Volume 13 Number 20 1985 Nucleic Acids Research

Nucleotide sequence of a B1 hordein gene and the identification of possible upstream regulatory elements in endosperm storage protein genes from barley, wheat and maize

B.G.Forde, A.Heyworth, J.Pywell and M.Kreis

Biochemistry Department, Rothamsted Experimental Station, Harpenden, Herts AL5 2JQ, UK

Received 19 July 1985; Revised and Accepted 30 September 1985

ABSTRACT

The B-hordeins are the major group of prolamin storage proteins in barley (Hordeum vulgare L.) and they are encoded by a small multigene family that is expressed specifically in the developing endosperm. We report the complete nucleotide sequence of a clone of one B-hordein gene (pBHR184). The cloned gene contains no introns and belongs to the B1 sub-family of B-hordein genes. Comparison of the 5'-flanking sequences of pBHR184 with those of related S-rich prolamin genes from wheat shows that several short sequences within 600 bp upstream of the translation initiation codon are strongly conserved. A sequence that is conserved at around -300 bp in the S-rich prolamins is also conserved at similar locations in genes encoding the two major classes of maize prolamin (the Z19 and Z21 zeins) and appears to be unique to prolamin genes. We discuss the possible role of this '-300 element' in the control of gene expression in the developing cereal endosperm.

INTRODUCTION

In most cereal species the major seed storage proteins are prolamins, a complex group of alcohol-soluble polypeptides that make up about half of the protein in the mature grain. In barley (Hordeum vulgare L.) they are classified into three main groups (B-, C- and D-hordeins), which are specified by separate compound genetic loci on chromosome 5 (1). Prolamins homologous to the hordeins are found in wheat (the gliadins and glutenins) and in rye (the secalins), while the major prolamins of the more distantly related maize (the zeins) seem to have evolved independently (1,2).

Synthesis of the prolamin polypeptides is endosperm-specific and is initiated coordinately at a relatively late stage of seed development (3-5). In barley, expression of some families and sub-families of hordein genes is modulated by the balance of nitrogen and sulphur nutrition (6) and by a mutation at an unlinked 'regulatory' locus (7,8). In maize, mutants that alter either the timing (9) or the rate (10) of zein deposition have been reported, some of which specifically affect synthesis of one or other of the two zein classes. Thus it appears that there are at least two types

of control operating on prolamin gene expression, one responsible for coordinate induction of the genes during endosperm development and another regulating the subsequent rate of prolamin accumulation, and these controls have the ability to act differentially on subsets of prolamin genes.

As part of our study of the organization and expression of the hordein gene families we now report the isolation and nucleotide sequencing of a B-hordein genomic clone. We discuss the possible significance of short upstream sequences that are conserved in B-hordein and α-gliadin genes and in genes encoding the two major classes of zeins.

METHODS

Screening a barley genomic library

A genomic library of barley DNA (*Hordeum vulgare* L., cv. Sundance) was generously provided by Dr. M. Murray and Dr. J. Slightom (Agrigenetics Corporation, Madison). The unamplified library (1 x 10^6 recombinant phage), which had been prepared by cloning a partial EcoRI digest of high molecular weight barley DNA in Charon 32, was screened by plaque hybridization (11) using as probe the nick-translated (12) cDNA inserts from pB7 and pB11 (13). Hybridizing clones were plaque-purified and phage DNA was prepared by a plate-lysate method (14).

Nucleotide sequencing

A 2.9 kb EcoRI fragment from λHvBH3.4 was sub-cloned in pUC9 for sequencing. Plasmid DNA of the subclone (pBHR184) was prepared from cells lysed with Triton X-100 (15) and further purified by banding twice in CsCl gradients. Fragments suitable for sequencing were generated by BAL-31 deletion. Three μg of pBHR184 was linearized with the appropriate restriction enzyme (see Fig. 1) and digested at a rate of 80 bp/min/end using 0.7 U BAL-31 at 37° in a 60 μl volume (16). Aliquots taken at 2 min intervals up to 18 min were phenol-extracted, digested with a second restriction enzyme and cloned in *E. coli* strain JM101 using as vectors M13 mp8 (17) or mp19 (supplied by Pharmacia Ltd.). Sequencing was by the dideoxy method (18) and sequences were assembled and analysed with the assistance Staden programs (19-21) operating on a VAX 11/750 computer.

