

BIOCHEMICAL STUDIES ON STARCH  
BRANCHING ENZYME IN  
DEVELOPING RICE SEEDS

1995

Kouichi Mizuno

DA

1481 (HG)

1995

BIOCHEMICAL STUDIES ON STARCH  
BRANCHING ENZYME IN  
DEVELOPING RICE SEEDS

Division of Applied Biochemistry  
Doctoral Program in Agricultural Sciences  
University of Tsukuba

Kouichi Mizuno



96302255

## ABBREVIATIONS

BE	branching enzyme
BSA	bovine serum albumin
b, kb	base, kilobases
bp, kbp	base pair, kilobase pair
CBB	Coomassie brilliant blue
cDNA	complementary DNA
CGTase	cyclodextrin glucanotransferase
Da, kDa	dalton, kilodalton
DEAE-cellulose	diethylaminoethyl-cellulose
EDTA	ethylenediaminetetraacetic acid
IPTG	isopropyl $\beta$ -D-thiogalactoside
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
SDS	sodium dodecyl sulfate
SSC	standard saline citrate
SSPE	standard saline phosphate-EDTA
TAA	Taka-amylase A ( <i>Aspergillus oryzae</i> $\alpha$ -amylase)
Tris	tris (hydroxymethyl) aminomethane

# CONTENTS

	PAGES
General Introduction	2
Chapter 1 <i>Identification of Multiple Forms of Branching Enzyme in Developing Rice Seeds, and Molecular Characterization of a Branching Enzyme Isoform, RBE1</i>	
Introduction	6
Materials and Methods	7
Results	12
Discussion	30
Chapter 2 <i>Characterization of an Isoform of Rice Branching Enzyme, RBE3, and Its Physiological Role in Starch Synthesis</i>	
Introduction	35
Materials and Methods	38
Results	42
Discussion	56
Chapter 3 <i>Molecular Characterization of Recombinant Branching Enzymes</i>	
Introduction	62
Materials and Methods	63
Results and Discussion	72
General Conclusion	95
Acknowledgment	98
References	99

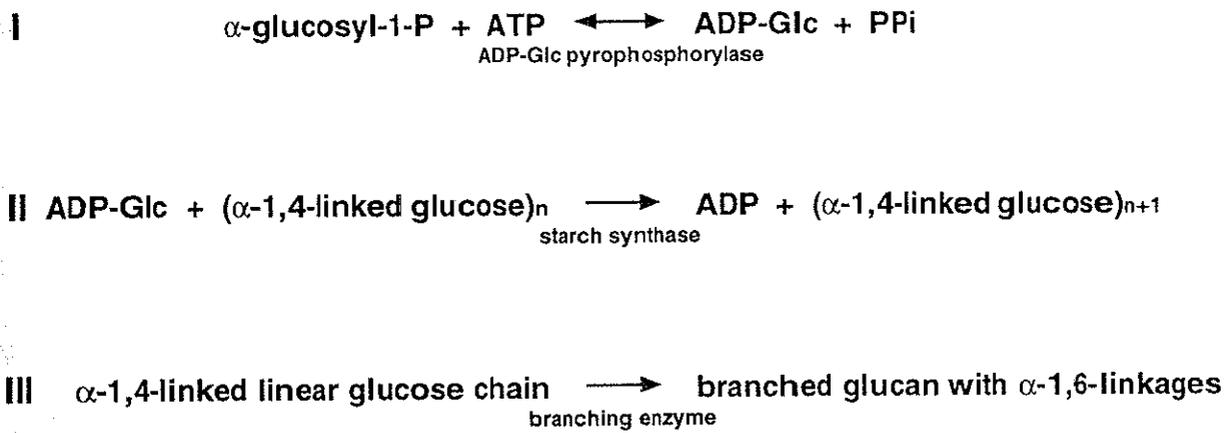
## General Introduction

In the storage organs of higher plants, starch is synthesized and accumulated in the organelles, amyloplast. Starch is the major reserve carbohydrate of most higher plants and is found as a water-insoluble granule which can vary in the size and shape depending on its plant origin and on the maturity or stage of development. The granules usually contain two different polysaccharides, amylose and amylopectin. Amylose is an essentially linear polymer with  $\alpha$ -1,4-glucosidic bonds. In contrast, amylopectin is a highly branched molecule containing  $\alpha$ -1,4- and  $\alpha$ -1,6-linkages.

An ADP-glucose (ADP-Glc) pathway consisting of the enzymatic reactions catalyzed by ADP-Glc pyrophosphorylase (EC 2.7.7.27; Fig. 1, reaction 1), starch synthase (EC 2.4.1.21; Fig. 1, reaction 2), and starch branching enzyme (EC 2.4.1.18; Fig. 1, reaction 3), is the predominant pathway, if not the sole pathway, for starch synthesis<sup>1,2</sup>. However, the synthetic mechanisms of amylose and amylopectin are poorly understood.

Starch branching enzyme acts on the already synthesized and/or elongating amylose chains to form  $\alpha$ -1,6-linked branch points. Thus, this enzyme is a key enzyme for amylopectin synthesis. Multiple forms of branching enzyme have been found in various plant tissues, e.g. spinach leaf<sup>3</sup>), maize leaf<sup>4</sup>), maize endosperm<sup>5</sup>), developing seeds of pea<sup>6</sup>), rice<sup>7,8</sup>), and germinating castor bean endosperm<sup>9</sup>). In the starchy endosperm of maize, three different branching enzymes (BE-I, BE-IIa, and BE-IIb) have been identified<sup>5,10,11</sup>). Recently, we have cloned the BE-I cDNA from maize<sup>12</sup>), and found that the deduced amino acid sequence of maize BE-I is highly divergent from that of the bacterial branching enzyme<sup>13</sup>). To elucidate starch synthesis in rice storage tissues, it is essential to characterize each of several forms of branching enzyme

and establish their primary structures. In Chapters 1 and 2, I report the identification of branching enzyme isoforms from developing rice seeds<sup>14)</sup> and isolation of their cDNA clones<sup>14,15)</sup>. Moreover, I have examined the structure/function relationships of two isoforms of rice branching enzyme, RBE1 and RBE3, in Chapter 3.



**Fig. 1.** Biosynthetic reactions in starch synthesis.

# Chapter 1

Identification of Multiple Forms of Branching  
Enzyme in Developing Rice Seeds, and Molecular  
Characterization of a Branching Enzyme Isoform,  
RBE1

## Introduction

Multiple forms of branching enzyme have been found in various plant tissues<sup>1,2)</sup>. There are at least two forms of branching enzyme, BE-I and BE-II, in maize kernels. These two forms are distinguishable from each other in the molecular size (82- and 80 kDa for BE-I and BE-II, respectively), amino acid composition, and peptide mapping<sup>5,11,16,17)</sup>. We have also identified the cDNA clone encoding maize BE-I from a maize kernel cDNA library<sup>12)</sup>. The deduced amino acid sequence reveals that maize BE-I shares a relatively limited sequence identity with bacterial branching enzyme. In rice seeds, two or three forms of branching enzyme have been reported to be present<sup>7)</sup>. However, the role of each of the isoforms in starch synthesis is still unclear. Also, little is known of the structure/function relationship for this enzyme. In order to elucidate the mechanism of starch synthesis in rice endosperm, we have purified and characterized each isoform of branching enzyme from developing seeds. Moreover, we have characterized the cDNA clones coding for an isoform of rice branching enzyme, RBE1, corresponding to maize BE-I.

## Materials and Methods

*Materials* --- Radioisotopes, [ $\alpha$ - $^{32}$ P]dCTP (3,000 Ci/mmol) and  $\alpha$ -D- [ $^{14}$ C]glucose 1-phosphate (335 mCi/mmol), were purchased from Bresatec and Amersham, respectively. Nitrocellulose and nylon membranes were purchased from Advantec (Tokyo) and Amersham (Hybond-N), respectively. All other reagents were of the highest purity available. Rice plants, *Oryza sativa* L. cv. Nipponbare, were field-grown in 1991 at the Life Science Laboratory of Mitsui Toatsu Chemicals, Mobarra, Chiba. The seeds were obtained at various stages after flowering, immediately frozen in liquid nitrogen, and stored at -80 °C until used.

*Purification of Rice Branching Enzyme* --- Frozen rice seeds (200 g, 10-15 days after flowering) were ground in 50 mM Tris/HCl, pH 8.5, containing 5 mM EDTA and 5 mM 2-mercaptoethanol (500 ml) in a mortar and pestle, and homogenized in an electric juicer. The homogenate was filtered through two layers of gauze and centrifuged at 10,000x g for 20 min. To the supernatant, solid ammonium sulfate was added to give 0-70 % saturation, and the mixture was stirred overnight. The pellet was collected by centrifugation at 18,000x g, dissolved in a minimum volume of a solution consisting of 50 mM Tris/HCl, pH 7.5, 5 mM EDTA, and 5 mM 2-mercaptoethanol (buffer A), and dialyzed thoroughly against buffer A. The dialyzed solution was applied to a DEAE-cellulose (Whatman DE-52) column previously equilibrated with buffer A. Proteins were eluted with a linear gradient of 0-0.3 M KCl (2 liters). Fractions (20 ml) were collected at a flow rate of 35 ml/h and assayed for branching enzyme activity. To the pooled fractions containing the enzyme activity, solid ammonium sulfate was added to give a 70 % saturation. The mixture was stirred for 6 h and

centrifuged at 10,000x g for 10 min. The precipitate was dissolved in 50 mM Tris/HCl, pH 7.5, containing 5 mM EDTA and 0.2 M KCl (buffer B), and then applied to a column of Toyopearl HW-55F previously equilibrated with buffer B. Proteins were eluted from the column with buffer B at a flow rate of 10 ml/h. Fractions (3 ml) were collected and assayed for the enzyme activity. All purification procedures were carried out at 0-4°C.

*Assay for Enzyme Activity* --- Branching enzyme activity was measured in the direction of stimulation of  $\alpha$ -glucan synthesis from  $\alpha$ -D-glucose 1-phosphate catalyzed by rabbit-muscle phosphorylase *a*, as described by Boyer and Preiss<sup>5)</sup>, with a minor modification. The reaction was conducted in a mixture (0.1 ml) containing 0.1 M sodium citrate, pH 7.0, 50 mM  $\alpha$ -D-[<sup>14</sup>C]glucose 1-phosphate (usually 50,000 cpm), 1 mM AMP, 10 mg of rabbit-muscle phosphorylase *a* (2x crystallized and lyophilized powder, Sigma), and an appropriate amount of enzyme. After incubation at 30°C for 60 min, the mixture was boiled for 3 min to terminate the enzyme reaction. To the mixture, 10 ml of glycogen solution (20 mg/ml) was added to co-precipitate the newly formed glucans with 75% methanol. The amount of the radioactive, methanol-insoluble material was measured using a liquid scintillation counter (Beckman LS5000TA). One unit of enzyme activity was defined as 1  $\mu$ mol of D-[<sup>14</sup>C]glucose incorporation from  $\alpha$ -D-[<sup>14</sup>C]glucose 1-phosphate into the methanol-insoluble material per min under the conditions used.

*Antibodies* --- Affinity-purified antibody against maize kernel BE-I was obtained as described previously<sup>12)</sup>. Antiserum against potato tuber type L phosphorylase was a kind gift from Dr. Toshio Fukui at the

Institute of Scientific and Industrial Research, Osaka University. An 87-kDa form of rice branching enzyme, termed RBE3, was separated by SDS-PAGE, as described below, and the gels were stained with Coomassie brilliant blue R-250. The protein band corresponding to RBE3 was cut from the gels, passed through a 21-gauge needle, and suspended in 0.5 ml of phosphate-buffered saline. The suspension was emulsified by sonication with 0.5 ml of Freund's complete adjuvant (Difco), and injected intradermally into rabbits. The injection procedure was repeated twice at 2-week intervals. The final injection was carried out using Freund's incomplete adjuvant. Antisera were collected every 7 days after the final injection. The antibodies were purified by fractionation with ammonium sulfate (0-40% saturation) followed by immunoaffinity chromatography on a column of Sepharose 4B which had been substituted with RBE3 by the cyanogen bromide procedure<sup>18</sup>. The antibodies retained on the column were eluted with 0.2 M glycine/HCl, pH 2.2, immediately neutralized with 0.4 M potassium phosphate, pH 8.0, and then dialyzed against phosphate-buffered saline.

*SDS-PAGE and Western Blot Analysis* --- SDS-PAGE was performed as described by Laemmli<sup>19</sup>. Briefly, protein samples were boiled for 3 min with an equal volume of a solution consisting of 20 mM Tris/HCl, pH 6.8, 2% SDS, 2% 2-mercaptoethanol, 16% glycerol, and 0.04% bromophenol blue. The gels were stained with Coomassie brilliant blue R-250. For Western blot analysis<sup>20</sup>, protein were separated by SDS-PAGE and transferred onto Immobilon-P polyvinylidenedifluoride (PVDF) membranes (Millipore), using a Sartorius Semi-Dry electroblotter. After blocking with 0.5 % skim milk, the blots were treated with affinity-purified antibodies against

maize BE-I or RBE3, and then with goat anti-rabbit IgG horseradish peroxidase conjugate. The immunoreactive bands were detected as purplish black bands after incubation with 0.05% 4-chloronaphthol and 0.0003% hydrogen peroxide in 20 mM Tris/HCl, pH 7.5, containing 0.15 M NaCl.

*Construction of a Rice Seed cDNA Library* --- Total RNA from developing rice seeds was prepared by the selective precipitation method with 8 M urea/3 M LiCl solution, as described previously<sup>12)</sup>. Polyadenylated RNA was selected by oligo(dT)-cellulose (Type 7, Pharmacia LKB Biotechnology) column chromatography. Prior to cDNA synthesis, the polyadenylated RNA (5 µg) was denatured in 10 mM methylmercury hydroxide for 5 min at room temperature. Double-stranded cDNA was synthesized from the denatured RNA using a commercial kit (Pharmacia LKB Biotechnology) according to the manufacturer's protocol. After *EcoRI/NotI*-adaptor ligation, the cDNAs were inserted into *EcoRI*-cleaved, dephosphorylated  $\lambda$ gt11 arms (Stratagene). The recombinant phage DNAs were then packaged using an *in vitro* packaging extract (Gigapack Gold, Stratagene). The original titer of the cDNA library was  $3.3 \times 10^6$  pfu/ml with 2% non-recombinants.

