Plant Molecular Biology **16**: 1099–1101, 1991.

Update section

Sequence

Nucleotide sequence of a wheat (*Triticum aestivum* L.) cDNA clone encoding the *waxy* protein

Joanna R. Clark, Morag Robertson[1] and Charles C. Ainsworth*

*Department of Biochemistry and Biological Sciences, Wye College (University of London), Wye, Kent TN25 5AH, UK (*author for correspondence)*; [1]*present address: Now at Centre for Genome Research, University of Edinburgh, Kings Buildings, West Mains Road, Edinburgh EH9 3JQ, UK*

Received 11 March 1991; accepted 13 March 1991

Key words: wheat, *Triticum aestivum*, cDNA clone, *waxy* protein

The *waxy* mutation in maize is characterised by the complete loss of amylase in the starch fraction of the kernel [1, 2, 3]. In addition, there is a significant reduction in the amount of granule-bound starch synthase activity in the mutant [2] which correlates with the absence of a protein of 60 kDa, which in wild-type kernels is the most abundant protein associated with starch granules [4, 5]. For this reason, it has been assumed that the 60 kDa *waxy* protein is the granule-bound starch synthase. Unequivocal evidence for this is lacking and in pea, evidence suggests that another protein, only distantly related to the *waxy* protein, is the granule-bound starch synthase [6]. It remains to be seen whether this is the case in cereals. The *waxy* locus has been cloned from a number of plants including maize [3], rice [7], barley [8] and potato [9]. We report here the isolation of a full-length cDNA clone for the *waxy* protein of hexaploid wheat and present the complete nucleotide sequence.

A cDNA library of was constructed in λgt10 from poly(A)$^+$ RNA isolated from wheat grains harvested at 20 days post anthesis from the hexaploid variety Chinese Spring. Isolation of RNA, selection of poly(A)$^+$ RNA and cDNA synthesis were carried out according to manufacturer's protocols (Amersham). After addition of linkers the cDNAs were ligated into the *Eco* RI site of λgt10 and packaged to produce 1.5×10^5 recombinants per μg poly(A)$^+$ RNA. The library was screened for *waxy* protein clones by hybridising plaque lifts with a barley *waxy* protein cDNA (pcwx27 [4]) and 9 clones with inserts that ranged in size from 0.5 to 4.4 kb were isolated and plaque-purified. The insert from a single cDNA clone of 2.2 kb was subcloned into pUC18 for further analysis since northern blots of wheat RNA hybridised with pcwx27 allowed the *waxy* protein mRNA size to be estimated at 2.2–2.3 kb (data not shown). The plasmid clone containing the 2.2 kb insert was designated pcSS22.

The complete nucleotide sequence of the cDNA insert of pcSS22 was determined by dideoxy chain termination sequencing [10] of *Bal* 31 deletions and restriction fragments cloned into pUC18. Synthetic oligonucleotide primers to the *waxy* cDNA were also used. Staden programs were used for sequence assembly [11] and the UWGCG programs for sequence analysis. The

