Plant Molecular Biology **30**: 97–108, 1996.

Two closely related cDNAs encoding starch branching enzyme from *Arabidopsis thaliana*

Dane K. Fisher [1], Ming Gao [1], Kyung-Nam Kim [1], Charles D. Boyer [2] and Mark J. Guiltinan [1,*]
[1] *Department of Horticulture, Intercollegiate Programs in Plant Physiology and Genetics and 519 Wartik Laboratory, Biotechnology Institute, Pennsylvania State University, University Park, PA 16802 USA (*author for correspondence);* [2] *Department of Horticulture, Oregon State University, Corvallis, OR 97331-7304, USA*

Received 30 May 1995; accepted in revised form 2 October 1995

Key words: *Arabidopsis thaliana*, cDNA, gene expression, starch biosynthesis, starch branching enzyme

Abstract

Two starch branching enzyme (SBE) cDNAs were identified in an *Arabidopsis* seedling hypocotyl library using maize *Sbe1* and *Sbe2* cDNAs as probes. The two cDNAs have diverged 5′ and 3′ ends, but encode proteins which share 90% identity over an extensive region with 70% identity to maize SBE IIb [12]. Genomic Southern blots suggest that the two cDNAs are the products of single, independent genes, and that additional, more distantly related SBE genes may exist in the *Arabidopsis* genome. The two cDNAs hybridize to transcripts which show similar expression patterns in *Arabidopsis* vegetative and reproductive tissues, including seedlings, inflorescence rachis, mature leaves, and flowers. This is the first report of the identification of cDNAs encoding two closely related starch branching enzymes from the same species.

Introduction

ADP glucose pyrophosphorylase (EC 2.7.7.27), starch synthase (EC 2.4.1.2.1) and starch branching enzyme (EC 2.4.1.18; SBE) are three key enzymes involved in biosynthesis of the plant starch granule [22, 28]. Recent evidence also shows that starch debranching enzyme has an important role in starch biosynthesis [16]. Starch branching enzymes (α-1,4-glucan:α-1,4-glucan-6-glycosyl transferase) catalyze the formation of α-1,6 linkages, and therefore are vital to the synthesis of the more highly branched fraction of starch, amylopectin. The reaction involves hydrolysis of a α-1,4 linkage followed by the reattachment of the released 1,4-glucan chain to the remaining or to another 1,4-glucan chain by a α-1,6 bond. The creation of branch points in the growing starch molecule results in additional nonreducing ends for further synthesis of starch. The nature of the

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers U18817 (*SBE2.1*) and U22428 (*SBE2.2*).

branched chains in amylopectin is thought to be responsible for many physical properties of the starch granule.

Multiple forms of SBEs have been characterized in many species, including spinach leaf, pea embryo, potato tuber, maize endosperm and leaf (reviewed by [22, 27, 28]), and recently in rice seed and other vegetative tissues [24, 25, 36]. In maize endosperm, three isoforms, SBE I, IIa, and IIb, have been identified [4, 15, 35]. Based on biochemical differences between these isoforms it is thought that SBE I may transfer longer glucan chains than SBE IIa and IIb during starch biosynthesis [15]. In this way, each starch branching enzyme isoform may be responsible for a unique aspect of amylopectin biosynthesis and structure.

Recently, cDNAs for SBE isoforms have been isolated from tissues of several species including pea embryo [3, 6], potato tubers [20, 26], rice seed [18, 23–25], cassava [30] and maize endosperm [1, 12, 13]. The starch branching isoforms characterized and cDNAs cloned to date can be grouped into two distinct families, as initially defined by the Commission on Plant Gene Nomenclature [9], and as suggested by Burton *et al.* [6]. One family (*Sbe2*) includes maize *Sbe2*, pea *Sbe1* and rice *Sbe3* which share 67% to 70% deduced amino acid sequence identity, as determined using Clustal analysis. The second family (*Sbe1*) includes maize *Sbe1*, rice *Sbe1*, pea *Sbe2* and potato *Sbe* which share 65% to 70% identity. Between families, only 45% to 48% identity has been conserved.

Mutants affecting starch branching enzyme isoforms have been identified in pea and maize. Mendel's famous wrinkled seed mutation in pea, the *rugosus* locus, has been identified as a lesion in pea SBE I [2]. In maize, the homologous *amylose-extender* mutation affects SBE IIb activity in endosperm, resulting in an increase in the apparent amylose content of endosperm starch and a partially collapsed kernel phenotype [10]. Recent molecular cloning data confirms that the *ae* mutant contains a lesion in the structural gene for maize endosperm SBE IIb [12, 34]. No other mutants have been found in other SBE isoform genes in any species. However, in *Arabidopsis*, several mutants affecting starch metabolism have been identified. These include mutants affecting phosphoglucomutase activity [7], ADP glucose pyrophosphorylase [21], and those affecting regulation of starch degradation [8].

In this study, the isolation and characterization of two cDNAs encoding isoforms of SBE from *Arabidopsis thaliana* (L.) Heynh ecotype Columbia is reported. These cDNAs, which share highest similarity to the *Sbe2* family, show similar expression patterns in vegetative and reproductive tissues. This is the first report of the cloning of two distinct members of the *Sbe2* family from the same species.

Materials and methods

cDNA library and plant materials

Size-fractionated cDNA libraries constructed from poly(A)$^+$ RNA isolated from 3-day-old seedling hypocotyls of *Arabidopsis thaliana* ecotype Columbia were graciously provided by J. Ecker [33]. The four size selected libraries screened were constructed in λZap II (Stratagene, La Jolla, CA), and propagated according to the manufacturer's instructions and standard methods [31].

A. thaliana ecotype RLD plants were grown at 24 °C in a controlled growth chamber with 16 h/8 h day/night photoperiod. Plant tissues were harvested, immediately frozen in liquid N2, and stored at −70 °C for use in RNA and genomic DNA extractions.

