Starch Branching Enzymes and their Genes in Sorghum

Joel Mutisya
Department of Plant Biology and Forest Genetics
Uppsala

Doctoral thesis
Swedish University of Agricultural Sciences
Uppsala 2004

Acta Universitatis Agriculturae Sueciae
Agraria 479

ISSN 1401-6249
ISBN 91-576-6768-3

Tryck: SLU Service/Repro, Uppsala 2004

Abstract

Mutisya, J.M. 2004. *Starch branching enzymes and their genes in sorghum*. Doctor's dissertation.
ISSN 1401-6249, ISBN 91-576-6768-3

Starch is an important raw material both for food and non-food purposes. It is synthesized and stored in source and sink tissues in plants. The starch deposited in amyloplasts of storage tissues possesses several physico-chemical properties, which makes it desirable for diverse applications. For industrial applications, starches high in either amylose or amylopectin are preferred in order to minimize chemical modification. The main purpose in our research is to understand starch synthesis in the sorghum endosperm, and use this information to generate transgenic sorghum with novel starches for both human diet and industrial application. In starch biosynthesis, the activity and expression profile of starch branching enzymes is important in determining both the yield and proportion of amylose and amylopectin. This thesis describes some of the characteristics of starch branching enzymes and their genes in sorghum.

The genes encoding two SBE isoforms, SBEIIa and SBEIIb, were cloned and characterised. Comparison of the SBEIIb amino acid sequence with sequences from related species revealed a conserved core stretch of amino acids believed to harbor the catalytic site of the enzyme. Spatial expression patterns showed that the activity of sorghum *sbeI* and *sbeIIb* are seed-specific, while sorghum *sbeIIa* is expressed in endosperm, embryonic and vegetative tissues. We demonstrated that endosperm-specific expression of barley *sbeIIb* is governed by the second intron of the barley *sbeIIb* gene, and that the same second intron can confer endosperm specific expression to both sorghum *sbeIIb* and barley *sbeIIa*. This implies that a similar regulation factor in barley, which interacts with sequences within the intron, is also present in sorghum. Deletion mutagenesis of the sorghum *sbeIIb* promoter showed that a, 1-kb fragment 5′ of the transcriptional start site constitutes the minimal promoter capable of high level expression. Diurnal expression profiles suggested that an endogenous oscillator in the endosperm controls expression of *sbeI*, *sbIIa* and *sbeIIb*.

Keywords: amylose, amylopectin, expression, regulation, second intron, transgenic sorghum.

Author′s address: Joel Mutisya. Department of Plant Biology and Forest Genetics, SLU, P. O. Box 7080, SE-750 07 Uppsala, Sweden.

Table of contents

To my parents

Appendix

Papers I-IV

The present thesis is based on the following papers, which will be referred to by their Roman numerals:

I. Mutisya, J., Sathish, P., Sun, C., Andersson, L., Ahlandsberg, S., Baguma, Y., Palmqvist, S., Odhiambo, B., Åman, P. & Jansson, C. 2003. Starch branching enzymes in sorghum (*Sorghum bicolor*) and barley (*Hordeum vulgare*): Comparative analysis of enzyme structure and gene expression. *Journal of Plant Physiology* 160, 921-930.

II. Mutisya, J., Sun, C., Palmqvist, S., Baguma, Y., Odhiambo, B. & Jansson, C. Transcriptional regulation of the *sbeIIb* genes in sorghum (*Sorghum bicolor*) and barley (*Hordeum vulgare*): Importance of the barley *sbeIIb* second intron. Submitted.

III. Mutisya, J., Sun, C., Venkateshaiah, L., Palmqvist, S., Baguma, Y., Odhiambo, B. & Jansson, C. Endogenous oscillation of *sbe* expression in sorghum endosperm. Submitted.

IV. Mutisya, J., Sun, C., Boren, M., Baguma, Y., Odhiambo, B. & Jansson, C. Comparative analysis of starch synthesis in white and brown sorghum: An initial characterization. Manuscript.

Other publications

1. Baguma, Y., Sun, C., Ahlandsberg, S., Mutisya, J., Palmqvist, S., Rubaihayo, P.R., Magambo, M.J., Egwang, T.G., Larsson, H. & Jansson, C. 2003. Expression patterns of the gene encoding starch branching enzyme II in the storage cassava (*Manihot esculenta* Crantz). *Plant Scienc*e 164, 835-839.

2. Baguma, Y., Sun, C., Boren, M., Olsson, H., Palmqvist, S., Mutisya, J., Rubaihayo, P.R. & Jansson, C. 2004. Sugar-mediated semidian oscillation of starch synthesis genes in the cassava storage root. *Submitted.*

Abbreviations

ADPglucose	adenosine diphosphate-D-glucose
AGPase	adenosine glucose pyrophosphorylase
BAC	bacterial artificial chromosome
DAP	days after pollination
DBE	debranching enzyme
GBSS	granular bound starch synthase
GFP	green fluorescent protein
SBE	starch branching enzyme
SS	starch synthase
SSS	soluble starch synthase

1. Introduction

1.1. Overview of Sorghum

Sorghum is the fourth most important cereal crop trailing behind rice, maize and wheat. However, it is ranked second to maize in supply of grain requirement within sub-Saharan Africa. Sorghum lags behind, compared to other cereals, in applications of biotechnological tools to improve yield, grain quality and yield stability. Over the years, conventional breeding has only offered limited improvement in sorghum production but its potential in grain productivity is yet to be exploited. Historically, sorghum was domesticated some 5000 years ago in Ethiopia and spread to all the six continents (Fig. 1). Available records indicate that world sorghum production was approximately 60 million tonnes in 2000 (FAO; http://www.fao.org/ag7magazine/0202sp2 accessed 2nd July 2004), most of it being produced in developing countries for food, feed and non-food applications. In most of these countries, sorghum is the main source of starch for human diet. To date, over 500 million people in the developing countries depend on sorghum as the main staple food. In other countries, sorghum starch is mainly used in livestock feed formulations and as a cheap source of raw material for industrial applications.

Anatomically, a mature sorghum grain has three basic features, namely, pericarp (6%), germ (10%) and endosperm (84%) (Hubbard et al., 1950). The pericarp is the outermost structural component and is composed of layers of thick cutin and pigments that give the grain a characteristic colour of brown or white. Most, if not all, biosynthesis and granule formation of storage starch occurs within the endosperm zones. Indeed, a direct comparisson between sorghum and maize starch was made to analyse and quantify its proportion in a grain. A starch analysis revealed that sorghum contains high amounts of starch per grain, constituting 63-68% of the grain volume, while maize contains 60-64% starch per grain. However, sorghum starch is less digestible compared to maize starch (Dreher, Dreher & Berry, 1984) which could be partly due to high levels of protein bodies and tannins which encapsulate starch granules, making the starch molecules partially inaccessible to enzymes responsible for degradation. Although starch granules within the peripheral cells of endosperm tissue are densely packed and embedded into protein bodies, these protein bodies decrease as starch content increases from the outer to the central region of the endosperm (Chandrashekar and Kirleis (1988). It has been shown that processing or cooking sorghum grain enables the proteins to impact positively on quality characteristics of functional foods such as composite breads, which have no gluten proteins abundant in wheat flour (Bugusu et al., 2001). These authors observed that the sorghum proteins have great potential as functional component in baking products and could be improved further through protein engineering.

The potential value of sorghum, a C4 plant, is derived from its ability to grow in marginal areas lacking sufficient moisture and fertility unfeasible to support production of maize, wheat or rice. Consequently, sorghum holds the potential to supply a greater share of the world's grain demand. This congruency makes

expansion of sorghum starch production and utility feasible as a main alternative to maize starch for food and non-food products. Besides, sorghum is one of the most versatile crops in terms of its utility, for example, the boiled grains is a perfect source of dietary fibre, flour for porridge, ugali, beer, cakes, chapatis animal feed, sugar syrups, liquid fuel and firewood from stem and roots.

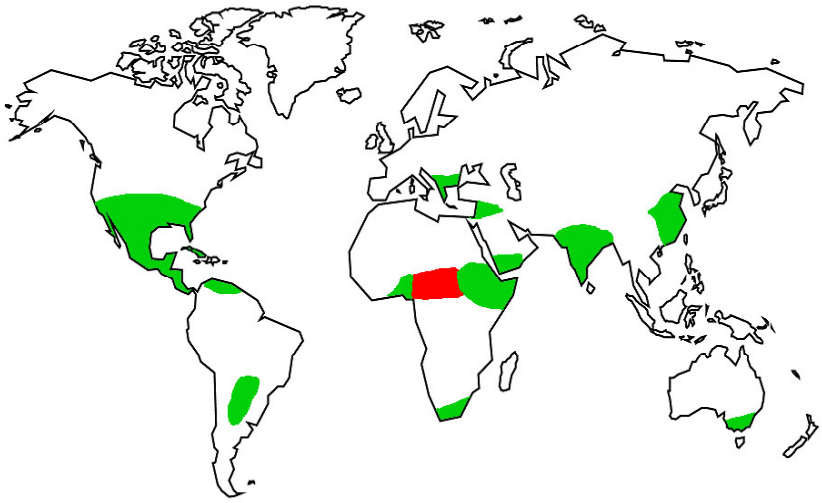

Figure 1. Schematic drawing of the global sorghum growing areas. Area marked with red represent the centre of origin and areas marked in green represents the main producing regions.

