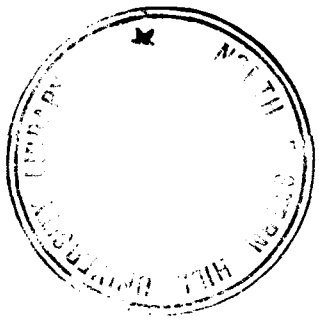
ISOLATION AND FUNCTIONAL CHARACTERIZATION OF THE *WAXY* LOCUS IN COMMON BUCKWHEAT (*FAGOPYRUM ESCULENTUM* MOENCH.)



ABSTRACT

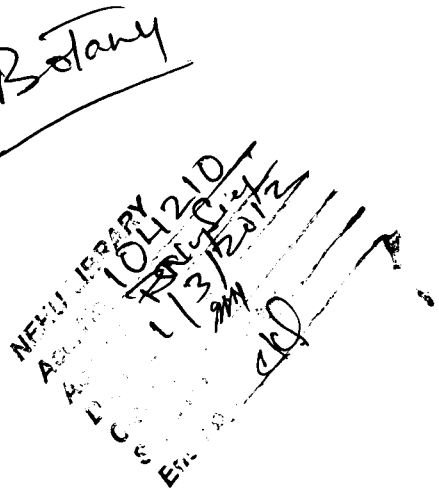
By

NABANITA DEVADASAN

THESIS SUBMITTED
IN FULFILMENT OF THE DEGREE OF
DOCTOR OF PHILOSOPHY IN BOTANY

NORTH EASTERN HILL UNIVERSITY
SHILLONG-793022, INDIA
2010

Botany



Starch is the main energy reservoir of higher plants and also a major source of dietary energy for humans and animals. Besides its nutritive value, starch is a very useful raw material with a wide range of applications in both food and non-food industries. Starch is synthesized in the form of granules with a partially crystalline texture. The morphology, chemical composition and super molecular structure of starch granules are known to be characteristic to each plant species (Banks and Greenwood, 1975). Starch owes much of its functionality to two major high molecular weight molecules *viz.*, amylose and amylopectin as well as to the physical organization of these molecules within the granule structure (French, 1984). Even though both amylose and amylopectin contain polymers of α-D-glucose units, they differ in their length and degree of branching. While amylose comprises largely of unbranched α-1, 4-linked glucan chains, amylopectin is a branched polymer in which linear chains of α-1, 4-linked glucoses are joined together by α-1, 6-linkages.

Amylose and amylopectin represent approximately 98-99% of the dry weight of starch granules. However, the relative proportion of amylose to amylopectin may vary from species to species, within species as well as between tissues from the same species (Jane *et al.*, 1992; Shujun *et al.*, 2006; Peroni *et al.*, 2006). The amount of amylopectin in starches is known to vary from 75% - 85% for normal starches (Manners, 1979) to even more than 99% for waxy starches which are essentially amylose free (Shure *et al.*, 1983; Sano, 1984; Song and Jane, 2000; Singh *et al.*, 2005; Zhu *et al.*, 2008).

The physicochemical properties of starches *viz.*, gelatinization temperature, retrogradation behavior, swelling power, viscosity, pasting properties, freeze/thaw stability, acid stability and gel strength have been shown to have an important bearing on the functional properties of starch (Dedeh and Sackey, 2002; Chen *et al.*, 2003; Amani *et al.*, 2004; Perez *et al.*, 2005; Peroni *et al.*, 2006; Sefa- Shujun *et al.*, 2006; Riley *et al.*, 2006). Therefore, characterization of starches for their physicochemical, functional and structural properties is essential in order to unravel their potential for use in industry. Since the physicochemical properties of starches are genetically determined, it is possible to manipulate the properties of starches by genetic manipulation. Therefore, discovery and characterization of the enzymes that affect the quality of starch are of notable worth. The biosynthesis of starch involves the coordinated interaction between a suite of starch biosynthetic enzymes including ADP glucose pyrophosphorylase (AGP), starch synthase (SS), starch branching enzyme (SBE) and starch debranching enzyme (DBE) with individual plant species having multiple forms of each enzyme.

One of the most intriguing challenges in understanding processes involved in starch biosynthesis is to understand the mechanisms involved biosynthesis of amylopectin and amylose. The existence of multiple forms of each enzyme involved in amylose/amylopectin biosynthesis makes the glucan biosynthetic system in plants complex. On the basis of their degree of association with starch granules, the granule-associated proteins have been categorized either as surface-associated proteins or as granule-bound proteins. Surface associated proteins have a small degree of association with the starch granule and can be separated from the granule by extensive washing in aqueous buffer containing a detergent and a reducing agent at a temperature below the gelatinization temperature for starch (Denyer *et al.*, 1993; Rahman *et al.*, 1995; Boren *et al.*, 2004). On the other hand, extraction of granule-bound proteins requires the gelatinization of starch granule in SDS in the presence of a reducing agent (Tsai, 1974; Echt and Schwartz, 1981; Denyer *et al.*, 1993)

While much work has been done on the characterization of starch from cereals and analysis of the regulation of its biosynthesis, not much information is available about the quality of starch in many other potentially important crops. One such group of crops belongs to pseudo cereals. Within this group, common buckwheat is an important pseudo cereal because of its high potential for use as a functional food. The National Bureau of Plant Genetic Resources (NBPGR), New Delhi has identified buckwheat as an important underutilized crop which has potential to become an important component of human diet. Common buckwheat (*Fagopyrum esculentum* Moench.) belongs to the family Polygonaceae. In India, cultivation of common buckwheat has been reported from the states of Hill areas of Jammu & Kashmir, Himachal Pradesh, Uttar Pradesh, Sikkim, Meghalaya,

Arunachal Pradesh, Assam, Manipur and Nagaland (Anonymous, 1987-2001). Buckwheat is also an important crop in China, Russia, Ukraine, and Kazakhstan, parts of Europe, Canada, Japan, Korea and Nepal.

Skrabanja *et al.* (2001) have shown that starch from buckwheat grains have a strong potential for use in the development of functional foods including its engineering for a modified course of starch digestion. However, despite, the wide utilization of buckwheat starch the flour obtained from buckwheat has poor dough making qualities due to its unfavorable amylose/amylopectin ratio. Since genetic manipulation of genes involved in starch biosynthesis pathway can eventually provide means of modifying the amylose/amylopectin ratio, the present investigation was carried out to isolate and characterize the granule bound protein(s) and the *Waxy* locus in common buckwheat.

The starch grains isolated from endosperm tissues of common buckwheat varied in shape from round to polygonal with size ranging from 3 to 12μm. While the grains from VL-7 and IC-13145 showed distinct polygonal shape, those from IC-188669, KBB-3, OC-2, Siva and Daria were round to spherical in shape. VL-7 and KBB-3 had the largest starch grains which ranged in size from 9.47 μm to 12.1μm. VL-7 is a high yielding and early maturing cultivar from Western Himalayas released by Indian Council for Agricultural Research, New Delhi. The range of variation in shape and size of starch grains in common buckwheat, as observed in the present investigation, as in agreement with other earlier reports on the shape and size of starch grains in common buckwheat (Kim *et al.*, 1977; Soral-Smietana *et al.*, 1984b; Acquistucci and Fornal, 1997; Qian and Kuhn, 1999b). Lindeboom *et al.* (2004) have classified starch grains as large (>25μm), medium (10-25μm), small (5-10μm) and

very small (<5μm). Starch grains isolated from the endosperm of buckwheat can be clearly classified under the small grain size category. Starches having small granules and a narrow granule size distribution have found application in fine printing paper and plastic sheets (Jane *et al.*, 1994; Wilhelm *et al.*, 1998), as a binder with orally active ingredients and as a carrier material in cosmetics (Whistler, 1995). Due to their small size, starch grains of common buckwheat may find similar applications. A distinct feature of starch grains isolated from Indian accessions/ varieties was the presence of pores on the surface of the grains. No such pores were visible on the surface of starch grains from European accessions/ cultivars. Pores were mostly observed on the smoother surface of starch granules. While, some granules had large number of pores on their surfaces, others had only few pores. Porosity and surface area are important characteristics of solid materials that determine their properties such as thermal conductivity, thermal diffusivity and mass diffusion coefficient. Scanning electron microscopy of partially digested starch grains showed a clear pattern of concentric rings thereby revealing the lamellar structure of starch grains. Such lamellar structures have been reported to represent the alternation of semi crystalline and amorphous zone within the matrix (French, 1984; Cameron and Donald., 1992). This alternation of semi crystalline and amorphous zones within the matrix of starch grains has been correlated with the presence of GBSS-I within the grains (Denyer *et al.*, 1995; Smith *et al.*, 1997). Our results on confocal laser scanning microscopy of buckwheat starch clearly indicated the localization of GBSS-I in the form of discrete internal rings within the matrix of buckwheat starch grains. Our observations on the localization of GBSS-I in starch grains are consistent with the assumption that amylose synthesis occurs within the core of the starch granules.

Irrespective of the buckwheat starch grain size and its variety, the percentage of apparent amylose in the starch grains varied between 47% to 51.9%. There were no marked differences in the rheological properties of starches isolated from different accessions of buckwheat studied in the present investigation. The starches showed a mean mean pasting temperature (*P*temp) of 68°C. The peak viscosity, minimum viscosity, breakdown viscoisty and final visocosity were 160 RVU, 102 RVU, 58 RVU and 210 RVU respectively. Similar results on peak, minimum, breakdown and final viscosities of buckwheat starch pastes were reported by Qian *et al.* (1998). Campbell (1995), Zheng *et al.* (1998) and Qian *et al.* (1998) have shown that compared to corn or wheat starches, buckwheat starches swelled faster, exhibited a greater final viscosity and formed stiffer and harder gels. Besides super molecular glucan structures, the high viscosity values for buckwheat starches can be explained by the fact that the starches exhibited a higher granule swelling and gelling tendency than cereal starches (Yoshimoto *et al.*, 2004).

SDS-PAGE profile of proteins associated with starch grains from European varieties/selections of buckwheat indicated varietal differences in the grain proteome amongst different varieties form eastern Europe. While the SDS-PAGE proflies of proteins associated with starch grains from some varieties revealed the presence of a single band corresponding to a moleuclar mass of 59.7 kDa, that of other varieties showed the presence of a duplex with molecular masses of 59.7 kDa and 56 kDa. On the other hand, SDS-PAGE profiles of proteins associated with starch granules isolated from the endosperm tissues of Indian accessions/ varieties of buckwheat showed the presence of only a single band corresponding to a molecular mass of 59.7 kDa. However, SDS-PAGE profile of the proteins associated with starch grains

isolated from leaves of common buckwheat revealed the presence of a single band corresponding to a molecular mass of 53 kDa. Starch granule associated proteins in endosperm of common wheat (*Triticum aestivum* L.) have been reported to include at least one major protein with a molecular weight of 61 kDa and six minor high molecular weight proteins. While the 61 kDa protein was identified as GBSS-I, all other proteins were identified as soluble starch synthase (Takaoka *et al.*, 1997). While the 59.7 kDa protein showed strong crossreactivity with antibodies raised against GBSS-I, the 56 kDa and 53 kDa proteins did not cross react with GBSS-I antibodies. These results indicate absence of serological homology between GBSS-I and the 53 and 56 kDa starch granule associated proteins. The antisera, however cross reacted with the 61 kDa GBSS-I of maize and the 60 kDa GBSS-I of rice and wheat thereby indicating serological homology between the GBSS-I of wheat, maize, rice and buckwheat. GBSS-I, the key enzyme responsible for amylose synthesis, has been identified as a 56, 58 or a 60 kDa protein in maize (Echt and Schwartz, 1981; Shure *et al.*, 1983; Gibbon *et al.*, 2003), a 60 kDa protein in rice (Sano, 1984), a 58 kDa protein in mung bean (Ko *et al.*, 2009), 68 kDa protein in grain amaranth (Konishi *et al.*, 1985), 59 kDa protein in pea (Dry *et al.*, 1992).

2D-PAGE of the endosperm starch granule associated proteins resolved the fraction into 10 spots with *pI* ranging from 5.2 to 6.2. All the spots showed an apparent molecular mass of 59.7 kDa. Imunoblotting of the 10 spots separated by 2D-PAGE with antisera raised against buckwheat GBSS-I identified two bands viz., spot no. 3 (*pI* 5.4) and spot no 8 (*pI* 6.1) which cross reacted with antibodies raised against buckwheat. Similar results have been reported for *waxy* proteins of maize, rice and barley (Nakamura *et al.*, 1993, 1995; Taira *et al.*, 1995; Chao *et al.*, 1985).

While the protein was detected as a single band of 60 kDa on 1D-PAGE, it resolved into 3 (barley) to 4 (maize, rice) isoforms, with *pI* ranging from 5.8 to 7.2 on 2D-PAGE. The protein corresponding to spot no. 3 was subjected to in-gel trypsin digestion, followed by Mass Spectrometry and database search with MASCOT (www.matrixscience.com). The peptides were identified by comparing them with the rice genomes. One of the tryptic fragment having the amino acid sequence "FNAPLAHLIMAGADVLAVPSR" with a predicted *pI* of 8.34 showed similarity with GBSS-I protein of rice.

BLASTp analysis of the N-terminal amino acid sequence for 25 residues of the 59.7 kDa granule associated protein, worked out in the present study, identified the protein as granule bound starch synthase-I. The sequence has been deposited in the SWISS-PROT protein data bank with Accession no. P84633. Multiple alignment of the N-terminal sequence of GBSS-I protein isolated from buckwheat with amino acid sequences of similar proteins available in protein data banks clearly identified the conserved KTGGL motif in the sequence. Furukawa *et al.* (1993), have demonstrated the ubiquitous presence of KTGGL domain in all the GBSS-I proteins. This motif has been identified as the ADP/ADPglucose binding site in the enzyme. The sequence showed 94% homology with GBSS-I from *Hordeum vulgaris*, *Triticum* spp. and *Phaseolus vulgaris.* The percentage homology varied between 94% to 88% with GBSS-I from other plants. Even though analysis of the sequence alignment revealed a clear diversification into monocotyledonous and dicotyledonous groups the protein sequence from buckwheat showed similarities with GBSS-I from both the groups. The protein sequence of buckwheat has Valine as the 5^{th} amino acid residue, as is also a case in the protein sequences of GBSS-I from

monocots. However, majority of the dicots' GBSS-I sequence had Isolucine at this position. The sequence of buckwheat also showed similarities with sequences from dicots in having Valine as the 11^{th} amino acid residue. GBSS-I protein sequence from monocots have methionine at this same position. The structural differences may imply differences in catalytic activity of the enzyme. Sequence analysis of buckwheat GBSS-I indicates similarities with both cereal as well as dicot GBSS-I sequences. It is possible that the protein from common buckwheat might be belonging to a catalytically distinct subclass and hence a possible candidate for altering amylose biosynthesis in both dicots as well as monocots. These results strongly indicate that the 59.7 kDa protein isolated from starch grains from endosperm tissues of common buckwheat is a GBSS-I class enzyme and hence, an isoform of the *waxy* protein. This is the first report on the identification of GBSS-I in common buckwheat.

Phylogenetic analysis of the N-terminal amino acid sequence of the GBSS type protein from common buckwheat reported here revealed a clear diversification into monocotyledonous and dicotyledonous groups'. Within the monocots, the sequences could be segregated into two groups. While one of the two groups (Group-I) was dominated by rice (*Oryza* spp.) and maize (*Zea mays*) the other group (Group-II) predominantly comprised sequences from *Triticum* spp, *Hordeum* spp, *Secale cereale*, *Elymus scaber* and *Aegilops speltoides*. The dicotyledons, on the other hand, resolved into three subgroups. While one of the groups comprised of sequences from *Fagopyrum esculentum*, *Nelumbo nucifera*, *Astragalus membranaceus* and *Amaranthus cruentus*, the other group comprised sequences from *Pisum sativum*, *Phaseolus vulgaris* and *Manihot esculenta*. The third group comprised of sequences

identified from *Ipomoea batatas* GBSS-I. These results are in conformity with the observations of Edwards *et al.* (2002) who have reported a clear division of GBSS-I proteins into those belonging to monocots and those to dicots.

Genomic DNA was isolated from etiolated seedlings of common buckwheat using a modified CTAB method. The modified CTAB method protocol used in the present study yielded high quality genomic DNA. The isolated genomic DNA was digested with to check the restriction digestion profile of the DNA isolated from shoot tissues of common buckwheat. The electrophoretic profile of *EcoR*1 digested DNA revealed a uniform smear ranging in size from 0.5 Kb to 14.7 Kb with three prominent bands showing apparent molecular masses of 3.7 Kb, 1.3 Kb and 1.15 Kb. *Hind*III digested DNA also resolved in the form of a uniform smear ranging in molecular mass from 0.5 Kb to 20.0 Kb. The electrophoretic profile of *Hind* III digested DNA showed four bands having molecular masses of 3.7 Kb, 1.04 Kb, 1.15 Kb and 0.534 Kb. The electrophoretic profile of *Nco*I digested DNA revealed that the enzyme could only partially digest the DNA. Even though *Nco*I could only partially digest buckwheat genomic DNA, the electrophoretic profile of *Nco*I digested DNA revealed six bands having molecular masses of 2.46 Kb, 1.87 Kb, 1.108 Kb, 0.983 Kb, 0.592 Kb and 0.488 Kb. The appearance of distinct bands in *Eco*R1, *Hind*III and *Nco*I digested gDNA indicates the presence of *Eco*R1, *Hind*III and *Nco*I repeats of varying lengths in buckwheat genomic DNA.

In the present study polymerase chain reaction was carried out with buckwheat genomic DNA as the template and combinations of oligonucleotide primer pairs designed to amplify the sequence of GBSS gene from buckwheat

genomic DNA. Out of the total 18 primer pairs used, 14 sets failed to amplify any sequence from the template DNA. Amplification of gDNA was, however, achieved with the primer combinations NDF2-NDR2, NDF3-NDR3, NDF7-NDR6 and NDF8-NDR6. PCR amplification with oligonucleotide primer pairs NDF2-NDR2 and NDF3-NDR3 amplified DNA fragments showing an apparent molecular mass of 0.7 Kb. BLASTn analysis of the nucleotide sequences of the two amplicons did not identify any of the two sequences with GBSS gene family. While amplification of gDNA wtih oligonucleotide primer pairs NDF7-NDR6 amplified a DNA fragment of showing an apparent molecular mass of 1.2 Kb, that with primer pair NDF8-NDR6 amplified a DNA fragment having an apparent molecular mass of 0.9 Kb. BLASTn analysis of the nucleotide sequences of the amplicons identified the same with genes encoding for granule bound starch synthase gene family. This is the first report of identification of GBSS-I gene in common buckwheat. ClustalW multiple alignment of the nucleotide sequence of 1.2 Kb amplicon with nucleotide sequences of granule bound starch synthase genes of other plants revealed two highly conserved regions. While one of the conserved regions, represented by the sequence "GGACATAGGGTTATGACAGTTGCTCCTCGTTATGATCAGTATAAAGATG GATGGGATACTAATGTACTAGTTCG", existed between between P'_{94} and P'_{168}, the other conserved region was represented by the sequence "CAGATACAAGTTG GGGAAAGAGTTGAGACTGTTCGGTTCTTTCACTGTACAAAGAGGAGTTAC CGGGTTTTCGTGGACACCCTATTCCTTGAGAAGGT" and was located between P'_{274} to P'_{377}. AUGUSTUS (version2.4) identified four exons and three introns within the nucleotide sequence of 1.2 Kb amplicon. Sequence analysis of the

1.2 Kb amplicon clearly revealed that the 1st and 2nd conserved domains were located within the 1st and 2nd exon respectively.

ClustalW alignment of the nucleotide sequence of the 0.9 Kb amplicon with nucleotide sequences of GBSS-I genes from other plants revealed three highly conserved regions between P'_{368}-P'_{388}, P'_{533}-P'_{602} and P'_{709}-P'_{823}. While the conserved region between P'_{368} and P'_{388} comprised of 21 bases represented by the sequence "TCAGGCTGCTCTAGAGGCACC", that between P'_{533} and P'_{602} comprised of 70 nucleotides represented by the sequence "ATGTTGTTTTCGTTCC AAATGACTGGCACACTGCTCTTGTTCCCTGTTACCTCAAATCTGTGTACCA ATC". The third conserved domain comprised of 115 bases from P'_{709} to P'_{823}. This domain was represented by the sequence "AGGTTGCATTCTGTATTCACAATATT TCATACCA AGGAAGATTTGCTTTTTCAGACTATTCGATGCTCAATTTGCCT GCAGAGTACAAGGGCTCGTTTGATTTCATTGATGGGTA".

AUGUSTUS software identified two exons within the nucleotide sequence of 896 bases in the 0.9 Kb DNA fragment amplified in the present study. While the 1st exon had a length of 101 bases and was located between P'_{527}-P'_{627}, the 2nd exon had a length of 114 bases and was located between P'_{711}-P'_{824}. The software identified the segment comprising of 101 bp as an internal exon. All the introns identified in our study were AT rich. Similiar repots of AT rich regions in the introns of *Oryza glaberrima waxy* gene have been reported by Umeda *et al.* (1991). Camirand *et al.* (1990) and Van der Leij *et al.* (1991) have observed that all introns in the GBSS-I gene also followed the universal GT-AG rule. In this context our observations on the intron/exon architechture in the nucleotide sequence of the 1.2 Kb amplicon are in

agreement with those of Umeda *et al.* (1991), Camirand *et al.* (1990) and Van der Leij *et al.* (1991).

While the coding region of of 604 bases of the nucleotide sequence of 1.2 Kb amplicon and of 287 bases of the nucleotide sequence of the 0.9 Kb amplicon showed 80-90% homology with the coding region of other plant species, the non-coding regions identified in the sequence did not show any significant homology with GBSS genes. Xu *et al.* (2009), have reported similar results on the sequence homology of rye *waxy* gene with *waxy* genes of other plants. These results indicate a higher level of sequence conservation of GBSS genes in the coding region than the non coding region. Similar results have been reported for GBSS genes in sorghum, rice and maize (Chen *et al.*, 1998). EMBOSS CpGplot detected a putative CpG island of 856 bases located between positions 310 to 1116 in the nucleotide sequence of the 1.2 kb amplicon and of 495 bases located between position 355 to 854 in the nucleotide sequence of the 0.9 Kb amplicon. Detection of a putative CpG plot is an indication that the gene could be under tight regulatory control. Since CpG islands are known to be associated with the 5' end of genes which are subjected to regulatory control (Larsen *et al.*, 1992), detection of CpG Island in the DNA amplified in the current study indicates that the region of DNA amplified could belong to the 5' end of the target genes.

NCBI scan detected twelve palindromic structures in the nucleotide sequences of both 1.2 Kb amplicon 0.9 Kb amplicon. Jangsutthivorawat and Holkaert (2009) have correlated the presence of (CT)n and (AATT)n repeats at the GBSS-I loci of rice with variations in the amylose contenent in the starch. Bao *et al.* (2006b) have also indicated a close correlation beween the nbumber of repeats with

the AAC in starch grains of North American and Chinese germplasm. Alleles with fewer repeats (n<12) were observed to be associated with higher apparent amylose content (AAC) and those with more repeats (n>12) with a lower AAC. Identification of such repeats in the partial nucleotide sequence of the buckwheat GBSS-I gene in the present study, could be used as markers for screening of buckwheat accessions for their AAC.

The deduced amino acid sequence for the 1,116 bp bases of the 1.2 Kb amplicon comprised of 199 amino acid residues with a predicted isoelectric point (p*I*) of 9.75 and calculated molecular weight of 22.3 kDa. Sequence similarity analysis of the sequence with BLASTp against non-redundant protein database, identified the deduced sequence with the granule bound starch synthase family. Domain search on the deduced amino acid sequence identified the putative substrate binding site comprising of the sequence Lys-Thr-Gly-Gly-Leu (K-T-G-G-L) between P'_4 and P'_8. KTGGL is a universal motif identified in GBSS-I of *Pisum sativum* (Edwards *et al.*, 2002), *Vigna radiata* (Ko *et al.*, 2009), *Solanum tuberosum* (Edwards *et al.*,1999), *Zea mays* (Harn *et al.*, 1997), grain amaranth (Park *et al.*, 2009), *Triticum aestivum* (Anisworth *et al.*, 1993; Baba *et.al.*, 1993) and *Oryza sativa* (Sano *et al.*,1984). The domain has also been reported to be present in GBSSII and GBSSII of cassava (Munyikwa *et al.*, 1997), SSII of *Pisum sativum* (Edwards *et al.*, 2002), *Solanum tuberosum* (Edwards *et al.*,1999) and *Zea mays* (Harn *et al.*, 1997). The motif has been suggested to be involved in substrate binding (Furukawa *et al.*,1990, 1993). Using an alignment that permitted maximum homology, the sequence showed a maximum of 84% homology with deduced amino acid sequences of GBSS-I of *Nelumbo nucifera* (acc.no. ACM78591). The sequence also showed

83%, 80%, 79% and 79% homolgy with deduced amino acid sequences of GBSS-I of *Gossypium hirsutum* (acc. no. ACJ11735), *Phaseolus vulgaris* (acc. no. BAA82346), *Ipomoea batatas* (acc. no. BAI83439) and *Pisum sativum* (acc. no.CAC69955) and 80% homology with GBSSII of *Oryza sativa* (acc. no. ACY56082) and *Triticum aestivum* (acc. no. AAF14233). The deduced amino acid sequence identified in the preent study showed an insertion of 9 residues "HLHVLILES" at P'_{17} and the presence of three highly conserved domains between P'_1-P'_{16}, P'_{33}- P'_{51} and P'_{66}-P'_{90}. While the conserved domain between P'_1-P'_{16} comprised of 16 residues represented by the sequence "PWSKTGGLGDVLAALP", that between P'_{33}-P'_{51} comprised of 20 residues represented by the sequence "GHRVMTVAPRYDQYKDGWDT". The third conserved domain comprised of 25 residues between P'_{66} to P'_{90}. This domain was represented by the sequence "VRFFHCYKRGVDRVFVDHPMFLEKV". The alignment also releaved a high level of sequence similairty within the first 95 amino acid residues aligned out of the total 119 deduced amino acid residues. This segment showed similarity with the glycosyltransferase enzmyes which catalyze the transfer of a monosaccharide unit from an activated nucleotide sugar to an glycosyl acceptor molecule forming glycosidic bonds in carbohydrate residues or other biopolymers (Breton *et al.*, 2006). Motif search on the deduced amino acid sequence identified, a starch synthase catalytic domain predominantly represented by the amino acid residues PWSKTGGLGDVLAALPHLHVILESPALAARGHRVMTVAPRYDQYK DGWDTNVLVQIQVGERVETVRFFHCYKRGVDRVFVDHPMFLEKVTGPLGY LEEHMIKIVIFRRVTTTTFASGTQICIVRGGAGRKGVMAY" between positions 1-136. The software also detected three protein kinase C phosphorylation sites represented by the sequences Thr-Val-Arg ($P'_{65\text{-}67}$), Thr-Gly-Lys ($P'_{141\text{-}143}$) and Thr-

Gly-Arg ($P'_{176-178}$). When the deduced amino acid sequence were plotted as a function of hydropathic index, the sequence showed a predominantly hydrophobic character. Based on the hydropathic index of Kyte and Doolittle (1982), the major regions of hydrophilic nature detected in the sequence were between residues 5-28, 50-57, 77-95 and 102-139.Wang *et al.* (2000), have reported similar results on the hydrophobicity of *Ipomoea batatas* granule bound starch synthase-I (GBSS-I). Statistical analysis by SAPS (Brendel *et al.*, 1992) predicted the sequence to have 28.1% nonpolar residues, 18.6% polar uncharged residues and 22.6% polar charged residues.

Phylogenetic analysis of the deduced amino acid sequence reported in the present study with amino acid sequences of granule bound starch synthases available in EMBL database revealed a clear diversification into monocotyledonous (GBSS-II) and dicotyledonous (GBSS-I) groups. According to Pan *et al.* (2009), the putative protein could have duplicated and diverged into two different forms viz., GBSS-I or GBSSIa and GBSS-II or GBSSIb during evolution and the diversification into monocotyledonous (GBSS-II) and dicotyledonous (GBSS-I) groups might be a consequence of this diversification process.

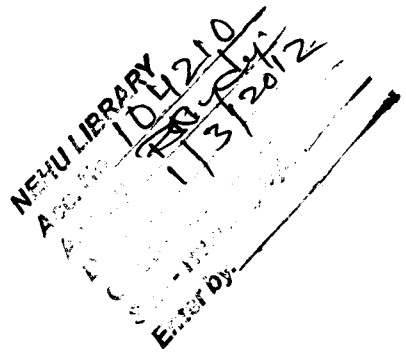
The deduced amino acid sequence for the 896 bases of the 0.9 Kb amplicon comprised of 70 amino acid residues with a predicted isoelectric point (*pI*) of 6.05 and a calculated molecular weight of 8.0 kDa. Sequence similarity analysis of the sequence with BLASTp against non-redundant protein database, identified the deduced sequence with the granule bound starch synthase family. Using an alignment that permitted maximum homology, the deduced amino acid sequence showed a maximum of 96% homology with granule bound starch synthase-I (GBSS-

I) of *Ipomoea batatas* (acc. no. BAI83439), *Ipomoea umbraticola* (acc. no. ABW83774) and *Solanum nemorense* (acc. no. AAY 63641) with a query coverage of 100%. The sequence also showed 95% homology with *Nicotiana tabacum* (acc. no. AAZ99063), *Hyoscyamus niger* (acc. no. AAZ99051), *Atropa belladonna* (acc. no. AAZ99047), *Jaborosa squarrosa* (acc. no. ABM46904), *Jaltomata procumbens* (acc. no. AAY63568), *Anisodus luridus* (acc. no. AAZ99048) and *Solanum johnstonni* (acc.no. ACV96018) and 93% homology with *Przewalskia tangutica* (acc. no. AAZ99055) respectively. BLASTp analysis of the sequence indicated that the entire 70 amino acid residues belongs to the glycosyltransferase superfamily. Motif search on the deduced amino acid sequence identified, the entire sequence as the catalytic domain of starch synthase.

A phylogenetic tree, describing the relationship of the deduced amino acid sequence reported in the present study with amino acid sequences of granule bound starch synthase proteins from other plants, was constructed by maximum parsimony method using the alignment matrix generated in the present study. The deduced amino acid sequence of 70 residues clustered together into one clad with amino acid sequences of *Ipomoea batatas* (acc. no. BAI83439) and *Ipomoea umbraticola* (acc.no. ABW83774).

Results obtained in the present study have revealed that the starch grains of common buckwheat showed a monomodal size distribution with size ranging from 3μm to 12μm. The grains can be clearly classified under the small grain size category and can thus have important applications in industrial processes. SEM of partially digested starch grains showed a clear pattern of concentric rings thereby revealing the lamellar structure of starch grains. This alternation of semi crystalline and amorphous

zones within the matrix of starch grains has been correlated with the presence of GBSS-I within the grains. Irrespective of grain size, the percentage of apparent amylose in buckwheat starch ranged between 47% to 51.9%. The starches showed a mean pasting temperature (*P*temp) of 68°C. Our results have revealed that a 59.7 kDa protein was associated with buckwheat starch grains as a granule bound protein. Based on N-terminal amino acid as well as serological analyses the protein has been identified as GBSS-I. This is the first report on the identification of GBSS-I in common buckwheat. Investigations carried out in the present study have led to the amplification and characterization of partial nucleotide sequences of GBSS-I genes from common buckwheat. The sequences showed features which were common to GBSS-I genes from dicots as well as monocots.



ISOLATION AND FUNCTIONAL CHARACTERIZATION OF THE *WAXY* LOCUS IN COMMON BUCKWHEAT (*FAGOPYRUM ESCULENTUM* MOENCH.)

By

NABANITA DEVADASAN

THESIS SUBMITTED
IN FULFILMENT OF THE DEGREE OF
DOCTOR OF PHILOSOPHY IN BOTANY

NORTH EASTERN HILL UNIVERSITY
SHILLONG-793022, INDIA
2010

CONTENTS

	Pages
Declaration	I
Acknowledgement	II - III
Dedication	IV
Abbreviation	V-VII
List of figures	VIII - XIII
List of tables	XIV

CHAPTERS

1. INTRODUCTION	1 - 7
2. REVIEW OF LITERATURE	8 - 24
3. MATERIALS AND METHODS	
(i) Materials	25
(ii) Methods	26 - 42
(iii) Chemicals and Solutions	43 - 47
4. ISOLATIONAL AND CHARACTERIZATION OF GRANULE BOUND STARCH SYNTHASE-I	
(i) Experimental	48 - 49
(ii) Results	49 - 55
(iii) Discussion	56 - 66

5. AMPLIFICATION AND SEQUENCING ANALYSIS OF *WAXY* LOCUS

(i) Experimental 67 - 68

(ii) Results 68 - 82

(iii) Discussion 82 - 91

6. GENERAL SUMMARY AND DISCUSSION 92 - 110

7. REFERENCES 111 - 134

The North Eastern Hill University

November, 2010

DECLARATION

I, Nabanita Devadasan, hereby declare that the subject matter of this thesis entitle "Isolation and functional characterization of the *waxy* locus in common buckwheat (*Fagopyrum esculentum* Moench.)" is the record of work done by me. The contents of this thesis did not form any basis of award of any previous degree to me or to the best of my knowledge to anybody else and that the thesis has not been submitted by me for any research degree in any other University/Institute.

This is being submitted to the North Eastern Hill University for the award of the degree of Doctor of Philosophy in Botany.

NABANITA DEVADASAN

Head of the Department
Head
Department of Botany
School of Life Sciences
N.E.H.U., Shillong-22

Supervisor 18/11/2010

ACKNOWLEDGEMENTS

I am deeply honoured to have been a part of NEHU and my six years of stay in this university have moulded my being today. I thank all the teaching staffs and non-teaching staffs of Botany Department, who are ever enthusiastic to render help.

I express my sincere gratitude to my supervisor Prof. N.K. Chrungoo for his constant encouragement and support and for motivating me in developing the innovativeness in my research work. I am in awe and admiration of his knowledge and ability to enunciate every little detail in the respective research areas. I am also thankful to him for examining the articles, reports and chapters of thesis without any complaints.

I am thankful to Prof. M.S. Dkhar and Prof. N.K. Chrungoo, the present and the past head, Department of Botany NEHU for providing the necessary facilities.

I would like to thank Prof. A.K. Misra, Prof. R.N. Sharan, and Prof. R. Sharma for providing all necessary help and the access to the laboratory facilities for executing my research work.

My special thanks to Dr. R. Lakadong who helped me in raising antibodies during the initial part of my research work.

I would like to thank all my Molecular Physiology labmates Dr. A. Rout, Dr. Sanga, Dr. S. Sangma, Dr. C. Jamir, N. Pathaw, A.G. Devi, R.K.J. Devi, L. Dohtdong, S. Barooah and T. Singh. Manoj deserves a special mention for always keeping things in place, letting my day start without a hitch.

I am forever indebted to my friends S. Das, E. Kharlyngdoh, P. Dkhar, Dr. P. Bhattacharjee, P. Bhattachajee, M. Burman, M. Elangbam, N. Nongpiur and R. Pramanik for their continuous help and cooperation.

Special thanks to A. M. Devi for her help and generosity in providing the printer willingly for printing my thesis.

All my work would have been incomplete without the prayers of Berean Baptist Church, California.

I would like to thank my parents, brother and in-laws for providing all the support needed. What I have inculcated through their faith in me has made my dream into reality.

I would like to thank my dear husband 'Jos' for his love, encouragement, understanding, and awesome support throughout the Ph.D. work.

It is quite possible that I might have missed names of few people in spite of their valuable assistance, both from a professional and personal perspective. I thank all of them.

I thank the University Grants Commission, New Delhi for providing financial support under the UGC Research Fellowship Science for Meritorious student, during my tenure as Research Scholar.

"Above all, I thank the Almighty".

TO MY PARENTS

ABBREVIATIONS

ADP	: Adinosine diphosphate
AGPase	: ADP glucose pyrophosphorylase
ADP-Glc	: ADP-glucose
BCIP/NBT	: 5-Bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium
BLAST	: Blast local alignment search tool
bp	: Base pair
CBB R-250	: Coomassie brilliant blue R-250
CH_3CN	: Acetonitrile
CHAPS	: 3 [3-cholamidopropyl] dimethylaminonia-1-propanesulsphonate
CLSM	: Confocal laser scanning microscope
CTAB	: Cetyl trimethyl ammonium bromide
dATP	: 2'Deoxyadenosine 5'- triphosphate
dCTP	: 2'Deoxycytosine 5'- triphosphate
dGTP	: 2'Deoxyguanosine 5'- triphosphate
DNA	: Deoxyribonucleic acid
DTT	: Dithiothreitol
dTTP	: 2'Deoxythymidine 5'- triphosphate
EDTA	: Ethylene di-amine tetra-acetic acid
EtBr	: Ethidium bromide
FAO	: Food and Agriculture Organization
GBSS	: Granule bound starch synthase
Glc-1-P	: Glucose-1-phosphate
I_2	: Iodine

IEF : Isoelectric focusing

IgG : Immunoglobulin G

IPG : Immobilized pH gradient

Kb : Kilo base

kDa : Kilo dalton

KI : Potassium iodide

mA : Milliampere

ME : β-Mercaptoethanol

MES : 2-[*N*-morpholino] ethanesulfonic acid

min : Minute

mM : Millimolar

M.wt. : Molecular weight

NH_4HCO_3 : Ammonium bicarbonate

PBS : Phosphate buffered saline

PCR : Polymerase chain reaction

pI : Isoelectric point

PMF : Peptide mass fingerprinting

PMSF : Phenyl methyl sulfonyl fluoride

PVDF : Polyvinylidene fluoride

PVP : Polyvinyl pyrrolidone

RDS : Rapidly digestible starches

Rf : Relative front/mobility

rpm : Revolutions per minute

RS : Resistant starch

SBE : Starch branching enzyme

SDE : Starch debranching enzyme

SDS : Slow digestible starches

SDS : Sodium dodecyl sulphate

SDS-PAGE : Sodium dodecyl sulphate-poly acrylamide gel electrophoresis

SEM : Scanning electron microscopy

SS : Starch synthases

SSSs : Soluble starch synthases

TBE : Tris-borate-EDTA

TBS : Tris-Cl buffered saline

TCA : Trichloro acetic acid

TEMED : N', N', N', N'-Tetramethylethylenediamine

TFA : Trifluoro acetic acid

Tm : Melting temperature

Tris : Tris (hydroxymethyl) aminomethane

TTBS : Tween–Tris-Cl buffered saline

UV/ VIS : Ultra violet/visible

μL : Microlitre

μM : Micrometer

LIST OF FIGURES

Fig. 1.1 : Schematic diagram of the structure of amylose and amylopectin.

Fig. 1.2 : Schematic model of starch biosynthesis in plants, known as the pre-amylopectin trimming model.

Fig. 2.1 : Utilization of starch in various food and non-food industries.

Fig. 2.2 : Schematic representation of the structure of a starch granule, with alternating amorphous and semi-crystalline regions constituting the growth ring.

Fig. 3.1 : A close up view of common buckwheat *(Fagopyrum esculentum* Moench.) plant at the flowering stage.

Fig. 3.2 : A close up view of dehulled seeds of common buckwheat (*Fagopyrum esculentum* Moench.).

Fig. 3.3 : An Ouchterlony immunodiffusion plate, after immune diffusion has taken place.

Fig. 3.4 : Schematic representation of a typical three step cycle parameter for PCR amplification.

Fig. 4.1 : Scanning electron micrographs of starch grains isolated from different Indian and European accessions/cultivars of buckwheat. (a)IC-188669, (b) KBB-3, (c) IC-13145, (d) VL-7, (e) OC-2, (f) Siva, and (g) Daria, (h) Starch granules covered with a proteinaceous membrane, (i) enlarged image of a starch granule, (j) Monomodal arrangement of buckwheat starch grains, (k, l) Mature starch grains fractured by sonication and incubated with 3u/ml of alpha amylase for 15 mins at room temperature.

Fig. 4.2 : Light microscopy image of matured starch granules stained with 0.2% KI, 0.4% I w/v.

Fig. 4.3 : RVA profile of normal buckwheat starch for viscosity and temperature as a function of time.

Fig. 4.4 : SDS-PAGE profile of granule associated proteins isolated from the endosperm tissues of different Slovenian varieties/selections of buckwheat.
(a):-{M}=Protein molecular weight marker, L1= Bf 06X12, L2= Gor 0619 Siva Min Selo, L3= 41 Min Gor05, L4=Siva Kontrola Rangus PR03, L5=53 Gor Siva Min 05, L6=Gor 0620 Siva Min

Selo.
(b):-{M}=Protein molecular weight marker, L1=Gor 1304 BF 0611, L2=63 Min Gor Siva 05, L3=Gor 0628 Siva Min Selo, L4=Gor 0413 BF 0618, L5=Gor 0413 BF 0618.
(c):-{M)=Protein molecular weight marker L1=Gor 11 2004 3 Crte, L2=Gor 2 2004 3 Crte, L3=Gor 4 2004 3 Crte, L4=Gor 10 2004 3 Crte, L5=191 2X rumena 3 Crte.
(d):-{M}=Protein molecular weight marker, L1=Gor 93 Crte, L2=7 Roza Crte, L3= 10 Rdeca 3 Crte.

Fig. 4.5 : SDS-PAGE profile of granule associated protein isolated from leaf tissues (L1) and endosperm tissues (L2) of Indian accessions/ varieties of buckwheat. M=Marker.

Fig. 4.6 : N-terminal amino acid sequence for the 59.7 kDa starch granule bound protein of common buckwheat.

Fig. 4.7 : Comparison of N-terminal amino acid sequences of the 59.7 kDa GBSS-I protein isolated from buckwheat starch granules with the N-terminal amino acid sequences of GBSS-I proteins from other crops.

Fig. 4.8 : Immunoblot profile of starch granule associated proteins separated by SDS-PAGE with antibodies raised against buckwheat GBSS-I. Cross–reacting proteins were visualized by specific anti-59.7 kDa antibody and HRP-goat anti-rabbit IgG conjugate.
(a) L1=Bf 06X12, L2=Gor 0619 Siva Min Selo, L3=41 Min Gor05, L4=Siva Kontrola Rangus PR03, L5=53 Gor Siva Min 05, L6=Gor 0620 Siva Min Selo.
(b) L1=Gor 1304 BF 0611, L2=63 Min Gor Siva 05, L3=Gor 0628 Siva Min Selo, L4=Gor 0413 BF 0618, L5=Gor 0413 BF 0618.
(c) L1=Gor 11 2004 3 Crte, L2=Gor 2 2004 3 Crte, L3=Gor 4 2004 3 Crte, L4=Gor 10 20043 Crte, L5=191 2X rumena 3 Crte.
(d) L1=Gor 93 Crte, L2=7 Roza Crte, L3=10 Rdeca 3 Crte.

Fig. 4.9 : Immunoblot profile of the starch granule associated proteins isolated from leaf (L1) and endosperm starch grains (L2) of common buckwheat with anti-59.7 kDa serum raised against buckwheat GBSS-I. Cross-reacting proteins were visualized by specific anti-59.7 kDa antibody and HRP-goat anti-rabbit IgG conjugate.

Fig. 4.10a : SDS-PAGE profile of GBSS-I proteins isolated from endosperm starch grains of maize (L1), rice (L2) and wheat (L3). M= protein

molecular weight marker.

Fig. 4.10b : Immunoblot profile of GBSS-I proteins isolated from starch grains of maize (L1), rice (L2) and wheat (L3) cross-reacted with anti-59.7 kDa serum raised against buckwheat GBSS-I. Cross-reacting proteins were visualized by specific anti-59.7 kDa antibody and HRP-goat anti-rabbit IgG conjugate.

Fig. 4.11 : SDS-PAGE profile of starch granule associated protein isolated from endosperm tissues of buckwheat and rice and leaf tissue of buckwheat.
L1 and L2= undigested 59.7 kDa and 60 kDa GBSS-I protein isolated from endosperm starches of buckwheat and rice.
L3= undigested 53 kDa starch granule associated protein isolated from leaf starches of common buckwheat.
L4 & L5= Trypsin digestion of 59.7 kDa and 60 kDa GBSS-I proteins isolated from endosperm starches of common buckwheat and rice.
L6= Trypsin digestion of 53 kDa starch granule associated protein isolated from leaf starches of common buckwheat.

Fig. 4.12a : 2D-PAGE profile of the endosperm starch granule associated proteins isolated from common buckwheat.

Fig. 4.12b : Immunoblot profile of the endosperm starch granule associated proteins isolated from common buckwheat with anti-59.7 kDa serum raised against buckwheat GBSS-I.

Fig. 4.13 : MS analysis of tryptic peptide mass fingerprinting of protein spot no3 of molecular mass 59.7 kDa isolated from common buckwheat starch granule. The upper values indicate the observed mass peaks. The lower values labelled the corresponding regions in the proteins.

Fig. 4.14 : 2D-PAGE profile of starch granule associated protein isolated from leaf tissues of common buckwheat.

Fig. 4.15 : Images of starch granules: (a) Bright field light micrographs of unstained starch granules; (b) Confocal laser-scanning micrographs of starch granules.Physiochemical localization of starch granule associated proteins were carried out by immunodetection with fluorescent labelled secondary antibody (Alexa-546 conjugated goat anti-rabbit IgG, Invitrogen).

Fig. 4.16 : Phylogenetic tree based on alignment matrix of N-terminal amino acid sequence of 59.7 kDa GBSS-I protein of common buckwheat with the amino acid sequences of GBSS type proteins available in

the protein data bank.

Fig. 5.1 : Total genomic DNA isolated from etiolated seedlings of buckwheat (*Fagopyrum esculentum* Moench.)
L1= isolated genomic DNA; M= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.2 : Restriction digestion profile of genomic DNA isolated from etiolated seedlings of buckwheat (*Fagopyrum esculentum* Moench.) (L1)=gDNA digested with *Eco*RI, (L2)= gDNA digested with *Hind*III, (L3)=gDNA digested with *Nco*I, (M)= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.3a : Electrophoresis profile of amplified fragment with buckwheat gDNA as template and primer pair NDF2/NDR2, M= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.3b : Nucleotide sequence of 690 bases for the 0.7 Kb fragment amplified with buckwheat gDNA as template and primer pair NDF2/NDR2.

Fig. 5.4a : Electrophoresis profile of amplified fragment with buckwheat gDNA as template and primer pair NDF3/NDR3, M= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.4b : Nucleotide sequence of 704 bases for the 0.7 Kb fragment amplified with buckwheat gDNA as template and primer pair NDF3/NDR3.

Fig. 5.5 : Electrophoresis profile of amplified fragment with buckwheat gDNA as template and primer pair NDF7/NDR6, M= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.6 : Nucleotide sequence of 1116 bases for the 1.2 Kb fragment amplified with buckwheat gDNA as template and primer pair NDF7/NDR6.The exons identified in the sequence are highlighted in green and their start and end positions are indicated with forward and backward arrows.

Fig. 5.7 : ClustalW (1.81) multi-alignment of the 1116 bp nucleotide sequences for the 1.2 kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF7/NDR6 with nucleotide sequences of GBSS genes available in the database. '*' denotes conserved residues.

Fig. 5.8 : Deduced amino acid sequences of the coding region of 1116 bp nucleotide sequences for the 1.2 Kb amplicon generated from buckwheat gDNA as the template and primer pair NDF7/NDR6.

Fig. 5.9 : ClustalW (1.81) multi-alignment of the deduced amino acid sequences for the 1.2 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF7/NDR6 with amino acid sequences of GBSS proteins of other plants available in the database. '*' denotes conserved residues and ':' denotes invariant or similar residues.

Fig. 5.10 : Screen shot of the BLAST of the 199 deduced amino acid sequence of the 1.2 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF7/NDR6 showing the conserved position of the 95 amino acids among GBSS protein sequences in different plant species.

Fig. 5.11a : Diagrammatic representation of the starch synthase domains positioned between 1 to 136 amino acid residues in the 199 deduced amino acid sequences as predicted by MOTIF SCAN.

Fig. 5.11b : Diagrammatic representation of the Protein Kinase C phosphorylation site in the 199 deduced amino acid sequence as predicted by MOTIF SCAN.

Fig. 5.12 : Hydropathy profile of the deduced 199 amino acid sequence for the 1.2 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF7/NDR6. The X-axis represents the residue number.

Fig. 5. 13 : Secondary structure prediction profile of the 199 deduced amino acid sequence for the 1.2 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF7/NDR6.

Fig. 5.14 : Phylogenetic tree based on the alignment matrix of the 199 deduced amino acid sequences with amino acid sequences of GBSS proteins of various plants available in the Genbank.

Fig. 5.15 : Electrophoresis profile of amplified fragment with buckwheat gDNA as template and primer pair NDF8/NDR6, M= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.16 : Nucleotide sequence of 896 bases for the 0.9 Kb fragment amplified with buckwheat gDNA as template and primer pair NDF8/NDR6.The exons identified in the sequence are highlighted in green and their start and end positions are indicated with forward and backward arrow.

Fig. 5.17 : ClustalW (1.81) multi-alignment of the 896 bp nucleotide sequences for the 0.9 Kb amplicon generated with buckwheat

genomic DNA as template and primer pair NDF8/NDR6 with nucleotide sequences of GBSS genes available in the database. '*' denotes conserved residues.

Fig.5.18 : Deduced amino acid sequences of the coding region of 896 bp nucleotide sequences for the 0.9 Kb amplicon generated from buckwheat gDNA as the template and primer pair NDF8/NDR6.

Fig. 5.19 : Screen shot of the BLAST of 70 deduced amino acid sequence of the 0.9 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF8/NDR6 showing the conserved position of the all the amino acids among GBSS protein sequences in different plant species.

Fig. 5.20 : ClustalW (1.81) multi-alignment of the deduced amino acid sequences for the 0.9 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF8/NDR6 with amino acid sequences of GBSS proteins of other plants available in the database. '*' denotes conserved residues and ':' denotes invariant or similar residues.

Fig. 5.21a : Diagrammatic representation of the starch synthase domains positioned between 1 to 70 amino acid residues in the 70 deduced amino acid sequence as predicted by MOTIF SCAN.

Fig. 5.21b : Diagrammatic representation of the N-glycosylation sites in the 70 deduced amino acid sequence as predicted by MOTIF SCAN.

Fig. 5.22 : Hydropathy profile of the deduced 70 amino acid sequence for the 0.9 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF8/NDR6. The X-axis represents the residue number.

Fig. 5. 23 : Secondary structure prediction profile of the 70 deduced amino acid sequence for the 0.9 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF8/NDR6.

Fig. 5.24 : Phylogenetic tree based on the alignment matrix of the 70 deduced amino acid sequences with amino acid sequences of GBSS proteins of various plants available in the Genbank.

LIST OF TABLES

Table 1.1 : World-wide starch production and the major crops sources.

Table 2.1 : Characteristics of starch granules from diverse botanical sources.

Table 2.2 : Physicochemical characteristics of amylose and amylopectin.

Table 4.1(a) : Amylose content, expressed as mg/100 mg starch in common buckwheat (*F. esculentum* Moench.) at different stages of seed maturation.

Table 4.1(b) : Amylose content, expressed as mg/100 mg in certified accessions/cultivars of *F. esculentum* L. procured from India and Europe.

Table 4.2 (a) : Rheological properties of buckwheat starch paste.

Table 4.2 (b) : Amylose content and rheological properties of starch obtained from different plant sources.

Table 4.3 : Identification of the Tryptic Fragment of GBSS-I (spot no.3 of molecular wt. 59.7 kDa) by comparison with rice genome using MASCOT software analysis tool.

Table 5.1 (a) : List of primers used for PCR amplification of GBSS-I protein genes.

Table 5.1(b) : List of primer combinations and PCR annealing temperatures used for the amplification of GBSS-I genes from buckwheat genomic DNA.

Table 5.2 : Palindromic structures detected in the 1116 bp nucleotide sequence of the 1.2 Kb amplicon generated with primer pair NDF7-NDR6.

Table 5.3 : Simple sequence repeates found in the 1116 bp nucleotide sequence of the 1.2 Kb amplicon generated with primer pair NDF7-NDR6.

Table 5.4 : Palindromic structures detected in the 896 bp nucleotide sequence of the 0.9 Kb amplicon generated with primer pair NDF8-NDR6.

Table 5.5 : Simple sequence repeates found in the 896 bp nucleotide sequence of the 0.9 Kb amplicon generated with primer pair NDF8-NDR6.

CHAPTER I

INTRODUCTION

Starch is one of the most important plant products to man. It is an important insoluble polymer that occurs in plant cells in the form of a complex granular mixture of amylose and amylopectin. While amylose comprises largely of unbranched α-1, 4-linked glucan chains, amylopectin is comprised of small size α-1, 4- linked chains that are clustered together by α-1, 6- linkages between adjoining straight glucan chains (Tetlow *et al.*, 2004a).

Starch granules constitute the major carbohydrate storage molecules for many plant cell types. In view of the importance of carbohydrates as an important storage reserve of plant cells, the biosynthesis of the starch polymers *viz.*, amylopectin and amylose are a central aspect of plant metabolism. The great majority of the granular mass, approximately 98-99%, is made up of two homopolymers of α-D-glucosyl units, amylose and amylopectin. In both the polymers, glucosyl units are joined by α-(1-4)

glycosidic linkages to form linear chains, with branch points introduced by α-(1-6) glycoside bonds (Fig. 1.1). In amylose, less than 1% of the glucose units participate in α-(1-6) bonds, so these molecules are essentially linear. Amylopectin, in contrast, is a moderately branched polymer having approximately 5-6% branches. The size, shape and structure of the granules vary substantially among botanical sources. Each starch granule has a highly organized structure defined by the succession of semi-crystalline and amorphous lamellae. The semi-crystalline and amorphous lamellae are determined mainly by the arrangement of amylopectin chains within granules (Hizukuri, 1986). Amylose content is considered the most important characteristics for predicting cooking and processing behaviour of world's important food crops. Texture and quality of white salted (udon) noodles is better in partial *waxy* wheat as compared with normal types (Wang and Seib, 1996; Batey *et al.*, 1997; Briney *et al.*, 1997). In bread, the rate of starch retrogradation and staling can be manipulated by adjusting the amylose content relative to amylopectin (Schoch, 1965). Expanded snack foods are also dependent on the content of amylose and amylopectin. In theory, the final amylose content of any product or process could be manipulated between 0% (using *waxy* starch) and ~25% (using normal starch) on a starch basis to obtain starch with different physical and chemical properties (Graybosch, 1998; Bligh, 1999).

Starch synthesis in higher plants occurs in specialized organelles, chloroplasts for transient starch and amyloplasts for storage starch. The process involves coordinated interactions between suites of starch biosynthetic enzymes including ADP glucose pyrophosphorylase (AGPase), starch synthases (SS), starch branching enzymes (SBE)

l Schematic diagram of the structure of amylose and amylopectin.