Heterologous screening of Arabidopsis *seedling libraries*

The *Arabidopsis* seedling libraries were screened with ^{32}P-labeled cDNA clones for maize *Sbe* 2 [12] and maize *Sbe1* [13], serially. The probes were labeled with [^{32}P]-dCTP by the random priming method and purified with Sephadex G-50 spun columns [31]. For each primary hybridiza-

tion, 1.2×10^5 plaque forming units were lifted onto duplicate nylon membranes (Hybond N+, Amersham). Prehybridization for one hour at 42 °C in 0.5 M $NaHPO_4$ pH 7.2, 7% (w/v) SDS was followed by hybridization at 42 °C for 18 h in identical buffer. After hybridization, all membranes were washed at 42 °C with 5% (w/v) SDS, 0.04 M sodium phosphate, 0.001 M EDTA for 15 min, and then twice at 42 °C with 1% (w/v) SDS, 0.04 M sodium phosphate, 0.001 M EDTA for 15 min, and exposed to X-ray film (Kodak XAR) at −80 °C with two intensifying screens. After three successive rounds of screening to purity, Bluescript SK^- phagemids were rescued into *Escherichia coli* strain SOLR using the supplier's protocol (Stratagene).

DNA sequence analysis

Double-stranded plasmid DNA was sequenced by the dideoxy chain termination method [32]. Sequence data was analyzed with programs in the DNASTAR package on an Apple Macintosh Centris 650 computer (DNASTAR, Madison, WI). Multiple sequence alignments and percent identity were calculated by the Clustal method (gap length penalty 10) implemented by the Megalign program. Pairwise sequence alignment for percent similarity values were calculated by the Clustal method. The dendrogram was constructed using the Clustal method with a PAM250 residue weight table.

Gene-specific probes

Gene-specific hybridization probes from the 3′-untranslated regions of the *Arabidopsis* SBE cDNAs were prepared by subcloning of a 172 bp *Eco* RI/*Pst* I fragment of clone *SBE2.1* and a 125 bp *Eco* RI/*Bam* HI fragment of clone *SBE2.2* into similarly cut plasmid Bluescript SK^-. Probes for the coding regions were full-length inserts of *SBE2.1* and *SBE2.2* lacking these same 3′-untranslated regions.

Genomic Southern analysis

Genomic DNA was isolated from *Arabidopsis* tissue by the method of Dellaporta *et al.* [11]. After restriction enzyme digestion and electrophoresis, 10 μg of DNA was transferred onto a nylon membrane (Hybond N+, Amersham) by capillary transfer [31]. Filters were air-dried for 1 h and fixed by a 30 s exposure under a UV lamp. The filters were prehybridized for one hour at 65 °C in 0.5 M sodium phosphate pH 7.2, 7% (w/v) SDS, and 80 μg/ml salmon sperm DNA. Hybridization was performed at 65 °C in 0.5 M sodium phosphate pH 7.2, 7% (w/v) SDS, 40 μg/ml salmon sperm DNA, and 2.5×10^7 cpm of probe (specific activity of 10^9 cpm/μg) for 18 h. After hybridization, all membranes were washed twice at 65 °C with 5% (w/v) SDS, 0.04 M sodium phosphate, 0.001 M EDTA for 20 min each. Autoradiography was performed as described above.

RNA analysis

RNA was isolated from vegetative and reproductive tissues of *Arabidopsis* as described [29], and poly(A)$^+$ RNA was selected on oligo-dT cellulose columns according to the manufacturer (Gibco-BRL). RNAs were fractionated on a 2.2 M formaldehyde, 1.3% (w/v) agarose gel after denaturation at 65 °C for 15 min in 30 μl of loading buffer containing 50% (v/v) formamide, 2.2 M formaldehyde, 20 mM MOPS pH 7.0, 5 mM sodium acetate, 1 mM EDTA, 4% (v/v) glycerol, 0.8% (w/v) bromophenol blue and 1 μg ethidium bromide. The denatured RNA samples were chilled on ice for 3–5 min, loaded and allowed to rest in wells for 5–10 min before running. They were run at 2–3 V/cm for 4–5 h until bromophenol blue neared the bottom of the gel. Fractionated RNA was blotted onto nylon membranes (Hybond N, Amersham) by the manufacturer's instruction, and fixed by a 30 s exposure under a UV lamp. Prehybridization was for one hour at 65 °C in 0.5 M sodium phosphate pH 7.2, 7% (w/v) SDS, and 50 μg/ml salmon sperm DNA.

Hybridization was performed at 65 °C in 0.5 M sodium phosphate pH 7.2, 7% (w/v) SDS, and 50 μg/ml salmon sperm DNA for 18 h. After hybridization, all membranes were washed twice at 65 °C with 5% (w/v) SDS, 0.04 M sodium phosphate, 0.001 M EDTA for 20 min each.

Results

Isolation of SBE cDNA clones from Arabidopsis

About 1.2×10^5 plaque forming units from the *Arabidopsis* seedling hypocotyl cDNA library were screened at low stringency using both the maize *Sbe2* [12] and the maize *Sbe1* cDNA probes [13]. With the maize *Sbe2* and *Sbe1* probes, six and nine clones were identified, respectively. Restriction mapping and partial sequence analysis of these clones revealed that they all grouped into two distinct classes. The largest clones of each cDNA class were identified in plasmids pABE421 and pABE423, and their cDNA inserts were assigned the clone names *Arabidopsis SBE2.1* and *SBE2.2*, respectively. The maize *Sbe2* probe identified three clones of both *SBE2.1* and *SBE2.2*, while the maize *Sbe1* probe identified four clones of *SBE2.1* and five clones of *SBE2.2*.