In recent years, sorghum starch has been used in industries such as pharmaceutical, textile and paper. In India, the second largest producer of sorghum in the world, the sorghum grain is largely used for animal feed, starch and alcohol industry. Development of technologies for industrial application of grain starch, such as biodegradable plastics, cosmetics, adhesives, surfactants, glues and as a biodegrade in fertilizer industries, will ultimately make sorghum starch more competitive in the industrial sector.

Based on these enormous benefits treasured by sorghum, current molecular approaches can be utilized to impact sorghum productivity for development and more importantly, contribute to improvement of livelihood of millions of farmers living in the marginal rural areas of the developing world.

1.2. Endosperm development

Endosperm is a complex plant organ comprising starchy endosperm, aleurone layer, transfer cells and cells surrounding the embryo. It is the principal heterotrophic storage organ for starch and other deposits of carbon in cereal plants. In the recent years, research efforts have been expended in understanding

the development of endosperm and its importance in anabolism of the developing seed. On this respect, several proposals have been advanced to explain the evolutionary origin of the endosperm, but the general agreement is that the nuclear endosperm originated from a modified second embryo. The endosperm has evolved over the years as a specialized independent storage organ (William & Friedman, 2002).

Initiation of endosperm development starts with a sequential process of double fertilization, resulting in a triploid organ, commonly referred as nuclear endosperm (Olsen 2001; Berger 2003). Starchy endosperm is the main and most important tissue in accumulation of plant products of economic importance such as starch, proteins and lipids. In general, the endosperm undergoes four major developmental stages; sycytrial, cellularization, differentiation and maturation. In early stages, during cellularization, division of nuclei characterizes the endosperm in absence of cell wall between the daughter nuclei results in multinucleated cells in contrast to single nuclei in somatic cells. This process continues for approximately three days in barley (Olsen 2004) and 5 days in wheat (Lan et al., 2004) before the nuclear division ceases and denucleation starts. Denucleation is the first stage of apoptosis when programmed cell death in endosperm starts in addition to mitochondrial degeneration. Lan et al. (2004) showed that, during nuclear degradation, starch synthase and starch branching enzymes in rice show an increased enzymatic activity, which implies that starch accumulation and endosperm cell initiation are distinct separate processes in endosperm development. The starchy endosperm cell differentiation leads to formation of two distinct morphological structures, inner endosperm cells and inner daughter cells of the aleurone layer, both of which are major sites of reserve storage (Jane, 2004). During grain development, endosperm cells are gradually filled with protein bodies and amyloplasts, the main site of starch biosynthesis. Several amyloplasts surrounds each nuclear in starch endosperm cells. In starch biosynthesis, amyloplasts essentially imports metabolites from the cytosol for ADP-glucose synthesis. This metabolite, ADP-glucose, is synthesised from glucose-1-phosphate and ATP, and it is the main precursor for starch synthesis. The ADP-glucose is consequently used as a substrate by isoforms of starch synthase enzymes in amylose and amylopectin synthesis. The pathway in starch synthesis and formation of starch granules in the amyloplasts is extensively discussed in this thesis.

In cereals, starch that accumulates in the endosperm occurs in two forms of granules, A- and B-types, shown in Fig. 2. However, the diferent granules are organised the same way. In wheat, initiation of A-type granules starts approximately four days earlier than the B-type granules (Emes et al., 2003), implying that formation of different granule types is developmentally regulated. It was suggested by Hizukuri (1986) that the crystalline and amorphous lamellae are determined mainly by arrangement of amylopectin chains within a granule (Fig. 3).

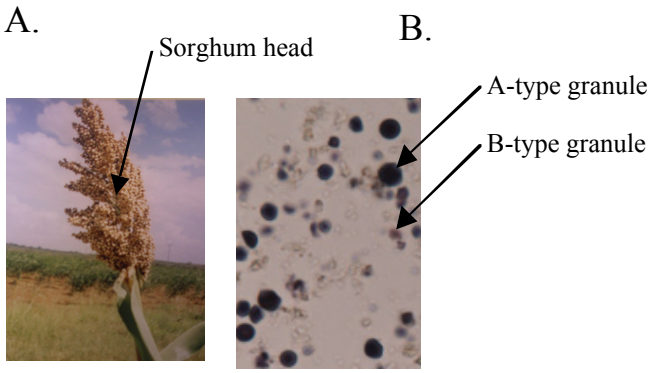

Figure 2. Schematic overview of sorghum and starch granule structure. **A**. Sorghum head showing developed grains **B**. Starch granule structure with A- and B-granules.

1.3. Starch structure and characteristics

Starch is the most abundant form of carbohydrates in the biosphere, next to cellulose. It is the primary source of energy for both animals and plants. In nearly all plants, two starch deposits exist, transitory starch and storage starch, synthesized in chloroplasts and amyloplasts, respectively.

1.3.1. Amylose and Amylopectin

Amylose and amylopectin are the two main components of starch (Fig. 3). In most plants, the total native starch consists of 20 to 30% amylose and 70 to 80% amylopectin. The proportion of amylose and amylopectin molecules and the length of glucan chains govern the size and structure of starch.

Anhydroglucose units linked in linear α-1,4 glycosidic linkages with a few α-1,6 bonds at the branching points forms amylose. However, presence of branching points does not alter the structural characteristics of amylose, which strictly remains linear. Compared to amylopectin, amylose molecules are smaller, generally made of 200 to 2000 glucose units and form a helical complex with iodine resulting in a characteristic blue colour.

The synthesis of amylose is mainly catalyzed by GBSSI in most storage tissues. Initiation of amylose synthesis requires an amylopectin chain as a primer, followed by elongation of the growing amylose chain (Denyer et al., 1996). These chains can be branched by SBE or cleaved by a hydrolytic enzyme, which determines the chain length. In cereals, mutant deficient in GBSSI activity has been identified with little or no amylose (Morrison et al., 1984). Similarly, mutants high in amylose have been reported in maize and barley and production of high-amylose potato through antisense inhibition has been attempted (Schwall et al., 2000).

Amylopectin is the main component of starch and it consists of a highly branched polyglucan. Structurally, amylopectin consists of linear α-1,4 linkages with α-1,6 bonds at the branching points. The branching points are formed approximately every 10 to 20 anhydroglucose units, while a single amylopectin molecule containing 10^3 to 10^4 anhydroglucose units. The multiplicity in branching of amylopectin leads to formation of A-, B- and C-chains as defined by Peat et al. (1952) (Fig. 3). The outer A-chains are linear and contain a potential non-reducing end. The B- and C-chains both carry one or more A- and /or B-chain (Buleon et al., 1998). Generally, there is only one C-chain containing the reducing end in each amylopectim molecule. The synthesis and packing of amylopectin molecules into granules is regulated by the enzymatic reactions governing starch biosynthesis.

1.3.2. Transitory and storage starch

Broadly, transitory starch is synthesized in chloroplasts of photosynthetic active leaves during the light period and degraded at night, resulting in diurnal fluctuation of leaf starch. Degradation of transitory starch is facilitated by the activity of phosphorylase and amylase, which are hydrolytic enzymes, to form small polysaccharides or simple sugars, which replenish metabolite requirement in sink tissues. Sucrose is the main form of sugars transported through the phloem tissue to the sink cells. It is hydrolyzed into hexose sugars, which form a part of the general metabolic pool of the cell. This inter-conversion between starch and sugars occurs during the whole life cycle of the plant (see review by Emes et al., 2003).

Storage starch is synthesised and deposited in amyloplasts of plant organs such as endosperm, fruits, embryo, tubers and roots. Unlike photosynthetic plastids, amyloplasts import or generate all assimilates and precursors for starch synthesis mainly from hexose sugars from the cytosol. However, the process of starch synthesis is similar in source and sink tissues, which is mainly controlled by four enzymatic reactions catalyzed by AGPase, starch synthases, starch branching enzymes and debranching enzymes (Smith et al., 1997). Generally, it is expected that the differential expression pattern of these enzymes determines the type of starch polymers synthesized in various plant tissues. Accordingly, precise activity of the exzymes involved in starch biosynthetic pathway generates amylose and amylopectin molecules, the principal components of starch. Comparatively, transitory starch has been shown to contain relatively high amylose:amylopectin ratio (Matheson, 1996). This phenomenon is largely because of reduced or lack of branching activity of enzymes such as SBEI and SBEIIb in photosynthetic tissues.

1.3.3. Granule structure and characteristics

Starch granules are synthesized in a semicrystalline form. The granules vary in size and shape and also differ in amylose:amylopectin ratio depending on their source. In potato, the granules measure upto 75 μm in diameter while sorghum has granules that are 3-27 μm in diameter. In cereals such as maize, barley and rye, starch granules are packed with a bimodal size distribution, with large lens-shaped

A-granules and small spherical B-granules (Fig. 2) (MacGregor & Fincher, 1993; Ellis et al., 1998). However, all the granules are formed in a similar pattern, with non-reducing ends of a starch components pointing outwards in an organized arrangement, leading to formation of crystalline and amorphous lamellae, alternating periodically at approximately 9 nm (Smith, Denyer & Martin, 1997). Generally, the structure of the granule is determined by the mode of packing of amylopectin molecules. A high degree of organization of amylopectin molecules and clustering of linear chains, permit the chains to align in parallel positions, forming a crystalline lamella. Consequently, the branching points of the amylopectin molecule forms the amorphous portion of the granule (Fig. 3).