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X57233.

```
1     CGGAGTTCCGAAGAGATCAGACCAGTCGTCTCTTGCTGCAGGTAGCCACACCCTGCGCGC  60

61    GCCATGGCGGCTCTGGTCACGTCCCAGCTCGCCACCTCCGGCACCGTCCTCAGCGTCACC  120
         M  A  A  L  V  T  S  Q  L  A  T  S  G  T  V  L  S  V  T
121   GACAGATTCCGGCGTCCAGGTTTTCAGGGCCTGAGGCCCCGGAACCCGGCGGATGCGGCG  180
         D  R  F  R  R  P  G  F  Q  G  L  R  P  R  N  P  A  D  A  A
181   CTCGGCATGAGGACTGTCGGAGCGAGCGCCGCCCCAAAGCAAAGCAGGAAACCGCACCGA  240
         L  G  M  R  T  V  G  A  S  A  A  P  K  Q  S  R  K  P  H  R
241   TTCGACCGGCGGTGCCTCTCCATGGTGGTGCGCGCCACGGGCAGCGGCGGCATGAACCTC  300
         F  D  R  R  C  L  S  M  V  V  R  A  T  G  S  G  G  M  N  L
301   GTGTTCGTCGGCGCCGAGATGGCCCCCTGGAGCAAGACTGGCGGCCTCGGCGACGTCCTC  360
         V  F  V  G  A  E  M  A  P  W  S  K  T  G  G  L  G  D  V  L
361   GGGGGCCTCCCCGCCGCCATGGCCGCCAACGGTCACCGGGTCATGGTCATCTCCCCGCGC  420
         G  G  L  P  A  A  M  A  A  N  G  H  R  V  M  V  I  S  P  R
421   TACGACCAGTACAAGGACGCCTGGGACACCAGCGTCATCTCCGAGATCAAGGTCGTTGAC  480
         Y  D  Q  Y  K  D  A  W  D  T  S  V  I  S  E  I  K  V  V  D
481   AGGTACGAGAGGGTGAGGTACTTCCACTGCTACAAGCGCGGGGTGGACCGCGTGTTCGTC  540
         R  Y  E  R  V  R  Y  F  H  C  Y  K  R  G  V  D  R  V  F  V
541   GACCACCCGTGCTTCCTGGAGAAGGTCCGGGGCAAGACCAAGGAGAAGATCTATGGACCC  600
         D  H  P  C  F  L  E  K  V  R  G  K  T  K  E  K  I  Y  G  P
601   GACGCCGGCACCGACTACGAGGACAACCAGCAGCGCTTCAGCCTTCTCTGCCAGGCAGCA  660
         D  A  G  T  D  Y  E  D  N  Q  Q  R  F  S  L  L  C  Q  A  A
661   CTTGAGGTGCCCAGGATCCTCGACCTCAACAACAACCCACACTTTTCTGGACCCTACGCC  720
         L  E  V  P  R  I  L  D  L  N  N  N  P  H  F  S  G  P  Y  A
721   ATGCTATGCCGTGCCGTGCCGCGCCGCGCAGGGGAAGACGTGGTGTTTGTGTGCAACGAC  780
         M  L  C  R  A  V  P  R  R  A  G  E  D  V  V  F  V  C  N  D
781   TGGCACACGGGCCTTCTGGCCTGCTACCTCAAGAGCAACTACCAGTCCAATGGCATCTAT  840
         W  H  T  G  L  L  A  C  Y  L  K  S  N  Y  Q  S  N  G  I  Y
841   AGGACGGCCAAGGTGGCATTCTGCATCCACAACATCTCGTACCAGGGCCGCTTCTCCTTC  900
         R  T  A  K  V  A  F  C  I  H  N  I  S  Y  Q  G  R  F  S  F
901   GACGACTTCGCGCAGCTCAACCTGCCTGACAGGTTCAAGTCGTCCTTCGACTTCATCGAC  960
         D  D  F  A  Q  L  N  L  P  D  R  F  K  S  S  F  D  F  I  D
941   GGCTACGACAAGCCGGTGGAGGGGCGCAAGATCAACTGGATGAAGGCCGGGATCCTGCAG 1020
         G  Y  D  K  P  V  E  G  R  K  I  N  W  M  K  A  G  I  L  Q
1021  GCCGACAAGGTGCTGACTGTGAGCCCCTACTATGCTGAGGAGCTAATCTCTGGCGAAGCC 1080