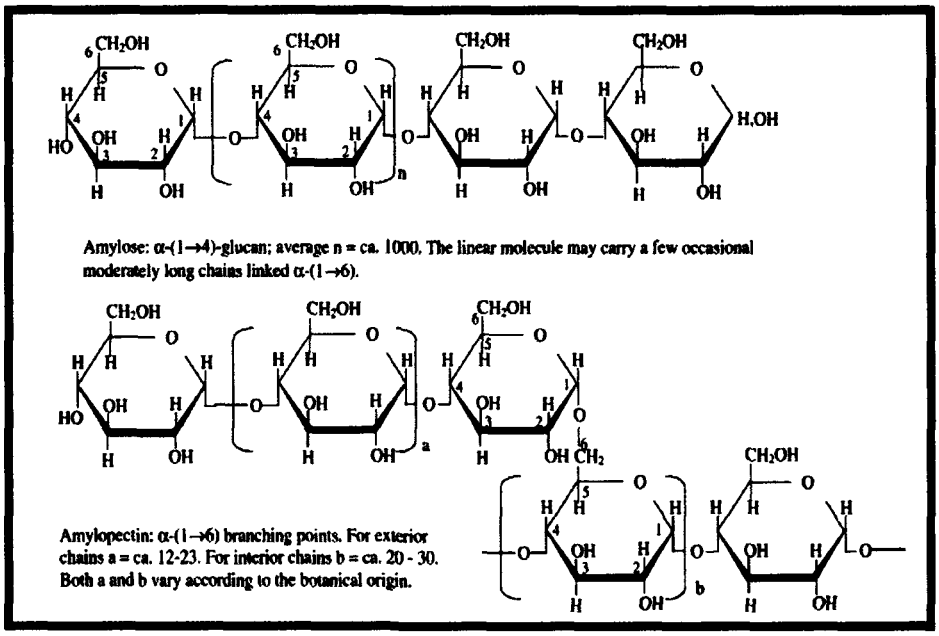

Amylose: α-(1→4)-glucan; average n = ca. 1000. The linear molecule may carry a few occasional moderately long chains linked α-(1→6).
Amylopectin: α-(1→6) branching points. For exterior chains a = ca. 12-23. For interior chains b = ca. 20 - 30. Both a and b vary according to the botanical origin.

Fig. 1.1

and starch debranching enzymes (DBE) with individual plant species having multiple forms of each enzyme. For example, plant contain a granule-bound SS (GBSSI), which is essentially for amylose synthesis and four or five soluble SS types (SSI-SSV), each of which plays a distinct role in amylopectin synthesis (Smith *et al.*, 1997; Myers *et al.*, 2000; Denyer *et al.*, 2001; Ball and Morell, 2003; James *et al.*, 2003; Hirose and Terao, 2004; Deschamps *et al.*, 2008) (Fig. 1.2). Many starch biosynthetic enzymes are distributed between the soluble fraction of the plastids and the insoluble starch granules themselves (Ball and Morell, 2003; Tetlow *et al.*, 2004a). An important aspect involving investigation of starch biosynthetic pathway is that most of the enzymes involved come in multiple forms, and differ in their physical and chemical properties along with the type of starch they produce. Multiple forms of both SSSs (SSI, SSII and SSIII) and GBSSs have been reported (Nelson and Rines, 1962; Boyer and Preiss, 1978; Macdonald and Preiss, 1985; Dang and Boyer, 1988). Although this increases the complexity of the process, and makes it more difficult to interpret the role of the different enzymes, it also increases the range of possibilities for bioengineering.

One of the most important intriguing challenges in understanding how starch is made is to explain how the two different polymers which constitute starch, namely amylopectin and amylose, are synthesized in the same place at the same time. The glucan biosynthetic system in plants is complex in the regard that multiple forms of each enzyme have been conserved in evolution and are co-expressed at the same site and time of starch accumulation. Granule-associated proteins can be divided into two categories according to their degree of association with starch granules. One group consists of

Fig. 1.2 Schematic model of starch biosynthesis in plants, known as the pre-amylopectin trimming model.

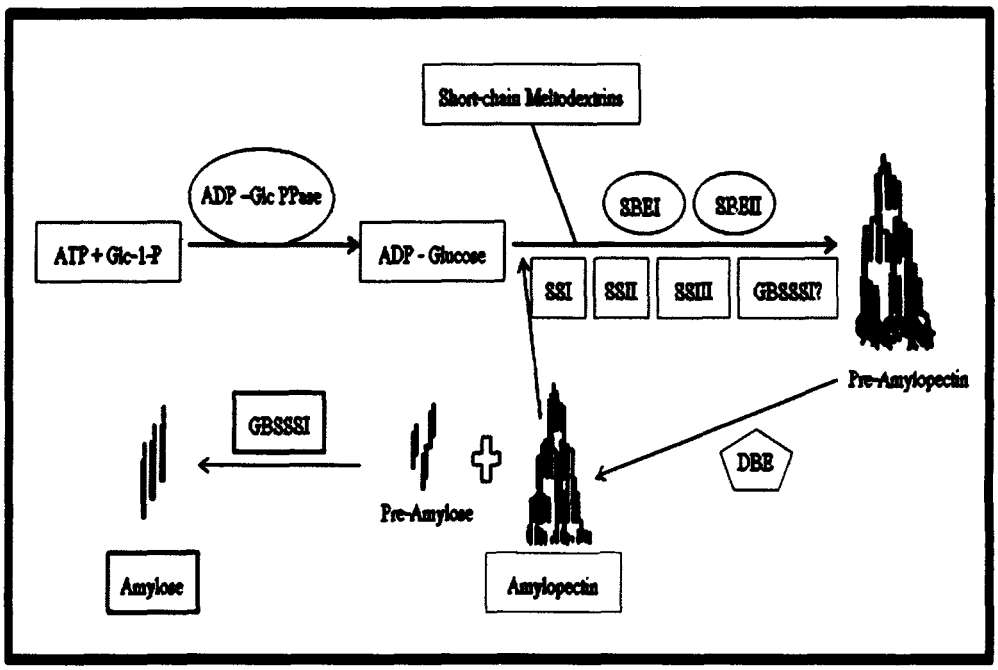

Short-chain Meltodextrins
ADP -Glc PPase
SBEI
SBEII
ATP + Glc-1-P
ADP - Glucose
SSI
SSII
SSIII
GBSSSI?
Pre-Amylopectin
GBSSSI
DBE
Pre-Amylose
Amylose
Amylopectin

Fig. 1.2

surface-associated proteins and the other comprises internal granule-associated proteins. Surface-associated proteins have a small degree of association with the starch granule and can be separated from the granule by protease digestion (with thermolysine, proteinase K) or by extensive washing in aqueous buffer containing a detergent and a reducing agent, at a temperature below the gelatinization temperature for starch (Denyer *et al.*, 1993; Rahman *et al.*, 1995; Boren *et al.*, 2004). By contrast, extraction of internal granule-associated proteins requires the gelatinization of starch granule in SDS in the presence of a reducing agent (Tsai, 1974; Echt and Schwartz, 1981; Denyer *et al.*, 1993). Studies in many plant species have shown that starch biosynthetic enzymes exist as both as both soluble proteins in the stroma and as internal granule-associated proteins.

The use of starch as a renewable and biodegradable polymer is becoming increasingly attractive because of the environmental concerns about the industrial wastes generated from petroleum products. Since, starch is one of the principal constituent of the harvestable organ of many agronomic plants; increase in starch quantity increases crop yield. Most of the starch utilized world-wide comes from a relatively small number of crops, the most important being maize, potato, wheat and tapioca with smaller amounts from rice, sorghum, sweet potato, arrowroot, sago and mung-beans (Table 1.1). In general, starches from tapioca and sorghum are used only for food and other, non-food purposes. Since the 1940's, demand for starch with desirable properties has led to the modification of starches for food and commercial

Table 1.1: World-wide starch production and the major crops sources.

	Quantity (10^6 t) in 2003		
PRODUCTS	**USA**	**EUROPE**	**OTHER COUNTRIES**
Maize starch	29.2	4.0	10.9
Wheat starch	0.5	3.0	1.1
Potato Starch	0.1	1.9	0.8
Other Starches	NA	0.1	2.5
TOTAL	29.8	9.0	15.3

Source: EU Commission (DG AGRI, Unit C2), United States Department of Agriculture (USDA), LMC Internal Database.

food products. With increasing industrial demand for starches, there is need to explore new and alternative sources of starch.

While much work has been done on the characterization of starch from cereals and analysis of the regulation of its biosynthesis, not much information is available about the quality of starch in many other potentially important crops. One such group of crops belongs to pseudo cereals. Among these groups of plants, common buckwheat is an important pseudo cereal because of its high potential for use as a functional food. The National Bureau of Plant Genetic Resources (NBPGR) New Delhi has identified buckwheat as an important underutilized crop which has potential to become an important component of human diet. Common buckwheat (*Fagopyrum esculentum* Moench.) belongs to the family Polygonaceae. In North eastern region of India, the distribution of *Fagopyrum* spp. has been reported from the states of Sikkim, Meghalaya, Arunachal Pradesh, Assam, Manipur and Nagaland (Anonymous, 1987-2001). According to Joshi (1999), buckwheat cultivation was confined to Arunachal Pradesh and Sikkim. Buckwheat is one of the nutritious crops of the mountain regions and it is the only crop, which can be grown up to 4500 m (Joshi and Paroda, 1991). Considering its various uses, early maturity, good adaptation and suitability for marginal production areas, the Indian Council of Agricultural Research (ICAR) started research work on this crop in 1982 under AICRP (All India Coordinated Research Project) on Under-Utilized Crops. Considering its high nutritive value, buckwheat was also included under Jai Vigyan Project under the Food and Nutrition Programme of AICRP on Under-Utilized

Crops. Buckwheat is an important crop in China, Russia, Ukraine, and Kazakhstan, parts of Europe, Canada, Japan, Korea and Nepal.

Buckwheat is one of the best sources of high quality, easily digestible protein in the plant kingdom. It has over 90% of the value of non-fat milk solids and over 80% of whole egg solids (Udesky, 1992). The balanced amino acid profile and a high level of essential amino acids allow buckwheat to be used in human diets, especially where shortages of lysine and sulfur containing amino acids appear. The specific ability of buckwheat proteins to soak up cholesterol from food and the relative proportions of its amino acids, make buckwheat the best cholesterol-lowering food substitute. The addition of buckwheat to human diets could also help cut calories and keep blood sugars at optimal levels. Buckwheat grain is digested more slowly than other carbohydrates (Buckwheat diet, 1989).

The dehulled buckwheat seed, or groat, is used in breakfast cereals and milled into grits (Robinson, 1980). Roasted groats which are called kasha, are sold in whole and granulated forms. Both Kasha and groats can be baked, steamed, or boiled for nutritious alternatives to potatoes and rice. It is used in pancake mixes as well as in various breads. It is often blended with wheat flour for use in bread, pasta products, and some breakfast cereals (Robinson, 1980). Studies have shown that up to 60% buckwheat flour mixed with wheat flour produced acceptable bread (Pomeranz, 1983). In general, buckwheat flour in bread mixes comprises only 30% to 40% of the total. Buckwheat flour may also be used in deserts, ice cream cones, dietetic foods, pancakes mixes, canned meat products, canned vegetable products, etc. Buckwheat can be used to

produce extruded cereal and snack products. Extruded buckwheat products are of very high nutritional quality when compared with products extruded from maize, wheat, or barley alone. The specific character of the proteins allows buckwheat to potentially be used in extruded products targeted to special nutritional needs.

Skrabanja *et al.* (2001), have shown that starch from buckwheat grains have a strong potential for use in the development of functional foods including its engineering for a modified course of starch digestion. However, despite the wide utilization of buckwheat starch in various food products the flour obtained from buckwheat has poor dough making qualities due to its unfavorable amylose/amylopectin ratio, adding to its less popularity as a flour-yielding crop. Since genetic manipulation of genes involved in starch biosynthesis pathway can eventually provide means of modifying the amylose/amylopectin ratio, the present investigation was carried out to isolate and characterize the *waxy* gene in common buckwheat. We assume that by targeting this enzyme it would be possible to alter the amylose/amylopectin ratio in buckwheat with a consequent alteration of its starch quality.

CHAPTER II

REVIEW OF LITERATURE

Starch is a naturally occurring polymer of α-D glucose. It is the main energy reservoir of higher plants and also a major source of dietary energy for humans and animals. As an additive for food processing, food starches are typically used as thickeners and stabilizers in foods such as puddings, custards, soups, sauces, gravies, pie fillings, and salad dressings, and to make noodles and pastas. Starch is also widely used in the preparation of foods such as bread, pancakes, cereals, noodles, pasta, porridge and tortilla (Eliason, 2004). On the basis of nutritional characteristics which is mainly based on the extent of digestibility, starch in foods can be classified as: (i) Digestible starches (Berry, 1986) and (ii) Resistant starch (Englyst *et al.*, 1982). Digestible starches include the starches digestible by body enzymes, namely the rapidly digestible starches (RDS) and the slow digestible starches (SDS). Digestible starches consists mainly of amorphous and dispersed starch, found in high amounts in starchy foods cooked by

moist heat, such as bread, potato and cereals. Resistant starch (RS) is not digested or absorbed in the small intestine and are fermented by bacterial microflora in the large bowel, affecting a number of physiological functions and thus having different effects on health, e.g., reduction of the glycemic and insulinemic responses to food, hypocholesterolemic action and protective effects against colorectal cancer (Asp *et al.*, 1996).

Besides its nutritive values, starch is a very useful raw material with a wide range of applications in both food and non-food industries (Fig. 2.1). Although the history of starch usage by man has been variously described for thousands of years, the properties of starch have been extensively studied and discussed over the last two centuries. The Romans used starch to whiten cloth as early as 100 B.C. and around 300 A.D. starch was used to stiffen cloth and mixed with dyes to color cloths (Robyt, 2008). Since then, the applications of starch in industries have rapidly increased, thereby increasing its commercial value. Starch contributes greatly to the textural properties of many foods and has many industrial applications including its use as a thickener, colloidal stabilizer, gelling agent, bulking agent, water retention agent and adhesive. Although native starch does have its industrial applications, most of the time the industry requires improved functionality from modified starch. Designer starches can be used for bulk or value-added food and feed and in non-food applications as in chemicals or as bioenergy (Morell and Myers, 2005). Thus, a major challenge would be to predict the effect of structural/chemical changes in starches on their functional properties.

Fig 2.1 Utilization of starch in various food and non-food industries.

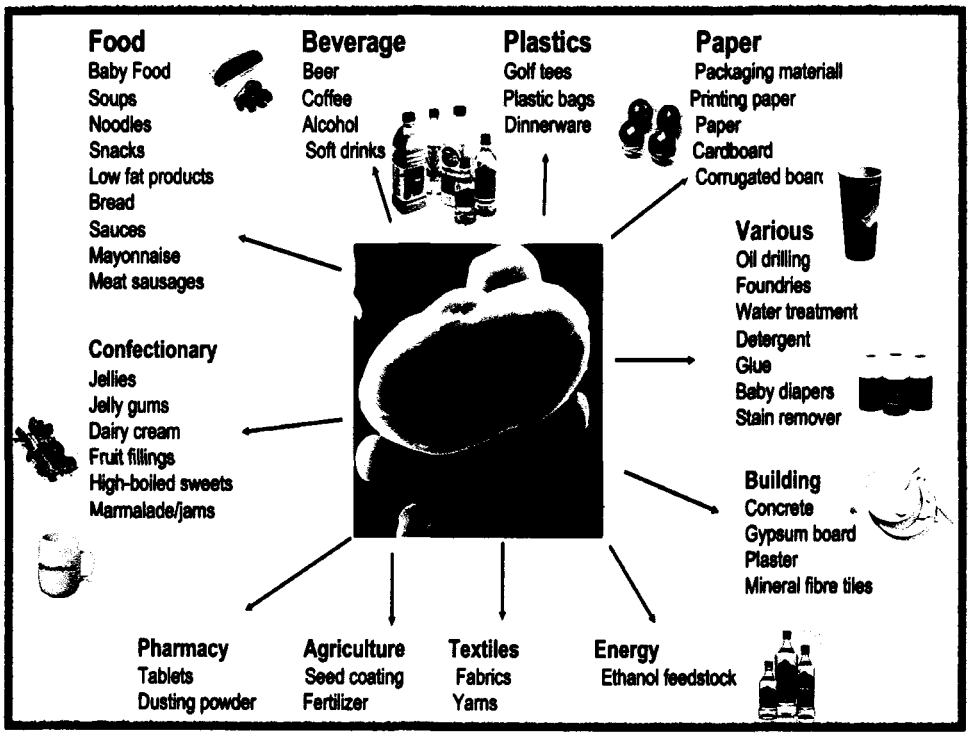

Food
Baby Food
Soups
Noodles
Snacks
Low fat products
Bread
Sauces
Mayonnaise
Meat sausages
Confectionary
Jellies
Jelly gums
Dairy cream
Fruit fillings
High-boiled sweets
Marmalade/jams
Beverage
Beer
Coffee
Alcohol
Soft drinks
Plastics
Golf tees
Plastic bags
Dinnerware
Paper
Packaging material
Printing paper
Paper
Cardboard
Corrugated board
Various
Oil drilling
Foundries
Water treatment
Detergent
Glue
Baby diapers
Stain remover
Building
Concrete
Gypsum board
Plaster
Mineral fibre tiles
Pharmacy
Tablets
Dusting powder
Agriculture
Seed coating
Fertilizer
Textiles
Fabrics
Yarns
Energy
Ethanol feedstock

Fig. 2.1

Starch is synthesized in the form of granules with a partially crystalline texture. The morphology, chemical composition and super molecular structure of starch granules are known to be characteristic to each plant species (Banks and Greenwood, 1975). Depending upon the species and the tissue, the grains vary in size from 1 to 100 μm in diameter and in shape from round to polygonal (Mäkelä and Laakso, 2006). In some cases even lenticular grains have been reported (Fredriksson *et al.*, 1998; Tester *et al.*, 2004). While most of the cereal and tuber starches have unimodal size distribution (Fredriksson *et al.*, 1998; Tester and Karkalas, 2002; Tester *et al.*, 2004; Mäkelä and Laakso, 2006), starch grains of wheat, barley, rye and triticale have been reported to show a bimodal size distribution (Evers, 1971, 1974; Simmonds and O'Brien, 1981; Dengate and Meredith, 1984; Morrison and Scott, 1986; Stoddard, 1999; Lindeboom *et al.*, 2004). The starch granules in these plants have been classified into two types *viz.*, "A" type which are lenticular in shape with a grain diameter ranging between 10-20 μm and "B" type which are spherical in shape with a diameter of <10μm (Eliason and Gudmundsson, 1996; Stoddard, 1999; Lindeboom *et al.*, 2004). Rice starch and oat starch granules tend to exist in clusters of individual granules. Such grains have been designated as compound granules (Juliano, 1984; Hoover *et al.*, 2003; Bao and Bergman, 2004; Tester *et al.*, 2004) (Table 2.1).

Starch owes much of its functionality to two major high molecular weight carbohydrate components *viz.*, amylose and amylopectin as well as to the physical organization of these molecules within the granule structure (French, 1984). Even though both amylose and amylopectin contain polymers of α-D-glucose units, they differ

Table 2.1: Characteristics of starch granules from diverse botanical sources

Source	Type	Distribution Pattern	Granule Shape	Mean Diameter (μm)
Hordeum vulgare	Cereal	Bimodal	Lenticular (A-type) Spherical (B-type)	15-20 2-5
Secale cereale	Cereal	Bimodal	Lenticular (A-type) Spherical (B-type)	10-40 5-10
Triticum aestivum	Cereal	Bimodal	Lenticular (A-type) Spherical (B-type)	15-35 2-10
Avena sativa	Cereal	Unimodal	Polyhedral	80
Oryza sativa	Cereal	Unimodal	Polyhedral	150
Sorghum bicolor	Cereal	Unimodal	Spherical	5-20
Triticosecale	Cereal	Unimodal	Spherical	1-30
Setaria italica	Cereal	Unimodal	Polyhedral	4-12
Zea mays	Cereal	Unimodal	Spherical/Polyhedral	2-30
Pisum sativum	Legume	Unimodal	Reniform	5-10
Solanum tuberosum	Tuber	Unimodal	Spherical/ Lenticular	5-100
Manihot esculenta	Tuber	Unimodal	Spherical/Lenticular	5-35

in their length and degree of branching. While amylose comprises largely of unbranched α-1, 4-linked glucans chains, amylopectin is a branched polymer in which linear chains of α-1, 4-linked glucoses are joined together by α-1, 6-linkages (Fig. 1.1). The branch points are arranged so that clusters of chains of about 12-20 glucose units occur at regular intervals of about 9 nm along the axis of the molecule. Differential packing of the amylopectin double helices gives rise to either the A-type or the B-type starch, the two polymorphs differing in the geometry of their single cell units, packing density of their double helices and the level of bound water within the crystal structure. While the A-type polymorph is relatively compact with the double helices packed in a monoclinic unit cell containing 12 glucose residues and 4 water molecules, the B-type polymorph has a more open structure with the double helices packed in a hexagonal unit cell containing 12 residues and 36 water molecules (Imberty *et al.*, 1991). C-type starches are considered to be a mixture of A and B-type polymorphs. This poly-modal distribution of glucan chain lengths, coupled by regular branch point clustering, forms the basis of the semi-crystalline nature of the matrix of the starch granule. Clustering of alternating regions of semi-crystalline and amorphous material give rise to the growth rings that are present in higher plant storage starches (Hall and Sayre, 1973; Pilling and Smith, 2003) (Fig. 2.2). The crystalline structure of starch granules is highly conserved in plants (Jenkins *et al.*, 1993). Thus, while the A-type starch is characteristic of cereals, C-type structures are characteristics of starch granules from legumes (Hizukuri, 1986; Imberty *et al.*, 1991; Cairns *et al.*, 1996; Wang *et al.*, 1998). B-type structures are

Fig. 2.2 Schematic representation of the structure of a starch granule, with alternating amorphous and semi-crystalline regions constituting the growth ring.

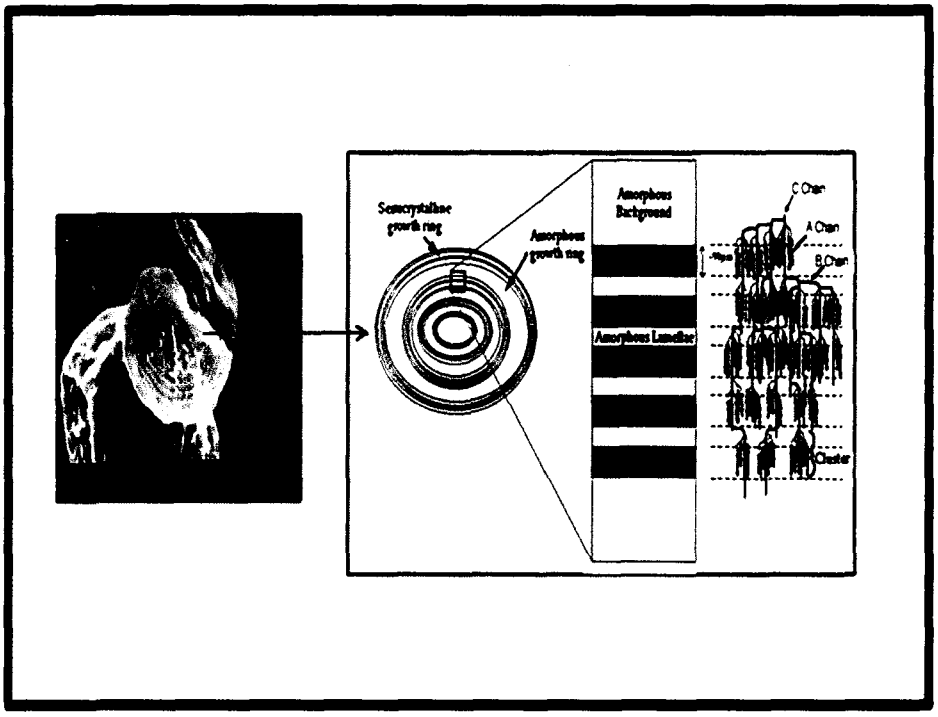

Amorphous Background
C Chain
A Chain
B Chain
Amorphous Lamellae

Fig. 2.2

characteristics for starches from roots and tuber tissues (Hizukuri, 1986; Imberty *et al.*, 1991).

Amylose and amylopectin represent approximately 98 - 99% of the dry weight of starch granules. However, the relative proportion of amylose to amylopectin may vary from species to species, within species, between tissues from the same species (Jane *et al.*, 1992; Shujun *et al.*, 2006; Peroni *et al.*, 2006). Amylopectin is the major component of most starches and its content in starch varies from 75 - 85% for normal starches (Manners, 1979) to even more than 99% for *waxy* starches, which are essentially amylose free (Shure *et al.*, 1983; Sano, 1984; Song and Jane, 2000; Singh *et al.*, 2005; Zhu *et al.*, 2008). On the other hand, amylose constitutes 15% - 25% of the total mass of starch. However, high amylose starches having 40% - 80% amylose content, have been reported in maize (Manners, 1979; Jane *et al.*, 1999; Zhu *et al.*, 2008). Similarly, starches with 35% - 45% amylose content have been reported in barley and rice (Fredriksson *et al.*, 1998; Song and Jane, 2000). Regardless of the source, starches are defined as *waxy* when the amylose content is <15%, normal when the amylose content ranges between 15- 35% and high amylose (amylo-) when the amylose content exceed 35% of the total biomass of the starch grains The general properties and functionalities of amylose and amylopectin are presented in Table 2.2. In addition to amylose and amylopectin, starch granules also contain minor non-carbohydrate components including lipids (0.01- 0.80%), proteins (0.10 -0.40%) and upto 0.5% minerals and salts. Lipids and proteins, both of which affect the functionality of starch, are the main minor

Table 2.2: Physicochemical characteristics of amylose and amylopectin.

Characteristics	**Amylose**	**Amylopectin**
Molecular structure	Linear (α-1,4)	Branched (α-1,4;α-1,6)
Molecular weight	~10^6 Daltons	~10^8 Daltons
Degree of Polymerization	1500-6000	3 *10^5-3*10^6
Helical complex	Strong	Weak
Iodine color	Blue	Red-purple
Dilute solutions	Unstable	Stable
Retrogradation	Rapidly	Slowly
Gel property	Stiff, irreversible	Soft, reversible
Film Property	Strong	Weak and brittle

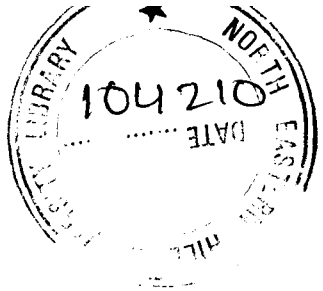
components of starch and are present either on the surface or are located within the matrix of starch granules (Appelqvist and Debet, 1997).

The performance of starch in any application is governed by its physical and chemical properties, which convey specific functionality to the molecule (Fegurson, 1994). According to FAO (1998), starch can be viewed as a set of functional properties suited to a particular application. The functional properties of starches vary considerably among starches from different sources and are therefore unique for each type. These differences are believed to arise from differences in amylose/amylopectin ratio, the molecular weight and mean starch grain size (Tian *et al.*, 1991; Jane and Chen, 1992; Shibanuma *et al.*, 1996; Freddriksso *et al.*, 1998; Sasaki and Matsuki, 1998; Lu *et al.*, 2005). Thus, while low amylose flour is good for preparation of oriental noodles as well as for extended shelf life of bread, flour with high amylose content is good for fried foods as high amylose starch has good filming properties (Graybosch, 1998; Bligh, 1999). Wang *et al.* (2008), have demonstrated that *waxy* starch from sorghum has higher ethanol conversion rates than the wild type. This makes *waxy* loci useful for development of high bioenergy grain sorghum. In addition, *waxy* grain has been shown to have increased starch and protein digestibility compared to wild-type (Rooney and Pflugfelder, 1986; Wong *et al.*, 2009).

The physicochemical properties of starches *viz.*, gelatinization temperature, retrogradation behavior, swelling power and solubility, viscosity, pasting properties, freeze/thaw stability, acid stability and gel strength have been shown to have an important bearing on the functional properties of starch (Sefa-Dedeh and Sackey, 2002;

Chen *et al.*, 2003; Amani *et al.*, 2004; Perez *et al.*, 2005; Peroni *et al.*, 2006; Riley *et al.*, 2006; Shujun *et al.*, 2006). Therefore, characterization of starches for their physicochemical, functional and structural properties is essential in order to unravel their potential for use in the food and non-food industries. Since the physicochemical properties are genetically determined, it is possible to manipulate the properties of starches by genetic manipulation. Therefore, discovery and characterization of the enzymes that affect the quality of starch are of notable worth. The biosynthesis of starch involves the coordinated action of three enzymes *viz.*, ADPglucose pyrophosphorylase (ADPase), starch synthases (SS) and starch branching enzyme (SBE). Kossmann and Lloyd (2000), have identified at least five isoforms of starch synthases which include one granule bound enzyme *viz.*, GBSS-I and four soluble starch synthases *viz.*, SS-I, SS-II, SS-III and SS- IV.

Many starch biosynthetic enzymes have been shown to be distributed between the soluble fraction of the plastids and the insoluble starch granules (Ball and Morell, 2003; Tetlow *et al.*, 2004a). GBSS which is localized exclusively within the starch granule is responsible for amylose biosynthesis (Shure *et al.*, 1983) and also has a role in the elongation of long chains in amylopectin (Delrue *et al.*, 1992; Denyer *et al.*, 1996; Maddelein *et al.*, 1994; Van de Wal *et al.*, 1998). However, the precise molecular mechanism by which this enzyme synthesizes amylose is still under debate (Wattebled *et al.*, 2002). Although this increases the complexity of the process and makes it difficult to interpret the role of different enzymes, it also increases the range of possibilities for



bioengineering. By targeting specific forms of an enzyme it is possible to exercise control over the physicochemical properties of the starch granules.

The first step in the starch biosynthetic pathway is the conversion of glucose-1-phosphate to the specific starch precursor ADP-Glucose by the enzyme ADPase (Preiss *et al.*, 1994; Keeling *et al.*, 1998). The ADP-Glucose units are then polymerized into the α-1, 4-glucosyl chains of starch by soluble starch synthases which synthesizes amylopectin and granule bound starch synthase that synthesizes amylose (Nakamura *et al.*, 1995). The linear amylopectin chain is then modified by other enzymes including the branching enzymes that transfer short α-1, 4-glucosyl chains of amylopectin to α-1, 6 configuration creating the branched amylopectin molecule (Boyer and Preiss, 1978; Guan and Preiss, 1993; Rahman *et al.*, 1995). The amylopectin molecules grow in alternating layers of tightly packed branch points and long linear sections of the polymer (Thompson, 2000). The amylopectin molecules have been reported to grow outward from a central point which has been predicted to be the core starch starter molecule amylogenin, like the precursor molecule of glycogen, glycogenin (Keeling *et al.*, 1994). Investigations carried out on starch composition mutants, including the *waxy* mutants which produce starch comprising solely of amylopectin, have implicated granule bound starch synthase I (GBSS-I) as the major enzyme responsible for amylose synthesis in starch grains in the endosperm (Smith *et al.*, 1997). On the other hand, the synthesis of amylopectin involves concerted activity of the starch synthases as well as branching and de-branching enzymes (Denyer *et al.*, 2001; Hirose and Terao, 2004). GBSS-I, the key enzyme responsible for amylose synthesis, has been identified as a 58 kDa (Shure *et al.*,

1983) or a 60 kDa (Echt and Schwartz, 1981) protein in maize, a 60 kDa protein in rice (Sano, 1984) and a 68 kDa protein in grain amaranth (Konishi *et al.*, 1985). Based on their molecular mass, two classes of GBSS have been reported in plants. While GBSS-I has a molecular mass ranging between 58-60 kDa (Echt and Schwartz, 1981; Shure *et al.*, 1983; Sano, 1984), GBSS-II has a molecular mass ranging between 77 to 79 kDa (Denyer *et al.*, 1997; Nakamura *et al.*, 1998; Vrinten and Nakumara, 2000; Edwards *et al.*, 2002).

In contrast to GBSS-I, the role of GBSS-II remains ambiguous. A reduction in the activity of this enzyme in potato tubers did not affect the ratio of amylose to amylopectin in the starch grains (Edwards *et al.*, 2002). However, studies on potato and *rugosus*5 (*rug*5) mutant of pea have indicated that the enzyme might have an important role in determining the structure of amylopectin and starch granule morphology (Craig *et al.*, 1998; Fulton *et al.*, 2002). While the amylose molecules are known to form single helices, the linear sections of amylopectin form double helices with adjacent amylopectin molecules (Davis, 1994; Ball *et al.*, 1996). While GBSS is known to be highly active enzyme, amylose production in starch grains has been suggested to be limited by the space available within the amylopectin matrix (Flipse *et al.*, 1996; Tatge *et al.*, 1999). While a reduction in the amount of active GBSS in the starch matrix lead to a corresponding reduction in the amount of amylose produced, the gene dose effects were not linear (Yamamori *et al.*, 1992; Yamamori and Quynh, 2000). It was suggested that since the GBSS gene expression is not regulated by feedback mechanism, the only

way to reduce the amount of amylose in the starch grains was by creating null alleles that do not express an active enzyme.

On the basis of their degree of association with starch grains, granule-associated proteins have been divided into two categories. While one group comprises of the surface-associated proteins, the other group comprises proteins localized within the granular matrix of the starch grains (Grimaud *et al.*, 2008). The surface-associated proteins have a small degree of association with the starch granule and can be separated from the granule by dissociation with a detergent or a reducing agent (Denyer *et al.*, 1993; Rahman *et al.*, 1995; Boren *et al.*, 2004). By contrast, extraction of internal granule-associated proteins requires the gelatinization of the starch granule in SDS in the presence of a reducing agent (Tsai, 1974; Echt and Schwartz, 1981; Denyer *et al.*, 1993). Studies in many plant species have shown that starch biosynthetic enzymes exist as both soluble proteins in the stroma and as internal granule-associated proteins. Enzymes that display this property in maize endosperm include SSI (Mu *et al.*, 1994, 2001), SSIIa (Zhang *et al.*, 2004) and BEIIb (Mu-Forster *et al.*, 1996).

Given the important role starch plays in food and non–food uses, several efforts are being made for the manipulation of its composition through modification of amylose/amylopectin ratio. Since Amylose-free or low-amylose starch is considered desirable for certain food and non-food industries (Morell *et al.*, 1995), most of the approaches aimed at modifying the quality of starch have focused on alteration of amylose/amylopectin ratio of the grains. Approaches used to achieve this goal are being pursued through manipulation of the genes coding for enzymes involved in the starch

biosynthetic pathway (Sestili *et al.*, 2010). The gene coding for the enzyme GBSS-I (called *waxy*) has been cloned from maize (Shure *et al.*, 1983; Klösgen *et al.*, 1986), potato (Hovenkamp-Hermelink *et al.*, 1987; Visser *et al.*, 1989; Van der Leij *et al.*, 1991), barley (Rohde *et al.*, 1988), rice (Wang *et al.*, 1990; Hirano and Sano, 1991; Okagaki, 1992), pea (Dry *et al.*, 1992), sorghum (Hsieh *et al.*, 1996a), cassava (Salehuzzaman *et al.*, 1993) and common wheat (Clark *et al.*, 1991; Mason-Gamer *et al.*, 1998; Murai *et al.*, 1999). Ample evidence exists to indicate that the *waxy* locus is a single copy gene (Klosgen *et al.*, 1986; Visser *et al.*, 1989; Wang *et al.*, 1990; Van der Leij *et al.*, 1991; Salehuzzaman *et al.*, 1993).

The term GBSS and *waxy* are used interchangeably since GBSS loci, genes and enzyme are all termed '*waxy*'. There are several known alleles of the GBSS gene which can differ by a single base pair, or many base pairs located at different polymorphic loci within the gene. GBSS alleles that express an active GBSS enzyme are termed as 'wild-type' and that does not express an active GBSS enzyme are termed as 'null' (Bradley, 2003). Based on the amount of GBSS protein and the proportion of amylose in the starch grain three major GBSS alleles (Wx^a, Wx^b and *wx*) located on chromosome 6 have been identified in rice (Sano *et al.*, 1986; Smith *et al.*, 1997). The Wx^a allele is primarily found in the *indica* subspecies, while Wx^b allele is found in *japonica* subspecies of *O. sativa* (Hirano and Sano, 1991). However, the genotypes with the recessive *wx* allele have opaque seeds that contain less amylose and are often referred to as glutinous rice varieties. According to Bligh *et al.* (1995), Ayers *et al.* (1997), Bergman *et al.* (2001) and Tan and Zhang (2001), different GBSS alleles can be distinguished on the basis of

the length of CT repeat present in the 5' untranslated region of the gene, thereby distinguishing low, intermediate and high amylose varieties. Thus, three homologous *Wx* genes (*Wx-A1a*, *Wx-B1a*, and *Wx-D1a*) located on chromosomes 7AS, 4AL, and 7DS have been shown to encode the *waxy* protein in hexaploid wheat (Nakamura *et al.*, 1993; Fujita *et al.*, 1996). On the other hand, four other wild-type alleles *viz.*, *Wx-B1c*, *Wx-B1d*, *Wx-B1e* and *Wx-B1f* including the most common *Wx-B1a* have been specifically identified on the 4AL locus (Yamamori *et al.*, 1995; Rodriguez-Quijano *et al.*, 1998; Yamamori and Quynh, 2000; Nieto-Taladriz *et al.*, 2000). Nieto-Taladriz *et al.* (2000), have reported that presence of the alleles *Wx-B1c* and *Wx-B1f* resulted higher amylose production and thereby starch with high levels of amylose. Pedersen *et al.* (2005), have characterized two *waxy* alleles, based on the absence (*wxa*) or presence (*wxb*) of the GBSS protein in the endosperm, of sorghum (*Sorghum bicolor* L. Moench).

The molecular structure of the *waxy* gene has been well studied by plant breeders and molecular biologists. Sequence information of the gene has been reported from maize (Klosgen *et al.*, 1986), rice (Wang *et al.*, 1990; Okagaki, 1992), barley (Rohde *et al.*, 1988), wheat (Murai *et al.*, 1999), potato (Van der Leij *et al.*, 1991), sweet potato (Kimura *et al.*, 2001), foxtail millet (Fukunaga *et al.*, 2002), rye (Xu *et al.*, 2009) and amaranth (Park *et al.*, 2009). The size of the *waxy* gene ranges between ~2.7 kb in foxtail millets, wheat and rye to ~5.3 kb in barley and rice. The *waxy* gene is comprised of 13 exons and 12 introns in rice (Olsen and Purugganan, 2002) and maize (Mason-Gamer *et al.*, 1998), 14 exons and 13 introns in potato (Van der Leij *et al.*, 1991), 11 exons and 10 introns in rye and 12 exons and 10 introns in barley (Rohde *et al.*, 1988).

Waxy mutants affecting either gene expression or function of GBSS have been isolated in several plants including pea (*Pisum sativum*), wheat (*Triticum aestivum*), rice (*Oryza sativa*), barley (*Hordeum vulgare*) and corn (*Zea mays*) (Denyer *et al.*, 1997; Nakamura *et al.*, 1995; Isshiki *et al.*, 1998; Larkin and Park, 1999; Wang *et al.*, 1995; Hylton *et al.*, 1996; Taira *et al.*, 1995; Shure *et al.*, 1983). Domon *et al.* 2002, have shown that the barley *waxy* allele has a 403 bp deletion mutation in the 5' terminal part of the GBSS-I structural gene. The deletion was found to include a TATA box and a deduced transcription starting point of the wild-type allele, thereby suggesting the influence of the mutation on the amylose content of starch grains in the endosperm. In rice, a G→T mutation at the splice site of intron 1 has been demonstrated to result in incomplete processing of pre-mRNA and consequently low levels of transcripts of *waxy* gene (Hirano and Sano, 1991; Wang *et al.*, 1995; Hirano *et al.*, 1998; Isshiki *et al.*, 1998). While in all the diploid crops the *waxy* phenotype has arisen by spontaneous mutations, *waxy* lines in hexaploid bread wheat have been produced by crossing partially-*waxy* cultivars with cultivars having spontaneous mutations in the three GBSS-I homeologues in the A, B and D genomes (Nakamura *et al.*, 1995). Characterization of the resulting *wx-A1b*, *wx-B1b* and *wx-D1b* alleles showed that the *wx-A1b* contained a 23 bp deletion with a 4 bp insertion of filler DNA at the end of exon1and beginning of intron 1 thereby affecting transcript splicing. On the other hand, the allele *wx-B1b* had deletion of the entire GBSS-I transcription unit and the *wx-D1b* allele had a 588 bp deletion at the 3' end with a novel 12 bp insertion at the same site which resulted in the formation of a

premature stop codon. Such a modification resulted in the resultant protein lacking 30 amino acids on the C-terminus. (Vrinten *et al.*, 1999).

The phylogenetic utility of the structural nuclear gene for GBSS-I was first explored by Mason-Gamer and Kellogg (1996) in the family Poaceae. Their studies on GBSS-I in the family Poaceae revealed that while phylogenetic relationships based on variations/similarities in the nucleotide sequences of introns could be used to differentiate even closely related species. Mason-Gamer *et al.* (1998), identified GBSS-I as a candidate for phylogenetic reconstruction because of evidence that it was a single-copy gene. Nucleotide sequence of GBSS-I has also been used in phylogenetic studies on *Liquidambar* (Li and Donoghue, 1999), *Ipomoea* (Miller *et al.*, 1999), *Triosteum* (Gould and Donoghue, 2000) and *Solanum lycopersicum* (Peralta and Spooner, 2001; Walsh and Hoot, 2001; Levin and Miller, 2005; Levin *et al.*, 2005, 2006; Smith and Baum, 2006). In contrast, Evans *et al.* (2000), documented the existence of two paralogous GBSS-I loci in Rosaceae and Rhamnaceae. Analyzed in combination, these loci proved useful for resolving higher-level relationships within Rosaceae and in particular, for understanding the origin of Maloideae (Evans *et al.*, 2000; Evans and Campbell, 2002). Based on analyses of one GBSS-I paralogue, Smedmark *et al.* (2003), suggested that several members of Rosaceae originated through allopolyploidy. Nucleotide sequence information of GBSS-I gene has been used to provide evidence for the origin of the allopolyploids in North American *Elymus* and African *Eragrostis* respectively (Mason-Gamer, 2001; Ingram and Doyle, 2003). Since low copy number genes are known to be subjected to different evolutionary processes than plastid genes or

highly repetitive nuclear markers, they provide a valuable source of independent phylogenetic evidence. Mason-Gamer and Kellogg (1998), have suggested that GBSS-I introns could provide more phylogenetic informative characters and higher levels of variation than the ITS region.

The wild type GBSS-I enzymes expressed from different loci are known to differ from each other in their structure and activity (Nakamura *et al.*, 1993; Yamamori *et al.*, 1994). Yamamori and Quynh (2000), have described the differences in the activities of GBSS-I transcribed from three different loci in wheat. It was noted that while starch from the wild type allele at 7A locus had an average of 19.8% amylose, the wild type allele at 7D locus had 22.5% amylose and 4A locus had 24.2% amylose. The higher activity in wild type allele at 4A locus was ascribed to the allele possibly expressing more copies of GBSS-I than the genes at other GBSS loci. Loss of GBSS function in plants causes the *waxy* endosperm phenotype, named for its altered texture and endosperm appearance (Denyer *et al.*, 2001). *Waxy* mutants contain very low or undetectable levels of amylose with their starches comprised nearly entirely of amylopectin (Denyer *et al.*, 2001). *Waxy* mutants affecting either gene expression or function of GBSS have been isolated in several plants including pea (Denyer *et al.*, 1996), wheat (Nakamura *et al.*, 1995), rice (Wang *et al.*, 1995; Isshiki *et al.*, 1998; Larkin and Park, 1999), barley (Taira *et al.*, 1995; Hylton *et al.*, 1996) and corn (Shure *et al.*, 1983). Because the *waxy* mutation alters starch composition by greatly reducing amylose content, the physiochemical properties of *waxy* starch are altered relative to normal starch (Pedersen *et al.*, 2007; Sang *et al.*, 2008). The isolation and analysis of

natural and insertion mutants has made an invaluable contribution towards our understanding of starch metabolism in higher plants.

Besides its nutritive values, starches are very useful raw material with a wide range of applications in both food and non-food industries. While much work has been done on the characterization of starch from cereals and analysis of the regulation of its biosynthesis, not much information is available about the quality of starch in many other potentially important crops. One such group of crops belongs to pseudo cereals. Amongst this group of plants, common buckwheat is an important pseudo cereal because of its high potential for use as a functional food. It is a highly nutritious crop mainly used for human consumption and holds expanded opportunity for maximum utilization (Kreft *et al.*, 2002). Skrabanja *et al.* (2001) have shown that starch from buckwheat grains have a strong potential for use in the development of functional foods including its engineering for a modified course of starch digestion. Buckwheat is also a rich source of high biological value protein, antioxidants and essential minerals. The specific ability of buckwheat proteins to soak up cholesterol from food and the relative proportions of its amino acids, make buckwheat the best cholesterol-lowering food substitute (Kreft *et al.*, 2002; Chrungoo *et al.*, 2005).

Common buckwheat (*Fagopyrum esculentum* Moench), which is considered as underutilized crop due to less investment in exploring its potential, is of significant economic importance in several countries of the world. The Indian Himalayas, Asia Pacific Region, Central and Eastern Europe are considered as important buckwheat producing regions of the world. Buckwheat flour is used extensively by the people living

in foothills of Indian Himalayas as also in other countries of the world including Japan, Korea, Poland, China, Russia, and Slovenia for preparation of noodles, pasta, bread, pancakes. The crop is also a rich source of proteins, carbohydrates, and glucosides including rutin. Even though there has been little progress in buckwheat breeding programs, the National Bureau of Plant Genetic Resources (NBPGR), New Delhi has identified buckwheat as an important underutilized crop which has the potential to become an important component of human diet. The Bureau has established a National Buckwheat Germ Plasm Bank and is maintaining more than 3000 accessions of buckwheat in its repository. However, because of its unfavorable amylose/amylopectin ratio, the buckwheat flour had poor dough making qualities. Since the ratio of amylose to amylopectin influences the texture and quality of flour, identification of the gene loci involved in the regulation of starch synthesis and deposition, would have an important bearing on devising effective breeding programs for improved starch quality. Even though such proteins have been isolated and their genes cloned from other conventional crops, none of the granule bound proteins in buckwheat have been identified nor have their genes been cloned.

CHAPTER III

MATERIALS AND METHODS

1. MATERIALS

1.1 Plant Material

Indian accessions/cultivars of buckwheat (*Fagopyrum* spp.) used in the present study were procured from the buckwheat germplasm of National Bureau of Plant Genetic Resources, Shimla, North Eastern Regional Station of NBPGR, Shillong and Vivekananda Laboratory for Hill Agriculture, Almora (India). The European accessions/cultivars of buckwheat were procured from the buckwheat germplasm collection at the University of Ljubljana, Slovenia (Europe).

1.2 Membranes and Filters

Positively charged PVDF membrane (0.45 μm pore size) from Fluka was used for Western blotting. Dialysis membrane-159 procured from Hi-media Laboratory was used for dialysis of the protein sample.

1 A close up view of common buckwheat *(Fagopyrum esculentum* Moench.) plant at the flowering stage.

2 A close up view of dehulled seeds of common buckwheat *(Fagopyrum esculentum* Moench.).

Fig. 3.1

Fig. 3.2

2. PROTOCOLS

2.1 Preparation of starch from Buckwheat

Starch granules were isolated from endosperm tissues of mature grains of common buckwheat as well as from the leaf tissues according to Takaoka *et al.* (1997). Healthy grains/fresh leaves from the collections of buckwheat procured from North Eastern Regional Station of NBPGR, Shillong and the buckwheat germplasm collection at the University of Ljubljana, Slovenia were washed with water followed by rinsing with distilled water. The hull portion of each of the grains was removed and the groat was made to a fine powder under liquid nitrogen in a prechilled pestle and mortar. The powder was subsequently suspended in distilled water and the suspension maintained at 4°C for 4 hours with occasional stirring. The suspension was filtered through a 75 μM mesh sieve and the filtrate centrifuged at 10000Xg at 4°C for 10 mins. The pellet was resuspended in 400 μL of SDS sample buffer containing 60 mM Tris-HCl, pH 6.8, 10% (w/v) sodium dodecyl sulphate (SDS), 3% β-mercaptoethanol and 10% (w/v) glycerol and centrifuged at 30, 000Xg for 10 mins. at 4°C. The supernatant was transferred to an eppendorf tube and the pellet suspended again in 400 μL of SDS sample buffer containing 60 mM Tris-HCl, pH 6.8, 10% (w/v) sodium dodecyl sulphate (SDS), 3% (v/v) β-mercaptoethanol and 10% (w/v) glycerol. The suspension was centrifuged at 30,000Xg for 10 mins. at 4°C. The supernatant was transferred to an eppendorf tube and pellet suspended again in 400 μL of SDS sample buffer containing 60 mM Tris-HCl, pH 6.8, 10% (w/v) sodium dodecyl sulphate (SDS), 3% β-mercaptoethanol and 10% (w/v) glycerol. The suspension was centrifuged again at 30,000Xg for 10 mins. at 4°C. The

pellet was finally washed thrice with deionized water followed by three washes with cold acetone. The acetone washed pellet, comprising of starch grains, was dried at room temperature and stored at -20°C for later use. For isolation of starch grains from the leaves, freshly harvested leaves were washed with distilled water, chopped into pieces and ground to a fine powder under liquid nitrogen in a prechilled pestle and mortar. Starch was isolated from the fine powder following the procedure of Takaoka *et al.* (1997) as described above.

2.2 Bright field microscopy of starch grains

Starch grains were visualized under a light microscope (Axio Image A1, Carl Zeiss, Germany) under bright field (100X magnification) after staining with LUGOL (5% (w/v) iodine, 10%(w/v) potassium iodide) :H_2O [1:2], according to Autio *et al.* (1992). An aliquot of the solution containing the stained grains was poured on a glass slide, covered with a cover slip and viewed.

2.3 Scanning Electron Microscopy of isolated Starch Granules

The sizes and the surface structure of the starch granules isolated from different accessions/cultivars of buckwheat by the above procedure were analyzed by using a scanning electron microscope (JOEL 1850 and JSM 6360, Tokyo, Japan), according to the method described by David *et al.* (2006). Dry starch was sprinkled on 'brass stubs' and placed on the specimen stand or the 'ion sputter' and coated with a thin layer of gold vapour (300A° layer). The starch granules were then observed in a JOEL-1850 SEM at magnification ranging between 1000X to 6000X operated at 15 kV.

To visualize the internal structure of the starch granules, starch grains were washed with acetone, air dried and ground in a prechilled pestle and mortar under liquid nitrogen to crack the granules after which they were suspended in 100 mM MES buffer (pH 6.0) containing 5 units of α-amylase and incubated for 30 mins. at 30°C to preferentially digest amorphous region of the starch granules (Pilling, 2001). The solution was centrifuged at 6000Xg for 5 mins. to pellet the partially digested starch grains. The pelletted grains were washed thrice with cold acetone, air dried and then viewed under the scanning electron microscope.

2.4 Determination of amylose content

Amylose content of the starch grains was measured according to the method described by Yamamori *et al.* (1992). A suitable mass (approx. 20 mg) of the purified starch grains was suspended in 5 ml of 0.75 N NaOH in 25% (v/v) ethanol. The suspension was incubated at room temperature for 12 hours. The sample was made up to 50 mL with distilled water, and amylose content was determined with an analyzer (Tecnicon, Bran-Lubbe Co., Tokyo, Japan) using KI-I_2 solution (0.005% [w/v] KI and 0.003% [w/v] I_2). An aliquot of starch suspension was pipette out and mixed with 0.5 ml of iodine solution. The reaction mixture was allowed to stand for 20 mins. in dark at room temperature. Absorbance of the coloured complex was recorded at 620 nm wavelength in Lambda 35 UV/VIS Spectrophotometer (Perkin Elmer, U.S.A) against a reagent blank. The λ_{max} of the iodine-starch complexes was determined according to the method of Konishi *et al.* (1985). The concentration of amylose in the suspension was calculated with reference to potato amylose (Sigma) as the standard.

The regression equation of the standard curve for determination of amylose % is given as:

$$y = a + bx$$

Where,

a = zero intercept = $\frac{\sum y}{n}$

b = slope = $\frac{\sum(y-\bar{y})(x-\bar{x})}{\sum(y-\bar{y})^2}$

n = number of points
x = free iodine
y = bound iodine
$\bar{x}$= average free iodine
$\bar{y}$ = average bound iodine

2.5 Determination of Rheological properties

Pasting properties of starches suspended in distilled water were determined by a Rapid Visco- Analyser (C-DG26.7/SS/QC, RheolabQC, Anton Paar. GmbH, Germany) following the method of Jayakody *et al.* (2007). A 7% (w/v) suspension of native buckwheat starch was prepared by dispersing dried starch grains in MilliQ deionized water with constant stirring. The slurry was heated to 40°C and held at this temperature for 1 min. after which it was heated to 90°C for 5 mins. The slurry was held at this temperature for 5 mins. followed by rapid cooling to 40°C for 5 mins. The slurry was held at this temperature for 1 min. The heating and cooling regimes were programmed as per the above mentioned cycle with constant rotational speed of 160 rpm. The pasting properties including peak viscosity, minimum viscosity, breakdown viscosity, final viscosity and pasting Temperature (°C) of each sample were inferred from acquired data.

2.6 Extraction of Starch Granule-Bound Proteins

Starch granule bound proteins were extracted from the isolated starch grains according to the method of Takaoka *et al.* (1997). A suitable mass of air dried starch was suspended in 10 volumes of the SDS sample buffer, and incubated in a water bath at 95°C for 10 mins. The suspension was gelatinized by rapid cooling to -80°C and maintained at this temperature for 1 hour. The frozen suspension was subsequently incubated for 10 mins. at 50°C in a water bath to allow the grains to thaw. After thawing at 50°C, the gelatinized starch was centrifuged at 30000Xg at 4°C for 10 mins. and the supernatant transferred to fresh tube. Proteins present in the supernatant were recovered by addition of 2 volumes of cold acetone followed by centrifugation at 30,000Xg at 4°C for 10 mins. The pellet containing starch granule-bound protein(s) was dried at room temperature and resuspended in appropriate volume of 0.1M Tris-Cl buffer pH for further use.

