Theor Appl Genet (1999) 98 : 1208–1216

Z. Li · S. Rahman · B. Kosar-Hashemi
G. Mouille · R. Appels · M. K. Morell

Cloning and characterization of a gene encoding wheat starch synthase I

Received: 5 June 1998 / Accepted: 29 September 1998

Abstract A cDNA clone, and a corresponding genomic DNA clone, containing full-length sequences encoding wheat starch synthase I, were isolated from a cDNA library of hexaploid wheat (*Triticum aestivum*) and a genomic DNA library of *Triticum tauschii*, respectively. The entire sequence of the starch synthase-I cDNA (wSSI-cDNA) is 2591 bp, and it encodes a polypeptide of 647 amino-acid residues that shows 81% and 61% identity to the amino-acid sequences of SSI-type starch synthases from rice and potato, respectively. In addition, the putative N-terminal amino-acid sequence of the encoded protein is identical to that determined for the N-terminal region of the 75-kDa starch synthase present in the starch granule of hexaploid wheat. Two prominent starch synthase activities were demonstrated to be present in the soluble fraction of wheat endosperm by activity staining of the non-denaturing PAGE gels. The most anodal band (wheat SSI) shows the highest staining intensity and results from the activity of a 75-kDa protein. The wheat SSI mRNA is expressed in the endosperm during the early to mid stages of wheat grain development but was not detected by Northern blotting in other tissues from the wheat plant. The gene encoding the wheat SSI (*SsI-D1*) consists of 15 exons and 14 introns, similar to the structure of the rice starch synthase-I gene. While the exons of wheat and rice are virtually identical in length, the wheat *SsI-D1* gene has longer sequences in introns 1, 2, 4 and 10, and shorter sequences in introns 6, 11 and 14, than the corresponding rice gene.

Communicated by G. E. Hart

Z. Li · S. Rahman · B. Kosar-Hashemi · G. Mouille
R. Appels · M. K. Morell (✉)
CSIRO Plant Industry, GPO Box 1600, Canberra,
ACT 2601, Australia
Fax: 61-2-6246-5000
E-mail: m.morell@.pi.csiro.au

Key words Starch · Wheat · Starch synthase · Gene structure · Gene expression

Introduction

Starch is a major component of wheat flour, accounting for 65–70% of the dry weight of the wheat grain. Starch is deposited in a granular form within the amyloplast of the endosperm. Wheat starch typically contains 20–30% amylose (an essentially linear α-1,4 glucan) and 70–80% amylopectin (a branched α-1,4 glucan containing about 5% α-1,6 linkages). Starch is synthesised by a complex pathway which involves the concerted action of a number of enzymes that include ADP glucose pyrophosphorylase (EC 2.7.7.27), starch synthases (EC 2.4.1.21), branching enzymes (EC 2.4.1.18) and de-branching enzymes (EC 3.2.1.41 and EC 3.2.1.68) (Mouille et al. 1996). Isoforms of these activities are found in the majority of plants and the specific contributions of the various isoforms of each activity to the amount and composition of starch deposited in the cereal endosperm remain to be fully defined.

Starch synthase extends starch polymers through the addition of the glucosyl moiety of its substrate, ADPglucose, to the non-reducing end of a pre-existing α-1,4 glucan. In plants an exclusively granule-bound starch synthase (GBSS) has been well characterised, and there are additional isoforms of starch synthase which are partitioned between the granule and the soluble phase. We concur with Knight et al. (1998) that it is inappropriate to refer to those starch synthases partitioned between the granule and the soluble fraction as "soluble" starch synthases, and therefore prefer the generic term "starch synthase" to describe those starch synthases which are not exclusively located in the granule.

While each starch synthase elongates α-1,4 glucans, mutation studies in maize, rice, barley, wheat, pea, *Chlamydomonas* and other species show that only GBSS is essential for amylose synthesis. While GBSS also contributes to amylopectin synthesis (Maddelein et al. 1994; Denyer et al. 1996), the remaining complement of starch synthases are thought to be primarily responsible for the elongation of amylopectin chains.

Multiple starch synthases have been described in the soluble fraction of maize, rice, pea, potato and *Chlamydomonas*, and these enzymes display differences in physical and kinetic properties (Dang and Boyer 1988; Denyer and Smith 1992; Baba et al. 1993; Edwards et al. 1995; Mouille et al. 1996; Craig et al. 1998). In potato (SSII and SSIII) and pea (SSI and SSII), two starch synthases have been characterised from soluble extracts of each species (Denyer and Smith 1992; Edwards et al. 1995; Abel et al. 1996). In cereals, both cDNA and genomic DNA encoding the SSI gene of rice have been cloned and analysed (Baba et al. 1993; Tanaka et al. 1995). Partial wheat starch synthase mRNAs have been deposited in Genbank (accession numbers: TAU48227, TAU66377). Recently the *dull1* locus in maize has been characterised and shown to encode a starch synthase related to the starch synthase III of potato (Gao et al. 1998).

In wheat, the full complement of starch synthases and their individual roles remain to be elucidated. Three distinct starch synthases, with molecular weights of 60 kDa, 75 kDa and 100–105 kDa, have been identified in the starch granule (Denyer et al. 1995; Rahman et al. 1995). The 60-kDa GBSS is the product of the *waxy* gene and is found only in the starch granule. The combination of null alleles from each of the wheat genomes (Nakamura et al. 1995) results in the amylose-free "waxy" phenotype found in other species. The 100–105 kDa putative starch synthases are found only in the granule; however, their role is unclear (Denyer et al. 1995; Rahman et al. 1995). In the soluble fraction, Denyer et al. (1995) reported the presence of two starch synthase activities.

In the present paper, we report the further identification of polypeptides with starch synthase activity from the soluble fraction of wheat endosperm extracts. A 75-kDa starch synthase, closely related to the starch synthase-I isoforms of other species, has been shown to be present in both the granule and the soluble fraction, and the cloning and characterisation of a full-length cDNA and a corresponding genomic DNA sequence encoding this starch synthase are reported. This 75-kDa wheat starch synthase I is identical to the protein designated Sgp-3 by Yamamori and Endo (1996), and we introduce the designation *SsI-A1*, *SsI-B1* and *SsI-D1* as more appropriate designations for the orthologous genes encoding this protein. We show that the mRNA encoded by these genes is expressed at the early to middle stage of endosperm development, deduce the intron-exon structure of the *SsI-D1* gene, and confirm the location of the wheat starch synthase-I genes in the group-7 chromosomes of the wheat genome.