*Screening of a Rice Seed cDNA Library* --- Screening of the cDNA library was carried out by the plaque hybridization method<sup>21)</sup>, using the cDNA insert of maize BE-I clone, MB9, as a probe<sup>12)</sup>. The library was plated at a density of  $1.5 \times 10^5$  plaques/plate (23 x 23 cm). Nitrocellulose lifts were denatured, neutralized, baked for 2 h at 80 °C, and prehybridized for 1 h at 42 °C in a solution containing 5x SSPE (1x SSPE=10 mM sodium phosphate, pH 7.7, 0.18 M NaCl, and 1 mM

EDTA), 0.02% Ficoll 400, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin, and 0.1% SDS. Hybridization was carried out for 16 h at 60°C in the prehybridization buffer containing 0.1 mg/ml denatured salmon testis DNA and 2 ng/ml DNA probe which had been <sup>32</sup>P-labeled by the random-oligonucleotide priming method<sup>22</sup>), using a commercial kit (Nippon Gene, Toyama). The filters were washed in 2x SSC (1x SSC=15 mM sodium citrate, pH 7.0, and 0.15 M NaCl) at room temperature, in 2x SSC containing 0.1% SDS at 60°C, and in 2x SSC at room temperature, dried, and then autoradiographed at -80°C. Positive clones were plaque-purified, and the cDNA inserts were obtained by digestion of the recombinant phage DNAs with *EcoRI*, and subcloned into the *EcoRI* site of pUC19.

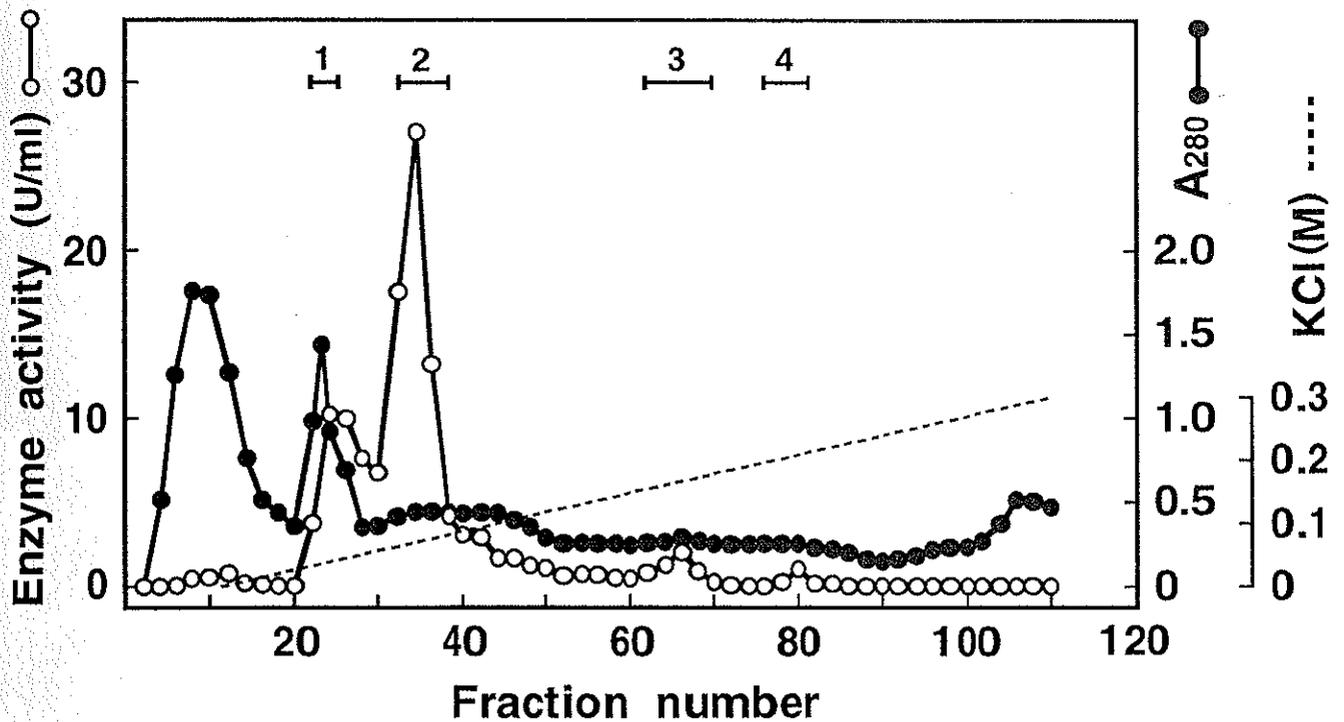
*Northern Blot Analysis* --- Total RNAs (usually 20 µg) were separated on formaldehyde-agarose gels, and transferred onto Hybond-N membranes. The blot was probed with <sup>32</sup>P-labeled DNA fragments. The conditions employed for hybridization were essentially similar to those described previously<sup>23</sup>). After washing, the blots were dried and analyzed using a BAS2000 Bio-image Analyzer (Fuji Photo Film, Tokyo).

*Analytical Procedures* --- Protein concentration was determined by the method of Hartree<sup>24</sup>) or Bradford<sup>25</sup>). Nucleotide sequence analysis was carried out by the dideoxy chain-termination method<sup>26</sup>), using a Sequenase Version 2.0 kit from U.S. Biochemical, in both directions. Computer-aided analysis of nucleotide and protein sequences was carried out using a GENETYX program (Software Development, Tokyo).

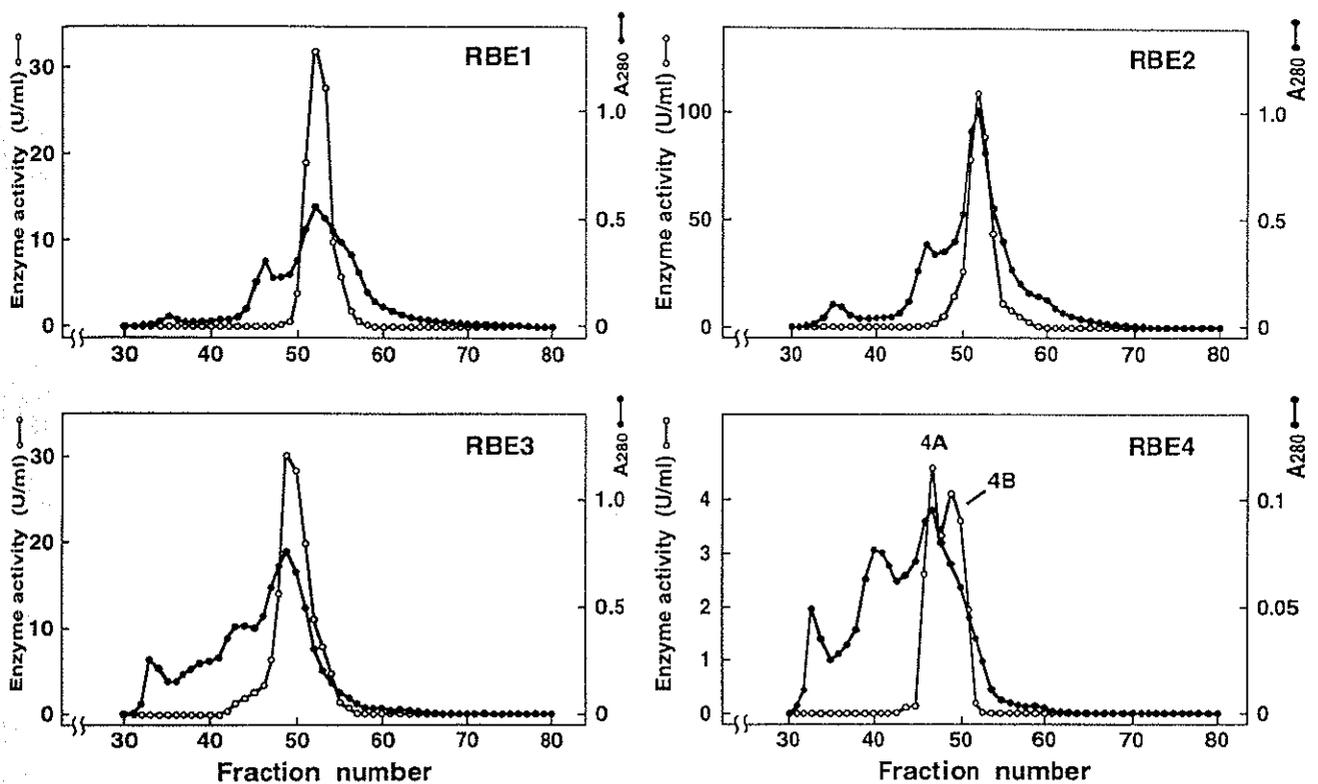
## Results

*Purification of Starch Branching Enzyme from Developing Rice Seeds* --- Crude extracts from developing rice seeds were fractionated by 0-70 % saturation of ammonium sulfate, and subjected to ion-exchange chromatography on a DEAE-cellulose column (Fig. 2). Four forms of branching enzyme, termed RBE1, RBE2, RBE3, and RBE4, were separated by a linear gradient of 0-0.3 M KCl. Further elution of the column with 1 M KCl yielded no enzyme activity (not shown). Fractions containing the enzyme activity were separately combined, and filtered through a Toyopearl HW-55F column (Fig. 3). A single peak of the enzyme activity was observed for RBE1, RBE2, and RBE3, whereas RBE4 was further separated into two peaks (RBE4A and RBE4B). It is important to note that, as judged from the elution pattern of branching enzyme on the DEAE-cellulose column, RBE1 and RBE2 are the predominant forms of the enzyme.

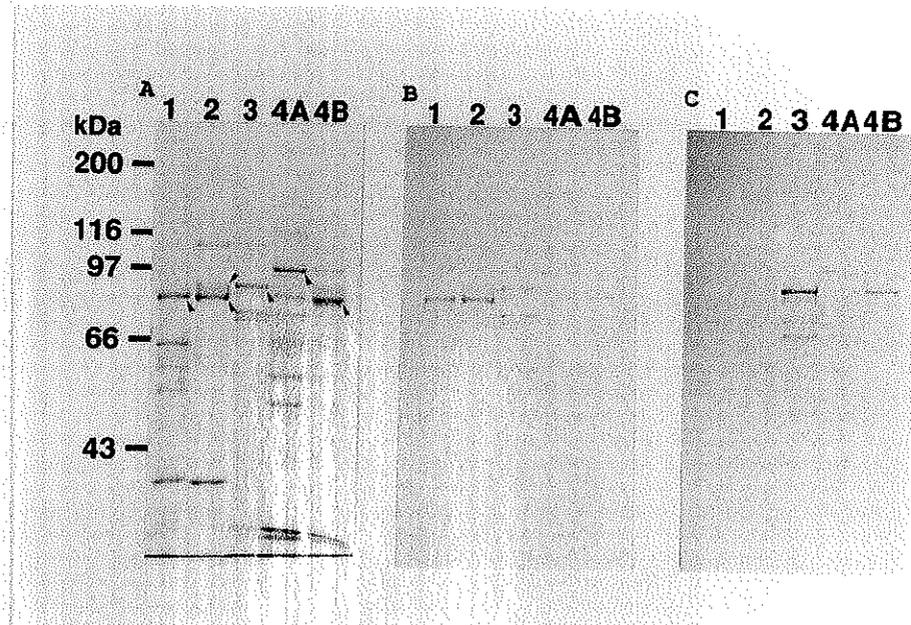
Western blot analysis using affinity-purified anti-maize BE-I antibody indicated that the molecular masses of the major immunoreactive proteins in the RBE1, RBE3, RBE4A, and RBE4B fractions were 82, 87, 93, and 83 kDa on SDS-PAGE, respectively (Fig. 4, Panels A and B). The RBE2 fraction gave two immunoreactive bands which migrated as proteins with molecular masses of 85 and 82 kDa (termed RBE2A and RBE2B, respectively) (Fig. 5). The immunoreactivities of RBE1, RBE2A, and RBE2B were greater than those of other forms of branching enzyme. Affinity-purified anti-RBE3 antibody reacted strongly with the 87-kDa RBE3, but not with the 82-kDa RBE1, 85-kDa RBE2A, or 82-kDa RBE2B (Fig. 4, Panel C). No immunoreactive protein band was detected when the blots were probed with anti-potato tuber type-L phosphorylase antiserum (Fig. 6). Other



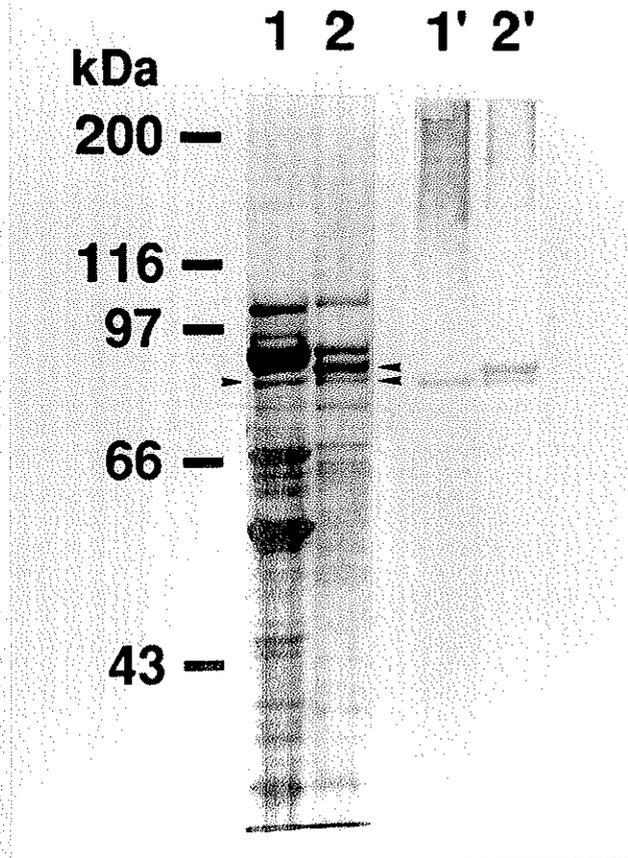
**Fig. 2. Chromatography of branching enzymes from developing rice seeds on DEAE-cellulose.** Crude extracts from rice seeds were fractionated by a 0-70% saturation of ammonium sulfate, and applied to a DEAE-cellulose column (1.9 x 40 cm). Proteins were eluted from the column with a linear gradient of 0-0.3 M KCl (-----, 2 liters). Fractions (20 ml) were collected and assayed for the enzyme activity (○) and absorbance at 280 nm (●). Numbers (1-4) indicate the fractions containing RBE1, RBE2, RBE3, and RBE4.