S1 mapping

Fragments of the cloned B1 hordein gene were prepared for S1 protection analysis as follows. Single-stranded phage DNA was prepared from two M13 clones that contained the 2.9 kb EcoRI fragment from pBHR184 in opposite orientations. Three μg of each phage was mixed in 10 μl of 50 mM Tris HCl

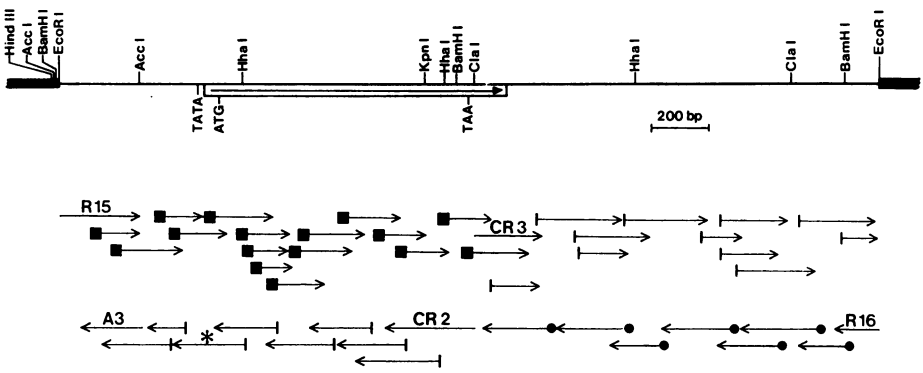
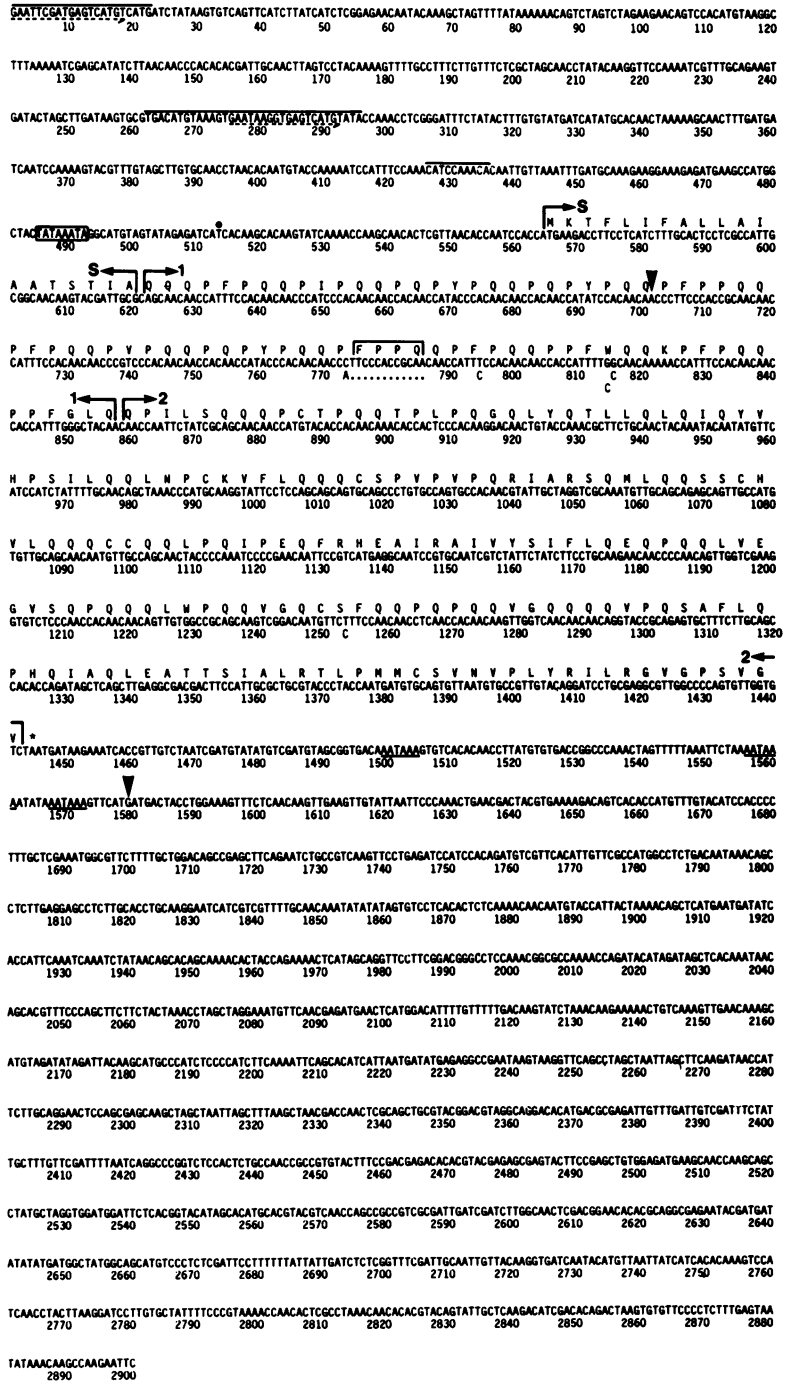