2.7 Gel electrophoresis

SDS-PAGE of the granule bound protein was carried out on 12% polyacrylamide gel following the method of Laemmli (1970). Suitable aliquots of the extracts, representing 50µg protein per sample, were mixed separately with 2X Laemmli buffers in the ratio 1:1. The mixture was heated for 5 mins. in a boiling water bath followed by a brief centrifugation to sediment the debris. The supernatant was loaded into the well of a 1.5mm thick 12% acrylamide gel. An aliquot of the mixture of standard molecular weight markers (Bangalore Genei, India) was also denatured similarly and loaded on the same gel to serve as reference for determination of

molecular mass of the resolved bands. Electrophoresis was carried out at a constant voltage of 100V for 5 hours. After electrophoresis the gel was removed and rinsed with distilled water. Proteins in the gel were fixed by immersing the gel for 30 mins. in methanol: glacial acetic acid: water (4:1:5). The gel was subsequently stained for 3 hours in 0.25% (w/v) Coomassie Brilliant Blue R-250v (CBB R-250) prepared in methanol: glacial acetic acid: water (4:1:5). Destaining was carried out in a destaining solution I composed of methanol: glacial acetic acid: water (4:1:5).The gels were further destained in destaining solution II composed of methanol: glacial acetic acid: water (4:5.3:0.7) till the background was cleared. Protein bands were visualized on KODAK Gel Logic 200 gel documentation system under white light. R_f of the bands was calculated as the ratio of the distance travelled by the band to the distance travelled by the dye front. For deriving a relationship between the molecular mass of the protein and its mobility on the polyacrylamide gel, the R_f value of the standard molecular weight markers was plotted against log molecular weight of the protein to obtain a standard curve. The molecular mass of the resolved proteins was determined by comparing their electrophoretic mobility of the resolved proteins with that of the standard molecular weight markers and by reference to the standard curve.

2. 8 Preparation of antiserum/polyclonal antibody

The protein suspension obtained after the extensive extraction procedure was transferred to a dialysis membrane and subjected to dialysis for 72 hours at 4°C against water with constant stirring and repeated changes of the water. The clear dialysate was collected in a pre-chilled glass vial. The purified 59.7 kDa protein was then used to raise

antibodies against the isolated protein according to the procedures described by Sharan *et al.* (2005). Immunization was carried out by injecting a young healthy rabbit (2.5-3.0 kg; 10-16 weeks of age) subcutaneously with the purified 59.7 kDa protein (30μg) redissolved in PBS buffer containing Freund's complete adjuvant (Sigma Aldrich, U.S.A) in the ratio of 1:1. Booster injections of 75-100μg of starch granule bound proteins dissolved in PBS buffer containing Freund's complete adjuvant in the ratio of 1:1 were administered every 2 weeks after the first doze. Antiserum was obtained by a number of consecutive bleedings between 60 and 100 days after the 2nd booster doze. The blood was centrifuged at 13,000Xg for 15 mins. at 4°C to pellet the coagulated blood and the resulting supernatant (antiserum containing IgG fraction of polyclonal antibody) mixed with 0.01% ammonium sulphate solution (Sharan *et al.*, 2005). All antiserum batches, referred to as anti-59.7 kDa antiserum, were pooled and stored at -30°C.

Specificity of the polyclonal antibody against GBSS-I antigen and its titer was determined according to Ouchterlony and Nilsson (1978). For this purpose, 1% solution of agarose was prepared in 0.9% NaCl containing 0.02% sodium azide and poured on a clean glass slide. The solution was allowed to polymerize for 15-20 mins. after which 3 wells, each 1cm apart, were punched on the gel. The antigen solution was loaded in the middle well, while the antibody solutions (anti -59.7 kDa antiserum) was loaded on one of the side wells. The other side well was loaded with equal volume of 0.01% ammonium sulphate. The glass slide was then placed inside a moist chamber (box containing wet cotton/wet filter paper) and incubated overnight at 37°C. During this time

the antigens in the sample extract and the antibodies diffuse out of their respective wells and cross reacts to form an immune complex if the antibodies recognize the antigens. The immune complex precipitates in the gel to give a thin white line/arc, which is visible sign of antigen recognition (Fig. 3.3).

2.9 Two Dimensional (2D) Gel Electrophoresis

2D-PAGE of the starch granule bound proteins, with IEF in the first direction and SDS-PAGE in the 2nd direction, was carried out according to the method of Berkelman and Stenstedt (2002), using the 2D gel electrophoresis system of GE Healthcare, AB, Uppsala, Sweden. Prior to IEF, the granule bound protein(s) isolated from starch grains of common buckwheat endosperm were solubilized in 350µl of rehydration stock solution composed of IPG buffer containing 8M urea, 2% (w/v) CHAPS(3[3- cholamidopropyl] dimethylaminonia-1- propanesulsphonate), 0.5% (v/v) Pharmalyte (pH 4- 6.5), 0.002% bromophenol blue and 0.3% DTT (DL-Dithiothreitol). The mixture was extensively vortexed and centrifuged to pellet insoluble material. A suitable aliquot of the supernatant containing 150 µg of protein, was loaded onto 13cm IPG strips with a non-linear *pI* gradient of 4-7 (GE Healthcare AB, Uppsala, Sweden), during the rehydration step which lasted for 14 hours at room-temperature. Following rehydration, IEF was carried out in Ettan IPGPhor 3 system (GE Healthcare AB, Uppsala, Sweden) at 500V for the first 45 mins. and then at 4500 V for next 5hr 45 mins. on a constant current of 15 mA/strip. Electrophoresis was carried out at a constant temperature of 20°C.

Fig. 3.3 An Ouchterlony immunodiffusion plate, after immune diffusion has taken place.

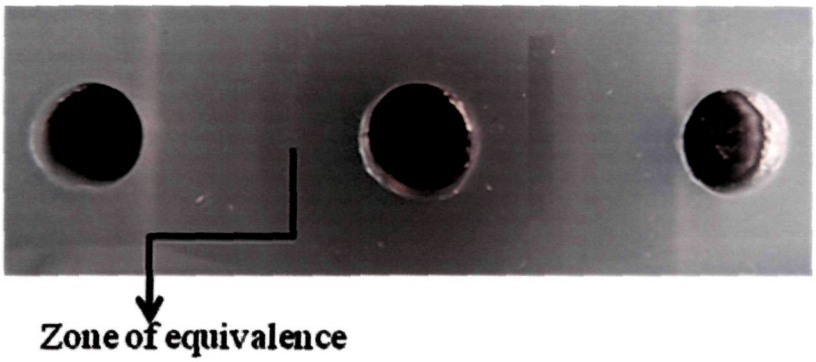


Fig. 3.3

Prior to the 2nd dimension run, the gel strips were equilibrated for 15 mins. in equilibration buffer (I) containing 0.05 M Tris-HCl buffer (pH 8.8), 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS and 1% (w/v) DTT. This was followed by a 2nd equilibration for 15 mins. in equilibration buffer (II) containing 0.05 M Tris-HCl buffer (pH 8.8), 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS and 2.5% iodoacetamide. The equilibrated strips were then applied to the polyacrylamide gel (12.0% resolving 5% stacking) and sealed with 0.5% agarose in SDS buffer containing bromophenol blue. Protein markers were loaded besides the strips before sealing. Electrophoresis was carried out for 30 mins. at 15 mA/gel and then at 20 mA/gel till the dye front reached the bottom of the gel. After electrophoresis, gels were visualized by staining with CBB R-250.

2.10 Western Blotting

The granule bound proteins extracted from starch grains of common buckwheat were subjected to SDS-PAGE under reducing conditions on a 12% polyacrylamide gel according to Laemmli (1970). Immediately after electrophoresis, the unstained gel was rinsed in Towbin transfer buffer for 15 mins. to saturate the gel with the buffer. Electrophoretic transfer of proteins from the gel to PVDF (0.46 μm) was carried out at 60 V for 1hr at room temperature in a semi-dry electro transfer blotting system (Amersham Biosciences) according to Towbin *et al.* (1979).

Following the transfer one lane from the PVDF membrane, having the transferred proteins, was cut and stained with India ink to check the quality and extent of transfer. For staining with India ink, the membrane was washed first in 1% Potassium hydroxide for 5 mins. followed by equilibration, under constant shaking over a rocking

platform, with PBS for 1 hour. The equilibrated membrane was stained overnight with 0.1% India ink prepared in TTBS. After destaining the membrane with PBS, the protein bands appeared as distinct black areas against a greyish white background. The other portion of the membrane carrying the transferred protein(s) was sent for N-terminal amino acid sequencing. The N-terminal amino acid sequencing of the mature GBSS-I protein was determined by sequential Edman degradation method.

2.11 Immunoblotting

For immunodetection of the granule bound proteins, the membrane carrying the transferred proteins, was incubated for 1 hour in TTBS blocking buffer comprised of 5% (w/v) defatted milk, 1X TBS and 0.1% Tween 20 followed by overnight incubation at 4°C with primary antibody (rabbit serum against the 59.7 kDa granule bound protein) isolated from common buckwheat at a dilution of 1:80,000. The excess of primary antibody was removed by four washes of 15 mins. each with the blocking buffer. The washed membrane was incubated for another 1 hour with secondary antibody (alkaline phosphatase-conjugated goat anti-rabbit serum, Bangalore Genei, India) at a dilution of 1:5,000. The excess secondary antibody was removed by three washes of 10 mins. each with the blocking buffer and three washes of 10 mins. each with the tris-sodium chloride buffer (50 mM Tris- HCl pH 7.5, 150 mM NaCl). Immunoreactive polypeptides were detected as blue bands by incubating the PVDF membrane with a substrate mix containing BCIP-NBT (5-bromo-4-chloro-3-indoylphosphate and nitroblue tetrazolium). The reaction was stopped with several rinses of double distilled water.

2.12 In-gel digestion

Tryptic digestion of selected protein spots were performed according to Shevchenko *et al.* (1996). The portion of the gel containing the target protein was chopped into pieces and destained at 40°C in 50 mM NH_4HCO_3 (ammonium bicarbonate) in 50% (v/v) CH_3CN (acetonitrile), the step being repeated until the gel was colorless. The gel pieces were dried and reduced by addition of 200 µl of 10 mM DTT in 100 mM NH_4HCO_3 for 1 hour at 56°C. Following reduction, the gel pieces were centrifuged to remove the supernatant. Alkalylation of the gel pieces were carried out by the addition of 200 µl of 55 mM iodoacetamide in 100 mM NH_4HCO_3 and incubated in the dark at room temperature for 45 mins. The gel pieces were washed with 50-100 µl aliquots of 100 mM NH_4HCO_3 for 10 mins. followed by dehydration with CH_3CN, reswelling in 100 mM NH_4HCO_3 and dehydrating again with CH_3CN. The gel pieces were then air dried and reswelled again at 4°C for 45 mins. with 150 µl of 50 mM NH_4HCO_3 containing trypsin (12.5 ng/ml of 50 mM NH_4HCO_3) and incubated overnight at 37°C. The peptides were collected by centrifugation, purified by dialysis and dried in speedvac till the desired volume.

The tryptic digests were analyzed by LC-MS (Water, USA) operating in the positive ion and reflector modes as following the method described by Sheoran *et al.* (2005, 2006). The spectra acquired in the 700–3000 m/z range were processed with Mascot Distiller 2.0 (www.matrixscience.com) and the resulting peak lists used to identify the corresponding proteins in NCBInr (non-redundant) and Swiss-Prot databases.

2.13 Confocal Laser Scanning Microscopy of isolated starch granules

For histochemical localization of the 59.7 kDa protein within the lamellar structure of starch grains, the partially digested grains of starch were incubated for one hour in 1X TTBS buffer containing 5% defatted milk and then centrifuged at 6000Xg for 10 mins. to pellet the grains. The pellet was washed twice with TBS to remove nonspecifically bound defatted milk and then incubated overnight at 4°C in 1X TBS containing antibodies raised against buckwheat GBSS-I at a dilution of 1:10000. The starch grains were subsequently washed thrice with 1X TTBS, each wash being for 5 mins. The washed grains were suspended in 1X TBS containing the fluorescent labelled secondary antibody [(Alexa Fluor 546 goat anti-rabbit IgG) Invitrogen, USA] and incubated for one hour at room temperature in the solution containing the secondary antibody. The starch grains were subsequently washed thrice with 1X TTBS and observed under the confocal microscope (Leica Microsystems CMS GMBH, Germany) for fluorescence.

2.14 Isolation of Genomic DNA

Total genomic DNA was isolated from 14 days old etiolated seedlings by a modified CTAB extraction protocol (Murray and Thompson, 1980). Healthy grains of common buckwheat were surface sterilized by immersing in 0.01% $HgCl_2$ for 5 mins. followed by repeated rinsing with sterile distilled water. The washed grains were germinated on sterile germination paper in a plant growth chamber maintained at 27°C±2°C and 75±10% relative humidity, under dark condition. The seedlings were harvested after 14 days of incubation, washed with sterile distilled water, wrapped in

aluminum foils, freeze dried in liquid nitrogen and stored at -80°C. The frozen tissue was crushed to fine powder in a sterile pestle and mortar under liquid nitrogen. 500 mg each of the powdered tissue was transferred to a 1.5 ml microcentrifuge tubes containing 500 µl of 2X CTAB buffer. The powder was mixed gently in the CTAB buffer to avoid formation of clumps and then incubated in a water bath at 65°C for 1 hour. The incubation was followed by addition of chloroform: isoamyl alcohol (24:1) in the ratio of 1:1, with gentle inversions to mix the cell lysate thoroughly with chloroform: isoamyl alcohol. The mixture was centrifuged for 5 mins. at 4°C in a Heraeus Biofuge (Fresco) table top refrigerated micro centrifuge to separate the aqueous phase, containing the nucleic acids, from the organic phase containing the proteins, carbohydrates and cell debris. The aqueous phase, containing the nucleic acids, was collected into a new tube and $1/10^{th}$ volume 3M sodium acetate (pH 5.2) and two volumes of 100% ethanol was added to it to precipitate the DNA. For complete precipitation, the mixture was allowed to stand at -20°C, overnight. The precipitated DNA was pelleted by centrifugation at 15,500Xg for 5 mins. at 4°C. The pellet was washed twice with 70% ethanol, vacuum dried and dissolved in nuclease-free ultra pure water. The isolated DNA was electrophoresed on 0.8% agarose gel at 80V for 1 hour in 1X TBE buffer (pH 7.5). After the electrophoresis, the gel was stained with ethidium bromide solution (0.5 µg ml^{-1}) for 15 mins. and destained in water till the background fluorescence disappeared. DNA was visualized as fluorescent bands under UV light on a KODAK Gel Logic 200 UV transilluminator. DNA was visualized as fluorescent bands under UV light on a UV transilluminator. DNA was quantified by visual observation of ethidium bromide stained

agarose gel on a transilluminator and comparison of the intensity of fluorescence of DNA bands with fluorescence of known amounts of λ DNA electrophoresed along with the isolated DNA sample (Sambrook *et al.*, 1989). Quantification of the DNA sample was also carried out spectrophotometrically. 10 µl of DNA solution was mixed with 990µl of ultrapure water and the absorbance of the solution recorded at 260 nm in Perkin Elmer Lambda 35 UV/VIS Spectrophotometer. Concentration of DNA in the solution was calculated using the equation:

A_{260} x dilution factor x 50 = µg/ml DNA.

2.15 Quality check of isolated DNA

For checking the quality of the isolated genomic DNA, 5µl of the DNA solution was electrophoresed on 0.8% agarose gel along with *Eco*RI- *Hind*III digested Lambda DNA. A single band showing molecular mass of >21kb on the agarose gel indicated a good quality preparation with no shearing of the isolated DNA. The purity of the DNA preparation was also checked by measurement of absorbance of the sample at 260 nm and 280 nm in Perkin Elmer Lambda 35 UV/VIS Spectrophotometer and working out the A_{260}/A_{280} ratio of the sample. Samples with A_{260}/A_{280} of >1.8 were used for further experiments.

A suitable aliquot of the isolated DNA was transferred to an eppendorf tube and 17 µl of sterile MilliQ pyrogen free water was added. 5 µl of appropriate buffer was added to the tubes following gentle mixing of the solutions by repeatedly inverting the tubes 3-5 times. The reaction was started by adding 50 units of the restriction enzyme to

the reaction mixture and incubating the tubes at 37°C for 4 hours in a circulatory water bath. A typical 50µl of reaction digestion mixture consisted of:

DNA (5 µg)	:	25.0µl
10X Restriction enzyme buffer	:	5.0 µl
Restriction enzyme	:	3.0 µl
Ultra pure water	:	17.0 µl
Total	:	**50.0 µl**

The reaction mixture was stopped by rapid heating of the reaction mixture at 65°C for 5 mins. in a dry heating bath. The Digested DNA was immediately size fractionated by electrophoresis on 0.8% agarose gel at 50V for 6 hours. DNA was visualized by exposure of ethidium bromide stained gel to UV light under UV transilluminator.

2.16 Polymerase Chain Reaction (PCR) Amplification

PCR amplification of the target DNA was carried out with buckwheat genomic DNA as the template and primers derived from the conserved regions of nucleotide sequences of granule bound starch synthase-like protein genes available in the databases of gene bank. The primers were designed using Primer 3 online tool (Rozen and Skaletsky, 2000). Amplification reactions were carried out in a Perkin Elmer thermal cycler with the reaction cycle comprising of one cycle of "hot start" (94°C, 5 mins.); 35 cycles of denaturation (94°C, 1 min.); annealing (63-68°C, 1 min.), and polymerization/ primer extension (72°C, 1 min.) and one cycle of chain elongation (72°C, 10 mins.). The PCR cycle is diagrammatically represented in Fig. 3.4. A typical 25µl reaction volume (in 0.2 ml reaction tube) contained the following components:

Fig 3.4 Schematic representation of a typical three step cycle parameter for PCR amplification.

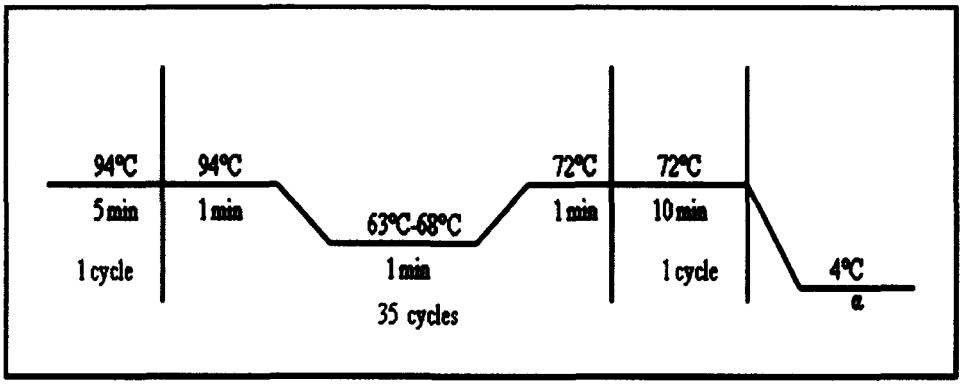

94°C
5 min
1 cycle
94°C
1 min
63°C-68°C
1 min
35 cycles
72°C
1 min
72°C
10 min
1 cycle
4°C
∞

Fig. 3.4

Sterile water	: 16.7 µl
10X reaction buffer	: 2.0 µl
$MgCl_2$ (25 mM Stock)	: 1.5 µl
10mM dNTP mix	: 0.5 µl
Primer 1 (5 mM stock)	: 1.0 µl
Primer 1 (5 mM stock)	: 1.0 µl
Genomic DNA template (~0.5 µg)	: 2.0 µl
Taq DNA Polymerase (3 unit/µl)	: 0.3 µl
Total	**: 25.0 µl**

The set up included positive (without primers) as well as negative (without template DNA) controls. After the amplification reaction, the entire reaction mixture was electrophoresed on 1.2% Agarose gel at 80 V for 3 hours. After the electrophoresis was complete, the gel was submerged for 15 mins. in ethidium bromide solution (0.5 µg ml^{-1}) followed by destaining in sterile distilled water. The DNA on the agarose gel was visualized under UV light on a transilluminator. Each amplification and subsequent agarose gel electrophoresis of the amplified DNA was repeated thrice to check reproducibility of the results. The region of agarose gel having the amplified DNA was excised from the gel with a sterile blade and the DNA was eluted from the agarose block with Genei Spin Gel Extraction Kit (Bangalore Genei, India) as per the manufacturer's protocol. The eluted DNA was lyophilized and stored at -80°C till it was used.

2.17 Nucleotide sequencing and sequence analysis

Nucleotide sequencing of the amplified DNA was carried out by automated sequencing service rendered by TCGA and Axygen, India. For this purpose, the amplified DNA was gel purified, lyophilized and sent for sequencing. Nucleotide

sequence of the coding as well as the non-coding strands of the double stranded DNA was carried out by using appropriate sequence specific primers.

2.18 Sequence analysis

The amino acid and nucleotide sequences were subjected to BLAST (http://www.ncbi.nlm.nih.gov/Blast) analysis to determine their identity with reference to sequence information available in GenBank/protein databases. ClustalW (www.ebi.ac.uk/ClustalW) software was used to align the sequences with the sequences available in the gene bank/protein databases. Coding regions of the nucleotide sequences were predicted using the AUGUSTUS software (Stanke and Morgenstern, 2005). The EMBOSS CpGplot software of European Bio Informatics Institute was used to detect putative CpG island and microsatellites repeat finder (http://Zlab.bu.edu/ repfind/) for detecting clustered, exact repeat search within the nucleotide sequences identified in the present study. Secondary structure of the deduced amino acid sequences was predicted with PSIPRED (McGuffin *et al.*, 2000) and GOR4 software (Garnier *et al.*,1996). Kyte Doolittle Hydropathy plot (Kyte and Doolittle, 1982) was used for determining the hydrophobicity of the deduced amino acid sequences and Statistical MOTIF SCAN software (http://myhits.isb-sib.ch/cgi-bin/motif_scan) was used for detection of motifs occurring in the sequence. Phylogenetic analysis of the deduced protein sequences were conducted with MEGA4 software (Kumar *et al.*, 2008).

3. CHEMICALS AND SOLUTIONS

3.1 Chemicals

Tris-HCl, glycerol, acetone, Iodine, Potassium iodide, sodium hydroxide, potato amylose, tween-20, ammonium sulphate, sodium azide, trichloro acetic acid (TCA), urea, glycine, ammonium bi-carbonate, trifluoroacetic acid (TFA), acrylamide, bisacrylamide, TEMED, Tris base, glycerol, EDTA, sodium chloride, calcium chloride, SDS, tri-sodium citrate, sucrose, sodium hydroxide, sarkosyl, ethidium bromide, potassium acetate, PVP, boric acid, bromophenol, β-mercaptoethanol, sorbitol, lithium chloride and formaldehyde were purchased from Hi-media Laboratories Ltd. Immobinline dry strip pH 4-7, CHAPS, DTT, Dry strip cover fluid, pharmalyte, iodoacetamide and acetonitrile were purchased from GE Healthcare Ltd. Isopropanol, phenol, chloroform, glacial acetic acid, isoamyl alcohol, acetone and other chemicals of routine use were purchased from Sisco Research Laboratories Pvt. Ltd. India. Agarose was purchased from Bioline and CTAB was procured from Bangalore Genei Pvt. Ltd. (India).

3.2 Molecular weight markers

Protein medium range molecular weight marker (97 kDa-14 kDa), DNA molecular weight markers including λ DNA *Eco*RI/*Hind*III double digest and 1kb ladder were purchased from Bangalore Genei Pvt. Ltd, India.

3.3 Reagent kits and enzymes

Alexa Fluor 546 goat anti-rabbit IgG (H+L) (fluorescent, labeled secondary antibody), BCIP/NBT substrate for alkaline phosphatase, standard potato amylose and

trypsin were procured from M/S Invitrogen, Bangalore Genei Pvt. Ltd. (India) and M/S Sigma–Aldrich Pvt. Ltd. Respectively.

Restriction endonucleases (with restriction buffers), ribonuclease A and spin gel extraction kit were purchased from Bangalore Genei Pvt. Ltd. (India).

3.4 Oligonucleotide primers

The oligonucleotide primers were synthesized at M/S TCGA, New Delhi, as lyophilized powders. Prior to use the oligonucleotide primers were dissolved in appropriate amount of nuclease-free water to make 100 µM stock solutions and aliquoted into 5 µM working solution. The aliquots were stored at -20°C.

3.5 Deoxynucleotied (dNTP) solutions

dATP,dTTP,dGTP and dCTP were procured from Bangalore Genei, India as 10 mM stock solutions. The stock solutions were mixed together and diluted with appropriate amount of nuclease free ultra pure water to give a 5 mM dNTP mix working solution.

3.6 *Taq* DNA polymerase and Assay Buffer

Taq DNA polymerase (3 units/µl) was purchased from M/S Bangalore Genei, India.

3.7 Buffers and Solutions

Starch Granule Extraction Buffer (SDS Lysis Buffer), pH 6.8
60 mM Tris, pH (6.8)
10% SDS
10% Glycerol
3% β-Mercaptoethanol

Tween-20 Tris Buffer Saline
100 mM Tris, pH (7.4)

0.9 % (w/v) Sodium Chloride
0.1% Tween-20

Tris Buffer Saline
100 mM Tris, pH (7.4)
0.9 % (w/v) Sodium Chloride

SDS electrophoresis buffer
25 mM Tris, pH (8.3)
192 mM Glycine
0.1% SDS

Towbin Transfer Buffer
25 mM Tris, pH (8.2-8.4)
192 mM Glycine
0.1% SDS
20 % Methanol

Phosphate buffer saline
137 mM NaCl
2.7 mM KCl
4.3 mM Na_2HPO_4
1.4 mM KH_2PO_4

Rehydration stock solution with IPG Buffer for 2D-PAGE
8M Urea
2% CHAPS
2% IPG Buffer
0.002% Bromophenol blue
1% DTT

SDS- Equilibration Buffer (I) for 2D-PAGE
50 mM Tris-HCl (pH 8.8)
6 M Urea
30% Glycerol
2% SDS
0.002 % Bromophenol blue
1% DTT

SDS- Equilibration Buffer (II) for 2D-PAGE
50 mM Tris-HCl (pH 8.8)
6 M Urea
30% Glycerol
2% SDS

0.002 % Bromophenol blue
2.5% Iodoacetamide

Agarose sealing solution
0.5% Agarose
0.002% (w/v) Bromophenol blue

2X Laemmli Buffer
0.625 M Tris HCl
2% SDS
10 % Glycerol
Increase the volume to 10 ml, add 0.001 M Bromophenol Blue

2X CTAB buffer
100 mM Tris-HCl, pH 8.0
20 mM EDTA.
5 M NaCl
2% CTAB
0.2% β-mercaptoethanol
1% PVP
5 mM glutamic acid
β-mercaptoethanol, PVP and glutamic acid were added just before use.

TE Buffer, pH 8.8
10 mM Tris-Cl, pH 8.0
1 mM EDTA pH 8.0

5X Tris-borate (TBE) buffer (pH 7.5)
800 ml water
20 ml of 0.5 M EDTA, pH 8.0
54 gm of Tris-Base
27.5 gm boric acid.

Gel loading buffer (5X)

25 gm of bromophenol blue was dissolved in 50 ml water and 50 ml of glycerol was added to it. The pH of the solution was adjusted to 7.0 and the final volume was made to 100 ml with water. The solution was filter sterilized and stored at -20°C.

Tris equilibrated phenol
Liquefied phenol, redistilled
50 mM Tris base (adjusted to ~pH 10.5)
TE buffer, pH 8.0

Distilled phenol was mixed with equal volume of 50 mM Tris and stirred continuously with repeated changes of 50 mM Tris till the pH of the phenol was brought to 8.0. The upper aqueous phase was removed from the solution and replaced by $1/10^{th}$ volume of TE buffer to form a thin layer of about 1cm on top of the phenol layer. The buffer saturated phenol was stored in dark bottles at 4°C.

CHAPTER IV

ISOLATIONAL AND CHARACTERIZATION OF GRANULE BOUND STARCH SYNTHASE-I

EXPERIMENTAL

Grains of common buckwheat (*Fagopyrum esculentum* Moench) were procured from North Eastern Regional Station of the National Bureau of Plant Genetic Resources, Shillong, Vivekananda Laboratory for Hill Agriculture, Almora and Buckwheat germplasm collection at the University of Ljubljana, Slovenia.

Healthy grains of uniform size were screened out and used for the present study. A portion of the collections was sown in the experimental garden of Department of Botany, North Eastern Hill University, Shillong for multiplication. The grains were dehulled, frozen in liquid nitrogen and stored at -80°C in an ultra freezer for further studies. Investigations carried out during the present study focused on isolation of the starch granules from the endosperm tissues of common buckwheat and the bound proteins associated with starch grains isolated from the endosperm tissues of seeds as well as from leaf tissues of different accessions/cultivars of buckwheat. The starch grains

were characterized on the basis of their morphology, amylose content, rheological properties and proteome profile.

RESULTS

Starch grains isolated from endosperm tissues of different European and Indian accessions/cultivars of buckwheat were analyzed for their crystalline structure and variations in shape and size by scanning electron microscopy (SEM). When viewed under a scanning electron microscope, buckwheat starch granules showed a monomodal size distribution within an accession/variety (Fig.4.1j). Marked variations were however, observed in the shape and size of starch grains isolated from different accessions/cultivars. The grains ranged in shape from round/spherical to polygonal. While the grains from VL-7 and IC-13145 showed distinct polygonal shape, those from IC-188669, KBB-3, OC-2, Siva and Daria were round to spherical in shape. The grains ranged in size from 3 μm to 12 μm. VL-7 and KBB-3 had the largest starch grains which ranged in size from 9.47 μm to 12.1 μm. The size of starch grains isolated from endosperm tissues of accessions IC-188669, IC-13145 and the variety OC-2 varied between 5.26 μm to 8.95 μm (Fig. 4.1 a, b, c, d, e). VL-7 having the largest size of starch grains is also a high yielding and early maturing cultivar from Western Himalayas released by Indian Council for Agricultural Research, New Delhi. The starch grains from the two European varieties Siva and Daria were distinctly smaller in size compared to that of Indian varieties/accessions. The size of the starch grains isolated from endosperm tissues of the varieties Siva and Daria ranged between 3.15 μm to 6.84 μm (Fig.4.1 f, g). A distinct feature of starch grains isolated from Indian accessions/varieties was the

presence of pores on the surface of the grains (Fig.4.1i). No such pores were visible on the surface of starch grains from European accessions/cultivars. The pores were however, observed to be present only on the smoother surface of starch granules. The grains were stacked in compact layers and were covered with a proteinaceous membrane (Fig. 4.1 h). Scanning electron microscopy of partially digested starch grains showed a clear pattern of concentric rings thereby revealing the lamellar structure of the starch grains (Fig. 4.1 k, l). The rings ranged in thickness from 0.16 to 0.25 μm. Such lamellar structures have been reported to represent the alternation of semi-crystalline and amorphous zone within the matrix. Under light miscroscope, however, the starch granules showed a conspicuous central core which stained blue with potassium iodide and iodine solution (0.2% KI, 0.4% I_2 w/v) (Fig. 4.2).

Irrespective of the variety and size of the starch grains, the percentage of apparent amylose in the starch grains varied between 47% to 51.9% of dry mass of starch. The content of amylose as percent of total dry mass of starch, showed a consistent increase with progressing grain development from milky mature stage to late maturation stage. While the amylose content of starch isolated from seeds harvested 10DAF was 13.5%, that harvested 15DAF, 20DAF and 28DAF showed an amylose content of 27.2%, 38.6% and 47.5% respectively. Thus, there was a more than 3 fold increase in the amylose content of the starch grains with progressive development from milky mature stage to late maturation stage. (Table 4.1 a, b).

There were no marked differences in the rheological properties of starches isolated from different accessions of buckwheat studied in the present investigation. The

Fig. 4.1 Scanning electron micrographs of starch grains isolated from different Indian and European accessions/cultivars of buckwheat. (a)IC-188669, (b) KBB-3, (c) IC-13145, (d) VL-7, (e) OC-2, (f) Siva, and (g) Daria, (h) Starch granules covered with a proteinaceous membrane, (i) enlarged image of a starch granule, (j) Monomodal arrangement of buckwheat starch grains, (k, l) Mature starch grains fractured by sonication and incubated with 3u/ml of alpha amylase for 15 mins at room temperature.

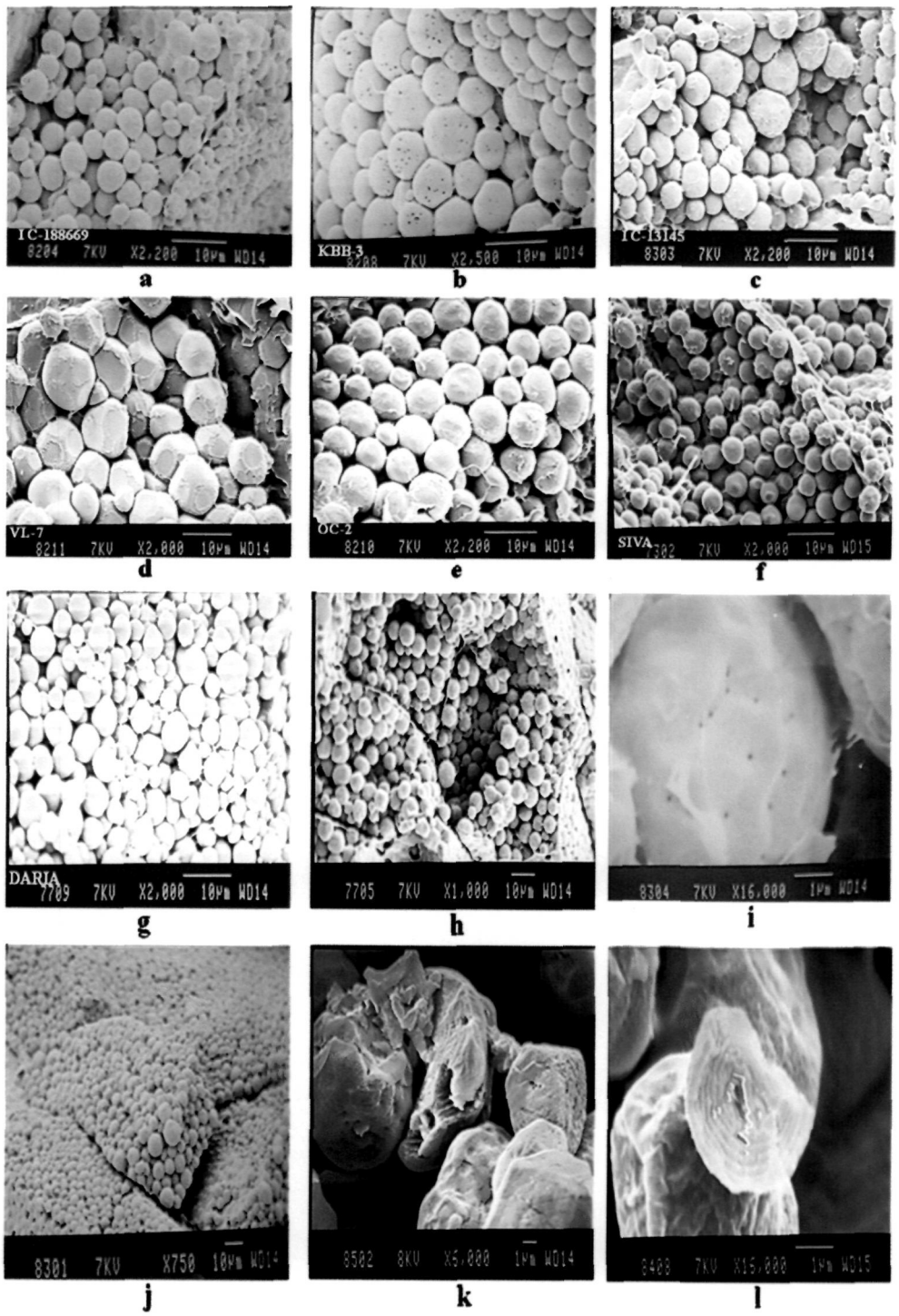

IC-188669
8204 7KV X2,200 10µm WD14
a
KBB-3
7KV X2,500 10µm WD14
b
IC-13145
8303 7KV X2,200 10µm WD14
c
VL-7
8211 7KV X2,000 10µm WD14
d
OC-2
8210 7KV X2,200 10µm WD14
e
SIVA
7KV X2,000 10µm WD15
f
DARJA
7709 7KV X2,000 10µm WD14
g
7705 7KV X1,000 10µm WD14
h
8304 7KV X16,000 1µm WD14
i
8301 7KV X750 10µm WD14
j
8502 8KV X5,000 1µm WD14
k
8408 7KV X15,000 1µm WD15
l

Fig. 4.1

Fig. 4.2 Light microscopy image of matured starch granules stained with 0.2% KI, 0.4% I w/v.

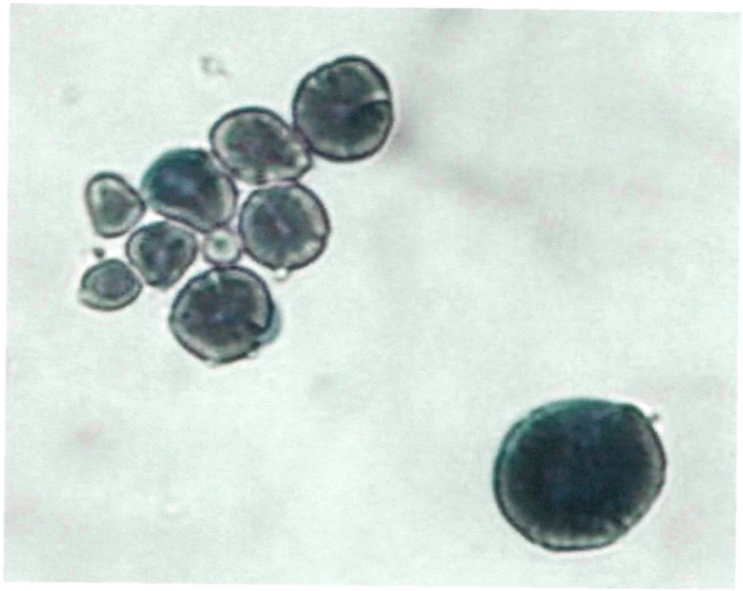


Fig. 4.2

Table 4.1(a): Amylose content ,expressed as mg/100 mg starch in common bucwheat (*F.esculentum* Moench.) at different stages of seed maturation.

Sl. No.	Stages of Seed Development	Amylose content (mg/100mg weight)
1	10 DAF	13.5
2	15 DAF	27.2
3	20 DAF	38.6
4	28 DAF	47.5

Table 4.1(b): Amylose content ,expressed as mg/100 mg starch in certified accessions/cultivars of *F.esculentum* L. procured from India and Europe.

Sl. No.	Accessions/ Cultivars	Amylose content (mg/100mg Starch)
1	IC-188669	51.93
2	IC-319595	48.40
3	OC-2	47.00
4	IC-324244	50.15
5	IC-13145	51.10
6	IC-363973	48.80
7	VL-7	49.20
8	KBB-3	49.56
9	Gor13 04 BF O611	50.11
10	63 Min Gor Siva 05	51.09
11	Gor 04 13BF 06 28	49.56
12	7 Roza 3 Crte	48.90
13	Gor 93 Crte	47.80

starches showed a mean pasting temperature (*P*temp) of 68°C. The peak, minimum, breakdown and final viscosities were found to be 160 RVU, 102 RVU, 58 RVU and 210 RVU respectively. (Fig. 4.3; Table 4.2a). The amylose percentage and rheological properties of starch pastes on heating, cooling and standing at different temperatures describes most of the potential end-use properties of starch in the food systems. Table 4.2b summarizes the RVA chracteristics of starch isolated from different plant sources.

In order to determine the profile of proteins associated with starch grains, SDS-PAGE in one dimensional (1D) as well as two dimensional (2D), with IEF in the 1st dimension and SDS-PAGE in the 2nd dimension, was carried out with granule associated proteins isolated from endosperm tissues of mature seeds. SDS-PAGE analysis of the granule bound proteins from European varieties/selections of buckwheat procurred from the Buckwheat germplasm collection at the University of Ljubljana, Slovenia revealed two types of electrophoretic profiles. The SDS-PAGE profiles of grain proteome of Bf 06X12, Gor 0619 Siva Min Selo, 41 Min Gor05, Gor 1304 BF 0611, Gor 0628 Siva Min Selo, Gor 0413 BF 0618, Gor 0413 BF 0618, Gor 2 2004 3 Crte, Gor 4 2004 3 Crte, Gor 10 20043 Crte and 10 Rdeca 3 Crte revealed the presence of a single band corresponding to molecular mass of 59.7 kDa, while that of Siva Kontrola, Rangus PR-03, 53 Gor Siva Min-05, Gor-06 20 Siva Min Selo, Gor-9 3 Crte, 7-Roza-3-Crte and 191 2X Rumena 3 Crte showed the presence of two bands in the form of a duplex. While one of the bands corresponded to a molecular mass of 59.7 kDa the other showed a molecular mass of 56 kDa (Fig. 4.4 a, b, c, d). On the other hand, SDS-PAGE profiles of proteins associated with starch granules isolated from the endosperm tissues of Indian

Fig. 4.3 RVA profile of normal buckwheat starch for viscosity and temperature as a function of time.

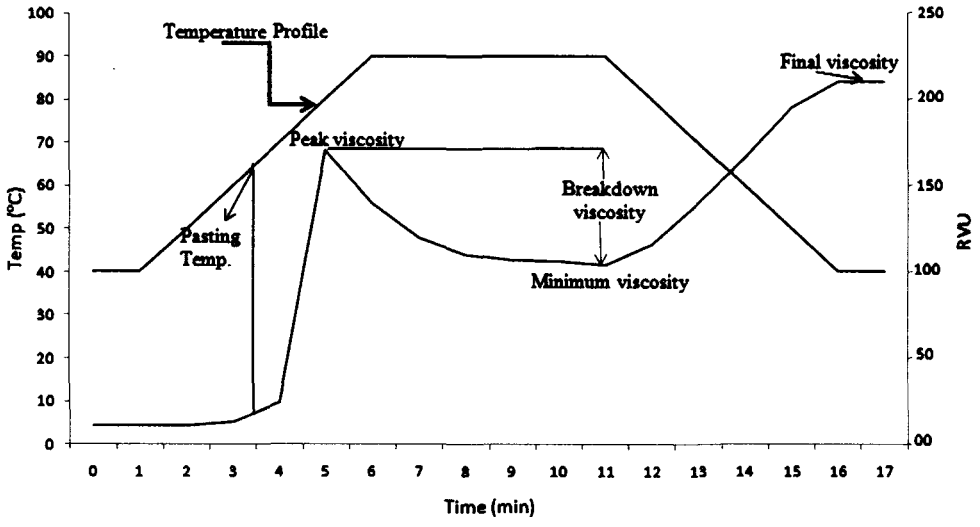

Temperature Profile
Final viscosity
Peak viscosity
Breakdown viscosity
Pasting Temp.
Minimum viscosity
Temp (°C)
RVU
Time (min)
100
90
80
70
60
50
40
30
20
10
0
250
200
150
100
50
00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

Fig. 4.3

Table 4.2(a): Rheological properties of buckwheat starch paste.

Waxy origin	Amylose content (mg/100mg of starch)	Pasting Temp. (°C)	Peak Viscosity (RVU)	Minimum Viscosity (RVU)	Breakdown Viscosity (RVU)	Final Viscosity (RVU)
Non-waxy	47-51	68(°C)	160	102	58	210

Table 4.2(b): Amylose content and rheological properties of starch obtained from different plant sources.

Source	Amylose content (mg/100mg of Starch)	Pasting Temp. (°C)	Peak Viscosity (RVU)	Minimum Viscosity (RVU)	Breakdown Viscosity (RVU)	Final Viscosity (RVU)
Zea mays (non-waxy)	29.4	82.0	113	75	40	138
Zea mays (waxy)	0.0	69.5	162	150	12	190
Triticum aestivum (non-waxy)	28.8	88.6	74	43	31	110
Oryza sativa (non-waxy)	25.0	79.9	113	96	17	160
Oryza sativa (waxy)	0.0	64.1	205	84	121	100
Hordeum vulgare (non-waxy)	25.5	87.8	58	15	43	24
Vigna radiata (non-waxy)	37.9	73.8	186	161	25	363
Amaranthus caudatus (waxy)	3.4	56.9	88	58.0	30	116
Manihot esculenta (non-waxy)	26.3	N.A	845	320	525	640
Solanum tuberosum (non-waxy)	36.0	64	202	165	37	231
F.esculentum (OC-2) (non-waxy)	47	68.2	160	102	58	210
F.tartaricum (KBB-3) (non-waxy)	49.5	68.0	161	102	59	210

Fig. 4.4 SDS-PAGE profile of granule associated proteins isolated from the endosperm tissues of different Slovenian varieties/selections of buckwheat.
(a):-{M}=Protein molecular weight marker, L1= Bf 06X12, L2= Gor 0619 Siva Min Selo, L3= 41 Min Gor05, L4=Siva Kontrola Rangus PR03, L5=53 Gor Siva Min 05, L6=Gor 0620 Siva Min Selo.
(b):-{M}=Protein molecular weight marker, L1=Gor 1304 BF 0611, L2=63 Min Gor Siva 05, L3=Gor 0628 Siva Min Selo, L4=Gor 0413 BF 0618, L5=Gor 0413 BF 0618.
(c):-{M)=Protein molecular weight marker L1=Gor 11 2004 3 Crte, L2= Gor 2 2004 3 Crte, L3=Gor 4 2004 3 Crte, L4=Gor 10 2004 3 Crte, L5=191 2X rumena 3 Crte.
(d):-{M}=Protein molecular weight marker, L1=Gor 93 Crte, L2= 7 Roza Crte, L3= 10 Rdeca 3 Crte.

Fig. 4.5 SDS-PAGE profile of granule associated protein isolated from leaf tissues (L1) and endosperm tissues (L2) of Indian accessions/ varieties of buckwheat. M=Marker.

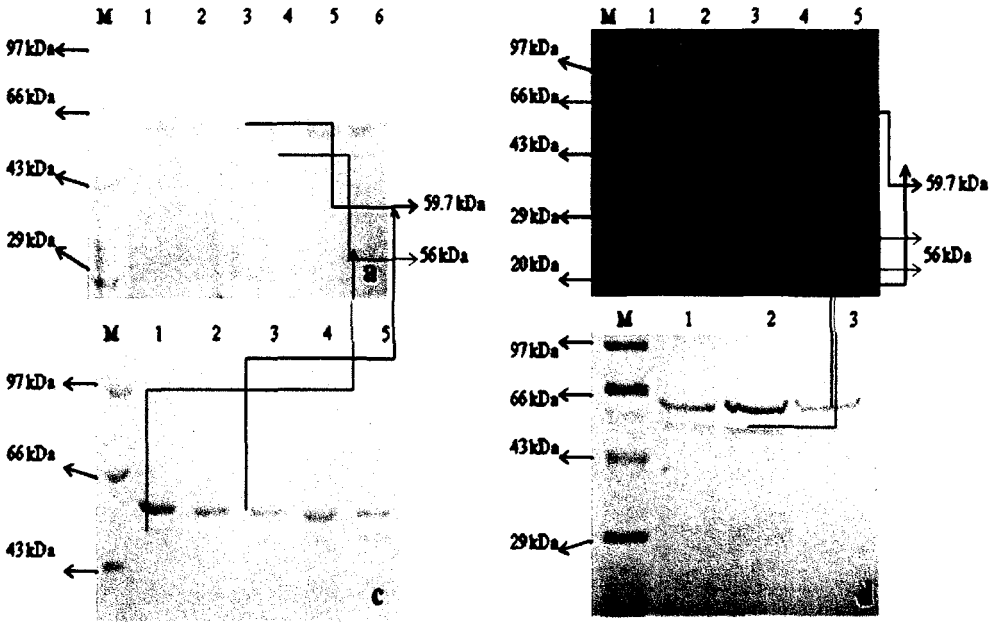

M 1 2 3 4 5 6
97kDa
66kDa
43kDa
29kDa
59.7kDa
56kDa
a
M 1 2 3 4 5
97kDa
66kDa
43kDa
c
M 1 2 3 4 5
97kDa
66kDa
43kDa
29kDa
20kDa
59.7kDa
56kDa
M 1 2 3
97kDa
66kDa
43kDa
29kDa

Fig. 4.4

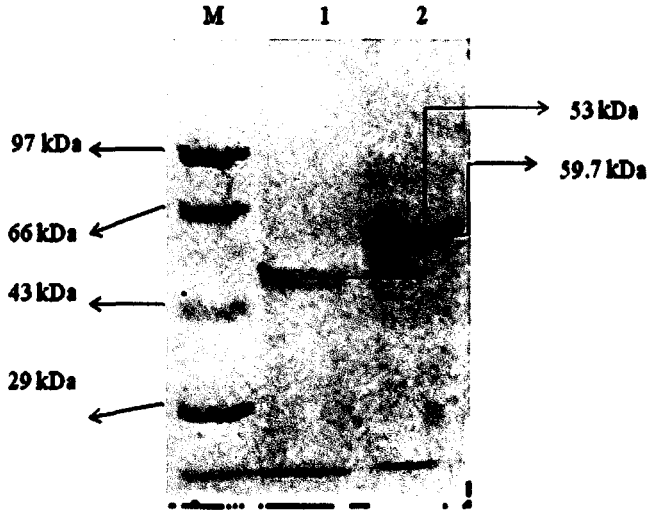

M 1 2
97 kDa
66 kDa
43 kDa
29 kDa
53 kDa
59.7 kDa

Fig. 4.5

accessions/varieties of buckwheat showed the presence of a single band corresponding to a molecular mass of 59.7 kDa (Fig. 4.5). SDS-PAGE profile of the proteins associated with starch grains isolated from leaves of common buckwheat revealed the presence of a single band corresponding to a molecular mass of 53 kDa (Fig. 4.5). The SDS-PAGE profiles of proteins associated with starch grains of varieties/accessions from Indian Himalayas as well as the European collections/selections did not show the presence of any high molecular weight bound protein, which have been reported from other plants such as pea and wheat.

BLASTp analysis of the N-terminal amino acid sequence identified it as a granule bound starch synthase-I (GBSS-I) enzyme. The sequence has been deposited in the Swiss-Prot protein data bank with accession no. P84633 (Fig. 4.6). Multiple alignment of the N-terminal sequence of GBSS-I protein isolated from buckwheat with amino acid sequences of similar proteins available in protein data banks clearly identified the conserved KTGGL motif, which has been reported as the ADP/ADPglucose binding site (Fig. 4.7). While the residue immediately preceding and following this motif in soluble starch synthase is usually basic, there are no basic residues in the immediate vicinity of this motif in the amino acid sequence identified in the present study. The 59.7 kDa protein identified in the present study showed 100% similarity from its 13^{th} to 25^{th} amino acid residues, with N-terminal amino acid residues at the same positions in GBSS-I from other plant species. Further, the region from the N-terminus to the 25^{th} residue of the 59.7 kDa GBSS-I protein showed 94% homology with GBSS-I from *Hordeum vulgaris*, *Triticum* spp. and *Phaseolus vulgaris*. The percentage

Fig. 4.6 N-terminal amino acid sequence for the 59.7 kDa starch granule bound protein of common buckwheat.

Fig. 4.7 Comparison of N-terminal amino acid sequences of the 59.7 kDa GBSS-I protein isolated from buckwheat starch granules with the N-terminal amino acid sequences of GBSS-I proteins from other crops .

Fig. 4.6

GBSS-I	1 5 10 15 20 25
F.esculentum	GMNLVFVGAEVAPWS[redacted]GDVLA
M.esculenta[b]	GMNLIFVGAEVAPWS[redacted]GDVLA
S.tuberosum[c]	GMNLIFVGAEVGPWS[redacted]GDVLA
P.sativum[d]	GMSLVFVGAEVAPWS[redacted]GDVLA
T.aestivum[e]	GMNLVFVGAE[redacted]APWS[redacted]GDVLA
H.vulgare[e,f]	GMNLVFVGAE[redacted]APWS[redacted]GDVLA
O.sativa[e,g]	GMNLVFVGAE[redacted]APWS[redacted]GDVLA
Z.mays[h]	GMNLVFVGAE[redacted]APWS[redacted]GDVLA
S.bicolor[i]	GMNLVFVGAE[redacted]APWS[redacted]GDVLA

[b]Salehuzzaman *et al.* (1993), [c]van der Leij *et al.* (1991), [d]Hsieh *et al.* (1996), [e]Taira *et al.* (1995),[f]Rohde *et al.* (1988), [g]Wang *et al* (1990),[h]Klosgen *et al* (1986), [i]Dry *et al.* (1992).

Fig. 4.7

homology with GBSS-I from other plant species, however, varied between 94% and 88%. Even though analysis of the sequence alignment revealed a clear diversification into monocotyledonous and dicotyledonous groups, the sequence from buckwheat showed similarities with GBSS-I from both the groups. While majority of GBSS-I sequences have Isoleucine as the fifth N-terminal residue, the residue in buckwheat GBSS-I is substituted by Valine as is also the case with GBSS-I of monocots. On the other hand, the N-terminal amino acid sequence of buckwheat GBSS-I has Valine at P_{11} as is the case with GBSS-I from dicots. This is in contrast with the N-terminal amino acid sequence of monocots which have methionine at P_{11}. These differences may imply differences in catalytic activity of the enzyme. The sequence analysis of buckwheat GBSS-I indicates similarities with both cereal as well as dicot GBSS-I sequences. It is possible that the protein from common buckwheat might belong to a catalytically distinct subclass and hence, a possible candidate for altering amylose biosynthesis in both dicots as well as monocots.

While immunoblotting of the endosperm starch granule associated proteins separated by SDS-PAGE with antibodies raised against buckwheat GBSS-I showed strong cross reactivity with the 59.7 kDa protein, there was no cross reactivity with the 56 kDa protein associated with the starch grains of European accessions/varieties (Fig. 4.8 a, b, c, d). Antisera raised against the 59.7 kDa protein also failed to crossreact with the 53 kDa protein associated with the starch grains isolated from leaf tissues of buckwheat, thereby, indicating absence of any serological homology between the 53 kDa protein and GBSS-I (Fig 4.9). The antisera, however cross reacted with the 61 kDa

Fig. 4.8 Immunoblot profile of starch granule associated proteins separated by SDS-PAGE with antibodies raised against buckwheat GBSS-I. Cross–reacting proteins were visualized by specific anti-59.7 kDa antibody and HRP-goat anti-rabbit IgG conjugate.
(a) L1=Bf 06X12, L2=Gor 0619 Siva Min Selo, L3=41 Min Gor05, L4=Siva Kontrola Rangus PR03, L5=53 Gor Siva Min 05, L6=Gor 0620 Siva Min Selo.
(b) L1=Gor 1304 BF 0611, L2=63 Min Gor Siva 05, L3=Gor 0628 Siva Min Selo, L4=Gor 0413 BF 0618, L5=Gor 0413 BF 0618.
(c) L1=Gor 11 2004 3 Crte, L2=Gor 2 2004 3 Crte, L3=Gor 4 2004 3 Crte, L4=Gor 10 20043 Crte, L5=191 2X rumena 3 Crte.
(d) L1=Gor 93 Crte, L2=7 Roza Crte, L3=10 Rdeca 3 Crte.