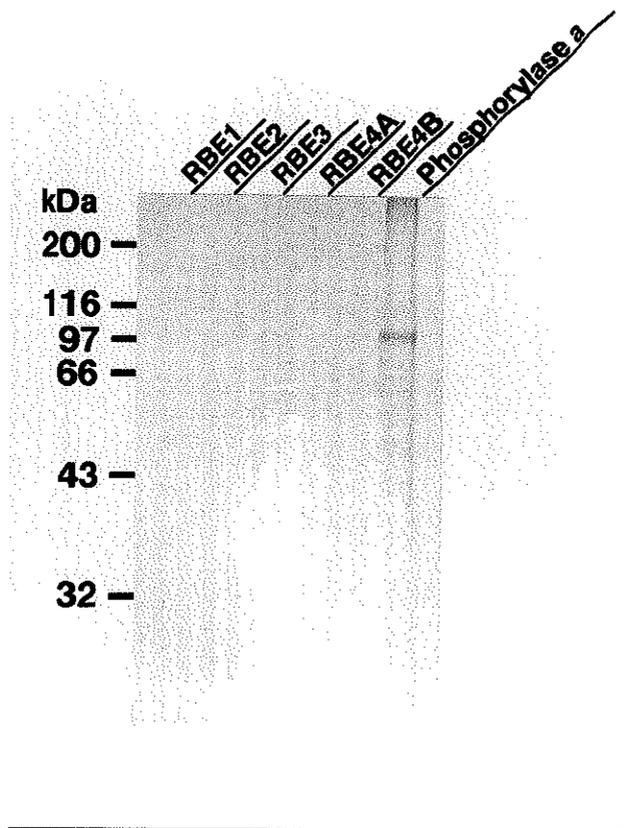
**Fig. 3. Elution patterns of branching enzymes on Toyopearl HW-55F.** Each of the four forms of branching enzyme, after DEAE-cellulose column chromatography, was filtered through a column (1.9 x 95 cm) of Toyopearl HW-55F. Fractions (3 ml) were collected at a flow rate of 10 ml/h and assayed for the enzyme activity (○) and absorbance at 280 nm (●). Note that approximately one-third of partially purified RBE1 and RBE2 from the DEAE-cellulose column chromatography (Fig. 2) was applied to the Toyopearl HW-55F column.



**Fig. 4. SDS-PAGE and Western blot analysis of partially purified branching enzymes.** The peak fractions of RBE1 (1), RBE2 (2), RBE3 (3), RBE4A (4A), and RBE4B (4B) from gel filtration on Toyopearl HW-55F (Fig. 3) were subjected to SDS-PAGE (panel A) followed by Western blot analysis (panels B and C). The gel was stained with Coomassie brilliant blue (panel A). The protein bands indicated by *arrows* (see panel A) corresponded to 82-kDa RBE1, 85-kDa RBE2A, 82-kDa RBE2B, 87-kDa RBE3, 93-kDa RBE4A, and 83-kDa RBE4B, all of which immunoreacted with affinity-purified anti-maize BE-I antibody (panel B). The blot was also probed by affinity-purified anti-87-kDa RBE3 antibody (panel C).



**Fig. 5. SDS-PAGE and Western blot analysis of RBE1 and RBE2 from DEAE-cellulose column chromatography.** SDS-PAGE (left panel) and subsequent Western blot analysis (right panel) of RBE1 (1 and 1') and RBE2 (2 and 2') from DEAE-cellulose column chromatography (Fig. 2). Arrows represent 82-kDa RBE1, 85-kDa RBE2A, and 82-kDa RBE2B. The blot was probed by affinity-purified anti-maize BE-I antibody. Note that the amount of 85-kDa RBE2A in the RBE2 fraction significantly decreased after gel filtration on Toyopearl HW-55F (see Fig. 4A).



**Fig. 6.** Western blot analysis of partially-purified branching enzymes. The peak fractions of RBE1, RBE2, RBE3, RBE4A, and RBE4B from gel filtration on Toyopearl HW-55F were subjected to Western blot analysis. The blot was probed by anti-potato tuber type L phosphorylase antiserum.

bands of proteins, which immunoreacted weakly with anti-maize BE-I and/or anti-RBE3 antibodies, were also found (Fig. 4, Panels B and C). These proteins may be partially degraded products from each form of branching enzyme. It should be noted that, as shown in Fig. 4, the ratio of 85-kDa RBE2A to 82-kDa RBE2B in the RBE2 preparation was significantly changed after gel-filtration on Toyopearl HW-55F.

Five proteins, 82-kDa RBE1, 85-kDa RBE2A, 82-kDa RBE2B, 87-kDa RBE3, and 93-kDa RBE4A, were separated by SDS-PAGE, transferred electrophoretically onto a PVDF membrane, and subjected to analysis of the amino-terminal amino acid sequences (Table 1). RBE1 possessed two amino-terminal amino acid sequences, Thr-Met-Val-Xaa-Val-Val-Glu-Glu-Val-Asp-His-Leu-Pro-Ile-Tyr- and Val-Xaa-Val-Val-Glu-Glu-Val-Asp-His-Leu-Pro-Ile-Tyr-Asp-Leu-, where Xaa was not identified. These two sequences were related to each other; the latter sequence lacked two amino acids of the former at the amino terminus. In addition, these sequences showed a high degree of identity with the amino-terminal sequence of maize BE-I<sup>12)</sup>, as described below. Unexpectedly, both RBE2A and RBE2B also had two amino-terminal sequences identical to those of RBE1. The amino-terminal sequences of RBE3 and RBE4A (Table 1) were not related either to each other or to those of RBE1. Thus, RBE1, RBE2A, and RBE2B are the same in the molecular size, immunoreactivity to anti-maize BE-I antibody, and amino-terminal amino acid sequence, except that the molecular size of RBE2A is 3-kDa larger than those of RBE1 and RBE2B (Fig. 4).

#### *Molecular Cloning of an Isoform of Rice Branching Enzyme, RBE1*

--- We have recently identified and characterized cDNA clones, termed MB9 and MB10, coding for maize BE-I<sup>12)</sup>. The cDNA insert of MB9 was <sup>32</sup>P-labeled and used to screen approximately  $7.5 \times 10^5$  recombinant

**Table 1. Comparison of amino terminal amino acid sequences of multiple forms of rice branching enzymes**

Cycle	RBE1		RBE2A		RBE2B		RBE3	RBE4
1	T(1.1)	V(6.3)	T(0.8)	V(0.4)			A(32.3)	S(3.3)
2	M(4.1)	X(-)	M(0.8)	X(-)			A(28.4)	G(15.0)
3	V(14.8)	V(14.8)	V(2.9)	V(2.9)			G(23.3)	A(20.2)
4	X(-)	V(6.5)	T(0.4)	V(0.7)	T(0.6)	V(0.7)	A(27.3)	P(18.4)
5	V(7.7)	E(6.9)	V(2.1)	E(1.2)	V(1.9)	E(6.7)	S(3.2)	G(9.7)
6	V(5.1)	E(6.0)	V(1.8)	E(1.3)	V(0.6)	E(8.9)	G(17.9)	K(10.8)
7	E(4.8)	V(6.7)	E(2.3)	V(0.9)	E(6.5)	V(4.9)	E(16.9)	V(12.4)
8	E(4.7)	D(3.6)	E(2.3)	D(0.9)	E(1.1)	D(7.6)	V(20.2)	L(12.9)
9	V(4.0)	H(2.1)	V(1.5)	H(0.5)	V(1.8)	H(4.2)	M(14.3)	V(11.6)
10	D(2.9)	L(4.5)	D(1.4)	L(0.8)	D(2.5)	L(7.9)	I(18.7)	P(11.0)
11	H(1.3)	P(2.2)	H(0.9)	P(0.3)	H(1.9)	P(6.7)	P(18.3)	G(18.6)
12	L(1.9)	I(2.1)	L(1.0)	I(0.1)	L(2.2)	I(4.8)	E(10.3)	G(16.3)
13	P(1.2)	Y(3.2)	P(0.6)	Y(0.7)	P(1.7)	Y(5.0)	G(16.9)	G(9.2)
14	I(2.0)	D(2.0)	I(0.4)		I(2.6)	D(5.2)	E(10.3)	S(1.0)
15	Y(2.4)	L(0.6)	Y(0.7)		Y(1.4)	L(1.5)	S(0.9)	D(9.1)
16	D(4.6)	D(4.6)	D(0.8)		D(4.2)	D(4.2)	D(11.0)	D(5.7)
17	L(1.0)	P(0.3)	X(-)		L(2.1)	P(0.2)	G(7.2)	L(8.2)
18	D(1.0)	K(0.2)	D(0.5)		D(2.5)	K(0.7)	M(4.7)	L(2.4)
19		L(0.5)	P(0.2)		P(1.4)	L(2.1)	P(6.3)	
20		E(2.7)					V(3.4)	

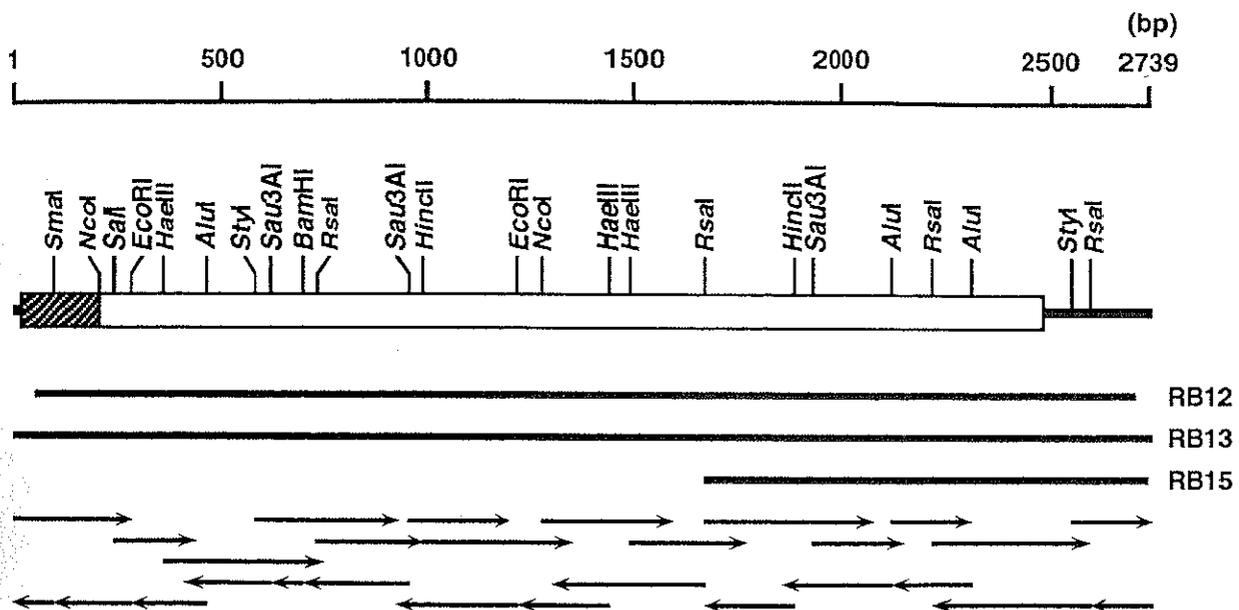
The values in the parenthesis indicate lag corrected pico mols.

phages from a rice seed cDNA library in  $\lambda$ gt11. Twenty positive clones were obtained after two more rounds of screening. Restriction mapping and Southern blot analysis of these cDNA inserts showed that they were all related. Therefore, we selected at random three clones, termed RB12, RB13, and RB15, for further characterization. The restriction map and sequencing strategy used are given in Fig. 7. Of these three clones selected, RB13 contained the longest cDNA insert with a size of approximately 2.7 kbp.

The nucleotide sequence of the overlapping cDNA inserts contains a 2,460-nucleotide open reading frame, which is flanked by 19-nucleotide 5'-untranslated and 257-nucleotide 3'-untranslated regions (Fig. 8). The ATG start codon is assigned on the basis of its similarity to the eukaryotic consensus sequence, as described by Kozak<sup>27</sup>. A putative polyadenylation signal, AATAAA, is located at nucleotides 2,683-2,688 and 2,708-2,713.

The 17-residue sequence at residues 1-17 matches the amino-terminal sequences of RBE1, RBE2A, and RBE2B determined experimentally (Fig. 8). This result confirms that the isolated cDNA clones encode RBE1 as well as RBE2A and RBE2B (to avoid complexity, the term, RBE1, is used for RBE1, RBE2A, and RBE2B hereafter). The cDNA-derived sequence indicates that rice RBE1 is initially synthesized as an 820-residue polypeptide, and that the mature protein contains 756 (or 754) amino acids, including 9 cysteine residues, with a calculated molecular mass of 86,734 (or 86,502) Da (Fig. 8). This value is consistent with that determined for RBE2A by SDS-PAGE (Fig. 4).

