to the maize gene. They share more than 86% and almost 90% similarity in nucleotide and amino acid sequences, respectively. However, exons not encoding the mature proteins (exon 1, 2 and most of exon 14) do not display any significant sequence similarity and vary in size. The carboxyl-terminal 64 (67 in rice) amino acids encoded by exon 14 in the maize gene is the only region which is not conserved in the mature protein. Unlike the exons, homology was not found in any of the introns other than the splice junction sequences. The large intron (2212 bp) present in the rice *SbeI* gene is not found in the maize gene.

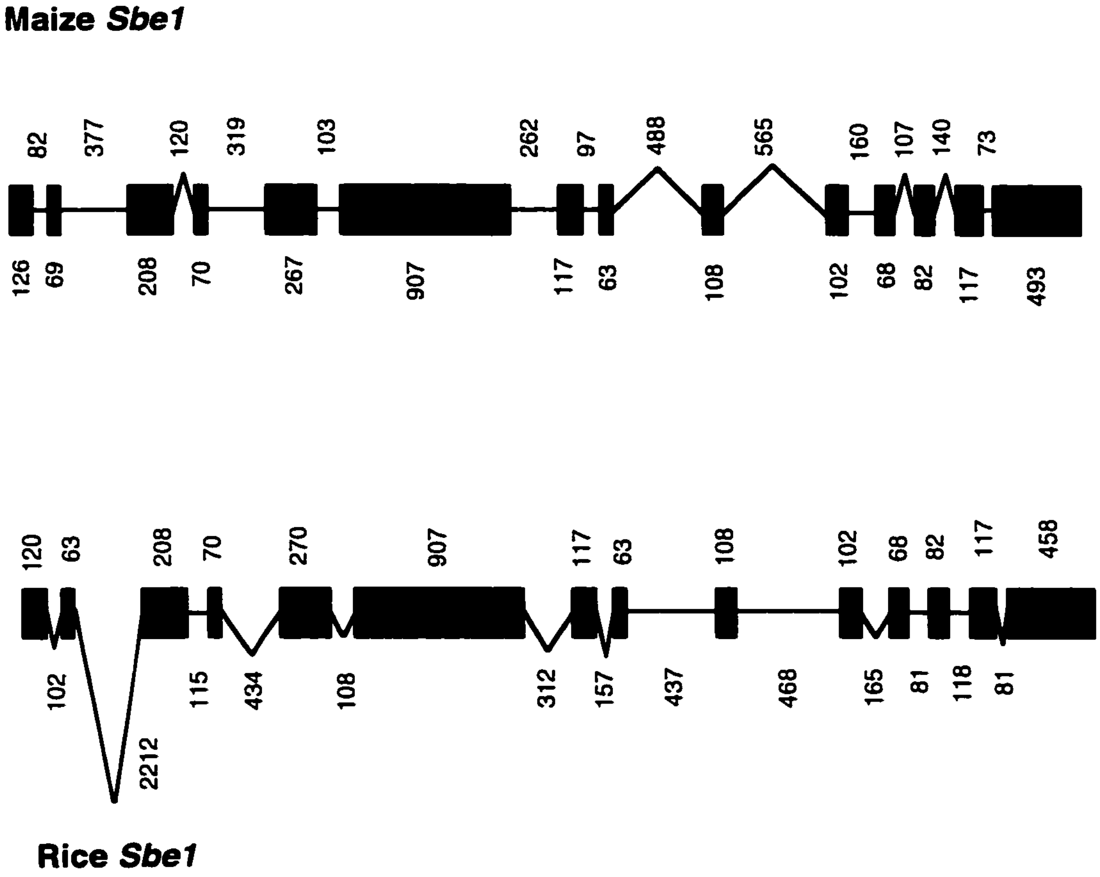


Figure 2.7. Structures of the maize and rice *SbeI* genes (Kawasaki et al.,1993a).

Black boxes indicate exons and thin lines denote introns. Bent lines indicate size differences between corresponding introns. Numbers represent the sizes of exons and introns.

2.3.4. Genomic Southern Blot Analysis

To determine the number of *SbeI* genes in the maize genome, Southern blot analysis was performed. When blots were probed with the full-length maize *SbeI* cDNA (Fisher et al., 1995) under high-stringency conditions, at least three hybridizing bands were observed in each lane as shown in Figure 2.8. Comparison of the hybridization patterns with the restriction map of the *SbeI* genomic clone revealed that not all the bands in the Southern blot corresponded to the genomic map, suggesting more than one *SbeI* gene is present in the maize genome.

To confirm this, a 0.6-kb genomic DNA probe which does not have any restriction enzyme sites used in the genomic blot was prepared from the genomic clone 5-1-1 by BamHI-HindIII digestion. The genomic probe will produce only one hybridizing band in every lane if there is a single copy of the *SbeI* gene in the maize genome. As shown in Figure 2.9, however, the genomic probe (probe 2) containing the central region of the *SbeI* cDNA detected in each lane one or two additional bands apart from the bands predicted by the genomic map. This indicates that along with the isolated *SbeI* gene, another *SbeI* gene or a gene very closely related to *SbeI* exists in the maize genome. Interestingly, when a 1.7-kb BamHI-PstI genomic fragment consisting of the *SbeI* promoter and transit peptide-coding region was used as a probe (probe 1), only the bands predicted from the identified *SbeI* genomic DNA sequence were detected as shown in Figure 2.9. Taken together, these results suggest that although two *SbeI* genes are present in the maize genome, their 5' flanking sequences and the 5'-end of the coding regions (at least in DNA level) are quite divergent from each other.

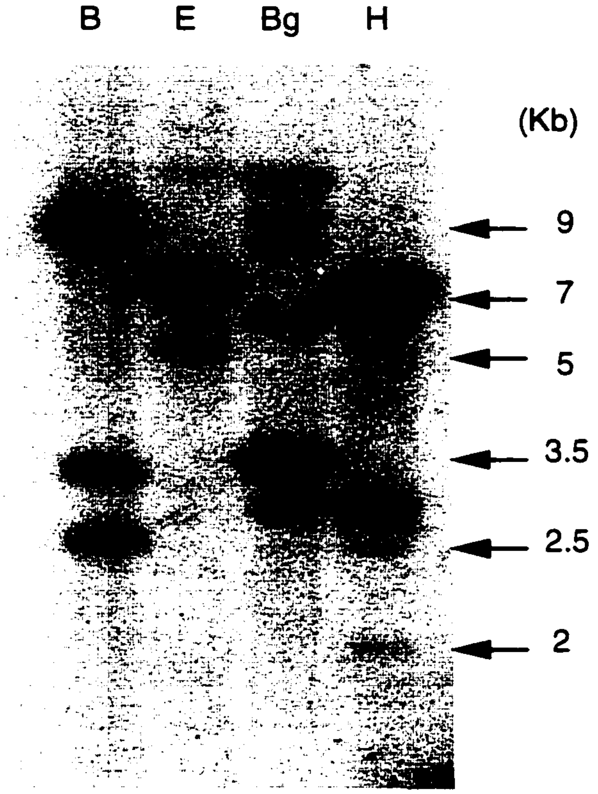


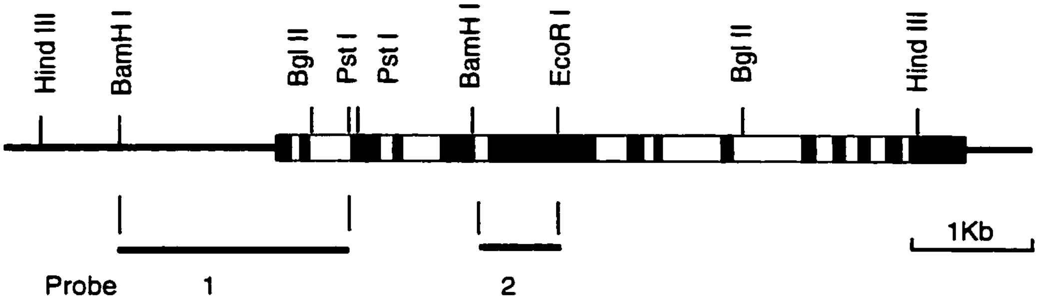
Figure 2.8. Southern blot analysis of maize genomic DNA probed with the full-length *SbeI* cDNA.

Each lane contains 10 μg of maize genomic DNA digested with the indicated restriction enzymes: B, BamH I; E, EcoR I; Bg, Bgl II; H, Hind III. Genomic DNA was prepared from etiolated maize seedlings (inbred B73). The maize *SbeI* full-length cDNA was ^{32}P-labeled using the random primed DNA labeling kit (Boehringer Mannheim) and used as a probe. Hybridization and washes were performed at high-stringency conditions. Arrows indicate the position of the DNA size markers in kilobases.

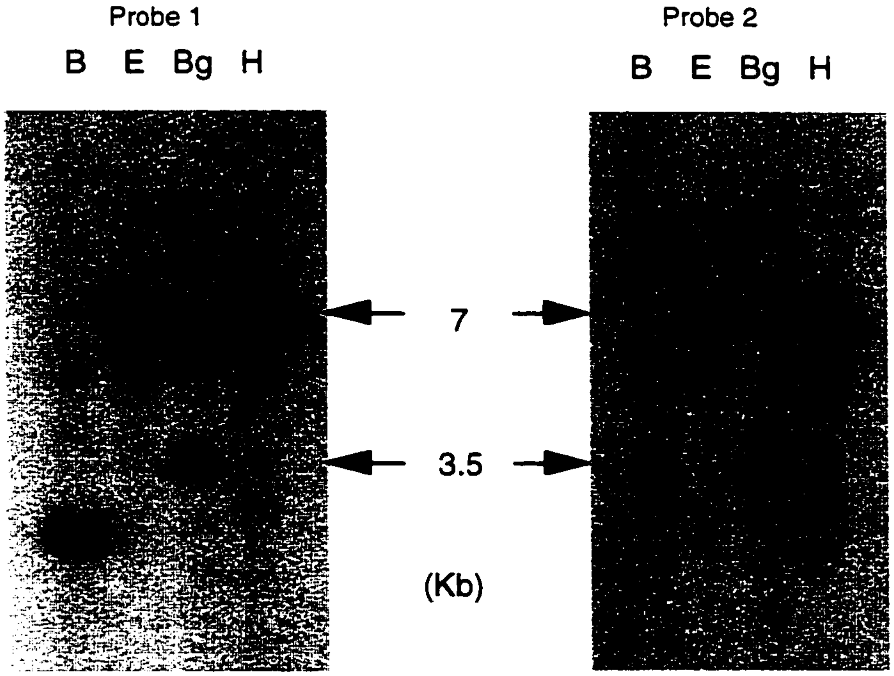
Figure 2.9. Southern blot analysis of maize genomic DNA probed with partial *SbeI* genomic DNA fragments.

(A) Schematic diagram of the *SbeI* genomic clone displaying the locations of restriction sites and the different probes used for genomic DNA analysis shown in (B).
(B) Autoradiography of genomic DNA gel blot hybridization analysis. Each lane contains 10 μg of maize genomic DNA digested with the indicated restriction enzymes: B, BamH I; E, EcoR I; Bg, Bgl II; H, Hind III. Blot 1 and 2 were probed at high stringency with the probe 1 and 2, respectively. Bands which were not predicted from the restriction map of the *SbeI* genomic clone are indicated by asterisks at right. Arrows indicate the position of the DNA size markers in kilobases.

(A)


Hind III
BamH I
Bgl II
Pst I
Pst I
BamH I
EcoR I
Bgl II
Hind III
1Kb
Probe
1
2

(B)


Probe 1
B
E
Bg
H
Probe 2
B
E
Bg
H
7
3.5
(Kb)

2.3.5. Genetic Mapping of *Sbe* Genes

As part of the mapping efforts of the University of Missouri-Columbia Maize RFLP Laboratory, two genomic loci have been mapped for *Sbe1* genes on chromosome 6, bin 6.01 and chromosome 10, 10.04 using a mapping population of 54 immortalized F2 individuals from a cross of Tx303 x CO159. This supports the conclusion made above based on genomic Southern analysis that two *Sbe1* genes exist in the maize genome.

2.3.6. Effect of Exon / Introns and 3'-flanking Regions on Expression in Maize Endosperm Suspension Cells.

To determine whether the 5' flanking sequence of the cloned *Sbe1* gene (λ5-1-1) has all of the necessary DNA elements to initiate transcription, a 2.2-kb fragment upstream of the translation start site (-2191 to +27) was fused to the luciferase (LUC) reporter gene in pUC119 (pKL101) as shown in Figure 2.10. The chimeric plasmid was then introduced into maize endosperm cells via particle bombardment along with a reference plasmid containing the cauliflower mosaic virus (CaMV) 35S promoter linked to a GUS gene (pBI221, Clontech) to correct for transfection efficiency. Promoter activity was assessed by measuring the level of LUC expression. However, only low levels of LUC activity were detected.

Since many reports indicated that DNA sequences within transcribed regions such as exons, introns and 3' flanking regions are involved in the expression of genes in either qualitative or quantitative manner (Callis et al., 1987; Hamilton et al., 1992; Fu et al., 1995; Ulmasov and Folk, 1995), three different types of translational fusion constructs were created to test the role of downstream elements on *SbeI* gene expression. First, as shown in Figure 2.10, the 5'-flanking sequence as well as the first exon and intron of the *SbeI* gene (-2190 to +228) were fused in-frame to the LUC reporter gene to make pKLN101. Second, the nopaline synthase (NOS) 3' sequence in the pKLN101 was replaced with the *SbeI* 3' flanking sequence (399 bp in length), which contains the translation stop codon and polyadenlylation signal to create pKLNS101. Finally, to determine whether an increase in the number of exon/introns enhances the gene expression, three more exons and introns from the *SbeI* gene were added to the pKLN101 to make pKLM101 (-2190 to +1617).

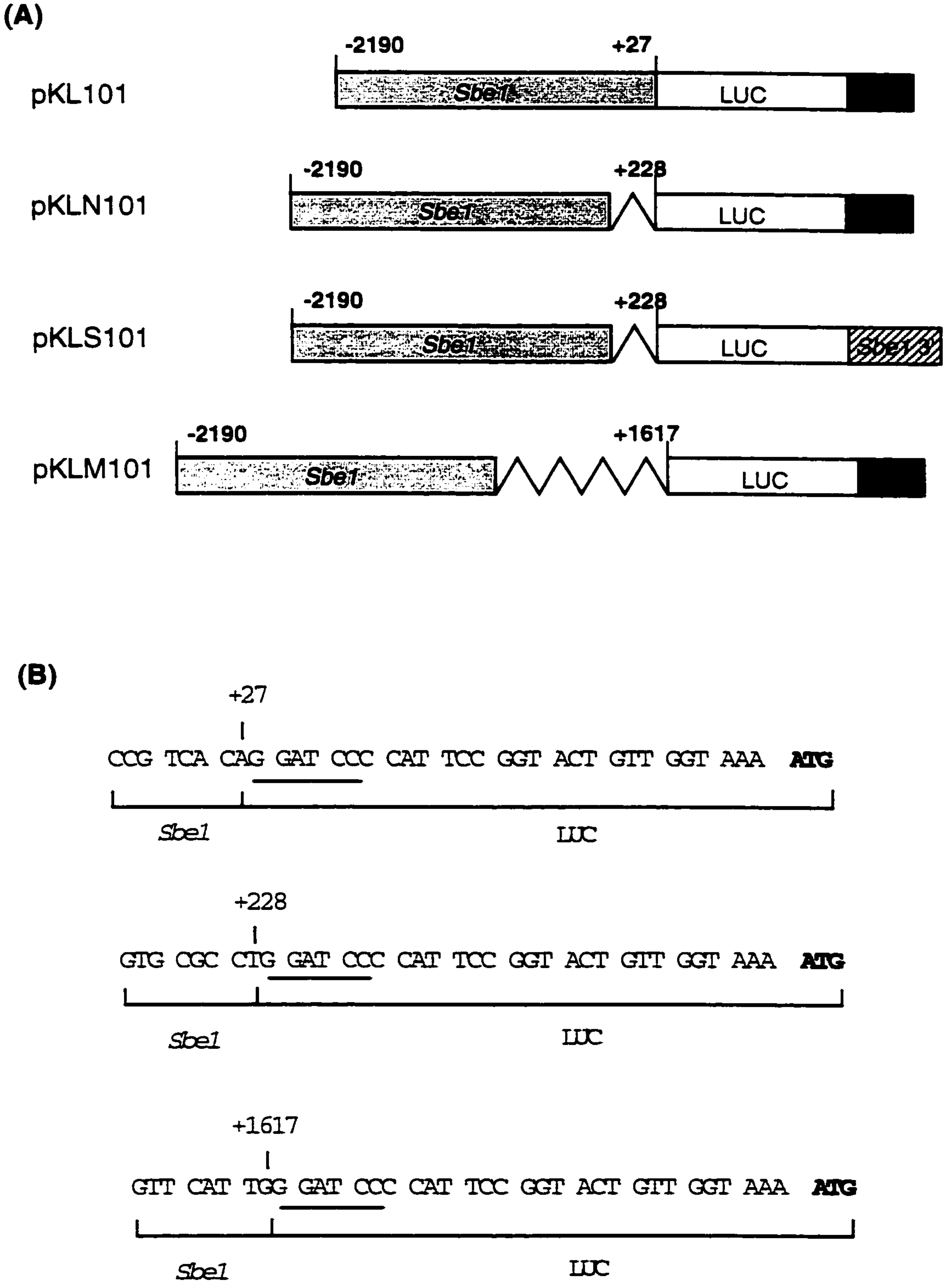
The results of transient expression assays using the chimeric constructs are shown in Figure 2.11. Including the DNA sequence (+28 to +228) containing the first exon and intron into the plasmid pKL101 dramatically increased the level of LUC expression. The LUC/GUS activity of pKLN101 was 14-fold higher than that of pKL101. This suggests that the first exon and intron region is required for high level expression of the *SbeI* gene in maize endosperm cells. At present, however, it is not known how the region containing the first exon and intron increases gene expression. Further experiments are necessary to determine whether the increase is the result of transcriptional or translational effects.

Replacement of the NOS 3' end in pKLN101 with the *SbeI* 3' region did not have a significant effect on the level of LUC expression, implying that the *SbeI* 3' untranslated region does not have indispensable control elements. However, it is still possible that the region may be important for *SbeI* gene expression in other cell types or inductive conditions.

Construct pKLM101 showed a slight reduction in LUC activity compared to pKLN101, indicating that additional exons and introns had an adverse effect on LUC expression in maize endosperm suspension cells. The adverse effect could be explained by inefficient splicing, resulting from the introduction of multiple copies of the plasmid into a single cell, or by formation of fusion protein consisting of the 5'-end of SBEI and luciferase, thus lowering LUC activity. However, it could also be actually due to the presence of negative *cis*-elements in the additional region.

Figure 2.10. Schematic diagram of chimeric *SbeI* promoter-luciferase constructs.

(A) Numbers indicate distance relative to the *SbeI* transcription start site. Translation initiation starts at a position, + 28. The light grey (stippled) boxes indicate the *SbeI* promoter region. Angled lines indicate exons and introns in the *SbeI* gene. Open and solid black boxes indicate luciferase (LUC) reporter gene and nopaline synthase 3' end sequences, respectively. The striped box indicates the *SbeI* 3' flanking sequence.
(B) The junction sequences between the *SbeI* gene and LUC. The BamHI sites used to join the two genes are underlined. The translation start site of *LUC* is indicated by boldface letters.


(A)
pKL101
-2190
+27
Sbe1
LUC
pKLN101
-2190
+228
Sbe1
LUC
pKLS101
-2190
+228
Sbe1
LUC
Sbe1 3'
pKLM101
-2190
+1617
Sbe1
LUC
(B)
+27
CCG TCA CAG GAT CCC CAT TCC GGT ACT GTT GGT AAA ATG
Sbe1
LUC
+228
GTG CGC CTG GAT CCC CAT TCC GGT ACT GTT GGT AAA ATG
Sbe1
LUC
+1617
GTT CAT TGG GAT CCC CAT TCC GGT ACT GTT GGT AAA ATG
Sbe1
LUC

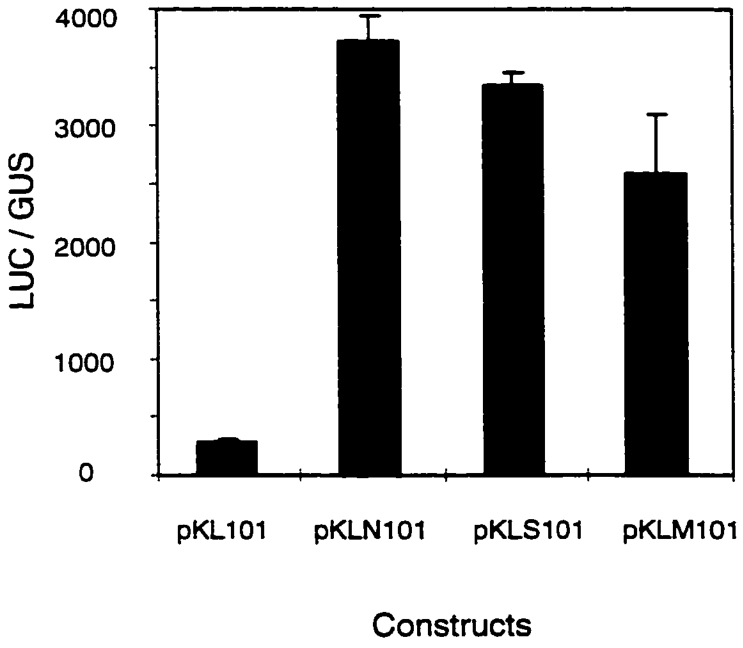


Figure 2.11. Effect of exon/introns and 3' end on the level of LUC expression driven by the *SbeI* promoter.

LUC/GUS ratios of the constructs described in Figure 8 were calculated as described in Methods. Each value represents the average of four independent shootings. Error bars indicate standard errors of the mean.

2.3.7. 5' Deletion Analysis

To identify promoter sequences critical for *Sbe1* expression in maize endosperm cells, a series of 5' deletion mutants were derived from pKLN101 as shown in Figure 2.12. The activity of each 5' deletion construct is presented in Figure 2.13. Removing the sequences to -1332 caused a decrease in the level of the LUC expression, while deletion of an additional 422 bp, to -910, resulted in an increase in the activity of the construct. This suggests that potential positive and negative distal *cis*-regulatory elements may be located in the regions from -2190 to -1332 and from -1332 to -910, respectively. Further deletions down to -315 did not significantly affect the promoter activity, but a severe reduction in the activity was observed when an additional 169 bp, to -145, was deleted. The -72 deletion construct produced a level of the LUC activity slightly over background, showing that the minimal promoter is functional.

