

---

**Molecular cloning and analysis of cDNA sequences derived from polyA<sup>+</sup> RNA from barley endosperm: identification of B hordein related clones**

---

B.G.Forde, M.Kreis, M.B.Bahramian, J.A.Matthews and B.J.Mifflin  
Biochemistry Department, Rothamsted Experimental Station, Harpenden, Herts. AL5 2JQ, and

R.D.Thompson, D.Bartels and R.B.Flavell  
Cytogenetics Department, Plant Breeding Institute, Maris Lane, Trumpington, Cambridge CB2 2LQ, UK

---

Received 6 November 1981

---

**ABSTRACT**

A collection of over 130 cDNA clones has been constructed in the bacterial plasmids pPH207 and pBR322 using as template the poly A<sup>+</sup> RNA from membrane-bound polysomes of barley endosperm (cv. Sundance). Fifty four B hordein cDNA clones have been identified by cross-hybridization analysis and *in vitro* translation of plasmid-selected mRNAs. Hybridization of II of the B hordein cDNA clones to Northern blots of size-fractionated RNA indicated that the B hordein mRNA is ca. 1300 nucleotides long. One cDNA clone, pHvE-cl6, has been partially sequenced and shown by comparison with C-terminal and other peptide sequences to be related to B1 hordein polypeptides. The results obtained from the analysis of the B hordein cDNA clones support the idea that the Hor 2 locus, which specifies the B hordeins, is complex and codes for a family of related mRNA species.

**INTRODUCTION**

The alcohol-soluble prolamins are the chief store of nitrogen in the barley seed. These proteins, termed hordeins, are found only in the endosperm and make up about 40% of the seed protein. They can be separated into two groups, termed 'B' and 'C', each of which is composed of a series of polypeptides that can be separated by various one and two dimensional gel electrophoretic systems (1). Although the two groups have many properties in common they differ in size and in the relative amounts of certain amino acids (2). Genetic studies have identified two linked structural loci, Hor-1 and Hor-2, which specify the C and B polypeptides respectively (3-5). Each of these loci could represent relatively long regions of the chromosome since the number of seeds analysed was insufficient to resolve loci closer than 0.01 centimorgans. Furthermore, two dimensional analysis of the B polypeptides from several cultivars, and comparison of the cyanogen bromide fragmentation pattern of individual polypeptides (6) supports the hypothesis that they are a

polymorphic series of proteins specified by a complex locus that has evolved by cycles of gene duplication and mutation. In order to test this further we have synthesized cDNA from poly A<sup>+</sup> RNA derived from the membrane-bound polysomes of developing endosperms. In this paper we report the construction and analysis of clones containing inserts of double-stranded cDNA and show that some of these clones contain DNA sequences directly related to B hordein polypeptides. The construction and preliminary identification of a B hordein cDNA clone has been reported previously (7).

#### METHODS

##### Preparation of RNA from membrane-bound polysomes

Barley plants (Hordeum vulgare L. cv. Sundance) were grown in the field and endosperms isolated 18-22 days after flowering as described previously (8). Etiolated barley shoots were grown in the dark for 5 days at 28°C. After harvest shoots and endosperms were frozen in liquid N<sub>2</sub> and stored at -80°C. Membrane-bound polysomes were isolated as described earlier (8,9). For purification of mRNA, the membrane-bound polysomes were disrupted in SDS buffer [0.2 M Tris pH 9, 0.1 M NaCl, 10 mM EDTA, 0.5% sodium dodecyl sulphate (SDS)], extracted with phenol/chloroform (10) and the RNA was precipitated at -20°C with 2.5 vol. of ethanol and 0.05 vol. of 4 M ammonium acetate. The polysomal RNA was subjected to chromatography on oligo (dT)-cellulose (11) and the poly A<sup>+</sup> RNA recovered by precipitation with ethanol. Unless otherwise stated the term poly A<sup>+</sup> RNA refers to this fraction from developing barley endosperms.