Sequence analysis of putative starch branching enzyme cDNA clones

The sequences of *SBE2.1* and *SBE2.2* were completely determined on both strands and the data deposited in GenBank. The two cDNA clones, *SBE2.1* and *SBE2.2*, are 2668 and 2542 bp long, respectively. Results from northern analysis show that each clone detects a transcript expressed in *Arabidopsis* of ca. 2.6 kb, indicating that each clone is nearly full-length (see below). However, open reading frames for both cDNAs begin with the first bases of the 5′ ends, encoding 854 (*SBE2.1*) and 800 (*SBE2.2*) amino acids (Fig. 2), indicating that additional bases at the 5′ ends of the cDNAs must exist on the native mRNAs. The *SBE2.1* and *SBE2.2* cDNAs have 3′-untranslated regions (3′-UTRs) of 102 bp and 139 bp, respectively. While both 3′-UTRs are AT-rich, neither clone has a poly(A)$^+$ tail. At least three other clones isolated from this library also do not contain poly (A)$^+$ tails [33], possibly as a result of exonuclease cleavage during cDNA synthesis.

DNA sequence identity between the two clones is 70%, while deduced amino acid identity is 76% (Fig. 1). The central 2070 bp of the two cDNAs shows nearly 90% identity, while the 5′ region of *SBE2.1* is 100 bp larger than and has below 50% identity with *SBE2.2*. Alignment of the 5′ regions indicates that there are short stretches of sequence identity within this diverged region (Fig. 1). The 3′-untranslated regions of the cDNAs are very different, showing only 35% identity (Fig. 1).

The largest open reading frames of the cDNAs encode proteins which share high levels of identity beginning at residue 148 for *SBE2.1* and residue 112 for *SBE2.2* (Fig. 2). High sequence conservation of both with the maize *Sbe2* cDNA also begins at this point (Fig. 2). The mature proteins calculated from transit peptide cleavage (see below) are 807 amino acids for *SBE2.1* (91.7 kDa) and 760 amino acids for *SBE2.2* (87.5 kDa), consistent with the deduced sizes of SBEs from other species which range from 84 to 98 kDa. The deduced amino acid sequences of the *SBE2.1* and *SBE2.2* mature proteins contain the conserved catalytic residues for all amylolytic enzymes [17]. In addition, residues predicted to fold into eight β-strands and eight α-helices of the proposed $(\beta\text{-}\alpha)_8$ barrel domain of the amylolytic enzyme superfamily occur in positions similar to that of other plant starch branching enzymes (Fig. 2) [6].

As suggested by Burton *et al.* [6], a shared feature of several SBE isoforms in the *Sbe2* family is stretch of three consecutive proline residues just preceding the region of high similarity among isoforms. These proline residues typically reside at the C-terminal end of the N-terminal domain of the mature SBE proteins. This domain typically contains a number of single or grouped serine residues. These characteristics are consistent with the predicted amino-termini of both *SBE2.1* and

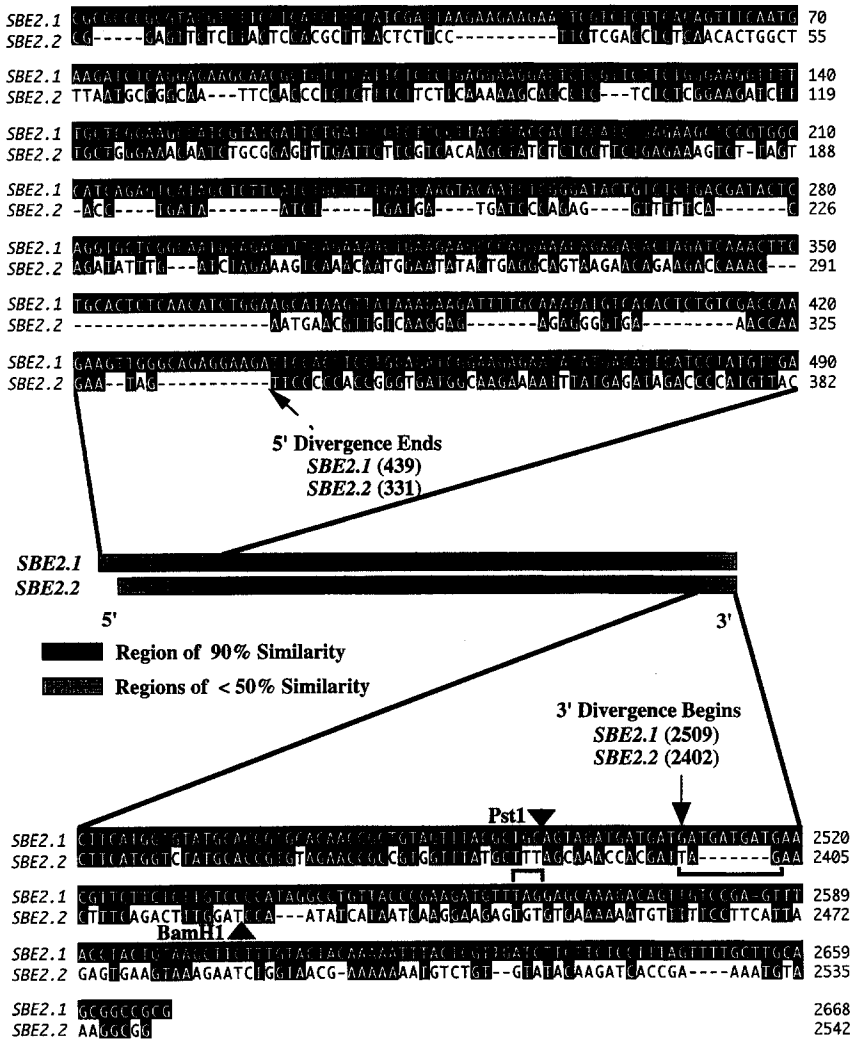

Fig. 1. Comparison of the 5′ and 3′ divergent DNA sequences of the *SBE2.1* and *SBE2.2* cDNAs. DNA sequence alignments were performed using the Clustal method (gap length penalty = 10) of the Megalign program (DNASTAR). Sequences which match the *SBE2.1* cDNA are shaded. Sequence positions where the 5′ region of low sequence similarity ends and the 3′ region of low similarity begins are indicated by an arrow and sequence numbers. *Pst* I (*SBE2.1*) and *Bam* HI (*SBE2.2*) restriction enzyme sites used to construct the gene-specific probes (GSP) are indicated by triangles. Translation stop codons are indicated in the 3′-regions of each cDNA by brackets.