To further delimit sequences essential for high level expression of the promoter, two additional 5' deletions with about 60-bp intervals were created between -315 and -145. The results are shown in Figure 2.14. A deletion to -255 severely reduced the expression of the LUC reporter gene, while a further deletion to -196 did not significantly change promoter strength. This indicates that a very strong positive regulatory element(s) is present in the 60-bp region between -315 and -255.

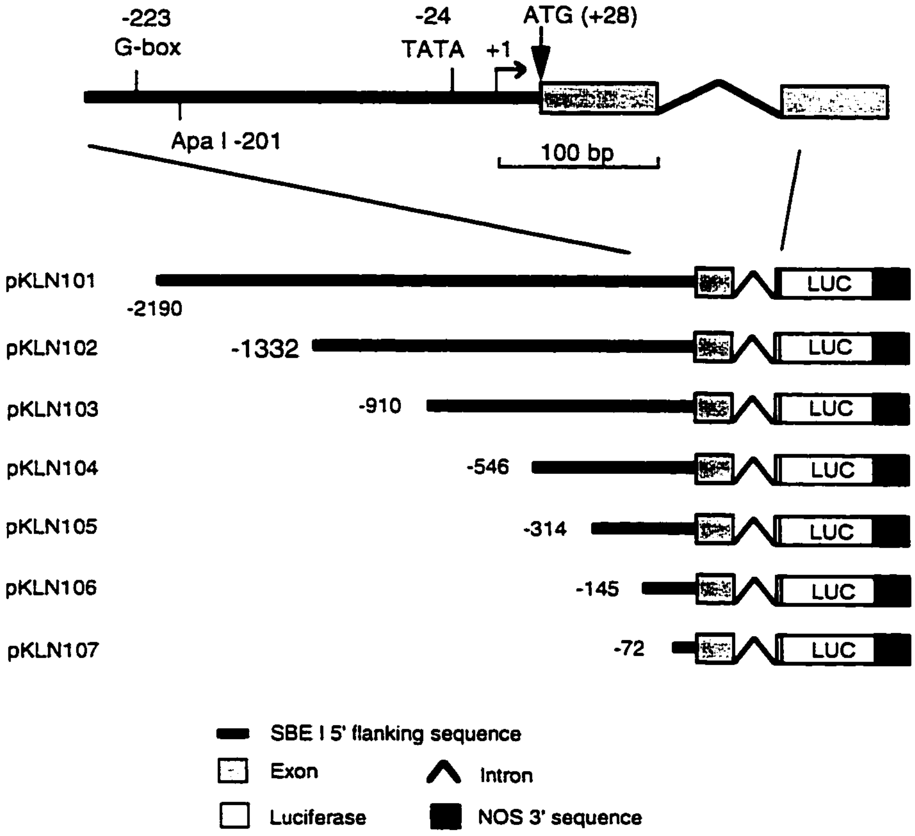


Figure 2.12. Schematic diagram of the 5' deletion chimeric constructs.

The thick black lines denote the *SbeI* promoter sequences. Numbers at left indicate deletion-end points relative to the transcription initiation site (+1) of the *SbeI* gene. Light grey (stippled) boxes and the thin black angled line represent the first exon and intron in the *SbeI* gene, respectively. Open boxes indicate the luciferase gene. The solid black boxes denote the nopaline synthase 3' end sequence.

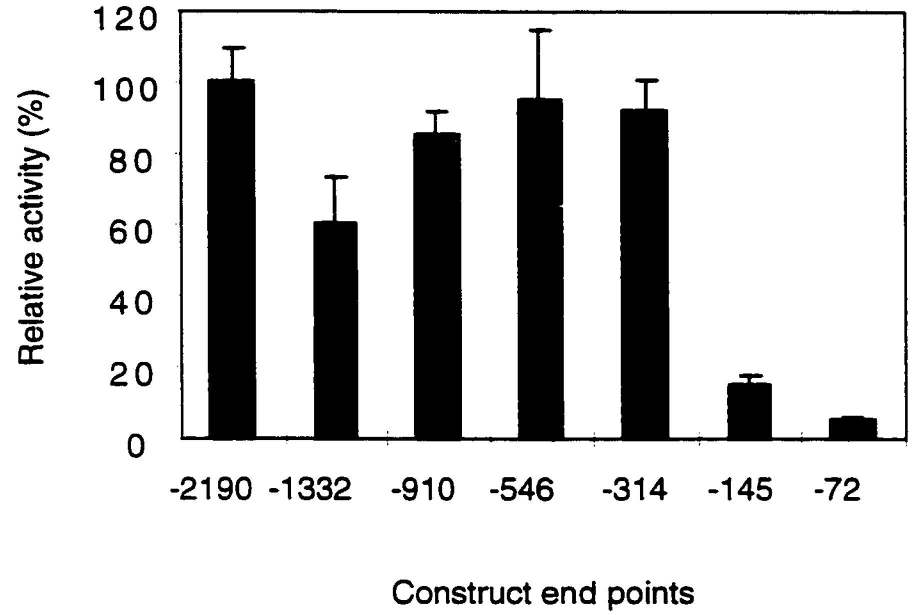


Figure 2.13. Effect of 5' deletions on *SbeI* promoter activity.

The relative activity values of the constructs described in Figure 2.12 are percentages of pKLN101 level. Each value represents the average of six to eight independent shootings. Error bars indicate standard errors of the means.

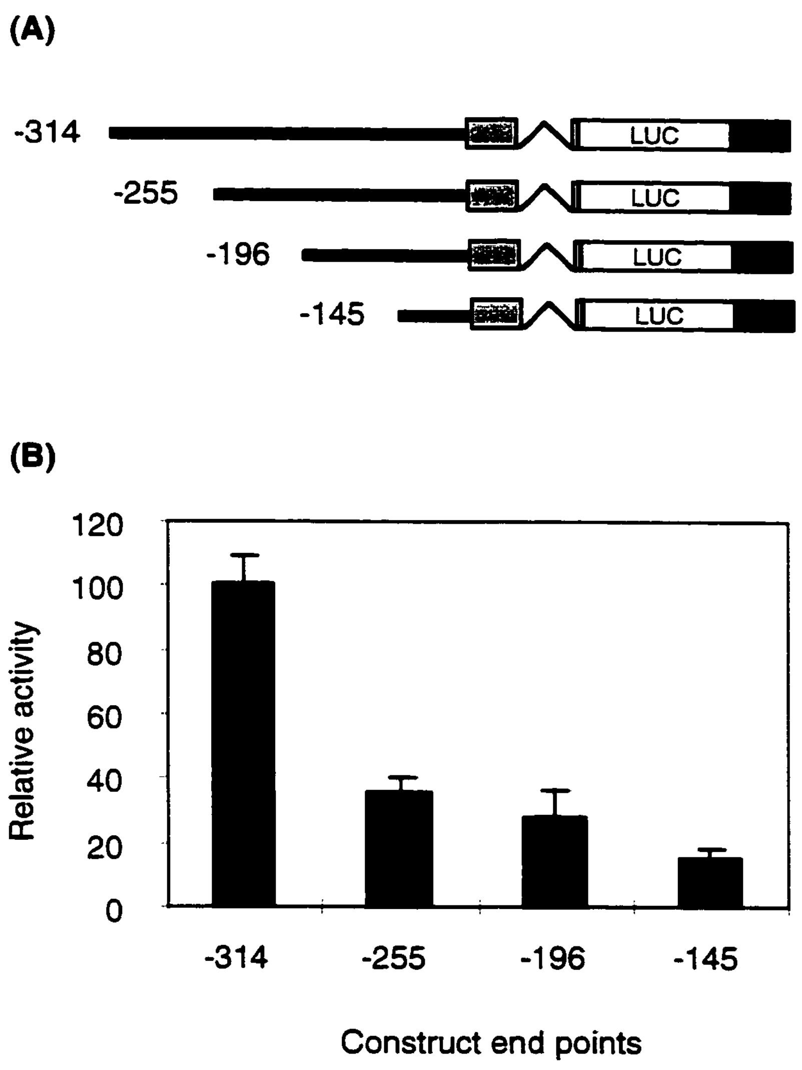


Figure 2.14. Further delimitation of *cis*-regulatory sequences in the *Sbe1* promoter.

(A) Schematic diagram of chimeric constructs. Numbers at left indicate deletion-end points relative to the transcription initiation site (+1) of the *Sbe1* gene. For an explanation of the other symbols, refer to the legend to Figure 2.12.

(B) Relative LUC activity levels of the constructs shown in (A). Relative activity values are percentages of construct pKLN105 (-314) level. Each value represents the average of four independent shootings. Error bars indicate standard errors of the means.

2.3.8. Linker-Scan Analysis

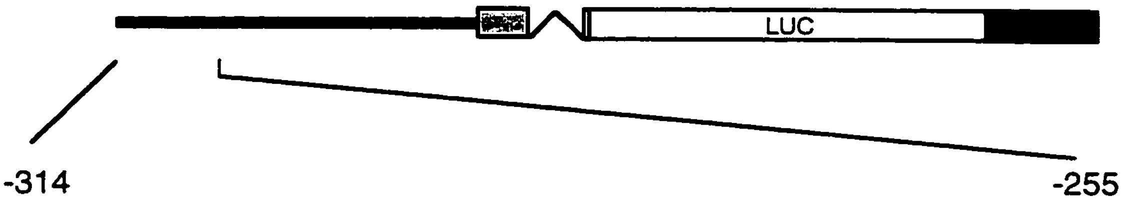
Since the 5' deletion analyses indicated that the region of the *SbeI* promoter from -314 to -255 is critical for the promoter activity, the 60-bp DNA fragment was further dissected by oligonucleotide-directed *in vitro* mutagenesis, as described by Kunkel et al. (1987). A series of six different substitution mutants, designated pLS1 to pLS6, were created by altering the wild-type DNA sequence of the *SbeI* promoter at 10-bp intervals.

The mutated constructs were tested for their promoter activity using the transient assay system, and the results of the experiments are shown in Figure 2.15. Mutations in the regions from -314 to -305 and -304 to -295, corresponding to pLS-1 and pLS-2, caused a decrease in the *SbeI* promoter activity to 60% and 72% of wild-type (pKLN105) expression, respectively. The pLS3 construct showed almost the same level of the LUC expression as the wild-type promoter, suggesting the nucleotides from -294 to -285 are not important for the promoter activity in maize endosperm cells. However, mutation in the pLS-4 region (-284 to -275) severely decreased promoter activity to 40% of the wild-type level. Also, other two mutants, pLS-5 and pLS-6, resulted in a reduction of the promoter activity to 55% and 50% of the wild-type promoter, respectively.

Figure 2.15. Linker-scan analyses of the 60-bp region in the *SbeI* promoter.

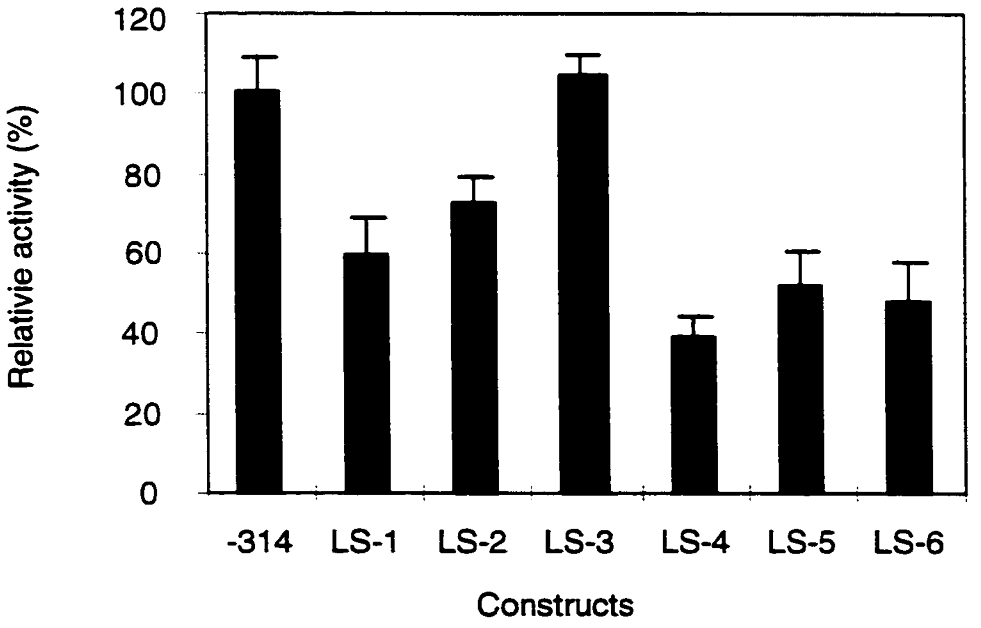
(**A**) Schematic diagram of the linker-scan constructs. The numbers indicate distance relative to the transcription initiation site (+1) of the *SbeI* gene. DNA sequence of the 60-bp region in the *SbeI* promoter is shown to the right of the wild-type construct pKLN105. The mutated bases in the linker-scan constructs are shown in lowercase letters. Dashes represent the unaltered nucleotides. For an explanation of the other symbols, refer to the legend to Figure 2.12.
(**B**) Relative LUC activity levels of the constructs shown in (A). The relative activity values are percentages of construct -314 level. Each value represents the average of four independent shootings. Error bars indicate standard errors of the means.

(A)



```
WT   ACATAAAATA AAAAAAGGCA AACCGGGAAG CGTGCAAGCC CAAAGGCCAA TCGGCCCAGA
LS-1 gtcctgc-gc ---------- ---------- ---------- ---------- ----------
LS-2 ---------- gcctgca-tc ---------- ---------- ---------- ----------
LS-3 ---------- ---------- gctatct-ga ---------- ---------- ----------
LS-4 ---------- ---------- ---------- tacctgcagt ---------- ----------
LS-5 ---------- ---------- ---------- ---------- tctgcagtcg ----------
LS-6 ---------- ---------- ---------- ---------- ---------- gatagt-gac
```

(B)



2.3.9. Proteins Interacting with the *Sbe1* Promoter

In order to detect a nuclear protein(s) interacting with the 60-bp *Sbe1* promoter fragment from -314 to -255, electrophoretic mobility shift assays were performed. The 60-bp fragment was ^{32}P end-labeled with Klenow fill-in reaction and then incubated with nuclear extract prepared from 30 DAP maize kernels (B73). The results are shown in Figure 2.16. Compared to the control (lane 1), four shifted bands were observed in lane containing the nuclear extract (lane 2). The intensities of the two upper bands (B1 and B2) were much weaker than those of the lower bands (B3 and B4). The bands were not detected after inclusion of proteinase K in the binding reaction (lane 7), indicating the shifted bands represent DNA-protein complexes.

Competition assays were conducted to determine whether or not the complexes are due to the binding of sequence-specific proteins. Including 10-fold and 100-fold excess of the unlabeled 60-bp fragment in the binding reaction significantly reduced formation of the complexes (lane 3 and 4), while the same amount of salmon sperm DNA, non-specific competitor, failed to compete for binding in the concentration range tested (lane 5 and 6). Thus, the complexes appear to be the results of sequence-specific interactions between a nuclear protein(s) and the DNA fragment.

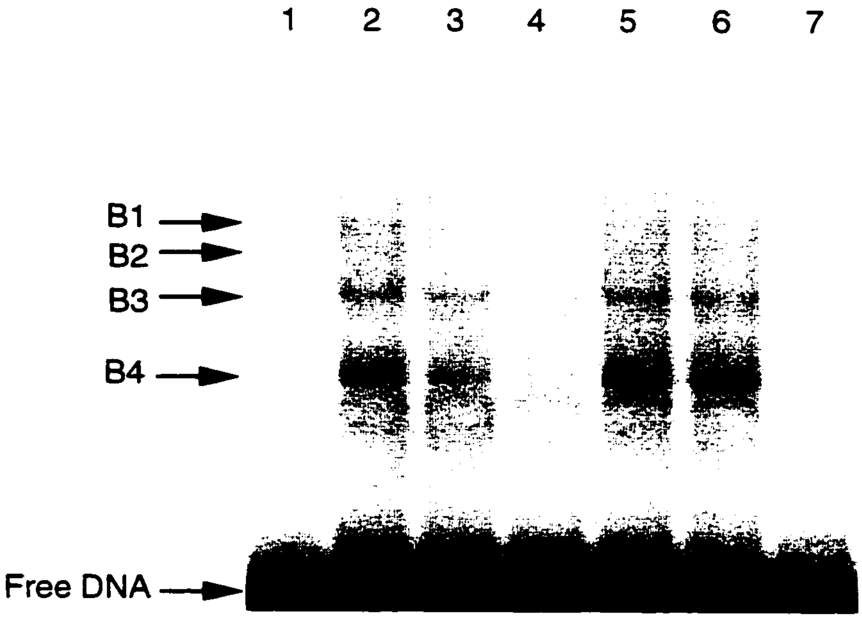


Figure 2.16. Interaction of nuclear proteins from maize kernels with the 60-bp *SbeI* promoter fragment from -315 to -255.

The 60-bp fragment was radiolabeled and1 ng of the probe was incubated with 10 μg of crude nuclear proteins prepared from maize kernels. After 20 min incubation, the samples were electrophoresed in a 4% polyacrylamide gel at 4°C for 2 hr. The gel was then dried and autoradiographed at -80°C with an intensifying screen. Lane 1, control reaction without nuclear extract; lane 2, control reaction with nuclear extract; lane 3 and 4, 10-fold and 100-fold excess of the 60-bp unlabeled fragment; lane 5 and 6, 10 ng and 100 ng of salmon sperm DNA; lane 7, 4 μl of 1 mg/ml proteinase K. Bands reduced in mobilities are indicated as B1, B2, B3 and B4.

2.3.10. Sucrose Effects on *Sbe1* Gene Expression

The SBEs are expressed in a coordinate fashion with the granule-bound starch synthase (GBSS) and ADP-glucose pyrophosphorylase during maize endosperm development (Gao et al., 1996). In addition, the ADP-glucose pyrophosphorylase gene (AGPase S) from potato and the genes encoding GBSS and SBE in cassava plants were induced by exogenous supply of sugars (Muller-Rober et al., 1990; Salehuzzaman et al., 1994). This led us to speculate that the *Sbe1* gene in maize may be also regulated by external sugar concentration.

To test this, maize endosperm suspension cells were incubated in MS media containing different concentrations of sucrose and their total endogenous RNAs were analyzed by Northern blot hybridization. Sucrose was used in preference to other metabolizable sugars, because it is known to be the major sugar unloading from the pedicel tissue of maize kernels (Porter et al., 1985). The results are shown in Figure 2.17. Increase in sucrose concentration from 0% to 9% elevated the *Sbe1* mRNA level by two-fold, and at higher concentrations the increase was reduced. Hexoses such as glucose, fructose and myo-inositol also increased the level of the transcript in a similar fashion (data not shown). However, glycerol and PEG 200 which have the same osmotic potential as 9% sucrose solution (263 mM) did not exhibit any effect, indicating the response is not an osmotic effect but a sugar-specific phenomenon. These results suggest that expression of the *Sbe1* gene in maize endosperm cells is regulated by sugar availability like other starch biosynthetic genes.

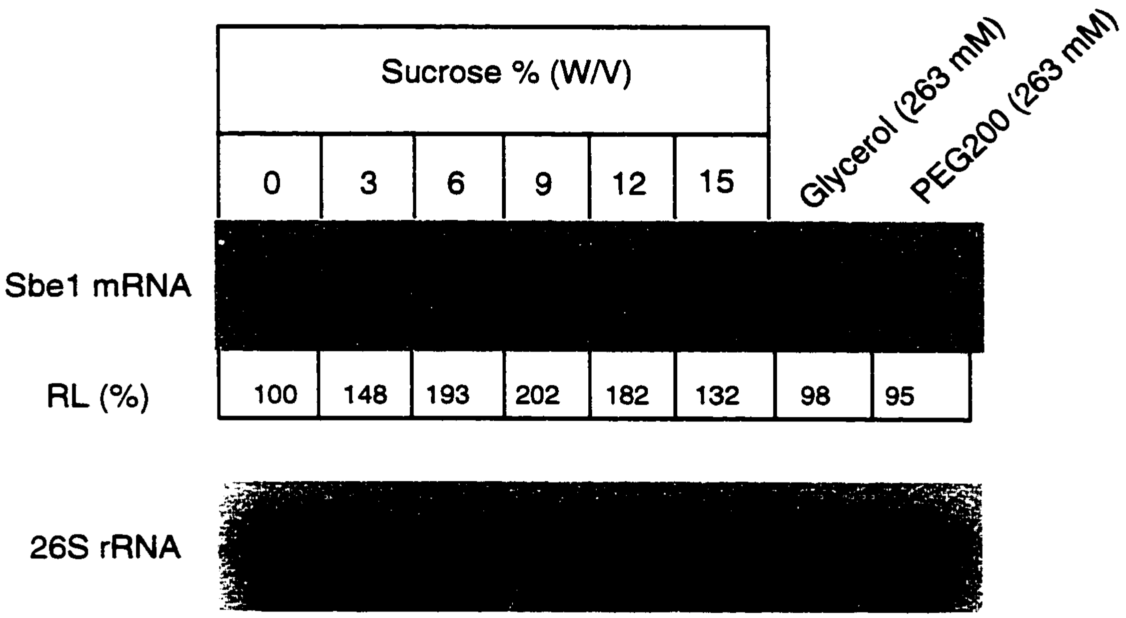


Figure 2.17. Sucrose effect on the *Sbe1* mRNA levels in maize endosperm suspension cells.

Total RNA was extracted from maize endosperm suspension cells incubated for 24 hr in MS medium supplemented with different amounts of sucrose, 0% to 15%. As osmotic controls, glycerol and PEG 200 (263 mM) were used instead of 9% sucrose (263 mM). RNA gel blots (10 μg per lane) were probed with the 32P-labeled full-length *Sbe1* cDNA, and were quantified with a PhosphorImager. The *Sbe1* mRNA levels were calibrated with 26S rRNA levels to correct for minor loading errors among the lanes. RL indicates the relative Sbe1 mRNA level. Each value is percentage of the *Sbe1* mRNA level in 0% sucrose.