Materials and methods

Plant material

Genetic stocks of wheat cv Chinese Spring with various chromosome additions and deletions were kindly supplied by Dr. E. Lagudah (CSIRO Plant Industry, Canberra) and derived from stocks described in Sears and Miller (1985). The hexaploid (*Triticum aestivum*) wheats "Rosella" and "Gabo" were grown in the field and in controlled growth cabinet conditions (18°C day and 13°C night, with a photoperiod of 16 h) respectively. Wheat leaves and florets prior to anthesis, and endosperm were collected over the grain-filling period, immediately frozen in liquid nitrogen and stored at −80°C until required.

Gel electrophoresis

SDS-PAGE and non-denaturing PAGE were carried out using 8% gels according to Laemmli (1970). Activity staining of non-denaturing PAGE gels was carried out according to Abel et al. (1996). Zymogram analysis of SDS denatured extracts was carried out according to Buléon et al. (1997).

Construction of genomic library and isolation of plaques

The genomic library used in this study, prepared from *Triticum tauschii*, var strangulata, accession number CPI 110799, has been described in Rahman et al. (1997). Of all the accessions of *T. tauschii* surveyed, DNA marker analysis suggests that the genome of CPI 110799 is the most closely related to the D genome of hexaploid wheat (Lagudah et al. 1991). Positive plaques in the genomic library were selected as those hybridising with the PCR-derived DNA fragment, wsssp, (refer to Results section). Hybridisation was carried out in 25% formamide, 6 × SSC, 0.1% SDS at 42°C for 16 h, then washed three times with 2 × SSC containing 0.1% SDS at 65°C for 1 h per wash.

Preparation of total RNA from wheat

Total RNA was isolated from the leaf, floret and endosperm tissues of wheat essentially as described by Higgins et al. (1976). RNA was quantified by UV absorption and by separation in 1.4% agarose-formaldehyde gels, which were then visualised under UV light after staining with ethidium bromide.

Cloning of specific cDNA regions of wheat starch synthase I using RT-PCR

The first-strand cDNAs were synthesised from 1 μg of total RNA (derived from endosperm of the cultivar Rosella, 12 days after anthesis) as described by Maniatis et al. (1982), and then used as templates to amplify one specific cDNA region of wheat starch synthase I by PCR. The primers used to obtain the cDNA clone, wsssp, amplify between nucleotide positions 1615–1919 [refer to GenBank accession number: (AFO91803)]. The primer sequences are;

sss5′ (5′ GTGCCTCTGATTGGCTTTATTG 3′) and sss3′; (5′ CCTTGACATGGTATAA CG 3′). The probe used for Northern and Southern analysis, wsss3′, was amplified by PCR from the wSSI-cDNA [corresponding to nucleotide positions 2025–2497, refer to GenBank accession number: (AFO91803)] using the primers sssA (5′ GGACAAGATGTTGTGG GCATTGC 3′) and sssB (5′ CAG-CTGTAACACGGACAGAGAG 3′). The amplification was performed using a FTS-1 or FTS4000 thermal sequencer (Corbett, Australia) for 1 cycle of 95°C for 2 min; 35 cycles of 95°C for 30 s, 60°C for 1 min, 72°C for 2 min and 1 cycle of 25°C for 1 min.

DNA and RNA analysis

DNA was isolated and analysed as described by Maniatis et al. (1982). Approximately 20 μg of DNA was digested with the restriction enzymes *Bam*HI and *Dra*I, separated on a 1% agarose gel and transferred to re-inforced nitrocellulose membranes (BioRad) and hybridised with the ^{32}P-labelled cDNA probe wsss3′, corresponding to nucleotide positions 2025–2497 [refer to Genbank accession number: (AFO91803)]. The hybridisation and wash conditions were performed as described in Rahman et al. (1997). For RNA analysis, 10 μg of total RNA was separated in a 1.4% agarose-formaldehyde gel and transferred to a Hybond N+ membrane (Amersham) (Maniatis et al. 1982), and hybridised with the cDNA probe at 42°C in Khandjian hybridisation buffer (Khandjian 1987). DNA fragments wsssp and wsss3′ (refer to Results section) were labelled with the Rapid Multiprime DNA Probe Labelling Kit (Promega) and used as probes. The membrane was washed once for 30 min at 65°C with 2 × SSC containing 0.1% SDS, and for 3 × 40 min at 65°C with 0.2 × SSC containing 1% SDS prior to overnight exposure at −80°C with Kodak MR X-ray film.

Construction and screening of the cDNA library

A cDNA library was constructed from endosperm RNA of the hexaploid *T. aestivum* cultivar Rosella as described by Rahman et al. (1997). The library was screened with a 305-bp cDNA fragment, wsssp (refer to Results section), employing the hybridisation conditions described earlier.

DNA sequencing and analysis

DNA sequencing was performed using the automated ABI system with dye terminators as described by the manufacturers. DNA sequences were analysed with the GCG suite of programs (Devereaux et al. 1984).

Results and discussion

Identification and localisation of wheat starch synthase polypeptides in the endosperm

The proteins found within the wheat starch granule after protease treatment were described by Rahman et al. (1995), and for reference are shown in Fig. 1 (Panel A, lane1). N-terminal sequence analysis indicated that the prominent 60-kDa band is the "waxy" GBSS protein, and strongly suggested that the 75-kDa protein was a starch synthase of the starch synthase-I class (Rahman et al. 1995). Denyer et al. (1995) reported