Fig. 1. Restriction map of a B-hordein genomic clone (pBHR184) and the sequencing strategy. The clone was constructed by sub-cloning a 2.9 kb EcoRI fragment from λHvBH3.4 into pUC9. The positions of the AccI and BamHI sites were determined experimentally and the locations of the other sites were obtained from the sequence. The open rectangle indicates the region corresponding to the mature mRNA (as deduced from subsequent analysis) and the arrow within it shows the direction of transcription. Arrows with closed squares indicate sequences obtained from M13 sub-clones generated by digestion of pBHR184 with, in turn, HindIII, BAL-31 and EcoRI (see Methods). For sequences indicated by arrows with closed circles the order was: EcoRI, BAL-31, ClaI; for those indicated by arrows with vertical bars it was: ClaI, BAL-31, EcoRI. The remaining sequences were obtained by sub-cloning restriction fragments without BAL-31 digestion. The asterisk indicates the M13 sub-clone (Cla42) that was used in Fig. 3 to provide size markers.

pH8.0, 50 mM NaCl and incubated at 60° for 30 min. The mixture was made to 10 mM $MgCl_2$ and the DNA digested with 4 U HhaI. Because the annealed phage molecules are only double-stranded for the length of their inserts the HhaI digest generates only three fragments. The HhaI fragments (3 µg) were end-labelled with ^{32}P using T4 polynucleotide kinase (Bethesda Research Laboratories Inc.). Hybridization to 7 µg poly A^+ RNA from barley endosperm (cv. Sundance), and subsequent treatment with S1 nuclease, were according to Berk and Sharp (22). Size markers were generated by performing sequencing reactions on an M13 sub-clone of pBHR184 (Cla42) that contains the 5'-flanking region of the gene and a short region of coding sequence, including the HhaI site immediately downstream from the translation initiation codon (see Fig. 1). To generate dideoxy-terminated fragments with the same 5' end as the protected fragment to be mapped, the products of the sequencing reaction were treated with HhaI and BamHI before electrophoresis. BamHI cuts within the polylinker region of Cla42 and was used to reduce the size of the fragments that contained the sequencing



primer, some of which would otherwise have co-migrated with the size-markers.

RESULTS AND DISCUSSION

Cloning and sequence analysis of a B1 hordein gene

A barley genomic library was screened for B-hordein genes using a mixed probe consisting of two cDNA clones, pB7 and pB11, which represent the two major sub-families of B-hordein mRNA (13). Of three clones that were plaque-purified and mapped by restriction digests, one was selected for detailed analysis. This clone, λHvBH3.4, contains a barley DNA fragment of 17.4 kb, which is made up of four EcoRI fragments of 7.4 kb, 3.6 kb, 3.5 kb and 2.9 kb. Southern blots (23) of the EcoRI digest revealed that only the latter fragment hybridized to the cDNA probe (not shown). Fig. 1 shows a restriction map of a sub-clone of the 2.9 kb fragment (pBHR184) and the sequencing strategy.

The nucleotide sequence of the 2.9 kb fragment is presented in Fig. 2. Nucleotides 564 to 1442 constitute an open reading frame that begins with an ATG codon and encodes a B-hordein-like polypeptide. Comparison with the sequences of the two cDNA clones that were used as probes shows that pBHR184 is much more closely related to pB11, a B1-type hordein cDNA clone, than to pB7, a B3-type (13). The alignment between pBHR184 and pB11 (which is incomplete at the 5' end) extends from nucleotides 701 to 1579, where the poly(A) tail in the cDNA clone begins. The coding sequence of the genomic fragment contains a 12 nucleotide sequence (positions 774-785) that is absent in pB11. There are only 4 other mismatches between the two sequences, and only one of these is a replacement substitution (Fig. 2). The differences between the cDNA and genomic sequences are consistent with previous evidence indicating nucleotide and amino acid sequence variations in the B-hordein multigene family (12, 24). As with other cereal prolamin

Fig. 2. Nucleotide sequence of pBHR184 and the derived amino acid sequence of a B1 hordein. The vertical arrowheads show the extent of the cDNA sequence in the B1 hordein cDNA clone, pB11 (13). The four nucleotides and one amino acid in the pB11 sequence that differ from pBHR184 are given below the main sequence. The four bracketed amino acid residues are absent in the pB11 sequence. The proposed extent of the signal peptide and the limits of domains 1 and 2 of the mature protein (13) are indicated by the right-angled arrows. A TATA box sequence (see Fig. 5) is enclosed by a rectangle, and the approximate transcription start site determined from S1 mapping (Fig. 3) is indicated by a dot. Three possible polyadenylation signal sequences are underlined; other putative control sequences discussed in the text are overlined. The dashed arrows indicate an imperfect repeat.