*Comparison of Transit Peptide Sequences of RBE1 with Several Amyloplastic Enzymes* --- The rice RBE1 precursor as well as the maize



**Fig. 7. Restriction map and sequencing strategy for the cDNA clones encoding an isoform of rice branching enzyme, RBE1.** The scale at the top designates nucleotide position in base pairs from the 5'-end of RB13. The *open box* represents the mature protein coding region, and the *hatched box* indicates a putative transit peptide required for translocation of branching enzyme into amyloplast. *Arrows* indicate the direction and extent of nucleotide sequence determined from each site.

		GCCA CCGACATCCG	14
15	CCGCA ATGCTGTGTCTCACCTCCTCTTCCCTCCTCGCGCCCGCTCCCGCTCCTCCTCTCGGTGATGACCGAGCCCGGAATCGGGGGGGGGT	112	
-64	M L C L T S S S S S A P A P L L P S L A D R P S P G I A G G G	-34	
113	GGCAATSTTCGCCGTAGCGTGGTTCTTCGCGCCGCGCGTCTGGCCGCGGAAATTCAGTTCCTGCGACTGCGGAAAAACAAA	211	
-33	G N V R L S V V S S P R R S W P G K V K T N F S V P A T A R K N K	-1	
212	ACCATGGTACTGTTGGAGGAGGTGCACACCTTCTATATATGATCTGGACCCTAAGTTGGAGGAATTCAGGATCAGTTCAACTATAGGATAAAA	310	
1	<u>T M V T V V E V D H L P I Y D L D P K L E E F K D H F N Y R I K</u>	33	
311	AGATACCTCGACCAAAATGCCTGATGTGAAAAACATGAGGGGGGCTTGAAGAATTTCTAAAGGCTATTGAAGTTTGGATTAAACAGTTGATGGT	409	
34	R Y L D Q K C L I E K H E G G L E E F S K G Y L K F G I N T V D G	56	
410	GCCACAATATATCGTCAATGGCGCCTCTGCACAAGAAGCACAGCTCATTGGTGAATCAATAACTGGAATGGTCAAAACACAAGATGGAGAAGGAT	508	
67	A T I Y R E W A P A A Q E A Q L I G E F N N W H G A K H K M E K D	99	
509	AAATTTGGCAATTTGGTCAATCAAGATTTACATGTCATGGGAAGCTGCCATCCCTCACAATTCAGAGTAAATTTGGCTTTGGGCTGGGGTGGGA	607	
100	K F G I W S I K I S H V N G K P A I P H N S K V K P R F R H G G G	132	
608	GCATGGGTGATCGTATTCGGCATGGATTCTGTATGCAACTTTTGTATGCCCTTAAATTTGGAGTCCATATGATGGTGTACACTGGGATCCTCCAGCC	706	
133	A W V D R I P A W I R Y A T F D A S K F G A P Y D G V H W D P P A	165	
707	TGTGAAGGTACGTGTTAAGCATCTCGACCTCCAAACCTGATGCTCCACCCATCTATGAGGCTCATGTTGGGATGACTGGTGAAGAGCCAGAAGTA	805	
166	C E R Y V F K H P R P P K P D A P R I Y E A H V G M S G E E P E V	198	
806	AGCACATACAGAGAATTTGCAGACAATGTGTACCAGGCATACGGGCAAAATACTACAACACAGTTCAAGTTAAGCAATCATGGAACATCCTACTAT	904	
199	S T Y R E F A D N V L P R I R A N N Y N T V Q L M A I M E H S Y Y	231	
905	GCTTCTTTGGGTATCAGTGCACAAATTTTTCGCGAGTCAGCAGCAGTACAGGAACACCAAGAGGATCTGAAATATCTTGTGACAAGGCACATAGTTTA	1003	
232	A S F G Y H V T N F F A V S S R S G T P E D L K Y L V D K H E S	265	
1004	GGATTACGAGTTCTGATGGATGTGTCCATAGCCATGCGAGTAATAATGTGACGATGCTCTAAATGGCTTGTGACCTTGGACAAAACACTCATGAGTCT	1102	
265	G L R V L H D V V H S H A S N N V T D G L N G Y D V G Q N T H E S	297	
1103	TATTTTCATACAGGAGATACGGGCTACCATAAACTCTGGGATAGTCTGTCTTCAACTATGCCAATTTGGGAGTCTTAAGATTTCTTCTTCTAAATTCG	1201	
298	Y F H T G D R G Y H K L W D S R L F N Y A N W E V L R F L L S N L	330	
1202	AGATATTGGATGGACCAATTCATGTTTGTATGGCTTCCGATTTGATGGGGTTACATCAATGCTATACCATCACCATGGTATCAATAAGGGATTTACTGGA	1300	
331	R Y W M D E F M F D G F R F D G V T S M L Y H H H G I N K G F T G	363	
1301	AACTACAAGGATATTTCAAGTTGGATACCGATGCGATGCAATGTGTTTACATGATGCTCCGAAACCATTTAATGATTAAGTCTTTCGCGAAGCAACT	1399	
364	N Y K E Y F S L D T D V D A I V Y M M L A N H L M H K L L P E A T	396	
1400	ATGTTGCTGAAGAATGTTTCGGGATGCGAGTCTTGTTCGGCCAGTTGATGAAGGTGGAGTAGGGTTTGAATTCGCGCTGGCAATGGCCATTCCTGAT	1498	
397	I V A E D Y L K N K E D R K W S M S E I V Q T L T N R R Y T K C I	429	
1499	AGATGGATTGACTACCTGAAGAACAAGAGGACCGCAATGGTCAATGAGTGAATAGTGCAAACTTTGACTAACAGGAGATATACAGAAAAATGCATT	1597	
430	R W I D Y L K N K E D R K W S M S E I V Q T L T N R R Y T K C I	462	
1598	GCCTATGCGAGAGCCATGATCAGTCCATGTTGGTGCACAAGACTATAGCATTTCTCTGATGACAAGGAATGTACACTGGCATGTGAGACTTGCAG	1696	
463	A Y A E S H D Q S I V G D K T I A F L L H D K E M Y T G M S D L Q	495	
1697	CCTGCTTACCATCAACCGTGGCATTCGACTCCAAAAGATGATTCACTTCAATACGATGGCCCTTGGAGGTGATGGCTACTTAAATTTATGGGC	1795	
496	P A S P T I N R G I A L Q K M I K F I T M A L G G D G Y L N F M G	528	
1796	AATGAGTTTGGCCATCCAGAATGGATTGACTTTCCAAAGAGAGGCAACACTGGAGCTATGATAAATGCAGAGCTGAGTGGGCTTGTGACACTGAT	1894	
529	N E F G H P E W I D F P R E G N N W S Y D K C R R Q W S L V D T D	561	
1895	CACCTTCGATACAAGTATATGAATGCAATTTGATCAAGCAATGAATGCACTCGAGGAGAAATTTCCCTTCCCTCATCATCAAGCAGATGTTTACGGAC	1993	
562	H L R Y K Y M N A F D Q A M N A L E E E F S F L S S S K Q I V S D	594	
1994	ATGAACGAGAAAGATAAGGTTATGTCTTTTGAACGTTGAGATTTTGTGTTTTCATTTTCAATTTTCAATCCCAACAAACTTACAAGGGTTACAAAGTCGGA	2092	
595	M N E K D K V I V F E R G D L V F V F N F H P N K T Y K G Y V G	627	
2093	TGTGACTTCCCGGAGTACAGAGTAGCTCTGAGCTGATGCTTTGGTCTTGGTGGCCATGGAAGAGTGGCCATGATGTTGATCAGTTCAGCTCT	2191	
628	C D L P G K Y R V A L D S D A L V F G G H G R V G H D V D H F T S	660	
2192	CCCGAGGAATGCCAGGAGTACCAGAAACAAATTTCAACAACCGCCCTAATCATCAAGTCTTTCCCGCCCGGTACCTGTGGCTTACTATCCG	2290	
661	P E G M P G V P E T N F N N R P N S F K V L S P P R T C V A Y Y R	693	
2291	GTTGATGAAGATCCGTAAGAGCTCAGGAGGGTGGAGCAGTCTGCTTCTGGAAGAATTTGTTACAGATATATGATGTTGAAGCAACAAGTGGGAGACT	2389	
694	V D E D R E E L R R G G A V A S G K I V T E Y I D V E A T S G E T	726	
2390	ATCTCTGGTGGTGGAGGGCTCCGAGAAGGACGATTTGGCAAGAAAGGATGAAGTTTGTGTTTGGTCTTCTGACGAGACTGCAATGA AGCAT	2487	
727	I S G G W K G S E K D D C G K K G M K F V F R S S D E D C K *	756	
2488	CAGATTTCTT GATCAGGAGC AACTGPIGGT GCCPIGTAA TCTGGAGATC CTGGCTGCC TTGGACTGG TTGTGTTCT TTAGCAGTTG	2577	
2578	CTATGACCT ATCTATGATA TGAACTTTAT GTATAGTTCG CCTTAAAGAA AGAATAAGCA GTGATGATGT GGCCTTAAAC CTGAGCTGCA	2667	
2668	CAAGCCTAAT GTA AAAATAA AGTTTCAGGC TTTCATCCAG AATAAAACAG CTGTTTCATT ACCATCTCAA AA	2739	

**Fig. 8. cDNA sequence of an isoform of rice branching enzyme, RBE1 and its deduced amino acid sequence.** The deduced amino acid sequence is shown below the nucleotide sequence numbered in the 5' to 3' direction. The amino-terminal sequences of RBE1, RBE2A, and RBE2B determined by protein analysis are underlined with a wavy line. The amino acids are numbered from the amino terminus, and the residues on the amino-terminal side of Thr at residue 1 are indicated by *negative numbers*. A putative polyadenylation signal is underlined with *broken lines*. Note that RBE1, RBE2A, and RBE2B possessed two amino-terminal amino acid sequences which started from Thr and Val at residues 1 and 3, respectively.

BE-I does not possess a typical signal peptide sequence (Figs. 8 and 9). Since branching enzyme is located in the amyloplast of seed endosperm, this enzyme probably contains a leader sequence necessary for transport into the amyloplast. The 64-residue sequence at residue -64 to -1 preceding the amino terminus of the mature RBE1 is rich in hydroxylic (23% as Ser and Thr) and basic (14% as Arg and Lys) amino acids, and contains an acidic amino acid (Asp) only at residue -44 (Figs. 8 and 9). Thus, the amino-terminal leader peptide of the RBE1 precursor is similar in amino acid composition to the transit peptide of potato phosphorylase precursor, which is nuclear-encoded and transported into the amyloplast<sup>28,29</sup>. The putative transit peptide of the RBE1 precursor shares a significant degree of sequence identity (53%) with that of the maize BE-I (Fig. 9), but not with those of ADP-glucose pyrophosphorylase<sup>30</sup> and waxy protein<sup>31</sup>.

*Branching Enzyme Belongs to a Family of Amylolytic Enzymes ---* Sequence alignment of mature rice RBE1 to maize BE-I shows 86% identity (Fig. 9). Interestingly, the sequence in the carboxyl terminal region is poorly conserved. RBE1 contains nine cysteine residues, five of which are conserved between the rice and maize proteins. Since plant branching enzyme requires free sulfhydryl groups for the activity<sup>32</sup>, the conserved cysteine residue(s) may play an important role in the enzymatic function without formation of disulfide bonds. Recently, we have demonstrated the sequence conservations of four regions, which constitute the catalytic sites of amylolytic enzymes such as  $\alpha$ -amylase, in maize BE-I<sup>12</sup>. The sequences of these four regions are identical between the rice and maize proteins (at residues 271-276, 341-349, 400-403, and 464-469 in the rice sequence, see Fig. 8), indicating that both rice RBE1 and maize BE-I belong to a family of amylolytic enzymes.

```

R : MLCLTSSSSSAPAPLLP-----SLADR-PSPGIACGGGNVRLSVVSS--SPRRSWPGKVKTN -12
M : LCLVSPSSS-PTPLPEPRRSRSHADRAAPPGIA-GGNVRLSVLSVQCKARRSGVRKVKSK -5
R : FSVPATARKN---KTMVTVVEVDHLPYDLDPKLEEFKDFNFYRIKRYLDQKCLIEKHEGG 48
M : F---ATAATVQEDKTMATAKGDVDHLPYDLDPKLEIFKDFRYRMKRFLQKGSIEENEES 55
R : LEEFSKGYLKEFGINTVDGATYREWAPAAQEAQLIGEPNNWNGAKHKMEKDKFGIWSIKISH 110
M : LESEFSKGYLKEFGINTNEDGTVYREWAPAAQEAELIGEPNDWNGANHKMEKDKFGVWSIKIDH 117
R : VNGKPAIPHNSKVKFRFRHGGGAVVDRIPAWIRYATFDASKFGAPYDGVHWDPPACERYVFK 172
M : VKGKPAIPHNSKVKFERELH-GGVVVDRIEALIRYATVDASKFGAPYDGVHWDPPASERYTFK 178
R : HPRPPKPDAPRIYEAHVGMSGEEPVEVSTYREFADNVLPRIRANNYNTVQLMAIMESYYASF 234
M : HPRPSKPAAPRIYEAHVGMSGEKPAVSTYREFADNVLPRIRANNYNTVQLMAVMEHSYYASF 240
R : GYHVTNFFAVSSRSRSGTPEDLKYLVDKAHSLGLRVLM DVVHSHA SNNVT DGLNGYDVGQNTHE 296
M : GYHVTNFFAVSSRSRSGTPEDLKYLVDKAHSLGLRVLM DVVHSHA SNNVT DGLNGYDVGQSTQE 302
R : SYFHTGDRGYHKLWDSRFLFNANWEVLRFLLSNLRYSWDEFMFDGFRFDGVTSMLYHHHGIN 358
M : SYFHAGDRGYHKLWDSRFLFNANWEVLRFLLSNLRYSWDEFMFDGFRFDGVTSMLYHHHGIN 364
R : KGFTGNYKEYFSLD TDVDAIVYMLLANHLMHKLLPEATIVAEDVSGMPVLCRPVDEGGVGF 420
M : VGFTGNYQKEYFSLD TAVDAVVYMLLANHLMHKLLPEATVVAEDVSGMPVLCRPVDEGGVGF 426
R : FRLAMAIPDRWIDYLNKKEEDRKWSMSEIVQTLTNRRYTEKCIAYAESHDQSI VGDKTIAFLL 482
M : YRLAMAIPDRWIDYLNKKEEDSEWSMGEIAHTLTNRRYTEKCIAYAESHDQSI VGDKTIAFLL 488
R : MDKEMYTGMSDLOPASPTINRGIALQKMIHFITMALGGDGYLNFMGNEFGHP EWIDFPREGN 544
M : MDKEMYTGMSDLOPASPTIDRGIALQKMIHFITMALGGDGYLNFMGNEFGHP EWIDFPREGN 550
R : NWSYDKCRRQWSLVD TDHLYRYKYMNAFDQAMNALEEFBSFLSSSKQIVSDMNEKDKVIVFER 606
M : NWSYDKCRRQWSLVD TDHLYRYKYMNAFDQAMNALDERFSELSSSKQIVSDMNEDEKVIIVFER 612
R : GDLVVFVNFHFNKTYKGYKVGCDLPGKYRVALDSDALVFGGHGRVGHVDVDFHFTSPEGMPGVP 668
M : GDLVVFVNFHFKKTYEGYKVGCDLPGKYRVALDSDALVFGGHGRVGHVDVDFHFTSPEGVPGVP 674
R : ETNFNNRPNSEKVLSPERTCVAYYRVDEEDREELRRGCAVA-SGK-IVTEYIDVEATSGETIS 728
M : ETNFNNRPNSEKVLSPERTCVAYYRVDE-AGAGRRLHAKAETGKTSPAESIDVKA-SRA-SS 733
R : GGWKGSEKDDCGKKGMKVFVFRSSDEDCK 756
M : KEDK--EATAGGKKGWKFARQPSDQDTK 759