Fig. 4.9 Immunoblot profile of the starch granule associated proteins isolated from leaf (L1) and endosperm starch grains (L2) of common buckwheat with anti-59.7 kDa serum raised against buckwheat GBSS-I. Cross-reacting proteins were visualized by specific anti-59.7 kDa antibody and HRP-goat anti-rabbit IgG conjugate.

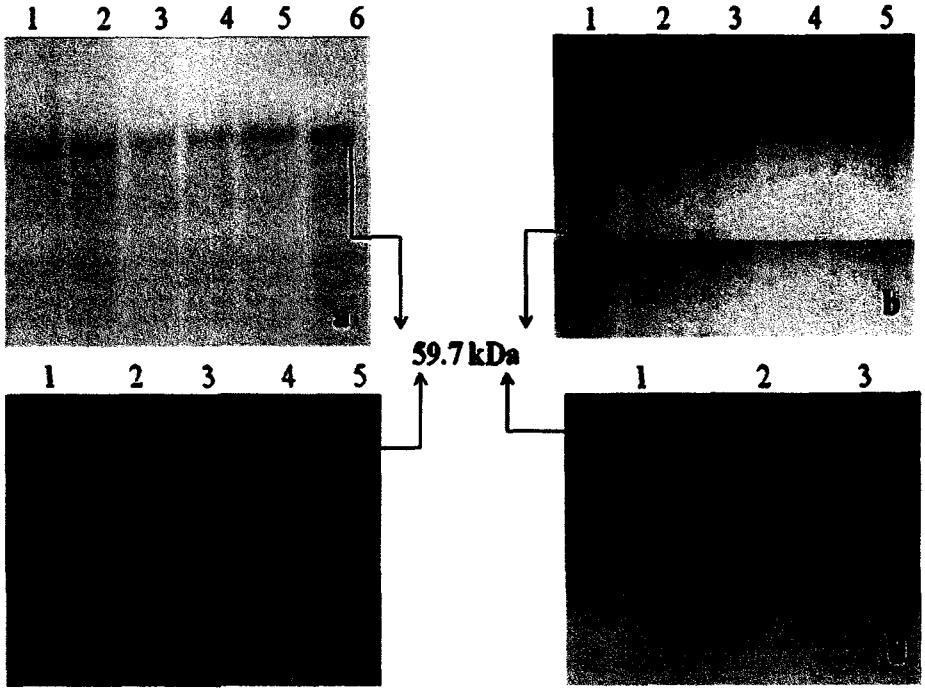


Fig. 4.8

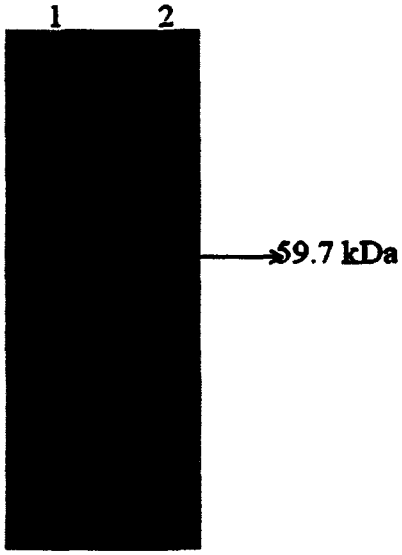


Fig. 4.9

kDa GBSS-I isolated from endosperm starches of maize and the 60 kDa GBSS-I isolated from endosperm starches of rice and wheat (Fig.4.10 a, b). These result indicate serological homology between the GBSS-I of wheat, maize, rice and buckwheat.

Trypsin digestion of the 59.7 kDa endosperm starch granule associated protein of buckwheat and the 60 kDa GBSS-I of rice and subsequent size fractionation of the tryptic digestion products on a 15% acrylamide gel resulted in generation of two fragments showing molecular mass of 31 kDa and 28 kDa for both the proteins. Trypsin digestion of the 53 kDa leaf starch granule associated protein of buckwheat resulted in generation of two fragments showing molecular mass of 27 kDa and 26 kDa (Fig. 4. 11).

2D-PAGE of the endosperm starch granule associated proteins resolved the fraction into 10 spots with *pI* ranging from 5.2 to 6.2 (Fig. 4.12 a). All the spots showed an apparent molecular mass of 59.7 kDa. Immunoblotting of the proteins separated by 2D-PAGE with antisera raised against buckwheat GBSS-I identified two bands [spot no 3 (*pI* 5.4) and spot no 8 (*pI* 6.1)] which cross reacted with antibodies raised against buckwheat GBSS-I (Fig. 4.12 b). The protein corresponding to spot no. 3 was subjected to in-gel trypsin digestion, followed by Mass Spectrometry and database search with MASCOT (www.matrixscience.com). Since the identification of peptide fragments relies on sequence information present in databases, this is a limitation for buckwheat, which has no sequence information on granule bound protein. Buckwheat proteins were thus identified by comparing them with the rice genomes (Table 4.3). However, one of the tryptic fragment having the amino acid sequence "FNAPLAHLIMAGADVLAVPSR" with a predicted *pI* of 8.34 showed similarity with GBSS-I protein of rice. The spectrum

Fig. 4.10a SDS-PAGE profile of GBSS-I proteins isolated from endosperm starch grains of maize (L1), rice (L2) and wheat (L3). M=protein molecular weight marker

Fig. 4.10b Immunoblot profile of GBSS-I proteins isolated from starch grains of maize (L1), rice (L2) and wheat (L3) cross-reacted with anti-59.7 kDa serum raised against buckwheat GBSS-I. Cross-reacting proteins were visualized by specific anti-59.7 kDa antibody and HRP-goat anti-rabbit IgG conjugate.

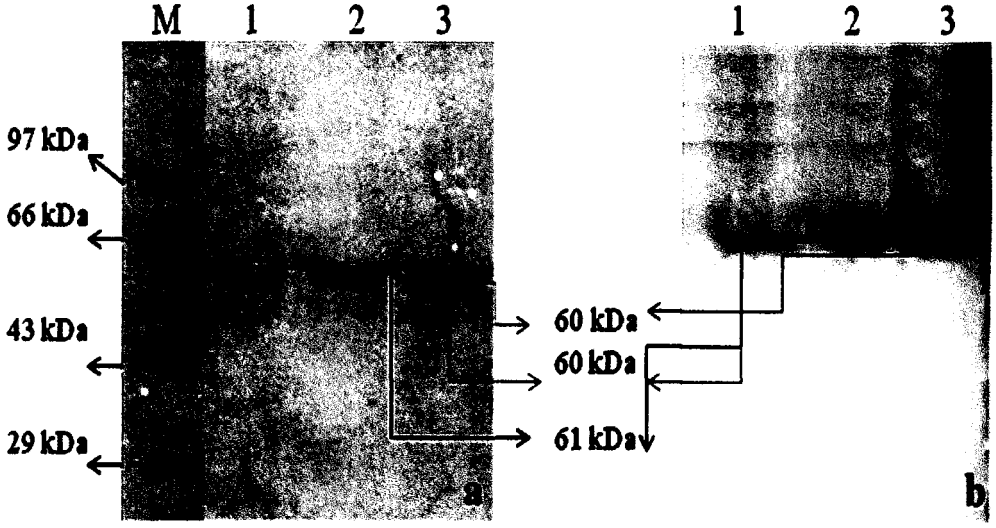


Fig. 4.10

Fig. 4.11 SDS-PAGE profile of starch granule associated protein isolated from endosperm tissues of buckwheat and rice and leaf tissue of buckwheat.
L1 and L2= undigested 59.7 kDa and 60 kDa GBSS-I protein isolated from endosperm starches of buckwheat and rice.
L3= undigested 53 kDa starch granule associated protein isolated from leaf starches of common buckwheat.
L4 & L5= Trypsin digestion of 59.7 kDa and 60 kDa GBSS-I proteins isolated from endosperm starches of common buckwheat and rice.
L6= Trypsin digestion of 53 kDa starch granule associated protein isolated from leaf starches of common buckwheat.

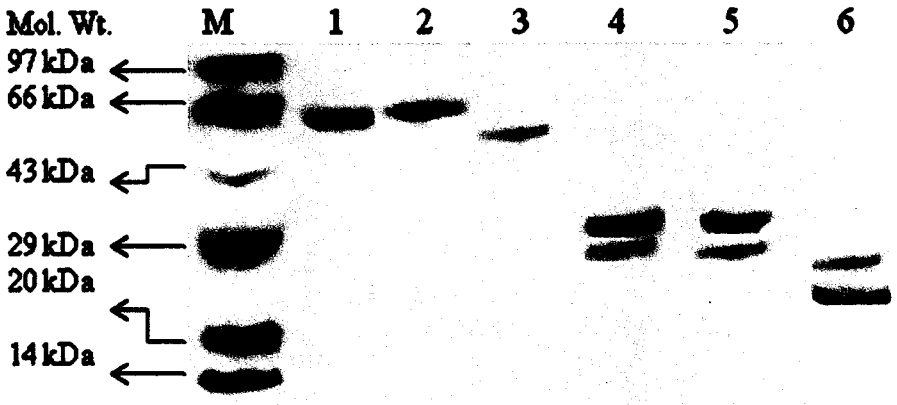


Fig. 4.11

Fig. 4.12a 2D-PAGE profile of the endosperm starch granule associated proteins isolated from common buckwheat.

Fig. 4.12b Immunoblot profile of the endosperm starch granule associated proteins isolated from common buckwheat with anti-59.7 kDa serum raised against buckwheat GBSS-I.

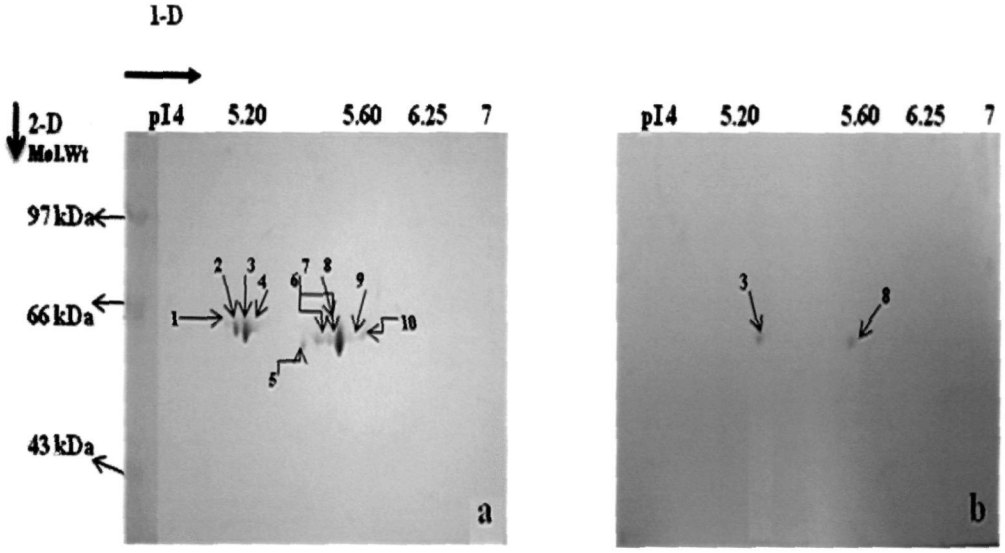

1-D
2-D
Mol.Wt
pI 4
5.20
5.60
6.25
7
97 kDa
66 kDa
43 kDa
1
2
3
4
5
6
7
8
9
10
a
pI 4
5.20
5.60
6.25
7
3
8
b

Fig 4.12

Table 4.3: Identification of the Tryptic Fragment of GBSS-I (spot no.3 of molecular wt. 59.7 kDa) by comparison with rice genome using MASCOT software analysis tool

No.	Rice Protein Name	Mol.Wt (kDa)	*pI*	Sequences
1	Granule-bound starch synthase 1, chloroplastic/amyloplastic	66434	8.34	FNAPLAHLIMAGADVLAVPSR
2	Granule-bound starch synthase 1, chloroplastic/amyloplastic	66281	8.34	FNAPLAHLIMAGADVLAVPSR
3	Alpha-amylase isozyme COS=Oryza sativa subsp. japonica	38454	11.74	MQVPIEHNGGEQTLLVPFR
4	Adenylate kinase A OS=Oryza sativa subsp.	26389	8.49	MAANLEDVPSMELMTELLR
5	Cysteine synthase OS=Oryza sativa subsp. japonica	34285	5.35	IYGVEPTESAILSGGRPGPHK
6	Cyclin-dependent kinase F-1 OS=Oryza sativa subsp. japonica	52040	4.62	WMLQVLEGVAACHSAGVVHR
7	Probable indole-3-acetic acid-amido synthetase	68976	5.32	VASGGEPSSALLCSDPITCLSR
8	Abscisic acid 8'-hydroxylase 3 OS=Oryza sativa subsp. indica	56122	9.11	WEIVGSSDEVEYSPFPVPK
9	Probable serine/threonine-protein kinase WNK4 OS=Oryza sativa subsp. japonica	68666	4.94	ELLQDPFLCSDNSSVLVGTK
10	Auxin response factor 1 OS=Oryza sativa subsp. japonica	77454	5.48	MSSQGAGGGVGDPELFAELWR
11	DEAD-box ATP-dependent RNA helicase 40 OS=Oryza sativa subsp. japonica	87927	10.00	EDEEEGMIDEDGEIADDPR
12	Coatomer subunit alpha-3 OS=Oryza sativa subsp. japonica	136478	6.43	SGAWDENGVFIYTTLNHIK
13	Lipoxygenase 7, chloroplastic OS=Oryza sativa subsp. japonica	102755	5.91	QLSEMHPIYQLLRPHFR
14	Beta-galactosidase 15 OS=Oryza sativa subsp. japonica	100942	5.66	SENYYPPLSAWSHLSSGR

shown in Fig.4.13 displayed the monoisotoped mass values of these tryptic fragments. On the other hand, 2D-PAGE of the leaf starch granule associated proteins resolved the fraction into 2 spots with *pI* ranging from 4.5 to 5.2 (Fig. 4.14).

Confocal laser scanning microscopy after cross reacting with antibodies raised against buckwheat GBSS-I, clearly revealed the localization of GBSS-I in the form of discrete internal rings within the matrix of buckwheat starch grains (Fig. 4.15b). The protein appeared to be arranged in the form of radial arrays emerging from the core of the grains.

Phylogenetic analysis of the N-terminal amino acid sequence of the GBSS type protein from common buckwheat reported here revealed a clear diversification into monocotyledonous and dicotyledonous groups (Fig 4.16). Within the monocots, the sequences could be segregated into two groups. While one of the two groups (Group-I) was dominated by rice (*Oryza* spp.) and maize (*Zea mays*), the other group (Group-II) predominantly comprised sequences from *Triticum* spp, *Hordeum* spp, *Secale cereale*, *Elymus scaber* and *Aegilops speltoides*. The dicotyledons on the other hand, were represented in three subgroups *viz.*, Group-III, Group-IV and Group-V. Group-III comprised of sequences from *Fagopyrum esculentum*, *Nelumbo nucifera*, *Astragalus membranaceus* and *Amaranthus cruentus*. Group-IV was represented by sequences from *Pisum sativum*, *Phaseolus vulgaris* and *Manihot esculenta* and Group V by *Ipomoea batatas*. These results are in conformity with the observations of Edwards *et al.* (2002), who have reported a clear division of GBSS-I proteins into those belonging to monocots and those to dicots.

Fig. 4.13 MS analysis of tryptic peptide mass fingerprinting of protein spot no3 of molecular mass 59.7 kDa isolated from common buckwheat starch granule. The upper values indicate the observed mass peaks. The lower values labeled the corresponding regions in the proteins

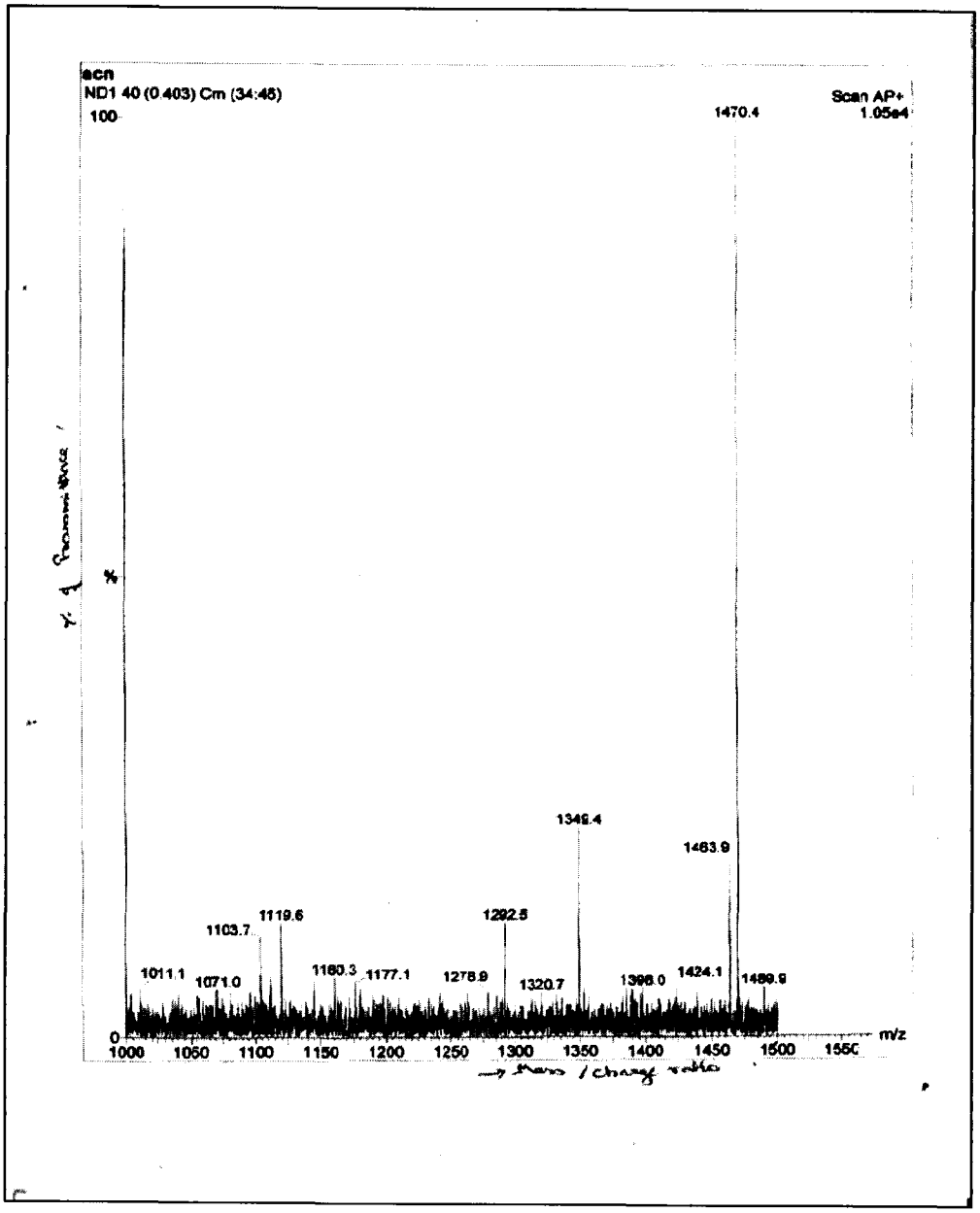

acn
ND1 40 (0.403) Cm (34:45)
100
Scan AP+
1.05e4
1470.4
%
1348.4
1463.9
1292.5
1119.6
1103.7
1011.1
1071.0
1180.3
1177.1
1278.9
1320.7
1398.0
1424.1
1489.9
0
m/z
1000
1050
1100
1150
1200
1250
1300
1350
1400
1450
1500
1550

Fig .4.13

Fig. 4.14 2D-PAGE profile of starch granule associated protein isolated from leaf tissues of common buckwheat.

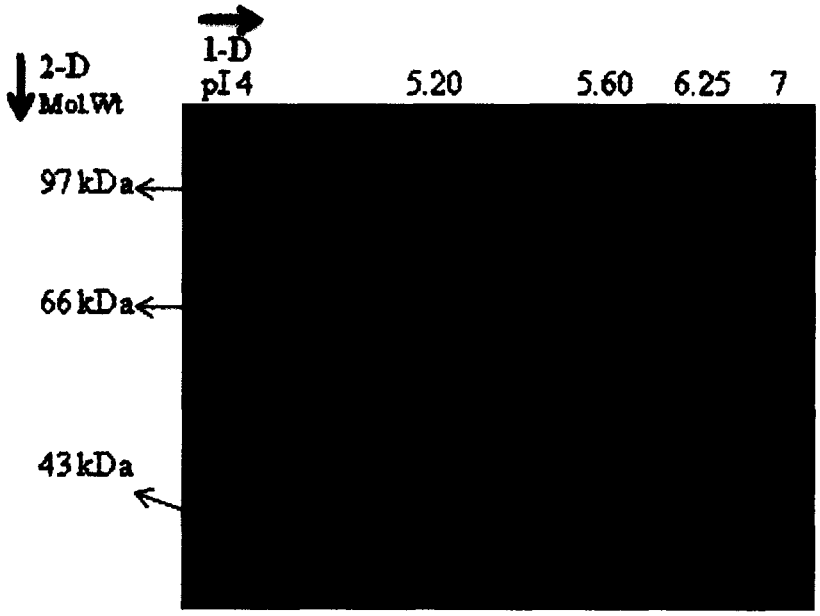


Fig 4.14

Fig. 4. 15 Images of starch granules:-(a) Bright field light micrographs of unstained starch granules; (b) Confocal laser-scanning micrographs of starch granules.Physiochemical localization of starch granule associated proteins was carried out by immunodetection with fluorescent labelled secondary antibody (Alexa- 546 conjugated goat anti-rabbit IgG, Invitrogen).

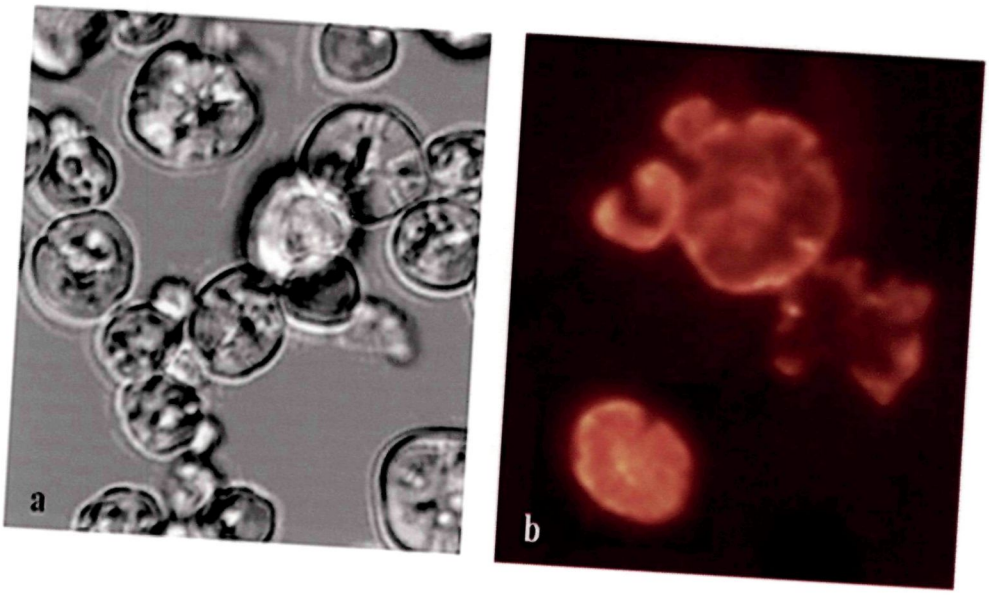


Fig . 4. 15

Fig. 4.16 Phylogenetic tree based on alignment matrix of N-terminal amino acid sequence of 59.7 kDa GBSS-I protein of common buckwheat with the amino acid sequences of GBSS type proteins available in the protein data bank.

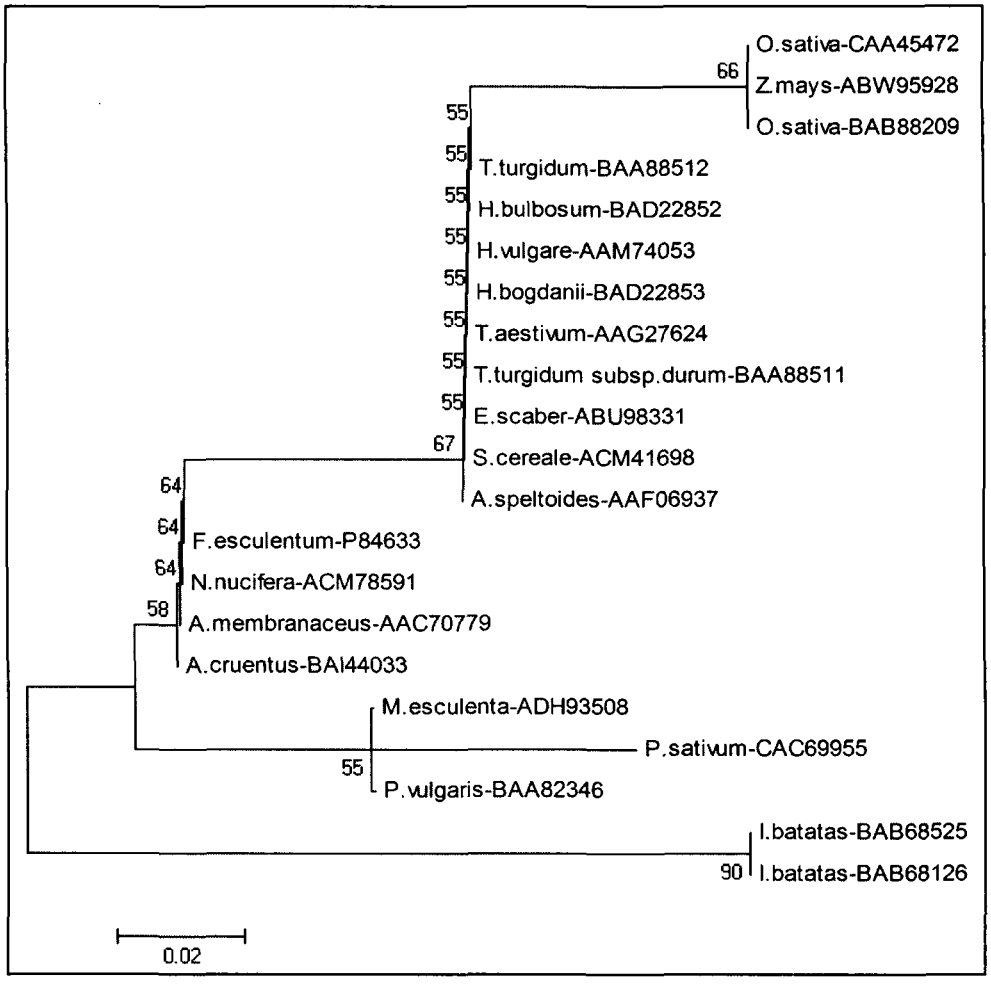

O.sativa-CAA45472
66
Z.mays-ABW95928
55
O.sativa-BAB88209
55
T.turgidum-BAA88512
55
H.bulbosum-BAD22852
55
H.vulgare-AAM74053
55
H.bogdanii-BAD22853
55
T.aestivum-AAG27624
55
T.turgidum subsp.durum-BAA88511
55
E.scaber-ABU98331
67
S.cereale-ACM41698
64
A.speltoides-AAF06937
64
F.esculentum-P84633
64
N.nucifera-ACM78591
58
A.membranaceus-AAC70779
A.cruentus-BAI44033
M.esculenta-ADH93508
P.sativum-CAC69955
55
P.vulgaris-BAA82346
I.batatas-BAB68525
90
I.batatas-BAB68126
0.02

Fig. 4. 16

DISCUSSION

Starch is an important ingredient in many food products and an important source of carbohydrate in the human diet (Ball *et al.*, 1996). The functional properties of starch such as gelatinization, pasting, retrogradation, viscosity, swelling and solubility vary considerably among starches from different sources and are therefore unique for each type (Perez *et al.*, 2005; Peroni *et al.*, 2006; Yuan *et al.*, 2007). These differences are dependent on composition and molecular structures of the starches which include amylose/amylopectin ratio, granular size, phosphorous content, molecular weight of the starches and chain length distribution of amylopectin (Fredriksson *et al.*, 1998; Sasaki and Matsuki, 1998; Lu *et al.*, 2005).

While much work has been done on the characterization of starch from cereals and analysis of the regulation of its biosynthesis, not much information is available about the quality of starch in many other potentially important crops. Amongst this group of plants, common buckwheat is an important pseudo cereal because of its high potential for use as a functional food. However, the flour obtained from buckwheat has poor dough making qualities due to its unfavorable amylose/amylopectin ratio, adding to its less popularity as a flour-yielding crop. Since the ratio of amylose to amylopectin influences the texture and quality of flour, identification of the granule bound starch synthase-I (GBSS-I) protein and its characterization would have an important bearing on improving starch quality. Even though such proteins have been isolated from other conventional crops, none of the granule bound proteins in buckwheat have been identified. In the present investigation isloation and characterization of the GBSS-I

protein from common buckwheat (*Fagopyrum esculentum*) was carried out for gaining insight in crop improvement programmes with improved amylose to amylopectin ratio for better quality of starch production.

Standard starch granule bound protein extraction protocol was followed for isolation of starch grains and starch grain bound proteins from endosperm of common buckwheat according to Takaoka *et al.* (1997). Marked variations were observed in the shape and size of starch grains isolated from different accessions/cultivars of common buckwheat. The grains ranged in shape from round/spherical to polygonal and showed a monomodal distribution with size from 3 μm to 12 μm. While the grains from VL-7 and IC-13145 showed distinct polygonal shape, those from IC-188669, KBB-3, OC-2, Siva and Daria were round to spherical in shape. VL-7 having the largest size of starch grains is also a high yielding and early maturing cultivar from Western Himalayas released by Indian Council for Agricultural Research, New Delhi. Compared to the Indian accessions, the starch grains from European accessions were distinctly smaller in size. Kim *et al.* (1977), Soral-Smietana *et al.* (1984), Acquistucci and Fornal, (1997) and Qian and Kuhn (1999b), have also reported similar variations in the shape and size of starch grains from endosperms tissues of common buckwheat. Lindeboom *et al.* (2004), have on the basis of their size, classified starch grains as large (>25μm), medium (10-25μm), small (5-10μm) and very small (>5μm). Starch grains isolated from the endosperm of buckwheat can be clearly classified under the small grain size category. Starches having small granules and a narrow granule size distribution have found application in fine printing paper and plastic sheets (Jane *et al.*, 1994; Wilhelm *et al.*, 1998), as a binder

with orally active ingredients and as a carrier material in cosmetics (Whistler, 1995). Aqueous dispersions of small starch granules are known to produce a creamy, smooth texture that exhibits fatmimetic properties (Daniel and Whistler, 1990; Biliaderis *et al.*, 1993; Jane *et al.*, 1994; Malinski *et al.*, 2003). Because of their small size, starch granules from grains of amaranth and quinoa have been reported to have applications as fat substitutes (Whistler 1997). Quinoa starch is also an important raw material in production of fine printing paper (Jane *et al.*, 1994; Wilhelm *et al.*, 1998), as a binder with orally active ingredients, and as a carrier material in the cosmetics and in textile and photographic industries (Biliaderis *et al.*, 1993, Whistler, 1995). Quinoa starch has also been reported to finds application as filler in biodegradable films as small granule size can substantially increase the level of starch that can be incorporated into these films while maintaining film quality (Lim *et al.* 1992; Ahamed *et al.*, 1996). Due to their small size, starch grains of common buckwheat may find similar applications.

Scanning electron microscopy of starch grains revealed the presence of pores on the surface of the grains. Similar observations on the presence of pores in buckwheat starch grains have been made by Qian *et al.* (1998). Fannon *et al.* (1993), have reported the presence of pores on the surface of corn, sorghum and millet starch granules as well as along the equatorial groove of large granules of wheat, rye and barley. They observed that the pores were the external openings of tube-like channels present within the grains. Qian *et al.* (1998), have ascribed the higher susceptibility of buckwheat starch to amylase action to the presence of pores on the surface. Porosity and surface area are important characteristics of solid materials that determine their properties e.g., thermal

conductivity, thermal diffusivity, mass diffusion coefficient. Mechanical and textural properties of food are also dependent on porosity (Marousis and Saravacos, 1990). Even though grain size has been described as an important parameter that determines the gelatinization behavior of starch (Zheng and Sosulski, 1997; Chiotelli and Meste, 2002; Lindeboom *et al.*, 2004), many other reports have shown absence of any correlation between the two (Goering and DeHaas, 1972; Eliason and Larsson, 1983; MacGregor and Bhatty, 1996; Myllarinen *et al.*, 1998; Chiotelli and Meste, 2002).

Scanning electron microscopy of partially digested starch grains showed a clear pattern of concentric rings thereby revealing the lamellar structure of starch grains. Such lamellar structures have been reported to represent the alternation of semi crystalline and amorphous zone within the matrix (French, 1984; Cameron and Donald, 1992). This alternation of semi crystalline and amorphous zones within the matrix of starch grains has been correlated with the presence of GBSS-I within the grains (Denyer *et al.*, 1995; Smith *et al.*, 1997). Confocal laser scanning microscopy of buckwheat starch however, clearly revealed the localization of GBSS-I in the form of discrete internal rings within the matrix of buckwheat starch grains. Even though many reports have indicated the localization of GBSS-I within the matrix of the starch grains rather than at its surface, most of these observations have been made on indirect evidences including organization of amylose within the granules (Han *et al.*, 2001). Han *et al.* (2001), Han and Hamaker (2002), have shown that the protein was present in the form of discrete internal concentric spheres in starch grains of potato, corn and wheat. Our results clearly demonstrate the localization of GBSS-I in the form of discrete internal rings within the

matrix of buckwheat starch grains. Our observations on the localization of GBSS-I in starch grains are consistent with the assumption that amylose synthesis occurs within the core of the starch granules (Tatge *et al.*, 1999). Han *et al.* (2001), have hypothesized that its granular location may impact gelatinized starch rheology. Strengthening of the gelatinized starch granule by an internal starch granule bound protein could reduce paste breakdown.

Irrespective of grain size, the percentage of apparent amylose in buckwheat starch ranged between 47% to 51.9%. The content of amylose in the grains showed a progressive increase with grain development. These results are in conformity with similar observations made in case of starches from other plant species by Shannon and Garwood (1984). The apparent amylose content of buckwheat starch has been reported to range between 15% to 52% (Morrison 1981, 1988; Campbell, 1997; Noda *et al.*, 1998; Qian and Kuhn, 1999b; Yoshimoto *et al.*, 2004). Thus, compared to that of wheat and corn buckwheat starch has a much higher content of amylose.

There were no marked differences in the rheological properties of starches isolated from different accessions of buckwheat studied in the present investigation. The starches showed a mean pasting temperature (*P*temp) of 68°C. Our result is in confirmity with Noda *et al.* (1998); Qian *et al.* (1998); Qian and Kuhn, (1999b); Yoshimoto *et al.* (2004), who reported that the pasting temperature of buckwheat starch ranged from 68.5°C to 71°C. The peak, minimum, breakdown and final viscosities were 160 RVU, 102 RVU, 58 RVU and 210 RVU respectively. Similar results on peak, minimum, breakdown and final viscosities of buckwheat starch pastes were reported by

Qian *et al.* (1998). The amylose percentage and rheological properties of starch pastes on heating, cooling and standing at different temperatures describes most of the potential end-use properties of starch in the food systems. Zheng *et al.* (1998) and Qian *et al.* (1998), have shown that compared to corn or wheat starches, buckwheat starches swelled faster, exhibited a greater set back viscosity and formed stiffer and harder gels. Besides super molecular glucan structures, the high viscosity values for buckwheat starches can be explained by the fact that the starches exhibited a higher granule swelling and gelling tendency than cereal starches (Yoshimoto *et al.*, 2004).

Even though the swelling and pasting properties of starch has been shown to be more related to granule structure and chemical composition than grain size (Zheng and Sosulski, 1997; Chiotelli and Meste, 2002), Lindeboom *et al.* (2004), have suggested that grain size was an important parameter in determining the gelatinization behaviour of starch grains. Lorenz (1990), has shown that quinoa starch, which had smaller size starch grains with a monomodal distribution, showed higher pasting temperatures than starches with large grain size.

SDS-PAGE profile of proteins associated with starch grains from different Slovenian varieties/selections of buckwheat revealed two types of electrophoretic profiles. While the SDS-PAGE profiles of starch grains isolated from varieties Bf 06X12, Gor 0619, Siva Min Selo, 41 Min Gor05, Gor 1304, BF 0611, Gor 0628, Siva Min Selo, Gor 0413 BF 0618, Gor 0413 BF 0618, Gor 2 2004 3 Crte, Gor 4 2004 3 Crte, Gor 10 20043 Crte and 10 Rdeca 3 Crte revealed the presence of a single band corresponding to molecular mass of 59.7 kDa, that of Siva Kontrola, Rangus PR-03, 53

Gor Siva Min-05, Gor-06 20 Siva Min Selo, Gor-9 3 Crte, 7-Roza-3-Crte and 191 2X Rumena 3 Crte showed the presence of a duplex with molecular masses of 59.7 kDa and 56 kDa. On the other hand, SDS-PAGE profiles of proteins associated with starch granules isolated from the endosperm tissues of Indian accessions/varieties of buckwheat showed the presence of a single band corresponding to a molecular mass of 59.7 kDa. However, SDS-PAGE profile of the proteins associated with starch grains isolated from leaves of common buckwheat revealed the presence of a single band corresponding to a molecular mass of 53 kDa. Our results on the SDS-PAGE profile of starch granule bound proteins of buckwheat are in conformity with similar observations made on granule associated proteins of potato and wheat (Takaoka *et al.*, 1997; Han and Hamaker, 2002). Starch granule associated proteins in endosperm of common wheat (*Triticum aestivum* L.) have been reported to include at least one major protein with a molecular weight of 61 kDa and six minor high molecular weight proteins. While the 61 kDa protein was identified as GBSS-I, all other proteins were identified as soluble starch synthase (Takaoka *et al.*, 1997). While immunoblotting of the endosperm starch granule associated proteins separated by SDS-PAGE with antibodies raised against buckwheat GBSS-I showed strong cross reactivity with the 59.7 kDa protein, there was no cross reactivity with the 56 kDa protein associated with the starch grains of European accessions/varieties. Antisera raised against the 59.7 kDa protein also failed to crossreact with the 53 kDa protein associated with the starch grains isolated from leaf tissues of buckwheat, thereby indicating absence of any serological homology between the 53 kDa protein and GBSS-I The antisera however, cross reacted with the 61 kDa GBSS-I of

maize and the 60 kDa GBSS-I of rice and wheat. These result indicate serological homology between the GBSS-I of wheat, maize, rice and buckwheat. GBSS-I, the key enzyme responsible for amylose synthesis, has been identified as a 56, 58 or a 60 kDa protein in maize (Echt and Schwartz, 1981; Shure *et al.*, 1983; Gibbon *et al.*, 2003), a 60 kDa protein in rice (Sano, 1984), 58 kDa protein in mungbean (Ko *et al.*, 2009), 68 kDa protein in grain amaranth (Konishi *et al.*, 1985), 59 kDa protein in pea (Dry *et al.*, 1992). Our results are in agreement with the known molecular mass of GBSS-I from other sources. The SDS-PAGE profile of buckwheat varieties/accessions both from Indian Himalayas and Slovenia did not show the presence of any high molecular weight bound protein, which were present in other plants like pea and wheat. This can be attributed to the extraction procedure which led to the complete loss of all the proteins present on the peripheral regions of the grains. Since GBSS-I is located in the core of the starch grains, it is retained during the extraction procedure.

2D-PAGE of the endosperm starch granule associated proteins resolved the fraction into 10 spots with *pI* ranging from 5.2 to 6.2. All the spots showed an apparent molecular mass of 59.7 kDa. On the other hand, 2D-PAGE of the leaf starch granule associated proteins resolved the fraction into 2 spots with *pI* ranging from 4.5 to 5.2. Imunoblotting of the 10 spots separated by 2D-PAGE with antisera raised against buckwheat GBSS-I identified two bands *viz.*, spot no. 3 (*pI* 5.4) and spot no 8 (*pI* 6.1) which cross reacted with antibodies raised against buckwheat. Similar results have been reported for *waxy* proteins of maize, rice and barley (Nakamura *et al.*, 1993, 1995; Taira *et al.*, 1995 and Chao *et al.*, 1989). While the protein was detected as a

single band of 60 kDa on 1D-PAGE, it resolved into 3 (barley) to 4 (maize, rice) isoforms, with *pI* ranging from 5.8 to 7.2 on 2D-PAGE. Bancel *et al.* (2010), have reported the presence of 150 coomassie blue stained spots out of which 10 spots were identified as charge variants of GBSS. It is possible that the other spots corresponded to soluble proteins associated with carbohydrate metabolism. The variation in the number of protein spots between starch grains from different sources can also be due to differences in extraction protocols. The protein corresponding to spot no.3 was subjected to in-gel trypsin digestion, followed by Mass Spectrometry and database search with MASCOT (www.matrixscience.com). Since the identification of peptide fragments relies on sequence information present in databases, this is a limitation for buckwheat,which has no sequence information on granule bound protein. Buckwheat proteins were thus, identified by comparing them with the rice genomes. However, one of the tryptic fragment having the amino acid sequence "FNAPLAHLIMAGADVLAVPSR" with a predicted *pI* of 8.34 showed similarity with GBSS-I protein of rice.

BLASTp analysis of the N-terminal amino acid sequence for 25 residues of the 59.7 kDa granule associated protein, identified it as granule bound starch synthase-I enzyme. Multiple alignment of the N-terminal sequence of GBSS-I protein isolated from buckwheat with amino acid sequences of similar proteins available in protein data banks clearly identified the conserved KTGGL motif in the sequence. Furukawa *et al.* (1993), have demonstrated the ubiquitous presence of KTGGL domain in all the GBSS-I proteins. This motif has been identified as the ADP/ADPglucose binding site

in the enzyme. Modelling based on threading of GBSSI onto the structure of glycogen phosphorylase has indicated that the N-terminal KTGGL motif of the protein may be close to the C-terminal KTGGL "look-alike" motif (Mason-Gamer *et al.*, 1998). It is interesting to note that while the residue immediately preceding and following this motif in soluble starch synthase is usually basic, there are no basic residues in the immediate vicinity of this motif in GBSS-I. Further, the region from the N-terminus to the 25th residue of the 59.7 kDa protein showed 94% homology with GBSS-I from *Hordeum vulgaris*, *Triticum* spp. and *Phaseolus vulgaris*. However, the percentage homology varied between 94% to 88% with GBSS from other cereal crops. Even though analysis of the sequence alignment revealed a clear diversification into monocotyledonous and dicotyledonous groups, the protein sequence from buckwheat showed similarities with GBSS-I from both the groups. The protein sequence of buckwheat has Valine as the 5th amino acid residue, as is also a case of the protein sequences of GBSS-I from monocots. However, majority of the dicots GBSS-I sequence had Isolucine at this position. Again, the protein sequence of buckwheat also showed similarities with protein sequences from dicots in having Valine as the 11th amino acid residue. GBSS-I protein sequence from monocots have methionine at this same position. It was also observed that the protein sequence of pea has Serine as the 3rd amino acid residue, while the rest of the monocots and dicots including buckwheat, had Asparagine at the same position. Potato sequence also differed from other monocots and dicots in having Glycine as its 12th amino acid residue. GBSS-I protein sequence from monocots and dicots have Alanine at this same position. The structural

differences may imply differences in catalytic activity of the enzyme. Sequence analysis of buckwheat GBSS-I indicates similarities with both cereal as well as dicot GBSS-I sequences. It is possible that the protein from common buckwheat might belong to a catalytically distinct subclass and hence, a possible candidate for altering amylose biosynthesis in both dicots as well as monocots. These results strongly indicate that the 59 kDa protein isolated from starch grains from endosperm tissues of common buckwheat is a GBSS-I class enzyme and hence, an isoforms of the *waxy* protein. This is the first report on the identification of GBSS-I in common buckwheat.

Phylogenetic analysis of the N-terminal amino acid sequence of the GBSS type protein from common buckwheat reported here, revealed a clear diversification into monocotyledonous and dicotyledonous groups. Within the monocots, the sequences could be segregated into two groups. While one of the two groups (Group-I) was dominated by rice (*Oryza* spp.) and maize (*Zea mays*), the other group (Group-II) predominantly comprised sequences from *Triticum* spp, *Hordeum* spp, *Secale cereale*, *Elymus scaber* and *Aegilops speltoides*. The dicotyledons on the other hand, were represented in three subgroups *viz.*, Group-III, Group-IV and Group-V. Group-III comprised of sequences from *Fagopyrum esculentum*, *Nelumbo nucifera*, *Astragalus membranaceus* and *Amaranthus cruentus*. Group-IV was represented by sequences from *Pisum sativum*, *Phaseolus vulgaris* and *Manihot esculenta* and Group V by *Ipomoea batatas*. These results are in conformity with the observations of Edwards *et al.* (2002), who have reported a clear division of GBSS-I proteins into those belonging to monocots and those to dicots.

CHAPTER V

AMPLIFICATION AND SEQUENCING ANALYSIS OF *WAXY* LOCUS

EXPERIMENTAL

Grains of common buckwheat *(Fagopyrum esculentum* Moench) were procured from the North Eastern Regional Station of National Bureau of Plant Genetics Resources, Shillong and multiplied in the experimental garden of Botany Department, North Eastern Hill University, Shillong.

The investigation carried out under the present study focused on amplification of the *waxy* locus by polymerase chain reaction using oligonucleotide primers designed from the conserved regions of GBSS gene sequences available in the Genbank database. For achieving the identified target, healthy grains of uniform size were germinated in dark in a seed germinator at 27°C and 85% R.H. The germinating seedlings were maintained in a plant growth chamber (Daihan Labtech Co.Ltd, Korea) for 14 days till the first leaves emerged fully. Genomic DNA was isolated from 14 days old etiolated seedlings of common buckwheat following the CTAB buffer extraction protocol. PCR

amplifications were carried out with primers designed from conserved regions of GBSS gene sequences in the database. The primers used for amplification of GBSS-I protein genes and the primer combinations along with their annealing temperatures are listed in Table 5.1 a, b.

RESULTS

The modified CTAB protocol used in the present investigation for isolation of genomic DNA yielded fairly good quality DNA from the etiolated seedlings of common buckwheat. Under UV light, the isolated DNA was detected on the agarose gel as a single fluorescent band having apparent molecular mass of 21 Kb (Fig. 5.1). There was no fluorescence due to RNA in the electrophoresed DNA sample. The electrophoretic profile of the isolated DNA did not reveal any degradation or shearing of the isolated DNA. The concentration of DNA in the preparation was quantified by visual observation of the ethidium bromide stained DNA in the agarose gel on a UV-transilluminator and comparison of the fluorescence of the band with fluorescence of the band representing a known amount of λ DNA electrophoresed alongside the isolated DNA, using the KODAK ID image analysis software. The concentration of DNA was also measured as a function of absorbance shown by the sample at 260 nm. The yield of DNA ranged between 20-25 μg DNA gm^{-1} tissues. Purity of the isolated DNA was assessed by measurement of absorbance at 260 and 280 nm and calculation of the ratio of absorbance at the two wavelengths. The ratio of absorbance at 260 and 280 nm (A_{260}/A_{280}) for the DNA samples isolated during the present investigation, ranged between 1.7 to 1.9 indicating that the isolated DNA was fairly pure for use in the investigation. Genomic

Table 5.1(a): List of Primers used for PCR amplification of GBSS-I protein genes

Primers	Sequences	Derived from
AFC	5'-TCGTGCTTCGTCGGCGCCGAGATGG-3'	Blast analysis
AR2	5'-CCGCGCTTGTAGCAGTGGAAGTACC-3'	Blast analysis
NDF1	5'-AGATCAAGGTTGCAGACAGG-3'	Blast analysis
NDR1	5'- GAGTGCTGCCTGCAGTAG-3'	Blast analysis
NDF2	5-AAACTGGTGGACTAGGTG-3'	Blast analysis
NDR2	5-CATTTCCTTTGCAAGAGGAG-3'	Blast analysis
NDF3	5'-AAGTTCCTCGCCGTCAACTA-3'	Blast analysis
NDR3	5'-AACTGGCAAGGAGGACTGAA-3'	Blast analysis
NDF4	5'-GCAAAACTGGTGGACTAGGTG-3'	Blast analysis
NDR4	5'-CCTTCTTTCACAGTGTCAAC-3'	Blast analysis
NDF5	5'- CCCTGGAGCAAGACCGGCGGCCT-3'	Blast analysis
NDF6	5'- CCTTGGAGCAAAACTGGTGGACT-3'	Blast analysis
NDF7	5'- CCCTGGAGCAAAACTGGTGGACT-3'	Blast analysis
NDR6	5'- TCCCAGCCTTCATCCAGTTGAT-3'	Blast analysis
NDR7	5'- TCCCGGCCTTCATCCAATTGAT-3'	Blast analysis
NDF8	5- GTGGATCACCCTATGTTCCTTGA-3'	Blast analysis
NDF9	5'- TGGGGGAAAACTGCATCGAAAA-3'	Blast analysis
NDF10	5'-TGACCACCCACCCACCCTTCT-3'	Blast analysis
NDR11	5'- GCGGCGACGTTCTCCTTGGCGA-3'	Blast analysis
NDF11	5'-CTCAAGAGCAACTACCAGTCCA-3'	Blast analysis

Table 5.1(b): List of primer combinations and PCR annealing temperatures used for the amplification of GBSS-I genes from buckwheat genomic DNA

Primer Combination	**Anneling Temp (°C)**	**Amplicon Size (Kb)**
NDF2/NDR2	54	0.7
NDF3/NDR3	57	0.7
AFC/AR2	56	No Amplification
NDF1/NDR1	55	No Amplification
NDF11/NDR11	61	No Amplification
NDF4/ NDR4	55	No Amplification
NDF5/NDR6	63	No Amplification
NDF5/ NDR7	63	No Amplification
NDF6/ NDR6	60	No Amplification
NDF6/NDR7	62	No Amplification
NDF7/NDR6	67	1.2
NDF7/ NDR7	64	No Amplification
NDF10/ NDR6	62	No Amplification
NDF10/ NDR7	63	No Amplification
NDF8/ NDR6	63	0.9
NDF9/ NDR6	60	No Amplification
NDF9/ NDR7	62	No Amplification

```
L.mespiloides-AF500438      AGTGAA-TTGTTTGAGTTTGATCTATGGATTTCA-------CATTTCATC 42
M.germanica-AF500444        GGTGTA-TTTTTTAAGTTTGATCTGTGAGTTTCA-------CATTTCGTC 42
F.alnus-AF500430            AGTGGACTTAATTAGGTCTCCTCATTACTTATTAACAGAGACTTACCATC 50
N.nucifera-FJ602702         -----ACTTTATAGAATTACTTAAGCACGTAT------------TTAGC 32
F.esculentum                --------CCCTGGAGCAAAACTGGTGGACTGGGGGACGTTTTAGCTGCT 42
                                    *

L.mespiloides-AF500438      TCTTTGGGTTTAATTGAATTG----TTGTGCCTGATTCAT-GTCTAGGCT 87
M.germanica-AF500444        TCACTGGATTTAATTGAACTGATTTTTGTACCTGATTTAT-GTCTAGGCT 91
F.alnus-AF500430            TTTTTTGTAGTATGGAACTTAACGTTTATGCCATGTCTATTGTCTAGGCT 100
N.nucifera-FJ602702         TCGCTGAAAT--AGCAATTTACTAACTCTGTATGATCCTT-TTCCAGGCC 79
F.esculentum                CTGCCCCATTTACATGTCTTGATACTTGAATCACCTGCCCTGGCT--GCC 90
                                          *     *       *      *   **

L.mespiloides-AF500438      AACGGGCACCGTGTTATGACCGTGTCTCCACGCTACGACCAGTACAAAGA 137
M.germanica-AF500444        AACGGGCACCGTGTTATGACGGTGTCTCCACGCTACGACCAGTACAAAGA 141
F.alnus-AF500430            AATGGGCACCGCGTTATGACAGTGTCTCCCCGCTATGACCAGTACAAAGA 150
N.nucifera-FJ602702         AACGGCCACCGTGTTATGACAGTAGCTCCACGTTATGACCAGTACAAAGA 129
F.esculentum                CGTGGACATAGGGTTATGACAGTTGCTCCTCGTTATGATCAGTATAAAGA 140
                               ** **  * ******** **  **** ** ** ** ***** *****

L.mespiloides-AF500438      TGCATGGGACACTAGTGTACTAGTTGAGGTAACATTTGCAGATTCTATAT 187
M.germanica-AF500444        TGCATGGGACACTAGTGTACAAGTTGAGGTAAAATTTGCAGATGCTCTAT 191
F.alnus-AF500430            CGCATGGGACACTGAGGTGCAAGTTGAGGTGTGTGATGCAGTTTATTCAT 200
N.nucifera-FJ602702         TGCATGGGATACAAGTGTTCTAGTTGAGGTATATTCTAGCAATCATGTTA 179
F.esculentum                TGGATGGGATACTAATGTACTAGTTCAGGTAAATT-TGGATTCATTGTAT 189
                             * ****** **    ** * **** ****      *        *

L.mespiloides-AF500438      GCTTCTTATACTTGAACAATCTTACAATGATTCGGAGATGCTTTTGCAGT 237
M.germanica-AF500444        GCTTCTTATACTTGAACAATATTACATTGTTTCGGAGATTCTTTCGCAGT 241
F.alnus-AF500430            GTCTCTA------GAGCAATTTT----TGGCTATGAGATTCTTT---AAT 237
N.nucifera-FJ602702         ATTTTCT------GAATAAA--------GCTTTCAATTTTCTTTTCCAGG 215
F.esculentum                AATTTTTAT-----GTTATCTAAGCAATATATTAATGCTGTTCAACTGGC 234
                               *            *             *      *  *

L.mespiloides-AF500438      TTCATGTTTCAGTGAGTAATGTATGCTTTGGTGGAAGTCCTCTTTGACTG 287
M.germanica-AF500444        TTCATGTTTCAGTGAGTAATGTATACTTTGGTGGAAATTCTCTTTGAATG 291
F.alnus-AF500430            TTCCCATTA--------AATGGA----------------------GGATG 257
N.nucifera-FJ602702         TGTTCATTCACCTG---AAGGGATTCATGCATTGGGATA-TGTTTGCTAA 261
F.esculentum                TTTTTATT---------AAGCTGT--------------------TTTTGG 255
                            *     **         **

L.mespiloides-AF500438      ATCAGAAGTTTTATGTATGCAGATAGAAGTTGGTGGTAGGGTTGAAACTG 337
M.germanica-AF500444        ATCAAAAGTTGTATGTATACAGATAGAAGTTGGTGGTAGGGTTGAAACTG 341
F.alnus-AF500430            ATTGAGATTTCTA------CAGATTAAAGTTGGTGATAGAATTGAAACCG 301
N.nucifera-FJ602702         AGTATACCTTTTA-ATATGCAGATAAAAGTTGGAGATAGAATTGAAACTG 310
F.esculentum                TGTTAAATTTATATA-ATGCAGATACAAGTTGGGGAAAGAGTTGAGACTG 304
                                   ** **      *****  ******* *  **  **** ** *]

L.mespiloides-AF500438      TTCGCTTCTTCCACTGCTACAAACGCGGAGTTGATCGTGTCTTTGTGGAT 387
M.germanica-AF500444        TTCGCTTCTTTCACTGCTACAAACGCGGAGTTGATCGTGTCTTTGTGGAT 391
F.alnus-AF500430            TTCGCTTCTTCCACTGCTACAGACGAGGAGTTGATCGTGTTTTTGTGGAT 351
N.nucifera-FJ602702         TTCGCTTCTTCCACTGCTACAAACGTGGAGTTGATCGTGTTTTTGTTGAT 360
F.esculentum                TTCGGTTCTTTCACTGCTACAAAAGAGGAGTTGACCGGGTTTTCGTGGAT 354
                            **** ***** ********** * * ******** ** ** ** ** ***

L.mespiloides-AF500438      CACCCACTGTTCCTTGAGAAGGTATGAAA--TTTATTTTCGT-TGTCTAA 434
M.germanica-AF500444        CACCCACTGTTCCTTGAGAAGGTATGAAA--TTTATTTCCAAATGTCTAA 439
F.alnus-AF500430            CACCCAATGTTCCTTGAGAAGGTTTGAGTGTTTTATTTATGAAACACTTG 401
N.nucifera-FJ602702         CACCCCATGTTCCTTGAGAAGGTTATTCA--TTGACCCAGGGCCATATAA 408
F.esculentum                CACCCTATGTTCCTTGAGAAGGTGACTGG----ACCATTAGGATATTTGG 400
                            *****  ****************                        *

L.mespiloides-AF500438      ---GGCTATTTGCTG-AAAGTATC------TAGGATATTATATGAATATG 474
M.germanica-AF500444        A--GGCTATTTGCTG-AAAGTTTC------CAGGATATTGTACGAATATG 480
F.alnus-AF500430            TTTGGCTCTGTGGCATAAAGTTTGAAATTGTAGAATATTGTATTTGTATG 451
N.nucifera-FJ602702         ATTGAGCTTCTTTATTGTAACTTCA-----CATGTTAACTGTTCAATAGT 453
F.esculentum                AG-GAGCATATGATT-AAAAT----------AGTAATCTTTAGGAGAG 436
                               *    * *       *               **
```

Fig. 5.7 ClustalW (1.81) multi-alignment of the 1116 bp nucleotide sequences for the 1.2 kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF7/NDR6 with nucleotide sequences of GBSS genes available in the database. '*' denotes conserved residues.