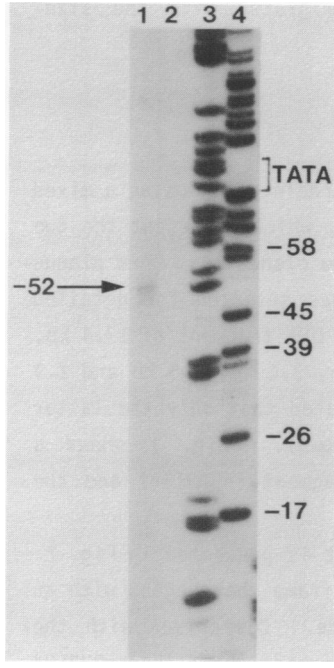


Fig. 3. S1 mapping of the mRNA cap site in pBHR184. HhaI fragments were generated from the B1 hordein clone as described in Methods. The fragments were end-labelled with ^{32}P and hybridized to poly A^+ RNA from barley endosperm. One aliquot of the hybridization was treated with 334 U S1 nuclease (lane 1) and a second aliquot with 835 U S1 nuclease (lane 2). The digestion products were analysed on a sequencing gel in parallel with two sequencing reactions, an 'A' track (lane 3) and a 'C' track (lane 4), and the gel was autoradiographed. To provide precise size markers the fragments synthesized in the sequencing reactions contained the same sequence and had the same 5' termini as the protected fragment (see Methods). The fragments in the 'C' track are numbered relative to the translation initiation codon (the first base upstream of the ATG codon being -1) and the location of the TATA box sequence is bracketed. The arrow indicates the position of the major protected fragment.

genes, there is no evidence for the presence of introns (25-32).

In the 3'-untranslated region of the gene there are several hexanucleotide sequences (AATAAA) that conform to the putative polyadenylation signal sequence (33) and there may therefore be several alternative polyadenylation sites. By analogy with pB11, which is identical to pBHR184 in the 3'-untranslated region, it is likely that there is a polyadenylation site at position 1579.

We have used an S1 protection assay to map the mRNA cap site in pBHR184. A HhaI digest of the cloned gene was end-labelled with ^{32}P and annealed to poly A^+ RNA from barley endosperm under R-looping conditions. The fragments were then treated with two concentrations of S1 nuclease and analysed on a sequencing gel (Fig. 3, lanes 1 and 2). Precise mapping of the cap site was achieved by using size markers that contained the same sequence as the protected fragments (Fig. 3, lanes 3 and 4). The major protected fragment migrates at the position corresponding to 52 bp upstream from the translation initiation codon. Allowing one base for the mRNA cap, we estimate the transcription start site to be at position -51 relative to the ATG codon (see Fig. 2). Assuming a poly(A) tail of 80 residues (34),

the predicted length of the mRNA is 1150 nucleotides, which is in good agreement with previous estimates based on Northern blots (35).

Primary structure of the pBHR184 gene product

The open reading frame that starts at nucleotide 564 (Fig. 2) encodes a protein with 293 residues (M_r 33,423). The amino terminal region of the protein has many of the characteristics of a signal peptide, including a charged residue near the N-terminus and a core of hydrophobic residues (36). The presence of a signal peptide would be consistent with the evidence that the B-hordeins are synthesized on the rough endoplasmic reticulum and deposited in protein bodies (34,37,38), and with the observations that they are synthesized *in vitro* as larger precursors (34,39). However, because the N-termini of the B-hordeins are blocked (40,41), the N-terminal sequence of the mature protein is not known and we cannot assign the site of signal peptide cleavage with certainty. Nevertheless, comparison with the nascent and mature sequences of the homologous α-gliadin storage proteins from wheat (42) suggests that cleavage would occur between residues 19 and 20. The same cleavage site is predicted by application of the rules formulated by von Heijne (43). The mature protein would therefore consist of 274 residues and have a M_r of 31,444, which agrees well with estimates based on direct analysis of the B-hordeins (40).