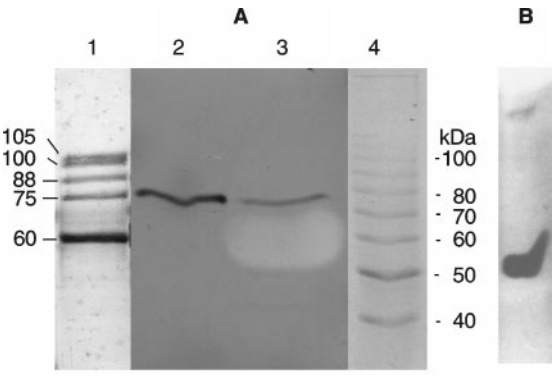

Fig. 1A, B Analysis of starch synthases from the wheat endosperm starch granule and soluble fractions. Panel **A** SDS-PAGE analysis. *Lane 1* wheat endosperm starch granule proteins revealed by silver staining; *lane 2* SDS-PAGE of wheat endosperm starch granule proteins. Starch synthase activity revealed by activity staining; *lane 3* SDS-PAGE of wheat endosperm soluble proteins. Starch synthase activity revealed by activity staining; *lane 4* standard molecular-weight ladder revealed by Coomassie blue staining. Panel **B** Activity staining of wheat endosperm soluble proteins separated by non-denaturing-PAGE

the co-migration of starch synthase activity with this band, the cross reaction of this band with antibodies to pea starch synthase II, and the presence of an immunoreactive 75-kDa protein in the soluble fraction of wheat endosperm extracts. The 75-kDa granule protein has also been investigated in wheat by Yamamori and Endo (1996) and Takaoka et al. (1997) who designated this protein Sgp-3, noted the homology of internal amino-acid sequences to rice SSI, and separated the granule forms encoded by the A, B and D genomes of wheat by 2-D gel electrophoresis. In Fig. 1 (panel A, lane 2) we show that following SDS-PAGE of wheat starch-granule extracts, staining for starch synthase activity after a protein renaturation procedure reveals a band which precisely co-migrated with the 75-kDa granule protein. Fig. 1 (panel A, lane 3) indicates that a 75-kDa starch synthase is also located in the soluble fraction of the wheat endosperm. Activity staining of non-denaturing PAGE gels for starch synthase activity revealed two prominent activities in wheat-endosperm soluble extracts (Fig. 1, panel B). Further analysis of the anodal band (marked SSI) by SDS-PAGE, followed by in-gel renaturation of starch synthase activity, demonstrated that this starch synthase had a molecular weight of 75 kDa (data not shown) and co-migrated in SDS-PAGE with the starch synthase activity shown in Fig. 1, panel A, lane 3. We conclude that the 75-kDa protein of the wheat starch granule and the soluble fractions are members of the SSI class and may be the products of the same orthologous genes. The cathodic starch synthase (Fig. 1, panel B) is designated SSx, to allow definitive classification of its gene at a later date. The molecular weight and electrophoretic mobility of this starch synthase (data not shown) is similar to the 180-kDa starch synthase II of maize, thought to be

encoded by the *Dull1* gene (Gao et al. 1998). This starch synthase did not cross-react in immuno-blotting experiments with monoclonal antibodies (Rahman et al. 1995) to the 100/105-kDa starch synthases of the starch granule (data not shown).

Amplification of a SSI cDNA sequence from wheat

Reverse transcription coupled with PCR (RT-PCR) was used to amplify wheat starch synthase cDNA sequences using primers (designated sss3′ and sss5′, see Materials and methods) based on the nucleotide sequences that were conserved between the starch synthase I of rice (Baba et al. 1993) and pea (Dry et al. 1992). A 305-bp RT-PCR product, wsssp, was obtained by amplification of cDNA prepared from wheat endosperm at 12 days post-anthesis. The sequence of the PCR product showed 88.9% identity at the nucleotide level with rice starch synthase I (Baba et al. 1993).

Cloning of a full-length wheat starch synthase-I cDNA

The 305-bp cDNA fragment of wheat starch synthase I was used as a probe for the hybridisation screening of a wheat endosperm cDNA library (Rahman et al., 1997). Eight positively hybridising cDNA clones were selected. The longest cDNA clone (2591 bp) was analysed and contained the 305-bp probe sequence, wsssp, between nucleotides 1615 and 1919 at 100% identity [refer to GenBank accession number: (AFO91803)]. The cDNA clone was designated wSSI-cDNA. A large open reading frame within this cDNA encoded a 647 amino-acid polypeptide starting at nucleotide 256 and terminating at nucleotide 2199. Figure 2 shows that the wSSI-cDNA encodes an amino-acid sequence which is identical to the N-terminal of a 75-kDa protein found in wheat starch granules and is highly homologous to the N-terminal of rice SSI (Rahman et al. 1995). In the 3′ non-coding region, two putative polyadenylation signals, AATAA and AAATA, were found at positions 2418 and 2584 downstream from the stop codon [refer to GenBank accession number: (AFO91803)].

Comparison of the wSSI-cDNA sequence with rice SSI showed that there is 80.8% identity at the amino-acid level and 74.7% identity at the nucleotide level. A partial wheat starch synthase-I sequence submitted to the Genbank database (accession number, U48227; M. Block, H. Loerz and S. Lutticke) has $> 98\%$ identity to wSSI-cDNA but is truncated at the N-terminus such that 411 bases are lacking from the 5′ end of the sequence to the start codon. Comparison of the amino-acid sequences of the starch synthases of wheat, rice, pea and potato are given in Fig. 3. There is a higher homology in the central region and lower homology in both N- and C-terminal regions of the starch synthases of these four crop species. Three highly conserved regions of amino-acid sequences in plant starch synthases, GBSS, and *Escherichia coli* glycogen synthase (van der Leij et al. 1991; Baba et al. 1993) were also found within the coding region of wSSI-cDNA and these are indicated as Domains I, II and III, respectively (Fig. 3). Domain I contains the K-Xaa-G-G sequence found in the ADP-glucose binding site (Furukawa et al. 1990; Baba et al. 1993).

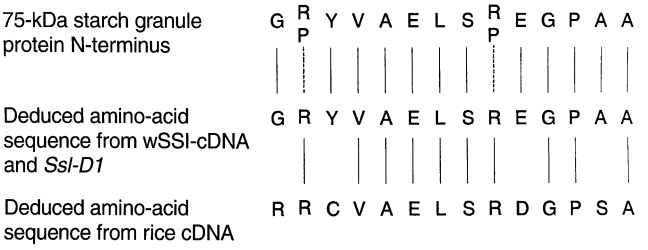

Fig. 2 Comparison of N-terminal sequences of wheat starch synthase-I and deduced amino-acid sequences of wheat and rice starch synthase-I cDNA. Identical residues between sequences are indicated by *vertical lines*. The sequences used are the N-terminal amino-acid sequence of the 75-kDa starch granule protein (Rahman et al. 1995; and further unpublished data), the deduced amino-acid sequence from the wSSI-cDNA (identical to the corresponding amino-acid sequence deduced from the *SsI-D1* DNA sequence) and the amino-acid sequence deduced from the rice SSI cDNA sequence (Baba et al. 1993). N-terminal amino-acid sequencing of the wheat starch granule 75-kDa protein yielded both R and P at positions 2 and 9, although R was obtained in a higher yield at this position. Both the cDNA sequence (wSSI-cDNA) and the genomic sequence (*SsI-D1*) encode arginine at this position. Proline may be encoded by either or both of the A or B genome homologues of the *SsI-D1* gene

The relationships between granule-bound and starch synthases from a variety of sources have been analysed at the amino-acid sequence level using the program PILEUP (Fig. 4). The comparison shows that wheat starch synthase I has highest homology to rice SSI. A second group of starch synthases was formed by the potato SSIII and maize SSII (*dull1*) sequences, and a third group by maize SSIIa, maize SSIIb, potato SSII and pea SSII. As has been noted previously (Abel et al. 1996), the granule-bound starch synthases form a distinct fourth cluster.