##### Synthesis of cDNA

The methods used were based on those of Efstratiadis et al. (12,13). Poly A<sup>+</sup> RNA (25 µg) was heated to 60°C for 4 min. Following addition of 2 µg oligo (dT)<sub>12-18</sub>, 0.5 µmol dithiothreitol (DTT), 1.5 µmol MgCl<sub>2</sub>, 15 µmol NaCl and 12.5 µmol 1M Tris-HCl pH 8.3, the primer and RNA were allowed to anneal in 85 µl for 15 min at 40°C. Before starting the reaction by addition of 40 U of avian myeloblastosis virus reverse transcriptase (the generous gift of Dr. J. Beard, Life Sciences, St. Petersburg, FL) the incubation mixture was supplemented with 0.5 µmol each of dCTP, dGTP and TTP, 50 nmol of α-[<sup>32</sup>P]dATP (1000 Ci/mol) and 17.5 µg of actinomycin D. The mixture, in a final volume of 250 µl, was incubated for 60 min at 40°C. The reaction was stopped by chilling on ice and adding EDTA to 4 mM. RNA was hydrolysed in 0.3 M NaOH at room

temperature for 60 min, and the solution neutralised with 0.5 vol. 0.6 M HCl, 0.6 M Tris-HCl pH 7.5. Following extraction with phenol/chloroform (1:1) the aqueous phase was passed through a Sephadex G50 column equilibrated with 0.3 M NaCl, 50 mM Tris-HCl pH 7.5. The [ $^{32}\text{P}$ ]cDNA in the void volume was precipitated at  $-20^{\circ}\text{C}$  by addition of 2.5 vol ethanol. The cDNA (ca. 2.5  $\mu\text{g}$ ) was made double-stranded by incubation with 175 U of *E. coli* DNA polymerase (Boehringer) at  $30^{\circ}\text{C}$  for 3 h in a final volume of 250  $\mu\text{l}$  containing 1.5  $\mu\text{mol}$  of  $\text{MgCl}_2$ , 17  $\mu\text{mol}$  of Tris-HCl pH 7.5, 18  $\mu\text{mol}$  of KCl, 1.5  $\mu\text{mol}$  of DTT, 12 nmol each of dGTP, dATP, TTP and [ $^3\text{H}$ ]dCTP (4.7 Ci/mmol). The reaction was stopped by the addition of EDTA and the DNA extracted, chromatographed and precipitated as before. The double-stranded cDNA (ca. 1  $\mu\text{g}$ ) was dissolved in 200  $\mu\text{l}$  of 200 mM NaCl, 1 mM  $\text{ZnCl}_2$ , 50 mM Na acetate pH 4.5 and treated with 60 U of  $S_1$  nuclease (Sigma) for 30 min at  $37^{\circ}\text{C}$ . The reaction was stopped with 60  $\mu\text{l}$  of 1 M Tris-HCl pH 8.3 and the DNA again extracted, chromatographed and precipitated. To remove low molecular weight material the double-stranded cDNA (ca. 700 ng) was loaded onto a 5 ml 5–20% (w/v) sucrose gradient (containing 50 mM Tris-HCl pH 7.5, 0.25 M NaCl) and centrifuged at 140,000 g av. for 90 h at  $4^{\circ}\text{C}$ . The leading fractions were pooled and the DNA (ca. 320 ng) precipitated with ethanol.