Immediately following the putative signal peptide there is a region of the protein that is extremely rich in proline and glutamine. This part of the protein was previously identified as domain 1, one of two domains of primary structure in B-hordein polypeptides, which is also characterized by a series of degenerate tandem repeats (13). The short N-terminal domain that precedes the proline-rich repeats in other S-rich prolamins is absent from this B-hordein polypeptide (2). The boundary between domain 1 and domain 2, as defined from the earlier analysis of the cDNA sequences (13), is indicated in Fig. 2. Domain 1 of the pBHR184-encoded protein contains 79 residues which are 39% glutamine, 39% proline, 10% phenylalanine and include no sulphur amino acids. The repeated motif (based on the prototype octapeptide Pro-Gln-Gln-Pro-Phe-Pro-Gln-Gln) is evident throughout this domain, including the N-terminal 27 residues not previously sequenced (see ref. 2). The remaining 198 residues of the protein make up domain 2, which is distinguished from domain 1 by being relatively proline-poor, S-rich and non-repetitive (13). Domain 2 is 27% glutamine, 11% proline, 4.1% cysteine and 1.5% methionine.

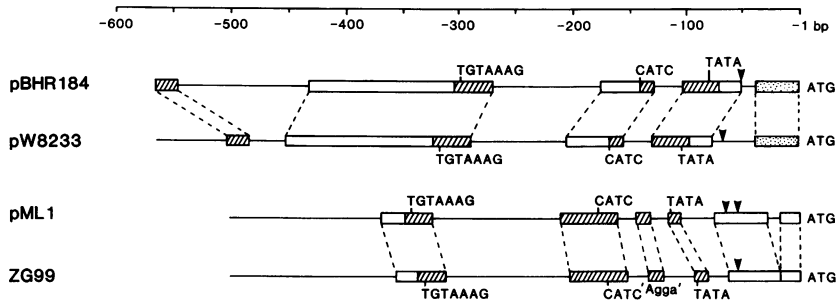


Fig. 4. Diagram showing the locations of conserved sequences in the 5'-flanking regions of two families of prolamin genes. The B1 hordein (pBHR184) and an α-gliadin gene from wheat (pW8233; ref. 30) are divergent representatives of the S-rich prolamin multi-gene family, while the Z21 (pML1; ref. 45) and Z19 (ZG99; ref. 25) genes are similarly divergent representatives of the zein multi-gene family. A graphic matrix homology plot was used to locate the most strongly conserved sequences in the 5'-flanking regions of each pair of genes and the sequences were then aligned manually to assess the extent of lower order homologies. The degree of homology is indicated by the shading within the rectangles. Diagonal shading: 80-92% identity; stippled: 74% identity; open: 56-62% identity. Sequences outside the rectangles show little or no homology (<40%). Short 'core' sequences that characterize the most strongly conserved blocks of homology are indicated. The conserved 'Agga box' sequence that was noted in a survey of the promoter regions of a number of plant genes (48), including the zein gene family, was not found in the B1 hordein or α-gliadin genes. The 5'-flanking regions of three other α-gliadin genes have been sequenced (31,32) and do not differ by more than 9% from the pW8233 gene that was used for these comparisons. Upstream sequences of at least 210 bp have been determined for two other Z21 genes, and these are 85% (Z7; ref. 29) and 90% (ZA1; ref. 27) homologous to pML1, and for three Z19 genes, which are 99% (Z4; ref. 26), 96% (ZE19; ref. 28) and 88% (ZE25; ref. 28) homologous to ZG99.

Conserved sequences in the 5'-flanking region

On the assumption that sequences important in gene expression are likely to be conserved among a group of genes with the same pattern of expression (44), we have carried out a detailed comparison between the flanking sequences of the B1 hordein gene and those of an α-gliadin gene from wheat. Although they are related proteins, the B1 hordeins and the α-gliadins are among the most divergent forms of the S-rich prolamins (2). A diagrammatic representation of the sequence homologies upstream of the two S-rich prolamin genes is shown in Fig. 4, along with a similar comparison between two maize genes that code for polypeptides belonging to the light chain (Z19) and heavy chain (Z21) classes of zein. Despite considerable divergence between the 5'-flanking sequences of the B1 hordein and α-gliadin

genes (<50% overall homology upstream of the cap sites) there are several short segments within 600 bp of the translation initiation codons that show more than 80% homology. Significantly, several of these most strongly conserved segments are related to sequences that are conserved at similar locations in the zein multi-gene family. Two of the sequences that are common to both multi-gene families are located at around -100 and -150 (relative to the ATG codon) and may be the counterparts of the TATA and CCAAT boxes that are components of the promoter region of animal genes (46, 47). A TATA box sequence has been found in almost all plant genes so far analysed (48), while a sequence similar to the CCAAT box-like segment in the prolamin genes is also found in a wide variety of other cereal genes (Fig. 5A) (although not in the maize Adh2 gene (53)). The CCAAT box-like sequence has been designated a 'CATC' box by virtue of the most strongly conserved tetranucleotide within the 11 bp segment. Sequences conforming to the CATC box consensus (Fig. 5A) were not found in a manual survey of the corresponding region upstream of the TATA box in 15 published dicot gene sequences. Despite the absence of a clearly identifiable CCAAT box-like sequence that is common to both monocots and dicots (see also ref. 48), the importance of sequences in this region for maximal gene expression in plants has been demonstrated for two genes (54, 55).