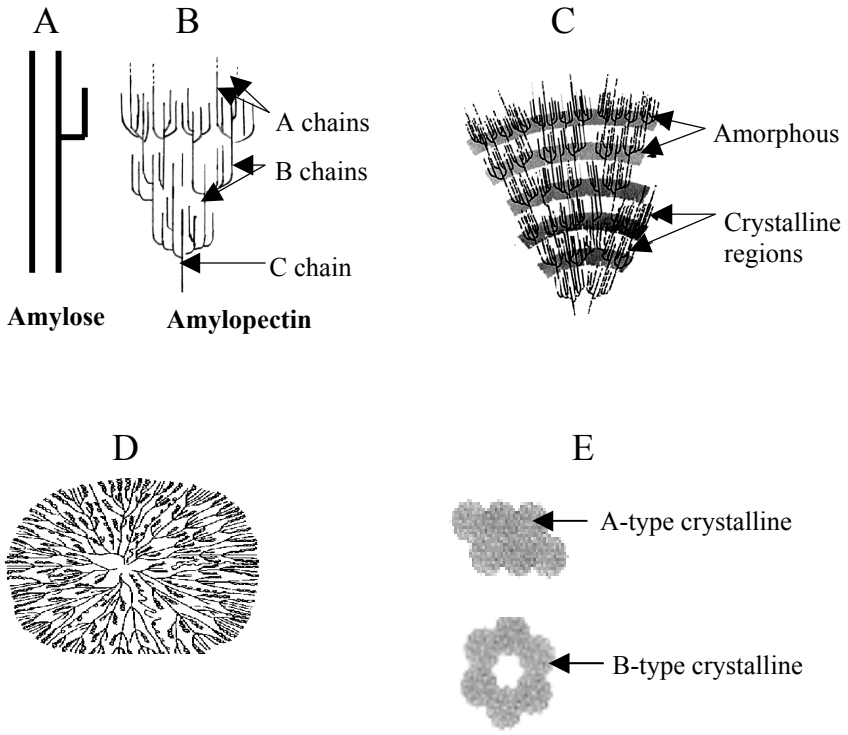

Figure 3. Schematic representation of different levels of starch granule arrangement. **A**. Amylose structure **B**. Amylopectin molecule showing A-, B- and C-chains. **C**. Arrangement of amylopectin clusters forming alternating crystalline and amorphous regions. **D**. Cross section of entire granule showing packing arrangement of amylopectin and amylose and **E.** Crystalline arrangement of double helices of A- and B-type crystallites.

In starch granules, there are two main types of crystalline structure, A- and B-types, which differ in symmetry and organization of amylopectin (Fig. 3). A-type, found in cereals, have densely packed short chains of amylopectin enabling the granule to accommodate minimal amounts of bound water (James, Denyer & Myers, 2003). On the contrary, B-type crystalline granules, observed in starches of potato and other tuber crops are believed to possess longer amylopectin chains than A-type, and contain cell structures that are capable of holding water molecules (Hizukuri, 1985). During the process of starch accumulation, other factors such as lipids, phosphate derivatives and environmental conditions are reported to affect the crystalline structure and, consequently, the functional properties of the starch (Buleon et al., 1998; Jane et al., 1999). The physical properties of starch are determined by the amylose:amylopectin ratio, the chain length distribution of amylopectins and complex formation with lipids and phosphates (Morrison et al., 1984).

1.4. Starch Biosynthesis

Starch is a complex polymer of glucose units linked together by glycosidic bonds. As decribed above, it consists of two main components, amylose molecules that mainly consists of linear glucose units linked by 1,4-glycosidic bonds with few 1,6-glycosidic linkages, and highly branched amylopectin molecules that have additional 1,6-glycosidic bonds. Amylopectin has distinct structural characteristics and its is synthesised by multiple isoforms of enzymes. Starch metabolism in both chloroplasts and amyloplasts is controlled mainly by five groups of enzymes: i) ADP Glucose Pyrophosphorylase, ii) Starch synthase, iii) Starch branching enzymes, iv) Debranching enzymes and v) Amylases and phosphorylases.

1.4.1. ADP Glucose Pyrophosphorylases

The first step in starch biosynthesis in all plant organs is initiated by adenosine glucose pyrophosphorylase (AGPase) that catalysis the formation of ADP-glucose, the substrate for the synthesis of starch polymers. Synthesis of ADP-glucose from glucose-1-phosphate requires ATP and results in subsequent release of inorganic phosphate (PPi). In most plant tissues, AGPase is allosterrically regulated by 3-phosphoglycreic acid as an activator and PPi as an inhibitor, a regulatory mechanism strongly believed to account for the rate limiting condition in starch biosynthesis.

The AGPase enzyme is tetrameric, consisting of two large and two small subunits, which share high similarities, but they are distinguishable by features of their amino acid sequences. A highly conserved region in the small subunit is believed essential for substrate specificity. Expression of AGPase sub-units may differ among plants. In barley, two genes encoding for the large subunit are differentially expressed in leaf and endosperm (Smith, Denyer & Martin, 1997). This suggests that the functional AGPase in leaves may differ from that in the endosperm, a phenomenon that offers selective advantage for specific target expression in genetic engineering.

The occurrence of AGPase activity in plastids is widely established. However, some studies have demonstrated additional occurrence of cytosolic AGPase activity in cereal endosperm but not in other sink tissues (Fig. 5) (Denyer et al., 1996; Beckles, Smith & ap Rees, 2001). Accordingly, it seems that cytosolic AGPase facilitates direct conversion of large proportion of sucrose in the endosperm into starch.

1.4.2. Starch synthases

Starch synthases (SS) catalyze the addition of glucosyl units to non-reducing ends, elongating an existing α1,4-glucan chain. To date, atleast six isoforms of SS have been identified in plants and are grouped into two main classes; granular bound starch synthase (GBSS) and soluble starch synthase (SSS) (Jiang et al 2004; Li et al., 2000, 2003; Tenorio et al., 2003). It is well established that GBSSI is the key enzyme responsible for amylose synthesis in sink tissues (Denyer et al., 1996; Ball, Wal & Visser, 1998; Tomlinson. & Denyer, 2003.). There is growing evidence that maltose oligosaccharides derived from modification and cleavage of irregular chains of amylopectin are the main substrates of GBSSI (Smith, Denyer & Martin, 1997). Though this enzyme has been shown to be tightly bound to the starch granule (Denyer et al., 1993; Mu-Forster et al., 1996), the matrix between amylopectin molecules is believed to be the site of amylose synthesis and deposition. From these observations, it is believed that amylose synthesis could be controlled by space availability within the amylopectin matrix and or substrate accessibility. Smith, Denyer & Martin (2000) postulated that increasing GBSS1 above the threshold may not increase amylose content beyond 20%. However, this opinion may hold in specific prevailing conditions but it can be complicated by environment-genotype interaction during gene expression. It is documented that during starch biosynthesis, high temperatures decrease metabolite availability and enzyme activity associated with the starch pathway (Hurkman et al., 2003). It was noticed that amylose concentration was slightly higher in wheat grains produced under high temperature conditions. This phenotypic variation in starch composition was strongly related to mRNA levels and duration of expression of genes involved in starch accumulation. In related studies, it was shown that oscillations of GBSSI mRNA levels in leaves of snapdragon (Merida et al., 1999) and Arabidopsis (Tenorio et al., 2003) were regulated by a circadian clock.

Mutants deficient in GBSSI activity, *waxy*, have been identified in several plants such as maize (Hixon & Brimhall, 1968), wheat (Nakamura et al., 1995), rice (Murata, Sugiyama & Akazawa, 1965), barley (Ishikawa, Ishihara & Itoh, 1994) and potato (Hovenkamp-hermelink et al., 1987), with little or no amylose starch in sink tissues, confirming GBSSI as responsible for amylose synthesis. However, amylose and different amylose:amylopectin ratios (compared to that of endosperm) have been detected in tissues such as pericarp and leaves of wheat (Nakamura et al., 1998) and pea (Tomlinson, Lloyd & Smith, 1997.), indicating that GBSSI is not the sole enzyme responsible for amylose synthesis. Interestingly, a second GBSS, GBSSII, has been cloned in wheat and has been shown to be responsible for synthesis of amylose in non-storage tissues of waxy mutants in cereals (Vrinten & Nakamura, 2000). It was observed that GBSSII expression was

abundant in *waxy* pericarp and leaf while GBSSI was preferentially expressed in wheat endosperm (Vrinten & Nakamura, 2000)

Unlike GBSS isoforms, SSS are partially granular bound or soluble proteins within the starch granules and they catalyse the transfer of glucosyl units, the products of the AGPase reaction to non-reducing end of glucan chains leading to extension of the amylopectin molecule. There are at least three SSS isoforms, SSI, SSII and SSIII, identified in higher plants. Although each isoform contributes to amylopectin synthesis, they possess different kinetic properties and seem to extend specific subset of glucan chains, as suggested by Li et al. (2000). In cereals, SSII and SSIII are the main contributors in amylopectin synthesis and their expression pattern vary among tissues. In wheat, the gene encoding SSIII is expressed earlier than those encoding SSI and SSII during endosperm development (Li et al., 1999a; 1999b; 2000). Whereas SSII activity dominates in cereal endosperm (Li et al., 2003, Morell et al., 2003) it constitutes barely 20% soluble synthase activity in tubers (Edwards et al., 1999). Interestingly, a SSIIa mutant in barley produced high levels of amylose and reduced amylopectin starch (Morell et al., 2003). In contrast, antisense inihibition of potato SSII or SSIII had no major effects on starch structure in the tubers (Edward et al., 1995). These observations suggest that the importance of SSII vary among crop species but its specific role in starch biosynthesis can not be complemented by other isoforms in cereals.