#### Construction and cloning of recombinant plasmid DNA

Double-stranded cDNA (1  $\mu\text{g}$ ) was incubated in the presence of 25  $\mu\text{Ci}$  [ $^3\text{H}$ ]dCTP (1.5 Ci/mmol), 100 mM cacodylic acid, 21 mM Tris base, 78 mM KOH, 0.1 mM DTT, 1 mM  $\text{CoCl}_2$ , and 115 U terminal transferase (TdT, Bethesda Research Laboratories) in a final volume of 100  $\mu\text{l}$  and a pH of 7.5. After 30 min at  $37^{\circ}\text{C}$  the reaction was stopped by addition of 5  $\mu\text{l}$  4 M  $\text{NH}_4\text{acetate}$  and extraction with phenol/chloroform. *Pst*I-digested pBR322 (200 ng) was incubated at  $37^{\circ}\text{C}$  under similar conditions, except that [ $^3\text{H}$ ]dGTP (4 Ci/mmol) was used. When *Hind* III-digested pPH207 (14) was used for cloning the plasmid was tailed with poly (dT) by incubation in the presence of 125  $\mu\text{M}$  TTP, 100 mM HEPES pH 7.0 1 mM DTT, 4 mM  $\text{MgCl}_2$  and 15 U TdT. After 5 min at  $37^{\circ}\text{C}$  the reaction was stopped by addition of 1  $\mu\text{l}$  200 mM EDTA and 60  $\mu\text{l}$  100 mM Tris-HCl pH 7.5. For tailing with poly (dA), double-stranded cDNA (20 ng) was incubated under similar conditions except that 50  $\mu\text{Ci}$   $^3\text{H}$ -dATP (7 Ci/mmol) was used in place of TTP.

Poly (dG)-tailed pBR322 DNA (125 ng) was mixed with poly (dC)-tailed double-stranded cDNA (60 ng) in 30  $\mu\text{l}$  of TEN (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 100 mM NaCl) and heated to  $60^{\circ}\text{C}$  for 60 min and then allowed to cool

slowly to room temperature overnight. After addition of 200  $\mu\text{g}$  tRNA as carrier the annealed DNA was chromatographed on G25 Sephadex equilibrated with TEN and 0.1% SDS. The excluded fractions were pooled, the DNA was precipitated with ethanol and dissolved in 100  $\mu\text{l}$  of 10 mM Tris-HCl pH 8.0, 1 mM EDTA. Poly (dT)-tailed pPH207 DNA (20 ng) and poly (dA)-tailed double-stranded cDNA (2 ng) were annealed in 50  $\mu\text{l}$  of TEN by heating to 65°C for 10 min, then gradually cooling to 45°C for 2 h and finally cooling to room temperature for 60 min.

Competent cells of *E. coli* HB101 were prepared by treatment with  $\text{Ca}^{2+}$  (15). The annealed plasmid and cDNA, in 100  $\mu\text{l}$  of 10 mM Tris-HCl pH 8.0., 1 mM EDTA, was mixed with 200  $\mu\text{l}$  of cell suspension and kept on ice for 20 min. After a 2 min heat shock at 42°C the cells were added to 0.5 ml L broth (1% bacto tryptone, 0.5% bacto yeast extract 0.5% NaCl) and incubated at 37°C with shaking for 30 min. The cells were plated out on L agar containing tetracycline at 5  $\mu\text{g}/\text{ml}$  (when pBR322 was the vector) or chloramphenicol at 50  $\mu\text{g}/\text{ml}$  (when pPH207 was the vector).

### Preparation of [ $^{32}\text{P}$ ]-labelled probes for colony hybridizations

Poly A<sup>+</sup> RNA (1-2  $\mu\text{g}$ ) was partially hydrolysed by heating to 100°C for 30 sec in 50  $\mu\text{l}$  of sterile distilled water. The RNA was then  $^{32}\text{P}$ -labelled by incubation in 20  $\mu\text{l}$  of a mixture containing 25  $\mu\text{Ci}$  [ $\gamma$ - $^{32}\text{P}$ ] ATP (2000 Ci/mmol), 50 mM Tris-HCl pH 9.0, 10 mM  $\text{MgCl}_2$ , 5 mM DTT, 5% glycerol and 4 U T4 polynucleotide kinase (P.L. Biochemicals). After 30 min at 37°C 1  $\mu\text{l}$  of 2 mM ATP was added and the incubation continued for a further 30 min. The reaction was stopped by the addition of 20  $\mu\text{l}$  TEN, 0.1% SDS.

Cloned cDNA sequences were excised by digestion of 10-20  $\mu\text{g}$  plasmid DNA with Pst 1. The DNA was then electrophoresed on a 6% acrylamide gel (16) and stained with ethidium bromide. The cDNA inserts were eluted from gel slices by diffusion (17) and  $^{32}\text{P}$ -labelled by nick translation (18). The radioactively labelled RNA or cDNA probes were chromatographed on a 1 ml Sephadex G50 column before use.