In order to determine whether or not expression of the isolated *Sbe1* gene is responding to external sucrose concentrations, a gene which is not regulated by sugar concentration was necessary for an internal control for the transient assay system. Since a CaMV 35S promoter has been used as a control in other studies investigating sucrose responsiveness of plant genes, the effect of sucrose on expression of the CaMV 35S promoter-GUS chimeric gene (pBI221) in maize endosperm cells was first investigated.

The plasmid pBI221 was bombarded into maize endosperm suspension cells supplemented with 0% sucrose or 9% sucrose media and incubated at 25°C in the dark. After 48 hr incubation, GUS activity and protein concentration were measured from each sample to calculate specific GUS activity (data not shown). The results showed that specific GUS activities of 9% sucrose samples were almost 2.5-fold higher than those of 0% sucrose samples, which is consistent with other reports (Graham et al., 1994; Grierson et al., 1994). Since similar results were obtained from a ubiquitin promoter (plasmid pACH18) and -64 CaMV 35S minimal promoter which does not have an activation sequence (*as*)-1, a binding site for the transcription factor TGA-1a (Katagiri et al., 1989), it appeared that the elevated levels of expression by the CaMV 35S and ubiquitin promoters in 9% sucrose may be a general phenomenon simply due to an increase in energy source rather than a sugar-specific effect. Therefore, we reasoned that if the chimeric construct pKLN101 (the *Sbe1* promoter-LUC) is sugar-modulated, it will further enhance the level of LUC expression beyond the general increase at the higher sucrose concentration.

As shown in Figure 2.18, after normalization to GUS activity driven by the CaMV 35S promoter the plasmid pKLN101 still showed approximately two-fold greater LUC activity in 9% sucrose media than in 0% sucrose media, which is consistent with the result of the endogenous RNA analysis. This indicates that the identified *SbeI* gene is regulated by sugar availability, and the nucleotide sequence containing a 2.2 kb 5'-flanking region and the first exon/intron of the *SbeI* gene is sufficient for conferring the sugar responsiveness in maize endosperm cells.

Next, to delimit a region(s) necessary for the response, two deletion constructs, pKLN105 and pKLN106 were also tested in the transient expression system (Figure 2.18). Like pKLN101, pKLN105 (deletion end point -314) responded to a high sucrose concentration (9%) and increased LUC expression by approximately two-fold. However, pKLN106 (deletion end point -145) showed similar levels of LUC expression in both low and high sucrose conditions. These results suggests that the region between -314 and -145 contains a *cis*-regulatory element(s) necessary for the sugar response in maize endosperm cells.

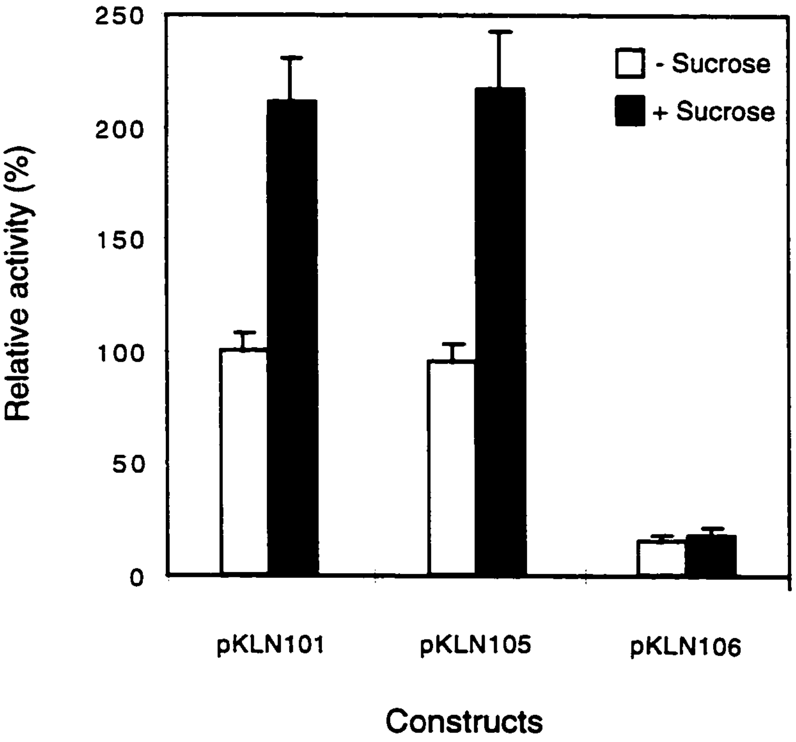


Figure 2.18. Transient expression analysis showing the sucrose effect on the *Sbe1* promoter-LUC chimeric constructs.

Each construct was bombarded onto the maize endosperm suspension cells supplemented with 0% (-) sucrose or 9% (+) sucrose media and incubated for 48 hr at 25°C in the dark.
The relative activity values are percentages of pKLN101 level in 0% sucrose. Each value represents the average of three independent shootings. Error bars indicate standard errors of the means.

2.3.11. Effect of mEmBP-1 overexpression on the *Sbe1* Gene Expression

The canonical G-box sequence, CCACGTGG (Giuliano et al. 1988), was also found in the 5'-flanking sequence of the maize *Sbe1* (-228 to -221) as well as the rice gene (-170 to -163), suggesting a possible role of the G-box motif in the regulation of gene expression. It is known that the G-box, a *cis*-acting DNA regulatory element, resides in the promoters of many plant genes responding to a variety of different environmental and physiological stimuli such as ABA, ethylene, methyl jasmonate, light, anaerobiosis and p-coumaric acid (Menken et al., 1995).

To test whether or not the G-box in the maize *Sbe1* is interacting with a G-box binding protein in maize (mEmBP-1) which is a homologue of the wheat EmBP-1 (Guiltinan et al., 1990), EMSA and DNase I footprint analyses were performed with purified mEmBP-1 protein. The analyses clearly showed that EmBP-1 interacts with the G-box sequence (data not shown). Since EmBP-1, a member of the basic leucine zipper (bZIP) transcription factors, is implicated in ABA-induced *Em* gene expression in wheat (Guiltinan et al., 1990), the data prompted me to ask two questions: First, is the *Sbe1* gene expression regulated by ABA concentration like the wheat *Em* gene? Second, can mEmBP-1 protein transactivate the *Sbe1* gene expression? The answer to the first question appears to be no. Transient expression assays failed to show a relationship between exogenous ABA concentrations (1 to 100 μM) and the *Sbe1* promoter activity in the maize endosperm suspension cells (data not shown), suggesting the G-box in the *Sbe1* promoter is not ABA-responsive.

To address the second question, a chimeric construct containing the cauliflower mosic virus (CaMV) 35S promoter fused to the full-length mEMBP-1 cDNA (35S-mEmBP-1) was created and co-introduced with the plasmid pKLN101(a full-length *SbeI* promoter-LUC) into the maize endosperm suspension cells. We predicted overexpression of mEmBP-1 protein would enhance the LUC expression driven by the *SbeI* promoter, since mEmBP-1 is known as a bZIP transcription activator. Contrary to the prediction, overexpression of mEmBP-1 protein actually resulted in a significant reduction (5-fold) of the *SbeI* promoter activity as shown in Figure 2.19. The effect was apparently selective for the *SbeI* promoter, since mEmBP-1 had little effect on expression of a LUC reporter gene linked to the ubiquitin promoter (pACH18).

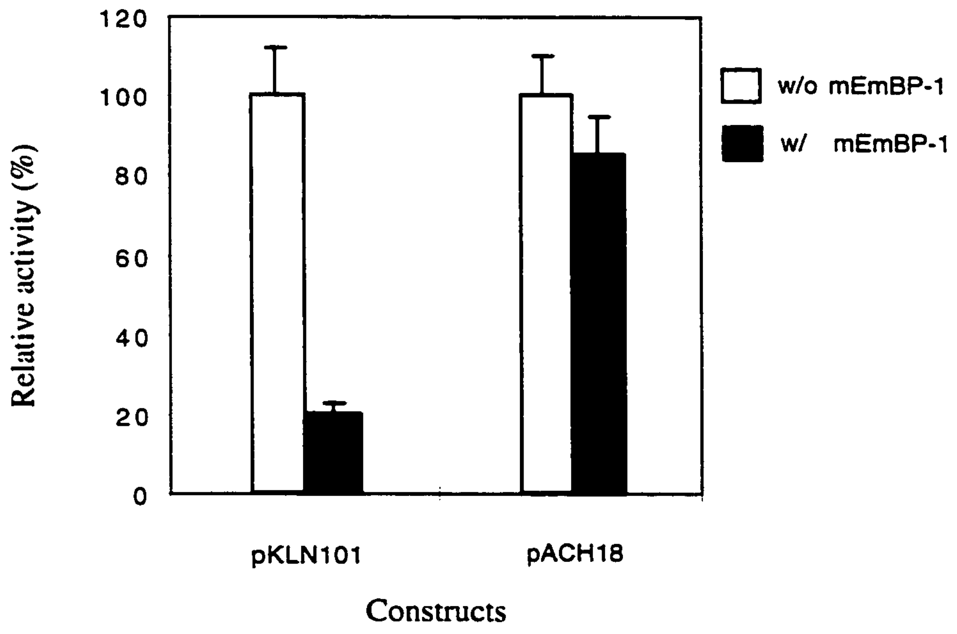


Figure 2.19 Effect of mEmBP-1 overexpression on the *SbeI* promoter.

4 µg each of reporter plasmid (*SbeI* promoter-LUC; pKLN101, or Ubiquitin-LUC; pACH18) and reference plasmid (CaMV 35S-GUS; pBI221) were co-precipitated onto gold particles with or without 4 µg of an effector plasmid (CaMV 35S-mEmBP-1). Maize endosperm suspension cells were bombarded wiht the gold particles and incubated at 25°C for 24 hr in the dark. The relative avtivity values are percentages of the pKLN101 or pACH18 levels without mEmBP-1 overexpression. Each value represents the average of two independent shootings. Error bars indicate standard errors of the means.

2.4. Discussion

The expression pattern of the maize *Sbe1* gene has been thoroughly investigated in almost all of maize tissues (Gao et al., 1996). The *Sbe1* gene was constitutively expressed at a low level in vegetative tissues, while its expression level continued to change in the kernel during its development. Especially in the endosperm, *Sbe1* mRNA began to accumulate to high levels at the onset of rapid starch deposition. These findings suggest that the expression of *Sbe1* is regulated by certain factors which vary in concentration or activity during the kernel development. Also, since the induced expression of the *Sbe1* gene appeared to be regulated in a coordinate way with that of other starch biosynthetic genes such as granule-bound starch synthase (*Wx*) and ADPG pyrophosphorylase (*Sh2* and *Bt2*) (Muller-Rober et al., 1990; Salehuzzaman et al., 1994), it is possible that these genes share some regulatory mechanisms for their induced gene expression. Therefore, knowledge of the regulatory mechanisms for one of the starch biosynthetic genes may be very helpful to understand how the other genes are controlled in plants. Unfortunately, however, no reports have been published regarding DNA sequence elements or transcription factors involved in the regulation of starch biosynthetic gene expression.

As a first step toward understanding the mechanisms regulating maize *Sbe1* gene expression, a maize genomic library was screened with a ^{32}P-labeled *Sbe1* genomic PCR product corresponding to the region from 438 to 764 of the maize *Sbe1* cDNA (Baba et al., 1991). Although 8 individual positive clones were isolated, further analyses demonstrated that they were all derived from a single locus, suggesting only one *Sbe1* gene might be present in the maize genome. However, comparison of the band patterns in the genomic Southern blot probed with a full-length *Sbe1* cDNA

(Fisher et al., 1995) with the restriction map of the *Sbe1* genomic clone (λ5-1-1) clearly indicated presence of another *Sbe1* gene which was not isolated by the screening. Genetic map data confirmed this conclusion, placing *Sbe1* genes on both chromosome 6 and 10. This discrepancy between the data can be explained by the results obtained from the Southern blots probed with two different genomic DNA fragments (Figure 2.9), which suggested the presence of two *Sbe1* genes with divergent DNA sequences around the 5'-end of the coding region as well as 5'-flanking region. Since the PCR genomic fragment located near the 5'-end of the *Sbe1* cDNA was used as a probe for the library screening, it possibly failed to pick up the other *Sbe1* gene. However, the possibility that the second *Sbe1* gene was not amplified during construction of the genomic library cannot be ruled out.

The complete genomic structure of one maize *Sbe1* gene was established by DNA sequence analysis using one of the positive full-length clones. The gene is composed of 14 exons and 13 introns, spanning about 5.7 kb in length. Primer extension analysis revealed a single transcription initiation site at a position 27 bp upstream of the translation start site (ATG).

Sequence comparison of the maize and rice *Sbe1* (Kawasaki, et al., 1993a) genomic DNAs revealed common structural features. First of all, both genes are composed of 14 exons and 13 introns. Second, the placement of all the introns are identical in both genes, although their sizes and sequences are not conserved. Third, the exons in general exhibit striking similarity in size and sequence. The major differences are confined to the terminal exons such as exon 1, 2 and 14, which are mostly not encoding mature protein. The carboxyl termini of the proteins are most divergent.

The conservation patterns exhibited by the exons and introns demonstrate that DNA sequences critical in gene function are evolutionarily conserved. Another noteworthy feature derived from the sequence comparison is the presence of two highly conserved regions in the 5'-flanking sequences, suggesting that the regions may have very important roles in the gene expression. In fact, deletion of the regions resulted in a decrease in the promoter activity in transient expression assays, indicating that they may act as distal positive regulatory elements in the *Sbe1* gene expression in maize endosperm cells.

Transient expression assays showed the level of the gene expression driven by the maize *Sbe1* promoter greatly depends on the presence of the DNA region spanning the first exon and intron of the maize *Sbe1*. Although there are many examples of plant genes which are regulated by DNA sequences within the transcribed region (Callis et al., 1987; Bruce et al., 1989; McRlroy et al., 1990; Fu et al., 1995), the first exon and intron sequences of the maize *Sh1* gene are one of the best examples studied so far (Vasil et al., 1985; Maas et al., 1991; Clancy et al., 1994). The *Sh1* exon appears to have two separate *cis*-elements which act independently to increase gene expression via different mechanisms. One of the elements may contain a novel promoter element which has the ability to interact with transcription factors binding upstream. The other acts possibly at the level of translation efficiency or mRNA stability. The enhancing effect of the *Sh1* intron is likely the result of an increase in the level of mature cytoplasmatic mRNA level like the maize *Adh1* first intron (Callis et al., 1987).

Overexpression of mEmBP-1 protein, a family of bZIP transcription activators, actually decreased *Sbe1* promoter activity (Figure 2.19). One of the possible explanations for the unexpected result can be found in the mechanism by which the human activating transcription factor 3 (ATF3) functions. ATF3, a member of the mammalian bZIP transcription factors, binds to a consensus DNA sequence (TGACGTCA) which is important for mediating transcriptional responses to a variety of signals (Chen et al., 1994). It has been demonstrated that ATF3 represses transcription when bound to DNA as homodimers, yet it can also form selective heterodimers with other transcription factors and activates transcription depending on the cellular conditions. In fact, many of G-box binding proteins in plants can bind to G-box sequences as homodimers and heterodimers *in vitro* (Armstrong et al., 1992; Schindler et al., 1992). Therefore, it is not hard to imagine that mEmBP-1, just like ATF3, may be able to form heterodimers with new properties distinct from those of the homodimers. If this is the case, we can assume that the overexpression of mEmBP-1 by the CaMV 35S promoter in maize cells forced the protein to form homodimers, thereby repressing *Sbe1* promoter activity. In conclusion, although a role for the G-box motif is not yet identified in the *Sbe1* gene expression, the above results provide a possibility that the G-box in the *Sbe1* promoter may play a critical role in the gene expression under different environmental conditions or in different tissues.

5' deletion analysis of the maize *Sbe1* promoter revealed several *cis*-regulatory elements affecting the promoter activity in maize endosperm cells. Of special interest was the identification of the 60-bp strong positive element located in the region from -314 to -255 relative to the transcription initiation site. Further investigation of the region using linker-scan analysis identified at least two separate regions, -314 to -295 and -284 to -255, which are critical for gene expression in maize endosperm cells.

Interestingly, as shown in Figure 2.20, the -314/-295 region has striking similarity with the sugar-response elements of the potato patatin-1 (Grierson et al., 1994) and sporamin promoters (Ohta et al., 1991), although it shows a little less similarity with the sugar-inducible regions of the potato sucrose synthase (Fu et al., 1995b) and proteinase inhibitor 2 (Kim et al., 1991). This finding along with the sugar enhanced expression shown in Figure 2.18 strongly suggest that the conserved sequences may be implicated in mediating sugar responsiveness of the *Sbe1* gene. Since high sucrose concentration media were used for the transient expression assays to maximize gene expression, it is understandable that mutation of this region decreased the level of LUC expression.

In potato and cassava plants, sugars have been shown to regulate expression of genes involved in starch biosynthesis (Muller-Rober et al., 1990; Salehuzzaman et al., 1994). Our results demonstrate that the maize *Sbe1* is also modulated by sugar concentration (Figure 2.17 and Figure 2.18). Such sugar effect was not due to change in osmotic potential, because glycerol and PEG which are osmotically active did not affect *Sbe1* gene expression. However, a surprising result was obtained when mannitol, a non-metabolizable substance, was used as an osmoticum. That is, mannitol did affect the *Sbe1* gene expression in maize endosperm cells to the same extent as metabolizable sugars such as sucrose, glucose and fructose (data not shown). A similar effect was observed in the maize ADPG pyrophosphorylase genes, *Sh2* and *Bt2* (Giroux et al., 1994). Since starch biosynthetic genes in other plants such as potato and cassava plants were not influenced by mannitol, the mannitol effect seems to be unique to the maize starch biosynthetic genes.

```
Sbe1     -314    ACATAAAATAAAAAAAGGCAA    -294
                 : : : |||| ||||||: ||
PS20     -177    ATAAAGAATAGAAAAAGGAAA    -151
                 ::    |||| ||||||  ||
SPO      -108    ATTCTTAATACAAAAAG--AA    -90
                  :: ::|||| ||||  |  ||
Sus4    -1303    TTTTTTAATA-AAAATG--AA    -1286
                   :  : ||  |||||   ||
PI-II    -546    GATAATTATTTAAAAAAACAA    -526
```

Figure 2.20. Sequence comparison of *cis*-regulatory regions found in sugar-modulated genes.

The maize *Sbe1* promoter sequence from -314 to -294 was aligned with the 5' sequences of the potato class I patatin (PS20), sporamin-A1 (SPO), sucrose synthase (Sus4) and proteinase inhibitor (PI-II) which are modulated by sugar concentration. The numbers indicate position of nucleotides from the relevant transcription initiation sites. The vertical bars and the dots represent conserved nucleotides among more than and less than four genes, respectively. Dashes indicate gaps to maximize alignment.

A possible explanation for this discrepancy is that unlike hexokinases in other plants, a maize endosperm hexokinase might be capable of phosphorylating mannitol, thereby triggering sugar-modulated gene expression. Alternatively, mannitol in maize endosperm cells may be first oxidized to mannose which is then phosphorylated by hexokinase. Recently, Jang et al. (1997) provided evidence that hexokinase is involved in sensing sugar concentration in higher plants, and sugar signaling mediated through hexokinase is uncoupled from sugar metabolism.

The other element present in the region from -284 to -255 includes a CCAAT sequence. The CCAAT motif is often found in mammalian promoters and is known to be crucial for basal levels of transcription (McKnight and Tjian, 1986). A multiplicity of proteins binding to the element was also identified (Dorn et al., 1987; Chodosh et al., 1988). In plants, however, the CCAAT sequence is not a common promoter component, suggesting it may have obtained new functions during plant evolution (Bucher, 1990). In fact, the CCAAT motif, present in the proximal promoter region (-132 to -128) of a *Lemna gibba Lhcb* gene encoding a light-harvesting chlorophyll a/b protein, was reported to be essential for phytochrome responsiveness (Kehoe et al., 1994). Interestingly, most of the short conserved segments between the rice and maize *Sbe1* proximal promoter sequences shown in Figure 2.6(B) are localized within the -284/-255 region, implying this region may also be involved in expression of the rice *Sbe1*.

Chapter 3

MOLECULAR ANALYSIS OF A GENOMIC FRAGMENT CONTAINING MAIZE (*ZEA MAYS* L.) STARCH BRANCHING ENZYME GENE IIB

3.1. Introduction

Purification and *in vitro* enzyme assays of the three isoform of starch branching enzymes (SBEI, IIa and IIb) in maize endosperm showed that SBEI has lower Km for amylose and preferentially transfers longer glucan chains than SBEIIa and IIb (Boyer and Preiss, 1978; Guan and Preiss, 1993; Takeda et al., 1993). The different catalytic properties of SBE isoforms led Guan and Preiss (1993) to suggest that SBEIIa and IIb use the slightly branched polysaccharides produced by SBEI as substrates for further branch formation. SBEI was also found to be distinct from SBE IIa and IIb in immunological properties, molecular weight, amino acid composition, and proteolytic digestive maps. While SBEIIa and IIb did not exhibit significant differences in such immunological and biochemical properties (Boyer and Preiss, 1978; Fisher and Boyer, 1983; Singh and Preiss, 1985), the report by Takeda et al. (1993) suggested that they are distinguishable isoforms of SBE. In the branching linkage assay, SBEIIa had more than 2-fold greater specific activity than SBEIIb, and the optimum temperature of SBEIIa and IIb are quite different, 25°C and 15-20°C, respectively.

Among the SBE isoforms in maize endosperm, a partial *Sbe1* cDNA was first cloned by screening kernel library with a rabbit antibody raised against SBEI protein (Baba et al., 1991) . Later, a full-length cDNA clone encoding SBEI was isolated by Fisher et al. (1995) using a PCR-amplified SBEI genomic fragment as a probe. Cloning of a gene encoding SBEIIb was reported by two separate laboratories which utilized different strategies. Fisher et al. (1993) identified a full-length *Sbe2b* cDNA by probing a maize endosperm library with the pea *rugosus* (*r*) gene which is one of the characters used by Mendel in his study on the laws of inheritance. The *r* gene determining the wrinkled versus round phenotype of mature seeds is known to be a structural gene for SBEI (Bhattacharyya et al., 1990). Stinard et al. (1993) cloned a dominant mutant allele of the amylose-extender (*ae*) locus (*Ae*-5180) through *Mu1* transposon tagging. Sequence comparison of the *ae* clone with *Sbe2b* cDNA clearly showed that they are the same gene, confirming that the SBEIIb is encoded by the *Ae* gene.