Characterization of SSI isoform has been done in a few plant species. In maize and barley, the genes encoding SSI isoforms are predominantly expressed in endosperm (Knight et al., 1998; Li et al., 2000) while in rice and potato, the transcripts are more abundant in leaves (Baba et al., 1993; Kossmann et al., 1999).

1.4.3. Starch branching enzymes

Starch branching enzymes (SBEs) form the branching points in amylopectin and rarely in amylose polymers. These enzymes hydrolyze α-1,4-glycosidic linkages and re-attach the chain to α-1,6-positions of the growing glucan chain. A single hydrolytic event accompanied by formation of an α-1,6-linkage leads to creation of new non-reducing ends that serve as potential sites for chain elongation by starch synthases (See review by Kossmann & Lloyd, 2000).

The middle region of the SBE amino acid sequences contains a conserved $(\beta/\alpha)_8$-barrel domain, common to all amylolytic enzymes. Previous studies by Jespersen et al. (1993) proposed that the domain harbors a catalytic centre while a more recent study by Kuriki, Stewart & Preiss (1997) suggests that the substrate specificity region of SBE is located within the carboxyl terminal.

Multiple isoforms of SBE have been reported in various plant species, such as barley (Sun et al., 1998), wheat (Morell et al., 1997), maize (Boyer & Preiss., 1978; Gao et al., 1996), rice (Mizuno et al., 1992), pea (Denyer et al., 1993), potato (Larsson et al., 1996) Arabidopsis (Fisher et al., 1996) and sorghum (Paper I) of this thesis. These isoforms are classified into two distinct families based on sequence homology (Burton at al., 1995). The families, SBEIs (family B) and SBEIIs (family A) differ in their enzymatic properties. In maize, it has been shown that SBEI has higher affinity for amylose while SBEII has a higher rate of

branching amylopectin than SBEI (Guan and Preiss., 1993). In comparison, SBEI transfers longer glucan chain than SBEII (Takeda, Guan & Preiss, 1993). Accordingly, it has been reported that changes in the ratio between SBEI and SBEII isoforms are responsible for differences in starch structure during seed development (Burton et al., 1995).

In major cereals such as maize, wheat, barley (Gao et al., 1997; Morell et al., 1997; Sun et al., 1998) and sorghum (Paper I and III), two isoforms of SBEII, SBEIIa and SBEIIb have been identified and characterized. Similarly in rice, the isoforms corresponding to SBEIIa and SBEIIb are denoted as SBE3 and SBE4, respectively (Mizuno et al., 1993). The two closely related isoforms share a high degree of homology (about 90%) in the central potion of the enzyme. A longer N-terminal in SBEIIb compared to SBEIIa is the main distinguishing feature between the two isoforms (Martin & Smith, 1995; Sun et al., 1998). In barley, this extended N-terminus is 94 amino acids longer than that of SBEIIa and its specific function remains elusive (Sun et al., 1998).

The genes encoding SBEI are generally expressed earlier than those encoding SBEII. In cereals, the spatial expression pattern shows that *sbeI* and *sbeIIb* are predominantly expressed in endosperm while *sbeIIa* is expressed in all tissues. However, *sbeIIb* in maize (Gao et al., 1997), is also expressed in the embryo unlike in rice (Yamanouchi & Nakamura, 1992) and barley (Sun et al., 1997) where *sbeIIb* expression is endosperm-specific. The reasons for having different isoforms of SBE in plants are poorly understood. It has been reported that mutations in *sbeIIb* genes resulted in an increased proportion of amylose in maize and rice (Mizuno et al., 1993; Kim et al., 1998). Interestingly, no mutants in SBEI and SBEIIa have been reported so far and this could be because the resultant phenotype is unable to survive environmental selection pressure.

In potato, antisense inhibition of *sbeI* had little effect on starch accumulation while *sbeII* antisense-inhibition resulted in a noticeable increase in amylose. Furthermore, when both *sbeI* and *sbeII* were down-regulated, a remarkable increase in amylose content was observed (Schwall et al., 2000). Put together, these observations lead us to suggest that the physiological activity of each SBE isoforms is distinct and cannot be complemented by another.

Upto date, the mechanism regulating the *sbe* genes is poorly understood. It is generally assumed that expression of these genes are controlled by a complex network of interactions among genes, metabolites, environmental conditions and transcription factors (Tenorio et al., 2003). A transcription factor family involved in the regulation of *sbe* genes was recently isolated (Sun et al., 2003).

1.4.4. Debranching enzymes

It is well established that DBEs play an important role in determining starch structure and granule characteristics during starch biosynthesis. In most plants, two families of DBEs namely, isoamylase- and pullulanase-types, have been identified and classified according to substrate specificity. Isoamylase can use glycogen as a substrate but not pullulan while pullulanase uses pullulans as substrate but not glycogen. However, both types of DBEs hydrolyze α-1,6

glycosidic linkages, a distinctive role important in trimming extensively irregular branches in amylopectin molecules (Ball & Morell, 2003).

Among the DBEs, isoamylase-types have been widely studied in several plant species. It is known that in maize, the *sugary1* mutant, lacking or deficient in DBE activity tends to accumulate water-soluble polysaccharides, termed phytoglycogen, in addition to starch (Pan & Nelson 1984). Similarly, orthologous mutations in rice, *sugarly1* (Nakamura, 1996), barley, *isa-1* (Burton et al., 2002), and Arabidopsis, *dbe1* (Zeeman et al., 1998) resulted in accumulation of phytoglycogen and reduced starch content. Reduction in isoamylase activity also correlates with reduced pullulanase activity. Based on these observations, Ball et al. (1996) proposed that amylopectin molecules are synthesized in a discontinuous mechanism accompanied by highly specific cleavage of glucan chains.

Recent studies on the role of isoamylase in starch metabolism has revealed that *sugary1* mutants in barley showed an increase in granule initiation sites, resulting in an increased number of granules during the early stages of endosperm development (Burton et al., 2002). It is believed that this increase in granule number could be a consequence of accumulation of large amounts of phytoglycogen polymers. At least three distinct isoamylases have been isolated in several plant species, such as Arabidopsis, wheat and potato. In potato, two isoforms are known to form a heterotetramer that inhibits starch granule initiation in tubers (Bustos et al., 2004).

1.4.5. Amylases and phosphorylases

Transitory starch or storage starch is degraded with the help of amylase and phosphorylase, when glucose supply is required for various biochemical processes in plants. Amylases hydrolyze α-1,4-glycosidic linkages of starch molecules into shorter oligosaccharides, which are further degraded to release glucose molecules. Amylases are sub-divided into α-amylase and β-amylase. The α-amylases degrade starch molecules by cleaving bonds within the chain, while β-amylases cleave the glycosyl linkages from the non-reducing end and stops the degradation activity on encountering a branching point. In maize, four isozymes of α-amylases, I, II, III, and IV, have been identified (Subbarao et al., 1998). Expression of α-amylase III is endosperm-specific while that of α-amylase II occurs in both aleurone layer and endosperm. Expression of α-amylases is induced by gibberellic acid secreted by the embryo. At the transcription level, the presence of abscissic acid dominantly inhibits α-amylase activity.

Phosphorylases likewise degrade starch by catalyzing insertion of phosphoryl groups from inorganic pyrophosphate into α-1,4-glycosidic bonds, releasing glucose-1-phosphate. This type of hydrolysis occurs in the cytosol resulting into formation of glucose-1-phosphate.

2. Present investigation

To gain insight in starch metabolism in sorghum, we undertook to investigate expression profiles of key enzymes crucial in determining starch structure and granule characteristics. In this work, we aimed at studying and characterizing starch branching enzymes and the genes that encode them. The long-term objectives are to contribute towards understanding regulation of starch biosynthesis in sorghum and to generate transgenic sorghum with modified starch content.

Most of the methodologies used in this study are presented in **Paper I** to **IV**.

2.1. Cloning of *sbeIIa* and *sbeIIb* (Paper I & III)

In order to characterize genes encoding starch branching enzymes in sorghum, we set out to isolate genomic clones for *sbeIIa* and *sbeIIb* genes from a Bacterial Artificial Chromosome (BAC) library containing the entire sorghum genome. Since it is known that genes encoding starch branching enzymes in cereals share a high degree of homology (Gao et al. 1997; Rahman et al. 2001), we used heterologous barley *sbeIIa* and *sbeIIb* probes obtained from our laboratory to identify clones for both *sbeIIa* and *sbeIIb* in sorghum.