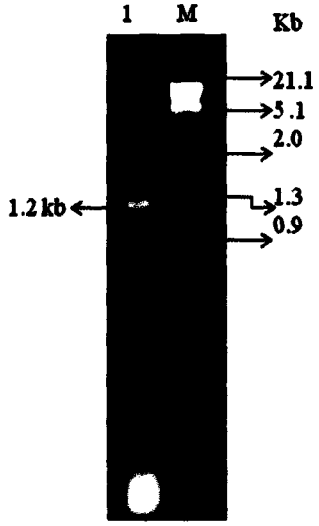


Fig. 5.5

```
1       CCCTGGAGCA AAACTGGTGG ACTGGGGGAC GTTTTAGCTG CTCTGCCCCA
51      TTTACATGTC TTGATACTTG AATCACCTGC CCTGGCTGCC CGT[redacted]ACATA
101     GGGTTATGAC AGTTGCTCCT CGTTATGATC AGTATAAAGA TGGATGGGAT
151     ACTAATGTAC TAGTTC[redacted]GT AAATTTGGAT TCATTGTATA ATTTTTATGT
201     TATCTAAGCA ATATATTAAT GCTGTTCAAC TGGCTTTTTA TTAAGCTGTT
251     TTTGGTGTTA AATTTATATA ATG[redacted]GATAC AAGTTGGGGA AAGAGTTGAG
301     ACTGTTCGGT TCTTTCACTG CTACAAAAGA GGAGTTGACC GGGTTTTCGT
351     GGATCACCCT ATGTTCCTTG AGAAG[redacted]GAC TGGACCATTA GGATATTTGG
401     AGGAGCATAT GATTAAAATA GTAATCTTTA GGAGAGTTAC TACTACCACC
451     TTCGCCTCGG GAACTCAAAT CTGTATTGTT CGGGGGGGGG TAAATACAAA
501     TCTAAGAGCC GTGGGCCTAA CGCGGAAACT GATTATGATG ACCACCCACC
551     CACCCTTCTT CCTCTTGTCA ATGGTGGTGA ATATAAAATA ACAAAGGAAT
601     ATGCAGCTTC CGCTTTGGCA CTGCACATCA TAAGCAGATC TTGTATCGTT
651     GATCGGTACA CACTCCGTCT GTGGTTTTGA TCCTCTAGCC CTCATCTTTG
701     TTTTTTTGTT GGCCCGGCCG AATTCCCTTA TTCTCTTTTT CTTTGGAAAT
751     TCTGACGAGC ACACAGAAGA AGAAACAAAG AATTGTTTTT TGGTGTTAAT
801     CAACTTGTCT CTTGCCCTCA G[redacted]AGGGCGC AAGGGTGTTA TGGCGTATAT
851     ACCCCACACC ACAGGAAAAA CAGCCCCTTG GGCGCCCGC[redacted]GTAAGTAAC
901     TATAGCGGCT CAGGGTTCGA CAGAAAATCT TTTCACCCTT AAAGTCTCCT
951     TCCGCAAAAA GAG[redacted]AGAAC GACCACCGTA TACCGAAAGA GAGAAGGAGA
1001    GGAGGAGGGC TGCCACCCTC CCGCTGAGCT GCGCCTAACT GGGCGCGGAG
1051    TAAGCACATA CAACCTAAAC AACCGTCTTT CTCGCCTCCG TCTGTGGACC
1101    GTAGGTGTAT [redacted]GTCT
```

Fig. 5.6

Fig. 5.5 Electrophoresis profile of amplified fragment with buckwheat gDNA as template and primer pair NDF7-NDR6, M= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.6 Nucleotide sequence of 1116 bases for the 1.2 Kb fragment amplified with buckwheat gDNA as template and primer pair NDF7-NDR6.The exons identified in the sequence are highlighted in green and their start and end positions are indicated with forward and backward arrows.

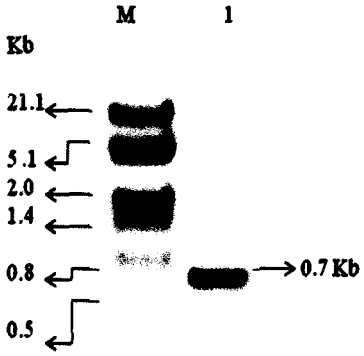


Fig. 5.4a

1	CTGATGTCTG TCGTGGTGAC GGTCCTTCCT GGTGGCGTAC GATTCGTGGC
51	CGGCTGCCGT CTGGGAGCCT ATGAGGAACA TTGTTAGGGA TATCTACTGC
101	CCTGTTTCCC TGAGGAGGGA AAGATTCCCG GATTTTTTTG GCGCTCCGTT
151	TTAATGTGAC CCGATTGCTG CAGTTAAAGC ACTGTCCACC TCCACCGGGG
201	GTCTGAACTG TAAAGTGGTA CCTGGACCTC CCCTGAGTCT GTCCTCTTCC
251	ATAACCTGCT ACTGCTGCAC TGCTACCACC ATGTCCTGTC CTCTGCCGCT
301	TATCCGGCCC ACTATACTGA GCTGCTTTCC TAGAATCTGG CCTATCTCTC
351	TCATAACAAG GCCCTATCCT GATAATCAAC AAGGTCTCAC ACAATCCATG
401	ATCGGTACTC TCTCTCACCT CTGAAAATGG TACGTTAAGG AGTTGGGGTC
451	AATTGCTAAT GTTGCCTAAA AATTTATGTG AAAGGGCGGG AGTAGTGGAA
501	GAAAGGAGTC TGGTACGGGG GGGTTTTCGC AGCACTCTCA CCATTTAATC
551	CATGTCTGGG CAGAAGGAGA ATCTCCAAGG AACACCGCCA ATCATAAAGA
601	ACAAAACCAA AACCCGCGAC AACCAAGAGT AACAATCAAA CGTAAATAGC
661	GAAACAAAAT CAATGCATGG TGCGAGACTA CAAGAAGGGC AGACATCACA
701	ATAG

Fig. 5.4b

Fig. 5.4a Electrophoresis profile of amplified fragment with buckwheat gDNA as template and primer pair NDF3-NDR3, M= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.4b Nucleotide sequence of 704 bases for the 0.7 Kb fragment amplified with buckwheat gDNA as template and primer pair NDF3-NDR3.

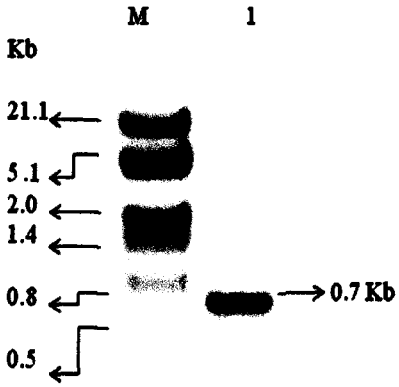


Fig. 5.3a

1	CATTCTTTCT TATTAATATA AGGGTTCGGA GACCGTCAGG CCCCACTCAC
51	AAACTAGCTG CCCTCTCACC TTAAGCATTT CTTCTATTTT GAAATTGTCC
101	ACTAGCCCCA TTCATTATTG GAATGCTCAA GCTCAATATC ATATTTTTTG
151	TGTTCTATCA ACGAGGCGAT CGAAGTGAGA AAGCACTGAG ACAGATAGAA
201	GTGTTCCGTA TAGGTTCGTA TATATACTGT GGTGCTATAT GATCTTTAGC
251	TATAGGTCTT GCGCAGTTAG TGTGTTTGCT AAAGTAAATA TACGAAGCAC
301	GGAGTAAGAG GAAGTGGTAA CGGTCGATGG ATGTTCTTAA GTATTTGCTG
351	ACTGAAGCCT TACATCAAGG TGACAGGATT CGAACCTATG GCCCTCTGTA
401	CCCAAAACAG ATGCGCTGAC CAGACTGCGC TACACCTCGT CTCACCCTCC
451	GAAGCATATA GATCTACCCG ATCGATGAAG ACTCGAACCG AACTCGCTCA
501	TCCTTTTTGT CAAGAAGGCC CCTCTTGAGG GGGAACCGCT TTTTTTTGGA
551	ATCTGCCTCC AGGGCCAAAA AGCCGTATAG TGCAGGAGCG CTCACATTTT
601	TTGGCTTCCT CTTTGACTTG CTTGAATTTC CAAAACCCGG CTCGGCTCCC
661	GGCTACGGTT TGGAACCCTT CGCCTTGCTT AAAAGTGGAT

Fig. 5.3b

Fig. 5.3a Electrophoresis profile of amplified fragment with buckwheat gDNA as template and primer pair NDF2-NDR2, M= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.3b Nucleotide sequence of 690 bases for the 0.7 Kb fragment amplified with buckwheat gDNA as template and primer pair NDF2-NDR2.

the primer combinations NDF2-NDR2, NDF3-NDR3, NDF7-NDR6 and NDF8-NDR6 (Table 5.1 a,b). PCR amplification with oligonucleotide primer pairs NDF2-NDR2 and NDF3-NDR3 amplified a DNA fragment showing an apparent molecular mass of 0.7 Kb on 0.8% agarose gel (Fig. 5.3 a, b). BLASTn analysis of the nucleotide sequences of the two amplicons did not identify any of the two sequences with GBSS gene family (Fig. 5.4 a, b). PCR amplification with oligonucleotide primer pair NDF7-NDR6 and amplification reaction comprising hot start at 94°C for 5 mins., 35 amplifications cycles comprising of denaturation at 94°C for 1 min., annealing at 67°C for 1 min. and chain extension at 72°C for 1 min. followed by one reaction of chain elongation at 72°C for 10 mins., amplified a DNA fragment showing an apparent molecular mass of 1.2 Kb on 0.8% agarose gel (Fig. 5.5). The nucleotide sequence comprising of 1,116 bases for the 1.2 Kb amplicon is presented in Fig. 5.6. BLASTn analysis of the sequence identified the sequence with genes coding for granule bound starch synthases (GBSS). The sequence showed a maximum of 85% homology with GBSS of *Hibiscus dasycalyx* (acc. no. AY341418), *Zanthoxylum schinifolium* (acc. no. AB288186), *Mespilus germanica* (acc. no. AF500444) and *Frangula alnus* (acc. no. AF500430). The sequence showed 84%, 83% and 72% homology with GBSS gene of *Lindleya mespiloides* (acc. no. AF500438), *Vauquelinia californica* (acc. no. AF285998) and *Nelumbo nucifera* (acc. no. FJ602702) respectively. ClustalW multiple alignment of the sequence with nucleotide sequences of GBSS genes from other plants is presented in Fig. 5.7. Using an alignment that permitted maximum homology, the sequence showed two highly conserved regions. While one of the conserved regions existed between P'_{94} and P'_{168},

DNA isolated from etiolated seedlings of common buckwheat was digested with *Eco*R1, *Hind*III and *Nco*I and size fractionated on 0.8% agarose gel. The electrophoretic profile of *EcoR*1 digested DNA revealed a uniform smear ranging in size from 0.5 Kb to 14.7 Kb with three prominent bands showing apparent molecular masses of 3.7 Kb, 1.3 Kb and 1.15 Kb. *Hind*III digested DNA also resolved in the form of a uniform smear ranging in molecular mass from 0.5 Kb to 20.0 Kb. The electrophoretic profile of *Hind* III digested DNA showed four bands having molecular masses of 3.7 Kb, 1.04 Kb, 1.15 Kb and 0.534 Kb. The electrophoretic profile of *Nco*I digested DNA revealed that the enzyme could only partially digest the DNA. Even though *Nco*I could only partially digest buckwheat genomic DNA, the electrophoretic profile of *Nco*I digested DNA revealed six bands having molecular masses of 2.46 Kb, 1.87 Kb, 1.108 Kb, 0.983 Kb, 0.592 Kb and 0.488 Kb. (Fig. 5.2).

Polymerase chain reaction comprising a hot start at 94°C for 5 mins. and 35 amplification cycles comprising of denaturation at 94°C for 1 min., annealing at 64°C for 1 min. and chain extension at 72°C for 1 min. followed by final reaction of chain elongation at 72°C for 10 mins. were carried out with buckwheat genomic DNA as the template and combinations of oligonucleotide primer pairs designed to amplify the sequence of GBSS gene from the buckwheat genomic DNA. Out of the total 18 primer pairs used, 14 pairs *viz.*, NDF1-NDR1, AFC-AR2, NDF4-NDR4, NDF5-NDR6, NDF5-NDR7, NDF6-NDR6, NDF6-NDR7, NDF7-NDR7, NDF8-NDR8, NDF9-NDR6, NDF9-NDR7, NDF10-NDR6, NDF11-NDR7 and NDF11-NDR11 failed to amplify any sequence from the template DNA. Amplification of gDNA was however, achieved with

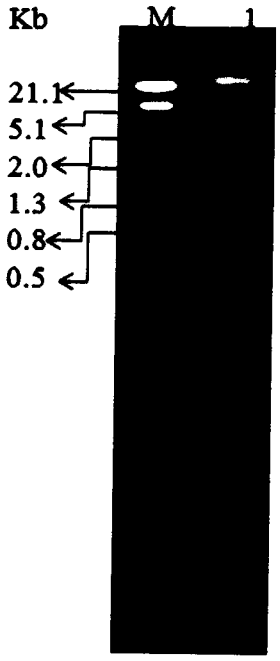


Fig. 5.1

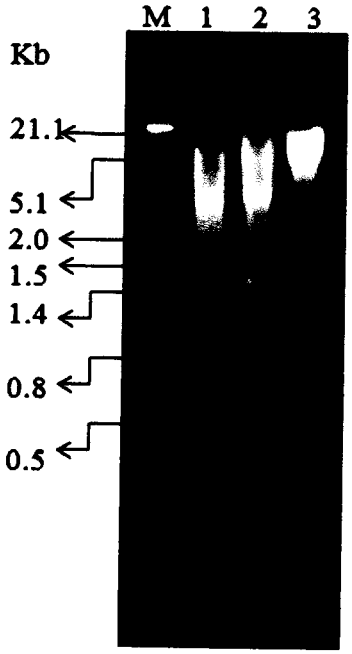
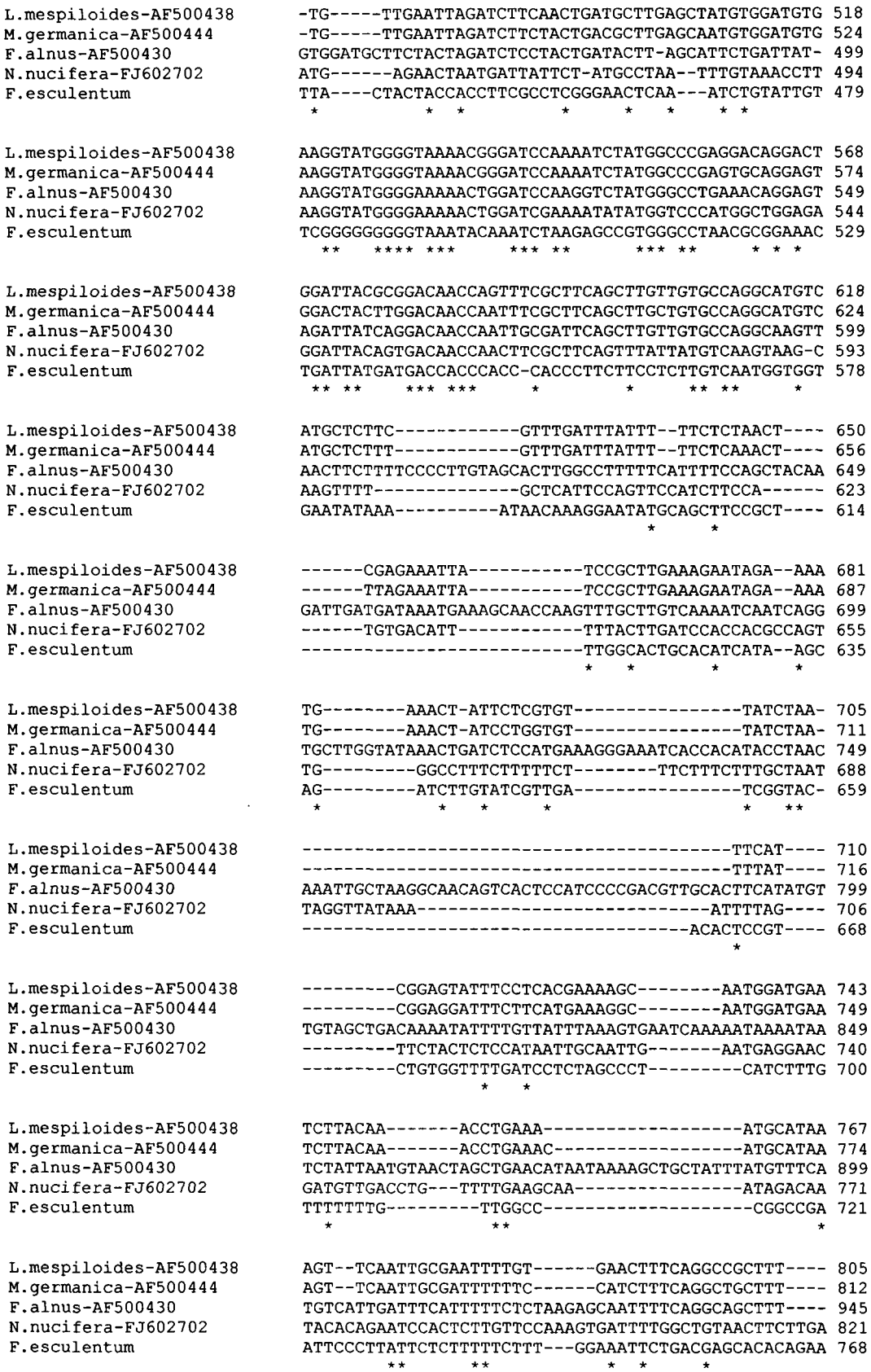


Fig. 5.2

Fig. 5.1 Total genomic DNA isolated from etiolated seedlings of buckwheat (*Fagopyrum esculentum* Moench.)
L1= isolated genomic DNA; M= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.2 Restriction digestion profile of genomic DNA isolated from etiolated seedlings of buckwheat (*Fagopyrum esculentum* Moench.)
(L1)=gDNA digested with *Eco*RI, (L2)= gDNA digested with *Hind*III, (L3)=gDNA digested with *Nco*I, (M)= *Eco*RI/ *Hind*III digested λ DNA.


L.mespiloides-AF500438 -TG-----TTGAATTAGATCTTCAACTGATGCTTGAGCTATGTGGATGTG 518
M.germanica-AF500444 -TG-----TTGAATTAGATCTTCTACTGACGCTTGAGCAATGTGGATGTG 524
F.alnus-AF500430 GTGGATGCTTCTACTAGATCTCCTACTGATACTT-AGCATTCTGATTAT- 499
N.nucifera-FJ602702 ATG------AGAACTAATGATTATTCT-ATGCCTAA--TTTGTAAACCTT 494
F.esculentum TTA----CTACTACCACCTTCGCCTCGGGAACTCAA---ATCTGTATTGT 479
L.mespiloides-AF500438 AAGGTATGGGGTAAAACGGGATCCAAAATCTATGGCCCGAGGACAGGACT 568
M.germanica-AF500444 AAGGTATGGGGTAAAACGGGATCCAAAATCTATGGCCCGAGTGCAGGAGT 574
F.alnus-AF500430 AAGGTATGGGGAAAAACTGGATCCAAGGTCTATGGGCCTGAAACAGGAGT 549
N.nucifera-FJ602702 AAGGTATGGGGAAAAACTGGATCGAAAATATATGGTCCCATGGCTGGAGA 544
F.esculentum TCGGGGGGGGGTAAATACAAATCTAAGAGCCGTGGGCCTAACGCGGAAAC 529
L.mespiloides-AF500438 GGATTACGCGGACAACCAGTTTCGCTTCAGCTTGTTGTGCCAGGCATGTC 618
M.germanica-AF500444 GGACTACTTGGACAACCAATTTCGCTTCAGCTTGCTGTGCCAGGCATGTC 624
F.alnus-AF500430 AGATTATCAGGACAACCAATTGCGATTCAGCTTGTTGTGCCAGGCAAGTT 599
N.nucifera-FJ602702 GGATTACAGTGACAACCAACTTCGCTTCAGTTTATTATGTCAAGTAAG-C 593
F.esculentum TGATTATGATGACCACCCACC-CACCCTTCTTCCTCTTGTCAATGGTGGT 578
L.mespiloides-AF500438 ATGCTCTTC------------GTTTGATTTATTT--TTCTCTAACT---- 650
M.germanica-AF500444 ATGCTCTTT------------GTTTGATTTATTT--TTCTCAAACT---- 656
F.alnus-AF500430 AACTTCTTTTCCCCTTGTAGCACTTGGCCTTTTTCATTTTCCAGCTACAA 649
N.nucifera-FJ602702 AAGTTTT--------------GCTCATTCCAGTTCCATCTTCCA------ 623
F.esculentum GAATATAAA----------ATAACAAAGGAATATGCAGCTTCCGCT---- 614
L.mespiloides-AF500438 ------CGAGAAATTA-----------TCCGCTTGAAAGAATAGA--AAA 681
M.germanica-AF500444 ------TTAGAAATTA-----------TCCGCTTGAAAGAATAGA--AAA 687
F.alnus-AF500430 GATTGATGATAAATGAAAGCAACCAAGTTTGCTTGTCAAAATCAATCAGG 699
N.nucifera-FJ602702 ------TGTGACATT------------TTTACTTGATCCACCACGCCAGT 655
F.esculentum ----------TTGGCACTGCACATCATA--AGC 635
L.mespiloides-AF500438 TG--------AAACT-ATTCTCGTGT----------TATCTAA- 705
M.germanica-AF500444 TG--------AAACT-ATCCTGGTGT----------TATCTAA- 711
F.alnus-AF500430 TGCTTGGTATAAACTGATCTCCATGAAAGGGAAATCACCACATACCTAAC 749
N.nucifera-FJ602702 TG---------GGCCTTTCTTTTTCT--------TTCTTTCTTTGCTAAT 688
F.esculentum AG---------ATCTTGTATCGTTGA----------TCGGTAC- 659
L.mespiloides-AF500438 ----------TTCAT---- 710
M.germanica-AF500444 ----------TTTAT---- 716
F.alnus-AF500430 AAATTGCTAAGGCAACAGTCACTCCATCCCCGACGTTGCACTTCATATGT 799
N.nucifera-FJ602702 TAGGTTATAAA----------ATTTTAG---- 706
F.esculentum ----------ACACTCCGT---- 668
L.mespiloides-AF500438 ---------CGGAGTATTTCCTCACGAAAAGC--------AATGGATGAA 743
M.germanica-AF500444 ---------CGGAGGATTTCTTCATGAAAGGC--------AATGGATGAA 749
F.alnus-AF500430 TGTAGCTGACAAAATATTTTGTTATTTAAAGTGAATCAAAAATAAAATAA 849
N.nucifera-FJ602702 ---------TTCTACTCTCCATAATTGCAATTG-------AATGAGGAAC 740
F.esculentum ---------CTGTGGTTTTGATCCTCTAGCCCT---------CATCTTTG 700
L.mespiloides-AF500438 TCTTACAA-------ACCTGAAA----------ATGCATAA 767
M.germanica-AF500444 TCTTACAA-------ACCTGAAAC----------ATGCATAA 774
F.alnus-AF500430 TCTATTAATGTAACTAGCTGAACATAATAAAAGCTGCTATTTATGTTTCA 899
N.nucifera-FJ602702 GATGTTGACCTG---TTTTGAAGCAA----------ATAGACAA 771
F.esculentum TTTTTTTG---------TTGGCC----------CGGCCGA 721
L.mespiloides-AF500438 AGT--TCAATTGCGAATTTTGT------GAACTTTCAGGCCGCTTT---- 805
M.germanica-AF500444 AGT--TCAATTGCGATTTTTTC------CATCTTTCAGGCTGCTTT---- 812
F.alnus-AF500430 TGTCATTGATTTCATTTTTCTCTAAGAGCAATTTTCAGGCAGCTTT---- 945
N.nucifera-FJ602702 TACACAGAATCCACTCTTGTTCCAAAGTGATTTTGGCTGTAACTTCTTGA 821
F.esculentum ATTCCCTTATTCTCTTTTTCTTT---GGAAATTCTGACGAGCACACAGAA 768

```
L.mespiloides-AF500438      GGAGGCACCAAGGGTTCTTAATTTGAACAGCAACAAATACTTCTCTGGAC 855
M.germanica-AF500444        GGAGGCACCAAGGGTTCTTAATTTGAACAGCAACAAATACTTCTCTGGAC 862
F.alnus-AF500430            GGAGGCACCAAGGGTTCTCAATTTGAACAGCAATAAATTTTTCTCAGGAC 995
N.nucifera-FJ602702         AAATGCTTCAAGACTAACGAATTC--TTTGAAACTTATTACTTCCTGAAC 869
F.esculentum                GAAGAAACAAAGAATTGTT--TTTTGGTGTTAATCAACTTGTCTCTTGCC 816
                              *      ***  *      **        **   *    * *   *

L.mespiloides-AF500438      CATATGGTGTGCATCTTTCTCTTTTTCTCAAAAACTTTTTCTTTCTAAAA 905
M.germanica-AF500444        CATATGGTGTGCATCTTTCTCTTTTCCTCAACAACTTTTTCTTAATAAAA 912
F.alnus-AF500430            CTTATGGTAAGGACCTTTCTCTGTT--------ACTTGTATT-------- 1029
N.nucifera-FJ602702         TGT-CAATGTTGTTATTTTGTTGCAATATAAGCAGGCAAGTC-----AAT 913
F.esculentum                CTCAGGCAGGGCGC-------------------AAGGGTGTT-------- 839
                                                           *

L.mespiloides-AF500438      AGTATAATTTCTTCAA-GAAATAGAATAACTAATGATGTATACCA-TGGG 953
M.germanica-AF500444        AGTACAATTTCTTCAAAGAAATAGAATAACTAACGATGTATATTAATCGG 962
F.alnus-AF500430            -GTAAAATTTTTGATATGAAGTA---TCACTAA-GGGATATGCT--TTGT 1072
N.nucifera-FJ602702         AGCACATTTGTTGCTTATGCATTGCTCAGCTGACACGGCTAGCTCATCAA 963
F.esculentum                -------ATGGCGTATATACCCCACACCACAGGAAAAACAGCCCCTTGGG 882
                                    *                    *                 *

L.mespiloides-AF500438      TACCAAAATGCAGGTGATGATGTTGTATTTATTGCTAATGATTGGCACAC 1003
M.germanica-AF500444        TACCACAATGCAGGTGATGATGTTGTATTCATTGCTAATGATTGGCACAC 1012
F.alnus-AF500430            TGCTTCAATGCAGGGGAGGATGTTGTATTTCTTACCAATGATTGGCACAC 1122
N.nucifera-FJ602702         -GCAACATT--AAATCCGTATCTCAAGTTGGTAAT--ACAAGGGACAAGA 1008
F.esculentum                CGCCCGCGGGTAAGTAACTATAGCG-GCTCAGGGT------TCGACAGAA 925
                              *       *       **       *              * **

L.mespiloides-AF500438      TGCTCTTCTTCCCTGCTACTTAAAGACCATCTACAGACCGAAGGGGATTT 1053
M.germanica-AF500444        TGCTCTTCTCCCCTGCTACTTAAACACCATCTACAAACCGAAGGGGATTT 1062
F.alnus-AF500430            AGCTCTTCTTCCATGCTACCTAAAGAGCATGTACAAATCTAGGGGGATTT 1172
N.nucifera-FJ602702         GTTGACCTTGCTCTAATAGCCACAGGC--TCTACATG-CTACCTTCATTT 1055
F.esculentum                AATCTTTTCACCCTTAAAGTC----TCCTTCCGCAAAAAGAGAAAGAACG 971
                                     *  *   *          *   **     *     *

L.mespiloides-AF500438      ACAAAAATGCGAAGGTAAGGTACCCTGCGTTTGAAGAGCACCATTTTAGA 1103
M.germanica-AF500444        ACAAAAATGCGAAGGTAAGGTACCCTGCGTTTGAAGAGTTCCATT--AGA 1110
F.alnus-AF500430            ACAAAAGTGCCAAGGTAATG-AGAGTAGGACTGAAGGAAACCATCAATAA 1221
N.nucifera-FJ602702         CCATAGGAGCAGCTGTGTAGTCCATCACCCAATATGGTCCTGATTATCCA 1105
F.esculentum                ACCACCGTATACCGAAAGAGAGAAGGAGAGGAGGAGGGCTGCCACCCTCC 1021
                             *                 *                *

L.mespiloides-AF500438      GTCTA--CAACA---ATTT-GTTTAAAGGAAAACTTAAACAATGCTGGAT 1147
M.germanica-AF500444        GTCTG--CAGCA---ATTTTGTTTAAATGAAAACTGAAACAATGCTGGAT 1155
F.alnus-AF500430            TTCTGGACAATATAGAATTATATTGGTTGAAATCTGAGACTTTGTTTGAT 1271
N.nucifera-FJ602702         CACCCTTTGGCT---TTTTATGCACAAAGTGGCCACTCTCAGTGTTGAAC 1152
F.esculentum                CGCTGAGCTGCGCCTAACTGGGCGCGGAGTAAGCACATACAAC-CTAAAC 1070
                              *               *         *    *     *     * *

L.mespiloides-AF500438      TAATGAATGGCAGGTTGCATTTTGCATCCACAACATAGCTTACCAGGGCA 1197
M.germanica-AF500444        TAATGAATGGCAGGTTGTGTTTTGCATCCACAACATAGCTTACCAGGGCA 1205
F.alnus-AF500430            GTATGA-TGGCAGGTTGCTTTTTGTATCCACAACATAGCCTACCAAGGCA 1320
N.nucifera-FJ602702         CCATGT------TCTCATGCTTGGCAGCC-CAAGATA--CCACCATTGCA 1193
F.esculentum                AACCGTCTTTCTCGCCTCCGTCTGTGGACCGTAGGTG---TATGAGTCT- 1116
                              *               * *    *   * *     * *
```

Fig. 5.7

the other existed between P'_{274} to P'_{377}. The conserved region between P'_{94} and P'_{168} comprised of 75 bases represented by the sequence "GGACATAGGGTTATGACAGTT GCTCCTCGTTATGATCAGTATAAAGATGGATGGGATACTAATGTACTAGTTC AG". The other conserved region comprised of 104 bases between P'_{274} and P'_{377}. This domain was represented by the sequence "CAGATACAAGTTGGGGAAAGAGTTGA GACTGTTCGGTTCTTTCACTGCTACAAAAGAGGAGTTGACCGGGTTTTCGTG GATCACCCTATGTTCCTTGAGAAGGT".

AUGUSTUS (version 2.4) (http://augustus.gobics.de/submission) identified four exons within the 1,116 bp nucleotide sequence amplified in the present study. While the 1st exon had a length of 168 bases and located between P'_1-P'_{168}, the 2nd exon had a length of 213 bases and located between P'_{277}-P'_{489}. While the 1st exon had the conserved domain of 74 bases comprising of the sequence "GGACATAGGGTTATGA CAGTTGCTCCTCGTT TGATCAGTATAAAGATGGATGGGATACTAATGTACTA GTTCAG" located between P'_{94} and P'_{168}, the 2nd exon had the conserved domain of 104 nucleotides comprising of the sequence "CAGATACAAGTTGGGGAAAGAGTT GAGACTGTTCGGTTCTTTCACTGCTACAAAAGAGGAGTTGACCGGGTTTTCG TGGATCACCCTATGTTCCTTGAGAAGGT" located between P'_{274} and P'_{377}. The 3rd and 4th exons had lengths of 70 and 149 bases respectively and were located beween P'_{822}-P'_{891} and P'_{964}-P'_{1112}. AUGUSTUS 2.4 identified the first three exons located between P'_1-P'_{168}, P'_{277}-P'_{489} and P'_{822}-P'_{891} as the internal exons in the sequence. The position of the four exons on the 1116bp nucleotide sequnce has been shown in Fig. 5.6. BLASTn analysis of the coding region of the 1,116 bp amplicon comprising of 604

bases showed a maximum of 90% homology with the coding region of GBSS-I of *Manihot esculenta* (acc. no. HM038439). The sequence showed 87%, 86%, 85%, 84%, 80% and 76% homology with the coding region GBSS-I gene of *Zanthoxylum schinifolium* (acc. no. AB288189), *Nelumbo nucifera* (acc. no. FJ602702), *Hibiscus dasycalyx* (acc. no. AY341419), *Photinia villosa* (acc. no. AF500451), *Ipomoea batatas* (acc. no. AB071604) and *Solanum tuberosum* (acc. no. EU403426). The sequence also showed 79% and 78% homology with the coding region of GBSS-II from *Sorghum bicolour* (acc. no. EF472254) and *Triticum aestivum* (acc. no. EF221764).

Analysis of the nucleotide sequence data for its G+C content revealed that the DNA fragment has a G+C content of 45%. The sequence showed specific A/T and G/C rich regions. While the segments between P'_{170} to P'_{220}, P'_{235} to P'_{253}, P'_{258} to P'_{272}, P'_{387} to P'_{398}, P'_{407} to P'_{430}, P'_{491} to P'_{504}, P'_{580} to P'_{595}, P'_{721} to P'_{751}, P'_{767} to P'_{791}, P'_{796} to P'_{806} and P'_{924} to P'_{933} were specifically rich in A/T content, those between P'_{15} to P'_{31}, P'_{74} to P'_{95}, P'_{331} to P'_{343}, P'_{508} to P'_{525}, P'_{446} to P'_{461}, P'_{814} to P'_{835}, P'_{880} to P'_{892} and P'_{1039} to P'_{1050} were rich in G/C content. The EMBOSS CpGplot software of European Bio Informatics Institute, detected a putative CpG island of 856 bases located between positions 310 to1116. The ratio of observed to expected CG composition in this segment of the sequence was > 0.60.

NCBIscan detected twelve palidromic structures in the nucleotide sequence of the 1.2 Kb amplicon. These incuded 'ACATGT' between P'_{54}-P'_{59}, 'TGATCA' between P'_{126}-P'_{131}, 'ACTAGT' between P'_{159}-P'_{164}, 'AAATTT' between P'_{171}-P'_{176}, 'AATATATT' between P'_{210}-P'_{217}, 'TAAATTTA' between P'_{259}-P'_{266}, 'CATATG'

between P'_{406}-P'_{411}, 'AGATCT' between P'_{536}-P'_{641}, 'CGGCCG' between P'_{715}-P'_{720}, 'TATATA' between P'_{846}-P'_{851}, 'GGCGCC' between P'_{881}-P'_{886} and GTATAC'between P'_{978} - P'_{983}(Table 5.2)

The sequence was subjected to microsatellite repeat finder (http://Zlab.bu. edu/ repfind/) for clustered, exact repeat search within the sequence. The software identified 12 repeats between position P'_{210} and P'_{998} on the sequence. The repeats identified included an $(AT)_3$ repeat at P'_{210}, a $(TA)_3$ repeat at P'_{264}, a $(TAC)_3$ repeat at P'_{437}, a $(GG)_4$ repeat at P'_{481}, a $(CCAC)_3$ repeat at P'_{541}, an $(AC)_3$ repeat at P'_{657}, two $(TT)_3$ repeats at P'_{700} and P'_{785}, a $(CA)_3$ repeat at P'_{759}, a $(GAA)_3$ repeat at P'_{765}, a $(AG)_3$ repeat at P'_{987} and a $(GAG)_3$ repeat at P'_{998} (Table 5.3).

The deduced amino acid sequence for the 1,116 bases of the 1.2 Kb amplicon comprising of 199 amino acid residues with a predicted isoelectric point (*pI*) of 9.75 and calculated molecular weight of 22.3 kDa is presented in Fig. 5.8. Sequence similarity analysis with BLASTp against non-redundant protein database identified the deduced sequence with the granule bound starch synthase family. ClustalW multiple alignment of the sequence with deduced amino acid sequences of genes coding for granule bound starch synthases of other plants is presented in Fig. 5.9. Using an alignment that permitted maximum homology, the sequence showed a maximum of 84% homology with the deduced amino acid sequence of GBSS-I of *Nelumbo nucifera* (acc. no. ACM78591). The sequence showed 83%, 80%, 79% and 79% homolgy with deduced GBSS-I sequences of *Gossypium hirsutum* (acc. no. ACJ11735), *Phaseolus vulgaris* (acc. no. BAA82346), *Ipomoea batatas* (acc. no. BAI83439) and *Pisum sativum* (acc. no.

Table 5.2: Palindromic structures detected in the 1116 bp nucleotide sequence of the 1.2 Kb amplicon generated with primer pair NDF7-NDR6.

Positions	Palindromic Structures
P'_{54} - P'_{59}	ACATGT
P'_{126} - P'_{131}	TGATCA
P'_{159} - P'_{164}	ACTAGT
P'_{171}- P'_{176}	AAATTT
P'_{210}- P'_{217}	AATATATT
P'_{259} - P'_{266}	TAAATTTA
P'_{406} - P'_{411}	CATATG
P'_{536}- P'_{641}	AGATCT
P'_{715} - P'_{720}	CGGCCG
P'_{846} - P'_{851}	TATATA
P'_{881} - P'_{886}	GGCGCC
P'_{978}- P'_{983}	GTATAC

Table 5.3: Simple sequence repeates found in the 1116 bp nucleotide sequence of the 1.2 Kb amplicon generated with primer pair NDF7-NDR6.

Positions	Type	Repeats
210	dinucleotide	$(AT)_3$
264	dinucleotide	$(TA)_3$
437	trinucleotide	$(TAC)_3$
481	dinucleotide	$(GG)_4$
541	tetranucleotide	$(CCAC)_3$
657	dinucleotide	$(AC)_3$
700	dinucleotide	$(TT)_3$
759	dinucleotide	$(CA)_3$
765	trinucleotide	$(GAA)_3$
785	dinucleotide	$(TT)_3$
987	dinucleotide	$(AG)_3$
998	trinucleotide	$(GAG)_3$

Fig.5.8 Deduced amino acid sequences of the coding region of 1116 bp nucleotide sequences for the 1.2 Kb amplicon generated from buckwheat gDNA as the template and primer pair NDF7-NDR6.

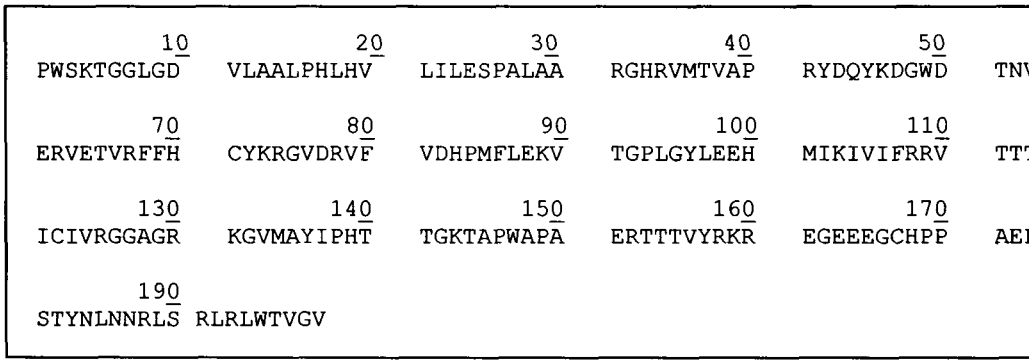

10 20 30 40 50
PWSKTGGLGD VLAALPHLHV LILESPALAA RGHRVMTVAP RYDQYKDGWD
70 80 90 100 110
ERVETVRFFH CYKRGVDRVF VDHPMFLEKV TGPLGYLEEH MIKIVIFRRV
130 140 150 160 170
ICIVRGGAGR KGVMAYIPHT TGKTAPWAPA ERTTTVYRKR EGEEEGCHPP
190
STYNLNNRLS RLRLWTVGV

Fig. 5.9 ClustalW (1.81) multi-alignment of the deduced amino acid sequences for the 1.2 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF7-NDR6 with amino acid sequences of GBSS proteins of other plants available in the database. '*' denotes conserved residues and ':' denotes invariant or similar residues.

```
O.s.indicaGBSSII-ACY56082       PWCKTGGLGDVLGGLP---------PALAAMGHRVMTIVPRYDQYKDAWD 41
Z.maysGBSS1b-ACG43100           PWCKTGGLGDVVGGLP---------PALAAMGHRVMTIAPRYDQYKDAWD 41
T.aestivumGBSSII-AF109395       PWCKTGGLGDVVGGLP---------PALAAMGHRVMTIAPRYDQYKDTWD 41
G.hirsutumGBSS-ACJ11735         PWSKTGGLGDVLGGLP---------PAMAAKGHRVMTVCPRYDQYKDAWD 41
N.nuciferaGBSS-ACM78591         PWSKTGGLGDVLGGLP---------PAMAANGHRVMTVAPRYDQYKDAWD 41
I.batatasGBSSI-BAI83438         PWCKTGGLGDVLGGLP---------PALAARGHRVMTVCPRYDQYKDAWD 41
P.vulgarisGBSSI-BAA82346        PWSKTGGLGDVLGGLP---------SALAEHGHRVMTVSPRYDQYKDAWD 41
P.sativumGBSS-CAC69955          PWSKTGGLGDVLGGLP---------PALSANGHRVMTVTPRYDQYKDAWD 41
F.esculentum                    PWSKTGGLGDVLAALPHLHVLILESPALAARGHRVMTVAPRYDQYKDGWD 50
                                **.********:..**          .*::  ******: ******** **

O.s.indicaGBSSII-ACY56082       TNVLVEVNIGDRTETVRFFHCYKRGVDRVFVDHPMFLEKVWGKTGAKLYG 91
Z.maysGBSS1b-ACG43100           TSVLVEVNIGDTVETVRFFHCYKRGVDRVFVDHPMFLEKVWGKTGAKLYG 91
T.aestivumGBSSII-AF109395       TNVLVEVIVGDRTETVRFFHCYKRGVDRVFVDHPMFLEKVWGKTGSKLYG 91
G.hirsutumGBSS-ACJ11735         TSVLVDLKVGDKVVTVRFFHCYKRGVDRVFVDHPMFLEKVWGKTASKIYG 91
N.nuciferaGBSS-ACM78591         TSVLVEIKVGDRIETVRFFHCYKRGVDRVFVDHPMFLEKVWGKTGSKIYG 91
I.batatasGBSSI-BAI83438         TCVVVELQVGDRIEPVRFFHSYKRGVDRVFVDHPMFLEKVWGKTGSMLYG 91
P.vulgarisGBSSI-BAA82346        TNVTVEVKVADRIETVRFFHCYKQGVDRVFVDHPCFLEKVWGKTGSKLYG 91
P.sativumGBSS-CAC69955          TNVTIEVKVGDRTEKVRFFHCFKRGVDRVFVDHPIFLEKVWGKTGTKLYG 91
F.esculentum                    TNVLVQIQVGERVETVRFFHCYKRGVDRVFVDHPMFLEKVTGPLG----- 95
                                * * ::: :.:    *****.:*:********** ***** *  .

O.s.indicaGBSSII-ACY56082       PTTGDDYRDNQLRFCLLCLAALEAPRVLNLNNSEYFSGPYGENVVFVAND 141
Z.maysGBSS1b-ACG43100           PTTGTDYRDNQLRFCLLCLAALEAPRVLNFNNSEYFSGPYGEDVVFVAND 141
T.aestivumGBSSII-AF109395       PTTGTDFRDNQLRFCLLCLAALEAPRVLNLNNSEYFSGPYGENVVFVAND 141
G.hirsutumGBSS-ACJ11735         PRAGLDYEDNQLRFSLLCQAALEAPRVLNLNSSKNFSGPYGEDVVFIAND 141
N.nuciferaGBSS-ACM78591         PMAGEDYSDNQLRFSLLCQAALEAPRVLNLNSSKHFSGPYGEDVVFICND 141
I.batatasGBSSI-BAI83438         PKAGKDYKDNQLRFSLLCQAALEAPRVLNLNSSKYFSGPYGEDVVFVAND 141
P.vulgarisGBSSI-BAA82346        PSAGVDYEDNQLRYSLLCQAALEAPRVLNLNSNKYFSGPYGEDVIFVAND 141
P.sativumGBSS-CAC69955          PAAGDDYQDNQLRFSIFCQAALEAARVLNLKSNKYFSGPYGEDVIFVAND 141
F.esculentum                    -----YLEEHMIKIVIFRR----------VTTTTFASG------------ 118
                                       :: ::  ::            ....   **

O.s.indicaGBSSII-ACY56082       WHTGVLPCYLKSIYQAKGMYVNAKVAFCIHNIAYQGRFAREDFELLNLPD 191
Z.maysGBSS1b-ACG43100           WHTAILPCYLKSMYKPNGIYKNAKVAFCIHNIAYQGRFARADFDLLNLPD 191
T.aestivumGBSSII-AF109395       WHTAVLPCYLKSMYKQNGIYVNAKVAFCIHNIAYQGRFPRVDFELLNLPE 191
G.hirsutumGBSS-ACJ11735         WHSALLPCYLKSMYQSRGIYMSAKVVFCIHNIAYQGRFAFADFKRLNLPE 191
N.nuciferaGBSS-ACM78591         WHTALLPCYLKTMYKSRGIYRNAKVAFCIHNIAYQGRFSFSDFSLLNLPD 191
I.batatasGBSSI-BAI83438         WHTALLPCYLKTMYQSRGIYMNAKVAFCIHNIAYQGRFAFSDFSLLNLPD 191
P.vulgarisGBSSI-BAA82346        WHTALLPCYLKSMYQTRGVYRNTKVAFCIHNISYQGRHPFEDFPLLNLPN 191
P.sativumGBSS-CAC69955          WHTALISCYMKSMYQSIGIFRNAKVVFCIHNIAYQGRFAFTDYSLLNLPD 191
F.esculentum                    ------------------------TQICIVRGG----------------- 127
                                                       . :** . .

O.s.indicaGBSSII-ACY56082       SFLPSFDFIDGHFKPVLGRKINWMKAGITECDLVMTVSPHYVKELTSGPD 241
Z.maysGBSS1b-ACG43100           SFLPSFDFIDGHVKPFLGRKLNWMKAGIIESDLVLTVSPHYVKELTSGPD 241
T.aestivumGBSSII-AF109395       SFMPSFDFVDGHVKPVVGRKINWMKAGITECDVVLTVSPHYVKELTSGPE 241
G.hirsutumGBSS-ACJ11735         RFKSSFDFIDGYNKPVMGRKINWMKAGILESHRVLTVSPYYAQELVSGED 241
N.nuciferaGBSS-ACM78591         EFKSSFDFIDGYDKPVKGRKINWMKAGILESDRVLTVSPYYAEELVSGIE 241
I.batatasGBSSI-BAI83438         EYKGSFDFIDGYDKPVKGRKINWMKAGIREADRVFTVSPNYAKELVSCVS 241
P.vulgarisGBSSI-BAA82346        EYRSAFDFTDGHLKPVRGRKINWMKAAILESDLVLTVSPYYAKELVSGED 241
P.sativumGBSS-CAC69955          QFKSSFDFLDGHVKPIVGRKINWMKAGIIESHRVLTVSPYYAQELVSGPD 241
F.esculentum                    ----------------AGRKG------------VMAYIPHTTGKTAP--- 146
                                               ***             *::  *  . : ..

O.s.indicaGBSSII-ACY56082       KGVELDGVLRTKPLETGIVNGMDVYEWNPATDKYISVKYDATTVTEARAL 291
Z.maysGBSS1b-ACG43100           KGVELDGVLRTKPLEIGIVNGMDVYEWDPSTDKYISVKYDATTVTEARAL 291
T.aestivumGBSSII-AF109395       KGVELDGVLRAKPLETGIVNGMDVVDWNPATDKYISVKYNATTVAEARAL 291
G.hirsutumGBSS-ACJ11735         KGVELDNIIR-KTGITGIVNGMDVQEWNPASDKYISVKYDATTIMKAKPL 290
N.nuciferaGBSS-ACM78591         KGVELDNIIR-KTGITGIVNGTDVQEWNPTTDKYISVKYDATTVMDAKPL 290
I.batatasGBSSI-BAI83438         KGVELDNHIR-DCGITGICNGMDTQEWNPATDKYLAVKYDITTVMQAKPL 290
P.vulgarisGBSSI-BAA82346        RGVELDNIIR-KTGVAGIVNGMDIREWSPKTDKFIDIHFDTTSVKEAKFL 290
P.sativumGBSS-CAC69955          KGVELDNILR-RVGVTGIVNGMDVQEWNPSTDKYISIKYDASTVLEGKAL 290
F.esculentum                    --------------------------WAPAER--------TTTVYRKREG 162
                                                          * *           :::   :

O.s.indicaGBSSII-ACY56082       NKEMLQAEVGLPVDSSIPLIVFVGRLEEQKGSDILIAAIPEFVEENAQII 341
Z.maysGBSS1b-ACG43100           NKESLQAEVGLPVDSSIPVIVFVGRLEEQKGSDILIAAIPEFVGENVQII 341
T.aestivumGBSSII-AF109395       NKEILQAEVGLPVDSSIPVIVFIGRLEEQKGSDILIAAIPEFLEENVQII 341
G.hirsutumGBSS-ACJ11735         LKEAIQGEVGLPCDRDVPLIGFIGRLEEQKGSDILAEAIPKLAAENCQIV 340
N.nuciferaGBSS-ACM78591         LKEALQSEVGLPVDRNIPVIGFIGRLEEQKGSDILAASIPKFIGENVQII 340
I.batatasGBSSI-BAI83438         LKEALQAAVGLPVDRNIPLIGFIGRLEEQKGSDILYAAISKFISMDVQIL 340
```

```
P.vulgarisGBSSI-BAA82346        LKEALQAEVGLPVNRDIPLIGFIGRLEEQKGSDILVEAIPKFIDQNVQII 340
P.sativumGBSS-CAC69955          LKEELQAEVGLPVDKNVPLIAFIGRLEEQKGSDILVEAIPQFIKENVQIV 340
F.esculentum                    EEEGCHPPAELRLTG---------------------------------- 177
                                 :*  :  . *

O.s.indicaGBSSII-ACY56082       VLGTGKKKMEEELILLEVKYPNNARGIAKFNVPLAHMMFAGADFIIVPSR 391
Z.maysGBSS1b-ACG43100           VLGTGKKKMEEELTQLEVKYPNNARGIAKFNVPLAHMMFAGADFIIVPSR 391
T.aestivumGBSSII-AF109395       VLGTGKKKMEEELMLLEAKYPQNARGIAKFNVPLAHMMFAGANFIIVPSR 391
G.hirsutumGBSS-ACJ11735         VLGTGKKAMENQIEQLEIQYPDNVRAVAKFNVPLAHMIIAGADYILVPSR 390
N.nuciferaGBSS-ACM78591         VLGTGKKAFEKQLEQLEIKYPDKARGVAKFNVPLAHMIIAGADFLLIPSR 390
I.batatasGBSSI-BAI83438         ILGTGKKKFEQQIEQLEVMYPDKARGVAKFNVPLAHMITAGADFMLIPSR 390
P.vulgarisGBSSI-BAA82346        ILGTGKKSMEKQIEQLEEIYPEKARGIAKFDGPLAHKIIAGSDFIMIPSR 390
P.sativumGBSS-CAC69955          ALGTGKKEMEKQLQQLEISYPDKARGVAKFNVPLAHMMIAGADFILIPSR 390
F.esculentum                    ------------------------RGVSTYNLNN---------------- 187
                                                         *.::.::

O.s.indicaGBSSII-ACY56082       FEPCGLIQLQGMRYGVVPICSSTGGLVDTVKEGVTGFHMGSFNVECETVD 441
Z.maysGBSS1b-ACG43100           FEPCGLIQLQGMRYGVIPICSSTGGLVDTVEEGVTGFHMGSFNVECETVD 441
T.aestivumGBSSII-AF109395       FEPCGLIQLQGMRYGVIPICSSTGGLVDTVSEGVTGFHMGSFNVEFETVD 441
G.hirsutumGBSS-ACJ11735         FEPCGLIQLHAMRYGTVPIVASTGGLVDTVKEGFTGFQMGAFNVECDEVD 440
N.nuciferaGBSS-ACM78591         FEPCGLIQLQTMPYGTIPLVSSTGGLVDTVKEGYTGFHMGAFNVDCDAVD 440
I.batatasGBSSI-BAI83438         FEPCGLIQLHAMRYGTPCICASTGGLVDTVKEGYTGFHMGAFNVDCETVD 440
P.vulgarisGBSSI-BAA82346        FEPCGLVQLHSMPYGTVPIVSSTGGLVDTVQEGFTGFHMGAFNVDCEAID 440
P.sativumGBSS-CAC69955          FEPCGLIQLQAMRYGTVPIVASTGGLVDTVKEGFTGFHMGSFNVKCDAVD 440
F.esculentum                    -------RLSRLRLWTVGV------------------------------ 199
                                       :*  :   .  :
```

Fig. 5.9

CAC69955). The alignment score also revealed 80% homology with GBSSII of *Oryza sativa* (acc. no. ACY56082) and *Triticum aestivum* (acc. no. AAF14233). One of the significant differences observed between the deduced amino acid sequences reported in the present study and the deduced amino acid sequences of granule bound starch synthase of other plant species, as observed in the alignment, was the insertion of 9 amino acid residues "HLHVLILES" at P'_{17} and the presence of three highly conserved domains between P'_1-P'_{16}, P'_{33}-P'_{51} and P'_{66}-P'_{90}. While the conserved domain between P'_1-P'_{16} comprised of 16 residues represented by the sequence "PWSKTGGLGDVL AALP", that between P'_{33}-P'_{51} comprised of 20 residues represented by the sequence "GHRVMTVAPRYDQYKDGWDT". The third conserved domain comprised of 25 residues between P'_{66} to P'_{90}. This domain was represented by the sequence "VRFFHCYKRGVDRVFVDHPMFLEKV". ClustalW multiple alignment also revealed a high level of sequence similarity within the first 95 amino acid residues aligned out of the total 199 deduced amino acid residues. This sequence showed similarity with the glycoslytransferase family of proteins (Fig. 5.10).