SBE2.2 deduced proteins, which contain proline residues at positions 148–150 and 112–113, respectively, and a large number of serine residues prior to these positions (Fig. 2).

Amino acid residues 45–48 in *SBE2.1* and 39–42 in *SBE2.2* show characteristics conserved with known chloroplast transit peptide cleavage motifs [14] (Fig. 2). Although a translation initiation codon for either cDNA cannot be identified, the putative transit peptides of 42 and 48 amino acids are encoded for *SBE2.1* (1–42) and *SBE2.2* (1–48), respectively, based upon the positions of the predicted transit peptide cleavage sites. These regions have the characteristics found in other chloroplast transit peptides with high serine/threonine and lysine/arginine content, and very few acidic amino acid residues (Fig. 2).

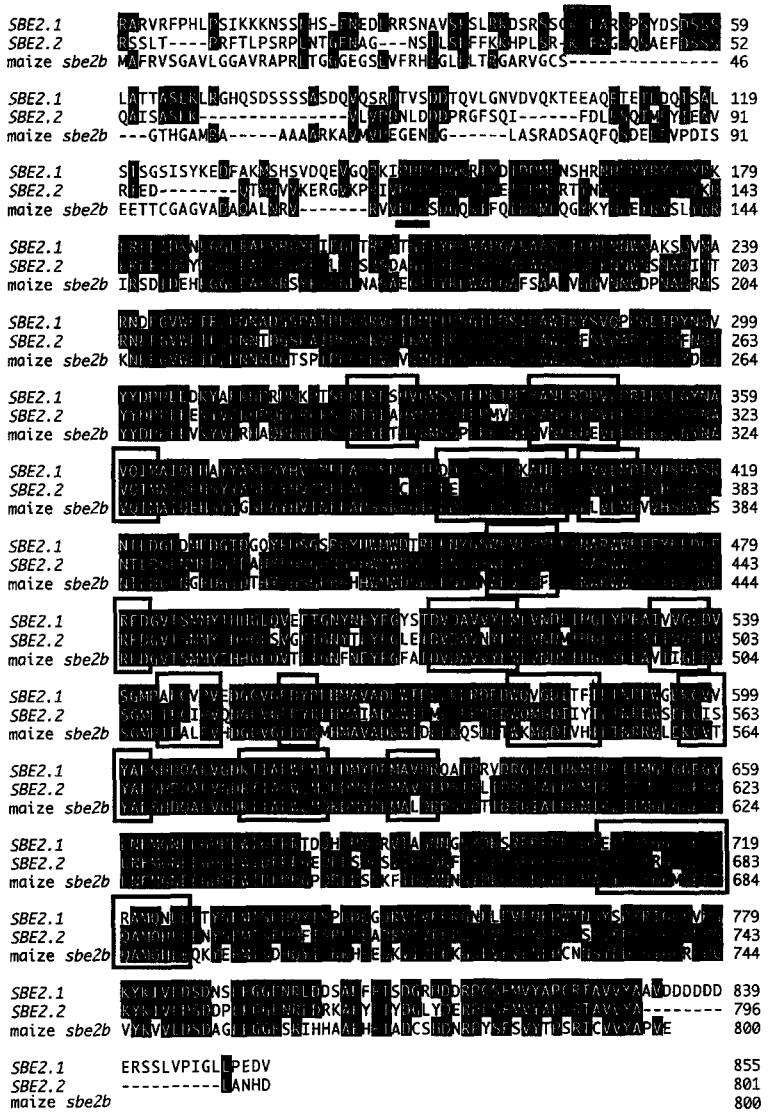

Fig. 2. Alignment of deduced amino acid residues between maize *Sbe2* and *Arabidopsis* isoforms. Amino acid residue positions are indicated on the right. Residues which match the consensus are shaded in black. Shaded in gray are the putative transit peptide cleavage sites for the *Arabidopsis* proteins [14] and the experimentally defined cleavage site for maize *Sbe2* [12]. The proline (P) residues found in the N-terminus of *Sbe2* family isoforms [6] are indicated with a thick black bar. The predicted positions of residues making up the $(\beta/\alpha)_8$ barrel domain of the amylolytic enzyme superfamily are indicated with boxes.

Phylogeny

Based on the features of the deduced *Arabidopsis* SBE proteins discussed above and on the results of sequence comparisons with other plant SBEs, the two cDNAs can be grouped into the *Sbe2* family of plant starch branching enzymes (Fig. 3 and [6, 9]). The sequence comparisons revealed that clone *SBE2.1* shares from 67% to 70% and *SBE2.2* shares from 71% to 75% identity to members of the *Sbe2* family. However, the two *Arabidopsis* cDNAs show only between 44% to 48% identity to the *Sbe1* family. Thus, based on sequence analysis, these *Arabidopsis* cDNAs ap-

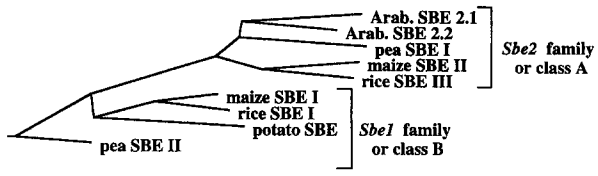

Fig. 3. Dendrogram showing a view of the phylogenetic relationships between nine plant SBE isoforms. The dendrogram was constructed using the deduced amino acid sequences analyzed by the Clustal method with the PAM250 residue weight table of the Megalign program (DNASTAR). SBE 2.1 and SBE 2.2 represent deduced amino acid residues from the *Arabidopsis SBE2.1* and *SBE2.2* cDNA clones, respectively. References for the sequences used in the analysis are indicated in the text.

pear to be members of the *Sbe2* family. However, enzymological analysis is necessary to confirm this at the functional level.