The isolation of *sbeIIb* gene was done through a series of screening procedures, involving restriction digests and repeated probing with a barley *sbeIIb* gene-specific fragment to detect positive genomic fragments. Four DNA fragments, totaling 7.8-kb, were cloned and sequenced. Blast search in the databases revealed that two overlapping fragments encompassed a 2.7-kb promoter sequence, first four exons and three introns of the sorghum *sbeIIb* gene (accession number AY304539). Further analyses on sequence comparison showed that sorghum *sbeIIb* shares a high nucleotide identity with *sbeIIb* genes from maize, rice, barley and wheat. The sorghum *sbeIIb* transcription start site was predicted using alignment with a maize *sbeIIb* cDNA. It was interesting to note that the genomic structure of *sbeIIb* is more closely related to that of maize *sbeIIb* than barley *sbeIIb*. A striking difference between maize, sorghum and barley *sbeIIb* genes is the presence of the long (2064-nt) second intron in barley *sbeIIb*. The functional relevance of this long second intron is discussed in detail in **Paper II**, in this thesis. we noted a high relatedness between the sorghum and maize *sbeIIb* genes, which leads us to suggest that the genes arose from a single sequence, conserved in the common ancestor of the two species. In agreement to our opinion, Hulbert et al. (1990) reported that the linkage relationship of polymorphism between maize and sorghum DNA fragments was highly conserved and the two genomes are very similar, although the maize genome has undergone rapid DNA duplication since the two species diverged.

Using the reverse transcription-PCR method, a 2664 nucleotide long fragment of sorghum *sbeIIb* cDNA was cloned and sequenced (accession number AY304540). The *sbeIIb* cDNA clone contained the entire coding sequence, encoding 803 amino acids. Alignment of the deduced amino acid sequence with sequences of barley SBEI and SBEII and sorghum SBEI, (Fig. 3 in **Paper I**), shows that the middle region of the SBEI and SBEII enzymes harbors the predicted catalytic sites of the conserved $(\alpha/\beta)_8$ barrel domain. From the alignment with maize SBEIIb, we predicted the cleavage site for the transit peptide in the sorghum SBEIIb, directing the enzyme to the plastid (Fig. 4 in **Paper 1**).

Sorghum *sbeIIa* genomic and cDNA sequences were cloned and sequenced using a similar process as done for sorghum *sbeIIb* and its corresponding cDNA. Based on Blast searches and alignment with related genomic sequences of Arabidopsis *sbe2.2* (accession number AL162506.1), maize *sbeIIb* (accession number AF072725) and wheat *sbeIIa* (accession number AF338431), it was found that the sorghum *sbeIIa* clone encompassed the first 9 exons and 8 introns.
Southern blot analysis was performed to determine gene copy number of *sbeIIa* and *sbeIIb* using four restriction enzymes that cut once within the probes. The results consistently showed two hybridizing bands for each digest, clearly indicating that each of the two genes exists as a single copy in the sorghum genome. This is consistent with findings from rice, barley, maize and wheat (Yamanouchi & Nakamura, 1997; Sun et al., 1998; Kim et al., 1998; Rahman et al., 2001).

2.2. Expression of SBE in sorghum (Paper I & III)

Expression profiles of genes involved in starch biosynthesis play an important role in determining starch structure and functional properties of the granules. In sorghum, we demonstrated that *sbeIIb* and *sbeI* genes are differentially expressed in the endosperm and embryo, the main storage organs for carbon reserve (Fig. 5, in **Paper I**). Unlike *sbeIIb*, the sorghum *sbeIIa* was detected in both photosynthetic and non-photosynthetic organs (Fig. 4).

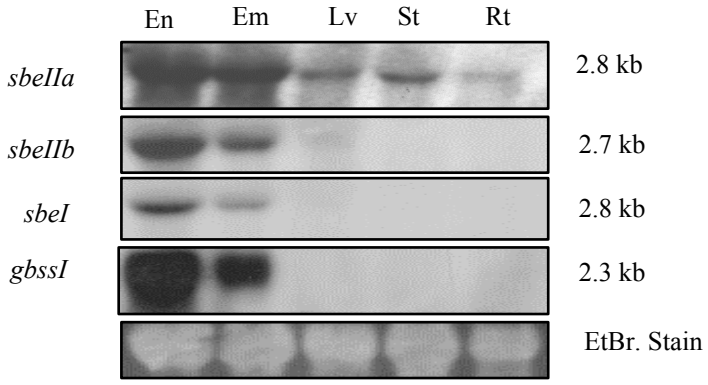

Figure 4. Spatial expression profiles for *sbe* genes in sorghum. Total RNA was isolated from endosperm (En), embryo (Em), leaf (Lv), stem (St) or roots (Rt) and probed with *gbss1*,*sbeI*, *sbeIIa* and *sbeIIb.* Sizes for the hybridizing transcripts are indicated.

The variations in spatial expression of the sorghum *sbe* genes implys that the enzyme combinations participating in starch synthesis differ among the various plant organs. From spatial expression shown in Fig. 4, it can be seen that *sbeIIb* transcripts are obviously abundant in endosperm. Therefore, it is tempting to suggest that the branching activity of starch in the endosperm is mainly contributed by SBEIIb since it is more active in branching amylopectin than SBEIIa. The pattern in gene expression concurs with results reported by Gao et al. (1996), who found that *sbeIIb* in maize is highly expressed in endosperm and embryo while *sbeIIa* is constitutively expressed in all tissues. On the contrary, spatial expression pattern for barley *sbeIIb* (Sun et al., 1998) and rice *sbeIIb* (Kawasaki et al., 1993) showed that the genes are endosperm specific. Whether sorghum and maize *sbeIIb* genes share the same regulatory mechanism, different from that regulating the barley *sbeIIb* or rice *sbeIIb* genes, is not yet known.

Temporal expression patterns for *sbeIIb* and *sbeI* revealed that the transcript levels peaked at 16-24 days after pollination (DAP) (Fig. 6A, **Paper I**). This peaking period deviates from the one observed for barley *sbeIIb* which occurred around 7 to 17 DAP, while *sbeI* peaked at later stages of endosperm development (Sun et al., 1998). Similar variation in *sbeI* and *sbeIIb* expression levels was also observed in maize (Gao et al., 1996). The expression profiles and maximum expression period of the *sbe* genes contribute greatly in determining quality and quantity of starch.

Based on our data, together with known information about spatial expression of genes involved in starch biosynthesis, we propose a list of possible enzyme composition responsible for starch biosynthesis in source and sink tissues in sorghum (shown Fig. 5).

2.3. Regulation of genes encoding SBEs

To identify the minimum functional promoter sequences critical for *sbeIIb* expression in endosperm and embryo cells, we created a series of 5′ deletions of the *sbeIIb* promoter by PCR amplification, and fused the fragments to chimeric plasmid constructs, containing *gfp* as a reporter gene (Fig. 5, **Paper II**). We introduced the resultant constructs into sorghum endosperm and embryos using biolistic transformation method. Deletions of the sequence upto -1023 had no noticeable effect on promoter activity, whereas further deletions down to -350 resulted in a significant reduction in promoter strength. Based on these transient assays, we established that the minimal functional promoter for sorghum *sbeIIb* expression is approximately 350-nt upstream of the transcriptional start site.

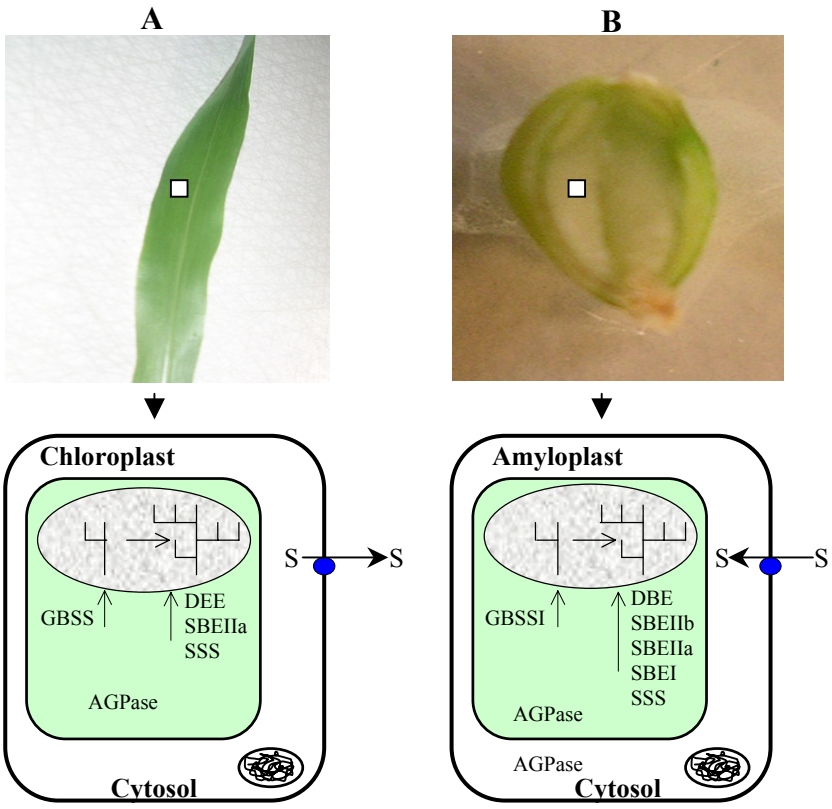

Figure 5. A diagrammatic representation of enzymes involved in starch biosynthesis in sorghum tissues. **A**. Leaf. **B**. Cross section of a grain showing the endosperm. S represent sucrose.