Potentially the most interesting conserved sequences in the 5'-flanking regions of the two families of prolamin genes are found about 300 bp upstream of the ATG codon (Fig. 4). These sequences, or '-300 elements', are aligned in Fig. 5B to illustrate the features that are common to both multi-gene families. In the α-gliadin gene the -300 element is imperfectly repeated about 200 bp further upstream and in the B1 hordein gene at least part of the element is imperfectly repeated about 270 bp upstream (Fig. 5B). Sequences homologous to the -300 element were not found in the other cereal genes for which extensive upstream sequence data are available (49, 50, 52, 53).

An indication of the very low frequency of random occurrence of sequences related to the -300 element was obtained by searching the GenBank nucleotide sequence database (release 18.0, 3 x 10^6 nucleotides) for all occurrences of a 28 bp consensus sequence: ANNTGTAAAGWWAATNNG ATGAWWCATG (where W = G or T, and invariant nucleotides are underlined). The consensus sequence was derived by aligning the -300 element and its repeats in the B1 hordein gene and in all published α-gliadin (30-32) and Z21 zein (27, 29, 44) sequences (without introducing gaps). The Z19

A

Clone	'CATC box'	'TATA Box'	Cap Site(s)
pBHR184:	-139 ACATCCAAACA.....-80	CTATAAATA.....(-51)....	
pW8233:	-165 GCATCCAAGCA.....-105	CTATAAATA.....(-67)....	
pML1:	-177 CCATCTATACC.....-112	GTATAAGCA....(-65,-52)..	
ZG99:	-168 TCATCTCTACC.....-90	GTATAAATA.....(-57)....	
p1S.1:	-167 CCATCTCTTCC.....-140	CTATATAAA.....(-99)....	
pTH012:	-201 CCATCTCGACC.....-97	CTATTTAAC.....(-62)....	
pWS4.3:	-113 CCATCCCAACC.....-79	CTATATATA.....(-42)....	
Consensus:	CCATC$^{TCT}_{CAA}$ACC	CTATA$^{T}_{A}$A$^{T}_{A}$A	
Animal Consensus:	GG$^{C}_{T}$CAATCT	TATA$^{T}_{A}$A	

B

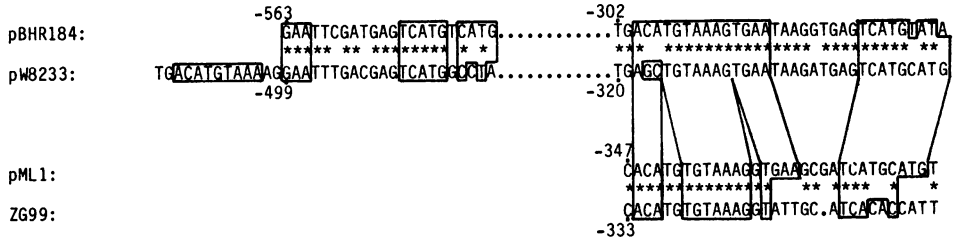


Fig. 5. Sequences common to the 5'-flanking regions of the B1 hordein (pBHR184), α-gliadin (pW8233), Z21 (pML1) and Z19 (ZG99) genes. Sequences that are common to all four genes analysed in Fig. 4 have been aligned here for comparison. (A) Conserved sequences located within 200 bp upstream of the ATG codon. Similar sequences are found at corresponding positions in a variety of other cereal genes (49-52) and a representative selection of these is also shown. p1S.1: maize alcohol dehydrogenase 1 (Adh1) gene (50); pTH012: wheat H3 histone gene (51); pWS4.3: wheat Rubisco small subunit gene (51). The consensus sequence for the TATA box is similar to that previously determined for a group of genes from both monocots and dicots (48). (B) Conserved distal sequences that are located at around position -300 in both families of prolamin gene and that are repeated further upstream in the α-gliadin gene and (at least partially) in the B1 hordein gene. The sequence of the repeat of the -300 element in the B1 hordein gene is incomplete because the repeat contains one of the two EcoR1 sites that define the boundaries of the sequenced 2.9 kb fragment. Asterisks indicate identical residues in each pair of aligned sequences, while those parts of the element that are common to the S-rich prolamin and zein genes are boxed. Numbering is relative to the ATG codon. The corresponding sequences in other α-gliadin (31, 32), Z21 (27, 29) and Z19 (26, 28) genes are similar or identical to the representative sequences in this figure.