```

**Fig. 9. Comparison of deduced amino acid sequences between rice RBE1 (R) and maize (M) branching enzyme-I precursors.** Amino acids are represented by the standard one-letter codes. Identical residues are indicated by shaded boxes. Dashes represent gaps introduced to maximize the similarity. Arrow heads indicate the amino-terminal amino acids of the rice and maize mature proteins. The boxed sequences of the rice and maize proteins share a high degree of identity with the sequences of four active-site regions of amylolytic enzymes (regions I, II, III, and IV)<sup>33,34</sup>. The maize sequence is that reported by Baba *et al.*<sup>12)</sup>

RBE1 shows only a limited sequence identity with *E. coli* branching enzyme<sup>13)</sup> (Fig. 10). However, the sequence corresponding to the four active-site regions of amylolytic enzymes, as described above, are highly conserved between these two proteins. The overall sequence identity of RBE1 is 23% with the *E. coli* enzyme. The amino acid sequence of potato branching enzyme has been recently reported<sup>35,36)</sup>. The sequence identity of the RBE1 was 81% with the potato branching enzyme (Fig. 10).

*Accumulation of RBE1 during endosperm development* --- Western blot analysis indicated that both rice RBE1 and RBE3 are produced specifically in the endosperm tissues of developing seeds (Fig. 11). Even when the same blot was re-probed by the 10-fold higher concentration of affinity-purified anti-maize BE-I or anti-RBE3 antibody, no immunoreactive protein band was detectable in the embryo extracts (not shown). Thus, it is likely that rice BE-I and RBE3 are endosperm-specific proteins.

Accumulation patterns of RBE1 and RBE3 during endosperm development were examined by Northern and Western blot analysis (Fig. 12). When the Northern blot was probed by the <sup>32</sup>P-labeled cDNA insert of RB13, one major and two minor mRNA signals with the sizes of 2.8, 2.0, and 1.6 kb, respectively, were detected (Fig. 12, Panel A). Of these three signals, the 2.8-kb mRNA corresponded to the true message for RBE1, since RNA blot hybridization with the cDNA fragments covering the 5'- and 3'-end regions gave only the 2.8-kb signal (not shown). The RBE1 message is abundantly present between the 7th and 15th days of seed development. The level of the message concentration decreases rapidly after the 15th day of development. The appearance of RBE1 at the protein level during seed development is significantly

```

R : 1 -----T M V T V E V --DH-LPIY D L D----- 18
E : 1 -----M S D R I D R D V I N A L I A G H F A D P F S V L G M H K T H A G L E V R A L L P D A T -D V W V I E P K T 53
P : 1 M E I N F K V L S K P I R G S F P S F S P K V S S G A S R N K I C F P S Q H S T G L K F G S Q E R S W D I S S T P K S R V R K D E R M K H S S A I S A V L T D D N S T M A P L E E D V R T E N - I G L L N L D----- 102

R : 19 ---E K L E---E F K D H F N Y R I K---R Y---L D Q K L I---E K H E G G L E E---E S K S Y L K F G I N T V D G A T I Y R---E W A P A O E A Q L I G E F N N W N G A 92
E : 54 G R K L A K L E C L D S R G E F S G V I P R R K N F F R Y Q L A V V W H G Q O N L I D D P Y R F G P L I Q E M D A W L L S S G T H L R P Y E T L G A H A D T M D G V T G T R F S V W A P N A R R V S V V G Q P N Y W D G R 162
P : 103 ---E T L E---P Y L D H F R H M K---R Y---V D Q K L I---E K Y E G P L E E---E A Q G Y L K F G F N R E D G C I V Y R---E W A P A O E D E V I G D F N G W N G S 176

R : 93 K H K M E K D K - E G I W S I R I - S H Y N G K P A R P H N S R V K F R F R H G G A M V D R I P A W I R Y A T P D A S - K F G A P Y D G V H W D P P A C E R Y V P K H P R P P K P D A P - R I Y E A H V G - - - M S G E 194
E : 163 R H P M R L R K E S G I W E L F I P G A H N G O L Y K Y E M I D A N G N L R - - L K S D P Y A F E A Q M R P E T - - A S L I C G L P E K V V - - - O T E S R L - - - K K A N Q F D A P I S I Y E V H L G S W R R H T D 258
P : 177 N H M M E K D Q - P G V W S I R I - P D V D S K E V L P H N S R V K F R F R H G N G V V D R I P A W I K Y A T A D A T - K P A A P Y D G V V W D P P S E R Y H F K Y P R P P K P R A P - R I Y E A H V G - - - M S S S 278

R : 195 E P E V S T Y R E F A D N V L P R I R A N N Y N T V O L M A I M H S Y Y A S F G Y H V T N F F A V S S R S G T F E D L K Y L V D K A H S L G L R V L M D V V H S H A S N N V T D G L N G V D V G O N T H E S Y F H T E D 303
E : 259 N N F W L S Y R E L A D Q L V P Y A K W M G F T H L E L L P I N S H P F D G S M G Y O P T G L Y A P T R R F G T R D D F R Y E I D A H A A G L N V I L D W V E G H - - F P T D F A L A E F D G T N - - L Y E H S D P 362
P : 279 E P R V N S Y R E F A D D V L P R I K A N N Y N T V O L M A I M H S Y Y G S F G Y H V T N F F A V S S R Y Q N P E D L K Y L I D K A H S L G L O V L V D V V H S H A S N N V T D G L N G F D I G G S Q E S Y F H A G E 387

R : 304 R I G Y H K L W D S R L F N Y A N W E V L R F L L S N L R Y W M D E F M E D G F R F D G V T S M L Y P H H G I N K G F T G N Y K E Y F S L D T D V D A I V Y M M L A N H L M H K L L E A T I V A E D V S G M P V L C R P 411
E : 363 R I E G Y H Q D W N T L I Y N Y G R R E V S N F L V G N A L Y W I S R E G I D A L R V D A V A S M I Y R D Y S R K E G - E W I P N E F E G G - R E N L E A I E F L R N T W R I L G E Q V S G A V T M A E S T D F P G V S R P 469
P : 388 R I G Y H K L W D S R L F N Y A N W E V L R F L L S N L R W M L E E Y N E D G F R F D G T S M L Y P H H G I N M G F T G N Y N E Y F S E A T D V D A V Y Y L M L A N L I H K I F P D A T V I A E D V S G M P F L G R P 495

R : 412 V D E G G V G F D F R L A M A T P D R W I D Y L R N K E D R K M S M S E I V O T L T N R R Y T E K C I A Y A E S H D O S I V G D K T I A F L L M D K E M Y T G M S D L O P A S P T I N R G T A L O K M I H E F T M A I G G 520
E : 470 Q D M G G E G F W Y K K N L G W M H D T L D Y K R L D P V Y R Q Y H H K L T F G I L Y N Y T E - N F V L P L S H D E V V H G K K S H - - - - - L D R M P G D A W O K F A - - - N R A Y Y G W M W A F P G K K - - - 564
P : 496 V S E G G I G F D Y R L A M A T P D R W I D Y L R N K N D E D M S M K E V T S S L T N R R Y T E K C I A Y A E S H D O S I V G D K T I A F L L M D K E M Y S G M S C L T D A S R V V D R Q Y A L H K M I H E F T M A L G G 604

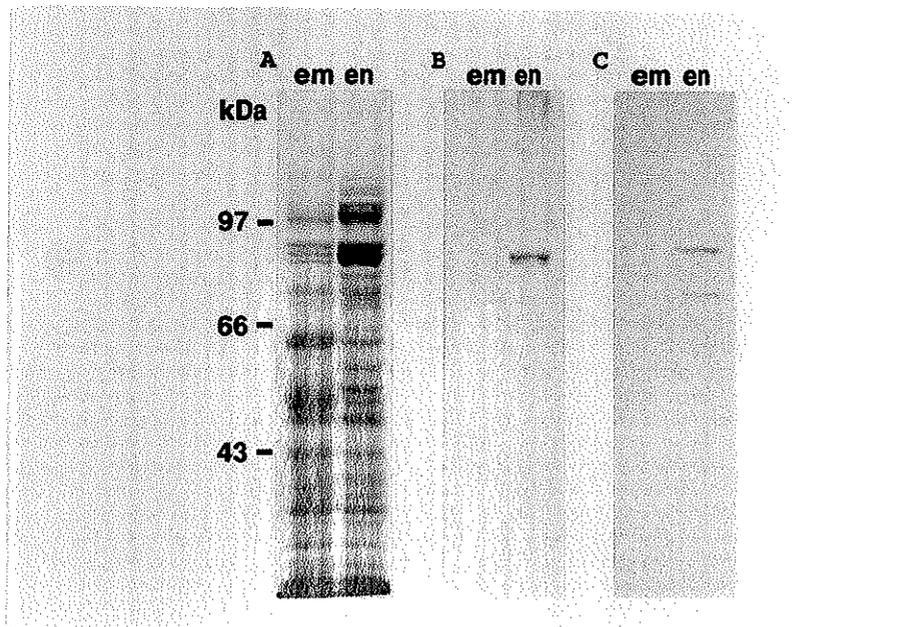
R : 521 D G Y L N F M G N E F G H P - E W I - D - - - - F P R E G - N N W S Y D K C R R O W S L V D T - D H L R Y K Y M N A F D Q A M N A L E E E F S F L S S S K O I V S D M N E R D K V I V F E R G D L V F V E N F H P N K T 620
E : 565 - - - L L F M G N E F A Q G R E W N H D A S L D W H L L E G G D N W H H G V Q R L V R D L N L T Y R H H K A M H E L D F D P Y G F E W L V V D D K E R S V L I F V R R D K E G N E I I V - - - - - A S N F T E - V P 661
P : 605 E G Y L N F M G N E F G H P - E W I - D - - - - F P R E G - N N W S Y D K C R R O W N L A D S - E H L R Y K F M N A F D R A M N S L D E K F S P L A S G K O I V S S M D D D N K V I V F E R G D L V F V E N F H P N T 704

R : 621 Y R K Y K V G C D L P G R Y R V A L D S D A L V F G G H G R V G H D V D H F T S P E G M P G V P E T N F N N R P N S F K - - V L - S P P R - - T C - V - A Y Y R V D E D R E E - L R R - - G G A V A - S - G K - I V T E Y 716
E : 662 R H D N R F G I N O P G R W R E I E N T D - - S M H Y H G - - - - - S N A G - - - - - N O G T V - - - - - 697
P : 705 Y E Y K V G C D L P G K Y R V A L D S D A W E F G G H G R V G H D V D H F T S P E G I P G V P E T N F N G R Q I P S K C C L L R E H V W L I T E L M N A C Q K L K I T R Q T F V V S Y Y Q Q P I S R R V T R N L K I R 813

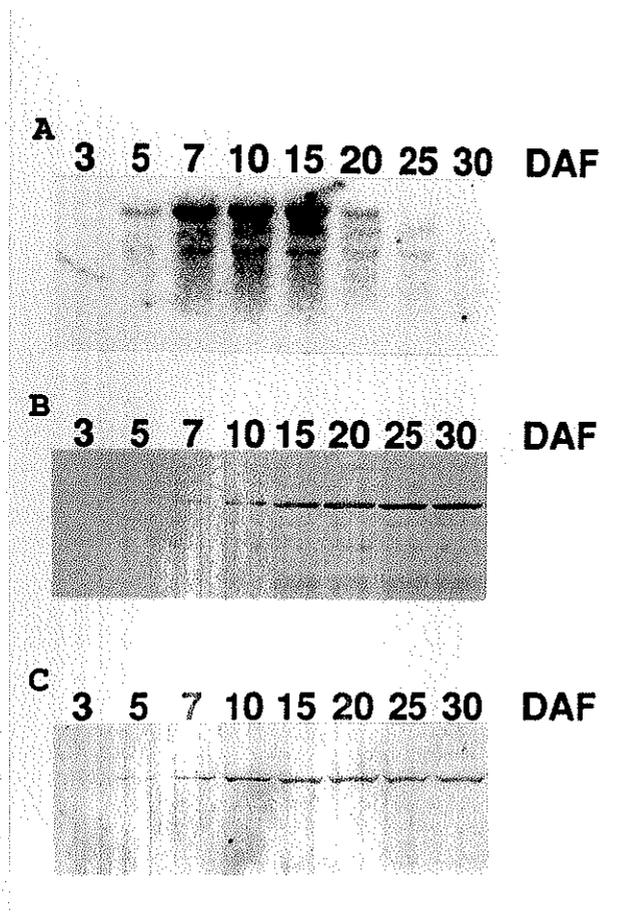
R : 717 I D V E - A - T S G - E T I S G Q W K G S E K D D C G K K G M K F V F R S - S E D E C K 756
E : 698 - - - - - H S D - E I A S H G R O M S L S L T L P P L A T I W L V R E - - A E 728
P : 814 L Q I S V T L E N A C Q R L K F T E O T F L V S Y Y Q Q P I L R R V T E K L K D S L S T N I S T 861

```

Fig. 10. Sequence alignment of rice RBE1 with *E. coli* and potato branching enzymes. The sequence of the rice mature RBE1 (R) is aligned with those of the *E. coli* (E) and potato (P) branching enzymes.<sup>13,35,36</sup> Dashes represent gaps introduced to optimize the alignment. Identical amino acids are indicated by shaded boxes.