Cloning and characterisation of a genomic fragment containing the *T. tauschii SsI-D1* gene

Six genomic DNA clones were obtained from the screening of the *T. tauschii* genomic DNA library using a 305-bp cDNA fragment, "wsssp", as a probe (see above). These genomic DNA clones were classified into two groups by Southern-blot hybridisation analysis after digestion with *Bam*HI or *Sac*I and electrophoresis on a 1% agarose gel. The groups differ in the length of the cloned sequences but have been shown to contain sequences from a single gene, *SsI-D1* (data not shown).

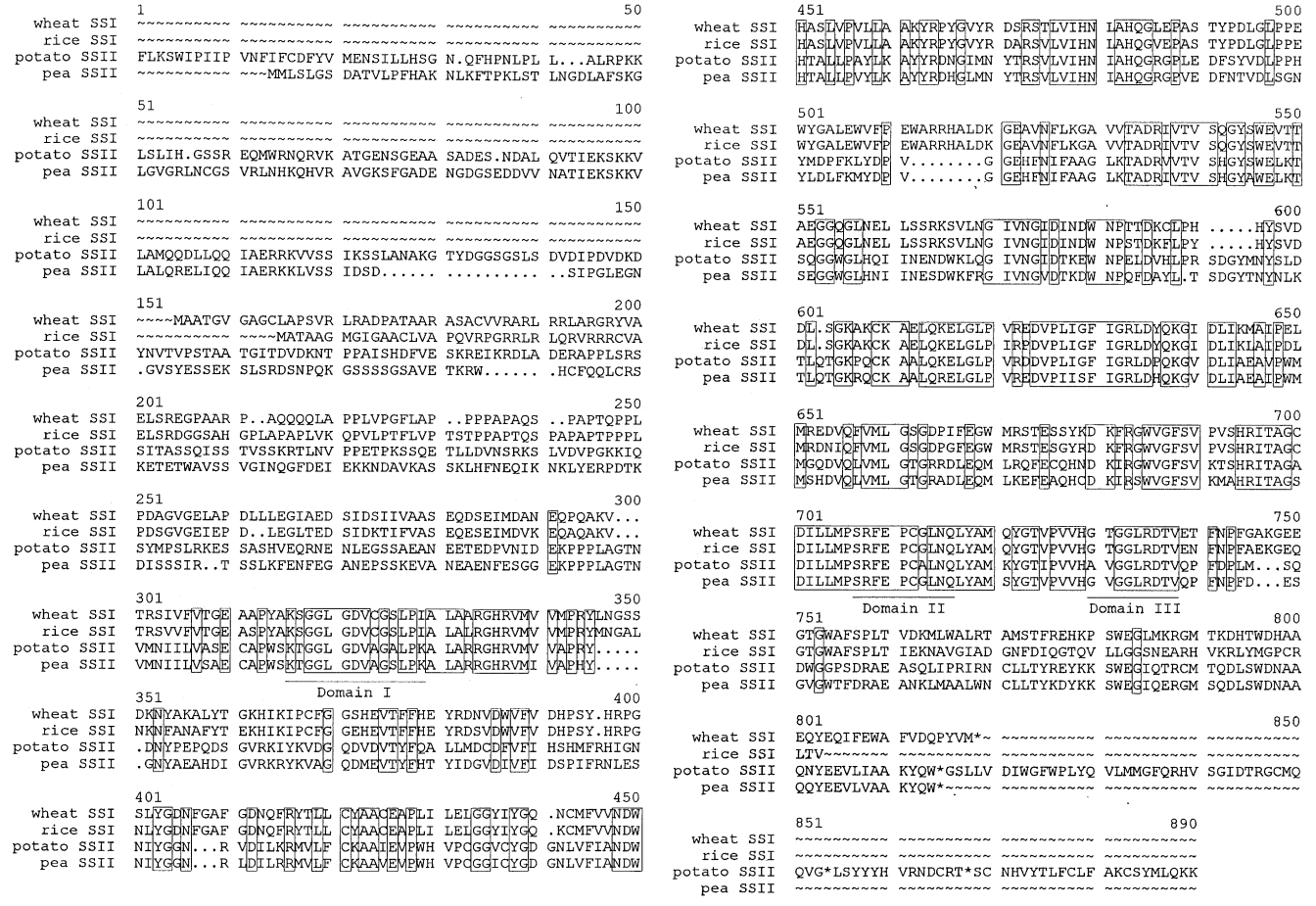

Fig. 3 A comparison of the deduced amino-acid sequences of starch synthases from wheat, rice, pea and potato. *WheatSSI* wheat starch synthase I (this paper); *riceSSI* rice starch synthase I (Baba et al. 1993); *potatoSSII* potato starch synthase II (van der Leij 1991); and *peaSSII* pea starch synthase II (Dry et al. 1992). Identical amino-acid residues among each of these sequences are *boxed* and three conserved domains are indicated

The genomic DNA clone designated SG3 was selected for further analysis on the basis of the largest total length of hybridising fragments (data not shown). SG3 was digested with *Bam*HI, *Kpn*I or *Sac*I, and fragments were subcloned into the plasmid pJKKm (Kirschman and Cramer 1988). A total of 22 clones were selected and sequenced from both ends. Nine of these subclones covered the intact wSSI-cDNA sequence region and 1 kb upstream containing at least part of the putative promoter region.