sequences were omitted because they deviate markedly from the consensus in the second half of the -300 element (see Fig. 5B). Even allowing up to three mismatches at the variant positions (the maximum deviation shown by

any of the sequences that were used to derive the consensus), no additional occurrences were found in the database. Therefore, in view of the apparently independent evolutionary origins of the S-rich prolamin and zein genes, it seems that the presence of this sequence at approximately the same location in both genes is highly unlikely to have arisen by chance. This suggests that the -300 element has some important function, which (in view of its location) is most likely to be related to the control of prolamin gene expression. Conserved 5'-flanking sequences have previously been noted in another group of developmentally co-regulated plant genes (56) and there is accumulating evidence that short upstream sequences are involved in the coordinate induction of unlinked eucaryotic genes (56-58). We are currently investigating the possibility that, by analogy with Drosophila heat shock genes (60), there is a specific interaction between DNA-binding proteins in endosperm nuclei and conserved upstream sequences in the B1 hordein gene.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. B. J. Miflin for many helpful discussions throughout the course of this work and for critical reading of the manuscript. The research was supported partly by contract no. GBI-4-023-UK(H) of the Biomolecular Engineering Programme of the Commission of the European Communities.

REFERENCES

1 Shewry, P.R., Miflin, B.J. and Kasarda, D.D. (1984) Phil. Trans. R. Soc. Lond. B 304, 297-308.
2 Kreis, M., Shewry, P.R., Forde, B.G., Forde, J. and Miflin, B.J. (1985) Oxf. Surveys Plant Molec. Cell Biol. 2, 253-317.
3 Greene, F.C. (1983) Plant Physiol. 71, 40-46.
4 Rahman, S., Shewry, P.R. and Miflin, B.J. (1982) J. Exp. Bot. 33, 717-728.
5 Misra, P.S. and Mertz, E.T. (1975) Cereal Chem. 52, 734-739.
6 Rahman, S., Shewry, P.R., Forde, B.G., Kreis, M. and Miflin, B.J. (1983) Planta 159, 366-372.
7 Thompson, R.D. and Bartels, D. (1983) Plant Sci. Lett. 29, 295-304.
8 Kreis, M., Shewry, P.R., Forde, B.G., Rahman, S., Bahramian, M.B. and Miflin, B.J. (1984) Biochem. Genet. 22, 231-255.
9 Manzocchi, L.A., Daminati, M.G. and Gentinetta, E. (1980) Maydica 25, 199-210.
10 Soave, C. and Salamini, F. (1983) in Seed Proteins, Daussant, J., Mosse, J. and Vaughan, J. Eds., pp. 205-218, Academic Press, London.
11 Benton, W.D. and Davis, R.W. (1977) Science 196, 180-182.
12 Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.