**Fig. 11. Accumulation of rice branching enzymes in embryo and endosperm tissues.** Soluble proteins (20 mg each) from embryo (*em*) and endosperm (*en*) tissues of rice seeds harvested at 10-15 days after flowering were separated by SDS-PAGE (panel A) and subjected to Western blot analysis (panels B and C). The gel was stained with Coomassie brilliant blue (panel A). The blots were probed by affinity-purified anti-maize BE-I and anti-RBE3 antibodies (panels B and C, respectively).



**Fig. 12. Accumulation of rice branching enzymes during endosperm development.** The seeds were harvested at various stages from 3 to 30 days after flowering (*DAF*). Northern blot analysis of total RNA was carried out using the RB13 cDNA encoding rice RBE1 as a probe (panel A). Moreover, soluble proteins were extracted from the rice seeds, and then subjected to Western blot analysis, using affinity-purified anti-maize BE-I and anti-RBE3 antibodies (panels B and C, respectively).

delayed, as compared with the accumulation pattern of the mRNA (Fig. 12, Panels B and C). Western blot analysis, using anti-RBE3 antibody, indicates somewhat earlier appearance of RBE3 than found in RBE1 during development (Fig. 12, Panel C).

## Discussion

Multiple isoforms of branching enzyme are present in developing rice seeds (Figs. 2 and 3). Three isoforms of the enzyme, RBE1, RBE2A, and RBE2B, however, are identical in molecular size, amino-terminal amino acid sequence, and reactivity with anti-maize BE-I antibody, except that RBE2A has a 3 kDa higher molecular mass than RBE1 and RBE2B (Fig. 4). The amino-terminal sequences of these three isoforms match the amino acid sequence deduced from the cDNA sequence encoding rice RBE1 (Figs. 7 and 8). These data provide evidence that rice RBE1, RBE2A, and RBE2B are products from the same gene, and correspond to BE-I in maize kernels.

Two forms of branching enzyme, BE-I and BE-II, were found in maize kernels<sup>5,11,16,17</sup>. Maize BE-I is separated from BE-II by DEAE-cellulose chromatography; BE-I is not bound to the column, but is eluted with the column wash. In the present study, rice BE-I was bound to DEAE-cellulose (Fig. 2). The similar observation was reported by Smyth<sup>7</sup>, who suggested a difference of surface charges between the rice and maize proteins. However, our data indicate that rice RBE1 shares a high degree of sequence identity with the maize protein (Fig. 9). Therefore, other factors, probably including aggregation or a hydrophobic interaction with other proteins, should account, at least in part, for the difference.

Although RBE1 is the same protein as RBE2B, as described above, these two proteins are apparently eluted with different concentrations of KCl from a DEAE-cellulose column (Fig. 2). This discrepancy may be because RBE2A is partly changed into RBE2B after elution of the RBE2A form. In fact, most of the RBE2A form appears to be converted into RBE2B after the subsequent purification on Toyopearl

HW-55F (Figs. 3, 4, and 5). Since the amino-terminal sequences of RBE2A and RBE2B are identical, and since the difference between these two forms is only in the molecular size (Fig. 4), the conversion must occur near the carboxyl terminus of the RBE2A form. In a separate study, we have constructed an expression plasmid carrying the rice RBE1 cDNA fragment at nucleotides 215-2739 (Fig. 8) in *E. coli* (see Chapter 3). Western blot analysis of the soluble extracts from the transformed *E. coli* cells showed two protein bands which are immunoreactive with anti-maize BE-I antibody and exhibit electrophoretic mobilities almost identical with those of RBE2A and RBE2B (RBE1). Thus, these data support the possibility that RBE2A is readily converted into RBE2B (RBE1) by removal of a carboxyl-terminal segment. However, it is not certain at present time whether the conversion is catalyzed by RBE2A itself, or by a processing enzyme or other factors.

Rice RBE1 possesses two amino-terminal sequences; one is two amino acids longer than another. An interesting possibility is that the heterogeneity results from a post-translational modification of the synthesized protein molecule. The alignment of the amino acid sequences of transit peptides between rice RBE1 and maize BE-I precursors, including the amino-terminal sequences of the mature proteins (Fig. 9), suggests that a cleavable site of the rice BE-I precursor is between Pro<sup>-8</sup> and Ala<sup>-7</sup> or between Ala<sup>-5</sup> and Arg<sup>-4</sup>. After cleavage of the peptide bond(s) at these positions, a proteolytic enzyme may delete the amino-terminal sequence of the mature protein. Alternatively, a processing enzyme, which is able to split the peptide bond between the rice mature protein and transit peptide, may differ in the recognition ability from the maize enzyme, resulting in formation of two amino termini; the processing enzyme for the rice RBE1 precursor may preferentially cleave the peptide bonds between Lys<sup>1</sup> and Thr<sup>1</sup>, and

between Met<sup>2</sup> and Val<sup>3</sup> rather than between Pro<sup>-8</sup> and Ala<sup>-7</sup> or Ala<sup>-5</sup> and Arg<sup>-4</sup>. These specificities are probably attributed to the stereochemical structure of the BE-I precursor, and to the charges and hydrophobicities near the cleavable region. Indeed, as compared with the maize BE-I precursor, the carboxyl-terminal sequence of the rice transit peptide at residues -4 to -1 is negatively charged (Fig. 8).

The deduced amino acid sequence of RBE1 indicates that the rice mature protein as well as the maize protein<sup>12)</sup> belongs to a family of amylolytic enzymes (Figs. 8, and 9); RBE1 conserves possible catalytic residues required for cleavage of  $\alpha$ -1,4-glucosidic linkages, including Asp, Glu, and Asp in regions II, III, and IV, respectively (residues 345, 400, and 469 in the rice RBE1 sequence, see Figs. 8 and 9, and Refs ). These data suggest that the central region of the mature RBE1 serves for cleavage of  $\alpha$ -1,4-glucosidic bonds of amylose and/or amylopectin, and that branching enzyme possesses two enzymatic functions: hydrolysis of  $\alpha$ -1,4-glucosidic linkages and addition (transfer) of the newly formed reducing-end residue to other  $\alpha$ -1,4-linked linear chains. The enzymatic properties of branching enzyme are similar to those of bacterial cyclodextrin glucanotransferase, which also belongs to a family of amylolytic enzymes, but mainly catalyzes the transfer reaction to form cyclodextrins<sup>37-40)</sup>. However, alignment of the sequences between rice RBE1 and bacterial cyclodextrin glucanotransferases shows no significant similarity in the amino- and carboxyl-terminal regions (not shown). In addition, other amylolytic enzymes share little sequence similarity in these regions, except that the 40-residue sequence of RBE1 in the amino-terminal region (residues 57-99, see Figs. 8) has a significant similarity with the sequence of *Klebsiella aerogenes* pullulanase<sup>41)</sup> at residues 293-366 (36% identity)<sup>42)</sup>. To elucidate the mechanism of the branching

reaction, it is essential to identify the functional domains of the enzyme, including those in the amino- and carboxyl-terminal regions.

## Chapter 2

Characterization of an Isoform of Rice Branching  
Enzyme, RBE3, and Its Physiological Role in Starch  
Synthesis

## Introduction

Among various endosperm mutants of plants, including maize, pea, and barley, the *amylose-extender* (high-amylose) mutant is characterized by an apparently increased content of amylose in the storage starch<sup>43)</sup>. The starch granules of this mutant are different from normal starch granules in the shapes<sup>44)</sup>, x-ray diffraction patterns<sup>45)</sup>, susceptibility to starch-hydrolyzing enzymes<sup>46)</sup>, and solubility in chemical reagents<sup>47)</sup>. In addition to these unusual properties, the *amylose-extender* starches contain two abnormal components, an amylopectin with inner and outer branches longer than those found in the usual amylopectin<sup>48-50)</sup> and an intermediate material with a lower molecular weight than that of amylose<sup>50,51)</sup>. The fine structure of the intermediate material in maize differs from those of normal amylose and amylopectin; this material has four or five branches with an average chain length of 50 that are linked to a main linear chain of 100-150 glucose units<sup>50)</sup>. Interestingly, the abnormal amylopectin and intermediate material possess high degrees of iodine binding capacity<sup>43,50,52,53)</sup>. Therefore, the *amylose-extender* starch should be characterized by its abnormal granular properties and components rather than by an increased content of amylose. For instance, a commercial high-amylose maize starch, Amylon 70, which appears to contain almost 70% amylose, is, in fact, composed of approximately 25% amylose, 20% abnormal amylopectin, and 55% intermediate material<sup>53)</sup>. In rice, Satoh and co-workers<sup>54,55)</sup> have induced various kinds of endosperm mutants and successively selected the high-amylose mutants. The starch properties of the rice high-amylose mutants in the unit-chain profile, x-ray diffractometry, photopastigraphy, and scanning electron microscopy are all consistent with those of the maize *amylose-extender* mutants<sup>55)</sup>.

The mechanism of starch biosynthesis is not fully understood<sup>2)</sup>. It is interesting to examine the role of the recessive *amylose-extender* allele in starch synthesis of plants, since the mechanism of branch formation on the amylopectin molecules is different between normal and *amylose-extender* starches. In normal maize kernels, there are three forms of branching enzyme, BE-I, BE-IIa, and BE-IIb<sup>5,10)</sup>. One of these three forms, BE-IIb is deficient in the kernels of *amylose-extender* mutant<sup>10)</sup>. Indeed, the total activity of the branching enzyme in the endosperm of the *amylose-extender* mutant is almost one-third of that present in normal maize during endosperm development<sup>17)</sup>. Hedman and Boyer<sup>56)</sup> reported that a near-linear relationship exists between the dosage of the recessive *amylose-extender* allele and the enzyme activity of the deficient BE-IIb form. On the basis of this observation, they<sup>56)</sup> concluded that the absence of BE-IIb results in the abnormal structure of the *amylose-extender* starch. However, Singh and Preiss<sup>11)</sup> have shown that the peptide map, amino acid composition, and immunoreactivity of BE-IIb are essentially identical to those of BE-IIa, implying that these two forms of maize branching enzyme are products of the same gene<sup>11)</sup>. Thus, further evidence seems to be required to establish that BE-IIb is solely responsible for synthesis of the abnormal, branched molecules in the *amylose-extender* mutant.

We have recently identified four apparent forms (RBE1, RBE2, RBE3, and RBE4) of branching enzymes in developing rice seeds<sup>14)</sup>. The amino-terminal amino acid sequence and immunoreactivity with anti-maize BE-I are identical between the 82-kDa RBE1 and 85-kDa RBE2A. The carboxyl-terminal region of RBE2A is readily processed during purification, and the amino-terminal amino acid sequence of the processed form (RBE2B) is also identical to those of RBE1 and RBE2A. Thus, RBE1, RBE2A, and RBE2B are products from the same gene.

Analysis of both genomic and complementary DNAs encoding RBE1 demonstrates that the rice branching enzyme shares a high degree of sequence identity (86%) with maize BE-I<sup>14,57</sup>). Thus, we have concluded that rice RBE1, RBE2A, and RBE2B correspond to maize BE-I (to avoid complexity, the term RBE1, denotes RBE1 itself, RBE2A, and RBE2B hereafter). An isoform of rice branching enzyme, termed RBE3, with a molecular mass of 87 kDa is distinguished from the RBE1 form by the amino-terminal sequence and immunoreactivity<sup>14</sup>). Although several forms of branching enzyme have been identified and partially characterized<sup>7,8,14</sup>), the role of each form in starch biosynthesis is not clear at the present time.

In this Chapter, I have focused on starch biosynthesis in the developing seeds of rice *amylose-extender* mutants. The results described here strongly suggest that the starch components with anomalous structures, characteristic for the *amylose-extender* mutants, are due to the lack of the RBE3 activity. The deduced primary structure of this enzyme is also reported here.