Cloning of the genes encoding SBE isoforms, SBEI, IIa and IIb, in maize endosperm has made it possible to investigate the genes at the molecular level. Recent studies showed that these genes are differentially expressed during kernel development and in various tissues, suggesting that they play distinct roles in starch biosynthesis (Gao et al., 1996 and 1997). *Sbe1* and *2a* are expressed in most maize tissues or organs including endosperm, embryos, leaf, stem, root and tassel, while *Sbe2b* is expressed only in endosperm and embryos as well as tassel. Unlike *Sbe1* and *2b*, *Sbe2a* is more highly expressed in embryos than in endosperm. Another noteworthy feature is that *Sbe1* and *Sbe2b* genes are differentially expressed during kernel development. That is, the timing of the highest transcript level of *Sbe2b* comes prior to that of *Sbe1* in developing endosperm and embryos, which may result in changes in the SBEI/SBEIIb ratio. Since SBEI and IIb have significantly different *in vitro* catalytic properties as mentioned above, such changes in the SBEI/SBEIIb ratio may cause differences in the starch synthesized during kernel development. During pea embryo development, in fact, changes in the SBE isoform ratio was accompanied by transition in branch lengths of amylopectin (Burton et al., 1995).

Therefore, it is important to understand the regulatory mechanisms involved in expression of SBE isoforms during kernel development for elucidating their relative roles in starch biosynthesis. In chapter 2, I described the isolation and characterization of a full-length *Sbe1* genomic clone containing the entire coding region of *Sbe1* as well as 5'-and 3'-flanking sequences. In that study, it was shown that the maize genome contains at least two *Sbe1* genes, and the complete structure of one of the *Sbe1* genes as well as the promoter regions critical for its expression in maize endosperm cells are determined.

To gain insight into *Sbe2b* gene regulation, I now describe the isolation and characterization of a maize genomic DNA containing this gene. As in the case of the maize *Sbe1* gene, the complete genomic organization of the maize *Sbe2b* gene is first established including the transcription initiation site, number of introns and the polyadenylation site. Then, I demonstrate that a single *Sbe2b* gene is present in the maize genome using restriction enzyme maps of the gene and genomic Southern blot analyses. A series of 5' deletion analyses reveal possible *cis*-regulatory elements in the promoter which control expression of this gene in maize endosperm cells.

3.2. Materials and Methods

3.2.1. Maize Genomic Library Screening and DNA Sequencing

Using ^{32}P-labeled full-length *Sbe2b* cDNA (Fisher et al., 1993) as probe, the same library used in chapter 2 was screened according to the methods used in chapter 2. Phage DNAs were digested with SalI to release the inserts from the EMBL-3 vector.

3.2.2. Primer Extension Analysis

An oligonucleotide, 5'-GATCGGATCGAACTGATCAG-3', which is complementary to the sense strand sequence of the *Sbe2b* cDNA from -35 to -16 relative to the translation start site (ATG) was designed and used to locate the transcription initiation site as described in chapter 2.

3.2.3. Genomic Southern Blot Analysis

A full-length *Sbe2b* cDNA (Fisher et al., 1993) was used as a probe to determine the number of *Sbe2b* genes in the maize genome. For details, refer to the Methods in chapter 2.

3.2.4. Construction of Plasmids

For a transcriptional fusion of the *Sbe2b* promoter to a luciferase (LUC) reporter gene, a BamHI restriction enzyme site was created just before the translation initiation site (ATG) of the *Sbe2b* gene. As shown in Figure 3.1, the DNA sequence between -14 and +100 of the *Sbe2b* gene was first amplified via polymerase chain reaction. *Pfu* DNA polymerase (Stratagene, La Jolla, CA), which has proofreading activity, was used to enhance the fidelity of PCR amplification. (*Pfu* DNA polymerase was used for all the following PCRs). The 5' primer (PII-2), 5'-CCTAATTGTAGCCCTGCAGTCA-3', is homologous to sequence of the *Sbe2b* gene from -10 to +12. A PstI restriction enzyme site (CTGCAG) is located immediately downstream of the 5' primer binding region of the *Sbe2b* promoter, +4 to +9. The 3' primer (PII-3), 5'-GACTGGATCCTCGCCTTCGCAGCCGGATCG-3', consists of a DNA sequence complementary to that of the *Sbe2b* gene from +80 to +100 and a BamHI restriction enzyme site (GGATCC) flanked with four random nucleotides (underlined). The PCR product was digested with PstI and BamHI, and the resulting 100-bp fragment was ligated to the 2,977-bp SalI-PstI *Sbe2b* promoter fragment and cloned into plasmid pLN cut with SalI and BamHI (promoterless LUC-NOS gene in pUC119) (Montgomery et al., 1993), thereby creating plasmid pKL201. This construct as well as all the following constructs were verified by DNA sequencing.

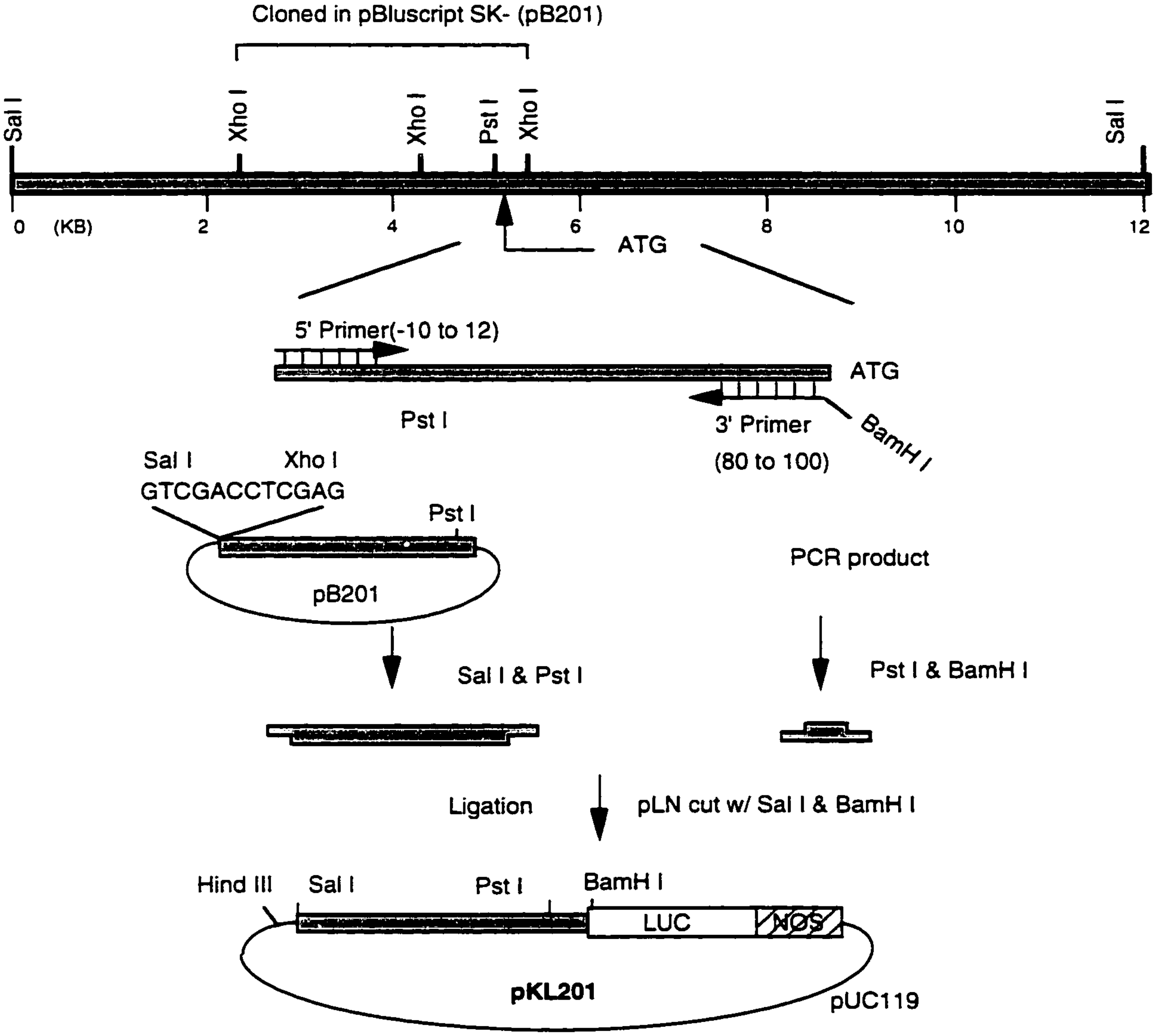


Figure 3.1. A schematic diagram of constructing a *Sbe2b*-LUC chimeric gene.

To construct a translational fusion of the *Sbe2b* promoter containing the first exon and intron to a LUC reporter plasmid, the *Sbe2b* genomic clone, 3-2-1, was digested with XhoI and the resulting 866-bp fragment was gel purified. The fragment was blunt ended by Klenow fill-in DNA synthesis and ligated with BamHI linkers (CGGGATCCCG). After complete digestion with PstI and BamHI, the 325-bp DNA fragment was isolated and used to replace the 100-bp PstI-BamHI region in pKL201. This construct was designated pKLN201.

A series of 5' deletion mutants were derived from pKL201 using available restriction enzyme sites and PCR techniques. To create pKL202, pKL203, pKL205 and pKL206, pKL201 was first digested with AccI (-1719 to -1714), SpeI (-1128 to -1123), XhoI (-537 to -532) and ApaI (-348 to -343), respectively. Then, each linearlized plasmid was separately gel purified and blunt ended by Klenow fragment. After SalI linker ligation, the modified plasmids were digested with SalI, and the larger DNA fragments from each reaction were isolated and self-ligated to produce the relevant plasmids carrying different deletion end points.

For a construction of pKL204, the *Sbe2b* promoter region between -755 and -500 was amplified with two primers. The 5' primer containing a SalI (GTCGAC) flanked with four extra nucleotides (underlined), 5'-GAAAGTCGACGAAGAGAGAATGAAAGCGAA-3', and the 3' primer, 5'-GCGCGGGTCCGTCGTGCCTTTT-3', were designed to anneal to DNA sequences of the *Sbe2b* gene from -755 to -736 and from -521 to 500, respectively. The amplified 266-bp product was digested with SalI and XhoI which was located 10 bp upstream of the 3' primer binding region, and the resulting 230-bp fragment was used to substitute the 2.4-kb SalI-XhoI fragment in pKL201.

To make pKL207, the *Sbe2b* promoter region between -160 and +100 was first amplified with 5' primer, 5'-GATAGTCGACCGACGCGCAACGGCCTGCCT-3', and 3' primer, 5'-GACTGGATCCTCGCCTTCGCAGCCGGATCG-3', which contain SalI and BamHI, respectively, along with four arbitrary extra bases (underlined). Next, the PCR product was gel purified and digested with SalI and BamHI. The resulting 267-bp DNA fragment was then used to replace the 3.1-kb SalI-BamHI fragment in pKL201. The same 3' primer and method were used for a construction of pKL208 except for the 5' primer, 5'-GATCGTCGACCGCTCGTCTCCGTCCTATAT-3', homologous to the DNA sequence of the *Sbe2b* gene from -49 to -32.

3.2.5. Transient Expression Assays

Particle Bombardment and Measurement of GUS and LUC activities were performed according to the method described in chapter 2.

3.3. Results

3.3.1. Cloning and Characterization of the Maize *Sbe2b* Gene

After screening approximately 3 x 10^5 plaque-forming units from a genomic library prepared from maize seedlings (inbred B73), 12 lambda clones that strongly hybridized to the full-length *Sbe2b* cDNA probe were isolated. These clones were further characterized using the N-terminal or C-terminal cDNA as probe; none of them contained both ends of the *Sbe2b* cDNA. Since they were initially classified into two groups based on restriction endonuclease maps, two clones from each group, λ3-2-1 and λ7-2-1, were selected, subcloned, and sequenced. DNA sequences of the two clones revealed that the λ3-2-1 clone containing the 5'-end of the *Sbe2b* gene had approximately 1.5-kb overlapping sequences with the λ7-2-1 clone containing the 3'-end of the gene, suggesting they originated from the same locus. As shown in Figure 3.2, a complete restriction map of the *Sbe2b* gene was constructed by combining the two overlapping genomic clones which turned out to contain the entire coding region of the gene as well as the 5'- and 3'- flanking sequences.

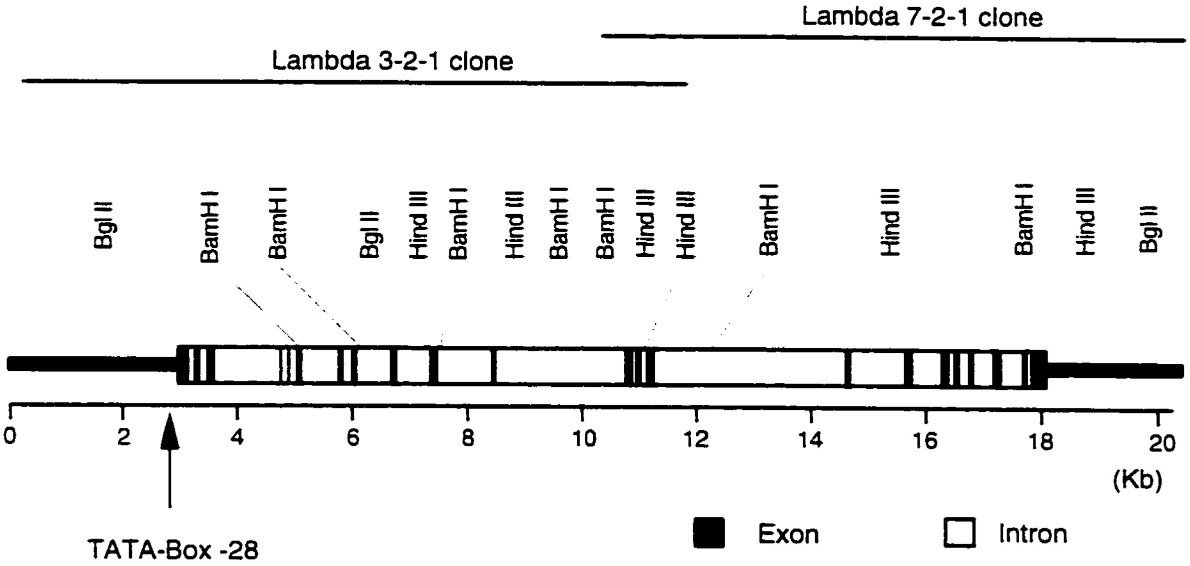


Figure 3.2. Genomic structure of the *Sbe2b* gene.

The complete structure of the *Sbe2b* gene was constructed using two overlapping genomic clones: lambda 3-2-1 and 7-2-1 clones. The thick black lines indicate the 5'- or 3'-flanking sequences of the *Sbe1* gene. The solid black boxes indicate exons and the open boxes denote introns. The number represents position of a putative TATA-box relative to the transcription initiation site (+1).

3.3.2. Genomic Organization of the Maize *Sbe2b* Gene

Primer extension analysis was conducted to determine the transcription initiation site of the *Sbe2b* gene. As shown in Figure 3.3, a major reverse transcription product was observed which co-migrated with a G residue in the sequencing ladder, indicating the transcription initiates mainly at a position which is located 28 bp downstream from a putative TATA box. The transcription initiation site was numbered +1 in the sequence shown in Figure 3.4.

The genomic structure of the *Sbe2b* gene was established by aligning the sequences of the two overlapping clones with the published sequence of *Sbe2b* cDNA (Fisher et al., 1993). The transcribed region of the gene consists of 22 exons and 21 introns distributed over 15,347 bp in length. Figure 3.2 summarizes the organization of the maize *Sbe2b* gene. The cDNA sequence is identical to the corresponding genomic sequence except for three nucleotides present in exon 4 and exon 6. This may be due to the different genetic stocks used in the two studies. Table 3.1 shows the sequences around the exon/intron junctions and a list of putative branch point consensus sequences, which was derived as described by Brown (1986). As in the maize *Sbe1* gene, the introns are relatively AT-rich (67%) compared to the exons (54%), and all of which have the conserved sequences at their 5' and 3' ends, following the "GT...AG" rule of plant introns (Brown, 1986). The introns vary in length from 76 bp (intron 4) to 3,051 bp (intron 14), and the exons vary in length from 43 bp to 303 bp. Exon 1 contains 100 bp of 5' untranslated DNA sequence, and exon 22 contains the translation stop codon (TGA) and 3' untranslated region. Although the canonical polyadenylation signal, AATAAA, was not found in the 3'-end of the gene, a similar sequence (AATTAAA) was observed 29 bp upstream of the polyadenylation site.

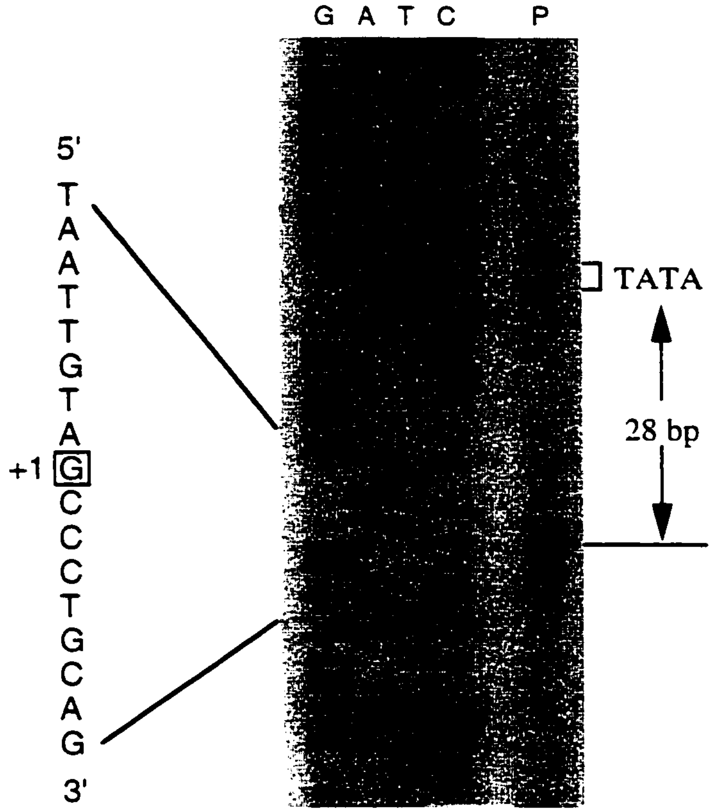


Figure 3.3. Primer extension analysis of the transcription initiation site of the *Sbe2b* gene.

An ^{32}P-end labeled antisense primer was annealed to 10 μg of total RNA isolated from maize kernels (30 DAP) and extended using reverse transcriptase. The extended cDNA product was then analyzed on a 5% sequencing gel (lane P) along with a *Sbe2b* 5'-flanking genomic DNA sequencing ladder (G, A, T, C) generated by the same primer. The number inside the vertical double-headed arrow on the right refers to distance from TATA-box to the primer extension product indicated by the horizontal arrow. The sense sequence around the product is shown on the left and the transcription initiation site is boxed and assigned +1.

Figure 3.4. Nucleotide sequence of the *Sbe2b* gene and 5'- and 3'-flanking regions.