To demonstrate the sequence relationships between the plant SBE proteins, a phylogenetic tree was constructed. Figure 3 shows the predicted relationship of nine plant SBE genes as determined by their amino acid sequence alignments. This phylogenetic analysis confirms the classification of Burton *et al.* [6], and illustrates the general grouping of plant SBEs into the *Sbe1* and *Sbe2* families [9], with the two *Arabidopsis* genes described here clearly falling into the *Sbe2* family. Within the *Sbe2* family cluster, maize SBE II and rice SBE III are highly related as are the two *Arabidopsis* isoforms, with the pea SBE I protein positioned between these four isoforms. It is clear that both monocots and dicots contain genes in the *Sbe1* and *Sbe2* families, evidence of the ancient origins of these distinct SBE isoforms (> ca. 150 million years ago).

Genomic Southern analysis

Gene-specific probes for each cDNA clone were constructed from their respective 3′ divergent untranslated regions to insure that genes corresponding to *SBE2.1* and *SBE2.2* could be identified separately in hybridization analyses. We used restriction sites in the divergent regions to produce gene specific probes as described in Fig. 1. The fragments produced share only 35% DNA sequence identity, and are 172 bp and 125 bp in length for *SBE2.1* and *SBE2.2*, respectively. Southern blots of plasmid DNA from each clone, digested to release the inserts, were probed separately with each probe at high stringency conditions, the same conditions used for genomic Southern and northern analysis (see below). Figure 4A shows that the probes were specific, with only minor cross-hybridization observed. Minor hybridization seen to several species of higher-molecular-weight DNA resulted from the incomplete cleavage of the plasmids.

Genomic Southern analysis under stringent conditions using both the coding region and gene

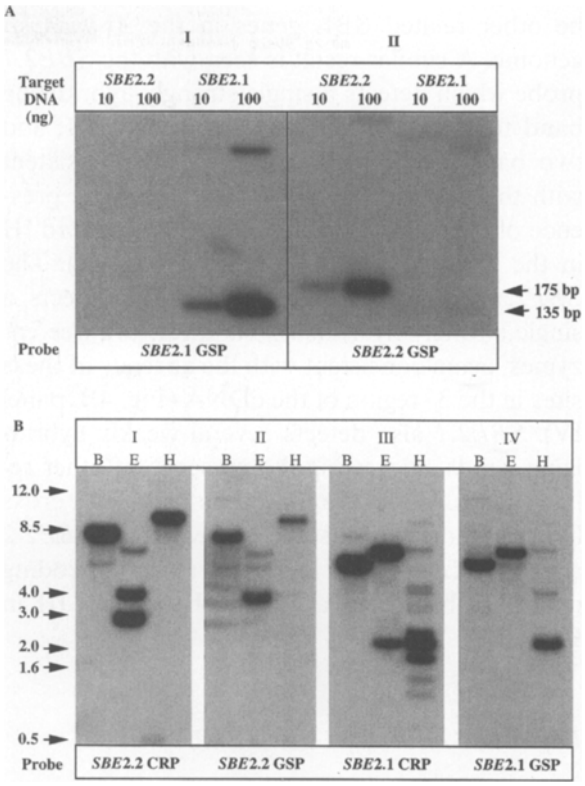

Fig. 4. Genomic Southern analysis with *Arabidopsis SBE2.1* and *SBE2.2* cDNA clones. A. Southern blot to verify the specificity of gene specific probes (GSP) using 10 and 100 ng of each cDNA clone (digested to release the inserts), probed with the *SBE2.1* GSP (panel I) and the *SBE2.2* GSP (panel II). B. Genomic Southern using *Arabidopsis* genomic DNA digested with *Bgl* II(B), *Eco* RI(E), and *Hin* cII(H), probed with *SBE2.2*- and *SBE2.1*-coding region probes (CRP) and GSPs. Sizes are indicated in kb.

specific probes revealed between one and two strongly hybridizing bands (Fig. 4B), suggesting that a single gene exists for each of these SBE isoforms in the *Arabidopsis* genome. The *SBE2.2*-coding region probe detects a single strongly hybridizing band in genomic DNA cut with *Bam* HI and *Hin*d III, and two bands for *Eco* RI, consistent with the absence of *Bam* HI and *Hin*d III sites and the presence of a single *Eco* RI site in the *SBE2.2* cDNA (Fig. 4B, panel I). The *SBE2.2* 3′ gene-specific probe detects a single strongly hybridizing band for all three enzymes, consistent with fact that the *Eco* RI site lies outside of the region spanned by the 3′ probe (Fig. 4B, panel II). Several weakly hybridizing bands are detected with the *SBE2.2* probes, indicating that there may be other related SBE genes in the *Arabidopsis* genome. A similar result is seen with the *SBE2.1* probe which detects a single strongly hybridizing band for genomic DNA cut with *Bam* HI, and two bands for *Eco* RI and *Hin*d III, consistent with the absence of *Bam* HI sites and the presence of single sites for both *Eco* RI and *Hin*d III in the *SBE2.1* cDNA (Fig. 4B, panel III). The *SBE2.1* 3′-untranslated region probe detects a single strongly hybridizing band for all three enzymes, again consistent with the absence of these sites in the 3′ region of the cDNA (Fig. 4B, panel IV). *SBE2.1* also detects several weakly hybridizing bands, suggesting the presence of other related genes. Under the conditions used, little cross hybridization was observed between the *SBE2.2* and *SBE2.1* genomic loci, even using the coding region probes (compare panel I to panel III in Fig. 4B).

Expression of SBE2.1 *and* SBE2.2

To compare the expression of the two SBE genes, blots of total RNA isolated from vegetative and reproductive tissues of *Arabidopsis* were probed using the *SBE2.1* cDNA, stripped and rehybridized with the *SBE2.2* cDNA. As shown in the genomic Southern blots, under the conditions used, the *SBE2.1* and *SBE2.2* probes (either coding region or untranslated region probes) do not cross hybridize significantly. Similarly, northern results for coding region probes were identical to those with 3′-untranslated region probes (not shown), therefore only results using coding region probes are presented in Fig. 5.