## Materials and Methods

*Materials* --- Rice cultivar, *Oryza sativa* L., cv. Kinmaze, and its mutant lines, EM10, EM72, EM129, EM145, and EM189, were cultivated at the Agronomy Farm of Kyushu University, Fukuoka, in 1992. Starches from the mature seeds of these *amylose-extender* mutants possessed an apparently increased content of amylose. Another strain of normal rice plant (cv. Koshihikari) was field-grown at the Life Science Laboratory of Mitsui Toatsu Chemicals, Mobara, Chiba. The seeds were obtained at various stages after flowering and stored at -80°C. Nitrocellulose and nylon (Hybond-N) membranes were purchased from Advantec (Tokyo) and Amersham Corp., respectively. Restriction endonucleases, T4 DNA ligase, and T4 polynucleotide kinase were purchased from Nippon Gene (Toyama). Radioisotopes, [ $\alpha$ -<sup>32</sup>P]dCTP (3,000 Ci/mmol) and [ $\gamma$ -<sup>32</sup>P]ATP (4,000 Ci/mmol), were purchased from Bresatec Ltd. (Australia). A cDNA library in  $\lambda$ gt11 was prepared from polyadenylated RNA of rice seeds harvested 10-15 days after flowering, as described previously (see Chapter 1)<sup>12,14</sup>. All other reagents were of the highest purity available.

*Antibodies* --- An 85-kDa form of rice branching enzyme, termed RBE1, was separated by SDS-PAGE, as described below, and the gels were stained with Coomassie Brilliant Blue R-250. The protein band corresponding to RBE1 was cut from the gels, passed through a 21-gauge needle, and suspended in 0.5 ml of phosphate-buffered saline. The suspension was emulsified by sonication with 0.5 ml of Freund's complete adjuvant (Difco), and injected intradermally into rabbits (New Zealand White). The injection procedure was repeated twice at 2-week intervals. The final injection was carried out using Freund's incomplete

adjuvant. Antisera were collected every 7 days after the final injection. The antibodies were purified by fractionation with ammonium sulfate (0-40% saturation) followed by immunoaffinity chromatography on a column of Sepharose 4B that had been substituted with RBE1 by the cyanogen bromide procedure<sup>58</sup>. The antibodies retained on the column were eluted with 0.2 M glycine/HCl, pH 2.2, immediately neutralized with 0.4 M potassium phosphate, pH 8.0, and then dialyzed against phosphate-buffered saline. Affinity-purified antibody against RBE3 was obtained described previously. GBSS was purified from maize starch granules by the method of Shure *et al.*<sup>59</sup> Antiserum to maize GBSS was produced by intradermal injection of the purified protein into rabbits.

*SDS-PAGE and Western Blot Analysis* --- Rice seeds were ground in 125 mM Tris/HCl, pH 6.8, containing 4% SDS, 4 M urea, 20% glycerol, and 5% 2-mercaptoethanol (0.5 ml of the buffer/seed) in a mortar and pestle. The homogenate was allowed to stand overnight at room temperature and was centrifuged at 13,000 rpm for 5 min. The supernatant solution was boiled for 3 min and then subjected to SDS-PAGE. The gels were stained with Coomassie Brilliant Blue R-250 (Sigma) or silver-stained using a commercial kit from Wako Pure Chemical Industries (Osaka). For Western blot analysis<sup>20</sup>, proteins were separated by SDS-PAGE and transferred onto Immobilon-P PVDF membranes (Millipore), using a Sartorius Semi-Dry electroblotter. The blot was blocked with 1 % skim milk, probed by affinity-purified antibody against rice RBE3 at room temperature for 1 h, and then incubated with goat anti-rabbit IgG horseradish peroxidase conjugate (Jackson Immunoresearch Laboratories). The immunoreactive bands were detected using an ECL Western blotting detection kit from

Amersham according to the manufacture's protocol. After detection, the probe was removed by incubation of the blot in 50 mM Tris/HCl, pH 6.8, containing 2% SDS and 0.1 M 2-mercaptoethanol at 50°C for 30 min. The blot was then reprobod with anti-RBE1 antibody followed by anti-GBSS antibody, described above.

*Assay for Enzyme Activity* --- Frozen rice seeds were homogenized at 4°C in 50 mM Tris/HCl, pH 7.5, containing 5 mM EDTA and 5 mM 2-mercaptoethanol (0.1 ml/seed). The homogenate was centrifuged at 12,000 rpm for 10 min at 4°C. As described previously<sup>5,14)</sup>, branching enzyme activity in the supernatant was assayed by the stimulation of phosphorylase *a*-catalyzed  $\alpha$ -D-glucan formation from  $\alpha$ -D-glucose 1-phosphate (Assay A). One unit of enzyme activity was defined as 1  $\mu$ mol of D-[<sup>14</sup>C]glucose incorporation from  $\alpha$ -D-[<sup>14</sup>C]glucose 1-phosphate per min under the condition described<sup>14)</sup>.

*Analysis of Amino-Terminal Amino Acid Sequence* --- Highly purified RBE3 was further purified by SDS-PAGE and transferred onto an Immobilon-P membrane, as described above. The membrane was stained with Commassie Blue. The protein band corresponding to RBE3 was cut from the membrane and analyzed by automated Edman degradation using an Applied Biosystems pulse-liquid Sequencer (model 477A/120A) equipped with an on-line phenylthiohydantoin analyzer.

*Isolation of cDNA Clones* --- A 38-mer oligonucleotide, 5'-ACIGGCATICC(A/G)TC(A/G)CA(C/T)TCICC(C/T)TCIGGIATCATIAC-3', corresponding to the amino-terminal amino acid sequence of RBE3 was synthesized using a DNA synthesizer (Applied Biosystems Inc.) and labeled with T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP. A cDNA library

of developing rice seed in  $\lambda$ gt11 was screened by the plaque hybridization method<sup>21)</sup> as described previously<sup>14)</sup>. Hybridization was carried out at 50°C for 16 h using the <sup>32</sup>P-labeled oligonucleotide as a probe. Positive clones were plaque-purified, and the cDNA inserts were subcloned into the *Eco*RI site of pUC19 or pUC119 for further characterization.

*Nothern Blot Analysis* --- Total cellular RNAs (20  $\mu$ g) prepared from various tissues<sup>12)</sup> were separated on 1.2% formaldehyde-agarose gels and transferred onto Hybond-N nylon membranes according to the manufacturer's protocol. The blots were probed by DNA fragments that had been labeled with [ $\alpha$ -<sup>32</sup>P]dCTP, using a random-priming DNA labeling kit (Nippon Gene). After washing, the blots were dried and analyzed using a BAS2000 Bio-Image analyzer (Fuji Photo Film Co., Tokyo)

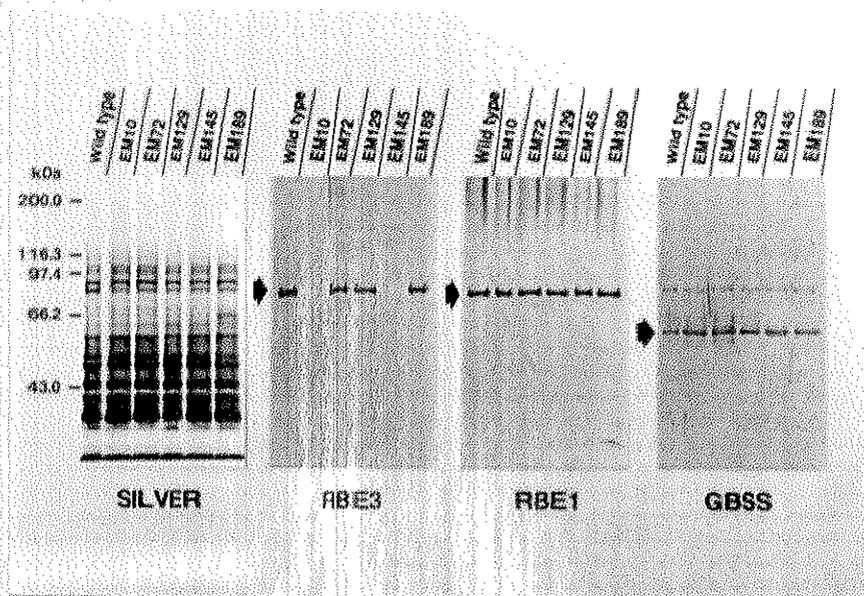
*Analytical Procedures* --- Nucleotide sequence analysis was carried out by the dideoxy chain termination method<sup>26)</sup>, using a *Bca*BEST dideoxy sequencing kit from Takara Shuzo (Kyoto). Protein concentration was determined by the method of Bradford<sup>25)</sup> using a commercial kit of Commassie protein assay reagent from Pierce Chemical Co. Computer-aided analysis of nucleotide and protein sequences was carried out using the GENETYX program (Software Development Co., Tokyo).

## Results

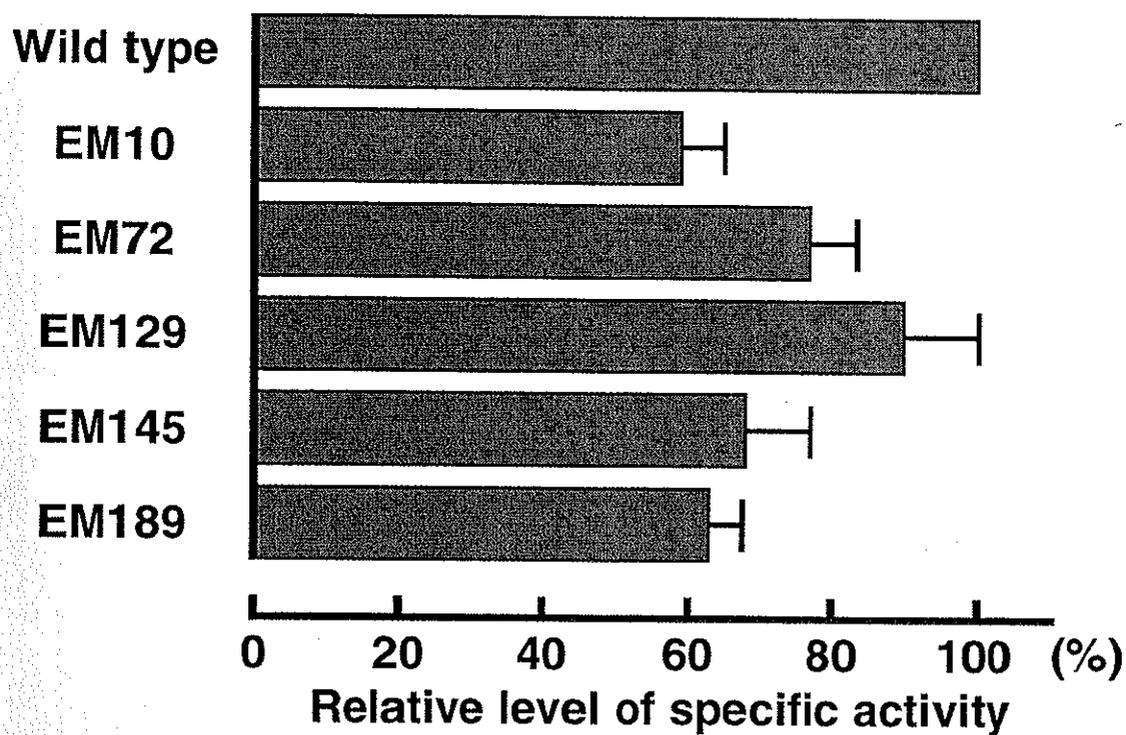
*Characterization of amylose-extender Mutants of Rice* --- To examine effects of starch-synthesizing enzymes on *amylose-extender* mutants of rice, crude extracts from the mature seeds were subjected to Western blot analysis (Fig. 13). When the blot was probed by antibody against RBE1 or GBSS, the immunoreactive bands corresponding to 85-kDa RBE1 and 60-kDa GBSS were detected in wild type and five *amylose-extender* mutants. Both the intensities of the immunoreactive bands and molecular sizes of the corresponding proteins were almost identical among these six samples. Two mutant lines, EM10 and EM145, gave no immunoreactive band when affinity-purified anti-RBE3 antibody was used as a probe. Thus, the seeds of the rice *amylose-extender* mutants lack the branching enzyme isoform, RBE3. However, the bands corresponding to 87-kDa RBE3 were found in another three mutants, EM72, EM129, and EM189, as well as in the wild type. The reason for this discrepancy is discussed under "Discussion."

As shown in Fig. 14, the activity of branching enzyme in soluble extracts of developing seeds harvested at 13-15 days after flowering was measured. The enzyme activity included those of RBE1 and RBE3. The level of the specific activity in the *amylose-extender* mutants was approximately 60-90% of that in the wild type. Western blot analysis of the soluble extracts from the developing seeds, using anti-RBE1 or anti-RBE3 antibody, gave similar results as in Fig. 13 (data not shown). Thus, the decrease of the enzyme activity in the *amylose-extender* mutants probably reflects the lack of the RBE3 activity.