The sequences from these nine subclones overlapped and were assembled into one contiguous sequence which included a 1-kb upstream sequence to the coding region and a 10-kb sequence spanning the coding region [refer to Genbank accession number: (AFO91802)]. By comparing the *SsI-D1* sequence with the wSSI-cDNA sequence the coding region of the *SsI-D1* gene was found to be composed of 15 exons and 14 introns (Table 1; Fig. 5). The intron-exon structure of the *SsI-D1* gene is similar to that of the rice SSI gene (Tanaka et al. 1995). Comparison of the nucleotide sequences of these introns and exons between wheat and rice SSIs showed that there is generally a high homology between analogous exons in the coding region, ranging from 58.9% (exon 1b) to 93.9% (exon 7), and a lower homology between the corresponding introns, ranging between 37.5% (intron 10) and 58.9% (intron 9) (Table 1). Introns 1, 2, 4 and 10 of the *SsI-D1* gene are longer than those in the rice gene, while introns 6, 11 and 14 are shorter than the corresponding introns of the rice starch synthase-I gene. The transit peptide and N-terminal sequence of the mature peptide encoded by *SsI-D1* were found in exon 1 (Fig. 6). Three highly conserved regions of amino-acid sequences in plant starch synthases, GBSS, and *E.coli* glycogen synthase (van der Leij et al. 1991; Baba et al. 1993), termed Domains I, II, and III, were detected in the analysis of the product encoded by the wSSI-cDNA for starch synthase (see Fig. 1). In the *SsI-D1* sequence, Domain I is found in exon 2 and Domains II and III in exon 12 (Fig. 5).

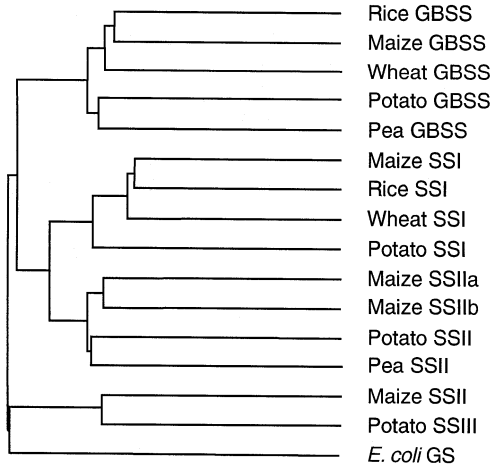

Fig. 4 Relationships between the primary amino-acid sequences of starch synthases. The dendrogram was generated by the program PILEUP (Devereaux et al. 1984). The amino-acid sequences used for the analysis are those deduced from the cDNAs of wheatSSI (this paper), riceSSI (Baba et al. 1993), maizeSSI (Knight et al. 1998), maize SSIIa and maize SSIIb (Harn et al. 1998), potatoSSII (Edwards et al. 1995), peaSSII (Dry et al. 1992), potatoSSIII (Abel et al. 1996), maizeSSII (Gao et al. 1998), riceGBSS (Okagaki, 1992), maizeGBSS (Kloesgen et al. 1986), wheatGBSS (Clark et al. 1991), potatoGBSS (van der Leij et al. 1991), peaGBSS (Dry et al. 1992) and *Escherichia coli* glycogen synthase (GS) (Kumar et al. 1986)

Table 1 Comparison of exons and introns of soluble starch synthase-I genes of wheat and rice

(1) Identity of exons of soluble starch synthase-I genes of wheat and rice

Exons	*SsI-D1*	rSSI	Identity (%)	Start site (*SsI-D1*)	Stop site (*SsI-D1*)
1a	255	113	57.52	−253	0
1b	316	298	58.92	1	316
2	356	356	82.87	1473	1828
3	78	78	92.31	2746	2823
4	125	125	90.40	2906	3028
5	82	82	89.02	4113	4194
6	174	174	93.10	4286	4459
7	82	82	93.90	4562	4643
8	92	92	92.39	4743	4835
9	63	63	90.48	4959	5021
10	90	90	82.22	5103	5192
11	125	125	88.80	8594	8718
12	109	109	91.74	8807	8915
13	53	53	81.13	8992	9044
14	40	41	80.00	9160	9199
15a	159	113	79.65	9499	9657
15b	392	539	46.46	9658	10098

(2) Identity of introns of soluble starch synthase-I genes of wheat and rice

Introns	*SsI-D1*	rSSI	Identity (%)	Start site (*SsI-D1*)	Stop site (*SsI-D1*)
1	1156	907	41.05	317	1472
2	917	851	41.65	1829	2745
3	82	87	45.12	2824	2905
4	1084	835	48.50	3029	4112
5	91	96	57.78	4195	4285
6	102	189	52.48	4460	4561
7	99	96	52.08	4644	4742
8	123	110	45.46	4836	4958
9	81	78	58.97	5022	5102
10	3401	663	37.56	5193	8593
11	88	124	56.82	8719	8806
12	76	81	48.68	8916	8991
13	115	135	45.22	9045	9159
14	299	830	45.80	9200	9498

Note: Exon 1a: non-coding region of exon 1. Exon 1b: coding region of exon 1
Exon 15a: coding region of exon 15. Exon 15b: non-coding region of exon 15
SsI-D1: wheat soluble starch synthase-I gene. rSSI: rice soluble starch synthase-I gene

Northern-blot hybridisation analysis of wheat starch synthase-I mRNA expression

The mRNA for starch synthase I could only be detected in wheat endosperm when RNA isolated from leaves, pre-anthesis florets and wheat endosperm was probed by Northern analysis (Fig 7, panel A). While expression of the *SsI-D1* gene in leaves could not be detected under the conditions employed, the rice SSI gene has been reported to be expressed in both seeds and leaves (Baba et al. 1993). This difference in gene regulation is unlikely to represent a major differences in starch biosynthetic patterns in wheat and rice leaves as both species partition a similar fraction (about 4%) of the carbon fixed in leaves to leaf starch compared to sucrose, and both accumulate low levels of leaf starch (J. Lunn, personal communication).

A Northern blot of mRNA extracted from endosperm of the wheat cultivar "Gabo" grown under defined temperature conditions (16-h daylength, day temperature 18°C, night temperature 13°C) is shown in Fig. 7, panel B. Under these comparatively low-temperature growth conditions, the message for the *SsI-D1* gene was present 6 days after anthesis and was strongly expressed 8 days after anthesis. Message levels remain high until 18 days after anthesis, after which time they decline. The expression pattern of the rice SSI gene in the endosperm is similar in that there is a low expression level detectable at 5 days post-anthesis, an increased expression level at 7–15 days post-anthesis, and a decline thereafter.