Out of the 199 residues compared, the position of 68 residues was found to be conserved among the deduced amino acid sequences amongst all the granule bound starch synthases analyzed in the present study. These included P_1, W_2, K_4, T_5, G_6, G_7, L_8, G_9, D_{10}, V_{11}, L_{15}, P_{16}, A_{27}, G_{32}, H_{33}, R_{34}, V_{35}, M_{36}, T_{37}, P_{40}, R_{41}, Y_{42}, D_{43}, Q_{44}, Y_{45}, K_{46}, D_{47}, W_{49}, D_{50}, T_{51}, V_{53}, V_{66}, R_{67} F_{68}, F_{69}, H_{70}, K_{73}, G_{75}, V_{76}, D_{77}, R_{78}, V_{79}, F_{80}, V_{81}, D_{82}, H_{83}, P_{84}, F_{86}, L_{87}, E_{88}, K_{89}, V_{90}, G_{92}, S_{137}, G_{138}, C_{178}, I_{179}, G_{218}, R_{219}, K_{220}, V_{234}, P_{239}, W_{277}, P_{279}, E_{303}, L_{311}, R_{375} and L_{409}. Other residues which were represented by amino

Fig. 5.10 Screen shot of the BLAST of the 199 deduced amino acid sequence of the 1.2 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF7-NDR6 showing the conserved position of the 95 amino acids among GBSS protein sequences in different plant species.

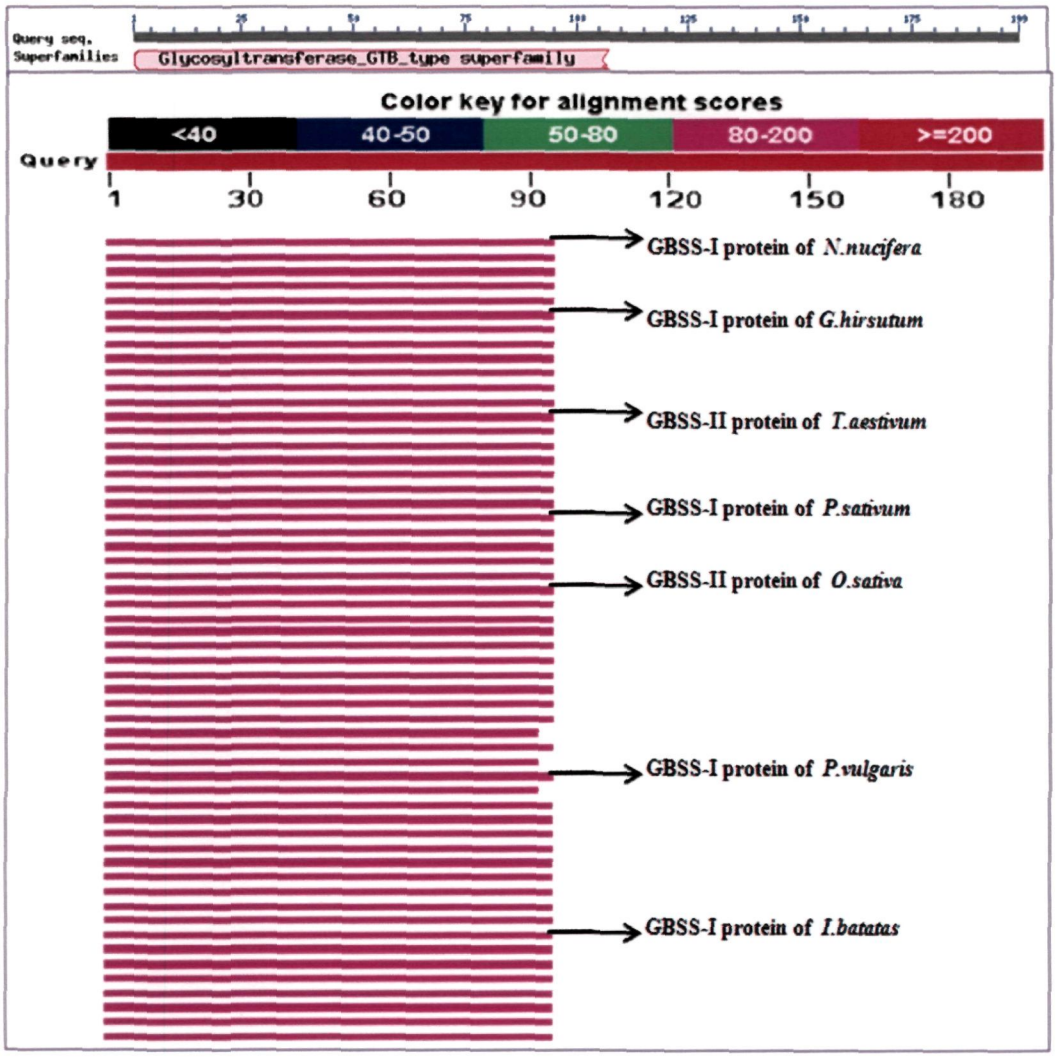

Query seq.
Superfamilies
Glycosyltransferase_GTB_type superfamily
Color key for alignment scores
<40
40-50
50-80
80-200
>=200
Query
1
30
60
90
120
150
180
GBSS-I protein of N.nucifera
GBSS-I protein of G.hirsutum
GBSS-II protein of T.aestivum
GBSS-I protein of P.sativum
GBSS-II protein of O.sativa
GBSS-I protein of P.vulgaris
GBSS-I protein of I.batatas

Fig. 5. 10

acids with invariant/similar functional groups included C/S_3, L/V_{12}, L/M_{28}, V/I_{38}, V/I_{55}, V/I/L_{57}, V/I_{59}, D/E_{61}, C/S_{71}, G/A_{95}, D/E_{109}, I/L_{112}, K/R_{113}, I/L_{116}, F/L_{117}, V/L/F_{130}, S/N/T_{132}, S/N/T_{133}, I/F_{177}, M/L/V_{235}, T/S_{292}, S/T_{293}, V/I_{294}, K/R_{298}, V/A_{309}, A/G_{376}, I/V_{377}, M/L_{412} and V/I/L_{419}.

Domain search on the deduced amino acid sequence identified the putative substrate binding site comprising of the sequence Lys-X-Gly-Gly-Leu (K-X-G-G-L) between P'$_4$ and P'$_8$. The sequence has been universally identified as a signature for the starch synthase family of proteins. The sequence was subjected to MOTIF SCAN (http://myhits.isb-sib.ch/cgi-bin/motif_scan) for identification of defined motifs within the sequence. Motif search on the deduced amino acid sequence identified a starch synthase catalytic domain predominantly represented by the amino acid residues "PWSKTGGLGDVLAALPHLHVLILESPALAARGHRVMTVAPRYDQYKDGWDT NVLVQIQVGERVETVRFFHCYKRGVDRVFVDHPMFLEKVTGPLGYLEEHMIKIV IFRRVTTTTFASGTQICIVRGGAGRKGVMAY" between positions 1-136 (Fig. 5.11a). The software also detected three protein kinase C phosphorylation sites represented by the sequence Thr-Val-Arg (P'$_{65-67}$), Thr-Gly-Lys (P'$_{141-143}$) and Thr-Gly-Arg (P'$_{176-178}$) (Fig. 5.11 b).

When the deduced amino acid sequence were plotted as a function of hydropathic index, the sequence showed a predominantly hydrophobic character (Fig. 5.12). Based on the hydropathic index of Kyte and Doolittle (1982), the major regions of hydrophobic nature detected in the sequence were located between residues 5-28, 50-57, 77-95 and 102-139. Statistical analysis by SAPS (Brendel *et al.*, 1992) predicted the

Fig. 5.11a Diagrammatic representation of the starch synthase domains positioned between 1 to 136 amino acid residues in the 199 deduced amino acid sequence as predicted by MOTIF SCAN.

Fig. 5.11b Diagrammatic representation of the Protein Kinase C phosphorylation site in the 199 deduced amino acid sequence as predicted by MOTIF SCAN.

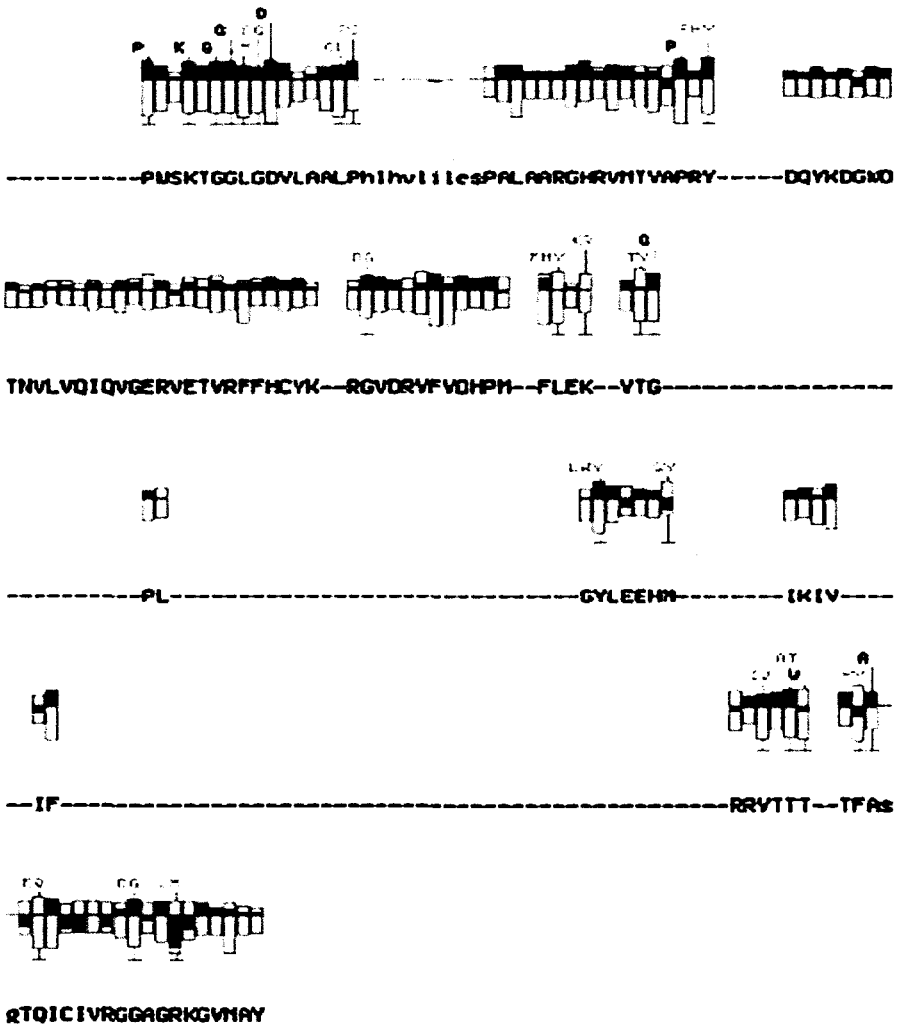


Fig. 5.11a

pos.: 65-67

pos.: 141-143

pos.: 176-178

Fig, 5. 11b

Fig. 5.12 Hydropathy profile of the deduced 199 amino acid sequence for the 1.2 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF7-NDR6. The X-axis represents the residue number.

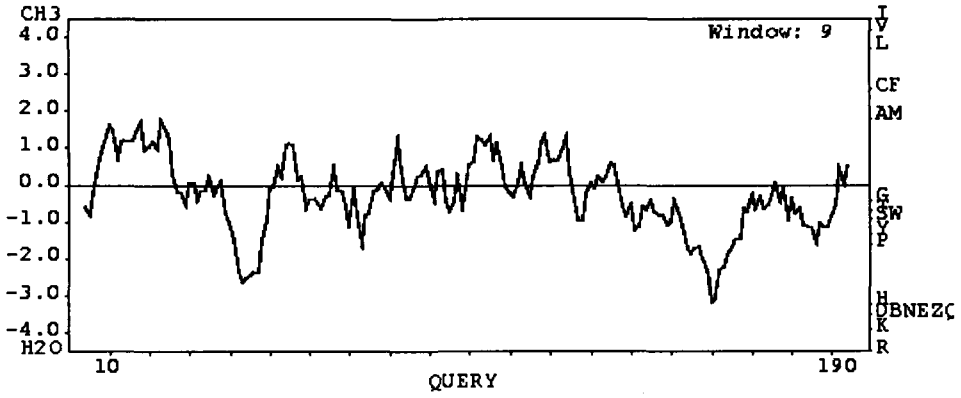


Fig. 5.12

sequence to have 28.1% nonpolar residues, 18.6% polar uncharged residues and 22.6% polar charged residues. Regions of high nonpolar residue concentration in the deduced protein sequence was correlated with the predicted hydrophobic amino acid residues in the hydropathy analysis. PSIPRED and GOR4 predicted 14.0% α-helix, 32.16 % extended strand and 53.77% random coil structure for the deduced amino acid sequence identified in the present study. Statistical analysis by PSIPRED indicated that the residues between 39-50, 82-98 and 126-152 would predominantly acquire a random coil secondary structure whereas residues between 99-106, 8-12 and 25-30 would predominantly fold into an α-helix (Fig. 5.13).

Phylogenetic analysis of the deduced amino acid sequence reported in the present study with amino acid sequences of granule bound starch synthases available in EMBL database, revealed a clear diversification into monocotyledonous (GBSS-II) and dicotyledonous (GBSS-I) groups. (Fig. 5.14). The monocotyledons were represented by Group-I and comprised of *Oryza sativa* (acc. no. ACY56082), *Triticum aestivum* (acc. No. AAF14233) and *Zea mays* (acc. no. ACG43100). The dicotyledons on the other hand, were represented by Group-II and comprised of *Ipomoea batatas* (acc. no. BAI83439), *Nelumbo nucifera* (acc. no. ACM78591), *Gossypium hirsutum* (acc. no. ACJ11735), *Pisum sativum* (acc. no. CAC69955), *Phaseolus vulgaris* (acc. no. BAA82346) and *Fagopyrum esculentum. Jabrossa squarrosa* (acc. no. ABM46904) was taken as an outgroup while constructing the phylogenetic tree.

Polymerase chain reaction comprising of hot start at 94°C for 5 mins., 35 amplification cycles comprising of denaturation at 94°C for 1 min., annealing at 64°C

Fig. 5. 13 Secondary structure prediction profile of the 199 deduced amino acid sequence for the 1.2 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF7-NDR6.

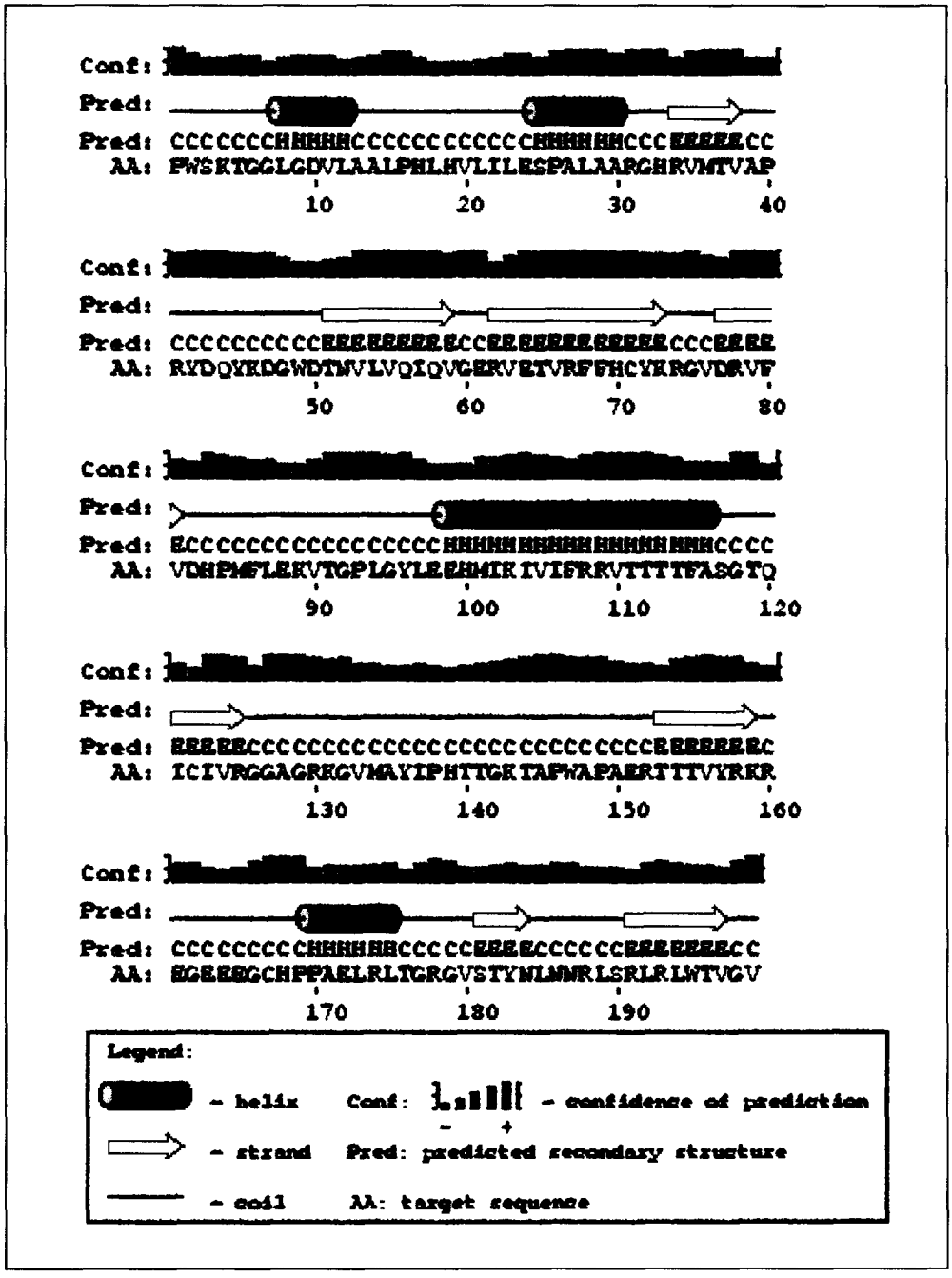

Conf:
Pred:
Pred: CCCCCCCHHHHHCCCCCCCCCCCHHHHHHCCCEEEEECC
AA: PWSKTGGLGDVLAALPHLHVLILESPALAARGHRVMTVAP
10 20 30 40
Conf:
Pred:
Pred: CCCCCCCCCCEEEEEEEECCEEEEEEEEEEEECCCEEEE
AA: RYDQYKDGWDTWVLVQIQVGERVETVRFFHCYKRGVDRVF
50 60 70 80
Conf:
Pred:
Pred: ECCCCCCCCCCCCCCCCHHHHHHHHHHHHHHHHCCCC
AA: VDHPMFLEKVTGPLGYLEEHMIRIVIFRRVTTTTFASGTQ
90 100 110 120
Conf:
Pred:
Pred: EEEEECCCCCCCCCCCCCCCCCCCCCCCCCEEEEEEEC
AA: ICIVRGGAGREGVMAYIPHTTGRTAPWAPAERTTTVYRKR
130 140 150 160
Conf:
Pred:
Pred: CCCCCCCCHHHHHHCCCCCEEEECCCCCCEEEEEEECC
AA: EGEEEGCHPPAELRLTGRGVSTYMLNNRLSRLRLWTVGV
170 180 190
Legend:
- helix
- strand
- coil
Conf: - confidence of prediction
Pred: predicted secondary structure
AA: target sequence

Fig. 5.13

Fig. 5.14 Phylogenetic tree based on the alignment matrix of the 199 deduced amino acid sequences with amino acid sequences of GBSS proteins of various plants available in the Genbank.

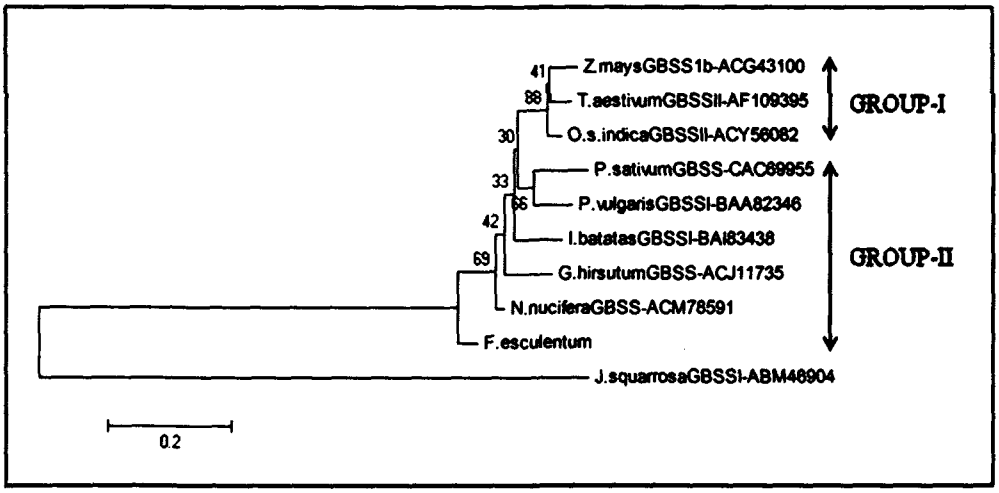

41
88
30
33
66
42
69
Z.maysGBSS1b-ACG43100
T.aestivumGBSSII-AF109395
O.s.indicaGBSSII-ACY56082
P.sativumGBSS-CAC69955
P.vulgarisGBSSI-BAA82346
I.batatasGBSSI-BAI83438
G.hirsutumGBSS-ACJ11735
N.nuciferaGBSS-ACM78591
F.esculentum
J.squarrosaGBSSI-ABM46904
GROUP-I
GROUP-II
0.2

Fig. 5.14

for 1min. and chain extension at 72°C for 1min. followed by one reaction of chain elongation at 72°C for 10 mins. with buckwheat genomic DNA as the template and oligonucleotide primer pair NDF8-NDR6, amplified a DNA fragment which resolved on 0.8% agarose gel as a fluorescent band showing an apparent molecular mass of 0.9 Kb (Fig. 5.15). The nucleotide sequence comprising of 896 bases for the 0.9 Kb is presentend in Fig. 5.16. BLASTn analysis of the sequence identified it with the nucleotide sequnces of GBSS genes. The sequence showed a maximum of 70% homology with granule bound starch synthase of *Nolana paradoxa* (acc. no. EU051872), *Schizanthus litoralis* (acc. no. DQ299433) and *Ipomoea batatas* (acc. no. AB524725).

The sequence was subjected to analysis by AUGUSTUS (version2.4) for determination of number and position of exons. AUGUSTUS 2.4 identified two exons within the nucleotide sequence of 896 bases in the 0.9 kb DNA fragment amplified in the present study. While the 1st exon had a length of 101 bases and located between P'$_{527}$-P'$_{627}$, the 2nd exon had a length of 114 bases and located between P'$_{711}$-P'$_{824}$. The software identified the segment comprising of 101 bp as an internal exon. (Fig.5.16).

ClustalW multiple alignment of the sequence with nucleotide sequences of granule bound starch synthase genes from other plant species is presented in Fig. 5.17. Using an alignment that permitted maximum homology, the sequence showed three highly conserved regions betweem P'$_{368}$-P'$_{388}$, P'$_{533}$ -P'$_{602}$ and P'$_{709}$-P'$_{823}$. While the conserved region between P'$_{368}$-P'$_{446}$ comprised of 21 bases represented by the sequence "TCAGGCTGCTCTAGAGGCACC" that between P'$_{533}$-P'$_{602}$, comprised of 70 nucleotides represented by the sequence "ATGTTGTTTTCGTTCCAAATGACTGGCA

Fig. 5.15 Electrophoresis profile of amplified fragment with buckwheat gDNA as template and primer pair NDF8-NDR6, M= *Eco*RI/ *Hind*III digested λ DNA.

Fig. 5.16 Nucleotide sequence of 896 bases for the 0.9 Kb fragment amplified with buckwheat gDNA as template and primer pair NDF8-NDR6.The exons identified in the sequence are highlighted in green and their start and end positions are indicated with forward and backward arrow.

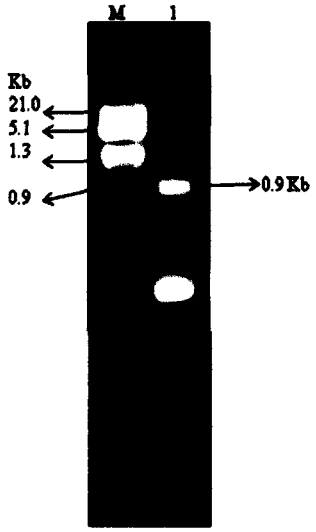


Fig. 5. 15

```
1     CTTCTAGTCA TCTAATCGGG GTACTCAAAT CGGTATTTTT CAGGTGTGGG
51    GTAACACTAG ATCAAAGATT TATGGGCCTA ATGCAGGAAC TGATTATGAT
101   GACAACCAAC TCAGATTCAG CTTATTGTCC CATGGTAATG CTTGATAATT
151   TCAATGAAAT ATGAAGCTTT TGTTTTGCAC TGCAACATTA TAATCAAATA
201   AGTATCTTTG ATCAGTAAAA ATACCTCGTT GGTTTGATAC TCTTGTCTAC
251   ACCTGCCTTT TTTGTTGAAC CTGTCTAAGT TGAGTATACT CTTTGCATGT
301   AAATCTGAAA TGTCACAACA AAATGAACAA AGAAAGATTA GTTGTATTAA
351   AAATCTTGTC ACTTTCTTCA GGCTGCTCTA GAGGCACCTA GGGTACTAAA
401   CCTGAACAAC GGAAAGCACC ACTCTGGCCC CTACGGTAAA GAGCTTTACG
451   CTTAATGTTT GACTGTATTC TGTAACCTTC AAATTCCTTC TGATATTATG
501   AAGTTCTTAC ATGTATACGA ATGCAG[redacted]GA AGATGTTGTT TTCGTTCCAA
551   ATGACTGGCA CACTGCTCTT GTTCCCTGTT ACCTCAAATC TGTGTACCAA
601   TCTCGGGGCA TCTACAAAAA TGCTA[redacted]GTG ATTTTCACTG AACTTGGCAT
651   AATTTGCCAT CGAAGTAGAG AAAAAACACA GTTTTGCTAA ATTCGATCAT
701   CCTATGGCAG [redacted]TGCATTCT GTATTCACAA TATTTCATAC CAAGGAAGAT
751   TTGCTTTTTC AGACTATTCG ATGCTCAATT TGCCTGCAGA GTACAAGGGC
801   TCGTTTGATT TCATTGATGG GT[redacted]GATTTT TTCGAAACCA GCATCTATAA
851   ATATATACTT GCCACTCTGA TCTCACGACG GCTAATTGAT CCTGCA
```

Fig. 5.16

Fig. 5.17 ClustalW (1.81) multi-alignment of the 896 bp nucleotide sequences for the 0.9 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF8-NDR6 with nucleotide sequences of GBSS genes available in the database. '*' denotes conserved residues.


S.litoralis-DQ299433 ----------CCAATATTCCTAG 13
N.paradoxa-EU051872 ----------GTAA 4
I.batatas-EU192919 CTTCCATTCATACAAACGCGG-AGTTGATCGCGTTTTTGTGGATCATCCTATGTTCCTGG 59
F.esculentum CTTCTAGTCATCTAATCGGGGTACTCAAATCGGTATTTTTCAGGTGTGGGGTAACACTAG 60
S.litoralis-DQ299433 AGAAAGTTAGTAT----------TTTCTGAAC--CTT-----TATACAACATAATATGAA 56
N.paradoxa-EU051872 GCATATTTTGAAT----------CCTTAAAAAGGTCC-----TGAGGGGCACAATGTGAA 49
I.batatas-EU192919 AGAAGGTTAGTATAGAGTATAGACTCATGGATTTTCCAGAGTTATGGTGTATGTTACCAG 119
F.esculentum ATCAAAGATTTATGGGCCTAATGCAGGAACTGATTATGATGACAACCAACTCAGATTCAG 120
S.litoralis-DQ299433 TGCACAGAAC----ACATCATATTGAAATT-----TGACTTTACGGCTTGCT-ACTACCC 106
N.paradoxa-EU051872 TACACGGAAC----ACATCATTTTGAATTTC-AGTTGACTTTACTGGT-GCT-TTTACCC 102
I.batatas-EU192919 TGAATTGTTT----GAATGGTTATAAGCTACTAATCGACATGAATGCTTTCGATCTGTGC 175
F.esculentum CTTATTGTCCCATGGTAATGCTTGATAATTTCAATGAAATATGAAGCTTTTGTTTTGCAC 180
S.litoralis-DQ299433 TTG-AAGGTTTTGGGCAAAACTGGTTCAAAAATTTATGGCCCCAAAGCTGGAAAAGATTA 165
N.paradoxa-EU051872 TTTTAAGGTTTGGGGCAAAACTGCTTCAAAAATCTATGGCCCCAAAGCTGGACAAGATTA 162
I.batatas-EU192919 TATTTAGGTTTGGGGAAAAACTGGATCTATGCTCTATGGCCCCAAGGCTGGGAAAGATTA 235
F.esculentum TGC-AACATTATAATCAAATAAGTATCTTTGATCAGTAA---AAATACCTCGTTGGTTTG 236
S.litoralis-DQ299433 TCTGGACAATGAACTTAGGTTCAGCTTGCTGTGTCAAG-TAAGTTGCTTGTAGTAGTACT 224
N.paradoxa-EU051872 TGTGGACAATGAACTTAGGTTCAGCTTGTTGTGTCAAG-TAAGTCACCTGTTGTACTGCT 221
I.batatas-EU192919 CAAGGACAACCAGTTGCGGTTCAGTTTGTTGTGCCAAG-TAA--TGCATAGTGTCTAATC 292
F.esculentum ATACTCTTGTCTACACCTGCCTTTTTTGTTGAACCTGTCTAAGTTGAGTATACTCTTTGC 296
S.litoralis-DQ299433 GTCTTGACTCTATACACCCAGCGGCTTTTTACTTGGTTATT-AATCCTTGTTTAACCTA- 282
N.paradoxa-EU051872 GTCTTGACTTTAC-------GTTGCATTTTACTTATGTATTGAATCATTTTTTAACCTTA 274
I.batatas-EU192919 GTCT--ATCTTAG----------GCTTCACATGAGTG----AACATTATTTAACATT- 333
F.esculentum ATGTAAATCTGAAATGTCACAACAAAATGAACAAAGAAAGATTAGTTGTATTAAAAATC- 355
S.litoralis-DQ299433 ---CTCCTGCCACT-TCAGGCAGCTCTAGAGGCACCTAGAGTTCTGAATTTGAATAGTAG 338
N.paradoxa-EU051872 TTTTTTTTGTCACTCTCAGGCAGCGCTAGAGGCACCTAGAGTTCTGAATTTGACCTGCAG 334
I.batatas-EU192919 TCACTTGTGTTGGTTTCAGGCAGCACTTGAGGCACCGAGAGTTTTGAATCTTAACTCCAG 393
F.esculentum ---TTGTCACTTTCTTCAGGCTGCTCTAGAGGCACCTAGGGTACTAAACCTGAACAACGG 412
S.litoralis-DQ299433 CAAATACTTCACAGGACCATATGGTAATACTTTGCAGTTTAAGAAAGTCTCCTTGGCAGT 398
N.paradoxa-EU051872 CAAATACTTCTCAGGACCCTATGGTAGCACCTCCCAGTTTCAGAAAG-CACCTTAGCAGT 393
I.batatas-EU192919 CAAATACTTCAGTGGACCTTATGGTA---TTTTTCTTCTTCTAGAGGTTTTGGAAGTACT 450
F.esculentum AAAGCACCACTCTGGCCCCTACGGTAAAGAGCTTTACGCTTAATGTTTGACTGTATTCTG 472
S.litoralis-DQ299433 CNTAGT-GTGTCCTAGTAGGTAATCATCTT----TGTTTTGCCTATTTCTGCAGGTGAGG 453
N.paradoxa-EU051872 CACAGTTATATCCTTGTAGGTAATCATCTT----TATTTTGCCTATTTCTGCAGGAGAGG 449
I.batatas-EU192919 TGAGATTAGAGTGTGGAAAGAATTAACATAATAATCTTTTATGAAATTGTAAAGGTGAGG 510
F.esculentum TAACCTTCAAATTCCTTCTGATATTATGAAGTTCTTACATGTATACGAATGCAGGAGAAG 532
S.litoralis-DQ299433 ATGTTCTCTTCATTGCCAATGATTGGCACACTGCTCTTCTTCCCTGTTATCTGAAGGCTA 513
N.paradoxa-EU051872 ATGTTCTCTTCATTGCCAATGATTGGCACACAGCTCTCATTCCCTGCTACCTGAAGTCTA 509
I.batatas-EU192919 ATGTTGTTTTCGTTGCCAATGATTGGCACACTGCTCTCCTTCCGTGCTATCTGAAAACCA 570
F.esculentum ATGTTGTTTTCGTTCCAAATGACTGGCACACTGCTCTTGTTCCCTGTTACCTCAAATCTG 592
S.litoralis-DQ299433 TGTACCAATCTAGAGGAATTTATGTGAATGCCAAGGTAAAACATCTTTTGAATAC----- 568
N.paradoxa-EU051872 TGTACCAGTCAAAAGGAATCTATCTGAATGCAAAGGTAAAATCTCTTT-GTATTC----- 563
I.batatas-EU192919 TGTACCAGTCGAGAGGAATCTACATGAACGCCAAGGTAATGCCTCTTCTGGATTGGACGA 630
F.esculentum TGTACCAATCTCGGGGCATCTACAAAAATGCTAAGGTGA--TTTTCACTGAACTTGGC-A 649
S.litoralis-DQ299433 -ACTTTATTGCACATTAGCTTGCC-AATCAAGAAGGTTATAT--ATATCTG-----AAAA 619
N.paradoxa-EU051872 -ACTTTACT----------TTGCG-AATCAAGAAGGTTGTATTAATATATG-----ATAA 606
I.batatas-EU192919 AAGTTTTCCACATTTAATCGTACCTAATCAAAAGGACTATACTTATATGTGTTCCCACTA 690
F.esculentum TAATTTGCC----------ATCGAAGTAGAGAAAAAACACAGTTTTGCT-----AAAT 692
S.litoralis-DQ299433 TTT-TCACACTGCCTCCAGGTTGTTTTCTGCATCCATAACATTGCCTACCAAGGAAGATT 678
N.paradoxa-EU051872 ATT-TCACATTGCCTCCAGGTCGCTTTCTGCATCCATAATATTGCCTACCAAGGTCGATT 665
I.batatas-EU192919 TTTGTTAAAATGTTGCCAGGTTGCTTTCTGCATTCACAACATTGCCTACCAAGGCAGATT 750
F.esculentum TCGATCATCCTATGGC-AGGTTGCATTCTGTATTCACAATATTTCATACCAAGGAAGATT 751

```
S.litoralis-DQ299433      TTCTTTTTCAGACTTCTCTCTTCTCAATCTGCCTGACGAATACAAGAGTTCTTTTGATTT 738
N.paradoxa-EU051872       TGCTTTCTCAGACTTCTCTCTTCTCAATCTGCCTGATGAGTACAGGAGTTCTTTTGATTT 725
I.batatas-EU192919        CGCCTTTTCAGACTTTTCTCTTCTGAATCTGCCTGATGAGTACAAGGGTTCTTTTGATTT 810
F.esculentum              TGCTTTTTCAGACTATTCGATGCTCAATTTGCCTGCAGAGTACAAGGGCTCGTTTGATTT 811
                            * ** *******  **  * ** *** ******  ** **** * * ** ********

S.litoralis-DQ299433      CATTGATGGGTAAGAATTTTTTGCTTGAAATCAGTCTACCAATTTTAGGACCTC---TTT 795
N.paradoxa-EU051872       CATTGATGGGTATGGATTTTATGCTTGGGACTAAACCACCTACTTCAGAAGCTC---TTT 782
I.batatas-EU192919        TATTGATGGGTAGGAATTAGACGCTTTTAGCTAAGATTATGAATTATGTATCTTGATTCT 870
F.esculentum              CATTGATGGGTAGGATTTTT----TCGAAACCAGCATCTATAAATATATACTTGCCACTC 867
                           ********** *  **      *       *        *  *    *  *

S.litoralis-DQ299433      GGCTTCTTGTTTAGTGAGTGTT--ATATTTTGCAGATATGCCAAGCCAGTTAAGGGTAG- 852
N.paradoxa-EU051872       TGATGTTAGTAAATTGAATTTTTAAAATTTTGCAGATCTGAAAAGCCTGTTCAGGGTAGG 842
I.batatas-EU192919        TGCTGCTTGATCTCTGAATTCAATTCATTCT-CAATAATGCCAGGTATG---------- 918
F.esculentum              TGATCTCACGACGGCTAATT-----GATCCTGCA-------------------------- 896
                           * *           * *       **  * **
```

Fig. 5.17

CACTGCTCTTGTTCCCTGTTACCTCAAATCTGTGTACCAATC". The third conserved domain comprised of 115 bases from P'_{709}-P'_{823}. This domain was represented by the sequence "AGGTTGCATTCTGTATTCACAATATTTCATACCAAGGAAGAT TTGCTTTTTCAGACTATTCGATGCTCAATTTGCCTGCAGAGTACAAGGGCTCG TTTGATTTCATTGATGGGTA". Sequence analysis showed that the 2nd conserved domain comprising of the sequence "ATGTTGTTTTCGTTCCAAATGACT GGCACACTGCTCTTGTTCCCTGTTACCTCAAATCTGTGTACCAATC" between P'_{533}-P'_{602} was located within the 1st exon (P'_{527}-P'_{627}). The other two conserved regions in the sequence were found to be part of the introns identified in the sequence. BLASTn analysis of the coding region of the sequence comprising of 287 bases showed a maximum of 80% homology with coding region of GBSS-I gene of *Ipomoea batatas* (acc. no. AB524727), 76% homology with coding region of GBSS-I gene of *Nelumbo nucifera* (acc. no. EU938541) and 75% homology with coding region of GBSS-I gene of *Kageneckia oblonga* (acc.no. DQ874893), *Lindleya mespiloides* (acc.no.DQ874930), *Perilla frutescens* (acc. no. AF210699) and *Photinia villosa* (acc. no. AF500452).

Analysis of the nucleotide sequence data for its G+C content revealed that the DNA fragment has a G+C content of 38%. The sequence showed specific A/T and G/C rich regions. While the segment between P'_{27} to P'_{40}, P'_{145} to P'_{165}, P'_{187} to P'_{223}, P'_{297} to P'_{311}, P'_{342} to P'_{368}, P'_{478} to P'_{485}, P'_{493} to P'_{509}, P'_{613} to P'_{626}, P'_{671} to P'_{693}, P'_{729} to P'_{739}, P'_{826} to P'_{836} and P'_{842} to P'_{860} was specifically rich in A/T content, that between P'_{383} to P'_{393}, P'_{416} to P'_{436} and P'_{874} to P'_{882} was rich in G/C content. The EMBOSS CpGplot software of European Bio Informatics Institute detected a putative CpG island of

495bases located between position 355 to 854. The ratio of observed to expected CG composition in this segment of the sequence was > 0.60.

NCBIscan detected twelve palidromic structures in the nucleotide sequence of the 0.9 Kb amplicon. These incuded 'CCATGG' between P'_{130}-P'_{135}, 'AAGCTT' between P'_{164}-P'_{169}, 'TTATAA' between P'_{188}-P'_{193}, 'TGATCA' between P'_{209}-P'_{214}, 'GTATAC' between P'_{284}-P'_{289}, 'CCTAGG' between P'_{387}-P'_{392}, 'ACATGT' between P'_{509}-P'_{514}, 'AATATT' between P'_{729}-P'_{734}, 'CCATGG' between P'_{740}-P'_{745}, 'CTGCAG' between P'_{783}-P'_{788}, 'TTCGAA' between P'_{830}-P'_{835} and 'ATATAT' between P'_{850}-P'_{855}. (Table 5.4)

The sequence was subjected to microsatellites repeat finder (http://Zlab.bu.edu/repfind/) for clustered, exact repeat search within the sequence. The software identified 4 repeats between positions P'_{275} and P'_{850} on the sequence. The repeats identified included a $(TT)_3$ repeat at P'_{275}, an $(AA)_3$ repeat at P'_{670}, a $(TT)_3$ repeat at P'_{826}, and an $(AT)_3$ repeats at P'_{850} (Table 5.5).

The deduced amino acid sequence for the 896 bases of the 0.9 Kb amplicon comprising of 70 amino acid residues with a predicted isoelectric point (*pI*) of 6.05 and calculated molecular weight of 8.0 kDa is presented in Fig. 5.18. Sequence similarity analysis with BLASTp against non-redundant protein database, identified the deduced sequence with the GBSS family of proteins. BLASTn analysis also identified the entire deduced amino acid with the glycosyltransferase super family (Fig. 5.19). ClustlW multiple alignment of the sequence with deduced amino acid sequences of genes coding for granule bound starch synthases from other plant species is presented in Fig. 5.20.

Table 5.4: Palindromic structures detected in the 896 bp nucleotide sequence of the 0.9 Kb amplicon generated with primer pair NDF8-NDR6.

Positions	Palindromic Structures
P'_{130} - P'_{135}	CCATGG
P'_{164} - P'_{169}	AAGCTT
P'_{188} - P'_{193}	TTATAA
P'_{209}- P'_{214}	TGATCA
P'_{284} - P'_{289}	GTATAC
P'_{387} - P'_{392}	CCTAGG
P'_{509} - P'_{514}	ACATGT
P'_{729} - P'_{734}	AATATT
P'_{740} - P'_{745}	CCATGG
P'_{783}- P'_{788}	CTGCAG
P'_{830} - P'_{835}	TTCGAA
P'_{850} - P'_{855}	ATATAT

Table 5.5: Simple sequence repeates found in the 896 bp nucleotide sequence of the 0.9 Kb amplicon generated with primer pair NDF8-NDR6.

Positions	Type	Repeats
275	dinucleotide	$(TT)_3$
670	dinucleotide	$(AA)_3$
826	dinucleotide	$(TT)_3$
850	dinucleotide	$(AT)_4$

Fig.5.18 Deduced amino acid sequences of the coding region of 896 bp nucleotide sequences for the 0.9 Kb amplicon generated from buckwheat gDNA as the template and primer pair NDF8-NDR6.

Fig. 5.19 Screen shot of the BLAST of 70 deduced amino acid sequence of the 0.9 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF8-NDR6 showing the conserved position of the all the amino acids among GBSS protein sequences in different plant species.

```
        10          20          30          40          50          60
EDVVFVPNDW  HTALVPCYLK  SVYQSRGIYK  NAKVAFCIHN  ISYQGRFAFS  DYSMLNLPAE

        70
YKGSFDFIDG
```

Fig. 5.18

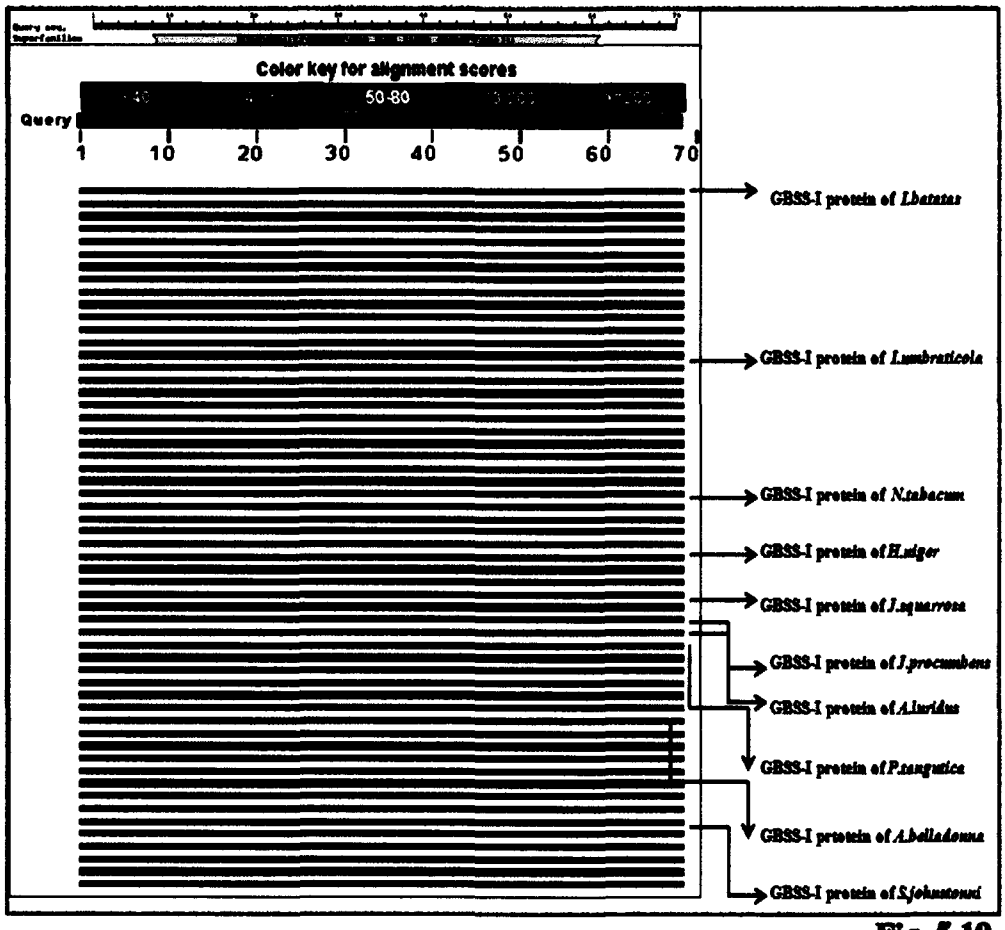


Fig. 5.19

Fig. 5.20 ClustalW (1.81) multi-alignment of the deduced amino acid sequences for the 0.9 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF8-NDR6 with amino acid sequences of GBSS proteins of other plants available in the database. '*' denotes conserved residues and ':' denotes invariant or similar residues.

```
I.batatas-BAI83436          EDVVFVANDWHTALLPCYLKTMYQSRGIYMNAKVAFCIHNIAYQGRFAFS 50
I.umbraticola-ABW83774      EDVVFVANDWHTALLPCYLKTMYQSRGIYMNAKVAFCIHNIAYQGRFAFS 50
A.luridus-AAZ99048          EDVLFIANDWHTALIPCYLKSMYQSRGLYMNAKVAFCIHNIAYQGRFAFS 50
A.belladonna-AAZ99047       EDVLFIANDWHTALIPCYLKSMYQSRGLYMNAKVAFCIHNIAYQGRFAFS 50
P.tangutica-AAZ99055        EDVLFIANDWHTALIPCYLKSMYQSRGLYMNAKVTFCIHNIAYQGRFAFS 50
H.niger-AAZ99051            EDVLFVANDWHTALVPCYLKSMYQSRGLYMNAKVAFCIHNIAYQGRFAFS 50
J.squarrosa-ABM46904        EDVLFIANDWHTALIPCYLKSMYQSRGIYMNAKVAFCIHNIAYQGRFAFS 50
S.nemorense-AAY63641        EDVLFIANDWHTALIPCYLKSMYQSRGIYLNAKVAFCIHNIAYQGRFAFS 50
S.johnstonii-ACV96018       EDVLFIANDWHTALIPCYLKSIYQSRGIYLNAKVAFCIHNIAYQGRFAFS 50
J.procumbens-AAY63568       EDVLFIANDWHTALIPCYLKSMYQSRGIYVNAKVAFCIHNIAYQGRFAFS 50
N.tabacum-AAZ99063          EDVVFIANDWHTALLPCYLKSMYQSRGIYMNAKVAFCIHNIAYQGRFAFS 50
F.esculentum                EDVVFVPNDWHTALVPCYLKSVYQSRGIYKNAKVAFCIHNISYQGRFAFS 50
                            ***:*:.*******:*****::*****:* ****:******:********

I.batatas-BAI83436          DFSLLNLPDEYKGSFDFIDG 70
I.umbraticola-ABW83774      DFSLLNLPDEYKGSFDFIDG 70
A.luridus-AAZ99048          DFSLLNLPDEYRSSFDFIDG 70
A.belladonna-AAZ99047       DFSLLNLPDEYRSSFDFIDG 70
P.tangutica-AAZ99055        DYSLLNLPDEYRSSFDFIDG 70
H.niger-AAZ99051            DFSLLNLPDEYRSSFDFIDG 70
J.squarrosa-ABM46904        DFSLLNLPDEYRSSFDFIDG 70
S.nemorense-AAY63641        DFSLLNLPDEFRGSFDFIDG 70
S.johnstonii-ACV96018       DFPLLNLPDEFRGSFDFIDG 70
J.procumbens-AAY63568       DFPLLNLPYEYRGSFDFIDG 70
N.tabacum-AAZ99063          DFSLLNLPDEYKSSFDFIDG 70
F.esculentum                DYSMLNLPAEYKGSFDFIDG 70
                            *:.:**** *::.*******
```

Fig. 5.20

Using an alignment that permitted maximum homology, the deduced amino acid sequence showed a maximum of 96% homology with deduced amino acid sequence of granule bound starch synthase-I (GBSS-I) of *Ipomoea batatas* (acc. no. BAI83439), *Ipomoea umbraticola* (acc. no. ABW83774) and *Solanum nemorense* (acc. no. AAY 63641) with a query coverage of 100%. The sequence also showed 95% homology with deduced amino acid sequence of granule bound starch synthase-I (GBSS-I) of *Nicotiana tabacum* (acc. no. AAZ99063), *Hyoscyamus niger* (acc. no. AAZ99051), *Atropa belladonna* (acc. no. AAZ99047), *Jaborosa squarrosa* (acc. no.ABM46904), *Jaltomata procumbens* (acc. no. AAY63568), *Anisodus luridus* (acc. no. AAZ99048) and *Solanum johnstonni* (acc. no. ACV96018) and 93% homology with the deduced amino acid sequence of granule bound starch synthase-I (GBSS-I) of *Przewalskia tangutica* (acc. no. AAZ99055).

Out of the 70 residues compared, the position of 55 residues was found to be conserved amongst the deduced amino acid sequences of the entire granule bound starch synthases analyzed in the present study. The conserved residues included E_1, D_2, V_3, F_5, N_8, D_9, W_{10}, H_{11}, T_{12}, A_{13}, L_{14}, P_{16}, C_{17}, Y_{18}, L_{19}, K_{20} Y_{23}, Q_{24}, S_{25}, R_{26}, G_{27}, Y_{29}, N_{31}, A_{32}, K_{33}, V_{34}, A_{35}, F_{36}, C_{37}, I_{38}, H_{39}, N_{40}, I_{41}, Y_{43}, Q_{44}, G_{45}, R_{46}, F_{47}, A_{48}, F_{49}, S_{50}, D_{51}, S_{53}, L_{55}, N_{56}, L_{57}, P_{58}, E_{60}, s_{64}, F_{65}, D_{66}, F_{67}, I_{68}, D_{69} and G_{70}. Other residues which were represented by amino acids with invariant/similar functional groups include V/I_4, V/I_6, P/A_7, V/I_{15}, T/S_{21}, V/M_{22}, L/I_{28}, M/L_{54} and R/K_{62}.

Motif search by MOTIFSCAN (http://myhits.isb-sib.ch/cgi-bin/motif_scan) identified the entire sequence as starch synthase catalytic domain (Fig. 5.21a). The

Fig. 5.21a Diagrammatic representation of the starch synthase domains positioned between 1 to 70 amino acid residues in the 70 deduced amino acid sequence as predicted by MOTIF SCAN.

Fig. 5.21b Diagrammatic representation of the N-glycosylation sites in the 70 deduced amino acid sequence as predicted by MOTIF SCAN.

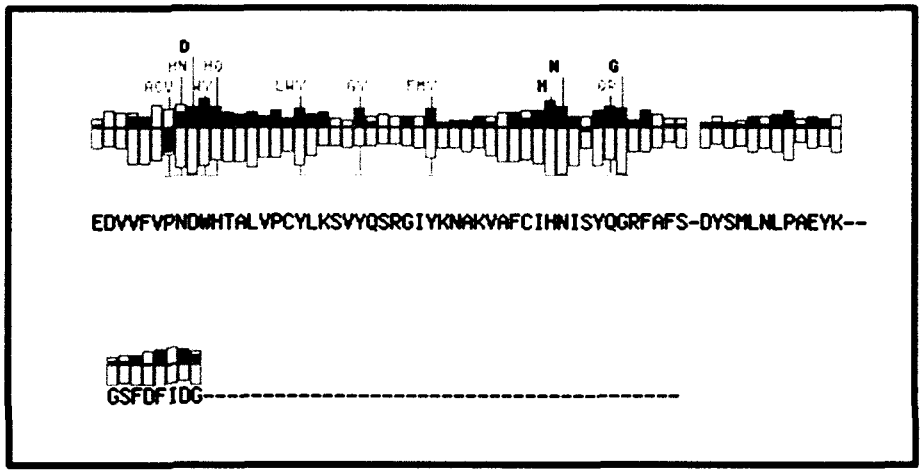


Fig. 5.21a

Fig. 5.21b

software also detected one putative N-glycosylation sites represented by the amino acids sequences Asn-Ile-Ser-Tyr' ($P'_{98\text{-}101}$) (Fig. 5.21b).