The flanking regions and introns are shown in lowercase letters, while exons are presented in uppercase letters. The deduced amino acid sequences are shown below the string of exon sequences. The numbers indicate the distance relative to the transciption start site (+1) which is indicated by the arrowhead. The consensus sequences of a putative TATA-box and putative polyadenylation signal are underlined. The asterisk and dot indicate the stop codon and possible polyadeylation site, respectively.

```
ctcgagaattatttttgataattcatcggaccgtcattgttcatggtcagtggcacacaagtccggaccg -2895
cctccaaaggatagtctttggcgtgcgcgggtgagccaactgttagctgacgtggcacccaatccagagg -2825
tgcactagactaatccggtgcccgtcagcgaaggaaagtgtcaaccagaatcctggactcatgacgcatg -2755
aaccgatccagtaccccgcaagtaagctgatatttgacatcttttgcaagaaaggagcaatgactatg -2685
gggtcttttggagctataaaagcacccccatggcgacctccttcagaacccaagcactccaagagtgtca -2615
catcactccgactctcttctgtaactcattctagtgatttagtgagatcaatgcgtcatttctgagctgt -2545
tcttgtgtatgtgtgactttgcgcttttgtgatcttctctcttgagttttgctgctacgttttggtcttg -2475
tatttgtaatctctcccctctcttgctttgagttgtaactctgatattgtgtacgactgcaagagactct -2405
aaattatggagattccttgcaaacagggatataagtgataaggaaggacgtggcactcaatttcgatctt -2335
tggatcacttgagagcggctgattgcaatccttgtcgttggcaaccacaacgtgaagtagtcaagcattt -2265
taggcttgaccgaaccacgagaaaaatcatcgtgtcttgtgtgcttatccacttgtgatttttctccctc -2195
tttctaagtctctcagattcacttgtaatattactcttaatatattatatatcttgaaagagcaaacaag -2125
ttgaagatgacttactttcttcttctctctatttaaacttggcttggtttcactaacttccatttccaga -2055
ccaaacttgtgttttagtgttgtttttgcagggtcacctattcacccccctctaggtgctctcaagaggt -1985
gttagctatgtttccttaggcgtgtggcggtcaccttttcgatagatatgaataagttcgaactttcttt -1915
agaaacaataagtcgcaaattcaacaaagtactaaatgcgctatacaacatgtctagtgacgtaattaag -1845
ctaaaggacgaccaatttcgtgagattcaccctagactattggaggtgcgtttttggcaacattgtaggg -1775
attgtatagaactatcgatggaagccactttcctgtaatcgtacctaggcggcatgtatacgcgtcgcaa -1705
aatgtcatggcagtatgcgacttcgatatgtggtttacctttgccgtgactggttggttgggctcagttc -1635
atgatctgtaggtactacaagagactctcattgccacgacaacatctttccccatccacctgaaggtaat -1565
gttttgggcgaacaaaaataattgtcacaaatgttataaatttgaatagcattgctaatgaacttgtttt -1495
ggtaatgcatgaatatattacttggtcgattccggtcagcgtaacaggttgggctactttgccccgtata -1425
agggacaaaagtatcacgttacagaacgacaacatgggagacatcccgtagatgagaaaggggtattcaa -1355
ctatgcacattcatccctaagaaatgtcatagagcggtcatttgatgtgtttgaaggtgaagtggcagat -1285
cttgcttagtctcccatgtttctcggtgcggaagcaataaaaaataattataacttgtatgaccctacat -1215
aactttattagggatagtgctatacacaaccacgacttggaaattttgtgtctgggaagcatgtaactga -1145
tgtaggggtaacagaaactagtagtggaccattaggtgagttagacacgagtgagtttccagatgacatt -1075
tatgctgcgctagtgacatgacctaccatttgtatttcatagatgtggcatgaacatgtttatgttgtaa -1005
tgttctacgtcagtgcattatgtaattgtgtaattaatttaaggttgccctagtagcaatgaaatacaaa -935
tctccatcgggatcggagggaattaaggtgaaaatgaacttattttctctccaacccсttcaatctcgaa -865
ggggatttgagtttccaaactagatcctaaaagcaagtataataaggtgatgtaggcgagctgtaggata -795
taacacatcagatttgtgatgatatgaaagaaaaaaaatgaagagagaatgaaagcgaactgttgctcac -725
ggactctaagatataatataagaggaagagatgagctaagtattaaatgtataaatatattttttagttag -655
ttatatatgataggacaatatcataaaactcactttgtgccatatcgttaaacttgctgtagcttcaccg -585
ttaatcgataaaaaatattaaaaaagcatctcccgtttgctcgactgctcgagtgtgcaaaaaaaaaggc -515
acgacggacccgcgcgctgacgcggtgagccgcaagtccgcaacggcgcggccgcgcgctaggaaaaaag -445
cctcgcccgtgaagcgaactccctctccttcgaccttcgttcttccactgcggcctgcgcgacccgtgca -375
gctgcgcgctccacctggccgcgctggggcccacaccgcctggcatctggagcattgcccccggacttcg -305
cgcggccgcccgcagccccgctccccaccgaaaagcgaagcgattgccatccccacgccaccgcgaagca -235
caaggtcccgccctgcacgatcagcaggacctcgccacgccgccgctggagctgcgcgtgcgcgtgtgcg -165
cttggaccgacgcgcaacggcctgcctcgaccgcccgtgcacgccactgctcatgcagccgtccgcctcg -95
cccccgccccgaactgccgaggtcgcgtgaacgcccactcccctcaccgctcgtctccgtgctatatagg -25
                         ▼                                        TATA-box
cagcccgcgcccctcctaattgta GCCCTGCAGTCACCCAGAGCAGACCCGGATTTCGCTCTTGCGGTCG   46
CTGGGGTTTTAGCATTGGCTGATCAGTTCGATCCGATCCGGCTGCGAAGGCGAGATGGCGTTCCGGGTTT 116
                                                        M  A  F  R  V
```

```
CTGGGGCGGTGCTCGGTGGGGCCGTAAGGGCTCCCCGACTCACCGGCGGCGGGGAGGGTAGTCTAGTCTT 186
S  G  A  V  L  G  G  A  V  R  A  P  R  L  T  G  G  G  E  G  S  L  V  F
CCGGCACACCGGCCTCTTCTTAACTC gtaatgatcctgcaactcctcctccctctctgatcaagtgtggg 256
  R  H  T  G  L  F  L  T
cctgattcgggtctgtatgcgagtgttgtggtggtgaactggtgaattggtgatgcatgcag GGGGTGCT 326
                                                                R  G  A
CGAGTTGGATGTTCGGGGACGCACGGGGCCATGCGCGCGGCGGCCGCGGCCAGGAAAGCGGTCATGGTTC 396
 R  V  G  C  S  G  T  H  G  A  M  R  A  A  A  A  A  R  K  A  V  M  V
CTGAGGGCGAGAATGATGGCCTCGCATCAAGGGCTGACTCGGCTCAATTCCAGTCGGATGAACTGGAG gt 466
P  E  G  E  N  D  G  L  A  S  R  A  D  S  A  Q  F  Q  S  D  E  L  E
tcgtcatccactcgtcactttcatgcattttatcacataattcacctgaaagtctacatctacatgcatt 536
tctgatttttacctctttttggatgctatttgagaacaatgagacacacgattagtgagatgcccaaacg 606
gttaaacgttccatggagctccacaagtctgtactgtaccaccatttagagttgtccatacaatcgtctt 676
ctgaacatttcgtctcttcggcacggttccag GTACCAGACATTTCTGAAGAGACAACGTGCGGTGCTGG 746
                                  V  P  D  I  S  E  E  T  T  C  G  A  G
TGTGGCTGATGCTCAAGCCTTGAACAGAGTTCGAGTGGTCCCCCCACCAAGCGATGGACAAAAAATATTC 816
  V  A  D  A  Q  A  L  N  R  V  R  V  V  P  P  P  S  D  G  Q  K  I  F
CAGATTGACCCCATGTTGCAAGGCTATAAGTACCATCTTGAGTATCG gtatgtattacttgcttctactg 886
 Q  I  D  P  M  L  Q  G  Y  K  Y  H  L  E  Y  R
caccactactgaaacatttcaagacgttcccgtctgtcatgatatatataagggtgtgtttggttgggga 956
gtgagggagaataaagtggttccattctaatttctggacgtttagtttccaacaaaataagagtggagc 1026
ggctcctggatctccatatggaaatttactataagtagttggaacgctcccgctccacaaaaacgacgga 1096
tgcgagcgctctctcggacgtgagcgctctcctggctctcctcccctctacactcatacggctctcaaac 1166
caaacaaagaatggaatgactccgctctattctcgtgacaagattctcgtgttatcttaattcttgcata 1236
cttttaggtgcgtttggttggggagccactgggatggagcggctccattccagtttctaaaaattgagcc 1306
atttcattctgtgtttgacaagcagaacggaaccacttcatttttttgtttggttgaagagtagaatagag 1376
cggagccgctccatttattgtttggttggagagccatatgactgttgcgtggaggagagggatgagagcg 1446
cttgcgtccgcgagagcgctcgcatcggatcgttttggtggagcgagagcatcccaaatatttgctattt 1516
atagactccaggaaccgctccgctctcctccatggtaaccaaacatcaaaaaaataagaatggagccgtt 1586
ccattcccctccgctcttcaaccaaacacattagtcaggttcttcatagtgttgctggaaggtcatatga 1656
atggcttagtcaggttcgttttcacttgaaatctttcatactgctttgatgcctaaagtattccttatgg 1726
tttcaccaagcattacaattgaacggacaagatatgttctctttcagaggatatttacagaaatgttagc 1796
cagacaatatagaggcaacagcgggttctaattctttggcaaccttcatttccatcccttctgtattctt 1866
cttataaggattcctatatataacccatgcgaacaccactatccatctcactgtgctttcaacttccttg 1936
aataaaagaacag GTACAGCCTCTATAGAAGAATCCGTTCAGACATTGATGAACATGAAGGAGGCTTGGA 2006
                Y  S  L  Y  R  R  I  R  S  D  I  D  E  H  E  G  G  L  E
AGCCTTCTCCCGTAGTTATGAGAAGTTTGGATTTAAT CGCAGgtattctttaacatgaagtgtttatctt 2076
  A  F  S  R  S  Y  E  K  F  G  F  N  R
ttatcctatatactagcaaataactttgtatcttattttccaatgtag CGCGGAAGGTATCACATATCGA 2146
                                                  S  A  E  G  I  T  Y  R
GAATGGGCTCCTGGAGCATTT gtatgatctttcttcaattattctaatcttattgcttatcatgatacaa 2216
  E  W  A  P  G  A  F
tactagttccatgtttcattatgagaatgatcactctcccagctacgaagctactcatatttaataaatc 2286
ttttaccttcaaaatatatacattgtcagacttagctcctagtgttattcagaattgactcctgctttct 2356
atcttag TCTGCAGCATTGGTGGGTGACTTCAACAACTGGGATCCAAATGCAGATCGTATGAGCAAA gta 2426
         S  A  A  L  V  G  D  F  N  N  W  D  P  N  A  D  R  M  S  K
```

```
tgcttatgccttcagtaaacattatatatttcacttattagattggtttgtttatctttttcattaactc 2496
aatatatgtgcactctatttaatcaactttccaaccggaaagtgcatatgttgcattcactggagtcagt 2566
tcctacaattgactctgcatataagtgtgctattagatacttgtgttgctgctgatactcttagaagtca 2636
atgtttccaaactgcacatgtgatacctgtattctaatgtctcgatatctttacaaattgcatgttccta 2706
taataggatgaattgaatactcatacttagtgtgcatggacttggagtaggtctaattcttaatgattct 2776
cgaagagttgtttgtttgtaagggttcttacccctttatttcttcttaatataataatgatacacaactc 2846
tcttgtatgttagcgaaaaaaaaagaaaacttactgatcctcgcatagtctaaatcgaggtactttgttc 2916
ttgcag AATGAGTTTGGTGTTTGGGAAATTTTTCTGCCTAACAATGCAGATGGTACATCACCTATTCCTC 2986
        N  E  F  G  V  W  E  I  F  L  P  N  N  A  D  G  T  S  P  I  P
ATGGATCTCGTGTAAAG gtagccgctttacctcattggttgttgtttttgctggggagcatctccaaaa 3056
H  G  S  R  V  K
ttaactggtgtgttttgaccatatatag GTGAGAATGGATACTCCATCAGGGATAAAGGATTCAATTCCA 3126
                              V  R  M  D  T  P  S  G  I  K  D  S  I  P
GCCTGGATCAAGTACTCAGTGCAGGCCCCAGGAGAAATACCATATGATGGGATTTATTATGATCCTCCTG 3196
 A  W  I  K  Y  S  V  Q  A  P  G  E  I  P  Y  D  G  I  Y  Y  D  P  P
AAGAG gtcttcctttcccatatttctgtacttacaaattttaaatatacctttgttcctactttaggcta 3266
E  E
gcaatttgtttcttggaaaataccgtttatttcattctggaatttctttacacccccttttggatccttgg 3336
aattgaattccattctaataatagtaatttaggcaaagatcaattaagctaatatggttttatgtgaaat 3406
atatttgtatactattattagcaatatgtgggggatatttatgtgctacattttactatagaggagtgag 3476
ccgaagagcgtcttataatttgcagagtataaacatattatgttgatacataaaattatttctcatactc 3546
caccctatgaatttgagataggcttatatcattgctttggaaagtggtaaaatgttaaattccaagctaa 3616
atagactactttattaagtaaattccaattcctccaaaatgaatagatctaaaagagccccttagtgtat 3686
ttcagttagtcttattttctgtggatctggatttactttctcacttgctactaattaagatattttac 3756
atgcatatgcag GTAAAGTATGTGTTCAGGCATGCGCAACCTAAACGACCAAAATCATTGCGGATATATG 3826
              V  K  Y  V  F  R  H  A  Q  P  K  R  P  K  S  L  R  I  Y
AAACACATGTCGGAATGAGTAGCCCG gtatgtcaccttttccttttgattctttagacactgctaggcat 3896
E  T  H  V  G  M  S  S  P
atgtaacacccggccactggggacgtcgtcaccccagggcgagtttttttttgactgctagccacctgca 3966
gttttagttaattataggcttcaacgatatttagatactatcatcttctcacattggtcttccttatcta 4036
agcgtatgtaatctaagtgtatatatttattttcataaatgttcactttgctactttgttagtgaagcca 4106
gcaacatttaattccaagttattctggaattcgtttaatgatgttgtcttgattgattgattccaatcgg 4176
atgtttgaacctctcaaaaatggatgttccaaagggagaatttttacattttcttgcaaaattagtaaca 4246
attttgaactgattaatgtaactttgaatgtcatttctgctccgctatatgtttctatagctcatagcat 4316
tatataacatttcaaaacaatattcatatgttgctcatgaatggtcatcaatatgttatatttcttctac 4386
tttagttgatacattttgcttaatatttaaaattaaattcatatttcccacaattactttgagttaaaaa 4456
aaataaacatggccatagcacagccacagtattgttactgtcaccctacatttgtttctaattgccatac 4526
gattttacatacatgttctttcactgatcattatccttttccggctatcattttgttatcttgtttgcac 4596
tagataattgcctgcattctcttgtgattag GAACCGAAGATAAACACATATGTAAACTTTAGGGATGAA 4666
                                 E  P  K  I  N  T  Y  V  N  F  R  D  E
GTCCTCCCAAGAATAAAAAAACTTGGATACAATGCAGTGCAAATAATGGCAATCCAAGAGCACTCATATT 4736
 V  L  P  R  I  K  K  L  G  Y  N  A  V  Q  I  M  A  I  Q  E  H  S  Y
ATGGAAGCTTTGG gtaatttcaggatccagttttgtttgtctttttcttctatatccactttatacctac 4806
Y  G  S  F  G
tggtacctgttgagtgtttgggaacttcatctttcatcgccctctttttagtgtcatttcatgcctctt 4876
```

```
tcactgttgcttctagtttgatagcattaaatgtacattttggtattgtcttcgataatctattgtattc 4946
agatgcatgcatagtaacaagactacacagacatatatgtaacgtcgtaacctcaactttatgaacatgc 5016
tctgattggattgatgattgcatgatactgtcaatatggcattacagaaactcgcaagtatcagtaccaa 5086
tgcttatagtctcctgtcataaaaacaagttgtttaggacaagatttagtcaaatcataaaaactacaaa 5156
tatcattatcttttaacttatttagtttgaaagcatgtaagttgacactaaaaaagcatgtaacttata 5226
tgtgtggaattgtattgaaaaatccttgcaaatctcatgcagtcgaaggattaagtggtcaaagatattc 5296
atcaaagaccatgcaatgttgaatcaagtattttgatagtaaggagtacatgtttccaacagcatgaatt 5366
ttgatacagtatccatctgttgagacgacggagatgataccatccatcttaaattttagttccgaatgat 5436
accttttcaattttgacaaacaacatgacaacgtcgatttaatattttgttacttctattgcag ATACCA 5506
                                                                  Y  H
TGTAACTAATTTTTTGCGCCAAGTAGTCGTTTTGGTACCCCAGAAGAATTGAAGTCTTTGATTGATAGA 5576
   V  T  N  F  F  A  P  S  S  R  F  G  T  P  E  E  L  K  S  L  I  D  R
GCACATGAGCTTGGTTTGCTAGTTCTCATGGATGTGGTTCATAG gtaattgataaattctttgttaataa 5646
 A  H  E  L  G  L  L  V  L  M  D  V  V  H  S
ctcatcttcctttaatacaagtatcttattgcatagttaaagctaaatttgatgatgtgtaattcatagc 5716
tccaagaaaaaggaaatacagtaagccttcatgcaaaattgaaacatttaaccctcgacttactggaccc 5786
gacctgcaactaatctacaacctctacaacatacatcagaagatgagaaataatcgccgtaacataacaa 5856
ttaccctgactaagatatatgacaagtcaaaatttagctgtaatgttcagtgttctaggcttaagcttt 5926
aacatcaggaaggaggggaagattatccaccaagtcccaaatcaagaatattcagaaagaacctgtactg 5996
tcaaagtccctttgatatatgaaaaaactggacatgcgtggggtaagacagtcccctggcattatattaa 6066
gaagacctcacacaggtcgagaaaacatccgaaccgtgccacacccatacacagcggcaccgtagcccat 6136
gtgagaaacgaccgcgtccgggaccggaccttagatatgtgctttggtatgtgatagatgaggggatttt 6206
tttaaccacagcctgaaattcgctcccacggtggtagaaacacctcggctgcgcagattataggagtgag 6276
tatcaagttggctggactagcgggagcgaaggggcgccgtgctgccgctcggctgtgaactcaggtgagt 6346
tgcggaacacaggtgaggcagggaacaatcacttcttagattgatcaatgcttcccaattgattacagac 6416
gcctgtgactttatagtccaactctaaacagatccttatctcaggaactctaacagattcttattaaaac 6486
agactccttaatctcaggatctctaaacagactctccttatctaaggatcctcctaatctttatctatac 6556
gtccctctataatgggccgactagagggctgtttaggccacctactataaaccataacacacggggagtc 6626
aaacccaggatctgaggagtgctactcagaccacctaaccaactcggctagaggccctttcgcaagcccc 6696
tttgatatcagccacccacatcttattttgaagtgcttgacatactgttcgctgtttttctgctctttt 6766
tggaaccaacttgatcaagtttggcaccaattctgccatggtcttgccattaagccagcgatcgaaccaa 6836
cattgagttctttaaaatttagtttaacaaacacttaattttagtttgattcatatttctttcttgaaca 6906
agcaggatagctgtgtatgtttgtattaacagagagaaatgcaaacaacccactttgtcttccttaggag 6976
aacaagggggttatggctgaggtttagagaaacgtacaatcacctgtaaataggataatatcaatacaa 7046
tgctactttaaactagcccctcaacccaacctagtcctagatcctggagtttacaggatccgcttgccaa 7116
acgaggccttcctcacccagattcctgctccttgtgtgactaagccggtgagaccattcagcacattcta 7186
attggttgtgtcttctctagacaaatttgggtcttgactttgcaatccttgcatctggaggccattacgc 7256
ccaatgggacagagttgggtttcttctcatggtgggcgcgctctggcaagatggttcccaagaacattcg 7326
aaaggggttgaacactctttgtattccgatagcctgagaaatttggaatttaggaacttgtgtgtgttca 7396
aagactcccaaccaaatgtacaactccttctacataggataggaagtgagggtcttctttggtgtgctgc 7466
agggttgcttgagctcctgaataggtcgttgtccccagcctgctaggtgctggtggccttggttgtcagc 7536
ttttgttgttgttttgtgtaaaacctagctggctagctgtttggactagggttgggtcattttgacctta 7606
ttcgctgtctttttctttcgtaaataaaatgacacgtagctctcctgcgtcgttcgagaaaaaaagtact 7676
cgtcagaatagaaaatttactctttgataaactataactacctaatcaagacaagaatataaaattta 7746
aatattggaaacctatttgtcgtggttcaaggttctttggcttccaagctttgggatgagatcgattata 7816
atctgttatgtacaaccttatgattatttgaagctctttggtttcatacctttaactttgcttttgtgtt 7886
```

```
acttgcag TCATGCGTCAAGTAATACTCTGGATGGGTTGAATGGTTTTGATGGTACAGATACACATTACT 7956
           H  A  S  S  N  T  L  D  G  L  N  G  F  D  G  T  D  T  H  Y
TTCACAGTGGTCCACGTGGCCATCACTGGATGTGGGATTCTCGCCTATTTAACTATGGGAACTGGGAA gt 8026
F  H  S  G  P  R  G  H  H  W  M  W  D  S  R  L  F  N  Y  G  N  W  E
acggaacaaaaatgctctatctccattaattttattctcctatttttctgcctgtatcgttccaacaatt 8096
ttatccgtatgcag GTTTTAAGATTTCTTCTCTCCAATGCTAGATGGTGGCTCGAGGAATATAAGTTTGA 8166
                V  L  R  F  L  L  S  N  A  R  W  W  L  E  E  Y  K  F  D
TGGTTTCCGTTTTGATGGTGTGACCTCCATGATGTACACTCATCACGGATTACAA gtaatttaagcttta 8236
  G  F  R  F  D  G  V  T  S  M  M  Y  T  H  H  G  L  Q
tgcctgttagtttatcttcacttgctaagtctgactggaatactggattatgcctgggaactagttttgt 8306
ttagtatcatatttgttatatatcattccttcttctaatctaaagtcatgcattttactttag GTAACAT 8376
                                                                 V  T
TTACGGGGAACTTCAATGAGTATTTTGGCTTTGCCACCGATGTAGATGCAGTGGTTTACTTGATGCTGGT 8446
F  T  G  N  F  N  E  Y  F  G  F  A  T  D  V  D  A  V  V  Y  L  M  L  V
AAATGATCTAATTCATGGACTTTATCCTGAGGCTGTAACCATTGGTGAAGAT gtaagtgctgagtttgct 8516
  N  D  L  I  H  G  L  Y  P  E  A  V  T  I  G  E  D
tgtcatttaatatgaattctcgcatatatttgtggaaatatttttgtagtcgaagttgcttttgtttatc 8586
tagacaagatactcctatttggttatgcagaagttaatttgaattttaatacgaagtgcacactaagtta 8656
ctggttaatattgttcttcatttcttcaagtttcagtctatttcaactccatttataataatgtgcttgg 8726
caagttactgttttaattttacattaatacacaatacaacaagatcatgctttacaacatgtgtgtatat 8776
tagataagtagttcatgcacatagagttgctacttttttgaagaaacatatagaattacttaaaaggaact 8846
atttgaaatacatgaagaaagttaatggctgaacctatattgtaatggacagaacactagaactttcggg 8916
tttatactgagcctggactataaagaagatttctcttgaagttcaatgagttctcgacttaaatattctt 8986
tctaaactactggatggataccaaagacacaaaaaattgaaattgtaagccactctgctcttgttttggc 9056
accagatttaatagaatataaataaaaattaataatggggagaagactccccttggattggctttaaaga 9126
gaggagtacagaacccttcccaactcgtccccgaaccaaactttgccaaaatcgggatccgcctcgtcct 9196
gtccccagtgctccttggaaccaaacacatgctgagggaatatactcctgggggatggagttatggaccg 9266
gttccaactcgtctccgaaccaatgatttcaagtcgtccgactaatcgcgattagtcgggctggtcggca 9336
attaggacacgatttgctgggcgactcgactagaagacctagtcgtcctggtcgttcgactaatcgtcga 9406
ctagggcgactagtcgttcgagttatgtgtcctggtcgtcccgactagggtttagtttattgggcctttt 9476
ttagcccatctacagtctaatttagagtaccctctcctctcctgttctaacctagccgccaacagtctcc 9546
acaacctacacgtctcctcttctctcctcccttggcatgccctgcagtcctgaatcctgattgccggcga 9616
ctacaccggcggctacagcaccctgcgtcccctcagggccggccctggacaggtgccggcggtgcggccg 9686
caccgggcctccgaaaaccagagggcccctccccgtgtatacgtgtatgtgtatatatatagaaaatgtt 9756
tatgtaaattacaacacaaagattgataaaaggccaaagcccatcaacgtacaatatgcatgcctgttta 9826
aactcgattacgtacttggcttctcgtggacgcttgagctgcaatcctctttgccttctcttggaagcct 9896
ccgctcatgttgtccaggctacaacaattgttcggcaccagcagtagcagctagcagcgacagacgggag 9966
ggatcatcttgtccgttgctctgtcatctgcccaaaccaggcttgccgcgcgccgtgcctagcggtagcg 10036
acatggtagcgtgccctgcctcggtaacttcgtcaagtctgctactctactttcgttgttcaattttcgc 10106
ctcaaaccaatctttactagttaaatgttcttttatttctgagtataggttattttgtgttctgatcacc 10176
acattccacatgcatctagtagtgaaaaaggaaaaaaacataaagataagttcaatgcatcgtagtattc 10246
atatttttttggtatttgtgatgaatggggccttaatttttgtctcgcaccgggccaccaaattctcagg 10316
gccggccctgcgtcccctgccccttctcccgcagtcctgctcgcccgctgcactgctgcagtctccaacc 10386
gatggtggctgcgagccagtgacgagatgagcatgtcatctgaaagtccttcctgctccccacttctccc 10456
cctaaaccctaaataaccctaaaccctaagcaacatgttactttatttatcctagttgtttgttctaatt 10526
tatataatgtatacaggtatatttatatatatacctatataagtataactactagtctaggacgaccagg 10596
```

```
gaccgactaggggtcgactagtcgccctagtcgtcgcctaatcgcgactagtcgcttggtcggtcacaga 10666
gttcgactaggcgactaggcgacttgaaatcattgctccgaaccaaacactcaggatcaccttgtcctgt 10736
cctggtggtccttggaaccaaacacacatttagggaagtaatcatctgacctgcaagcacctgggttcat 10806
gcaggtggataggatctcagctagaggctctaccaaccaagttattgttcgattctcagttctcatgctc 10876
cgttgtgcaatctaatgtaaataaatgtaccaccgtgttgtacgcatataccagcgtaatctccaacaag 10946
ggctgtcatgtggaatcaatagaggtctctgtcataatctgcaacggcaacaaaggacataccaatcttt 11016
aacacgaagagaattaactccttcccatctgtagttagatattagaatattagaattgcatatttttatg 11086
ggttcttgaaggaagtggcaaacctcaatattgaactctaccaacacatgtataatatgaaaaataaagt 11156
tatggtatatttacgagtaaatatttctatgccttataacgggaggaaaaatatcgaattttgatattga 11226
cttaataggacattgtcacgttattgtcaatattaacttatattatggatgtcatcacaaaataactata 11296
ccttttagttttaaagaaagggtaaaacatacatatatttgtaacgtttccgtccattttccattaatg 11366
tttgattgtttttttctctcttgttgcaacatagtgacatgtttgctagtttgacaaaattagggcaaga 11436
tcattggcttacataaaccaataggaaatttgaatttatctaaaataatttgtgagttttcttggttatt 11506
caaaaattatataacttgttcag GTTAGTGGAATGCCTACATTTGCCCTTCCTGTTCACGATGGTGGGGT 11576
                         V  S  G  M  P  T  F  A  L  P  V  H  D  G  G  V
AGGTTTTGACTATCGGATGCATATGGCTGTGGCTGACAAATGGATTGACCTTCTCAA gtaagtgtttcat 11646
  G  F  D  Y  R  M  H  M  A  V  A  D  K  W  I  D  L  L  K
atgtatgtggacgatcattattttgttttattgggctgcttatttaaataattatttttttggcataaa 11716
tgttataagtacatacactaatgtttaattgcatgggccatgtgcatatgtatttttttgttgtatttgtc 11786
accacttttgtggtttactttacatagtcataatgagataatgttttatgcacatttgcatgtggctgca 11856
cttatgattaacgacacaaaatgtactcgagattgatgtatattttcaaacttgaaactaatgaacacat 11926
atgatatacattgtacactccatgttacaaattataagacgtttttgcattttttagatatattttttttac 11996
tatgtatctagacaaagtgtatatttaagtgcattgcaaaggcaatgcatctagaaaagccaaaatgcct 12066
ttgcatttggaacggaaggagtaatattgttgcacaaatcttgaagttttctatgcataggatattgcac 12136
ataccagcaatatatcagtggtgcatatatatttgtataactaacattacaccattcactgtgttcata 12206
tgttttgtgtttaatttctatacaaaaaatctcttgaatctggtactcaaatcatttgtataagctttt 12276
gtttcaaacgggcaacgctctatctgtaatatgccatctgcgtaaaagacaagtatgtaattttgtgatc 12346
ctttggttgtgtggttgtcggtgtaaccgggattctttaacctttatccttcttttatataatgataca 12416
caaactctcctgtgcgttcgagagaaaaaatgccttctgcattcactttgagatatgtggtgagtttcaa 12486
tttctatttaaccgcacag GCAAAGTGATGAAACTTGGAAGATGGGTGATATTGTGCACACACTGACAAA 12556
                      Q  S  D  E  T  W  K  M  G  D  I  V  H  T  L  T  N
TAGGAGGTGGTTAGAGAAGTGTGTAACTTATGCTGAAAGTCATGATCAAGCATTAGTCGGCGACAAGACT 12626
  R  R  W  L  E  K  C  V  T  Y  A  E  S  H  D  Q  A  L  V  G  D  K  T
ATTGCGTTTTGGTTGATGGACAAG gttaccctactcattaattttcttggtgtacttattgggacataga 12696
 I  A  F  W  L  M  D  K
tcatgtttcacgtattgtttttttacaatgattaattctatttgtttcttcaagctcaaggtgtattcatg 12966
gttcccacaaaacaaatgttttattaaaagcaaagagttagaacttttgttagttttctttaatttggac 13036
ttgtgtcactgtttctctgtcatgactcatgagcattatagttgcaatttcacctaagtgagttcctgtt 13106
tttggacagttcagagtgaactacgacttattgttttaatacttcattgagtttgtagaacaagtattgg 13176
ctctttctcatcctatactttcaaaagtaggatttggcctttttcttccactattctttgaaaaatacat 13246
gtagaaacacggaacaataataaatggtaacatgagaacctcactggttctatttatgcag GATATGTAT 13316
                                                              D  M  Y
GATTTCATGGCCCTCGATAGACCTTCAACTCCTACCATTGATCGTGGGATAGCATTACATAAGATGATTA 13386
 D  F  M  A  L  D  R  P  S  T  P  T  I  D  R  G  I  A  L  H  K  M  I
```

```
GACTTATCACAATGGGTTTAGGAGGAGAGGGCTATCTTAATTTCATGGGAAATGAGTTTGGACATCCTG g 13456
R  L  I  T  M  G  L  G  G  E  G  Y  L  N  F  M  G  N  E  F  G  H  P
tgagatttaactactttgtttcatttaaccttcgttgagtcttatagaacagtacctcatccaacaatta 13526
tcttgcaatttatcttttgttagttatattagtgttgaggacttgaggtcattttcttcttattattttg 13596
cag AATGGATAGATTTTCCAAGAGGTCCGCAAAGACTTCCAAGTGGTAAGTTTATTCCAGGGAATAACAA 13666
   E  W  I  D  F  P  R  G  P  Q  R  L  P  S  G  K  F  I  P  G  N  N  N
CAGTTATGACAAATGTCGTCGAAGATTTGACCTG gtaaactttctttgattgtgcaaaagtccaagtttg 13736
  S  Y  D  K  C  R  R  R  F  D  L
tatttacttttaccactgatccagtgctttaatcagcaaggtgccattataatagttccttctttattca 13806
tattagcatgttccagaagtaaaaatattactacctttgtaaaagttttctttaatatatgtcgctttgt 13876
ttggtcaataattcgtcattatcggaattgtgttatttttacattgtcag GGTGATGCAGACTATCTTAG 13946
                                                   G  D  A  D  Y  L  R
GTATCATGGTATGCAAGAGTTTGATCAGGCAATGCAACATCTTGAGCAAAAATATGAA gtatgttctttt 14016
  Y  H  G  M  Q  E  F  D  Q  A  M  Q  H  L  E  Q  K  Y  E
tttactttttgatttggttctgcaaggtttccacaaacatcatatttgttgtgcattctacttgtaatg 14086
tcattttaaaaaaatcattcctcagttttactgagcttttaagcaatgaaggtttcattatgaattctt 14156
tcatgttgcatcaacaactcttaggtattttcatgatcattaatagtactctggagacagcacgccataa 14226
tggtaacgaaaaatttctgatgaaatttgctgtgtaattgcag TTCATGACATCTGATCACCAGTATAT 14296
                                            F  M  T  S  D  H  Q  Y  I
TTCCCGGAAACATGAGGAGGATAAGGTGATTGTGTTCGAAAAGGGAGATTTGGTATTTGTGTTCAACTTC 14366
  S  R  K  H  E  E  D  K  V  I  V  F  E  K  G  D  L  V  F  V  F  N  F
CACTGCAACAACAGCTATTTTGACTACCGTATTGGTTGTCGAAAGCCTGGGGTGTATAAG gtatgcatct 14436
 H  C  N  N  S  Y  F  D  Y  R  I  G  C  R  K  P  G  V  Y  K
atcttgcattccctatgctcaaagtgcatttcttttcttgataaatgagttagatatacgtactatcatg 14506
ctgcaatttatcaagtgtcattattgatctctttctacggtgaagctaggagcagctaagctgttggtgt 14576
cagcaattcatgttgtagttaatttaatttgcttgaaaacgtaggacgctagatttggatttttccaatt 14646
tttaggctgcacgagcaggtaaaaggtagcaaaatactagggcgccatgtttacatgtataaaaaaaaca 14716
aaacaaaaaagaactaggagttcctgtgacggatagccgcatgctcgttctcttggcgtctctgatattg 14786
gagcacatccgttgttcgaaaactagccggagaagtttctcaaaatcccactagcggaggtgctgacagc 14856
atttgctatttgtaccag GTGGTCTTGGACTCCGACGCTGGACTATTTGGTGGATTTAGCAGGATCCATC 14926
                    V  V  L  D  S  D  A  G  L  F  G  G  F  S  R  I  H
ACGCAGCCGAGCACTTCACCGCC gtaagttttgtggcacgtgatactgctctaggtacgcagatgtccac 14996
H  A  A  E  H  F  T  A
ttgttcctgacagaggtgaactaacttctgttatggccattacttgcag GACTGTTCGCATGATAATAGG 15066
                                                   D  C  S  H  D  N  R
CCATATTCATTCTCGGTTTATACACCAAGCAGAACATGTGTCGTCTATGCTCCAGTGGA GTGATAGCGGG 15136
 P  Y  S  F  S  V  Y  T  P  S  R  T  C  V  V  Y  A  P  V  E  *
GTACTCGTTGCTGCGCGGCATGTGTGGGGCTGTCGATGTGAGGAAAAACCTTCTTCCAAAACCGGCAGAT 15206
GCATGCATGCATGCTACAATAAGGTTCTGATACTTTAATCGATGCTGGAAAGCCCATGCATCTCGCTGCG 15276
TTGTCCTCTCTATATATTTAAGACCTTCAAGGTGTCAATTAAACATAGAGTTTTCGTTTTTCGCTTTCCT 15346
 ●                                  Poly (A) signal
AAtgcttgatggctgattgtttgcacttgtttcattccgttgggcactgatggtcttagagttagacaat 15416
cggctgcagcgcataggtttcaagctgggggttgcatgtccgactacagagcagccagcaatgtggccc 15486
ctgctgcctgctctgcttacttttaaatgccacccctcccgattaccgactcactcagattcagacacgc 15556
```