The *SsI-D1* promoter region

The promoter regions of the wheat *SsI-D1* gene and rice starch synthase-I gene were compared because of the similarity in function of the two genes in the endosperm and because they drive the expression of highly

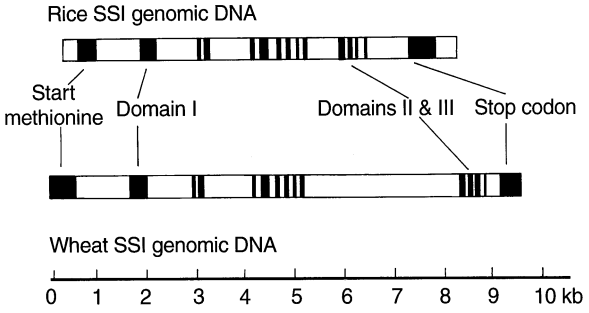

Fig. 5 Comparison of the structure of starch synthase-I genes of wheat and rice. The exons are indicated as *dark black boxes* and the introns are shown as *open boxes*. The size scale is indicated below the genes in base pairs: wheat SSI genomic DNA, wheat starch synthase gene I, rice SSI genomic DNA, rice starch synthase gene I. The location of the translation start codon, stop codon and three highly conserved amino-acid sequence regions (*Domains I, II and III*) are indicated

homologous cDNA sequences. However, little sequence similarity was found. The wheat *SsI-D1* gene contains no exact match to the EM-box motif described by Forde et al. (1985); however, a motif (TGAAAAC) similar to the EM box (TGTAAAG) is found between nucleotides −300 and −326 upstream of the first methionine, a region consistent with the location of other endosperm boxes. A putative TATA box is located from −358 to −361. However, as in the case of the promoter of the gene encoding wheat starch branching enzyme I (Rahman et al. 1998), there is no GCN4 box. Two directly repeated sequences were found in the putative promoter region (Fig. 5). An 11-nucleotide motif, (motif 1: "GAGTTTCACTA") was located at −643, −676 and −705, respectively, while a second motif (motif 2: "CCG(T/C)CCGTCCG") was located at −280 and −296. Motif 2 was also found in a similar position in the promoter region of the gene encoding the wheat starch branching enzyme-II gene from *T. tauschii* (Rahman et al. in preparation) and within the putative exon 1 of the GBSS gene from wheat (C. Konik et al. personal communication). A search of Genbank revealed that both motifs occurred in a number of plant promoters. While motif 1 exists in the promoter of the *Arabidopsis* tufA gene for elongation factor Tu (Genbank accession number: X52256), motif 2 is found in the promoter sequence of the barley *waxy* gene locus (Rohde et al. 1988), the maize gene for sucrose synthase (Werr et al. 1985) and the maize gene for extensin-like protein (Genbank accession number: Z34465). The significance of these motifs in promoter function remains to be determined.

Chromosomal localization of the wheat *SsI-1* genes

This study has shown that the Sgp-3 proteins investigated by Yamamori and Endo (1996) are encoded by the *SsI-1* genes. Previously, Yamamori and Endo

```
AAGGGTTCAGCGGCAACAACTGCGGCTCCAGAGCGATGGTCCTTAAGGGCACGTGCACGA   -941
AGACTTCACGGCTGTTATCGACAAGGTCAAGCCGGCTCCGATAGGGGAGCAGCGACAGCG   -881
GCGCGTCAACCGCTCGTTCTGGCGGCAGTAGTGGTCGTTCGGTGCTCTCGGAACCTCGAT   -821
GTAATTTTTATGATTTTAGAGATGCTTTGTACTTCCGATCGATGAACTCTGATAATAGAT   -761
ATCTCTTCTCTCGCAAAAAAAGAGAGTTTTCAACTGAAAACAAAAGAGTTTCACTAGTTC   -701
TTCTTTTAGAAACAGAGTTTCACTAGCACTTTTTTTTGCGAGAAGTCGAGTTTCACTAAG   -641
TACTAAACCCACGCAATTATTCTCAAAAAAAAAACCCACGCAACTGTCTGGATCCATCTT   -581
CGTTTTTTCCCCGAGAATCGTCTGGATCCATTTTCGTGTGCGAGGCATCCTCTCATTTTG   -521
CACGGCCCAGCTCTCTTCTCGCCGGCGTACGCTGCTACATGTCGGCACTCCACGCAAACA   -461
AAAAGAAGCCCAACCGAAAACGCACGCGCCTTTCCAGGCTCACCACGGAAAAAAATACCA   -401
CGCGCCGCTCACGAGCAAACCGTGACAACAGCCAGCCAGATATGGCAACGGAGGCACGGG   -341
CCGCACACAGCCACTGAAAACCGCAGCTGCTCTTCCGTCCGTCCGTCCCTCCGCCCGTCG   -281
GCGCCACTCCACTCGCCTTGCCCCACTCCCACTCTTCTCTCCCCGCGCACACCGAGTCGG   -221
CACCGGCTCATCACCCATCACCTCGGCCTCGGCCACCGGCAAACCCCCCGATCCGCTTTT   -161
GCAGGCAGCGCACTAAAACCCCGGGGAGCGCGCCCCGCGGCAGCAGCAGCACCGCAGTGG   -101
GAGAGAGAGGCTTCGCCCCGGCCCGCACCGAGCGGGGCGATCCACCGTCCGTGCGTCCGC    -41
                                                    ***
ACCTCCTCCGCCTCCTCCCCTGTCCCGCGCGCCCACACCCATGGCGGCGACGGGCGTCGG     20
                                                 M  A  A  T  G  V  G
                                                 (EXON 1)

CGCCGGGTGCCTCGCCCCCAGCGTCCGCCTGCGCGCCGATCCGGCGACGGCGGCCCGGGC     80
 A  G  C  L  A  P  S  V  R  L  R  A  D  P  A  T  A  A  R  A

GTCCGCTTGCGTCGTCCGCGCGCGGCTCCGGCGCTTGGCGCGGGGCCGCTACGTCGCCGA    140
 S  A  C  V  V  R  A  R  L  R  R  L  A  R↓G  R  Y  V  A  E

GCTCAGCAGGGAGGGCCCCGCGGCGCGCCCCGCGCAGCAGCAGCAACTGGCCCCGCCGCT    200
 L  S  R  E  G  P  A  A  R  P  A  Q  Q  Q  Q  L  A  P  P  L

CGTGCCAGGCTTCCTCGCGCCGCCGCCGCCCGCGCCCGCCCAGTCGCCGGCCCCGACGCA    260
 V  P  G  F  L  A  P  P  P  P  A  P  A  Q  S  P  A  P  T  Q

GCCGCCCCTGCCGGACGCCGGCGTGGGGGAACTCGCGCCCGACCTCCTGCTCGAAGGTAA    320
 P  P  L  P  D  A  G  V  G  E  L  A  P  D  L  L  L  E  G
```