A plot of position of each residue as a function of hydropathic index revealed the predominantly hydrophilic character of the sequence with the major regions of hydrophilic nature located between residues 20-32, 42-56 and 61-67. (Fig. 5.22). Sequence analysis by SAPS predicted the presence of 30.8% residues with non polar functional side chains, 35.7% residues with uncharged polar groups and 18.6% residues with charged polar group in the sequence. The regions with predominantly nonpolar residues correlated with the hydropathy analysis carried out on the sequence. PSIPRED and GOR4 predicted 5.71% α-helix, 30.00% extended strand and 64.29% random coil secondary structure for the deduced amino acid sequence. Statistical analysis by PSIPRED indicated that the residue between 41-70 would predominantly acquire random coils secondary structure whereas residues between 10-24 would predominantly fold into a α-helix (Fig. 5.23).

A phylogenetic tree, describing the relationship of the deduced amino acid sequence reported in the present study with amino acid sequences of granule bound starch synthase proteins from other plants, was constructed by maximum parsimony method using the alignment matrix generated in the present study. The deduced amino acid sequence of 70 residues clustered together into one clad with amino acid sequences of *Ipomoea batatas* (acc.no. BAI83439) and *Ipomoea umbraticola* (acc.no. ABW83774). While the amino acid sequences of granule bound starch synthase protein from *Nicotiana tabacum* (acc. no. AAZ99063), *Hyoscyamus niger* (acc. no. AAZ99051),

Fig. 5.22 Hydropathy profile of the deduced 70 amino acid sequence for the 0.9 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF8-NDR6. The X-axis represents the residue number.

Fig. 5.23 Secondary structure prediction profile of the 70 deduced amino acid sequence for the 0.9 Kb amplicon generated with buckwheat genomic DNA as template and primer pair NDF8-NDR6.

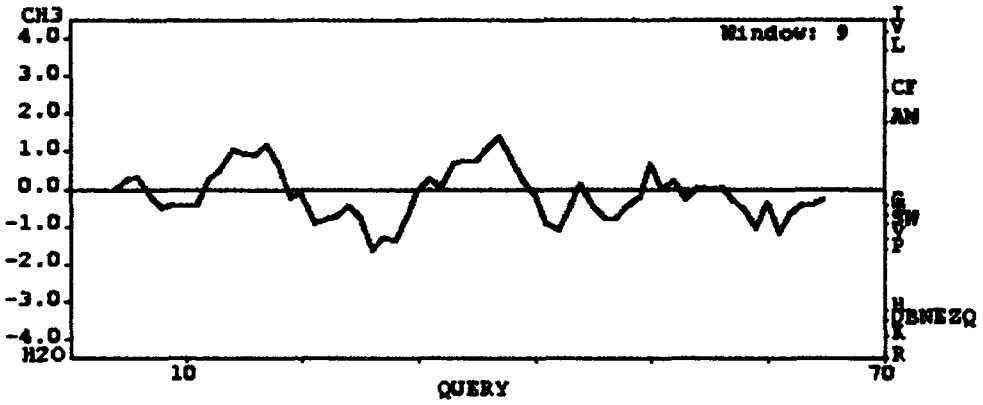


Fig .5.22

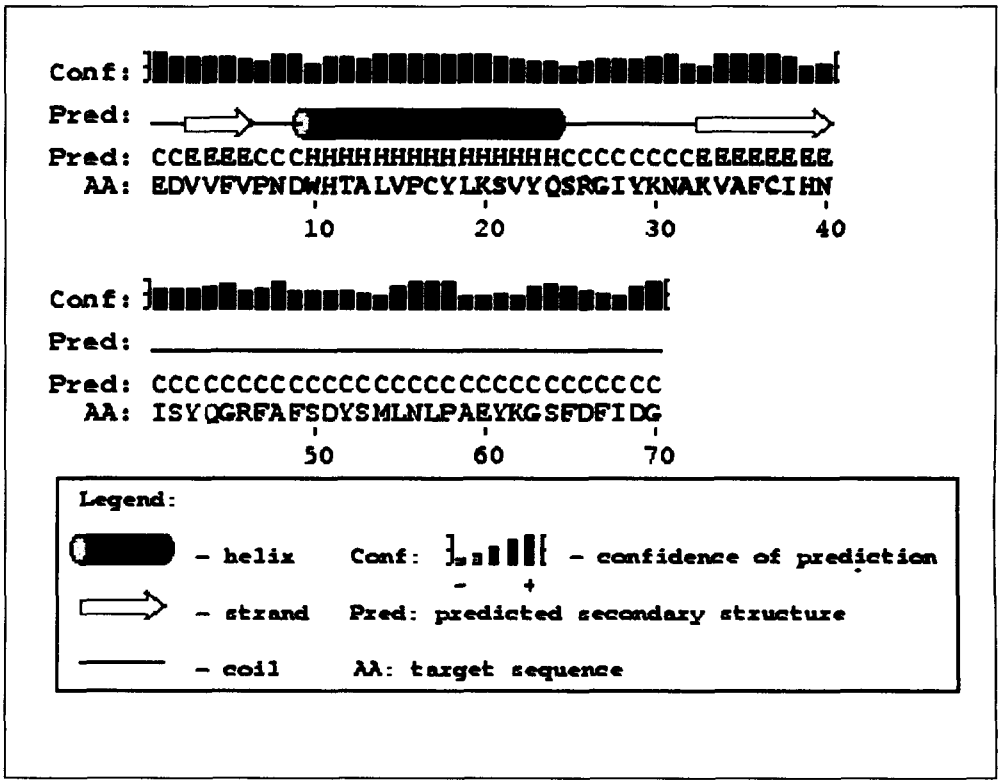


Fig. 5.23

Fig. 5.24 Phylogenetic tree based on the alignment matrix of the 70 deduced amino acid sequences with amino acid sequences of GBSS proteins of various plants available in the Genbank.

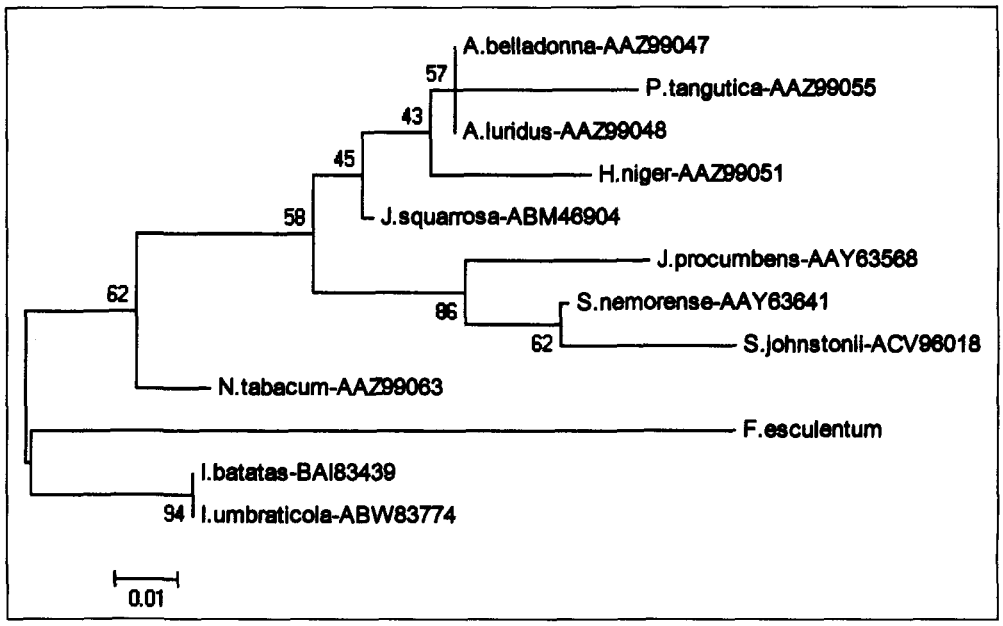

A.belladonna-AAZ99047
57
P.tangutica-AAZ99055
43
A.luridus-AAZ99048
45
H.niger-AAZ99051
58
J.squarrosa-ABM46904
J.procumbens-AAY63568
62
86
S.nemorense-AAY63641
62
S.johnstonii-ACV96018
N.tabacum-AAZ99063
F.esculentum
I.batatas-BAI83439
94
I.umbraticola-ABW83774
0.01

Fig. 5.24

Anisodus luridus (acc. no. AAZ99048), *Atropa belladonna* (acc. no. AAZ99047), *Jaborosa squarrosa* (acc. no.ABM46904), *Jaltomata procumbens* (acc. no. AAY63569), *Solanum johnstonni* (acc. no. ACV96018), *Solanum nemorense* (acc. no. AAY 63641),*Przewalskia tangutica* (acc. no. AAZ99055) clustered into another clad. (Fig. 5. 24).

DISCUSSION

The modified CTAB protocol used in the present investigation for isolation of genomic DNA yielded fairly good quality DNA from the etiolated seedlings of common buckwheat. The ratio of absorbance at 260 and 280 nm (A_{260}/A_{280}) for the DNA samples isolated during the present investigation ranged between 1.7 to 1.9 indicating that the isolated DNA was fairly pure. One of the biggest obstacles in isolation of high quality DNA from starch/protein rich plant tissues is the interferences in purification procedures due to the presence of carbohydrates, proteins and even polyphenolic compounds. While *EcoR*1 and *Hind*III digested buckwheat genomic DNA fully, *Nco*I could only partially digest buckwheat genomic DNA. The electrophoretic profile of *EcoR*1 and *Hind*III digested DNA revealed a uniform smear ranging in size from 0.5 Kb to 20 Kb with prominent bands indicating the presence of *EcoR*1/*Hind*III repeats of varying lengths in the genomic DNA of common buckwheat. Similar results have been observed by Bharali (2002), on the restriction digestion profile of buckwheat genomic DNA. Bharali (2002), has reported 6 bands representing *Eco*R1 repeats in genomic DNA of common buckwheat. Even though, *Nco*I could only partially digest buckwheat genomic DNA, the electrophoretic profile of *Nco*I digested DNA revealed six bands having molecular

masses of 2.46 Kb, 1.87 Kb, 1.108 Kb, 0.983 Kb, 0.592 Kb and 0.488 Kb. The appearance of distinct bands in *Eco*R1, *Hind*III and *Nco*I digested gDNA indicates the presence of *Eco*R1, *Hind*III and *Nco*I repeats of varying lengths in buckwheat genomic DNA.

Even though isolation of genes from genomic DNA libraries by screening of the library with appropriate probes, offers a reliable system for cloning full length genes, PCR amplification using oligonucleotide primers designed from the nucleotide sequences of conserved regions of 5' upstream and 3' downstream regions of the genes offers a much quicker way of isolation of full length genes directly from genomic DNA isolated from the target plants. In the present study, polymerase chain reaction was carried out with buckwheat genomic DNA as the template and combinations of oligonucleotide primer pairs designed to amplify the sequence of GBSS gene from buckwheat genomic DNA. Out of the total 18 primer pairs used, 14 sets failed to amplify any sequence from the template DNA. Amplification was, however, achieved with the primer combinations NDF2-NDR2, NDF3-NDR3, NDF7-NDR6 and NDF8-NDR6. PCR amplification with oligonucleotide primer pairs NDF2-NDR2 and NDF3-NDR3 amplified a DNA fragment showing an apparent molecular mass of 0.7 kb on 0.8% agarose gel. Even though this indicated the presence of primer specific homologous sequences in the buckwheat genomic DNA, BLAST analysis of the nucleotide sequences of the two amplicons did not identify any of the two sequences with GBSS gene family.

While amplification of gDNA wtih oligonucleotide primer pairs NDF7-NDR6 amplified a DNA fragment showing an apparent molecular mass of 1.2 Kb that with primer pair NDF8-NDR6 amplified a DNA fragment having an apparent molecular mass of 0.9 Kb. BLASTn analysis of the nucleotide sequences of the amplicons identified the same with genes encoding for granule bound starch synthase gene family. This is the first report of identification of GBSS-I gene in common buckwheat. ClustalW multiple alignment of the nucleotide sequence of 1.2 Kb amplicon with nucleotide sequences of granuel bound starch synthase genes of other plants revealed two highly conserved regions. While one of the conserved regions existed between P'_{94} and P'_{168}, the other existed between P'_{274} to P'_{377}. The conserved region between P'_{94} and P'_{168} comprised of 74 bases represented by the sequence "GGACATAGGGTTATGACAGTTGCTCCTCG TTTGATCAGTATAAAGATGGATGGGATACTAATGTACTAGTTCAG". The other conserved domain comprised of 104 bases between P'_{274} and P'_{377}. This domain was represented by the sequence "CAGATACAAGTTGGGGAAAGAGTTGAGACTGTTC GGTTCTTTCACTGCTACAAAAGAGGAGTTGACCGGGTTTTCGTGGATCACCA TGTTCCTTGAGAAGGT" AUGUSTUS (version 2.4) identified four exons and three introns within the nucleotide sequence of 1.2 Kb amplicon and two exons/two introns in the nucleotide sequence of the 0.9 Kb amplicon. Sequence analysis of the 1.2 kb amplicon clearly revealed the presence of 1st and 2nd conserved domains within the sequence in the 1st and 2nd exon respectively. The sequences at the 5' and 3'ends of the three introns were found to follow the universal GT-AG rule (Breathnach and Chambon, 1981). Our results are in conformity with the observations of Xu *et al.* (2009), on the

nucleotide sequence of *waxy* gene of rye wherein, the intron/exon architecture has been demonstrated to follow the universal GT-AG rule. All the three introns identified in our study were shown to have AT rich regions. Similiar reports of AT rich regions in the introns of *Oryza glaberrima waxy* gene have been reported Umeda *et al.* (1991). Camirand *et al.* (1990) and Van der Leij *et al.* (1991), have observed that all introns in the GBSS-I gene also followed the universal GT-AG. In this context our observations on the intron/exon architecture in the nucleotide sequence of the 1.2 Kb amplicon are in agreement with those of Umeda *et al.* (1991), Camirand *et al.* (1990) and Van der Leij *et al.* (1991).

While the coding region of 604 bases of the nucleotide sequence of 1.2 Kb amplicon showed a maximum of 90% homology with the coding region of granule bound starch synthase-I of other plant species, the noncoding regions identified in the sequence did not show any significant homology with GBSS genes. Xu *et al.* (2009), have reported similar results on the sequence homology of rye *waxy* gene with *waxy* genes of other plants. While the coding region of rye *waxy* gene showed a maximum of 97.4% homology with coding regions of *waxy* genes of other plants, the maximum homology observed for the non coding region was only 53%. These results indicate a higher level of sequence conservation of GBSS genes in the coding region than the non coding region. Similar results have been reported for GBSS genes in sorghum, rice and maize (Chen *et al.*, 1998).

EMBOSS CpGplot detected a putative CpG island of 856 bases located between positions 310 to 1116 in the nucleotide sequence of the 1.2 Kb amplicon and a

CpG island of 495bases located between position 355 to 854 in the nucleotide sequence of the 0.9 kb amplicon. The ratio of observed to expected CG composition in this segment of the sequence was > 0.60. Detection of a putative CpG plot is an indication that the gene could be under tight regulatory control. Since CpG islands are known to be associated with the 5' end of genes which are subjected to regulatory control (Larsen *et al.*, 1992), the detection of CpG Island in the DNA amplified in the current study indicates that the region of DNA amplified could belong to the 5' end of the target gene.

NCBI scan detected twelve palindromic structures in the nucleotide sequences of both 1.2 Kb amplicon 0.9 Kb amplicon. Jangsutthivorawat and Volkaert (2009), have correlated the presence of (CT)n and (AATT)n repeats at the GBSS-I loci of rice with the variations in the amylose contenent in the starch. Bao *et al.* (2006b), have also indicated a close correlation between the number of repeats with the AAC in starch grains of North American and Chinese germplasm. Alleles with fewer repeats (n<12) were observed to be associated with higher apparent amylose content (AAC) and those with more repeats (n>12) with a lower AAC. Domon *et al.* (2002), have identified three SSRs *viz.*, $(TCGA)_3$, $(AT)_9$ and $(GGAT)_3$ *waxy* gene of barley. Identification of such repeats in the partial nucleotide sequence of the buckwheat GBSS-I gene, identified in the present study could be used as markers for screening of buckwheat accessions for their AAC.

The deduced amino acid sequence for the 1,116 bp bases of the 1.2 Kb amplicon comprised of 199 amino acid residues with a predicted isoelectric point (p*I*) of 9.75 and calculated molecular weight of 22.3 kDa. Sequence similarity analysis of the

sequence with BLASTp against non-redundant protein database, identified the deduced sequence with the granule bound starch synthase family. Domain search on the deduced amino acid sequence identified the putative substrate binding site comprising of the sequence Lys-Thr-Gly-Gly-Leu (K-T-G-G-L) between P'_4 and P'_8. KTGGL is a universal motif identified in GBSS-I of *Pisum sativum* (Edwards *et al.*, 2002), *Vigna radiata* (Ko *et al.*, 2009), *Solanum tuberosum* (Edwards *et al.*, 1995), *Zea mays* (Harn *et al.*, 1997), grain amaranth (Park *et al.*, 2009), *Triticum aestivum* (Anisworth *et al.*, 1993; Baba *et.al.*, 1993) and *Oryza sativa* (Sano, 1984). The domain has also been reported to be present in GBSSII of cassava, (Munyikwa *et al.*, 1997), SSII of *Pisum sativum* (Edwards *et al.*, 2002), *Solanum tuberosum* (Edwards *et al.*,1995) and *Zea mays* (Harn *et al.*, 1997). The motif has been suggested to be involved in substrate binding (Furukawa *et al.*,1990; 1993). Based on affinity labelling and site directed mutagenesis, it was shown that the lysine residue in conserved motif KTGGL, formed part of the ADPG binding site of *E.Coli* glycogen synthase (Furukawa *et. al.*, 1990, 1993). ADPG is the glucosyl donor in the reaction catalysed by glycogen and starch synthase. Hence, it was assumed that this motif could be the ADPG binding site in the plant enzmes.

Using an alignment that permitted maximum homology, the sequence showed a maximum of 84% homology with deduced amino acid sequences of GBSS-I of *Nelumbo nucifera* (acc. no. ACM78591). The sequence also showed 83%, 80%, 79% and 79% homolgy with deduced amino acid sequences of GBSS-I of *Gossypium hirsutum* (acc. no. ACJ11735), *Phaseolus vulgaris* (acc. no. BAA82346), *Ipomoea batatas* (acc. no. BAI83439) and *Pisum sativum* (acc. no. CAC69955) and 80% homology with GBSSII

of *Oryza sativa* (acc. no. ACY56082) and *Triticum aestivum* (acc. no. AAF14233). The deduced amino acid sequence identified in the present study showed an insertion of 9 residues "HLHVLILES" at P'_{17} and the presence of three highly conserved domains between P'_{1}-P'_{16}, P'_{33}-P'_{51} and P'_{66}-P'_{90}. While the conserved domain between P'_{1}-P'_{16} comprised of 16 residues represented by the sequence "PWSKTGGLGDVLAALP", that between P'_{33}-P'_{51} comprised of 20 residues represented by the sequence "GHRVMTVA PRYDQYKDGWDT". The third conserved domain comprised of 25 residues between P'_{66}-P'_{90}. This domain was represented by the sequence "VRFFHCYKRGVDRVFVDH PMFLEKV". ClustalW multiple alignment also revealed a high level of sequence similarity within the first 95 amino acid residues aligned out of the total 119 deduced amino acid residues. This sequence showed similarity with the glycoslytransferase enzmyes which catalyze the transfer of a monosaccharide unit from an activated nucleotide sugar to a glycosyl acceptor molecule forming glycosidic bonds in carbohydrate residues or other biopolymers (Breton *et al.*, 2006). Motif search on the deduced amino acid sequence identified a starch synthase catalytic domain predominantly represented by the amino acid residues "PWSKTGGLGDVLAALP HLHVLILESPALAARGHRVMTVAPRYDQYKDGWDTNVLVQIQVGERVETVRFF HCYKRGVDRVFVDHPMFLEKVTGPLGYLEEHMIKIVIFRRVTTTTFASGTQICIV RGGAGRKGVMAY" between positions 1-136. The software also detected three protein kinase C phosphorylation sites represented by the sequence Thr-Val-Arg (P'_{65-67}), Thr-Gly-Lys ($P'_{141-143}$) and Thr-Gly-Arg ($P'_{176-178}$). Phosphorylation has been shown to play a major role in modulating the function and DNA-binding activity of many nuclear

proteins, including transcription factors and proteins involved in chromatin organization (Dang *et al.*, 1994; Armstrong *et al.*, 1997; Hoffmann *et al.*, 1998). The identification of such sites in the deduced amino acid sequence is indicative of its role in signal transduction processes during seed maturation.

When the deduced amino acid sequence were plotted as a function of hydropathic index, the sequence showed a predominantly hydrophobic character. Based on the hydropathic index of Kyte and Doolittle (1982), the major regions of hydrophobic nature detected in the sequence were between residues 5-28, 50-57, 77-95 and 102-139. Wang *et al.* (2000), have reported similar results on the hydrophobicity of *Ipomoea batatas* GBSS-I. Statistical analysis by SAPS (Brendel *et al.*, 1992) predicted the sequence to have 28.1% nonpolar residues, 18.6% polar uncharged residues and 22.6% polar charged residues. The absence of any transmembrane helices in the deduced amino acid sequence of the putative GBSS protein indicates presumably the protein functions within the cytosolic or nuclear compartments. This assumption is supported by the detection of putative sites for post translational modification in the amino acid sequence of the putative buckwheat GBSS like protein identified in the present study.

Phylogenetic analysis of the deduced amino acid sequence reported in the present study with amino acid sequences of granule bound starch synthases available in EMBL database, revealed a clear diversification into monocotyledonous (GBSS-II) and dicotyledonous (GBSS-I) groups. The monocotyledons were represented by Group-I and comprised of *Oryza sativa* (acc. no. ACY56082), *Triticum aestivum* (acc. no. AAF14233) and *Zea mays* (acc. no. ACG43100). The dicotyledons on the other hand,

were represented by Group-II and comprised of *Ipomoea batatas* (acc. no. BAI83439), *Nelumbo nucifera* (acc. no. ACM78591), *Gossypium hirsutum* (acc. no.ACJ11735), *Pisum sativum* (acc. no.CAC69955), *Phaseolus vulgaris* (acc. no. BAA82346) and *Fagopyrum esculentum*. According to Pan *et al.* (2009), the putative protein could have duplicated and diverged into two different forms *viz.*, GBSS-I or GBSSIa and GBSS-II or GBSSIb during evolution and the diversification into monocotyledonous (GBSS-II) and dicotyledonous (GBSS-I) groups might be a consequence of this diversification process. Similar observations have been made by Dian *et al.* (2003).

The deduced amino acid sequence for the 896 bases of the 0.9 Kb amplicon comprised of 70 amino acid residues with a predicted isoelectric point (*pI*) of 6.05 and calculated molecular weight of 8.0 kDa. Sequence similarity analysis of the sequence with BLASTp against non-redundant protein database, identified the deduced sequence with the GBSS family. Using an alignment that permitted maximum homology, the deduced amino acid sequence showed a maximum of 96% homology with GBSS-I of *Ipomoea batatas* (acc. no. BAI83439), *Ipomoea umbraticola* (acc. no. ABW83774) and *Solanum nemorense* (acc. no. AAY 63641) with a query coverage of 100%. The sequence also showed 95% homology with *Nicotiana tabacum* (acc. no. AAZ99063), *Hyoscyamus niger* (acc. no. AAZ99051), *Atropa belladonna* (acc. no. AAZ99047), *Jaborosa squarrosa* (acc. no. ABM46904), *Jaltomata procumbens* (acc. no. AAY63568), *Anisodus luridus* (acc. no. AAZ99048) and *Solanum johnstonni* (acc. no. ACV96018) and 93% homology with *Przewalskia tangutica* (acc. no. AAZ99055) respectively. BLASTp analysis of the sequence indicated that the entire 70 amino acid

residues belongs to the glycosyltransferase superfamily, which catalyzes the transfer of sugar moieties from activated donor molecules to specific acceptor molecules, forming glycosidic bonds in carbohydrate residues or other biopolymers (Breton *et al.*, 2006). Similar results were also reported in rye *waxy* gene (Xu *et al.*, 2009). Motif search on the deduced amino acid sequence identified the entire sequence as starch synthase catalytic domain. The starch synthase catalytic domain is also found in glycosyltransferases that uses only ADP glucose as substrate (Leterrier *et al.*, 2008).

A phylogenetic tree, describing the relationship of the deduced amino acid sequence reported in the present study with amino acid sequences of GBSS proteins from other plants, was constructed by maximum parsimony method using the alignment matrix generated in the present study. The deduced amino acid sequence of 70 residues clustered together into one clad with amino acid sequences of *Ipomoea batatas* (acc. no. BAI83439) and *Ipomoea umbraticola* (acc. no. ABW83774). On the other hand, the amino acid sequences of granule bound starch synthase protein from *Nicotiana tabacum* (acc. no. AAZ99063), *Hyoscyamus niger* (acc. no. AAZ99051), *Anisodus luridus* (acc. no. AAZ99048), *Atropa belladonna* (acc. no. AAZ99047), *Jaborosa squarrosa* (acc. no.ABM46904), *Jaltomata procumbens* (acc. no. AAY63569), *Solanum johnstonni* (acc. no.ACV96018), *Solanum nemorense* (acc.no. AAY 63641), *Przewalskia tangutica* (acc. no. AAZ99055) clustered into another clad.

CHAPTER VI

GENERAL SUMMARY AND DISCUSSION

INTRODUCTION

Starch is one of the most important plant products to man. It occurs in the form of a complex granular mixture of amylose and amylopectin. While amylose comprises largely of unbranched α-1, 4-linked glucan chains, amylopectin is comprised of small size α-1, 4-linked chains that are clustered together by α-1, 6-linkages between adjoining straight glucan chains (Tetlow *et al.*, 2004).

Starches are typically used as thickeners and stabilizers in foods such as puddings, custards, soups, sauces, gravies, pie fillings, and salad dressings and to make noodles and pastas. Starch is also widely used in the preparation of foods such as bread, pancakes, cereals, noodles, pasta, porridge and tortilla (Elliason, 2004). On the basis of nutritional characteristics which are mainly based on the extent of digestibility, starch has been classified either as: (i) Digestible starch (Berry, 1986) or (ii) Resistant starch (Englyst *et al.*, 1982). Besides its nutritive values, the industrial applications of starch

include its use as a thickener, colloidal stabilizer, gelling agent, bulking agent, water retention agent and adhesive. Designer starches can be used for bulk or value-added food and feed and in non-food applications as in chemicals or as bioenergy (Morell and Myers, 2005). Thus, a major challenge would be to predict the effect of structural/chemical changes in starch which determines their functional properties. The functional properties of starch such as gelatinization temperature, pasting, retrogradation, water absorption capacity, swelling power, and solubility vary considerably from one botanical source to another and also contribute to use of starch for industrial applications (Dedeh and Sackey, 2002; Chen *et al.*, 2003; Amani *et al.*, 2004; Perez *et al.*, 2005; Peroni *et al.*, 2006; Riley *et al.*, 2006; Shujun *et al.*, 2006; Yuan *et al.*, 2007) These properties also depend on amylose/amylopectin ratio, chain length distribution of amylopectin, phosphorus content, granular size and molecular weight of the starches (Lu *et al.*, 2005; Sasaki and Matsuki, 1998; Fredriksson *et al.*, 1998; Shibanuma *et al.*, 1996; Jane and Chen, 1992; Tian *et al.*, 1991). Therefore, unraveling the potential of starches for use in the food and non-food industries calls for a better understanding of their unique physicochemical, functional and structural properties. Since the physicochemical properties of starches are genetically determined, it is possible to manipulate the properties of starches by genetic manipulation. Therefore, discovery and characterization of the enzymes that affect the quality of starch are of notable worth. The biosynthesis of starch involves the coordinated action of three enzymes *viz.*, ADPglucose pyrophosphorylase (ADPase), starch synthases (SS) and starch branching enzyme (SBE). Kossmann and Lloyd (2000), have identified at least

five isoforms of starch synthases which include one granule bound enzyme *viz.*, GBSS-I and four soluble starch synthases *viz.*, SS-I, SS-II, SS-III and SS-IV. Many starch biosynthetic enzymes have been shown to be distributed between the soluble fraction of the plastids and the insoluble starch granules (Ball and Morell, 2003; Tetlow *et al.*, 2004a). GBSS, which is localized exclusively within the starch granule, is responsible for amylose biosynthesis (Shure *et al.*, 1983). On the other hand the synthesis of amylopectin involves concerted activity of the starch synthases as well as branching and de-branching enzymes (Denyer *et al.*, 2001; Hirose and Terao., 2004). GBSS-I, the key enzyme responsible for amylose synthesis, has been identified as a 58 kDa (Shure *et al.*, 1983) or a 60 kDa (Echt and Schwartz, 1981) protein in maize, a 60 kDa protein in rice (Sano, 1984) and a 68 kDa protein in grain amaranth (Konishi *et al.*, 1985).

Since Amylose-free or low-amylose starch is considered desirable for certain food and non-food industries (Morell *et al.*, 1995), most of the approaches aimed at modifying the quality of starch have focused on alteration of amylose/amylopectin ratio of the grains. Approaches used to achieve this goal are being pursued through manipulation of the genes coding for enzymes involved in the starch biosynthetic pathway (Sestili *et al.*, 2010). The gene coding for the enzyme GBSS-I (called *waxy*) has been cloned from maize (Shure *et al.*, 1983; Klösgen *et al.*, 1986), potato (Hovenkamp-Hermelink *et al.*, 1987; Visser *et al.*, 1989; Van der Leij *et al.*, 1991), barley (Rohde *et al.*, 1988), rice (Wang *et al.*, 1990; Hirano and Sano, 1991; Okagaki, 1992), pea (Dry *et al.*, 1992), sorghum (Hsieh *et al.*,1996a), cassava (Salehuzzaman *et al.*, 1993) and common wheat (Clark *et al.*, 1991; Mason-Gamer *et al.*, 1998 and Murai *et al.*, 1999).

While much work has been done on the characterization of starch from cereals and analysis of the regulation of its biosynthesis, not much information is available about the quality of starch in many other potentially important crops. One such group of crops belongs to pseudo cereals. Amongst this group of plants, common buckwheat is an important pseudo cereal because of its high potential for use as a functional food. However, because of its unfavorable amylose/amylopectin ratio, the buckwheat flour had poor dough making qualities. Since the ratio of amylose to amylopectin influences the texture and quality of flour, identification of the gene loci involved in the regulation of starch synthesis and deposition would have an important bearing on devising effective breeding programs for improved starch quality. These loci could be used as markers for identification of genotypes with low amylose content of starch in the grains. Even though such proteins have been isolated and their genes cloned from other conventional crops, none of the granule bound proteins in buckwheat have been identified nor have their genes been cloned.

The present work focused on isolation and characterization of granule bound starch synthase protein and its genes from maturing grains of common buckwheat.

RESULTS

Marked variations were observed in the shape and size of starch grains isolated from different accessions/cultivars of common buckwheat. The grains ranged in shape from round/spherical to polygonal with a monomodal distribution with in size from 3μm to 12μm. While the grains from VL-7 and IC-13145 showed distinct polygonal shape, those from IC-188669, KBB-3, OC-2, Siva and Daria were round to spherical in shape.

VL-7 and KBB-3 had the largest starch grains which ranged in size from 9.47 μm to 12.1μm. VL-7 is a high yielding and early maturing cultivar from Western Himalayas released by Indian Council for Agricultural Research, New Delhi.

The range of variation in shape and size of starch grains in common buckwheat, as observed in the present investigation, are in agreement with other earlier reports on the shape and size of starch grains in common buckwheat (Kim *et al.*, 1977; Soral-Smietana *et al.*, 1984; Acquistucci and Fornal, 1997; Qian and Kuhn, 1999b). Lindeboom *et al.* (2004), have classified starch grains as large (>25μm), medium (10-25μm), small (5-10μm) and very small (>5μm). Starch grains isolated from the endosperm of buckwheat can be clearly classified under the small grain size category. Starches having small granules and a narrow granule size distribution have found application in fine printing paper and plastic sheets (Jane *et al.*, 1994; Wilhelm *et al.*, 1998), as a binder with orally active ingredients and as a carrier material in cosmetics (Whistler, 1995). Due to their small size, starch grains of common buckwheat may find similar applications.

A distinct feature of starch grains isolated from Indian accessions/varieties was the presence of pores on the surface of the grains. No such pores were visible on the surface of starch grains from European accessions/cultivars. Pores were mostly observed on the smoother surface of starch granules. While, some granules had large number of pores on their surfaces, others had only few pores. Similar observations on the presence of pores in buckwheat starch grains have been made by Qian *et al.* (1998). Fannon *et al.* (1993), have reported the presence of pores on the surface of corn, sorghum and millet

starch granules as well as along the equatorial groove of large granules of wheat, rye and barley. Qian *et al.* (1998), have ascribed the higher susceptibility of buckwheat starch to amylase action to the presence of pores on the surface. Porosity and surface area are important characteristics of solid materials that determine their properties e.g., thermal conductivity, thermal diffusivity, mass diffusion coefficient. Mechanical and textural properties of food are also dependent on porosity (Marousis and Saravacos, 1990). The grains were stacked in compact layers and were covered with a proteinaceous membrane. Scanning electron microscopy of partially digested starch grains showed a clear pattern of concentric rings thereby revealing the lamellar structure of the starch grains. Such lamellar structures have been reported to represent the alternation of semi crystalline and amorphous zone within the matrix (French, 1984; Cameron and Donald, 1992). This alternation of semi crystalline and amorphous zones within the matrix of starch grains has been correlated with the presence of GBSS-I within the grains (Denyer *et al.*, 1995; Smith *et al.*, 1997). Our results on confocal laser scanning microscopy of buckwheat starch clearly indicated the localization of GBSS-I in the form of discrete internal rings within the matrix of starch grains. Han *et al.* (2001), and Han and Hamaker (2002), have shown that the protein was present in the form of discrete internal concentric spheres in starch grains of potato, corn and wheat. Our observations on the localization of GBSS-I in starch grains are consistent with the assumption that amylose synthesis occurs within the core of the starch granules. Irrespective of grain size, the percentage of apparent amylose in buckwheat starch ranged between 47% to 51.9%. The content of amylose in

the grains showed a progressive increase with progressing seed development from milk mature stage to late mature stage.

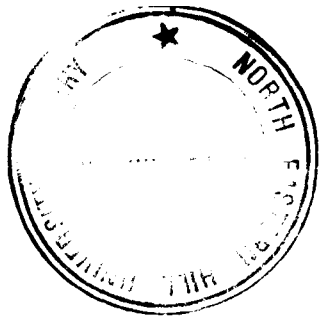
There were no marked differences in the rheological properties of starches isolated from different accessions of buckwheat studied in the present investigation. The starches showed a mean pasting temperature (*P*temp) of 68°C. The peak viscosity, minimum viscosity, breakdown viscosity and final viscosity were 160 RVU, 102 RVU, 58 RVU and 210 RVU respectively. These results are in conformity with the observations of Qian *et al.* (1998). Compared to corn or wheat starches, buckwheat starches have been shown to swell faster, exhibit a greater set back viscosity and form stiffer and harder gels (Campbell, 1995; Zheng *et al.*, 1998; Qian *et al.*, 1998). Besides super molecular glucan structures, the high viscosity values for buckwheat starches can be explained by the fact that the starches exhibited a higher granule swelling and gelling tendency than cereal starches (Yoshimoto *et al.*, 2004). SDS-PAGE profile of proteins associated with starch grains from European varieties/selections of buckwheat indicated varietal differences in the grain proteome amongst different varieties. While the SDS-PAGE proflies of proteins associated with starch grains from some varieties revealed the presence of a single band corresponding to a moleuclar mass of 59.7 kDa, that of other varieties showed the presence of a duplex with molecular masses of 59.7 kDa and 56 kDa. On the other hand, SDS-PAGE profiles of proteins associated with starch granules isolated from the endosperm tissues of Indian accessions/varieties of buckwheat showed the presence of only a single band corresponding to a molecular mass of 59.7 kDa. However, SDS-PAGE profile of the proteins associated with starch grains isolated from

leaves of common buckwheat revealed the presence of a single band corresponding to a molecular mass of 53 kDa. These results are in conformity with similar observations on granule associated proteins of potato and wheat (Takaoka *et al.*, 1997; Han and Hamaker, 2002). Starch granule associated proteins in endosperm of common wheat (*Triticum aestivum* L.) have been reported to include at least one major protein with a molecular weight of 61 kDa and six minor high molecular weight proteins. While the 61 kDa protein was identified as GBSS-I, all other proteins were identified as soluble starch synthase (Takaoka *et al.*, 1997). While the 59.7 kDa protein showed strong crossreactivity with antibodies raised against GBSS-I, the 53 kDa and 56 kDa proteins did not cross react with GBSS-I antibodies. These results indicate absence of serological homology between GBSS-I and the 53 and 56 kDa starch granule associated proteins. The antisera, however, cross reacted with the 61 kDa GBSS-I of maize and the 60 kDa GBSS-I of rice and wheat thereby indicating serological homology between the GBSS-I of wheat, maize, rice and buckwheat. GBSS-I, the key enzyme responsible for amylose synthesis, has been identified as a 56, 58 or a 60 kDa protein in maize (Echt and Schwartz, 1981; Shure *et al.*, 1983; Gibbon *et al.*, 2003), a 60 kDa protein in rice (Sano, 1984), a 58 kDa protein in mungbean (Ko *et al.*, 2009), 68 kDa protein in grain amaranth (Konishi *et al.*, 1985), 59 kDa protein in pea (Dry *et al.*, 1992). Our results are in agreement with the known molecular mass of GBSS-I from other sources.

2D-PAGE of the endosperm starch granule associated proteins resolved the fraction into 10 spots with *pI* ranging from 5.2 to 6.2. All the spots showed an apparent molecular mass of 59.7 kDa. Imunoblotting of the 10 spots separated by 2D-PAGE with

antisera raised against buckwheat GBSS-I identified two bands *viz.*, spot no. 3 (*pI* 5.4) and spot no. 8 (*pI* 6.1) which cross reacted with antibodies raised against buckwheat. Similar results have been reported for *waxy* proteins of maize, rice and barley (Nakamura *et al.*, 1993, 1995; Taira *et al.*, 1995 and Chao *et al.*, 1985). While the protein was detected as a single band of 60 kDa on 1D-PAGE, it resolved into 3 (barley) to 4 (maize, rice) isoforms, with *pI* ranging from 5.8 to 7.2 on 2D-PAGE. The protein corresponding to spot no. 3 was subjected to in-gel trypsin digestion, followed by Mass Spectrometry and database search with MASCOT (www.matrixscience.com). Since, the identification of peptide fragments relies on sequence information present in databases; this is a limitation for buckwheat, which has no sequence information on granule bound protein. Buckwheat proteins were thus identified by comparing them with the rice genomes. However, one of the tryptic fragment having the amino acid sequence "FNAPLAHLIMAGADVLAVPSR" with a predicted *pI* of 8.34 showed similarity with GBSS-I protein of rice.

BLASTp analysis of the N-terminal amino acid sequence for 25 residues of the 59.7 kDa granule associated protein, worked out in the present study, identified the protein as granule bound starch synthase-I. Multiple alignment of the N-terminal sequence of GBSS-I protein isolated from buckwheat with amino acid sequences of similar proteins available in protein data banks clearly identified the conserved KTGGL motif in the sequence. Furukawa *et al.* (1993), have demonstrated the ubiquitous presence of KTGGL domain in all the GBSS-I proteins. This motif has been identified as the ADP/ADPglucose binding site in the enzyme. It is interesting to note that while the



residue immediately preceding and following this motif in soluble starch synthase is usually basic there are no basic residues in the immediate vicinity of this motif in sequence identified in the present study. The sequence showed 94% homology with GBSS-I from *Hordeum vulgaris*, *Triticum* spp. and *Phaseolus vulgaris*. The percentage homology varied between 94% to 88% with GBSS-I. Even though analysis of the sequence alignment revealed a clear diversification into monocotyledonous and dicotyledonous groups. The protein sequence from buckwheat showed similarities with GBSS-I from both the groups. The protein sequence of buckwheat has Valine as the 5th amino acid residue, as is also a case of the protein sequences of GBSS-I from monocots. However, majority of the dicots GBSS-I sequence had Isoleucine at this position. The sequence of buckwheat also showed similarities with sequences from dicots in having Valine as the 11th amino acid residue. GBSS-I protein sequence from monocots have Methionine at this same position. The structural differences may imply differences in catalytic activity of the enzyme. Sequence analysis of buckwheat GBSS-I indicates similarities with both cereal as well as dicot GBSS-I sequences. It is possible that the protein from common buckwheat might belong to a catalytically distinct subclass and hence a possible candidate for altering amylose biosynthesis in both dicots as well as monocots. These results strongly indicate that the 59 kDa protein isolated from starch grains from endosperm tissues of common buckwheat is a GBSS-I class enzyme and hence, an isoforms of the *waxy* protein. This is the first report on the identification of GBSS-I in common buckwheat.

Phylogenetic analysis of the N-terminal amino acid sequence of the GBSS type protein from common buckwheat reported here revealed a clear diversification into monocotyledonous and dicotyledonous groups'. Within the monocots, the sequences could be segregated into two groups. While one of the two groups (Group-I) was dominated by rice (*Oryza* spp.) and maize (*Zea mays*), the other group (Group-II) predominantly comprised sequences from *Triticum* spp, *Hordeum* spp, *Secale cereale*, *Elymus scaber* and *Aegilops speltoides*. The dicotyledons, on the other hand, resolved into three subgroups. While one of the groups comprised of sequences from *Fagopyrum esculentum*, *Nelumbo nucifera*, *Astragalus membranaceus* and *Amaranthus cruentus*, the other group comprised sequences from *Pisum sativum*, *Phaseolus vulgaris* and *Manihot esculenta*. The third group comprised of *Ipomoea batatas* GBSS-I. These results are in conformity with the observations of Edwards *et al.* (2002), who have reported a clear division of GBSS-I proteins into those belonging to monocots and those to dicots.

The modified CTAB protocol used in the present study for isolation of genomic DNA yielded fairly good quality DNA from the etiolated seedlings of common buckwheat. The ratio of absorbance at 260 and 280 nm (A_{260}/A_{280}) for the DNA samples isolated during the present investigation ranged between 1.7 to 1.9 indicating that the isolated DNA was fairly pure. While *EcoR*1 and *Hind*III digested buckwheat genomic DNA fully, *Nco*I could only partially digest buckwheat genomic DNA. The electrophoretic profile of *EcoR*1 and *Hind*III digested DNA revealed a uniform smear ranging in size from 0.5 Kb to 20 Kb with prominent bands indicating the presence of *EcoR*1/*Hind*III repeats of varying lengths in the genomic DNA of common buckwheat.

Similar results have been observed by Bharali (2002), on the restriction digestion profile of buckwheat genomic DNA. Bharali (2002), has reported 6 bands representing *Eco*R1 repeats in genomic DNA of common buckwheat. The appearance of distinct bands in *Eco*R1 and *Hind*III digested gDNA indicates the presence of *Eco*R1, *Hind*III and *Nco*I repeats of varying lengths in buckwheat genomic DNA.

While amplification of gDNA wtih oligonucleotide primer pairs NDF7-NDR6 amplified a DNA fragment of showing an apparent molecular mass of 1.2 Kb that with primer pair NDF8-NDR6 amplified a DNA fragment having an apparent molecular mass of 0.9 Kb. BLASTn analysis of the nucleotide sequences of the amplicons identified the same with genes encoding for granule bound starch synthase gene family. This is the first report of identification of GBSS-I gene in common buckwheat. ClustalW multiple alignment of the nucleotide sequence of 1.2 kb amplicon with nucleotide sequences of GBSS genes of other plants revealed two highly conserved regions. While one of the conserved regions, represented by the sequence "GGACATAGGGTTATGACAGTTGC TCCTCGTTATGATCAGTATAAAGATGGATGGGATACTAATGTACTAGTTCAG ", existed between P'_{94}-P'_{168}, the other represented by the sequence "CAGATACAAG TTGGGGAAAGAGTTGAGACTGTTCGGTTCTTTCACTGCTACAAAAGAGGAGT TGACCGGGTTTTCGTGGATCACCCTATGTTCCTTGAGAAGGT", existed between P'_{274}-P'_{377}. AUGUSTUS (version 2.4) identified four exons and three introns within the nucleotide sequence of 1.2 Kb amplicon and two exons/two introns in the nucleotide sequence of the 0.9 Kb amplicon. Sequence analysis of the 1.2 Kb amplicon clearly revealed the presence of 1st and 2nd conserved domains within the sequence in the 1st and

2nd exon respectively. The sequences at the 5' and 3'ends of the three introns were found to follow the universal GT-AG rule (Breathnach and Chambon, 1981). These results are in agreement with the observations of Xu *et al.* (2009), on the nucleotide sequence of *waxy* gene of rye. All the three introns identified in our study were shown to have AT rich regions. Similiar reports of AT rich regions in the introns of *Oryza glaberrima waxy* gene have been reported by Umeda *et al.* (1991). Camirand *et al.* (1990) and Van der Leij *et al.* (1991), have observed that all introns in the GBSS-I gene also followed the universal GT-AG. In this context, our observations on the intron/exon architecture in the nucleotide sequence of the 1.2 Kb amplicon are in agreement with those of Umeda *et al.* (1991), Camirand *et al.* (1990) and Van der Leij *et al.* (1991).

While the coding region of 604 bases of the nucleotide sequence of 1.2 Kb amplicon showed a maximum of 90% homology with the coding region of GBSS of other plant species, the non-coding regions identified in the sequence did not show any significant homology with GBSS genes. Xu *et al.* (2009), have reported similar results on the sequence homology of rye *waxy* gene with *waxy* genes of other plants. These results indicate a higher level of sequence conservation of GBSS genes in the coding region than the non coding region. Similar results have been reported for GBSS genes in sorghum, rice and maize (Chen *et al.*, 1998).

EMBOSS CpGplot detected a putative CpG Island of 856 bases located between positions 310 to 1116 in the nucleotide sequence of the 1.2 Kb amplicon and of 495 bases located between position 355 to 854 in the nucleotide sequence of the 0.9 Kb amplicon. The ratio of observed to expected CG composition in these segments of the

sequence was > 0.60. Detection of a putative CpG plot is an indication that the gene could be under tight regulatory control. Since CpG islands are known to be associated with the 5' end of genes which are subjected to regulatory control (Larsen *et al.*, 1992), the detection of CpG Island in the DNA amplified in the current study indicates that the region of DNA amplified could belong to the 5' end of the target genes.

NCBI scan detected twelve palindromic structures in the nucleotide sequences of both 1.2 Kb amplicon 0.9 Kb amplicon. Jangsutthivorawat and Holkaert (2009), have correlated the presence of (CT)n and (AATT)n repeats at the GBSS-I loci of rice with the variations in the amylose contenent in the starch. Bao *et al.* (2006b), have also indicated a close correlation beween the nbumber of repeats with the apparent amylose content (AAC) in starch grains of North American and Chinese germplasm. Alleles with fewer repeats (n<12) were observed to be associated with higher AAC and those with more repeats (n>12) with a lower AAC. Identification of such repeats in the partial nucleotide sequence of the buckwheat GBSS-I gene that was identified in the present study could be used as markers for screening of buckwheat accessions for their AAC.

The deduced amino acid sequence for the 1,116bp bases of the 1.2 Kb amplicon comprised of 199 amino acid residues with a predicted isoelectric point (*pI*) of 9.75 and calculated molecular weight of 22.3 kDa. Sequence similarity analysis of the sequence with BLASTp against non-redundant protein database, identified the deduced sequence with the GBSS family. Domain search on the deduced amino acid sequence identified the putative substrate binding site comprising of the sequence Lys-Thr-Gly-Gly-Leu (K-T-G-G-L) between P'_4 and P'_8. KTGGL is a universal motif identified in GBSS-I of

Pisum sativum (Edwards *et al.*, 2002), *Vigna radiata* (Ko *et al.*, 2009), *Solanum tuberosum* (Edwards *et al.*, 1999), *Zea mays* (Harn *et al.*, 1997), grain amaranth (Park *et al.*, 2009), *Triticum aestivum* (Anisworth *et al.*, 1993; Baba *et al.*, 1993) and *Oryza sativa* (Sano *et al.*, 1984). The domain has also been reported to be present in GBSSII of cassava (Munyikwa *et al.*, 1997), SSII of *Pisum sativum* (Edwards *et al.*, 2002), *Solanum tuberosum* (Edwards *et al.*, 1999) and *Zea mays* (Harn *et al.*, 1997). The motif has been suggested to be involved in substrate binding (Furukawa *et al.*, 1990, 1993). Based on affinity labelling and site directed mutagenesis, it was shown that the Lysine residue in conserved motif KTGGL formed part of the ADPG binding site of *E. coli* glycogen synthase (Furukawa *et al.*, 1990, 1993). ADPG is the glucosyl donor in the reaction catalysed by glycogen and starch synthase. Hence, it was assumed that this motif could be the ADPG binding site in the plant enzymes.

Using an alignment that permitted maximum homology, the sequence showed a maximum of 84% homology with deduced amino acid sequences of GBSS-I of *Nelumbo nucifera* (acc.no. ACM78591). The sequence also showed 83%, 80%, 79% and 79% homolgy with deduced amino acid sequences of GBSS-I of *Gossypium hirsutum* (acc. no. ACJ11735), *Phaseolus vulgaris* (acc. no. BAA82346), *Ipomoea batatas* (acc. no. BAI83439) and *Pisum sativum* (acc. no. CAC69955) and 80% homology with GBSSII of *Oryza sativa* (acc. no. ACY56082) and *Triticum aestivum* (acc. no. AAF14233). The deduced amino acid sequence identified in the present study showed an insertion of 9 residues "HLHVLILES" at P'_{17} and the presence of three highly conserved domains between P'_{1}-P'_{16}, P'_{33}-P'_{51} and P'_{66}-P'_{90}. While the conserved domain between P'_{1}-P'_{16}

comprised of 16 residues represented by the sequence "PWSKTGGLGDVLAALP", that between P'_{33}-P'_{51} comprised of 20 residues represented by the sequence "GHRVMTVAPRYDQYKDGWDT". The third conserved domain comprised of 25 residues between P'_{66} to P'_{90}. This domain was represented by the sequence "VRFFHCYKRGVDRVFVDHPMFLEKV". ClustalW multiple alignment also revealed a high level of sequence similarity within the first 95 amino acid residues aligned out of the total 119 deduced amino acid residues. This sequence showed similarity with the glycoslytransferase enzmyes which catalyze the transfer of a monosaccharide unit from an activated nucleotide sugar to a glycosyl acceptor molecule forming glycosidic bonds in carbohydrate residues or other biopolymers (Breton *et al.*, 2006).

Motif search on the deduced amino acid sequence identified a starch synthase catalytic domain predominantly represented by the amino acid residues "PWSKTGGLG DVLAALPHLHVLILESPALAARGHRVMTVAPRYDQYKDGWDTNVLVQIQVGE RVETVRFFHCYKRGVDRVFVDHPMFLEKVTGPLGYLEEHMIKIVIFRRVTTTTFA SGTQICIVRGGAGRKGVMAY" between positions 1-136. The software also detected three protein kinase C phosphorylation sites represented by the sequences Thr-Val-Arg ($P'_{65\text{-}67}$), Thr-Gly-Lys ($P'_{141\text{-}143}$) and Thr-Gly-Arg ($P'_{176\text{-}178}$). When the deduced amino acid sequence were plotted as a function of hydropathic index, the sequence showed a predominantly hydrophobic character. Based on the hydropathic index of Kyte and Doolittle (1982), the major regions of hydrophobic nature detected in the sequence were between residues 5-28, 50-57, 77-95 and 102-139. Wang *et al.* (2000), have reported similar results on the hydrophobicity of *Ipomoea batatas* GBSS-I. Statistical analysis by

SAPS (Brendel *et al.*, 1992) predicted the sequence to have 28.1% nonpolar residues, 18.6% polar uncharged residues and 22.6% polar charged residues.

Phylogenetic analysis of the deduced amino acid sequence reported in the present study with amino acid sequences of granule bound starch synthases available in EMBL database revealed a clear diversification into monocotyledonous (GBSS-II) and dicotyledonous (GBSS-I) groups. According to Pan *et al.* (2009), the putative protein could have duplicated and diverged into two different forms *viz.*, GBSS-I or GBSSIa and GBSS-II or GBSSIb during evolution and the diversification into monocotyledonous (GBSS-II) and dicotyledonous (GBSS-I) groups might be a consequence of this diversification process.

The deduced amino acid sequence for the 896 bases of the 0.9 Kb amplicon comprised of 70 amino acid residues with a predicted isoelectric point (*pI)* of 6.05 and a calculated molecular weight of 8.0 kDa. Sequence similarity analysis of the sequence with BLASTp against non-redundant protein database, identified the deduced sequence with the GBSS family. Using an alignment that permitted maximum homology, the deduced amino acid sequence showed a maximum of 96% homology with GBSS-I of *Ipomoea batatas* (acc. no. BAI83439), *Ipomoea umbraticola* (acc. no. ABW83774) and *Solanum nemorense* (acc. no. AAY63641) with a query coverage of 100%. The sequence also showed 95% homology with *Nicotiana tabacum* (acc. no. AAZ99063), *Hyoscyamus niger* (acc. no. AAZ99051), *Atropa belladonna* (acc. no. AAZ99047), *Jaborosa squarrosa* (acc. no. ABM46904), *Jaltomata procumbens* (acc. no. AAY63568), *Anisodus luridus* (acc. no. AAZ99048) and *Solanum johnstonni* (acc. no.

ACV96018) and 93% homology with *Przewalskia tangutica* (acc. no. AAZ99055) respectively. BLASTp analysis of the sequence indicated that the entire 70 amino acid residues belong to the glycosyltransferase superfamily. Motif search on the deduced amino acid sequence identified, the entire sequence as the catalytic domain of starch synthase.

A phylogenetic tree, describing the relationship of the deduced amino acid sequence reported in the present study with amino acid sequences of granule bound starch synthase proteins from other plants, was constructed by maximum parsimony method using the alignment matrix. The deduced amino acid sequence of 70 residues clustered together into one clad with amino acid sequences of *Ipomoea batatas* (acc. no. BAI83439) and *Ipomoea umbraticola* (acc. no. ABW83774).