To explore the possible regulatory elements responsible for *sbeIIb* expression, we analyzed the sorghum genomic DNA fragment containing the promoter region and identified several putative sequences important for gene expression in endosperm and embryo. Among these were SURE elements (A/TAAANA), W-boxes (TGAC) and CAAAT motifs (CAAAAT) (Table I, **Paper II**). It has been proposed that many genes involved in starch biosynthesis are regulated by sugar availability and that a SURE element confers sugar responsiveness.

Work done in our laboratory has shown that a sugar singling transcription factor in barley (SUSIBA2) binds to the SURE *cis* element in the isoamylase promoter, leading to activation in gene expression (Sun et al., 2003). However, the SURE-like sequence in the sorghum *sbeIIb* promoter was located at a far distal position (-1594 relative the transcriptional start site) compared to –253 in barley *sbeIIb*. As we suspected, deletion of the region containing the SURE element in sorghum did not affect the *sbeIIb* promoter strength in transient *gfp* expression assays. To investigate the sugar responsiveness of sorghum *sbeIIb*, we conducted ectopic *sbeIIb* expression assays in barley and sorghum leaves and examined the transcripts by northern blot analyses. The analysis indicated that sorghum *sbeIIb* was not sugar-inducible and suggested that the SURE-like sequence located at the distal position is not an important regulatory element. The promoter deletion studies further indicated that the w-box and CAAAT motif was not required for high-level expression of the *sbeIIb* gene in sorghum endosperm or embryo.

Previous studies in our laboratory (Sun et al., 2003; Ahlandsberg, Sun & Jansson 2002) indicated that the long second intron in barley *sbeIIb* is partly responsible for the endosperm-specificity in gene expression. Based on these studies, we made three chimeric plasmids containing functional promoters of sorghum *sbeIIb* and barley *sbeIIb* and *sbeIIa*, with or without the second intron, and introduced the constructs into barley and sorghum embryo and endosperm using the biolistic transformation system. Interestingly, the presence of the intron in all constructs abolished *gfp* expression in both sorghum and barley embryos, while embryo and endosperm cells transformed with constructs lacking the intron showed *gfp* expression. These findings confirm that the barley *sbeIIb* second intron is responsible for conferring endosperm specificity in *sbeIIb* gene expression and, that, possibly, the same repressor factor acting in barley embryo is also present in sorghum embryo.

The above data prompted us to search for a sorghum orthologs to the SUSIBA2 transcription factor (Sun et al., 2003). In an effort to address this task, we cloned and characterized a novel WRKY transcription factor gene, which is predominantly expressed in leaf. Alignment of the amino acid sequence deduced from the cloned cDNA fragment with other WRKY proteins revealed presence of two WRKY domains and one zinc finger-like motif, confirming that the protein belongs to class 1 of the WRKY transcription factors. Its relatedness with a rice repressor factor and SUSIBA2 leads us to suspect that the SbWRKY1 may be playing a role in regulation of genes involved in carbohydrate metabolism. This work opens room for further characterization to establish the specific role of SbWRKY1 in plant developmental processes.

In **Paper IV**, we undertook to examine expression profiles of genes involved in starch biosynthesis and starch structure in different sorghum cultivars in order to

gain more insight on their starch metabolism in endosperm and embryo tissues. The cultivars used in this study were KariMtama1 (white cultivar) and Serena and Seredo (brown cultivars). From the analysis, we observed that *gbssI*, *sbeI*, *sbeIIa* and *sbeIIb* were highly expressed in KariMtama1 and barely undetectable in Serena and Seredo. Since the enzymes encoded by these genes are involved in starch synthesis in the endosperm, it is therefore likely that the observed gene expression levels contribute to higher yields in KariMtama1.

Preliminary results on starch analysis showed that majority of the granules in KariMtama1, unlike in Serena and Seredo were A-type and contained high levels of amylose. However, it has been noted that most farmers in east African region prefer brown to white cultivars. The reason could be partly due to traditional beliefs. According to some African cultures, brown or pigmented sorghum is preferred to produce certain foods for special occasions. In addition, thick porridge prepared from brown sorghum is believed to stay longer in the stomach and it is also thought to be a curative against diseases of the digestive system. The brown characteristic of the sorghum grains is a result of high levels of phytochemicals such as tannins, phenolic compounds and anthocyanins. Although these phytochechemicals particularly tannins, reduces digestibility of sorghum starch, they offer beneficial health characteristics such as reducing risk of certain cancer types. However, further characterization of brown and white sorghums is required to determine more desirable characteristics, which contribute to farmer's preference on cultivars for production and utility.

2.4. Diurnal oscillation of *sbe* genes

In the study described in **Paper III**, we showed that expression of the *sbe* genes in sorghum endosperm exhibit a diurnal rhythm with a maximum level of activity 6-8 h after the onset of light. The oscillation was maintained also in dark-adapted cultured spikes. From these findings, we conclude that sorghum endosperm contain an oscillator that regulates the activity of the *sbe* genes, possibly in anticipation of the influx of sucrose from the source.

2.5. Conclusion

The genomic and cDNA sequences of *sbeIIa* and *sbeIIb* genes were isolated and characterized.

The spatial and temporal expression profiles of sorghum *sbe* genes were determined. Sorghum *sbeI* and *sbeIIb* exhibit a seed-specific expression while *sbeIIa* is constitutively expressed in source and sink tissues. The activity of all these genes show late onset with a peak around 22 days after pollination.

Using deletion mutagenesis of the sorghum *sbeIIb* upstream region, it was found that a 350 bp-long sequence, relative to start of transcription, is the minimal functional promoter size that could support low-level gene expression.

Transient expression assays of *gfp* reporter constructs introduced into sorghum and barley embryo and endosperm cells demonstrated that the barley *sbeIIb* second intron confers endosperm-specific expression in both homologous and heterologous systems.

Diurnal expression patterns of sorghum *sbeI*, *sbeIIa* and *sbeIIb* indicate that an oscillatory mechanism regulates the *sbe* genes in sorghum endosperm.

2.6. Future perspectives

Grain yield and quality are the most important factors limiting food sufficiency in the world to date. Sorghum, as already indicated, has an opportunity with input from molecular and biotechnological approaches to supply a greater share of the world's grain demand, and as an inexpensive source of starch. To be able to improve starch quality in sorghum through genetic approaches, it is probably rewarding to perfect regeneration and transformation protocols for elite cultivars. This study, together with a powerful approach to control gene expression in specific cell types, will be an invaluable gain towards future improvement in sorghum productivity.

Further understanding of interactions between environment and genetic factors, related to starch accumulation, needs to be gained in order to maximise the genetic potential of sorghum. It is also important to study the interactions among enzymes such as starch synthases and starch branching enzymes during endosperm development in order to shed some light on the best choice of enzymes to target during transformation events.

In contributing to our knowledge regarding starch metabolism, it is worthwhile to explore and quantify the contribution of individual enzymes towards grain yield and starch composition in sorghum.

3. References

Ahlandsberg S., Sun, C. & Jansson, C. 2002. An intronic element directs endosperm-specific expression of the *sbeIIb* gene during barley seed development. *Plant Cell Report* 20, 864-868.

Baba, T., Nishihara, M., Mizuno, K., Kawasaki, T., Shimada, H., Kobayashi, E., Ohnishi, S., Tanaka, K. & Arai, Y. 1993. Identification, cDNA cloning, and gene expression of soluble starch synthase in rice (*Oryza sativa* L.) immature seeds. *Plant Physiology* 103, 565-573.

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

Ball, S.G., van de Wal, M. H.B.J. & Visser, R.G.F. 1998. Progress in understanding the biosynthesis of amylose. *Trends in Plant Science* 3, 462-467.

Ball, S.G.& Morell, M.K. 2003. From bacterial glycogen to starch: Understanding the biogenesis of the plant starch granule. *Annual Review in Plant Biology* 54, 207-233.

Beckles, D.M., Smith, A.M. & Rees, T.A. 2001. A Cytosolic ADP-Glucose pyrophosphorylase is a feature of graminaceous endosperm, but not of other starch-storing organs. *Plant Physiology*. 125, 818-827.

Beleon, A., Colonna, P., Planchort V. & Ball, S. 1998. Starch granule: structure and biosynthesis. *Biological Macromolecules* 23, 85-112.

Berger, F. 2003. Endosperm. Current Opinion Plant Biology 6, 42-50.

Boyer, C.D., & Preiss, J. 1978. Multiple forms of starch branching enzyme of maize. Evidence for independent genetic control. *Biochem. Biophys. Res. Commun. 80*, 169-175.

Bugusu, B.A., Campanella, O. & Hamaker, B.R. 2001. Improvement of sorghum-wheat composite dough rheological properties and breadmaking quality through zein addition. *Cereal Chemistry* 78,31-35.

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

Burton, R.A., Jenner, H., Carrangis, L., Fahy, B., Fincher, G.B., Hylton, C., Laurie, D.A., Parker, M., Waite, D., van Wegen, S., Verhoeven, T. & Denyer, K. 2002. Starch granule initiation and growth are altered in barley mutants that lack isoamylase activity. *Plant Journal* 31(1), 97-112.