Table 3.1. List of introns and sequences of exon/intron borders in the *Sbe2b* gene.

Intron number	Exon/	Putative intron branch point	/Exon	Intron size(bp)	GC content (%)
1	ACTC	GTAA.....GTGAA..24..GCAG	GGGG	106	51.9
2	GGAG	GTTC.....CTGAA..29..CCAG	GTAC	244	41.4
3	ATCG	GTAT.....TTCAA..21..ACAG	GTAC	1086	43.9
4	GCAG	GTAT.....ATAAC..24..GTAG	CGCG	76	25.0
5	ATTT	GTAT.....TTCAG..25..TTAG	TCTG	196	31.1
6	CAAA	GTAT.....CTAAA..22..GCAG	AATG	499	33.5
7	AAAG	GTAG.....TTAAC..23..ATAG	GTGA	81	40.7
8	AAGA	GTCT.....TTAAG..21..GCAG	GTAA	567	30.3
9	CCCG	GTAT.....ATAAT..23..TTAG	GAAC	774	32.0
10	TTGG	GTAA.....TTAAT..21..GCAG	ATAC	751	33.2
11	ATAG	GTAA.....TTAAC..22..GCAG	TCAT	2274	42.3
12	GGAA	GTAC.....TTTAT..49..GCAG	GTTT	86	34.9
13	ACAA	GTAA.....CTAAA..19..TTAG	GTAA	148	31.1
14	GTAA	GTGC.....TTCAA..20..TCAG	GTTA	3051	41.8
15	TCAA	GTAA.....TTCAA..19..ACAG	GCAA	872	31.8
16	CAAG	GTTA.....ATGAG..25..GCAG	GATA	457	32.6
17	CCTG	GTGA.....GTCAT..21..GCAG	AATG	144	30.6
18	CCTG	GTAA.....ATTAT..28..TCAG	GGTG	226	30.1
19	TGAA	GTAT.....ATGAA..19..GCAG	TTCA	266	31.2
20	TAAG	GTAT.....CTGAC..21..CCAG	GTGG	448	40.0
21	CGCC	GTAA.....CTAAC..24..GCAG	GACT	96	45.8

Consensus sequences between introns are underlined.

3.3.3. Genomic Southern Blot Analysis

Southern blot analyses were performed to determine the number of genes in the maize genome that are similar to *Sbe2b*. When blots were probed with the full-length maize *Sbe1* cDNA (Fisher et al., 1993) under high-stringency conditions, at least two strongly hybridizing bands were observed in each lane as shown in Figure 3.5. The band patterns agreed with the restriction map of the *Sbe2b* genomic DNA, suggesting all the bands were derived from a single genetic locus. Therefore, we concluded that a single copy of the *Sbe2b* is present in the maize genome. To confirm this, the blots were probed with a small fragment of the *Sbe2b* genomic DNA which does not have any restriction enzyme sites for BamHI, EcoRI, BglII and HindIII. The result is shown in in Figure 3.6. As expected, only a single band was observed in every lane, supporting the conclusion.

3.3.4. Analysis of the 5'-flanking Region of the *Sbe2b* Gene

To identify the 5'-flanking regions necessary for *Sbe2b* gene expression, we utilized a transient expression assay system in maize endosperm suspension cells. In an initial experiment, a transcriptional chimeric construct containing the *Sbe2b* gene fragment between -2964 and +100 linked to a luciferase (LUC) reporter gene in pUC119 was created and called pKL201. In the case of maize *Sbe1* gene, transient expression assays demonstrated that the level of LUC expression driven by the *Sbe1* promoter greatly depended on the presence of the DNA region spanning the first exon and intron of the maize *Sbe1*.

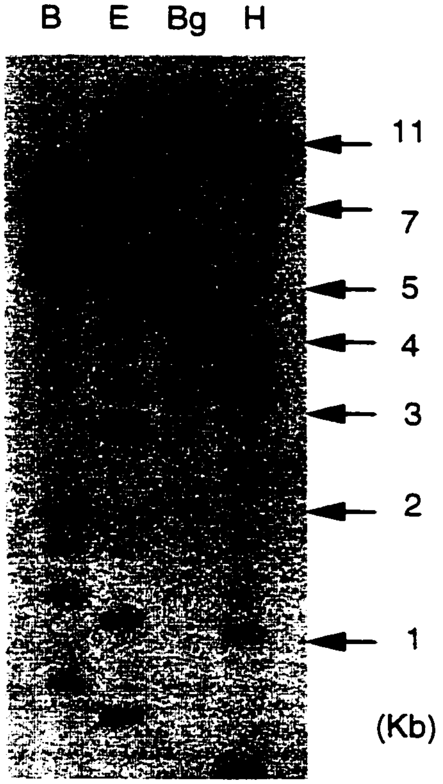


Figure 3.5. Southern blot analysis of maize genomic DNA probed with the full-length *Sbe2b* cDNA.

Each lane contains 10 μg of maize genomic DNA digested with the indicated restriction enzymes: B, BamH I; E, EcoR I; Bg, Bgl II; H, Hind III. Genomic DNA was prepared from etiolated maize seedlings (inbred B73). The maize *Sbe2b* full-length cDNA was 32P-labeled using the random primed DNA labeling kit (Boehringer Mannheim) and used as a probe. Hybridization and washes were performed at high-stringency conditions. Arrows indicate the position of the DNA size markers in kilobases.

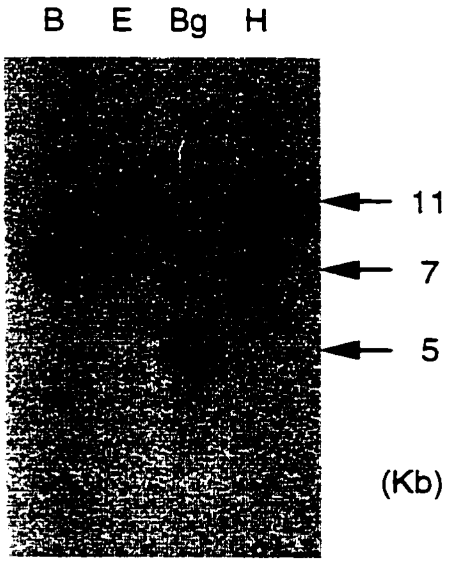


Figure 3.6. Southern blot analysis of maize genomic DNA probed with the small *Sbe2b* genomic DNA fragment.

The 326-bp PstI and XhoI fragment in the maize *Sbe2b* genomic DNA was isolated, labeled and used as a probe. Each lane contains 10 μg of maize genomic DNA digested with the indicated restriction enzymes: B, BamH I; E, EcoR I; Bg, Bgl II; H, Hind III. Genomic DNA was prepared from etiolated maize seedlings (inbred B73). Hybridization and washes were performed at high-stringency conditions. Arrows indicate the position of the DNA size markers in kilobases.

To test effect of the first exon and intron regions of the maize *Sbe2b* gene on gene expression, a translational fusion construct (pKLN201) was created by including an additional *Sbe2b* DNA from +101 to +329 into the plasmid pKL201. These plasmids were then tested by assaying LUC activity that accumulated after introduction of the DNA into maize endosperm suspension cells by particle bombardment. Plasmid pBI221 containing the CaMV 35S promoter linked to GUS gene was used as an internal control to correct for transfection efficiency. The results showed that levels of LUC expression driven by the two constructs were almost same (data not shown). This suggested that unlike the maize *Sbe1* gene, the first exon and intron region of the *Sbe2b* is not necessary for gene expression in maize endosperm cells.

To define the promoter sequences important for the *Sbe2b* expression in maize endosperm cells via transient expression assays, a series of 5' deletion mutants as shown in Figure 3.7 was derived from the plasmid pKL201 using available restriction enzyme sites and PCR techniques. The activity of each 5' deletion construct is presented in Figure 3.8. Two consecutive deletions of the *Sbe2b* 5'-flanking sequence down to -1128 decreased the level of the LUC expression to approximately 40% of the full-length promoter level. However, the removal of an additional 353 bp, to -775, restored the LUC activity. This suggests that potent positive and negative distal *cis*-regulatory elements may be located in the regions from -2964 to -1129 and from -1128 to -776, respectively. Although further deletions down to -160 did not significantly change the levels of LUC expression, a dramatic reduction in the promoter strength was observed when an additional 111 bp, to -49, was deleted. This indicates that a very strong positive regulatory element(s) is located between -160 and -49. The -49 deletion construct (pKL208) still exhibited promoter activity even though it was very low.

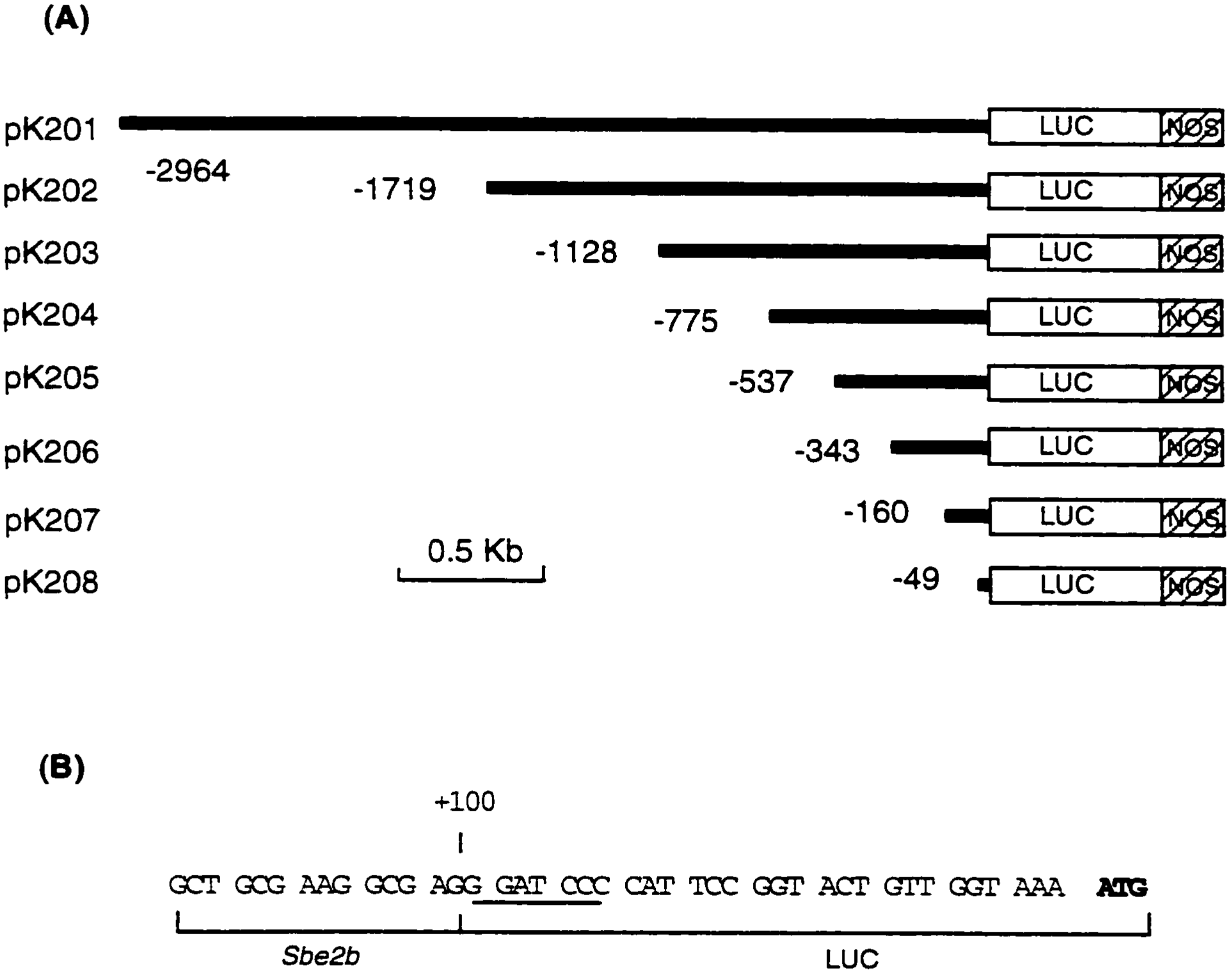


Figure 3.7. Schematic diagram of the 5' deletion chimeric constructs.