SBE2.1 appears to expressed in all tissues analyzed, including 12 and 20 DAE seedling, 35 DAE mature leaf, inflorescence rachis, inflorescence leaf, and flowers (Fig. 5, panel A and B). The highest level of expression appears in the inflorescence (leaf and rachis are similar; panel A, lanes 4–5), followed by mature leaf, and then flower. Seedlings at 12 and 20 DAE express the

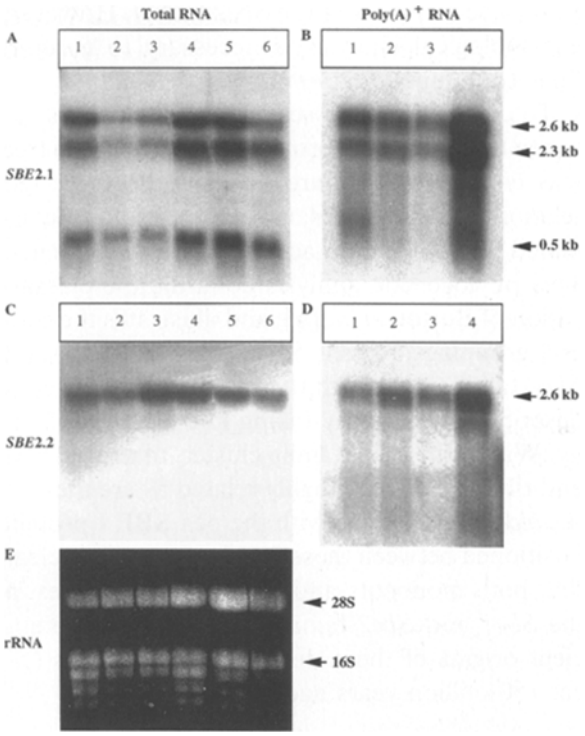

Fig. 5. RNA gel blot analysis of *SBE2.1* and *SBE2.2* expression in *Arabidopsis* tissues. A. Total RNA from (1) 35 DAE leaf, (2) 20 DAE seedling, (3) 12 DAE seedling, (4) inflorescence leaf, (5) inflorescence rachis and (6) flower were blotted and probed with the *SBE2.1* cDNA coding region probe (CRP) under high-stringency conditions. B. Poly(A)$^+$ RNA from (7) 35 DAE leaf, (8) 20 DAE seedling, (9) 12 DAE seedling and (10) inflorescence rachis: probed with the *SBE2.1* cDNA. Total RNA blot in A was stripped and reprobed with the *SBE2.2* cDNA CRP. D. Poly(A)$^+$ blot in B was stripped and reprobed with the *SBE2.2* CRP. Sizes are indicated in kb. E. Relative amount of ribosomal RNA in total RNA loadings.

lowest levels of the *SBE2.1* transcript of the tissues tested. The *SBE2.1* probe detects a transcript of 2.6 kb, but also two additional bands of size 2.3 kb and 0.5 kb, in apparently near-stoichiometric amounts. This transcript pattern is interesting in that it may represent some form of transcript processing or degradation. These bands do not result from non-specific mRNA degradation, as single bands are detected by the *SBE2.2* probe on the same blots (see below). To investigate these multiple transcripts further, we isolated poly(A)$^+$ RNA from several of the same tissues (Fig. 5, panel B). The *SBE2.1* transcript shows a similar pattern of accumulation in poly(A)$^+$ RNA as in total RNA, very high in inflorescence rachis (Fig. 5, panel B, lane 4) and mature leaf (panel B, lane 1), with lower levels of expression in 12 and 20 DAE seedlings. The 2.6 and 2.3 kb transcripts are still detected in poly(A)$^+$ RNA, while the 0.5 kb transcript is absent for all tissues analyzed (a weak 0.5 kb signal in lane 1 most likely results from inefficient poly(A) selection). This indicates that the 0.5 kb product does not contain a poly(A)$^+$ tail, and could be a relatively stable product of RNA splicing or degradation. The 2.3 and 0.5 kb products are not simply the result of a single cleavage of the full-length message, as the gene specific probe hybridizes to both fragments and the presence of a poly(A)$^+$ tail on the larger fragment precludes the possibility that the cleavage site lies within the region spanned by the probe. Alternatively, the 0.5 kb band could represent a high-abundance non-polyadenylated RNA which cross hybridizes to this probe.

In Fig. 5, panels C and D show the expression patterns of the *Arabidopsis SBE2.2* gene. This gene recognizes a single 2.6 kb transcript on identical blots previously probed by the *SBE2.1* gene. *SBE2.2* appears to be expressed in all tissues analyzed, with only a slightly different pattern than *SBE2.1*. Highest expression in total RNA is seen in 12 DAE seedlings and inflorescence leaves, followed by inflorescence rachis, 35 DAE leaf, and flower (Fig. 5, panel C). The poly(A)$^+$ RNA blot shown in panel D probed by *SBE2.2* shows a 2.6 kb single transcript with expression levels in each tissue similar to the results with total RNA. Panel E in Fig. 5 shows an image of the gel stained with ethidium bromide just prior to capillary transfer, demonstrating that relatively equal amounts of total RNA (30 μg) were loaded for each tissue.

Discussion

In this paper we describe the characterization of two very closely related starch branching enzyme cDNAs from *Arabidopsis* which show DNA and deduced amino acid sequence similarity to the *Sbe2* family which includes maize *Sbe2*, pea *Sbe1*, and rice *Sbe3*. The classification of plant starch branching isoforms into the distinct *Sbe1* and *Sbe2* families was initially based largely on enzymological properties of the maize isoforms [9]. Biochemical characteristics of other plant SBE isoforms suggests that they also fall into two major families [6, 25, 36]. More recently, molecular cloning has revealed that isoforms from rice [18, 23] and pea [6] can also be classified into the *Sbe1* and *Sbe2* families based upon DNA and deduced amino acid sequences of their cDNAs. Burton *et al.* [6] suggest that members of the *Sbe2* family share primary sequence similarities and a common structural feature of a flexible N-terminal domain just prior to three conserved proline residues. Our two *Arabidopsis* cDNAs share these features, suggesting that they encode proteins with evolutionary and functional relatedness to the *Sbe2* family.