Bustos, R., Fahy, B., Hylton, C.M., Seale, R., Nebane, N.M., Edwards, A., Martin, C. & Smith, A. M. 2004. Starch granule initiation is controlled by a heteromultimeric isomylase in potato tubers. *PNAS* 101, 2215-2220.

Chandrashekar, A. & Kirleis, A.W. 1988. Influence of protein on starch gelatinization in sorghum. *Cereal Chemistry* 65, 457-462.

Denyer K., Sidebottom, C., Hylton, C.M. & Smith, A.M. 1993. Soluble isoforms of starch synthase and starch-branching enzyme also occur within starch granules in developing pea embryos. *Plant Journal* 4, 191-8.

Denyer, K., Clark, B., Hylton, C., Tatge, H. & Smith, A.M. 1996. The elongation of amylose and amylopectin chains in isolated starch granules. *Plant Journal* 10, 1135-1143.

Dreher M.L, Dreher, C.J. & Berry, J.W. 1984. Starch digestibility of foods: a nutritional perspective. Crit Rev *Food Science Nutrition*. 20, 47-71.

Edwards, A., Marshall, J., Denyer, K., Sidebottom, C., Visser, R.G.F., Smith, A.M. & Martin, C. 1995. Biochemical and molecular characterization of a novel starch synthase from potato tubers. *Plant Journal* 8, 283-294.

Edwards, A., Fulton, D.C., Hylton, C.M., Jobling, S.A., Gidley, M., Rossner, U., Martin, C. & Smith A.M. 1999. A combined reduction in activity of starch synthases II and III of potato has novel effects on the starch of tubers. *Plant Journal* 17, 251-261.

Ellis R.P., Cochrane, M.P. Dale, M.F.B., Duffus, C.M., Lynn, A., Morrison, I.M., Prentice, R. D.M., Swanston, J.S. & Tiller, S.A. 1998. Starch production and industrial use. *Journal of the Science of Food and Agriculture* 77, 289-311.

Emes M.J., Bowsher, C.J., Hedley, C.M., Burrell, M., Scrase-Field, E.S.F. & Tetlow, I.J. 2003. Starch synthesis and carbon partitioning in developing endosperm. *Experimental Botany* 54, 569-575.

FAO Statistic*al data.* http://www.fao.org/ag7magazine/0202sp2 accessed 2nd July 2004

Fisher, D.K., Gao, M., Kim, K-N., Boyer, C.D. & Guiltinan, M.J. 1996. Two closely related cDNAs encoding starch branching enzyme from *Arabidopsis thaliana*. *Plant Molecular Biology* 30, 97-108.

Gao, M., Fisher, D.K., Kim, K-N., Shannon, J.C. & Guiltinan, M.J. 1996. Evolutionary conservation and expression patterns of maize starch branching enzyme I and IIb genes suggests isoform specialization. *Plant Molecular Biology* 30, 11223-32.

Gao, M., Fisher, D.K., Kim, K-N., Shannon, J.C. & Guiltinan, M.J. 1997. Independent genetic control of maize starch-branching enzymes IIa and IIb – isolation and characterization of a *sbe2a* cDNA. *Plant Physiology* 114, 69-78.

Hixon, R.M & Brimhall, B, 1968. Waxy cereals and red iodine starches. In JA Radley, ed, Starch and its Derivatives, *Ed 4. Chapman and Hall, London*, pp 203-246.

Hizukuri, S. 1995. Relationship between the distribution of the chain length of amylopectin and the crystalline structure of starch granule. *Carbohydrate Research* 141, 295-306.
Hizukuri, S. 1996. Polymodal distribution of the chain lengths of amylopectin, and its significance. *Carbohydrate Research*, 147, 342-347.
Hovenkamp-Hermelink, J.H., Jacobsen, E., Ponstein, R., Visser, R.G.F., Vos-Scheperkeuter, G.H., Bijmolt, E.W., De Vries, J., Witholt, B. & Feenstra, W. 1987. Isolation of an amylose-free starch mutant of the potato (*Solanum Tuberosum* L). *Theoritical Applied Genetics* 75, 217-221
Hubbard, J.E., Hall, H.H. & Earle, F.R. 1950. Composition of the component parts of the sorghum kernel. *Cereal Chemistry* 27, 415-420.
Hulbert, S.H., Richter, T.E., Axtell, J.D. & Bennetzen, J.L. 1990. Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. *Proc. Natl. Acad. Science* 87, 4251-4255.
Hurkman, W.J., McCue, K.F., Altenbach, S.B., Korn, A., Tanaka, C.K., Kothari, K.M., Johnson, E.L., Bechtel, D.B., Wilson, J.D., Anderson, O.D. & DuPont, F.M.. 2003. Effect of temperature on expression of genes encoding enzymes for starch biosynthesis in developing wheat endosperm. *Plant Science* 165, 873-881.
Ishikawa, N,. Ishihara, J. & Itoh, M. 1994. Artificial induction and characterization of amylose-free mutants of barley. *Barley Genet Newsletter* 24, 49-53.
James G.M., Denyer, K. & Myers, A.M. 2003. Starch biosynthesis in the cereal endosperm. *Current Opinion in Plant Biology* 6, 215-222.
Jane, J., Chen, Y.Y., Lee, L.F., McPherson, A.E., Wong, K.S., Radosavljevic, M. & Kasemsuwan, T. 1999. Effects of amylopectin branch chain length and amylose content on the gelatinization and pasting properties of starch. *Cereal Chemistry* 76, 629-637.
Jane, W-N. 2004. Ultrastructure of endosperm development in *Arundo formonsana* Hack (Poaceae) from differentiation to maturity. *Bot. Bull. Acad. Sin*. 45, 69-85.
Jansson, C. 2004. Sugar signaling mutants in *Arabidopsis*. *Unpublished.*
Jespersen, H.M., MacGregor, E.A., Henrissat, B., Sierks, M.R. & Svensson, B. 1993. Starch- and glycogen-debranching and branching enzymes: Prediction of structural features of the catalytic $(\beta/\alpha)_8$–barrel domain and evolutionary relationship to other amylytic enzymes. *Journal of Protein Chemistry* 12, 791-805.
Kawasaki, T., Mizuno, K., Baba, T. & Shimada, H. 1993. Molecular analysis of the gene encoding a rice starch branching enzyme. *Mol. Gen. Genet*. 237, 10-6.
Khoshnoodi, J., Larsson, C.T., Larsson, H. & Rask, L. 1998. Differential accumulation of Arabidopsis thaliana sbe2.1 and sbe2.2 transcripts in response to light. *Plant Science* 135, 183-193.
Kim, K-N., Fischer, D.K. & Guiltinan, M.J. 1998. Molecular cloning and characterization of the maize gene encoding starch branching enzyme IIb. *Plant Molecular Biology* 38, 945-956.
Knight, M.E., Harn, C., Lilley, C.E.R., Guan, H.P., Singletary, G.W., MuForster, C.M., Wasserman, B.P. & Keeling, P.L. 1998. Molecular cloning of starch synthase I from maize (w64) endosperm and expression in *Escherichia coli. Plant Journal*. 14, 613-622.
Kossmann, J. & Lloyd, J. 2000. Understanding and influencing starch Biochemistry. Critical reviews in *Plant Science* 19, 171-226.
Kossmann, J., Abel, G.J.W., Springer, F., Lloyd, J.R. & Willmitzer, L. 1999. Cloning and functional analysis of a cDNA encoding a starch synthase from potato (*Solanum tuberosum* L.) that is predominantly expressed in leaf tissue. *Planta* 208, 503-511.
Kuriki, T., Stewart, D.C. & Preiss, J. 1997. Construction of chimeric enzymes out of maize endosperm branching enzymes II and I: Activity and properties. *Journal of Biological Chemistry*. 272, 28999-29004.
Lan, S.Y., Zhong F.X, Yang, Z.M., Jin, de M. & Xu Z.X. 2004. The starch endosperm denucleation by a process of programmed cell death during rice grain development. *Shi Yan Wu Xue Bao* 37, 34-44.
Larsson, C-T., Hofvander, P., Khoshnoodi, J., Ek, B., Rask, L. & Larsson, H. 1996. Three isoforms of starch synthase and two isoforms of branching enzyme are present in potato tuber starch. *Plant Science* 117, 9-16.

Li Z., Chu, X., Mouile, G., Yan, L., Kosar-Hashemi, B., Hey, S., Napier, J., Shery, P., Clarke, B., Appels, R. & Morell, M.K. 1999a. The localization and expression of the class II starch synthases of wheat. *Plant physiology* 120: 1147-1156.

Li Z., Rahman, S., Kosar-Hashemi, B., Mouile, G., Appels, R. & Morell, M.K. 1999b. Cloning and characterization of a gene encoding wheat starch synthase I. *Theor. Appl. Genet*. 98, 1208-1216.

Li, Z., Mouille, G., Kosar-Hashemi, B., Rahman S., Bryan, C., Kevin, R.G., Rudi, A. & Morell, M.K. 2000. The structure and expression of the wheat starch synthase III gene. motifs in the expression gene define the lineage of the starch synthase III gene family. *Plant Physiology* 123, 613-624.