(**A**) The thick black lines denote the *Sbe2b* promoter sequences. The numbers at left indicate deletion-end points relative to the transcription initiation site (+1) of the *Sbe2b* gene. The open and stripped boxes indicate luciferase gene and nopaline synthase 3' end sequences, respectively.
(**B**) The junction sequences between the *Sbe1* gene and LUC. The BamHI site used to join the two genes is underlined. The translation start site of *LUC* is indicated by boldface letters.

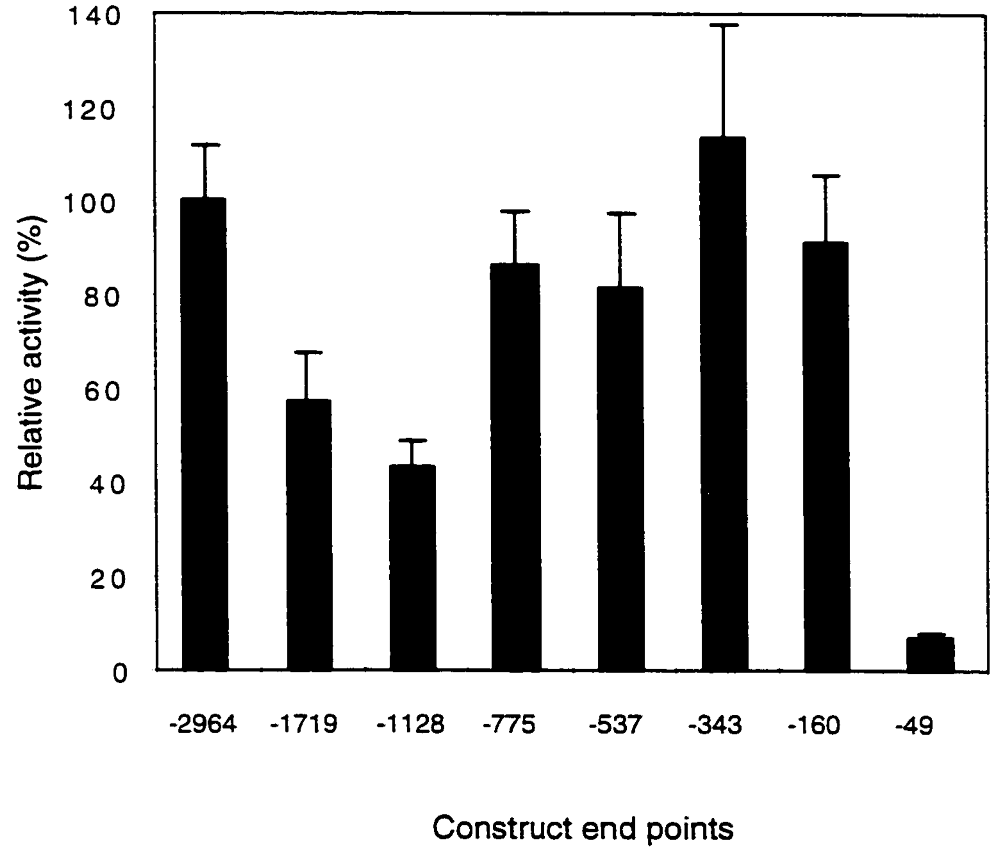


Figure 3.8. Effect of 5' deletions on *Sbe2b* promoter activity.

The relative activity values of the constructs described in Figure 3.7 are percentages of pKL201 level. Each value represents the average of three independent shootings. Error bars indicate standard errors of the means.

3.4. Discussion

In this study, we established the complete genomic organization of the maize *Sbe2b* gene, which contains 22 exons and 21 introns. Primer extension analysis demonstrated that transcription mainly initiates at a nucleotide 100 bp upstream of the translation start site. The consensus TATA-box sequence is located at the appropriate position, 28 bp upstream from the transcription initiation site. Alignment of the genomic sequence with the *Sbe2b* cDNA (Fisher et al., 1993) revealed that the poly(A) addition occurs at a position 219 bp downstream of the translation stop codon (TGA). This indicated that the transcribed region of the *Sbe2b* is 15,347 bp in length which is almost three times longer than the *Sbe1* gene.

Although the coding regions of the genes *Sbe1* and *Sbe2b* have approximately 60% similarity in DNA level, the genomic organization of the maize *Sbe2b* gene was very different from that of the maize *Sbe1* gene. The *Sbe2b* gene is interrupted with almost twice as many introns as the *Sbe1* gene, and the total length of introns (12,448 bp) of the *Sbe2b* gene is almost 4.5 times longer than that (2,796 bp) of the *Sbe1* gene. Comparison of their 5'-flanking sequences exhibited little similarity, which possibly account for different expression patterns between the two genes. The G-box motif present in the *Sbe1* promoter was not found in the corresponding region of the *Sbe2b*.

Transient expression assays of the *Sbe2b*-LUC chimeric constructs in maize endosperm cells showed that the 5'-flanking 3.0-kb region of the *Sbe2b* gene has promoter activity, and unlike the *Sbe1* gene, the DNA region containing the first exon and intron of the *Sbe2b* gene does not have the ability to increase the level of gene expression. Among the possible *cis*-regulatory elements identified by the 5' deletion analysis, the 122-bp region between -160 and -49 relative to the transcription initiation site appears to be the most important region for the gene expression in maize endosperm cells.

Many maize mutants defective in the amylose-extender (*ae*) locus have been isolated (Moore and Creech, 1972; Garwood et al., 1976; Hedman and Boyer, 1983; Stinard et al., 1993). The typical phenotype of *ae* maize mutants is a glassy, tarnished endosperm containing reduced amounts of starch with a higher proportion of amylose (up to 70%)(Shannon and Garwood, 1984). In addition, the *ae* mutant synthesize an amylopectin with a longer than average chain length (fewer branch points) in the endosperm. Among these mutants, the identification of a dominant mutant allele of the *ae* locus (*Ae-5180*) led to the cloning of *Ae* gene which is now synonymous with the *Sbe2b* gene. Unfortunately, no mutants lacking either SBEI or SBEIIa activity have been reported so far. Possible explanations for this are that at least two genes encoding each SBEI and IIa may be present in the maize genome and/or their roles in starch biosynthesis can be compensated for by other SBE isoforms. In fact, I demonstrated that two *Sbe1* genes are present in the maize genome (Chapter 2). Gao et al (1997) have shown that more than one *Sbe2a* genes are also present.

Analysis of the *ae* mutants in maize endosperm demonstrated the importance of SBEIIb activity in starch biosynthesis. Endosperm extracts from the *ae* mutant contained no detectable amounts of SBEIIb, while SBEI and IIa activities remained unaltered in the mutant (Boyer and Preiss, 1981; Hedman and Boyer, 1983). Hedman and Boyer (1982) also demonstrated that as the number of the functional *ae* alleles increases, SBEIIb activity increases almost linearly without affecting SBEI and IIa activities. These findings suggested independent genetic control of SBEIIa and IIb.

Despite these overwhelming genetic data suggesting that SBEIIa and IIb are encoded by two separate genes, it remained controversial as to whether they are the products of one or two genes. According to the single-gene hypothesis which mainly relies on the similarities in immunological and biochemical properties between SBEIIa and IIb, the differences in the two isoforms are the result of posttrancriptional or posttranslational modification of the *Sbe2a/2b* transcript or protein (Singh and Preiss, 1985; Preiss, 1991). Recently, however, analysis of *ae* mutants and cloning of *Sbe2a* cDNA provided solid evidence showing that SBEIIa and IIb are encoded by separate genes in maize endosperm. For example, Fisher et al. (1996) showed that *ae*-B1 endosperm contains a virtually undetectable level of the SBE IIb transcript but has SBEIIa enzymatic activity, and Gao et al. (1997) has isolated a near full-length *Sbe2a* cDNA clone which has a DNA sequence that is distinct from *Sbe2b* cDNA.

Now, to provide definitive evidence for independent genetic control of the SBEIIa and IIb in maize endosperm cells, we compared the complete genomic sequence of the *Sbe2b* gene to the recently cloned maize *Sbe2a* cDNA (Gao, et al. 1997). The result shows that the *Sbe2a* cDNA is not an alternatively spliced product of the *Sbe2b* gene, indicating that maize SBEIIa and IIb are the products of distinct genes.

Stinard et al. (1993) showed that the *Ae-5180* mutation is associated with two *Mu1* insertions flanked by complex rearrangements of *ae*-related sequences. Interestingly, this allele is dominant and appears to affect both SBEIIa and IIb activity (J. Preiss, personal communication). Due to lack of information on the *Sbe2b* genomic structure as well as DNA sequence, however, they were not able to derive a clear understanding as to the organization of this mutation. As shown in Figure 3.9, structural data of the *Sbe2b* gene was applied to published restriction maps of the two HindIII fragments, 12 kb and 12.5 kb, cloned from the *Ae-5180* mutant. It clearly shows that both *Mu1* insertions occurred in the 5'-flanking region of the *Sbe2b* gene. The *Mu1* element in the 12-kb genomic fragment is flanked by inverted repeats of *Sbe2b* DNA with a small region of asymmetry (150 bp or less) located immediately next to the right side of the *Mu1* element. Interestingly, the repeated regions contain at least 400-bp of the *Sbe2b* proximal promoter sequences which exhibited nearly wild-type transcriptional activity according to our 5' deletion analysis (Figure 3.8). It is possible that the three copies of the promoter region in *Ae-5180* may still be functional. Since the two HindIII fragments were reported to be tightly linked genetically, there is a possibility that the *Ae-5180* allele actually produces antisense mRNA of the *Sbe2b* gene. This provides a possible molecular mechanism for a dominant gene action of the *Ae-5180* which suppresses the accumulation of the *Sbe2b* transcript from wild-type *Ae* alleles. This also explains why wild-type revertants of the *Ae-5180* were accompanied with the loss of the 3.0-kb XhoI fragment in the 12-kb clone containing the inverted promoter regions, but not with the 2.3-kb XhoI fragment in the 12.5-kb clone (Figure 3.9). However, as discussed in Stinard et al.'s report, we cannot rule out the possibility that the *Ae-5180* actually suppresses in *trans* the expression of wild-type *Ae* alleles via a mechanism that allows related sequences to recognize and interact with each other, and results in gene silencing.

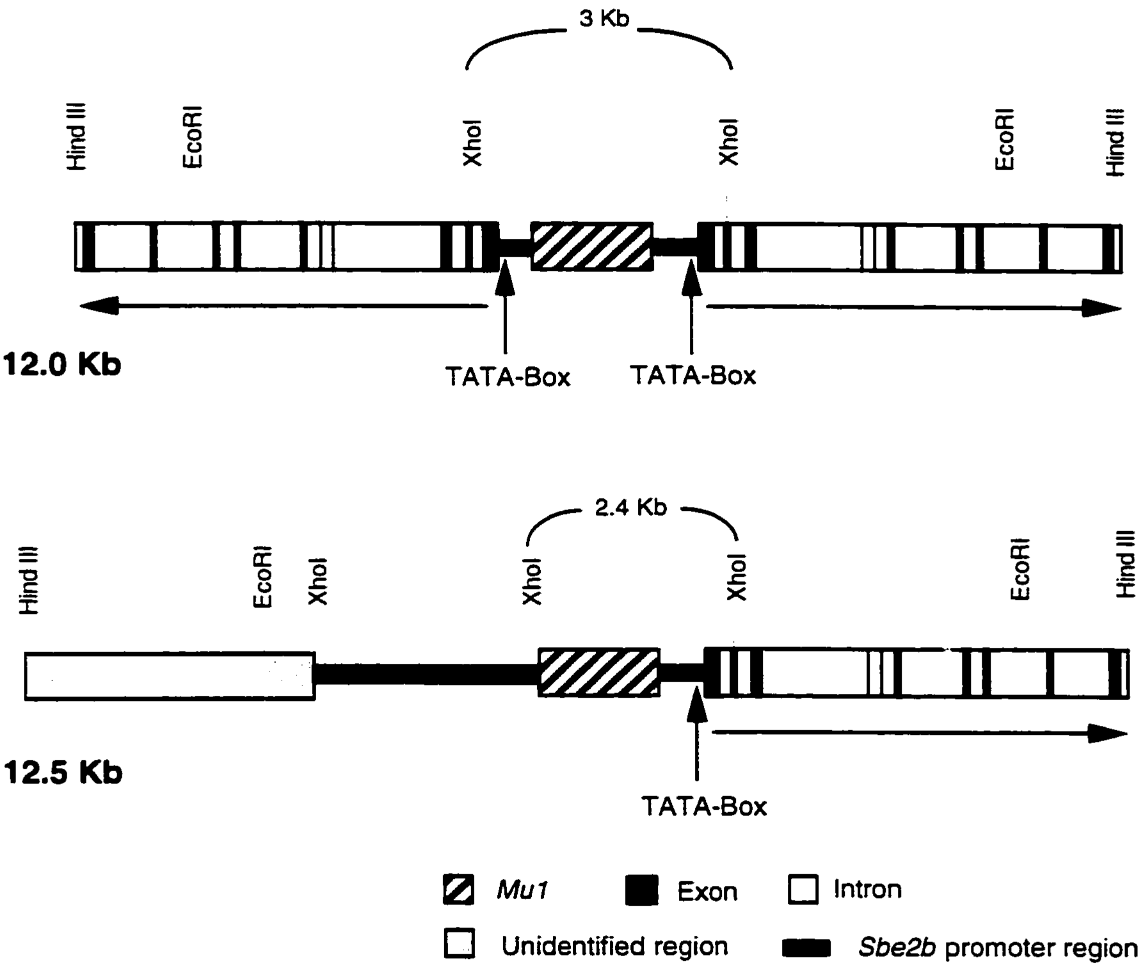


Figure 3.9. Structure of the 12.0- and 12.5-kb HindIII *Ae-5180* fragments cloned by Stinard et al. (1993).

Exons and introns of the *Sbe2b* gene are indicated by the solid black boxes and the open boxes, respectively. The thick black lines denote the *Sbe2b* promoter region. The segments representing the 1.4-kb *Mu1* inserts are indicated by hatched boxes. The stippled box indicates unidentified region. The arrows represent directions of the transcription.

Chapter 4

CONCLUSIONS AND FUTURE STUDIES

Starch is the major form of carbon and energy reserve in plants and provides a significant portion of the food for the human population of the world. It is also an important industrial commodity. Although our understanding of starch biosynthesis in plants has greatly improved over the past three decades, little is known about regulatory mechanisms controlling the genes involved in starch biosynthesis. In order to better understand regulation of starch biosynthesis in higher plants, our laboratory has been working on molecular characterization of the maize genes encoding starch branching enzymes. Our ultimate goals are to define the individual roles of each isoform of SBE in starch biosynthesis and to identify the regulatory DNA elements and transcription factors involved in their gene expression. Such knowledge will not only establish a basis for unraveling the regulatory mechanisms of starch biosynthesis, but will also allow us to generate transgenic plants that might enhance total starch production or produce unique starches for special industrial purposes.

The Maize *Sbe* Gene Family

As a step toward these goals, maize genomic fragments containing the entire coding region of *SbeI* and *SbeIIb*, respectively, as well as 5'-and 3'-flanking sequences were first isolated and their genomic structures were completely established. After determination of the number of *SbeI* and *Sbe2b* genes in the maize genome, DNA sequences important for their gene expression in maize endosperm cells were defined.

In order to fully understand the individual roles of each isoform of SBE in starch biosynthesis, characterization of all the *Sbe* genes in the maize genome must first be completed. Therefore, it is necessary to isolate and characterize the second *Sbe1* gene. Since genomic blot analysis clearly displayed the sizes of fragments containing the second *Sbe1* gene in each restriction enzyme digestion, the second gene could be easily cloned by screening sub-genomic libraries constructed from maize genomic DNA cut with restriction enzymes, size selected on agarose gels and ligated via linkers into lambda ZAPII vector. In addition, genomic clones containing the *Sbe2a* gene could also be isolated by screening the maize EMBL-3 genomic library used in this research with a gene-specific cDNA probe identified by Gao et al. (1997). Once positive clones are isolated, they can be restriction mapped and sequenced. Also, DNA sequences important for their expression in maize endosperm cells could be tested by the transient expression assay system which was developed in this research.

After identification of all of the genes encoding SBE isoforms, a transposon tagging system could be used to obtain genetic and biochemical data which would definitively prove the identity and functional importance of each *Sbe* gene in starch biosynthesis. In the past, transposon tagging only allowed us to clone genes which exhibit identifiable mutant phenotypes. However, the Trait Utility System for Corn, a *Mutator* (*Mu*) based system recently developed by Pioneer Hi-Bred International, Inc. has provided a valuable tool with which we can perform reverse genetics of any cloned gene.

Downstream Control Element

Addition of the DNA sequence (+28 to +228) containing the first exon and intron of the *Sbe1* gene into the transcriptional chimeric construct (pKL101) increased reporter gene expression in maize endosperm suspension cells up to 14-fold. Since such DNA sequences containing transcriptional stimulating effects are useful in investigations of gene expression in plant cells and for plant genetic engineering, it will be necessary to determine whether or not the DNA sequence has the ability to increase gene expression under the control of other promoters. The sequence could be fused to the ubiquitin or CaMV 35S promoters and then tested by transient expression assays. In addition, further experiments will be required to answer the following questions. Which part of the sequence is necessary and sufficient for the increased gene expression? Does inclusion of multiple tandem copies of the sequence have an additional effect on gene expression? How does placement and orientation of this sequence effect expression levels? What is the mechanism responsible for the stimulating effect exerted by the DNA sequence within the transcribed region?

Sugar Responsive Gene Expression

Northern blot analysis (Figure 2.17) and transient expression assays (Figure 2.18) indicated that expression of the *Sbe1* gene in maize endosperm is regulated by sugar concentration. Therefore, it would be interesting to test whether other *Sbe* genes are also sugar-modulated. This could be easily carried out by stripping off the radioactive probe from the membrane used in Figure 2.17 and then sequentially hybridizing with gene-specific probes for other *Sbe* genes.

The 5'-flanking sequences proximal to the protein-coding region of rice and maize *Sbe1* genes are highly divergent from each other, but both genes contain the canonical G-box sequence (CCACTGG) which are located in similar positions relative to the corresponding transcription initiation sites. In addition, co-expression of mEmBP-1 protein with a reporter gene driven by the *Sbe1* promoter showed that mEmBP-1 protein down-regulates transcription of the *Sbe1* promoter in the transient assay system. Taken together, these findings suggested that the G-box might be involved in regulation of the *Sbe1* gene expression even though 5' deletion analyses failed to show the importance of G-box in the *Sbe1* promoter activity. Since the G-box is known to be a *cis*-acting DNA regulatory element present in the promoters of many plant genes responding to a variety of different environmental and physiological stimuli, effect of light and plant hormones on the *Sbe1* gene expression should be investigated in the future. Using pKLN101 and a mutant chimeric construct created by substituting the wild-type G-box sequence in pKLN101 with mutated nucleotides, this investigation can be carried out via a transient assay system. If the *Sbe1* gene responds to a certain factor (for example, sugar) and the G-box would be implicated in such response, this would provide a valuable tool with which we can demonstrate involvement of mEmBP-1 or other G-box binding proteins in a signal transduction pathway.

Recent studies showed that expression of *Sbe* genes is differentially regulated in a tissue-specific manner (Gao et al., 1996): *Sbe1* and *2a* are expressed in both vegetative and reproductive tissues, while *Sbe2b* is mainly expressed in endosperm and embryos. Therefore, it will be very useful and informative if DNA elements involved in the tissue-specific expression of *Sbe* genes are discovered. In order to gain this information, the

Sbe promoter chimeric constructs described in this thesis could used to make transgenic plants to identify *cis*-regulatory elements responsible for the tissue-specificity.

All the *cis*-acting elements identified by the loss-of-function experiments will be further tested by gain-of-function experiments to confirm the functional significance of each element. One, two, and three tandem copies of the various DNA elements will be fused to a -64 CaMV 35S minimal promoter and the gene fusions will be tested for transcriptional activity using a particle bombardment-mediated transient expression system.

Using EMSA and DNase I footprinting (or methylation interference), the interaction of nuclear proteins with the DNA elements identified above can be tested. Nuclear extracts will be prepared from various maize tissues such as endosperm suspension cells incubated in different conditions, leaf and root. The DNA elements which form specific DNA-protein complexes could be used as probes to clone cDNAs encoding nuclear proteins involved in the transcriptional regulation of the *Sbe* genes by screening maize expression cDNA libraries. Knowledge of the *Sbe* promoter elements and their associated regulatory proteins may lead to a better understanding of the relationships between the regulation of the *Sbe* genes and of other genes, especially genes encoding storage proteins in maize endosperm. Giroux et al. (1994) showed that expression of genes involved in starch and storage protein synthesis of the maize endosperm are coordinated. Mutations altering synthetic events in one biosynthetic pathway may affect expression of genes in both pathway. Thus, knowledge gained in this thesis and in the future may have broad significance to our understanding of plant biology and to the use of biotechnology for the production of novel plants.

REFERENCES

Ardita, F., and Tandecarz, J.S. (1992). Potato tuber UDP-glucose; protein transglucosylase catalyzes its own glucosylation. Plant Physiol. **99,** 1342-47.

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

Baba, T., Kimura, K., Mizuno, K., Etoh, H., Ishida, Y., Shida, O., and Arai, Y. (1991). Sequence conservation of the catalytic regions of amylolytic enzymes in maize branching enzyme-I. Biochemical and Biophysical Research Communication **181,** 87-94.

Bae, J.M., Giroux, M., and Hannah, L.C. (1990). Cloning and characterization of the brittle-2 gene of maize. Maydica. **35,** 317-322.

Ball, S., Guan, H.-P., James, M., Myers, A., Keeling, P., Mouille, G., Buleon, A., colonna, P., and Preiss, J. (1996). From glycogen to amylopectin: A model for the biogenesis of the plant starch granule. Cell **86,** 349-352.