Biochemical evidence for multiple *Sbe2* isoforms have been reported in maize [5] and rice [36], but to our knowledge, this is the first report of the cloning of two distinct members of the *Sbe2* family within the same species. Several possibilities must be addressed which could explain this finding. The first is that the two cDNAs are the result of heterozygosity at a single *Sbe2* locus. It is known that the Columbia ecotype originated from a heterozygous population of seeds [19]. If *Sbe2.1* and *Sbe2.2* are allelic, then the extensive regions of divergence found in the 5′ and 3′ regions have maintained their uniqueness in spite of

gene conversion and recombination over many generations of self-crossing. In this case, RFLPs must be invoked for each of the enzymes tested in our Southern analysis, since the bands hybridizing are different in each case for the two probes. This would be an exceptionally high level of polymorphism at this locus. A second possibility raised is that one of the two cDNAs actually encodes a SBE I-like isoform. Since both pea and maize have both *Sbe1* and *Sbe2* genes, it is clear that they have diverged prior to the monocot-dicot evolutionary split. Thus, it would be predicted that there is an *Sbe1*-like gene in *Arabidopsis*. Our results indicate that the two *Arabidopsis* clones contain regions of sequence homology to both *Sbe2* and *Sbe1* families of starch branching enzymes. This is consistent with previous sequence alignments of plant starch branching enzyme cDNAs, which indicate regions of sequence similarity among all SBE isoforms [6]. Thus, the low-stringency conditions used in our library screening allowed the detection of these clones with both maize *Sbe1* and *Sbe2* cDNA probes. However, both *SBE2.1* and *SBE2.2* are much more similar to the *Sbe2* family and both encode the N-terminal proline repeat and flexible domains, making it unlikely that either encodes an SBE-I-like protein. Our failure to recover a cDNA with high similarity to SBE I could be due to its low representation in the library we screened, or to the possibility that it does not exist in *Arabidopsis*. A final possibility is that the two cDNAs are the products of separate genes at independent loci. This scenario fits best with the genomic Southern data that show completely different patterns for the two probes, and with the large regions of high sequence divergence we observed. Even though isolation of multiple genes for any of the SBE isoforms has not been reported, two independent loci for maize *Sbe1* have been genetically mapped, and genomic Southern data also indicates multiple *Sbe2* genes in the maize genome (M. Guiltinan, unpublished and University of Missouri-Columbia 1995 maize RFLP map). Thus, it seems most likely that the two cDNAs represent the products of two independent genes.

Our genomic Southern blots suggest that additional genes closely related to clones *SBE2.1* and *SBE2.2* in *Arabidopsis* do not exist, however weakly hybridizing fragments suggest that other more distantly related SBE isoforms may exist. One divergent *Arabidopsis* cDNA has been isolated by Kawasaki *et al.* [18], who reported a partial deduced amino acid translation from a 0.6 kb PCR amplified genomic fragment from *Arabidopsis* that showed sequence similarity to the *E. coli* branching enzyme. This cDNA fragment showed limited similarity to plant SBE genes and was not characterized further.

Analysis of the expression of SBE genes in plant tissues other than specialized starch storage organs has been limited to date. In this paper, we describe two SBE cDNAs from *Arabidopsis* which are expressed in vegetative and reproductive tissues. *Arabidopsis SBE2.1* is expressed particularly highly in inflorescence rachis and leaf tissues, while *SBE2.2* is expressed similarly in all vegetative and reproductive tissues analyzed (whole seedling, leaf, inflorescence rachis and leaf, and flower). The differences between the expression patterns of the genes corresponding to these two cDNAs are very slight. In general these two SBE isoforms probably play a role in the synthesis of transitory starch used for energy demands in the growth of these tissues. In particular, *SBE2.1* may be highly expressed in tissues such as the inflorescence rachis and leaf for a build-up of transitory starch which may be used to shuttle energy at a later time to developing reproductive organs and the seed. It remains to be determined if these genes are also expressed during seed and/or pollen development in *Arabidopsis*.

Evidence presented in this paper describes two SBE cDNAs which encode proteins highly similar to members of the *Sbe2* family, and which are expressed at the mRNA level within the same tissues. Evidence explaining how these two isoforms interact with other SBE isoforms in starch biosynthesis within *Arabidopsis* vegetative tissues awaits further analysis. Enzymological analysis will be important in defining their biochemical properties. As more knowledge of the properties and genetic control of plant SBE isoforms is

gained, the biosynthesis of starch in both storage and transitory tissues will be better understood. It is intriguing to postulate how multiple isoforms of starch branching enzyme may interact with other enzymes (starch synthase, ADP glucose pyrophosphorylase and starch debranching enzyme) to synthesize the amylopectin fraction of the plant starch granule.

Acknowledgements

This work was supported by a grant from Pioneer Hi-Bred to M.J.G., C.D.B. and Jack Shannon and by the Penn State Agricultural Experiment Station (project 3303). The authors thank J. Ecker (University of Pennsylvania) for providing the cDNA libraries. We also appreciate the assistance of Jack C. Shannon and Richard N. Arteca in reviewing the manuscript, and Lech Michalczuk in contributing to this project. The technical support of Chris Dervinis and Jeri Searfoss is appreciated.