         A  D  K  V  L  T  V  S  P  Y  Y  A  E  E  L  I  S  G  E  A
1081  AGGGGCTGCGAGCTCGACAACATCATGCGCCTCACTGGGATCACCGGCATCGTCAACGGC 1140
         R  G  C  E  L  D  N  I  M  R  L  T  G  I  T  G  I  V  N  G
1141  ATGGACGTCAGCGAGTGGGACCCCATCAAGGACAAGTTCCTCACCGTCAACTACGACGTC 1200
         M  D  V  S  E  W  D  P  I  K  D  K  F  L  T  V  N  Y  D  V
1201  ACCACCGCGTTGGAGGGGAAGGCGCTGAACAAGGAGGCGCTGCAGGCCGAGGTGGGGCTG 1260
         T  T  A  L  E  G  K  A  L  N  K  E  A  L  Q  A  E  V  G  L
1261  CCGGTGGACCGGAAGGTGCCCCTGGTGGCGTTCATCGGCAGGCTGGAGGAGCAGAAGGGC 1320
         P  V  D  R  K  V  P  L  V  A  F  I  G  R  L  E  E  Q  K  G
1321  CCCGACGTGATGATCGCCGCCATCCCGGAGATCGTGAAGGAGGAGGACGTCCAGATCGTT 1380
         P  D  V  M  I  A  A  I  P  E  I  V  K  E  E  D  V  Q  I  V
1381  CTCCTGGGCACCGGGAAGAAGAAGTTTGAGCGGCTGCTCAAGAGCGTGGAGGAGAAGTTC 1440
         L  L  G  T  G  K  K  K  F  E  R  L  L  K  S  V  E  E  K  F
1441  CCGACCAAGGTGAGGGCCGTGGTCAGGTTCAACGCGCCGCTGGCTCACCAGATGATGGCC 1500
         P  T  K  V  R  A  V  V  R  F  N  A  P  L  A  H  Q  M  M  A
1501  GGCGCCGACGTGCTGGCGGTCACCAGCCGCTTCGAGCCCTGCGGCCTCATCCAGCTCCAG 1560
         G  A  D  V  L  A  V  T  S  R  F  E  P  C  G  L  I  Q  L  Q
1561  GGAATGCGCTACGGAACGCCGTGCGCCTGCGCGTCGACAGGCGGGCTCGTCGACACTATC 1620
         G  M  R  Y  G  T  P  C  A  C  A  S  T  G  G  L  V  D  T  I
1621  GTGGAAGGCAAGACCGGGTTCCACATGGGCCGCCTCAGCGTTGACTGCAACGTGGTGGAG 1680
         V  E  G  K  T  G  F  H  M  G  R  L  S  V  D  C  N  V  V  E
1681  CCGGCCGACGTGAAGAAGGTGGTCACCACCCTGAAGCGCGCCGTCAAGGTCGTCGGCACG 1740
         P  A  D  V  K  K  V  V  T  T  L  K  R  A  V  K  V  V  G  T
1741  CCGGCGTACCATGAGATGGTCAAGAACTGCATGATACAGGATCTCTCCTGGAAGGGGCCT 1800
         P  A  Y  H  E  M  V  K  N  C  M  I  Q  D  L  S  W  K  G  P
1801  GCCAAGAACTGGGAGGACGTGCTTCTGGAACTGGGGGTGGAGGGGAGCGAGCCGGGCATC 1860
```

```
      A  K  N  W  E  D  V  L  L  E  L  G  V  E  G  S  E  P  G  I
1861  GTCGGCGAGGAGATCGCGCCGCTCGCCCTGGAGAACGTCGCCGCTCCCTGAAGAGAGAAA 1920
      V  G  E  E  I  A  P  L  A  L  E  N  V  A  A  P
1921  GAAGAGGAGCTTCTGGTGCATGGAGCATCCATCCAATCTGCAGGGTTCTCGTATGGGGAG 1980

1981  ATAGCCGCTTGTTGTAGTGAAGAAGGGCCGATATATATATATAATATAGACTAATAAGTA 2040

2041  CTTTAACTTTTGTTGTGCCGCTTGCCTCTTTTACAAACAAAAAAGGAGTTAGGGGTTGTG 2100

2101  CCTATGATAGTGTGCTGAATTGTGCTTGCATTTTGGTGTGGTATATTGCAATAAACAAAG 2160

2161  GATTTGTTAAAAAAAAAAAAAAAAA                                     2186
```