Fig. 6 A partial DNA sequence of the wheat starch synthase-I gene, *SsI-D1*, which includes the 1-kb sequence upstream of the translation start codon and the first exon of the sequence of the *SsI-D1*-containing clone. The nucleotide sequence and deduced amino-acid sequence are numbered in the 5′ to 3′ direction. The deduced amino-acid sequence is shown below the nucleotide sequence. The translation start codon (ATG at nucleotide 1–3) is labelled with *asterisks and in bold*. The N-terminal sequence (at amino acid 41 to 55, determined by Rahman et al. 1995), is *underlined and in bold*. The transit polypeptide sequence (at amino acids 1–40) is in *italic form* and the putative peptide cleavage site is indicated by an *arrow*. An endosperm-specific sequence, TGAAAAC, and one putative TATA box are *underlined with single line and in bold or not in bold*, respectively. Two highly conserved repeats, GAGTTTCACTA and CCG(T/C)CCG TCCG, are enclosed within *boxes and in bold or not in bold*, respectively

(1996) demonstrated that these genes (referred to by these authors as the *Sgp-3* genes) are located in the short arm of chromosome 7A, chromosome 7B and chromosome 7D. This result has been confirmed in this study by Southern-hybridisation analysis of Chinese Spring homoeologous group-7 compensating nullisomic-tetrasomic lines (data not shown). Yamamori and Endo (1996) further showed that the genes for the Sgp-3 proteins were in the short arm of each of the group-7 chromosomes and that the order of genes is centromere/*Sgp-1*/*Sgp-3*/*Wx*, where *Sgp-1* encodes a group of proteins of molecular weight between 100 and 105 kDa (Denyer et al. 1995; Rahman et al. 1995) and *Wx* encodes the granule-bound starch synthase (GBSS).

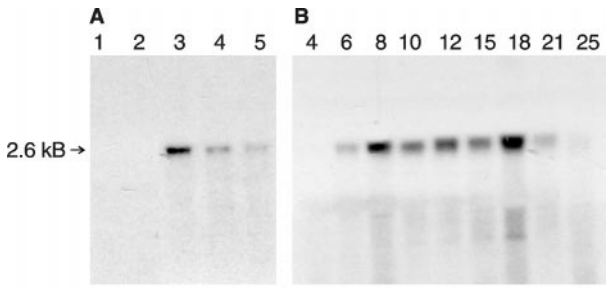

Fig. 7A, B Northern-blot analysis of the expression of starch synthase-I mRNA in wheat. Panel **A** shows Northern blotting of mRNA extracted from the wheat cultivar "Rosella" grown under field conditions and probed with the wsssp DNA fragment. *Lane 1* total wheat leaf RNA; *lane 2* pre-anthesis floret total RNA; *lane 3* total endosperm RNA 12 days after anthesis; *lane 4* total endosperm RNA 15 days after anthesis; *lane 5* total endosperm RNA 18 days after anthesis. Panel **B** shows Northern blotting of total RNA from the endosperm of the wheat cultivar "Gabo", grown under defined conditions (16 h daylength, 18°C day, 13°C night), and probed with the wsss3′ DNA fragment. The number of days after anthesis is given above the lanes in panel **B**

Conclusion

Four groups of wheat endosperm proteins have been reported by various groups to possess starch synthase activity, the 60-kDa GBSS, the 75-kDa granule-bound and soluble proteins, the 100/105 granule proteins and, in this report, a starch synthase with low mobility in non-denaturing PAGE (referred to as SSx). To-date, only the role of the 60-kDa GBSS has been defined with respect to starch biosynthesis in the wheat endosperm. Nakamura et al. (1995) have shown that the absence of the products of all three orthologous GBSS genes in wheat produces the classical "waxy" phenotype well known from a wide range of diploid species.

The 75-kDa starch synthase is found in both the granule and in the soluble fraction of the endosperm (Denyer et al. 1995); however, this enzyme was considered by these authors to contribute only a minor fraction of the starch synthase activity in the soluble fraction on the basis of activity measurements following anion-exchange separation of two peaks of starch synthase activity. In the present report, we have used non-denaturing PAGE to identify polypeptides with starch synthase activity in the wheat endosperm soluble fraction because this method is comparatively rapid and reduces interference from starch-modifying or degradative activities. This analysis confirms that there is a minimum of two forms of starch synthase in the soluble fraction of the wheat endosperm; however, in this assay format SSI activity is the predominant activity. The molecular cloning of cDNA and genomic sequences for the 75-kDa proteins reported here will facilitate a more definitive investigation of the role of the 75-kDa starch synthase I in starch synthesis in the wheat endosperm through genetic analysis or manipulation.

Acknowledgements The authors acknowledge the financial support provided by Goodman Fielder Limited and by Groupe Limagrain through the Cooperative Research Centre for Plant Science and also thank Dr. Bryan Clarke, Ms. Mandy Watson and Dr. Bill Taylor for assistance in the preparation of wheat mRNAs. The authors appreciate the expert assistance of Prof. Gary E. Hart with nomenclature issues. All recombinant DNA experiments complied with current Australian regulations.