References

1. Baba T, Kimura K, Mizuno K, Etoh H, Ishida Y, Shida O, Arai Y: Sequence conservation of the catalytic regions of amylolytic enzymes in maize branching enzyme-I. Biochem Biophys Res Comm 181: 87–94 (1991).
2. Bhattacharyya MK, Martin C, Smith A: The importance of starch biosynthesis in the wrinkled seed shape character of peas studied by Mendel. Plant Mol Biol 22: 525–531 (1993).
3. Bhattacharyya MK, Smith AM, Ellis THN, Hedley C, Martin C: 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 (1990).
4. Boyer CD, Preiss J: Multiple forms of (1,4)-α-D-glucan-6-glucosyl transferase from developing *Zea mays* L. kernels. Carbohydr Res 61: 321–334 (1978).
5. Boyer CD, Preiss J: Multiple forms of branching enzyme of maize: evidence for independent genetic control. Biochem Biophys Res Commun 80: 169–175 (1978).
6. Burton RA, Bewley JD, Smith AM, Bhattacharyya MK, Tatge H, Ring S, Bull V, Hamilton WDO, Martin C: Starch branching enzymes belonging to distinct enzyme families are differentially expressed during pea embryo development. Plant J 7: 3–15 (1995).
7. Caspar T, Huber SC, Somerville C: Alterations in growth, photosynthesis, and respiration in a starchless mutant of *Arabidopsis thaliana* (L.) Heynh deficient in chloroplast phosphoglucomutase activity. Plant Physiol 79: 1–7 (1986).
8. Caspar T, Lin TP, Kakefuda G, Benbow L, Preiss J, Somerville C: Mutants of *Arabidopsis* with altered regulation of starch degradation. Plant Physiol 95: 1181–1188 (1991).
9. Commission on Plant Gene Nomenclature: Nomenclature of sequenced plant genes. Plant Mol Biol Rep 12: S1–S109 (1994).
10. Deatherage WL, MacMasters MM, Vineyard ML, Bear RP: A note of starch of high amylose content from corn with high starch content. Cereal Chem 31: 50 (1954).
11. Dellaporta SL, Wood J, Hicks JB: Maize DNA miniprep. In: Molecular Biology of Plants: A Laboratory Course Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1985).
12. Fisher DK, Boyer CD, Hannah LC: Starch branching enzyme II from maize endosperm. Plant Physiol 102: 1045–1046 (1993).
13. Fisher DK, Kim K-N, Gao M, Boyer CD, Guiltinan M: A cDNA encoding starch branching enzyme I from maize endosperm. Plant Physiol 108: 1313–1314 (1995).
14. Gavel Y, van Heijne G: A conserved cleavage-site motif in chloroplast transit peptides. FEBS Lett 261: 455–458 (1990).
15. Guan HP, Press J: Differentiation of the properties of the branching isozymes from maize (*Zea mays*). Plant Physiol 102: 1269–1273 (1993).
16. James MG, Robertson DS, Myers AM: Characterization of the maize gene *sugaryl*, a determinant of starch composition in kernels. Plant Cell 7: 417–429 (1995).
17. Jespersen HM, MacGregor EA, Henrissat B, Sierks MR, Svensson B: Starch- and glycogen-debranching and branching enzymes: prediction of structural features of the catalytic (beta/alpha)8-barrel domain and evolutionary relationship to other amylolytic enzymes. J Pro Chem 12: 791–805 (1993).
18. Kawasaki T, Mizuno K, Baba T, Shimada H: Molecular analysis of the gene encoding a rice starch branching enzyme. Mol Gen Genet 237: 10–16 (1993).
19. Koncz C, Chua N-H, Schell J (eds) Methods in Arabidopsis Research. World Scientific, River Edge, NJ (1992).
20. Kossmann J, Visser RGF, Muller-Rober B, Willmitzer L, Sonnewald U: Cloning and expression analysis of a potato cDNA that encodes branching enzyme:evidence for co-expression of starch biosynthetic genes. Mol Gen Genet 230: 39–44 (1991).
21. Lin TP, Caspar T, Somerville C, Preiss J: Isolation and characterization of a starchless mutant of *Arabidopsis thaliana* (L.) Heynh, lacking ADPglucose pyrophosphorylase activity. Plant Physiol 86: 1131–1135 (1988).
22. Martin C, Smith AM: Starch biosynthesis. Plant Cell 7: 971–985 (1995).

23. Mizuno K, Kawasaki T, Shimada H, Satoh H, Kobayashi E, Okumura S, Arai Y, Baba T: 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: (1993).
24. Mizuno K, Kimura K, Arai Y, Kawasaki T, Shimada H, Baba T: Starch branching enzymes from immature rice seeds. J Biochem 112: 643–651 (1992).
25. Nakamura Y, Takeichi T, Kawaguchi K, Yamanouchi H: Purification of two forms of starch branching enzyme (Q-enzyme) from developing rice endosperm. Physiol Plant 84: 329–335 (1992).
26. Poulsen P, Kreiberg JD: Starch branching enzyme cDNA from *Solanum tuberosum*. Plant Physiol 102: 1053–1054 (1993).
27. Preiss J: Biosynthesis of starch and its regulation. In: Preiss J (ed) The Biochemistry of Plants, pp. 181–249. Academic Press, San Diego (1988).
28. Preiss J: Biology and molecular biology of starch synthesis and its regulation. Oxford Surv Plant Mol Cell Biol 7: 59–114 (1991).
29. Rochester DE, Winer JA, Shah DM: The structure and expression of maize genes encoding the major heat shock protein, hsp70. EMBO 5: 451–458 (1986).
30. Salehuzzaman SNIM, Jacobsen E, Visser RGF: Cloning, partial sequencing and expression of a cDNA coding for branching enzyme in cassava. Plant Mol Biol 20: 809–819 (1992).
31. Sambrook J, Fritsch EF, Maniatis T: Molecular cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).
32. Sanger F, Nicklen S, Coulsen AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
33. Schindler U, Menkens AE, Beckmann H, Ecker JR, Cashmore AR: Heterodimerization between light-regulated and ubiquitously expressed *Arabidopsis* GBF bZIP proteins. EMBO J 11: 1261–1273 (1992).
34. Stinard PS, Robertson DS, Schnable PS: Genetic isolation, cloning, and analysis of *Mutator*-induced, dominant antimorph of the maize *amylose-extender1* locus. Plant Cell 5: 1555–1566 (1993).
35. Takeda Y, Guan HP, Preiss J: Branching of amylose by the branching isoenzymes of maize endosperm. Carbohydr Res 240: 253–263 (1993).
36. Yamanouchi H, Nakamura Y: Organ specificity of isoforms of starch branching enzyme (Q-enzyme) in rice. Plant Cell Physiol 33: 985–991 (1992).