Fig. 1. Complete nucleotide sequence of the wheat *waxy* protein cDNA insert of pcSS22. The derived amino acid sequence is shown underneath the nucleotide sequence. The translation initiation and termination codons and the putative polyadenylation signal are underlined.

cDNA insert in pcSS22 is 2186 nucleotides and includes an open reading frame of 1845 nucleotides from the initiating ATG at position 64 to the termination codon, TGA, at position 1909 (Fig. 1). The cDNA includes a 5′ untranslated region of 63 nucleotides and a 3′ untranslated region of 278 nucleotides. A putative polyadenylation signal at position 2150 precedes a poly(A) tail of 18 residues. The deduced protein contains 615 amino acids with a calculated molecular weight of 67.7 kDa. The mature *waxy* proteins of all species investigated to data are 58–60 kDa as is the case in wheat (data not shown) which suggests the existence of a 7 kDa transit peptide. A 7 kDa transit peptide which targets the maize *waxy* protein to the amyloplast is also 7 kDa [3, 12]. The cDNA sequence of pcSS22 exhibits a high degree of homology with the sequence of the genes encoding the *waxy* proteins of barley, maize and rice.

Acknowledgements

We are grateful to Dr W Rohde (Max-Planck Institut für Züchtüngsforschung, D-5000 Köln 30, FRG) for kindly providing the plasmid pcwx27.

References

1. Echt CS, Schwartz D: Evidence for the inclusion of controlling elements within the structural gene at the *waxy* locus in maize. Genetics 99: 275–284 (1981).
2. Nelson OE, Chourey PS, Chang MT: Nucleoside diphosphate sugar-starch glucosyl transferase activity of *wx* starch granules. Plant Physiol 62: 383–386 (1978).
3. Shure M, Wessler S, Federoff N: Molecular identification and isolation of the *waxy* locus in maize. Cell 35: 225–233 (1983).
4. MacDonald FD, Preiss J: Partial purification and characterisation of granule-bound starch synthases from normal and *waxy* maize. Plant Physiol 78: 849–852 (1985).
5. Hovenkamp-Hermelink JHM, Jacobsen E, Ponstein AS, Visser RGF, Vos-Scheperkeuter GH, Bijmolt EW, de Vries JN, Witholt B, Feenstra WJ: Isolation of an amylose-free mutant of the potato (*Solanum tuberosum* L.). Theor Appl Genet 75: 217–221 (1987).
6. Smith AM: Evidence that the 'waxy' protein of pea (*Pisum sativum* L.) is not the major starch-granule-bound starch synthase. Planta 182: 599–604 (1990).
7. Wang Z, Wu Z, Xing Y, Zheng F, Guo X, Zhang W, Hong M: Nucleotide sequence of the rice *waxy* gene. Nucl Acids Res 18: 5898 (1990).
8. Rohde W, Becker D, Salamini F: Structural analysis of the *waxy* locus from *Hordeum vulgare*. Nucl Acids Res 16: 7185–7186 (1988).
9. Visser RGF, Hergersberg M, van der Leij FR, Jacobsen E, Witholt B, Feenstra WJ: Molecular cloning and partial characterisation of the gene for granule-bound starch synthase from a wild-type and an amylose-free potato (*Solanum tuberosum* L.). Plant Sci 64: 185–192 (1989).
10. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain terminating inhibitors. proc Natl Acad Sci USA 74: 5463–5467 (1977).
11. Staden R: Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucl Acids Res 10: 4731–4751 (1982).
12. Klösgen RF, Saedler H, Weil J: The amyloplast-targeting transit peptide of the *waxy* protein of maize also mediates protein transport *in vitro* into chloroplasts. Mol Gen Genet 217: 155–161 (1989).