References

Abel GJW, Springer F, Willmitzer L, Kossmann J (1996) Cloning and functional analysis of a cDNA encoding a novel 139-kDa starch synthase from potato (*Solanum tuberosum* L.). Plant J 10:981–991

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

Buléon A, Gallant DJ, Bouchet B, Mouille G, D'Huist C, Kossman J, Ball S (1997) *Chlamydomonas reinhardtii* as a model microbial system to investigate the biosynthesis of the plant amylopectin crystal. Plant Physiol 115:949–957

Clark JR, Robertson M, Ainsworth CC (1991) Nucleotide sequence of a wheat (*Triticum aestivum* L.) cDNA clone encoding the waxy protein. Plant Mol Biol 16:1099–1101

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

Dang PL, Boyer CD (1988) Maize leaf and kernel starch synthases and starch branching enzymes. Phytochemistry 27:1255–1259

Denyer K, Smith AM (1992) The purification and characterisation of two forms of soluble starch synthase from developing pea embryos. Planta 186:609–617

Denyer K, Hylton CM, Jenner CF, Smith AM (1995) Identification of multiple isoforms of soluble and granule-bound starch synthase in developing wheat endosperm. Planta 196:256–265

Denyer K, Clarke B, Smith AM (1996) The elongation of amylose and amylopectin chains in isolated starch granules. Plant J 10:1135–1143

Devereaux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Dry I, Smith A, Edwards A, Bhattacharyya M, Dunn P, Martin C (1992) Characterisation of cDNAs encoding two isoforms of granule-bound starch synthase which show differential expression in developing storage organ of pea and potato. Plant J 2:193–202

Edwards A, Marshall J, Sidebottom C, Visser RGF, Smith AM, Martin C (1995) Biochemical and molecular characterization of a novel starch synthase from potato tubers. Plant J 8:283–294

Forde BG, Heyworth A, Pywell J, Kreis M (1985) Sequence of a B1 hordein gene and the identification of possible upstream regulatory elements in the storage protein genes from barley, wheat and maize. Nucleic Acids Res 13:7327–7339

Furukawa K, Tagaya M, Inouye M, Preiss J, Fukui T (1990) Identification of lysine 15 at the active site in *Escherichia coli* glycogen synthase. J Biol Chem 265:2086–2090

Gao M, Wanat J, Stinard PS, James MG, Myers AM (1998) Characterization of *dull1*, a maize gene coding for a novel starch synthase. Plant Cell 10:399–412

Harn C, Knight M, Ramakrishnan A, Guan H, Keeling PL, Wasserman BP (1998) Isolation and characterization of the zSSIIa and zSSIIb starch synthase cDNA clones from maize endosperm. Plant Mol Biol 37:639–649

Higgins TJV, Zwar JA, Jacobsen JV (1976) Gibberellic acid enhances the level of translatable mRNA for a-amylase in barley aleurone layers. Nature 260:166–168

Khandjian EW (1987) Optimized hybridisation of DNA blotted and fixed to nitrocellulose and nylon membranes. Bio/Technology 5:165–167

Kirschman JA, Cramer JH (1988) Two new tools: multi-purpose cloning vectors that carry kanamycin or spectinomycin/streptomycin resistance markers. Gene 68:163–165

Kloesgen RB, Gierl A, Schwarz-Sommer ZS, Saedler H (1986) Molecular analysis of the waxy locus of *Zea mays*. Mol Gen Genet 203:237–244

Knight ME, Harn C, Lilley CER, Guan H, Singletary GW, Mu-Forster C, Wasserman BP, Keeling PL (1998) Molecular cloning of starch synthase from maize (W64) endosperm and expression in *Escherichia coli*. Plant J 14:613–622

Kumar A, Larsen CE, Preiss J (1986) Biosynthesis of bacterial glycogen: primary structure of *Escherichia coli* ADP-glucose: alpha-1,4-glucan, 4-glucosyltransferase as deduced from the nucleotide sequence of the *glg*A gene. J Biol Chem 261: 16256–16259

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685

Lagudah ES, Appels R, McNeil D (1991) The *Nor-D3* locus of *Triticum tauchii*: natural variation and genetic linkage to markers in chromosome 5. Genome 34:387–395

Leij van der FR, Visser RGF, Ponstein AS, Jacobsen E, Feenstra WJ (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. Mol Gen Genet 228: 240–248

Maddelein M-L, Libbessart N, Bellanger F, Delrue B, D'Hulst C, Ven den Koorhuyse N, Fontaine T, Wieruszeski JM, Decq A, Ball SG (1994) Toward an understanding of the biogenesis of the starch granule: determination of granule-bound and soluble starch-synthase functions in amylopectin synthesis. J Biol Chem 269:25150–25157

Maniatis T, Fritsch EF, Maniatis J (1982) *Molecular cloning*: *a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Mouille G, Maddelein M-L, Libessart N, Talaga P, Decq A, Delrue B, Ball S (1996) Preamylopectin processing: a mandatory step for starch biosynthesis in plants. Plant Cell 8:1353–1366

Okagaki RJ (1992) Nucleotide sequence of a long cDNA from the rice waxy gene. Plant Mol Biol 19:513–516

Nakamura T, Yamamori M, Hirano H, Hidaka S, Nagamine T (1995) Production of waxy (amylose-free) wheats. Mol Gen Genet 248:253–259

Rahman S, Kosar-Hashemi B, Samuel M, Hill A, Abbott DC, Skerritt JH, Preiss J, Appels R, Morell M (1995) The major proteins of wheat endosperm starch granules. Aust J Plant Physiol 22:793–803

Rahman S, Abrahams S, Abbott D, Mukai Y, Samuel M, Morell M, Appels R (1997) A complex arrangement of genes at a starch branching enzyme-I locus in the D-genome donor of wheat. Genome 40:465–474

Rahman S, Li Z, Abrahams S, Abbott D, Appels R, Morell MK (1998) Characterisation of a gene encoding wheat endosperm starch branching enzyme I. Theor Appl Genet 98: (in press)

Rohde W, Becker D, Salamini F (1988) Structural analysis of the waxy locus from *Hordeum vulgare*. Nucleic Acids Res 16:7185–7186

Sears ER, Miller TG (1985) The history of Chinese spring wheat. Cereal Res Commun 13:261–263

Takaoka M, Watanabe S, Sassa H, Yamamori M, Nakamura T, Sasakuma T, Hirano H (1997) Structural characterisation of high-molecular-weight starch granule-bound proteins in wheat (*Triticum aestivum* L). J Agric Food Chem 45:2929–2934

Tanaka K, Ohnishi S, Kishimoto N, Kawasaki T, Baba T (1995) Structure, organisation, and chromosomal location of the gene encoding a form of rice soluble starch synthase. Plant Physiol 108:677–683

Werr W, Frommer WB, Maas C, Starlinger P (1985) Structure of the sucrose synthase gene on chromosome 9 of *Zea mays* L. EMBO J 4:1373–1380

Yamamori M, Endo TR (1996) Variation of starch granule proteins and chromosome mapping of their coding genes in common wheat. Theor Appl Genet 93:275–281