Bhattacharyya, M.K., Smith, A.M., Ellis, T.H.N., Hedley, C., and Martin, C. (1990). The wrinkled-seed character of pea described by mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme. Cell **60,** 115-122.

Bhave, M.R., Lawrence, S., Burton, C., and Hannah, L.C. (1990). Identification and molecular characterization of *shrunken-2* cDNA clones of maize. Plant Cell **2,** 581-588.

Boyer, C.D., and Fisher, M.B. (1984). Comparison of soluble starch synthases and branching enzymes from developing maize and teosinte seeds. Phytochemistry **23,** 733-737.

Boyer, C.D., and Preiss, J. (1978). Multiple forms of (1.4)-a-D-glucan, (1.4)-a-D-glucan-6-glucosyl transferase from developing Zea mays L. kernels. Carbohydrate Research **61,** 321-334.

Boyer, C.D., and Preiss, J. (1981). Evidince for independent genetic control of the multiple forms of maize endosperm branching enzymes and starch synthases. Plant Physiol. **67,** 1141-1145.

Brown, J.W.S. (1986). A catalogue of splice junction and putative branch point sequences from plant introns. Nucleic Acids Research **14,** 9549-9559.

Bruce, W.B., H., C.A., Klein, D., Fromm, M., and H., Q.P. (1989). Photoregulation of a phytochrome gene promoter from oat transferred into rice by particle bombardment. Proc. Natl. Acad. Sci. USA **86,** 9692-9696.

Bryce, W.H., and Nelson, O.E. (1979). Starch-synthesizing enzymes in the endosperm and pollen of maize. Plant Physiol. **63,** 312-317.

Bucher, P. (1990). Weigh matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequence. J. Mol. Biol. **212,** 563-578.

Burton, R.A., Bewley, J.D., Smith, A.M., Bhattacharyya, M.K., Tatge, H., Ring, S., Bull, V., Hamilton, W.D.O., and Martin, C. (1995). Starch branching enzymes belonging to distinct enzyme families are differentially expressed during pea embryo development. Plant J. **7,** 3-15.

Callis, J., Fromm, M., and Walbot, V. (1987). Introns increase gene expression in cultured maize cells. Genes & Development **1,** 1183-1200.

Cao, H., Sullivan, T.D., Boyer, C.D., and Shannon, J.C. (1995). *Bt1*, a structural gene for the major 39-44 kDa amyloplast membrane polypeptides. Phyiologia Plantarum **95,** 176-186.

Chen, B.P.C., Liang, G., Whelan, J., and Hai, T. (1994). ATF3 and ATF3ΔZip. J. Biol. Chem. **269,** 15819-26.

Chisholm, D. (1988). A convenient moderate-scale procedure for obtaining DNA from bacteriophage Lamda. Biotechniques **7,** 255-257.

Chodosh, L.A., Baldwin, A.S., Carthew, R.W., and Sharp, P.A. (1988). Human CCAAT-binding proteins have heterologous subunits.

Chourey, P.S., and Nelson, O.E. (1976). The enzymatic deficiency conditioned by the *shrunken-1* mutations in maize. Biochem. Genet. **14,** 1041-55.

Chourey, P.S., and Nelson, O.E. (1979). Interallelic complementation at the *sh* locus in maize at the enzyme level. Genetics **91,** 317-25.

Chourey, P.S., and Taliercio, E.W. (1994). Epistatic interaction and functional compensation between the two tissue- and cell-specific sucrose synthase genes in maize. Proc. Natl. Acad. Sci. USA **91,** 7917-7921.

Chourey, P.S., Tallercio, E.W., and Im, K.H. (1993). Sucrose phosphate synthase (SPS) in developing kernels of maize. Plant Physiol. **102 (Suppl.),** 6 (Abstr.).

Church, G.M., and Gilbert, W. (1984). Genomic sequencing. Proc. Natl. Acad. Sci. U.S.A. **81,** 1991-1995.

Clancy, M., Vasil, V., Hannah, L.C., and Vasil, I.K. (1994). Maize *Shrunken-1* intron and exon regions increase gene expression in maize protoplasts. Plant Science **98,** 151-161.

Dang, P.L., and Boyer, C.D. (1989). Comparison of soluble starch synthases and branching enzymes from leaves and kernels of normal and *Amylose-extende*r maize. Biochemical Genetics **27,** 521-532.

Denyer, K., Dunlap, F., Thorbjornsen, T., Keeling, P., and Smith, A.M. (1996). The major form of ADP-glucose pyrophosphorylase in maize endosperm is extra-plastidial. Plant Physiol. **112,** 779-785.

Dorn, A., Bollekens, J., Staub, A., Benoist, C., and Mathis, D. (1987). A multiplicity of CCAAT box-binding proteins. Cell **50,** 863-872.

Edwards, J., Green, J.H., and T, A.R. (1988). Activity of branching enzyme as a cardinal feature of the *ra* locus in *Pisum sativum.* Phytochemistry **27,** 1615-1620.

Entwistle, G., and ap Rees, T. (1990). Lack of fructose-1,6-bisphosphatase in a range of higher plants that store starch. Biochem. J. **271,** 467-472.

Fisher, D.K., Boyer, C.D., and Hannah, L.C. (1993). Starch branching enzyme II from maize endosperm. Plant Physiol. **102,** 1045-1046.

Fisher, D.K., Gao, M., Kim, K.-N., Boyer, C.D., and Guiltinan, M.J. (1996). Allelic analysis of the maize *amylose-extender* locus suggests that independent genes encode starch-branching enzymes IIa and IIb. Plant Physiol. **110,** 611-619.

Fisher, D.K., Kim, K.-N., Gao, M., Boyer, C.D., and Guiltinan, M.J. (1995). A cDNA encoding starch branching enzyme I from maize endosperm. Plant Physiol. **108,** 1313-1314.

Fisher, M.B., and Boyer, C.D. (1983). Immunological characterization of maize starch branching enzymes. Plant Physiol. **72,** 813-816.

Fu, H., Kim, S.Y., and Park, W.D. (1995). A potato *Sus*3 sucrose synthase gene contains a context-dependent 3' element and a leader intron with both positive and negative tissue-specific effects. Plant Cell **7,** 1395-1403.

Fu, H., Kim, S.Y., and Park, W.D. (1995b). High-level tuber expression and sucrose inducibility of a potato Sus4 sucorse synthase gene require 5' and 3' flanking sequences and the leader intron. Plant Cell **7,** 1387-1394.

Gao, M., Fisher, D.K., Kim, K.-N., Shannon, J.C., and Guiltinan, M.J. (1996). Evolutionary conservation and expression patterns of maize starch branching enzyme I and IIb genes suggests isoform specialization. Plant Mol. Biol. **30,** 1223-1232.

Gao, M., Fisher, D.K., Kim, K.-N., Shannon, J.C., and Guiltinan, M.J. (1997). Independent genetic control of maize starch-branching enzymes IIa and IIb. Plant Physiol. **114,** 69-78.

Garwood, D.L., shannon, J.C., and Creech, R.G. (1976). Starches of endosperms possessing different alleles at the *amylose-extender* locus in *Zea mays* L. Cereal Chem. **53**, 355-364.

Giroux, M.B., Smith-White, B., Gilmore, V., Hannah, L.C., and Preiss, J. (1995). The large subunit of the embryo isoform of ADP glucose pyrophosphorylase from maize. Plant Physiol. **108**, 1333-35.

Giroux, M.J., Boyer, C., Feix, G., and Hannah, C. (1994). Coordinated transcriptional regulation of storage product genes in the maize endosperm. Plant Physiol. **106**, 713-722.

Giroux, M.J., and Hannah, L.C. (1994). ADP-glucose pyrophosphorylase in shrunken2 and brittle2 mutants of maize. Mol. Gen. Genet. **243**, 400-408.

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

Grierson, C., Du, J.-S., Zabala, M.T., Beggs, K., Smith, C., Holdsworth, M., and 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.

Griffin, H.L., and Wu, Y.V. (1968). Isolation and characterization of the potato α-1,4-glucan 6-glucosyltransferase. Biochemistry **7**, 3063-3075.

Guan, H.P., and Preiss, J. (1993). differentiation of the properties of the branching isozymes from maize (*Zea mays*). Plant Physiol. **102**, 1269-1273.

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

Hamilton, D.A., Roy, M., Rueda, J., Sindhu, R.K., Sanford, J., and Mascarenhas, J.P. (1992). Dissection of a pollen-specific promoter from maize by transient transformation assays. Plant Mol. Biol. **18**, 211-218.

Hannah, L.C., and Nelson, O.E. (1975). Characterization of adenosine diphosphate glucose pyrophosphorylase from developing maize seeds. Plant Physiol. **55**, 297-302.

Hatzfeld, W.-D., and Stitt, M. (1990). A study of the rate of recycling of triose phosphates in heterotrophic *Chenopodium rubrum* cells, potato tubers, and maize endosperm. Planta **180**, 198-204.

Hedman, K.D., and Boyer, C.D. (1982). Gene dosage at the *amylose-extender* locus of maize: Effects on the levels of starch branching enzymes. Biochemical Genetics **20**, 483-491.

Hedman, K.D., and Boyer, C.D. (1983). Allelic studies of the *amylose-extender* locus of *Zea mays* L.: Levels of the starch branching enzymes. Biochemical Genetics **21,** 1217-1222.

Hodges, H.F., Creech, R.G., and Loerch, J.D. (1969). Biosynthesis of phytoglycogen in maize endosperm: The branching enzyme. Biochimica et Biophysica Acta **185,** 70-79.

James, M.G., Robertson, D.S., and Myers, A.M. (1995). Characterization of the maize gene sugary1, a determinant of starch composition in kernel. Plant Cell **7,** 417-429.

Jang, J.-C., Leon, P., Zhou, L., and Sheen, J. (1997). Hexokinase as a sugar sensor in higher plants. Plant Cell **9,** 5-19.

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

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

Junghans, H., and Metzlaff, M. (1990). A simple and rapid method for the preparation of total plant DNA. Biotechniques **8,** 176.

Kainuma, K. (1988). Structure and chemistry of the starch granule. In J. Preiss, eds. The biochemistry of plants **vol 14. Academic Press, San Diego,,** pp 141-180.

Katagiri, F., Lam, E., and Chua, N.-H. (1989). Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. Nature **304,** 727-730.

Kawasaki, T., Hayashida, N., Baba, T., Shinozaki, K., and Shimada, H. (1993b). The gene encoding a calcium-depedent protein kinase located near the sbe 1 gene encoding starch branching enzyme I is specifically expressed in developing rice seeds. Gene **129,** 183-189.

Kawasaki, T., Mizuno, K., Baba, T., and Shimada, H. (1993a). Molecular analysis of the gene encoding a rice starch branching enzyme. Mol. Gen. Genet. **237,** 10-16.

Kehoe, D.M., Degenhardt, J., Winicov, I., and Tobin, E.M. (1994). Two 10-bp regions are critical for phytochrome regulation of a Lemna gibba Lhcb gene promoter. Plant Cell **6,** 1123-1134.

Khoshnoodi, E.B., Rask, L., and Larsson, H. (1993). Characterization of the 97 and 103 kDa forms of starch branching enzyme from potato tubers. FEBS **332,** 132-138.

Kim, S.-R., Costa, M.A., and G., A. (1991). Sugar response element enhance wound response of potato proteinase inhibitor II promoter in transgenic tobacco. Plant Mol. Biol. **17**, 973-983.

Kleczkowski, L.A., Villand, P., Luthi, E., Olsen, O.-A., and Preiss, J. (1993). Insensitivity of barley endosperm ADP-glucose pyrophosphorylase to 3-phosphoglycerate and orthophosphate regulation. Plant Physiol. **101**, 179-186.

Kunkel, T.A., Roberts, J.D., and Zakour, R.A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**, 367-382.

Li, H.-M., Sullivan, T.D., and Keegstra, K. (1992). Information for targeting to the chloroplastic inner envelope membrane is contained in the mature region of the maize *Bt1*-protien. J. Biol. Chem. **267**, 18999-19004.

Liu, K.C., Boyer, C.D., and Shannon, J.C. (1992). Carbohydrate transfer into isolated maize amyloplasts (abstract No. 234). Plant Physiol. **99**, S-39.

Maas, C., Laufs, J., Grant, S., Korfhage, C., and Werr, W. (1991). The combination of a novel stimulatory element in the first exon of the maize *Shrunken-1* gene with the following intron 1 enhances reporter gene expression up to 1000-fold. Plant Mol. Biol. **16**, 199-207.

Martin, C., and Smith, A.M. (1995). Starch biosynthesis. Plant Cell **7**, 971-985.

Matters, G.L., and Boyer, C.D. (1981). Starch Synthases and Branching Enzymes from *Pisum satium*. Phytochemistry **20**, 1805-1809.

McCarty, D.R., Shaw, J.R., and Hannah, L.C. (1986). The cloning, genetic mapping and expression of the constitutive sucrose synthase locus of maize. Proc. Natl. Acad. Sci. USA **83**, 9099-9103.

McKnight, S., and Tjian, R. (1986). Transcriptional selectivity of viral genes in mammalian cells. Cell **46**, 759-805.

McRlroy, D., Zhang, W., Cao, J., and Wu, R. (1990). Isolation of an efficient actin promoter for use in rice transformation. Plant Cell **2**, 163-171.

Menken, A.E., Schindler, U., and Cashmore, A.R. (1995). The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZip proteins. TIBS **20**, 506-510.

Miller, M.E., and Chourey, P.S. (1992). The maize invertase-deficient miniature-1 seed mutation is associated wisth aberrant pedicel and endosperm development. Plant Cell **4**, 297-305.

Mizuno, K., Kawasaki, T., Shimada, H., Satoh, H., Kobayashi, E., Okumura, S., Arai, Y., and Baba, T. (1993). Alteration of the structural properties of starch components by the lack of an isoform of starch branching enzyme in rice seeds. J. Biol. Chem. **268**, 19084-19091

Mizuno, K., Kimura, K., Arai, Y., Kawasaki, T., Shimada, H., and Baba, T. (1992). Starch branching enzymes from immature rice seeds. J. Biochem. **112**, 643-651.

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

Moore, C.W., and Creech, R.G. (1972). Genetic fine structure analysis of the amylose-extender locus in Zea mays L. Genetics **70**, 611-619.

Muller-Rober, B.T., Kobmann, J., Hannah, L.C., Willmitzer, L., and Sonnewald, U. (1990). One of two different ADP-glucose pyrophosphorylase genes from potato responds strongly to elevated levels of sucrose. Mol. Gen. Genet. **224**, 136-146.

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

Nakamura, Y., Takeichi, T., Kawaguchi, K., and Yamanouchi, H. (1992a). Purification of two forms of starch branching enzyme (Q-enzyme) from developing rice endosperm. Physiologia Plantarum **84**, 329-335.

Nelson, O., and Pan, D. (1995). Starch synthesis in maize endosperms. Annu. Rev. Plant Physiol. Plant Mol. Biol. **46**, 475-96.

Ohta, S., Hattori, T., Morikami, A., and Nakamura, K. (1991). High-level expression of a sweet potato sporamin gene promoter:beta-glucuronidase (GUS) fusion gene in the stems of transgenic tobacco plants is conferred by multiple cell type-specific regulatory elements. Mol. Gen. Genet. **225**, 369-378.

Ozbun, J.L., Hawker, J.S., and Preiss, J. (1971). Adenosine diphosphoglucose-starch glucosyl transferases from developijng kernels of *waxy* maize. Plant Physiol. **78**, 765-69.

Pan, D., and Nelson, O.E. (1984). A debranching enzyme deficiency in endosperms of the *sugary-1* mutants of maize. Plant Physiol. **74**, 324-328.

Pien, F.-M., Boyer, C.D., and Shannon, J.C. (1993). Compartmentation of carbohydrate intermediates in the endosperm of starch-deficient maize genotypes (abstract No. 275). Plant Physiol. **102**, S52.

Porter, G.A., Knievel, D.P., and Shannon, J.C. (1985). Sugar efflux from maize (*Zea mays* L.) pedicel tissue. Plant Physiol. **77**, 524-531.

Pozueto-Romero, J., Ardila, F., and Akazawa, T. (1991a). ADP-glucose transport by the chloroplast adenylate translocator is linked to starch biosynthesis. Plant Physiol. **97**, 1565-1672.

Pozueto-Romero, J., Frehner, M., Viole, A.M., and Akazawa, T. (1991b). Direct transport of ADP-glucose by and adenylate tranlocator is linked to starch biosynthesis in amyloplasts. Proc. Natl. Acad. Sci. USA **88,** 5769-5773.

Preiss, J. (1991). Biology and Molecular Biology of Starch Synyhesis and its Regulation. Oxford Surveys of Plant Molecular & Cell Biology **7,** 59-114.

Prioul, J.-L., Jeannette, E., Reyss, A., Gregory, N., Giroux, L., Hannah, L.C., and Causse, M. (1994). Expression of ADP-glucose pyrophosphorylase in maize grain and source leaf during grain filling. Plant Physiol. **104,** 179-187.

Rogers, S.O., and Bendich, A.J. (1985). Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol. Biol. **5,** 69-76.

Salehuzzaman, S.N.I.M., Jacobsen, E., and Visser, R.G.F. (1994). Expression patterns of two starch biosynthetic genes in in vitro cultured cassava plants and their induction by sugars. Plant Science **98,** 53-62.

Salisbury, F.B. (1993). Gravitropism: changing ideas. Hortic. Rev. **15,** 233-278.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-termination inhibitors. Proc. Natl. Acad. Sci. USA **74,** 5463-5467.

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

Shannon, J.C. (1968). Carbon-14 distribution in carbohydrates of immature *Zea mays* kernels following $^{14}CO_2$ treatment of intact plants. Plant Physiol. **43,** 1215-20.

Shannon, J.C., and Garwood D.L. (1984) Genetics and physiology of starch development. In RL whistler, J.N. BeMiller and E.F. Paschall, eds, Starch Chemistry and Technology, vol Academic Press Inc., pp 25-86

Shannon, J.C., and Liu , J.W. (1977). A simplified medium for the growth of maize (Zea mays) endosperm tissue in suspension culture. Physiol. Plant **40,** 285-291.

Singh, B.K., and Preiss, J. (1985). Starch branching enzymes from maize. Plant Physiol. **79,** 34-40.

Smith, A.M. (1988). Major Differences in isofroms of Starch-Branching Enzyme between Developing Embryos of Round- and Wrinkled-Seeded Peas (*Pisum sativum* L.). Planta **175,** 270-279.

Stinard, P.S., Robertson, D.S., and Schnable, P.S. (1993). Genetic isolation, cloning, and analysis of *Mutator-* induced, dominant antimorph of the maize *amylose-extender*1 locus. Plant Cell **5,** 1555-1566.

Stitt, M., Huber, S., and Kerr, P. (1987). Control of photosynthetic sucrose formation. In : Hatch M., Boardman N. (eds) The biochemistry of plants **Academic Press, New York, pp 327-409,**

Sullivan, T.D., and Kaneko, Y. (1995). The maize brittle1 gene encodes amyloplast membrane polypeptides. Planta **196,** 477-484.

Sullivan, T.D., Strelow, L.I., Illingworth, C.A., Phillips, R.L., and Nelson, O.E. (1991). The maize *brittle-1* locus: molecular characterization based on DNA clones isolated using the *dSpm*-tagged *brittle-1-mutable* allele. Plant Cell **3,** 1337-1348.

Takeda, Y., Guan, H.-P., and Preiss, J. (1993). Branching of amylose by the branching isoenzymes of maize endosperm. Carbohydrate Research **240,** 253-263.

Takeda, Y., Shitaozono, T., and Hizukuri, S. (1988). Molecular structure of corn starch. Starch/Staerke **40,** 51-54.

Takeda, Y., Shitaozono, T., and Hizukuri, S. (1990). Structures of sub-fractions of corn amylose. Carbohydr. Res. **199,** 207-214.

Tobias, R.B., Boyer, C.D., and Shannon, J.C. (1992). Alterations in carbohydrate intermediates in the endosperm of starch-deficient maize (*Zea mays* L.) genotypes. Plant Physiol. **99,** 146-152.

Ulmasov, B., and Folk, W. (1995). Analysis of the role of 5' and 3' flanking sequence elements upon in vivo expression of the plant $\mathrm{tRNA}^{\mathrm{Trp}}$ genes. Plant Cell **7,** 1723-1734.

Vasil, V., Clancy, M., Ferl, R.J., Vasil, I.K., and Hannah, L.C. (1985). Increased gene expression by the first intron of *maize Shrunken-1* Locus in grass species. Plant Physiol. **91,** 1575-1579.

Vries, S.D., Hoge, H., and Bisseling, T. (1988). Isolation of total and polysomal RNA from plant tissues. Plant Mol. Biol. Manual **B6,** 1-13.

White, P.J. (1994) Properties of corn starch. In A.R. Hallauer. eds. Specialty Corns. vol CRC Press. Ann Arbor. pp 29-54

Xu, J., McCarty, D.R., and Koch, K. (1993). Maize invertase gene expression is carbohydrate responsive. Plant Physiol. **102 (supple.),** 44 (Abstr.).

Yamanouchi, H., and Nakamura, Y. (1992). Organ specificity of isoforms of starch branching enzyme (Q-Enzyme) in rice. Plant Cell Physiol. **33,** 985-991.

Zobel, H.F. (1988). Molecules to granules: a comprehensive starch review. Starch **40,** 44-50.

VITA

KYUNG-NAM KIM

BORN: July 18. 1965. Seoul. Korea.

EDUCATION:

8/1992-present	Ph. D. candidate in Intercollege Graduate Program in Plant Physiology at The Pennsylvania State University
3/1988-2/1990	Master degree in Genetic Engineering Program at Korea University. Seoul. Korea.
3/1984-2/1988	Bachelor degree in Department of Genetic Engineering at Korea University. Seoul. Korea.

RECENT PUBLICATIONS:

Gao. M., Fisher. D. K., Kim. K.-N., Shannon. J. C., and Guiltinan. M. J. (1997). Independent genetic control of maize starch-branching enzymes IIa and IIb. Plant Physiol. 114. 69-78.

Gao. M., Fisher. D. K., Kim. K.-N., Shannon. J. C., and Guiltinan. M. J. (1996). Evolutionary conservation and expression patterns of maize starch branching enzymes I and IIb genes suggests isoform specialization. Plant Mol. Biol. 30. 1223-1232.

Fisher. D. K., Kim. K.-N., Gao. M., Boyer. C. D., Guiltinan. M. J. (1995). A cDNA encoding starch branching enzyme I from maize endosperm. Plant Physiol. 108. 1313-1314.