Results obtained in the present study have revealed that the starch grains of common buckwheat showed a monomodal size distribution with size ranging from 3μm to 12μm. The grains can be clearly classified under the small grain size category and can thus have important applications in industrial processes. SEM of partially digested starch grains showed a clear pattern of concentric rings thereby revealing the lamellar structure of starch grains. This alternation of semi crystalline and amorphous zones within the matrix of starch grains has been correlated with the presence of GBSS-I within the grains. Irrespective of grain size, the percentage of apparent amylose in buckwheat starch ranged between 47% to 51.9%. The starches showed a mean pasting temperature (*P*temp) of 68°C. Our results have revealed that a 59.7 kDa protein was associated with buckwheat starch grains as a granule bound protein. Based on N-terminal amino acid as

well as serological analyses, the protein has been identified as GBSS-I. This is the first report on the identification of GBSS-I in common buckwheat. Investigations carried out in the present study have led to the amplification and characterization of partial nucleotide sequences of GBSS-I genes from common buckwheat. The sequences showed features which were common to GBSS-I genes from dicots as well as monocots.

CHAPTER VII

REFERENCES

Acquistucci, R. and Fornal, J. (1997) Italian buckwheat (*Fagopyrum esculentum*) starch: Physico-chemical and functional characterization and *in vitro* digestibility. Food/Nahrung., 41(5): 281-284.

Ahamed, N. T., Singhla, R. S., Kulkarni, P. R., Kale, D. D. and Pal, M. (1996) Studies on chenopodium quinoa and amaranthus paniculatatas starch as biodegradable fillers in LDPE films. Carbohydrate Polymers., 157-160.

Ainsworth, C., Clark, J. and Balsdon, J. (1993) Expression, organisation and structure of the genes encoding the *waxy* protein (granule-bound starch synthase) in wheat. Plant Molecular Biology., 22: 67-82.

Amani, N. G., Buleon, A., Kamenan, A. and Colonna, P. (2004) Variability in starch physicochemical and functional properties of yam (*Dioscorea* sp.) cultivated in Ivory Coast. Journal of the Science of Food and Agriculture., 84: 2085–2096.

Anonymous. (1987-2001) Annual Report, NBPGR Regional Station, Barapani, Meghalaya.

Appelqvist, I. A. M. and Debet, M. R. M. (1997) Starch-biopolymer interactions—A review. Food Reviews International., 13: 163-224.

Armstrong, S. A., Barry, D. A., Leggett, R. W. and Mueller, C. R. (1997) Casein Kinase II-mediated phosphorylation of the C terminus of Sp1 decreases its DNA binding activity. Journal of Biological Chemistry., 272: 13489-13495.

Asp, N. G., Van Amelsvoort, J. M. M. and Hautvast, J. G. A .J. (1996) Nutritional implications of resistant starch. Nutrition Research Review., 9: 1–31.

Autio, K., Poutanen, K., Suortti, T. and Pessa, E. (1992) Heat-induced structural changes in acid-modified barley starch dispersions. Food Structure., 11: 315-322.

Ayres, N. M., McClung, A. M., Larkin, P. D., Bligh, H. F. J., Jones, C. A. and Park, W. D. (1997) Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germplasm. Theoretical and Applied Genetics., 94: 773–781.

Baba, T., Nishihara, M. K., Kawasaki, T., Shimada, T., Kobayashi, E., Ohnishi, S., Tanaka, K. and 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. G. and Morell, M. K. (2003) From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. Annual Review of Plant Biology., 54: 207–233.

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

Bancel , E., Rogniaux, H., Debiton, C., Chambon, C. and Branlard, G. (2010) Extraction and proteome analysis of starch granule-associated proteins in mature wheat kernel (*Triticum aestivum* L.). Journal of Proteome Research., 9(6): 3299-3310.

Banks, W. and Greenwood, C. T. (1975) Fractionation of the starch granule, and the fine structure of its components. In: Starch and Its Components. W. Banks, and C. T. Greenwood, eds. Edinburgh University Press: Edinburgh., pp 5-66.

Bao, J. and Bergman, C. (2004) The functionality of rice starch. In: Starch in Food: Structure, Function and Applications. A.-C. Eliasson, ed. Woodhead Publishing: Cambridge, UK., pp. 152-184.

Bao, J. S., Corke, H. and Sun, M. (2006b) Microsatellites, single nucleotide polymorphism and a sequence tagged site in starch-synthesizing genes in relation to starch physicochemical properties in nonwaxy rice (*Oryza sativa* L.). Theoretical and Applied Genetics., 113: 1185-1196.

Batey, I. L., Gras, P. W. and Curtin, B. M. (1997) Contribution of the chemical structure of wheat starch to Japanese noodle quality. Journal of the Science and Food Agriculture., 74: 503-508.

Bergman, C. J., Delgado, J. T., Mcclung, A. M. and Fjellstrom, R. G. (2001) An improved method for using a microsatellite in rice *waxy* gene to determine amylose class. Cereal Chemistry., 78: 257-260.

Berkelman, T. and Stenstedt, T. (2002) 2-D Electrophoresis using immobilized pH gradients. Principles and Methods. Amersham Biosciences, Uppsala, Sweden.

Bradley, B. (2003) The granule bound starch synthase genes of wheat. Ph. D. Thesis. Western Australian State Agricultural Biotechnology Centre.School of Biological Sciences and biotechnology. Murdoch University.

Berry, C. S. (1986) Resistant starch. Formation and measurement of starch that survives exhaustive digestion with amylolytic enzymes during the determination of dietary fiber. Journal of Cereal Sciences., 4: 301–14.

Bharali, S. (2002). Isolation, cloning and molecular analysis of the legumin gene of common buckwheat (*Fagopyrum esculentum* Moench.). Doctoral thesis, North Eastern Hill University, Shillong, India.

Biliaderis, C. G., Mazza, G. and Przybylski, R. (1993) Composition and physic-chemical properties of starch from cow cockle (*Saponaria vaccaria* L.) seeds. Starch/Strake., 45: 121- 127.

Bligh, F. J. (1999) Genetic manipulation of starch biosynthesis: Progress and potential; Biotechnology and Genetic Engineering Reviews., 16: 177-201.

Bligh, F. J., Till, R. I. and Jones, C. A. (1995) A microsatellite sequence closely linked to the *waxy* gene of *Oryza sativa.* Euphytica., 86: 83–85.

Boren, M. M., Larson, H., Falk, A. and Jansson, C. (2004) The barley starch granule proteome-internalized granule polypeptides of the mature endosperm. Plant Science., 166: 617-626.

Boyer, C. D. and Preiss, J. (1978). Multiple forms of (1-4)-α-D-glucan, (1-4)-α-D-glucan-6-glycosyl transferase from developing *Zea mays* L. kernel. Carbohydrate Research., 61: 321-334.

Breathnach, R. and Chambon, P. (1981) Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochemistry., 50(1): 349-283.

Brendel, V., Bucher, P., Nourbakshsh, I., Blaisdell, B. E. and Karlin, S. (1992) Methods and algorithms for statistical analysis of protein sequences. Proceedings of the National Academy of Sciences. U.S.A., 89: 2002-2006.

Breton, C., Snajdrova, L., Jeanneau, C., Koca, J. and Imberty, A. (2006) Structures and mechanisms of glycosyltransferases. Glycobiology, 16: 29R-37R.

Briney, A., Wilson, R., Potter, R. H., Crosbie, G. and Jones, M. G. K. (1997) A molecular marker of the selection of starch and potential noodle quality in wheat. In: Cereals '96: Proc. 46th Australian Cereal Chemistry Conference. C. W. Wrigley, ed. RACI: Melbourne, Australia., pp: 133-137.

Buckwheat diet. AGWEEK News, Aug. 14, 1989.

Cairns, P., Morris, V. J., Botham, R. L. and Ring, S. G. (1996) Physicochemical studies on resistant starch *in vitro* and *in vivo*. Journal of Cereal Sciences., 23: 265-275.

Cameron, R. E. and Donald, A. M. (1992) A small-angle X-ray scattering study of the annealing and gelatinization of starch. Polymer., 3 (12): 2628-2625.

Camirand, A., St-Pierre, B., Marineau, C. and Brission, N. (1990) Occurance of a copia-like transposable element in one of the introns of the potato starch phosphorylase gene. Molecular Genetics and Genomics., 224: 33-39.

Campbell, C. G. (1997) Buckwheat (*Fagopyrum esculentum* Moench). Promoting the conservation and use of underutilized and neglected crops. 19. International Plant Genetic Resources Institute, Rome, Italy.

Chao, S., Sharp, P. J., Worland, A. J., Warham, E. J., Koebner, R. M. D. and Gale, M. D.(1989). RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. Theoretical and Applied Genetics., 78: 495.

Chen, M., SanMiguel, P. and Bennetzen, J. L. (1998) Sequence organization and conservation in Sh2/a1-homologous regions of sorghum and rice. Genetics., 148(1): 435-443.

Chen, Z., Schols, H. A. and Voragen, A. G. J. (2003) Physicochemical properties of starches obtained from three different varieties of chinese sweet potatoes. Journal of Food Science., 68: 431-437.

Chiotelli, E. and Meste, L. M. (2002) Effect of Small and Large Wheat Starch Granules on thermomechanical Behavior of Starch. Cereal Chemistry., 79: 286-293.

Chrungoo, N. K., Bharali, S. and Jamir, C. (2005) Seed storage proteins: Structure, properties and approaches for improvement by genetic engineering. Advances in Plant Physiology., 8: 201-220.

Clark, J. R., Robertson, M. and Ainsworth, C. C (1991) Nucleotide sequence of a wheat (*Triticum aestivum* L.) cDNA clone encoding the *waxy* protein. Plant Molecular Biology., 16: 1099-1101.

Craig, J., Lloyd, J. R., Tomlinson, K., Barber, L., Edwards, A., Wang, T. L., Martin, C., Hedley, C. L. and Smith, A. M. (1998) Mutations in the gene encoding starch synthase II profoundly alter amylopectin structure in pea embryos. Plant Cell., 10: 413–426.

Dang, J, M. C. (2004) Genotype and Environmental Influences on Pasting Properties of Rice Flour, Les Copeland. Cereal Chemistry., 81(4): 486.

Dang, P. L. and Boyer, C. D. (1988) Maize leaf and kernel starch synthases and starch branching enzymes. Phytochemistry., 27: 1255-1259.

Dang, Q., Alghisi, G. C. and Gasser, S. (1994) Phosphorylation of the C-terminal domain of yeast topoisomerase II by casein kinase II affects DNA- protein interaction. Journal of Molecular Biology., 243: 10-24.

Daniel, J. R. and Whistler, R. L. (1990) Fatty sensory qualities of polysaccharides. Cereal Food world., 35: 825.

David, G. S., Russel, K. D., Jay-lin, J. and George, E. I. (2006) Structures and functional properties of starch from seeds of three soyabean (*Glycine max* (L.) Merr.) varieties. Starch/Starke., 58: 509-519.

Davis, E. A. (1994) Wheat Starch. Cereal Foods World., 39 (1) : 34-36.

Dedeh, S. and Sackey, K-A. E. (2002) Starch structure and some properties of cocoyam (*Xanthosoma sagittifolium* and *Colocasia esculenta*) starch and raphides. Food Chemistry., 79: 435-444.

Delrue, B., Fontaine, T., Routier, F., Decqm, A., Wieruszeski, J. M. and Ball, S. (1992) Waxy *Chlamydomonas reinhardtii*: monocellular algal mutants defective in amylose biosynthesis and granule-bound starch synthase activity accumulate a structurally modified amylopectin. Journal of Bacteriology.,174: 3612–3620.

Dengate, H. and Meredith, P. (1984) Variation in size distribution of starch granules from wheat grain. Journal of Cereal Science., 3: 83-90.

Denyer, K., Sidebottom, C., Hylton, C. M. and Smith, A. M. (1993) Soluble isoforms of starch synthase and starch-branching enzyme also occur within starch granules in developing pea embryos. The Plant Journal., 4: 191-198.

Denyer, K., Barber, L. M., Burton, R., Hedley, C. L., Hylton, C. M., Johnson, S., Jones, D. A., Marshall, J., Smith, A. M. and Tatge, H. (1995) The isolation and characterization of novel, low-amylose mutants of *Pisum sativum* L. Plant Cell and Environment., 18: 1019-1026.

Denyer, K., Clarke, B., Hylton, C., Tatge, H. and Smith, A. M. (1996) The elongation of amylose and amylopectin chains in isolated starch granules. Plant Journal., 10: 1135–1143.

Denyer, K., Barber, L. M., Edward, E. A., Smith, A. M. and Wang, T. L. (1997) Two isoforms of the GBSS-I class of granule-bound starch synthase are differentially expressed in the pea plant (*Pisum sativum* L.) Plant Cell and Environment., 20: 1566-1572.

Denyer, K., Johnson, P., Zeeman, S. and Smith, A. M. (2001) The control of amylose synthesis. Journal of Plant Physiology., 158: 479-487.

Deschamps, P., Moreau, H., Worden, A. Z., Dauville'e, D. and Ball, S. G. (2008) Early gene duplication within *Chloroplastida* and its correspondence with relocation of starch metabolism to chloroplasts. Genetics., 178: 2373–2387.

Dian, W. M., Jiang, H. W., Chen, Q. S., Liu, F. Y. and Wu, P. (2003) Cloning and characterization of the granule-bound starch synthase II gene in rice: gene expression is regulated by the nitrogen level, sugar and circadian rhythm. Planta., 218: 261-268.

Domon, E., Fujita, M. and Ishikawa, N. (2002) The insertion/ deletion polymorphisms in the *waxy* gene of barley genetic resources from East Asia. Theoretical and Applied Genetics., 104: 132-138.

Dry, I., Smith, A. M., Edwards, E. A., Bhattacharyya, M., Dunn, P. and Martin, C. (1992) Characterization of cDNAs encoding two isoforms of granule bound starch synthase which show differential expression in developing storage organs. Plant Journal., 2: 193-202.

Echt, C. S. and Schwartz. D. (1981) Evidence for the inclusion of controlling elements within the structural gene at the waxy locus in maize. Genetics., 99: 275-284.

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

Edwards, A., Vincken, J. P., Suurs, L. C., Visser, R. G., Zeeman, S., Smith, A. and Martin, C. (2002) Discrete forms of amylose are synthesized by isoforms of GBSSI in pea. The Plant Cell., 14: 1767-1785.

Eliason, A. C. and Larsson, K. (1983) Physiochemical behavior of the components of wheat flour, in Cereals in Breadmaking. A Molecular Colloid Approach, (Ed. O. R. Fennema) Marcel Dekker, New York .U.S.A., pp. 325 - 370.

Eliason, A. C. and Gudmundsson, M. (1996) Starch: Physicochemical and functional aspects. In A.-C. Eliasson (Ed.), carbohydrates in food New York, N.Y: Marcel Dekker., pp. 431–503.

Eliason, A. C. (2004) Starch in food :Structure, function and applications. Woodhead Publishing Limited Abington Hall, Abington Cambridge CB1 6AH England.

Englyst, H. N., Wiggins, H. S. and Cummings, J. H. (1982) Determination of the non-starch polysaccharides in plant foods by gas-liquid chromatography of constituent sugars as alditol acetates. Analyst., 107: 307–18.

Evans, R. C., Alice, L. A., Campbell, C. S., Kellogg, E. A. and Dickinson,T. A. (2000) The granule-bound starch synthase (GBSS-I) gene in the Rosaceae: multiple loci and phylogenetic utility. Molecular Phylogenetics and Evolution., 17: 388–400.

Evans, R. C. and Campbell, C.S. (2002) The origin of the apple sub-family (Maloideae; Rosaceae) is clarified by DNA sequence data from duplicate GBSS-I genes. American Journal of Botany., 89: 1478–1484.

Evers, A. D. (1971) Scanning electron microscopy of wheat starch. III. Granule development in the endosperm. Starch., 23: 157.

F. A. O. (1998) Carbohydrates in Human Nutrition: Report of a Joint FAO/WHO Expert Consultation, Rome. F.A.O. Food and Nutrition Paper No. 66., pp. 14-18.

Fannon. J. E., Shull, J. M and BeMiller, J. N. (1993) .Interior channels of starch granules. Cereal Chemistry., 70: 611-613.

Fegurson, V. (1994) High amylose and waxy corns, in Specialty Corns (Ed. A.R. Hallauer) CRC Press, Bpca Raton, pp. 55-77.

Flipse, E., Keetels, C. J. A. M., Jacobsen, E. and Visser, R. G. F. (1996) The dosage effect of the wild type GBSS allele is linear for GBSS activity but not for amylose content: Absence of amylose has a distinct influence on the physico-chemical properties of starch. Theoretical and Applied Genetics., 92: 121-127.

Fredriksson, H., Silverio. J., Andersson, R., Eliason, A. C. and Aman. P. (1998) The influence of amylose and amylopectin characteristics on gelatinization and retrogradation properties of different starches. Carbohydrate Polymers., 35: 119-134.

French, D. (1984) Organisation of starch granules. In R.L. Whistler, J. N. Bemiller and Paschal, E. F (Eds.) (2nd ed). Starch chemistry and Technology, New York: Academic Press., pp. 182-247.

Fujita, N., Wadano, A., Kozaki, S., Takaoka, K., Okabe, S. and Taira, T. (1996) Comparison of the primary structure of *waxy* proteins (granule bound starch synthase) between polyploid wheats and related diploid species. Bichemical Genetics., 34: 403-413.

Fukunaga, K., Kawase, M. and Kato, K. (2002) Structural variation in the Waxy gene and differentiation in foxtail millet [*Setaria italica* (L.) P. Beauv.]: implications for multiple origins of the waxy phenotype. Molecular Genetics and Genomics., 268: 214-222.

Fulton, D. C., Edwards, A., Pilling, E., Robinson, H. L., Fahy, B., Seale, R., Donald, A. M., Geigenberger, P., Martin, C. and Smith, A. M. (2002) Role of granule bound starch synthase in determination of amylopectin structure and starch granule morphology in potato. Journal of Molecular Biology., 277: 10834–10841.

Furukawa, K., Tagaya, M., Inouye, M., Preiss, J. and Fukui, T. (1990) Identification of lysine 15 at the active site in *Escherichia coli* glycogen synthase. Conservation of Lys-X-Gly-Gly sequence in the bacterial and mammalian enzymes. Journal of Biological Chemistry., 265: 2086–2090.

Furukawa, K., Tagaya, M., Tanizawa, K. and Fukui, T. (1993) Identification of Lys227 at the active site of *Escherichia coli* glycogen synthase. Journal of Biological Chemistry., 268: 23837-23842.

Garnier, J., Gibrat, J-F. and Robson, B. (1996) GOR method for predicting protein secondary structure from amino acid sequence. Methods in Enzymology., 2006: 540-543.

Gibbon, B. C., Wang, X. and Larkins, B. A. (2003) Altered starch structure is associated with endosperm modification in quality protein maize. Proceedings of the National Academy of Sciences., U. S. A., 100: 15329-15334.

Goering, K. J. and DeHaas, B. (1972) New starches. VIII. Properties of the small granule-starch from Colocasia esculenta. American Association of Cereal Chemistry., pp. 712-719.

Gould, K. R. and Donoghue, M. J. (2000) Phylogeny and biogeography of Triosteum (Caprifoliaceae). Harvard Papers in Botany., 5: 157–166.

Graybosch, R. A. (1998) Waxy wheats: Origin, properties, and prospects. Trends in Food Science and Technology., 9: 135-142.

Grimaud, F., Rogniaux, H., James, M. G., Myers, A.M. and Planchot .V. (2008) Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis. Journal of Experimental Botany., 59 (12): 3395-3406.

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

Hall, D. M. and Sayre, J. G. (1973) A comparison of starch granules as seen by both scanning and ordinary light microscopy. Starch-Stärke., 25: 292–297.

Han, X. Z., Campanellae, O. H. and Hamaker, B. R. (2001) Location of granule bound starch synthase (GBSS) in starch granules and significance of GBSS on the rheological properties of starch pastes. Proceedings of the IFT annual meeting, New Orleans, Louisiana.

Han, X. Z. and Hamaker, B. R. (2002) Association of starch granule proteins with starch ghosts and remnants revealed by confocal laser scanning microscopy. Cereal Chemistry., 79: 892-896.

Harn, C., Knight, M. E., Ramakrishnan, A., Guan, H. P., Keeling, P. L. and Wasserman, B. P. (1997) Isolation and characterization of the zSSIIa and zSSIIb starch synthase cDNA clones from maize endosperm. Plant Molecular Biology., 37: 639-649.

Hirano, H. and Sano, Y. (1991) Molecular characterization of the *waxy* locus of rice *(Oryza sativa).* Plant Cell Physiology., 32: 989-997.

Hirano, H. Y., Eiguchi, M. and Sano, Y. (1998) A single base change altered the regulation of the waxy gene at the post-transcriptional level during the domestication of rice. Molecular Biology Evolution., 15: 978-987.

Hirose, T. and Terao, T. (2004) A comprehensive expression analysis of the starch synthase gene family in rice (*Oryza sativa* L.). Planta., 220: 9–16.

Hizukuri, S. (1986) Polymodal distribution of the chain lengths of amylopectins, and its significance. Carbohydrate Research., 147: 342-347.

Hoffmann, R., Craik, D. J., Pierens, G., Bolger, R. E. and Otvos, Jr. L. (1998) Phosphorylation of the C-terminal site of human p53 reduces non-sequence-specific DNA binding as modeled with synthetic peptides. Biochemistry., 37: 13755-13764.

Hoover. R., Smith, C., Zhou, Y. and Ratnayake. R. M. W. S. (2003) Phy-sicochemical properties of Canadian oat starches. Carbohydate Polymers., 52:253-261.

Hovenkamp-Hermelink, J. H. M., Jacobsen, E., Ponstein, A. S., Visser, R. G. F., Vos-Scheperkeuter, G. H, Bijmolt, E. W., de Vries, J. N., Witholt, B. and Feenstra, W. J (1987) Isolation of an amylose-free starch mutant of the potato (*Solanum tuberosum* L.). Theoretical and Applied Genetics., 75: 217-221.

Hsieh, J., Liu, C. and Hsing, Y. C (1996) Molecular cloning of a sorghum cDNA (acc. no. U23945) encoding the seed *waxy* protein (PGR96-119). Plant Physiology., 112: 1735.

Hylton, C. M., Denyer, K., Keeling, P. L., Chang, M .T. and Smith, A. M. (1996) The effect of *waxy* mutations on the granule-bound starch synthases of barley and maize endosperms. Planta., 198: 230-237.

Imberty, A., Buléon, A., Tran, V. and Peréz, S. (1991). Recent advances in knowledge of starch structure. Starch/Staerke 43: 375-384.

Ingram, A. L. and Doyle, J. J., (2003). The origin and evolution of *Eragrostis tef* (Poaceae) and related polyploids: evidence from nuclear *waxy* and plastid *rps*16. American Journal of Botany., 90: 116–122.

Isshiki, M., Morino, K. M., Nakajima, M., Okagaki, R. J., Wessler, S. R., Izawa, T. and Shimamoto, K. (1998). A naturally occurring functional allele of the rye waxy locus has a GT to TT mutation at the 5' splice site of the first intron. Plant Journal., 15: 133-138.

James, M. G., Denyer, K. and Myers, A. M. (2003) Starch synthesis in the cereal endosperm .Current Opinion in Plant Biology., 6: 215-222.

Jane, J. L. and Chen, F. J. (1992) Effect of amylose molecular size and amylopectin branch chain length on paste properties of starch. Cereal Chemistry., 69: 60-65.

Jane, J. L., Kasemsuwan, T., Leas, S., Zobel, H. and Robyt, J.F. (1994) Anthology of starch granule morphology by scanning electron microscopy. Starch/Strake., 46: 121-129.

Jane, J., Shen, L., Lim, S., Kasemsuwan, T. and Nip, W. K. (1992) Physical and chemical studies of taro starches and flours. Cereal Chemistry., 69: 528-535.

Jane, J., Chen, Y.-Y., Lee, L. F., McPherson, A. E., Wong, K.-S., Radosavljevic, M. and 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.

Jangsutthivorawat, S. and Volkaert, H. (2009) Nucleotide polymorphism of starch synthesis genes in Thai rice. The Kasetsart Journal (Natural Science)., 43: 154-163.

Jayakody, L., Hoover, R., Liu, Q. and Donner, E. (2007) Studies on tuber starches. II. Molecular structure, composition and physicochemical properties of yam (*Dioscorea sp*) starches grown in Sri Lanka. Carbohydrate Polymers., 69: 148-163.

Jenkins, P. J., Cameron, R. E. and Donald, A. M. (1993) A universal feature in the structure of starch granules from different botanical sources. Starch-Starke., 45: 417–420.

Joshi, B. D. and Paroda, R. S. (1991) Buckwheat in India. NBPGR, Shimla Science Monograph. No. 2 pp. 1-117.

Joshi, B. D. (1999) Status of buckwheat in India. Fagopyrum., 16: 7-11.

Juliano, B. O. (1984) Rice starch: Production, properties, and uses.In: Starch Chemistry and Technology. R. L. Whistler, J. N.BeMiller, and E. F. Paschall, eds. Academic Press: Orlando, FL., pp. 507-528.

Keeling, P. L., Banisadr, R., Barone, L., Wasserman, B. P. and Singletary, G. W. (1994) Effect of temperature on enzymes in the pathway of starch biosynthesis in

developing wheat and maize grain. Australian Journal of Plant Physiology., 21: 807-827.

Keeling, P. L., Wood, J. R., Tyson, R. H. and Bridges, I. G. (1998) Starch biosynthesis in developing wheat grain. Plant Physiology., 87: 311-319.

Kim, S., Hahn, T. R., Kwon, T. W. and D'Appolpnia, B.L. (1977) Physicochemical properties of buckwheat starch. Korean Journal of Food Science and Technology., 9: 138-143.

Klosgen, R. B., Gierl, A., Schwartz-Sommer, Z. and Seadler, H. (1986) Molecular analysis of the *waxy* locus of maize. Molecular Genetics and Genomics., 203: 237-244.

Ko, Y. T., Dong, Y. L., Hsieh, Y. F. and Kuo, J. C. (2009) Morphology, associated protein analysis and identification of 58 kDa starch synthase in Mungbean (*Vigna radiate* L. cv. KPS1) starch granule. Journal of agriculture and food chemistry., 57 (10): 4426–4432.

Konishi ,Y., Nojima, H., Okuno, K., Asaoka, M. and Fuwa, H.(1985) Characterization of starch granules from *waxy*, *nonwaxy* and hybrid seeds of *Amaranthus hypochondriacus* L. Agriculture and Biological Chemistry., 49: 1965-1971.

Kossmann, J. and Lloyd , J. (2000) Understanding and influencing starch biochemistry. Critical Reviews in Plant Sciences., 19: 171–226.

Kreft, S., Strukelj, B.,Gaberscik, A. and Kreft, I. (2002) Rutin in buckwheat herbs grown at different UV-B radiation levels: Comparison of two UV spectrophotometric and HPLC method. Journal of Experimental Botany., 53 (375): 1801-1804.

Kumar, S., Dudley, J., Nei, M. and Tamura, K. (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Briefings in Bioinformatics., 9: 299-306.

Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology., 157(1): 105-32.

Laemmli, K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature., 22: 680-685.

Larkin, P. D. and Park, W. D. (1999) Transcript accumulation and utilization of alternate and non-consensus splice sites in rice granule- bound starch synthase are

temperature- sensitive and controlled by a single-nucleotide polymorphism. Plant Molecular Biology., 40: 719-727.

Larsen, F., Gundersen, G., Lopez, R. and Prydz, H. (1992) CpG islands as gene markers in the human genome. Genomics., 13(4): 1095-1107.

Leterrier, M., Holappa, L. D., Broglie, K. E. and Beckles, D. M. (2008) Cloning, characterization and comparative analysis of a starch synthase IV gene in wheat: functional and evolutionary implications. BCM Plant Biology., 8: 98-119.

Levin, R. A., Watson, K. and Bohs, L. (2005) A four-gene study of evolutionary relationships in Solanum section Acanthophora. American Journal of Botany., 92: 603–612.

Levin, R. A., Myers, N. R. and Bohs, L., (2006). Phylogenetic relationships among the "Spiny Solanum" (Solanum subgenus Leptostemonum, Solanaceae). American Journal of Botany., 93: 157–169.

Li, J. and Donoghue, M. J. (1999) More molecular evidence for interspecific relationships in Liquidambar (Hamamelidaceae). Rhodora.,101: 87–91.

Lim, T. S., Jane, J. L., Rajagopalan, S. and Seib, P.A. (1992) Effect of starch granule size on physical properties of starch –filled polyethylene films. Biotechnology Progress., 51-57.

Lindeboom, N., Chang, P. R. and Tyler, R. T. (2004) Analytical, Biochemical and Physicochemical Aspects of starch granule size, with emphasis on small granule starch: A Review. Starch/Starke., 56: 89-99.

Lorenz, K. (1990) Quinoa (*Chenopodium quinoa*) starch: Physiochemical properties and functional characteristics. Starch 42: 81-86.

Lu, T-J., Chen, J-C., Lin, C-L. and Chang, Y-H. (2005) Properties of cocoyam (*Xanthosoma sagittifolium*) tubers planted in different seasons. Food Chemistry., 91: 69-77.

Macdonald. F. D and Preiss, J. (1985) Partial purification and characterization of granule bound starch synthases from normal and waxy maize. Plant Physiology., 78(4): 849-852.

MacGregor, A. W. and Bhatty, R. S. (1996) Barley: Chemistry and Technology. American Association of Cereal Chemists Inc., St. Paul, M.N., pp. 355-417.

Maddelein, M. L., Libessart, N., Bellanger, F., Delrue, B., D'Hulst, C. and Ball, S. (1994) Towards an understanding of the biogenesis of the starch granule: determination of granule-bound and soluble starch synthase functions in amylopectin synthesis. Journal of Biological Chemistry., 269: 25150–25157.

Mäkelä, M. J. and Laakso, S. (2006) Studies on Oat Starch with a Celloscope: Granule Size and Distribution. Starch-Starke., 36(5): 159-163.

Malinski, E., Daniel, J. R., Zhang, X. X. and Whistler, R.L. (2003) Isolation of small granules and determination of their fat mimic characteristics. Cereal Chemistry., 80: 1- 4.

Manners, D. J. (1979) The enzymatic degradation of starches. In J.M.V. Blanshard and J.R. Mitchell, Polysaccharides in Food London :Butterworths., pp. 5-91.

Marousis S. N. and Saravacos G.D. (1990) Density and porosity in drying starch materials. Journal of Food Science., 55(5): 1367-1372.

Mason-Gamer, R. J. and Kellogg, E. A. (1996) Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). Systematic Biology., 45: 524-545.

Mason-Gamer, R. J., Weil. C. F. and Kellogg, E. A. (1998) Granule bound starch synthase: structure, function and phylogenetic utility. Molecular Biology and Evolution., 15: 1658-1673.

Mason-Gamer, R. J. (2001) Origin of the North American Elymus (Poaceae: Tritiaceae) allotetraploids based on granule-bound starch synthase gene sequences. Systematic Botany., 26: 757–768.

McGuffin, L. J., Bryson, K. and Jones, D. T. (2000) The PSIPRED protein structure prediction server. Bioinformatics., 16: 404-405.

Miller, R. E., Rausher, M. D. and Manos, P. S. (1999) Phylogenetic systematic of *Ipomoea* (Convolvulaceae) based on ITS and *Waxy* sequences. Systematic Botany., 24: 209–227.

Morell, M. K., Rahman, S., Abrahams, S. L., and Appels, R. (1995) The biochemistry and molecular biology of starch synthesis in cereals. Australian Journal of Plant Physiology., 22: 647-660.

Morell, M. K. and Myers, A. M. (2005) Towards the rational design of cereal starches. Current Opinion in Plant Biology., 8: 204-210.

Morrison, W. R. (1981) Starch lipids: A reappraisal. Starch/Staerke., 33: 408-410.

Morrison, W. R. and Scott, D. C. (1986) Measurements of the dimensions of wheat starch granule populations using a Coulter Counter with 100-channel analyzer. Journal of Cereal Science., 4: 13-21.

Morrison, W. R. (1988) Lipids in cereal starches: A review. Journal of Cereal Science., 8: 1-15.

Mu, C., Harn, C., Ko, Y. T., Singletary, G. W., Keeling, P. L. and Wasserman, B. P. (1994) Association of a 76 kDa polypeptide with soluble starch synthase I activity in maize (cv. B73) endosperm. The Plant Journal., 6: 151-159.

Mu, H. H., Yu, Y., Wasserman, B. P. and Carman, G. M. (2001) Purification and characterization of the maize amyloplast stromal 112-kDa starch phosphorylase. Archives of Biochemistry and Biophysics., 388: 155-164.

Mu-Forster. C., Huang, R. M., Powers, J. R., Harriman, R. W., Knight, M., Singletary, G. W., Keeling, P. L. and 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.

Munyikwa, T. R. I., Langeveld, S., Salehuzzaman, S. N. I. M., Jacobsen, E. and Visser, R. G. F. (1997) Cassava starch biosynthesis: new avenues for modifying starch quantity and quality. Euphytica., 96: 65-75.

Murai, J., Taira, T. and Ohta, D. (1999) Isolation and characterization of the three *Waxy* genes encoding the granule-bound starch synthase in hexaploid wheat. Gene., 234: 71-79.

Murray, M. G. and Thomson, W. F. (1980) Rapid isolation of high molecular weight plant DNA, Nucleic Acids Research., 8: 4321-4325.

Myers, A. M., Morell, M. K., James, M. G. and Ball, S. G. (2000) Recent progress toward understanding biosynthesis of the amylopectin crystal. Plant Physiology., 22: 989-997.

Myllarinen, P., Autio, K., Schulman, A.H. and Poutanen, K. (1998) Heat induced structural changes of small and large barley starch granules. Jounral of the Institute of Brewing., 104: 343- 349.

Nakamura. T., Yamamori, M., Hiurano, H. and Hidaka, S. (1993) The *waxy* (Wx) protein of maize, rice and barley. Phytochemistry., 33: 749-753.

Nakamura, T., Yamamori, M., Hirano, H., Hidaka, S. and Nagamine, T. (1995) Production of *waxy* (amylose-free) wheats. Molecular Genetics and Genomics., 248: 253–259.

Nakamura, T., Virnten.P., Hayakawa, K. and Ikeda, J. (1998) Characterization of a granule bound starch synthase isoform found in the pericarp of wheat. Plant Physiology., 118: 125-132.

Nelson, O. E and Rines, H. W. (1962) The enzymatic deficiency in the *waxy* mutant of maize. Biochemical and Biophysical Research Communications., 9: 297-300.

Nieto-Taladriz, M. T., Rodriguez-Quijano, M. and Carrillo, J. M. (2000) Polymorphism of waxy proteins in Spanish durum Wheats. Plant Breeding., 119: 277-279.

Noda, T., Takahata, Y., Sato, T., Suda, I., Morishita, T., Ishiguro, O. and Yamakawa, O. (1998) Relationships between chain length distribution of amylopectin and gelatinization properties within the same botanical origin for sweet potato and buckwheat. Carbohydrat. Polymers., 37: 153-158.

Okagaki, R. J. (1992) Nucleotide sequence of a long cDNA from the rice waxy gene. Plant Molecular Biology., 19: 1453-1462.

Olsen, K. M. and Purugganan, M. D. (2002) Molecular evidence on the origin and evolution of glutinous rice. Genetics., 162: 941-50.

Ouchterlony, O. and Nilsson, L. A. (1978) Immunodifussion and immunoelectrophoresis.In; Weir.D.M.(Ed.), Handbook of experimental immunology,3rd edn, oxford. Blackwell Scientific Publications,19: 16-19.23

Pan, X., Tang, Y., Li, M., Wu, G. and Jiang, H. (2009) Isoforms of GBSSI and SSII in four legumes and their phylogenetic relationship to their orthologs from other angiosperms. Journal of Molecular Evolution., 69: 625-634.

Park, Y-J., Nemoto, K., Nishikawa, T., Matsushima, K., Minami, M. and Kawase, M. (2009) Molecular cloning and characterization of granule bound starch synthase

I cDNA from a grain amaranth (*Amaranthus cruentus* L.). Breeding Science., 59: 351-360.

Pedersen, J. E., Bean, S. R.. Graybosch, R. A., Park, S. H. and Tilley, M. (2005) Characterization of waxy grain sorghum lines in relation to granule-bound starch synthase. Euphytica., 144: 151-156.

Pedersen, J. F., Graybosch, R. A. and Funnell, D.L. (2007) Occurrence of the *waxy* alleles *wxa* and *wxb* in waxy sorghum plant introductions and their effect on starch thermal properties. Crop Science., 47: 1927-1933.

Peralta, I. and Spooner, D. M. (2001). Granule-bound starch synthase (GBSS-I) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*). American Journal of Botany., 88: 1888–1902.

Pérez, E., Schultz, F. S. and de Delahaye, E. P. (2005) Characterization of some properties of starches isolated from *Xanthosoma sagittifolium* (tannia) and *Colocasia esculenta* (taro). Carbohydrate Polymers., 60: 139-145.

Peroni, F., Rocha, T. and Franco, C. (2006) Some structural and physicochmical characteristics of tuber and root starches. Food Science and Technology International., 12(6): 505-513.

Pilling. E. (2001) The origins of growth rings in starch granules. PhD thesis. University of East Anglia, UK.

Pilling, E. and Smith, A. M. (2003) Growth ring formation in the starch granules of potato tubers. Plant Physiology., 132: 365–371.

Pomeranz, Y. (1983) Buckwheat: Structure, composition, and utilization. Critical Reviews in Food Science and Nutrition., 19(3): 213-250.

Preiss, J., Stark, D., Barry, G. F., Guan, H. P., Libal-Weksler, Y., Sivak, M. N., Okida, T. W. and Kishore, G. M. (1994) Prospects for the production of cereals with improved starch properties. Improvement of Cereal Quality by Genetic Engineering. Edited by R.J. Henry and J.A. Ronalds. Plentum Press, New York. pp. 115-127.

Qian, J., Duarte, P. R. and Grant, L. (1998) Partial characterization of Buckwheat (*Fagopyrum esculentum*) Starch. Cereal Chemistry., 75(3): 365-373.

Qian, J. Y. and Kuhn, M. (1999b) Physical properties of buckwheat starches from various origins. Starch., 51: S81-S85.

Rahman, S., Kosar- Hashemi, B., Samuel, M., Hill, A., Abbott, D. C., Skerritt, J. H., Preiss, J., Appels, R. and Morell, M. (1995) The major proteins of wheat endosperm starch granules. Australian Journal of Plant Physiology., 22: 793-803.

Riley, C. K., Wheatley, A. O. and Asemota, H.N. (2006) Isolationand characterization of starches from eight Dioscorea alata cultivars grown in Jamaica. African Journal of Biotechnology., 5 (7): 1528-1536.

Robinson, R. G. (1980) The buckwheat crop in Minnesota. Agr. Exp. Sta. Bul. 539, Univ. Minnesota, St. Paul.

Robyt, J. F. (2008) "Biosynthesis of Polysaccharides," in Carbohydrate-active Enzymes: Structure, Function, and Applications (K-H. Park, ed.), CRC Press, Boca Raton, FL .,pp. 3-19.

Rodriguez-Quijano, M., Nieto-Taladriz, M.T. and Carrillo, J.M. (1998) Polymorphism of waxy proteins in Iberian hexaploid wheats. Plant Breeding., 117: 341-344.

Rohde, W., Becker, D. and Salamini, F. (1988) Structural analysis of the *waxy* locus from *Hordeum vulgure*. Nucleic Acids Research., 16: 7185-7186.

Rooney, L. W. and Pflugfelder, R. L. (1986) Factors affecting starch digestibility with special emphasis on sorghum and corn. Journal of Animal Science., 63: 1607-1623.

Rozen, S. and Skatelsky, H. J. (2000) In: Bioinformatics Methods and Protocols: Methods in Molecular Biology (Krawetz S, Misener S eds.) Humana Press Totowa New Jersey pp 365-386.

Salehuzzaman, S. N. I. M., Jacobsen, E. and Visser, R. G. F. (1993) Isolation and characterization of a cDNA encoding granule-bound starch synthase in cassava (*Manihot esculenta* Crantz) and its antisense expression in potato. Plant Molecular Biology., 23: 947-962.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) "Molecular Cloning: A Laboratory Manual."Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sang, Y., Bean, S., Seib, P. A., Pedersen, J. and Shi, Y. C. (2008) Structure and functional properties of sorghum starches differing in amylose content. Journal of Agriculture and Food Chemistry., 56: 6680-6685.

Sano, Y. (1984) Differential regulation of *waxy* gene expression in rice endosperm. Theoretical and Applied Genetics., 68: 467–473.

Sano, Y., Katsumata, M. and Omura, K. (1986) Genetic studies of speciation in cultivated rice. Inter- and intraspecific differentiation in the *waxy* gene expression of rice. Euphytica., 35: 1–9.

Sasaki, T. and Matsuki, J. (1998) Effect of wheat starch structure on swelling power. Cereal Chemistry., 75: 525-529.

Schoch, T. J. (1965) Starch in bakery products. Baker's Dig., 39(2): 48-52, 54-57.

Sestili, F., Botticella, E., Bedo, Z., Phillips, A. and Lafiandra, D. (2010) Production of novel allelic variation for genes involved in starch biosynthesis through mutagenesis. Molecular Breeding., 25: 145-154.

Shannon, J. C and Garwood. D. L. (1984) Genetics and physiology of starch development.Starch: Chemistry and Technology, R. L. Whistler, J. N. Be-Miller, E. F. Paschell, Academic Press, Orlando, FL: 25.

Sharan. R. N., Devi, B. J., Humtsoe, J. O., Saikia,J. R and Kma, L. (2005) Detection and quantification of poly-ADP-ribosylated cellular proteins of spleen and liver tissues of mice in vivo by slot and western blot immonoprobing using polyclonal antibody against mouse-ADP-ribose polymer. Molecular and Cellular Biochemistry., 278(1-2): 213-221.

Sheoran, I. S., Sproule, K. A., Olson, D. J. H., Ross, A. R. S. and Sawhney, V. K. (2006) Proteome profile and functional classification of proteins in Arabidopsis thaliana (Landsberg erecta) mature pollen. Sexual Plant Reproduction., 19: 185-196.

Sheoran, I. S., Ross, A. R. S., Olsen, D. J. H. and Sawhney, V. K. (2007) Proteomic analysis of tomato (*Lycopersicon esculentum*) pollen. Journal of Experimental Botany., 58 (13): 3525-3535.

Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Analytical Chemistry., 68: 850–858.

Shibanuma, K., Takeda, Y. and Hizukuri, S. (1996) Molecular and pasting properties of some wheat starches. Carbohydrate Polymer., 29: 253-261.

Shujun, W., Hongyan, L., Wenyuan, G., Haixia, C., Jiugao, Y. and Peigen, X. (2006) Characterization of new starches separated from different chinese yam (*Dioscorea opposite* Thumb.) cultivars. Food Chemistry., 99: 30-37.

Shure, M., Wessler, S. and Federoff, N. (1983) Molecular identification and isolation of the *waxy* locus in maize. Cell., 35: 225-233.

Simmonds, D. H. and O'Brien, T. P. (1981) Morphological and biochemical development of wheat endosperm. Advanced Cereal Science Technology., 4: 5-14.

Singh, N., Kaur, L., Ezekiel, R. and Gurraya, H. S. (2005) Microstructural, cooking and textural characteristics of potato (*Solanum tuberosum* L.) tubers in relation to physico-chemical and functional properties of their flours. Journal of the Science of Food and Agriculture., 85: 1275–1284.

Skrabanja, V., Elmstahl, L., Helena, G. M., Kreft I. and Bjork, I. M. E. (2001) Nutritional properties of starch in buckwheat Products: studies *in vitro* and *in vivo*. Journal of Agriculture and Food Chemistry., 49 (1): 490-496.

Smedmark, J. E. E., Eriksson, T., Evans, R. C. and Campbell, C.S. (2003) Ancient allopolyploid speciation in Geinae (Rosaceae): evidence from nuclear granule-bound starch synthase (GBSS-I) gene sequences. Systemic Biology., 52: 374–385.

Smith, A. M., Denyer, K. and Martin, C. (1997) The synthesis of the starch granule. Plant Molecular Biology., 48: 67-87.

Smith, S. D. and Baum, D. A. (2006) Phylogenetics of the florally diverse andean clade Iochrominae (Solanaceae). American Journal of Botany., 93: 1140–1153.

Song, Y. and Jane, J. (2000) Characterization of barley starches of waxy, normal, and high amylose varieties. Carbohydrate Polymers., 41: 365-377.

Soral- Smietana, M., Fornal, L. and Fornal, J. (1984) Characteristics of buckwheat grain starch and effect of hydrothermal processing upon its chemical composition, properties and structure. Starch/ Staerke., 36: 153-158.

Stanke, M. and Morgenstern, B. (2005) "AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints", Nucleic Acids Research., 33: W465-W467.

Stoddard F. L. (1999) Survey of starch particle-size distribution in wheat and related species. Cereal Chemistry., 76: 145–149.

Taira, T., Fujita, N., Takaoka, K., Uematsu, M., Wadano, A., Kozaki, S. and Okabe, S. (1995) Variation in the primary structure of *waxy* proteins (granule-bound starch synthase) in diploid cereals. Biochemical Genetics., 33: 269–281.

Takaoka, M., Watanabe, S., Sassa. H., Yamamori, M., Nakamura, T., Sasakuma, T. and Hirano, H. (1997) Structural characterization of high molecular weight starch granule-bound proteins in wheat (*Triticum aestivium* L.) Journal of Agriculture and Food Chemistry., 45: 2929-2934.

Tan, Y. F. and Zhang, Q. F. (2001) Correlation of simple sequence repeat variants in the leader sequence of the waxy gene with amylose content of the grain in rice. Acta Botanica Sinica., 43: 146–150.

Tatge, H., Marshall, J., Martin, C., Edwards, E. A. and Smith, A. M. (1999) Evidence that amylose synthesis occurs within the matrix of starch granules I potato tubers. Plant cell and environment., 22: 543-550.

Tester, R. F and Karkalas, J. (2002) Polysaccharides. II. Polysaccharides from eukaryotes. In: Vandamme EJ, De Baets S, Steinbuchel A, eds. Starch in biopolymers, Weinheim: Wiley-VCH., 6: 381–438.

Tester, R. F., Karkalas, J. and Qi, X. (2004) Starchcomposition, fine structure and architecture :(Review). Journal of Cereal Science., 39: 151-165.

Tetlow, I. J., Morell, M. K. and Emes, M. J. (2004a) Recent developments in understanding the regulation of starch metabolism in higher plants. Journal of Experimental Botany., 55: 2131-2145.

Thompson, D.B. (2000) On the non-random nature of amylopectin branching. Carbohydrate Polymers., 43: 223-239.

Tian, S. J., Richard, J. E. and Blanshard, J. M. V. (1991) Physicochemical properties of sweet potato starch. Journal of the Science of Food and Agriculture., 57: 459-491.

Towbin, H., Staehelin, T. and Gordan, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proceedings of the National. Academy of Sciences, U.S.A., 76: 4350-4354.

Tsai. C-Y. (1974) The function of the *waxy* locus in starch synthesis in maize endosperm. Biochemical Genetics., 11: 83-96.

Udesky, J. (1992) The book of soba. Harper & Row, New York.

Umeda, M., Ohtsubo, H. and Ohtsubo, E. (1991) Diversification of the rye *waxy* gene by insertion of mobile DNA elements into introns. Japanese Journal of Genetics. 66: 569-586.

Van de Wal. M., D'Hulst, C., Vincken, J-P., Buléon, A., Visser, R. and Ball, S. (1998) Amylose is synthesised *in vitro* by extension of and cleavage from amylopectin. Journal of Biological Chemistry., 273: 22232–22240.

Van der Leij, F. R., Visser, R. G. F., Ponstein, A. S., Jacobsen, E. and Feenstra, W. J. (1991) Sequence of the structural gene for granule – bound starch synthase of potato (*Solanum tuberosum* L.) and evidence for a single point deletion in the *amf* allele. Molecular Genetics and Genomics., 228(1-2): 240-248.

Visser, R. G. F., Hergersberg, M., Van Der Leij, F. R., Jacobsen, E., Witholt, B. and Feenstra, W. J. (1989) Molecular cloning and partial characterization of the genes for granule-bound starch synthase from a wildtype and an amylose-free potato (*Solanum tuberosum* L.). Plant Science., 64(2): 185-192.

Vrinten, P., Nakamura, T. and Yamamori, M. (1999) Molecular characterization of waxy mutations in wheat. Molecular Genetics and Genomics., 261: 463-471.

Vrinten, P. L. and 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: 255-264.

Walsh, B. M and Hoot, S. B. (2001) Phylogenetic relationships of capsicum (Solanaceae) using DNA sequences from two noncoding regions: the chloroplast atpB-rbcL spacer region and nuclear *waxy* introns. International Journal of Plant Sciences., 162: 1409-1418.

Wang, D., Bean, S., McLaren, J., Seib, P., Madl, R., Tuinstra, M., Shi, Y., Lenz, M.,Wu, X. and Zhao, R. (2008) Grain sorghum is a viable feedstock for ethanol production.Journal of Industrial Microbiology and Biotechnology., 35: 313-320.

Wang, L. T., Bograchava, T. Y. and Hedley, C. (1998) Starch: as simple as A, B, C: Review: Journal of Experimental Botany., 49(320) : 481-502.

Wang, L., and Seib, P. A. (1996) Australian salt-noodle flours and their starches compared to U.S. wheat flours and their starches. Cereal Chemistry., 73: 167-175.

Wang, S-J., Yeh, K-W. and Tsai, C-Y. (2000) An expression of the Sweet Potato GBSSI gene in *Escherichia coli*. Taiwania., 45 (4): 296-304.

Wang, Z., Wu, Z., Xing, Y., Zheng, F., Guo, X., Zhang, W. and Hong, M. (1990) Nucleotide sequence of rice *waxy* gene. Nucleic Acids Research., 18: 5898.

Wang, Z-Y., Zheng, F-Q., Shen, G-Z., Goa, J-P., Snustad, D. P., Li, M-G., Zhang, J-L. and Hong, M-M. (1995) The amylose content in rice endosperm is related to the post- transcriptional regulation of the *waxy* gene. Plant Journal., 7: 613-622.

Wattebled, F., Buleon, A., Bouchet, B., Ral, J. P., Lienard, L., Delvalle, D., Binderup, K., Dauvillee, D., Ball, S. and D'Hulst, C. (2002) Granule-bound starch synthase I. A major enzyme involved in the biogenesis of B-crystallites in starch granules. Eurpean Journal of Biochemistry., 269: 3810–3820.

Whistler, R. L. (1995) U.S. Pat. 5453281. Lafayette Applied Chemistry Inc.

Whistler, R. L. (1997) U.S. Pat. 5651828. Lafayette Applied Chemistry Inc.

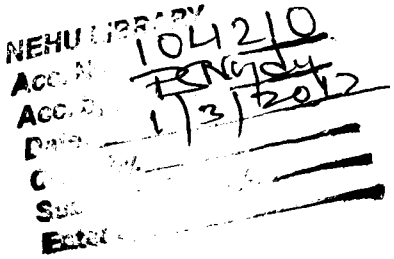
Wilhelm, E., Themeier, W. and Lindhauer, M. G. (1998) Small granule starches and hydrophilic polymers as components for novel biodegradable two-phase compounds for special applications, 1: Separation and refinement techniques for small granule starches from amaranth and quinoa. Starch/Strake., 50: 7-13.

Wong, J. H., Lau. T., Cai, N., Singh. J., Pedersen, J. E., Vensel, W. H., Hurkman, W. J., Wilson, J. D., Lemaux, P. G. and Buchanan, B. B. (2009) Digestibility of protein and starch from sorghum (*Sorghum bicolor*) is linked to biochemical and structural properties of grain endosperm. Journal of Cereal Science., 49: 73-82.

Xu, J., Frick, M., Laroche, A., Ni, Z-Fu., Li, B-Y. and Lu, Z-X. (2009) Isolation and characterization of the rye *waxy* gene. Genome., 52: 658-664.

Yamamori, M., Nakamura, T. and Kuroda, A. (1992) .Variations in the content of starch granule bound protein among several Japanese cultivars of common wheat (*Triticum aestivum* L.). Euphytica., 64: 251-219.

Yamamori, M. and Quynh, N.T. (2000) Differential effects of Wx -A1, -B1 and -D1 protein deficiencies on apparent amylose content and starch pasting properties in common buckwheat. Theoretical and Applied Genetics., 100: 32-38.



Yamamori, M., Nakamura, T., Endo, T. and Nagamine, T. (1992) Variations in the content of starch-granule bound protein among several Japanese cultivars of common wheat (*Triticum aestivum* L.) Euphytica., 64: 215-219.

Yamamori, M., Nakamura, T., Endo, T.R. and Nagamine, T. (1994). *Waxy* protein deficiency and chromosomal location of coding genes in common wheat. Theoretical and Applied Genetics., 89 : 179-184.

Yamamori, M., Nakamura, T. and Nagamine, T. (1995) Polymorphism of two *waxy* proteins in the emmer group of tetraploid wheat, *Triticum dicoccoides*, *T.dicoccum*, and *T.durum*. Plant Breeding., 114: 215-218.

Yoshimoto, Y., Egashira, T., Hanashiro, I., Ohinata, H., Takase, Y. and Takeda, Y. (2004) Molecular structure and some physicochemical properties of buckwheat starches. Cereal Chemistry., 81(4): 515–520.

Yuan, Y., Zhang, L., Dai, Y. and Yu. J. (2007) Physicochemical properties of starch obtained from *Dioscorea nipponica* Makino comparison with other tuber starches. Journal of Food Engineering., 82: 436-442.

Zhang, X. L., Colleoni, C., Ratushna, V., Sirghle-Colleoni, M., James, M. G. and Myers, A. M. (2004) Molecular characterization demonstrates that the Zea mays gene sugary2 codes for the starch synthase isoform SSIIa. Plant Molecular Biology., 54: 865-879.

Zheng, G. H. and Sosulski, F.W. (1997) Physico-chemical properties of small starch granules. American Association of Cereal Chemists. Annual Meeting, San Diego.

Zheng, G. H. and Sosulski F. W. (1998) Determination of Water Separation from Cooked Starch and Flour Pastes after Refrigeration and Freeze-thaw. Journal of Food Science., 63(1): 134-139.

Zheng, G. H., Sosulski, F. W. and Tyler, R. T. (1998) Wet-milling, composition and functional properties of starch and protein isolated from buckwheat groats. Food Research International., 30: 493-502.

Zhu, T., Jackson, D. S., Wehling, R. L. and Geera, B. (2008) Comparison of amylose determination methods and the development of a dual wavelength iodine binding technique. Cereal Chemistry., 85: 51-58.