Li, Z., Xu, F., Chu, X., Mukai,Y., Yamamoto, M., Ali, S., Rampling, L., Kosar-Hashemi, B., Rahman, S. & Morell, M.K. 2003. The structural organization of the gene encoding class II starch synthase of wheat and barley and the evolution of the genes encoding starch synthases in plants. *Funct. Integr. Genomics* 3, 76-85.

MacGregor, A.W. & Fincher, G.B. 1993. Carbohydrates of the barley grain. In Barley: Chemistry and Technology. Edited by A. W. MacGregor and R. S. Bhatty. *American Association of Cereal Chemists Inc., St Paul*, pp 73-130.

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

Matheson, N.K. 1996. The chemical structure of amylose and amylopectin fractions of starch from tobacco leaves during development and diurnally-nocturnally. *Carbohydrate Research* 282, 247-62.

Merida A., Rodriguez-Galan, J.M., Vincent, C. & Romero, J.M. 1999. Expression of the granular-bound starch synthase I (*Waxy*) gene from snapdragon is developmentally and circadian clock regulated. *Plant Physiology* 120, 401-409.

Mizuno, K., Kamura, K., Arai, Y., Kawasaki, T., Shimada, H. & Baba, T. 1992. Starch branching enzymes from immature rice seeds. *Journal of Biochemistry*. 112:643-51.

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

Morell, M.K., Blennow, A., Kosar-Hashemi, B. & Samuel, M.S. 1997. Differential expression and properties of starch branching enzyme isoforms in developing wheat endosperm. *Plant Physiology* 113, 201-8.

Morell, M.K. 2001. Comparison of starch-branching enzyme genes reveals evolutionary relationships among isoforms. Characterization of a gene for starch-branching enzyme IIa from the wheat D genome donor *Aegilops tauschii. Plant Physiology* 125, 1314-1324.

Morell, M.K., Kosar-Hashemi, B., Cmiel, M., Samuel, M.S., Chandler, P., Rahman, S., Buleon, A., Batey I.L.& Li, Z. 2003. Barley sex6 mutants lack starch synthase IIa activity and contain a starch with novel properties. *Plant Journal* 34,173-85.

Morrison, W.R., Milligan, T.P., & Azudin, M.N. 1984. A relationship between the amylose and lipid contents of starches from diploid cereals. *Journal of Cereal Science*, 2 257-271

Mu-Forster, C., Huang, R.M., Powers, J.R., Harriman, R.W., Knight, M., Singletary, G.W., Keeling, P.L. & Wasserman, B.P. 1996. Physical association of starch biosynthetic enzymes with starch granules of maize endosperm – Granule-associated forms of starch synthase I and starch branching enzyme II. *Plant Physiology* 111, 821-829.

Murata, T., Sugiyama, T. & Akazawa, T. 1965. Enzymatic mechanism of starch synthesis in glutinous rice grain. *Biochem Biophys Res Commun* 18, 371-376.

Nakamura, Y., Yamamori,, M., Hirano, H., Hidaka, S. & Nagamine, T. 1995. Production of Waxy (Amylose-free) Wheat. *Molecular Gen. Genetics* 248, 253-259.

Nakamura, Y. 1996. Some properties of starch debranching enzymes and their possible role in amylopectin biosynthesis. *Plant Science* 121,1-18.

Nakamura, T., Vrinten, T., Hayakawa, K. & Ikeda, J. 1998. Characterization of a granular-bound starch synthase isoform found in the pericarp of wheat. *Plant Physiology* 118, 451-459.

Olsen, O-A. 2001. Endosperm development: Cellularrization and cell fate specialization. Annu. Review. *Plant Molecular. Biology*. 52, 233-267.

Olsen, O-A. 2004. Nuclear endosperm development in cereal and Arabidopsis thaliana. Plant Cell 16, 214-227.

Pan, D. & Nelson, O.E. 1984. A debranching enzyme deficiency in endosperm of the *sugary-1* mutants of maize. *Plant Physiology* 74, 324-328.

Peat, S., Whelan, W.J. & Thomas, G.J. 1952. Evidence of multiple branching in waxy maize starch. *Journal of the Chemical Society*, 4546-4548.

Rahman, S., Regina, A., Li, Z., Mukai, Y., Yamamoto, M., Kosar-Hashemi, B., Abrahams, S & Morel, K. 2001. Comparison of starch-branching enzyme genes reveals evolutionary relationships among isoforms. Characterization of a gene for starch-branching enzyme IIa from wheat D genome donor *Aegilops tauschii*. *Plant Physiology* 125, 1314-1324.

Smith, A. M., Denyer, K. & Martin, C. 1997. The synthesis of the starch granule. Annu. Rev. *Plant Molecular Biology* 48, 67-87.

Subbarao, K.V., Datta, R. & Sharma, R. 1998. Amylases synthesis in scutellum and aleurone layer of maize seeds. *Phytochemistry* 49, 657-66.

Schwall, G.P., Safford, R., Westcott, R., Jeffcoat, R., Tayal, A., Shi, Y-C., Gidley, M.J. & S.A. Jobling. 2000. Production of a very-high amylose potato starch by inhibition of SBE A and B. *Nature Biotechnology* 18, 551-554.

Sun, C., Sathish, P., Ahlandsberg, S. and Jansson, C. 1997. Identification of four starch-branching enzymes in barley endosperm: partial purification of forms I, IIa and IIb. *New Phytol*. 137, 215-222.

Sun, C., Sathish, P., Ahlandsberg, S., Deiber, A. & Jansson, C. 1998. The two genes encoding starch-branching enzymes IIa and IIb are differentially expressed in barley. *Plant Physiology* 118, 37-49.

Sun, C., Palmqvist, S., Olsson, H., Boren, M., Ahlandsberg, S. & Jansson, C. 2003. A novel WRKY transcription factor, SUSIBA2 participates in sugar signaling in barley by binding to the sugar-responsive elements of the *iso* promoter. *Plant Cell* 15:2076-2092.

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

Tenorio, G., Orea, A., Romero, J.M. & Merida, A. 2003. Oscillation of mRNA level and activity of granular-bound starch synthase I in Arabidopsis leaves during the day/night cycle. *Plant Molecular Biology* 51, 949-958.

Tomlinson, K. & Denyer, K. 2003. Starch synthesis in cereal grains. *Adv. Bot. Res.* 40. 2-61.

Tomlinson, K.L., Lloyd, R.J. & Smith, A.M. 1997. Importance of iso-forms of starch-branching enzyme in determining the structure of starch in pea leaves. *Plant Journal* 11, 31-43.

Vrinten, P.L. & Nakamura, T. 2000. Wheat granule-bound starch synthase I and II are encoded by separate genes that are expressed in different tissues. *Plant Physiology*. 122, 255-263.

Williams, J.H. & Friedman, W.E. 2002. Identification of diploid endosperm in an early angiosperm lineage. *Nature* 425, 522-526.

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

Zeeman, S.C., Umemoto, T., Lue, W-L., Au-Yeung, P., Martin, C.,A.M. & Chen, J. 1998. A mutant of Arabidopsis lacking a chloroplastic isoamylase accumulations both starch and phyoglycogen. *Plant Cell* 10, 1699-1711.

4. Acknowledgements

I wish to thank my supervisor, Prof. Christer Jansson for accepting me into his CJ group and introducing me to molecular biology. Thank you for your supervision, support, guidance and patience during these years. I immensely thank you for your concern and encouragement during my PhD period.

Thank you my co-supervisor Dr. Odhiambo for your special understanding and contribution to this work. I wish to confer my sincere appreciation to my employer, KARI-Kenya, for giving me the study leave and logistical support needed to complete the training. Thanks to you all members of KARI staff who contributed in one way or another to this work and imparted success of BIO-EARN program at KARI.

Special thanks to BIO-EARN secretariat at SEI, regional co-ordinators and focal points for effective co-ordination of the program. Thank you Benita and Ivar for your efficiency, tireless effort in ensuring smooth running of the BIO-EARN program and taking care of our welfare extremely well. This program could not have been such a success without you, so thank you Benita once again.

I want to extend my sincere gratitude to the CJ group; Sara, Per, Sun, Laxmi, Hanna, Helena, Yona, Haile, Mats, Staffan, Ronney, Isabel and all of you colleques for your contribution into this work and being true friends. I highly treasure the home-made lunch, Thursday cakes and friendly atmosphere we shared during those years. Thank you Sun for all the special tips of molecular biology I learnt from you and most importantly, reminding me of the apparent alternative when a “gel” breaks into pieces

In addition I want to thank all the people at Genetics Centre especially at the Department of Plant Biology and Forest Genetics and for your valuable interaction during these years. Thank you Mona for exceptional efficiency in processing orders, Ingrid for timely sequencing of my plasmids, Urban and his team for tendering my plants, Bjorn and Lena for solving computer related problems, Lars-Olof Hansson for huge generosity in rendering library services, Per Bergman for being a good leader for the department. The list is too long to name all those who contributed to this work. Thank you all.

Last but not least, thanks to you my parents; Mutisya, Ngene and children; Irene, Brian, lynn and the rest of family members and friends for your endurance and support to the end of my study period. Thank you Mutindi, Bibiana, Norman, Uncle Ben and family for your prayers, understanding and encouragement during times of tribulation.

Finally, thanks to SIDA/SAREC for financial support in this study, BIO-EARN program and contribution in fostering Biotechnology, Biosafety and Biotechnology Policy applications in east Africa region.