13 Forde, B.G., Kreis, M., Williamson, M.S., Fry, R.P., Pywell, J., Shewry, P.R., Bunce, N. and Miflin, B.J. (1985) EMBO J. 4, 9-15.
14 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, New York.
15 Clewell, D.B. and Helinski, D.R. (1970) Biochemistry 9, 4428-4440.
16 Guo, L-H., Yang, R.C.A. and Wu, R. (1983) Nucl. Acids Res. 16, 5521-5540.
17 Messing, J. (1983) Meth. Enzymol. 101, 20-78.
18 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
19 Staden, R. (1982) Nucl. Acids Res. 10, 2951-2961.
20 Staden, R. (1982) Nucl. Acids Res. 10, 4731-4751.
21 Staden, R. (1984) Nucl. Acids Res. 12, 505-519.
22 Berk, A.J. and Sharp, P.A. (1977) Cell 12, 721-732.
23 Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
24 Rasmussen, S.K., Hopp, H.E. and Brandt, A. (1983) Carlsberg Res. Commun. 48, 187-199.
25 Pedersen, K., Devereux, J., Wilson, D.R., Sheldon, E. and Larkins, B.A. (1982) Cell 29, 1015-1026.
26 Hu, N-T., Peifer, M.A., Heidecker, G., Messing, J. and Rubenstein, I. (1982) EMBO J. 1, 1337-1342.
27 Spena, A., Viotti, A. and Pirotta, V. (1983) EMBO J. 2, 1589-1594.
28 Spena, A., Viotti, A. and Pirotta, V. (1983) J. Mol. Biol. 169, 799-811.
29 Kridl, J.C., Vieira, J., Rubenstein, I. and Messing, J. (1984) Gene 28, 113-118.
30 Rafalski, J.A., Scheets, K., Metzler, M., Peterson, D.M., Hedgcoth, C. and Soll, D.G. (1984) EMBO J. 3, 1409-1415.
31 Anderson, O.D., Litts, J.C., Gautier, M.-F. and Greene, F.C. (1984) Nucl. Acids Res. 12, 8129-8144.
32 Sumner-Smith, M., Rafalski, J.A., Sugiyama, T., Stoll, M. and Soll, D. (1985) Nucl. Acids Res. 13, 3905-3916.
33 Proudfoot, N.J. and Brownlee, G.G. (1976) Nature 263, 211-214.
34 Matthews, J.A. and Miflin, B.J. (1980) Planta 149, 262-268.
35 Forde, B.G., Kreis, M., Bahramian, M.B., Matthews, J.A., Miflin, B.J., Thompson, R.D., Bartels, D. and Flavell, R.B. (1981) Nucl. Acids Res. 9, 6689-6707.
36 Watson, M.E.E. (1984) Nucl. Acids Res. 12, 5145-5164.
37 Cameron-Mills, V., Ingversen, J. and Brandt, A. (1978) Carlsberg Res. Commun. 43, 91-102.
38 Cameron-Mills, V. (1980) Carlsberg Res. Commun. 45, 557-576.
39 Brandt, A. and Ingversen, J. (1978) Carlsberg Res. Commun. 43, 451-469.
40 Shewry, P.R., Field, J.M., Kirkman, M.A., Faulks, A.J. and Miflin, B.J. (1980) J. Exp. Bot. 31, 393-407.
41 Schmitt, J.M. and Svendsen, I. (1980) Carlsberg Res. Commun. 45, 143-148.
42 Kasarda, D., Okita, T.W., Bernardin, J., Baecker, P.A., Nimmo, C., Lew, E., Dietler, M. and Greene, F.C. (1984) Proc. Natl. Acad. Sci. USA 81, 4712-4716.
43 Von Heijne, G. (1983) Eur. J. Biochem. 133, 17-21.
44 Davidson, E.H., Jacobs, H.T. and Britten, R.J. (1983) Nature 301, 468-470.
45 Langridge, P. and Feix, G. (1983) Cell 34, 1015-1022.
46 Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) Nucl. Acids Res. 8, 127-142.
47 Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell,

C., Spritz, R.A., De Riel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. and Proudfoot N.J. (1980) Cell 21, 653-668.
48 Messing, J., Geraghty, D., Heidecker, G., Hu, N-T., Kridl, J. and Rubsenstein, I. (1983) in Genetic engineering of Plants, Kosuge, T., Meredith, C.P. and Hollaender, A. Eds., pp. 211-227, Plenum Press, New York.
49 Dennis, E.S., Gerlach, W.L., Pryor, A.J., Bennetzen, J.L., Inglis, A., Llewellyn, D., Sachs, M.M., Ferl, R.J. and Peacock, W.J. (1984) Nucl. Acids Res. 12, 3983-4000.
50 Tabata, T., Fukasawa, M. and Iwabuchi, M. (1984) Mol. Gen. Genet. 196, 397-400.
51 Broglie, R., Coruzzi, G., Lamppa, G., Keith, B. and Chua, N-H. (1983) Biotechnology 1, 55-61.
52 Tabata, T., Sasaki, K. and Iwabuchi, M. (1983) Nucl. Acids. Res. 11, 5865-5875.
53 Dennis, E.S., Sachs, M.M., Gerlach, W.L., Finnegan, E.J. and Peacock, W.J.(1985) Nucl. Acids Res. 13, 727-743.
54 Shaw, C.H., Carter, G.H., Watson, M.D. and Shaw, C.H. (1984) Nucl. Acids Res. 12, 7831-7846.
55 Odell, J.T., Nagy, F. and Chua, N-H. (1985) Nature 313, 810-812.
56 Mauro, V.P., Nguyen, T., Katinakis, P. and Verma, D.P.S. (1985) Nucl. Acids. Res. 13, 239-249.
57 Pelham, H.R.B. (1982) Cell 30, 517-528.
58 Donahue, T.F., Daves, R.S., Lucchini, G. and Fink, G.R. (1983) Cell 32, 89-98.
59 Miller, A.M., MacKay, V.L. and Nasmyth, K.A. (1985) Nature 314, 598-603.
60 Wu, C. (1984) Nature 311, 81-84