*Molecular Cloning of an Isoform of Starch Branching Enzyme, RBE3* --- Highly purified RBE3 was separated by SDS-PAGE,



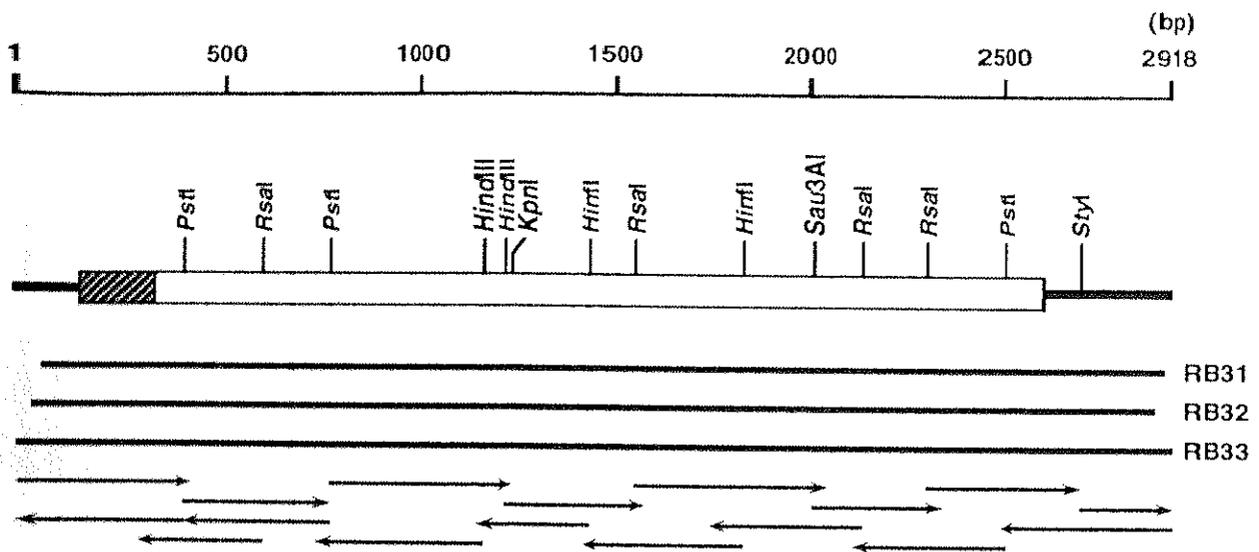
**Fig. 13.** Two out of five *amylose-extender* mutants of rice lack a branching enzyme isoform, RBE3, in the seeds. Proteins (5  $\mu$ g each) in mature seed extracts of wild type and *amylose-extender* mutants, EM10, EM72, EM129, EM145, and EM189, were separated by SDS-PAGE and subjected to Western blot analysis as described under "Materials and Methods." Proteins on the gel were visualized by silver staining (*SILVER*). The blot was probed by affinity-purified anti-RBE3 antibody (*RBE3*) and then re-probed by anti-RBE1 antibody (*RBE1*) followed by antibody against granule-bound starch synthase (*GBSS*). The immunoreactive bands corresponding to the 87-kDa RBE3, 85-kDa RBE1, and 60-kDa GBSS are indicated by *arrows*.



**Fig. 14.** A decreased level of branching enzyme activity in developing seeds of *amylose-extender* mutants. Soluble extracts from developing seeds, harvested at 13-15 days after flowering, of wild type and five *amylose-extender* mutants were assayed for branching enzyme activity, as described under "Materials and Methods." Protein concentrations in the extracts were also determined to calculate specific activity. *Error bars* indicate S.D. ( $n \geq 4$ ).

electrophoretically transferred onto a PVDF membrane, and subjected to amino-terminal amino acid sequence analysis. The amino-terminal sequence was determined to be Ala-Ala-Gly-Ala-Ser-Gly-Glu-Val-Met-Ile-Pro-Glu-Gly-Glu-Ser (or Cys) -Asp-Gly-Met-Pro-Val-. This sequence shared no significant similarity with the amino-terminal sequences of RBE1 and *E. coli* branching enzyme. To isolate cDNA clones coding for RBE3, a 38-mer oligonucleotide corresponding to the amino-terminal amino acid sequence of RBE3 at residues 8-20 was synthesized and used to screen approximately  $3 \times 10^5$  plaques from a developing seed cDNA library of normal rice plant. Three positive clones, termed RB31, RB32, and RB33, were obtained. Restriction mapping (Fig. 15) and Southern blot analysis of these cDNA inserts demonstrated that they were all related. Since the inserts of these three cDNA clones were almost identical in length (approximately 2.9 kbp), RB32 was selected for complete nucleotide sequence analysis. The cDNA inserts of RB31 and RB33 were partially sequenced.

The composite nucleotide sequence of the overlapping cDNA inserts encodes a 2,475-nucleotide open reading frame that is flanked by 5'- and 3'-untranslated regions of 127 and 316 nucleotides, respectively (Fig. 16). The translation initiation site is assigned to an ATG initiator codon at nucleotides 128-130 because of the similarity of the sequence surrounding this codon to the eucaryotic consensus sequence described by Kozak<sup>27)</sup>. In fact, an in-frame stop codon, TAG, is located 57 nucleotides upstream from the ATG codon. There is no typical sequence of a polyadenylation signal, AATAAA, near the 3'-end. The open reading frame encodes a polypeptide of 825 amino acids with a calculated molecular mass of 92,756 Da. The deduced amino acid sequence at residues 1-20 matches the amino-terminal sequence of RBE3 determined by protein analysis, as described before, confirming the



**Fig. 15.** Restriction map and sequencing strategy for RBE3 cDNAs. The scale at the top designates nucleotide position in base pairs from the 5'-end of RB33. The *open box* represents the mature protein coding region, and the *hatched box* indicates a putative transit peptide required for translocation of branching enzyme into amyloplast. *Arrows* indicate the direction and extent of nucleotide sequence determined from each site.

1	CGCGG CACACCCACA CACCACCAC CAGGCGCG	35
36	CTCCTCGCTT TGGCTCTCG GTGAGGAGG TTTAGTGG A GCAGAGCG GGGGGTTGCC GGGGGATCCG ATCCGGCTGC GGTGGGGCG	125
126	AG ATGGCGGCGCGGGGCTTCGGGTTCCGGGAGCGGGGGGGCTACGGCGGGGGCGTGGGTTCCCGCGTCCAGCCGGGGCCCGAGCTGGCGT	224
-65	M A A P A S A V P G S A A G L R A G A V R P P V P A G A R S W R	-32
225	GCGGCGGCGGACCTCCGAGCTCCGGTCCGCTGCTCCGGCCGAGATTCGCCGTTCCGGTTCGGGTTCCGGGGGCGCGTGGCGCTGCGC	322
-33	A A A E L P T S R S L L S G R R F P G A V R V G G S G G R V A V R	-1
323	GCGGCGGCGGCTCAGGAGGATGATGCCCGAGGCGAGAGCGAGGATGCCGTTTCAACAGGTTTCAGACGATCTGCAGTTGCCAGCTTAGAT	421
1	<u>A A G A S G E V M I P E D E S D G M P V</u> S A G S D D L Q L P A L D	33
422	GATGAATTAAGCACGGAGTTGGAGCTGAAGTTGAGATTGAGTCATCTGGACCAAGTACGTTGAAGCCGTTGAAGAGATGGTTCAAGAAATTAAGTGCCT	520
34	D E L S T E V G A E I E S S G A S D V R G V K R V V E E L A A	66
521	GAGCAGAAACCAGGATGTTCCACCACACAGGAGATGGGCAAAAATATTCACAGTGGACTCTATGCTTAATGGCTATAAGTACCATCTTGAATATCGA	619
67	E Q K P R V V P P T G D G Q K I T P Q M D S M L N G Y K Y H L E Y R	99
620	TATAGCTATATAGGAGACTGCGTTTCAGACATTTGATCAGTATGAAAGGAGACTGGAACATTTTCTCGCGTTATGAGAAGTTGGATTAAATCACAGT	718
100	Y S L Y R R L R S D I D Q Y E G G L E T F S R G Y E K F G F N H S	132
719	GCTGAAGGTGCTCACTTATCCAGAAATGGGCTCCCGGGCACATTTCTGAGCATTAGTAGGTGACTTCAACAATTTGAAATCCAAATGCAGACCCCATGAGC	817
133	A E G V T Y R E W A P G A H S A A L V G D F N N W N P N A D R M S	165
818	AAAAATGAGTTGGTGTGGGAGATTTTCTGCCTAACAACTGCTGAGGCTCACTCTCTATTCACATGCGTCAAGTCTAAAGGTGCGAATGSAAACT	916
166	K N E F G V W E I P L P N N A D G S S P I P H G S R V K V R M E T	198
917	CCATCTGGTATAAAGGATTTCTTCTGCTCGCTGATCAAGTACTCTGTGCGAGCCGAGGAGAAATCCCATACAATGGAATATATTATGATCCCTCGTAA	1015
199	P S G I K D S I P A W I K Y S V Q A A G E I P Y N G I Y Y D P P E	231
1016	GAGGAGAAGTACATATCAAGCATCTCAACCTAAAAGACCAAAGTCAATGCGATATACGAAACTCATGTTGGAATGAGTAGCCAGGACGCAAGATC	1114
232	E E K Y I F K H P Q P K R P K S L R I Y E T H V G M S S T E P K I	265
1115	AACACGTATGCAAACTTTAGGGATGAGGTGCTTCCAAGAATCAAAAAGCTTGGATACAATGCGAGTCAAAATATGGCAATTCAGAGCATGCATATTAT	1213
265	N T Y A N F R D E V L P R I K K L G Y N A V Q I M A I Q E H A Y Y	297
1214	GGAAGCTTTGGTACCATGTCACCAATTTCTTTCACCAAGTAGTCTGTTGGGACCCAGAAAGATTTAAAGTCATTGATTGATAAAGCTCATGAGCTT	1312
298	G S F G Y H V T N F F A P S S R F G T P E D L K S L I D K A H E L	330
1313	GGTTTACTTGTCTCATGGATGTTGTTTCACAGCCATGCGTCAAATAATACCTAGATGGTTGAACGTTTGGATGGTACAGATACCCATTTACTTTCAT	1411
331	G L V L M D V V H S H A S N N T L D D G L N G F D G T D T H Y F H	363
1412	AGTGGTTCACGCGCCATCTGGATGTTGGGATCTCGCCTTTTCAACTATGGGAATGGGAAGTCTAAGATTTCTACTATCCAATGCAAGATGGTGG	1510
364	S G S R G H W M W D S R L F N Y G N W E V L R F L L S N A R W W	396
1511	CTCGAGGATATAAGTTTGGATGGTTTTCAGATTGACCGTGTAACTCAATGATGACACTCATGATGCAAGTACAGTACAGTACAGTACAGT	1609
397	L E E Y K F D G F R F D G V T S M M Y T H H G L Q V A F T G N Y S	429
1610	GAATACTTTGGATTTGCCACTGATGCTGATGCTGATGCTTACTTGAAGCTGGTAAATGATTTAATTCATGAGCTTTAAGTGGCCATAACCATCGGT	1708
430	B Y F G F A T D A D A V V Y L M L V N D L I H G L Y P E A I T I G	462
1709	GAAGATGTCAGTGGAAATGCTTACATTTGCCCTTCTGTTTCAAGATGGTGGGTTGGTTTGGATTATCCGCTTCAATGGCTTCCCTGCACAAATGGATT	1807
463	<u>E D V S G M P T F A L P V Q D G V G F D Y R L H M A V P D K W I</u>	495
1808	GAACTCTCAAGCAAAGTCAATGAACTTGGAAAGTGGTGTATGTTGTCACACACTGACTAACAGAAGGTGCTCAGAGAAGTGTGTTACTTATGCTGAA	1906
496	E L L K Q S D E S W K M G D I V H T L T N R R W S E K C V T Y A E	528
1907	AGTCATGATCAAGCAGTACTGGTGACAAAATATTCATCTGCTGGTTGATGACAGGATATGATGATTTTATGGCTCTGGACAGACCGGCAACACCT	2005
529	<u>S H D Q A L V G D K T I A F W L M D K D M Y D F M A L D R P A T P</u>	561
2006	AGCATTCATCGTGAATAGCATTCATATAAAATGATTAGACTTATCAACAATGGGTTAGGAGGAGAGGCTATCTTAACTTTATGGAAATGAGTTCCGA	2104
562	S I D R G I A L H K M I R L I T M G L G G E G Y L N F M G N E F G	594
2105	CATCTGAATGGATTGATTTTCCAGAGCTCCACAAGTACTTCCAAATGGTAAATTCATCCAGGGAATAACAACAGTTATGATAAATGCCCTCGAAGA	2203
595	H P E W I D F P R A P Q V L P N G K F I P G N N N S Y D K C R R R	627
2204	TTTGACCTGGGTGATGGGACTATCTTAGGTATCGTGGCATGCTAGAGTTTGACCCCGGATGCGAGTCTCTCGAGGAAAAATGGGTTTCATGACATCA	2302
628	F D L G D A D L E S R W L E F D R A M Q S L E E K Y G F M T S	660
2303	GACCACAGTACATATCTGAAAGCATGAAGAGGATAAGATGATTATATTTGAGAAGGAGATCTGGTATTTGTGTTCAACTTCCATTGGAGTAAACAGC	2401
661	D H Q Y I S R K H E E D K M I I F E K G D L V F V F N F H W S N S	693
2402	TATTTGACTACCGTGGTTGGTTTAAAGCCAGGAAATATAAGTGGTCTTGGACTCAGATCTGGACTCTTTGGTGGATTGGCAGGATCCATCAC	2500
694	Y F D Y R V G C L K P G K Y K V V L D S D A G L F G G F G R I H H	726
2501	ACTGCAGAGCACTTCACTCCGATTTGTTACATGACAAACAGCCCTACTCGTTCAGTTTATTTCTCAGAGAACCTGCTGCTGCTATGCTCCAGCG	2599
727	T A E H F T A D C S H D N R P Y S F S V Y S P S R T C V V Y A P A	759
2600	GAATGA GAAC ACCAAGAGGC AGCATGCAAG TGTGTGCGGC TGCTAGTGGC AAGGAGCAAG AAAAAGTACT TGCCAGCAAT CTGTGAACCG	2689
760	E *	
2690	CTTCTAGG TTTGCTTCG ATGAATGCCG GATAGACTAG ACAGTTGCT TTTGTCGTT GCGTCCCAA TTTGATGTT TAGTTGTGA	2779
2780	GGGAAAGAAA CGTTTATTG TAATTATCTA TGCTGTGCA ACGGCGACGA AACCATGAAC CCGTATATT TGTGGTACC GTTCGAACTG	2869
2870	CCAGTTATAC ATAGTTCTGC ACTTCTGAC ATCTTGTGAT GCTTGAATC	2918

**Fig. 16. Nucleotide and deduced amino acid sequences of the cDNA encoding an isoform of rice branching enzyme, RBE3.** The deduced amino acid sequence is shown below the nucleotide sequence numbered in the 5' to 3' direction. The amino-terminal amino acid sequence of RBE3, determined experimentally, is *underlined* with a wavy line. Amino acids are numbered from the amino terminus of the mature protein, and the residues at the amino-terminal side from residue 1 are represented by *negative numbers*. The conserved sequences of RBE3 with the four regions of the catalytic sites of amylolytic enzymes are *underlined*.