### Colony hybridizations

The procedure used was based on that of Grunstein and Hogness (19). Bacterial colonies were grown overnight on nitrocellulose filters (Schleicher and Schuell BA85 0.45  $\mu\text{M}$ ) and the filters treated in turn with 0.5 M NaOH, 1 M Tris HCl pH 7.5 (twice) and 1.5M NaCl, 0.5 M Tris-HCl pH 7.5, for 5 min each and dried under reduced pressure at 80°C for 2 h. Before hybridization the filters were incubated for 60 min at 65°C in pre-

hybridization buffer containing 0.2% Ficoll, 0.2% polyvinylpyrillidone, 100 mg/ml bovine serum albumin, 0.1% SDS, 2 x SSC (0.3 M NaCl, 30 mM Na citrate), 100 µg/ml denatured Salmon sperm DNA and 100 µg/ml poly A. Hybridization was for 24 h at 65°C in 2 ml of pre-hybridization buffer to which 200 µl of <sup>32</sup>P-labelled probe in TEN 0.1% SDS had been added. Filters were washed overnight at 65°C in 3 x 1 l of 2 x SSC, dried and autoradiographed at -70°C using Kodak X-omat H X-ray film.

#### Preparation of plasmid DNA

Hybrid plasmids derived from pBR322 were amplified by chloramphenicol treatment and plasmid DNA was purified from cells lysed with Triton X-100 (20).

#### Restriction digests

Pst I and Hae III digests were carried out at 37°C for 2 h in 50 mM NaCl, 10 mM MgCl<sub>2</sub>, 10 mM Tris-HCl pH 7.4, 10 mM dithiothreitol and 20-250 µg/ml DNA. Other restriction enzymes, obtained commercially from Bethesda Research Laboratories or New England Biolabs, were used under the conditions recommended by the suppliers.

#### Electrophoresis of RNA and hybridization of cloned DNA to Northern transfers

Poly A<sup>+</sup> RNA was electrophoresed in horizontal 1% agarose slab gels containing 6 mM CH<sub>3</sub>HgOH (21). Molecular weight marker RNA's were run in parallel tracks and visualised by staining with 1 µg/ml ethidium bromide. Preparation of DBM paper, transfer of RNA from the gel to the paper, prehybridization treatment and hybridization to <sup>32</sup>P-labelled plasmid DNA was carried out according to Alwine *et al* (22,23). The plasmid DNA was labelled by nick-translation (18) and both prehybridization and hybridization were done in the presence of 50 µg/ml poly (A). After hybridization the paper was washed as follows: (i) 6 x 20 min at 42°C with 5 x SSC, 0.2% SDS, 50 mM Na phosphate pH 6.5, 50% (v/v) formamide. (ii) 0.1 x SSC, 0.2% SDS, 1 mM Na phosphate at 42°C, 50°C and 55°C for 60 min each. (iii) 0.1 x SSC, 0.2% SDS, 1 mM Na phosphate pH 6.5, 2 mM EDTA at 55°C for 60 min. (iv) 50 mM Tris-HCl pH 7.5, 0.1% SDS, 50% formamide at 40°C for 10-60 min. When the RNA-DBM paper was re-used for hybridization the already hybridized DNA was first removed by twice boiling in water for 5 min.

#### Hybrid-release translation

10 µg plasmid DNA was relaxed with S<sub>1</sub> nuclease (Boehringer) (24). The relaxed DNA, in 20 µl 5 mM Tris-HCl pH 8.6, was boiled for 1 min, rapidly cooled in liquid N<sub>2</sub>, made to 4 x SSC and immobilized on

nitrocellulose filters (Schleicher and Schuell BA 85/1). The filters were washed for 3 h with 4 x 500  $\mu$ l of 5 x SSC at room temperature. The filter was dried under reduced pressure at 80°C for 2 h. Prehybridization treatment was for 1 h at 52°C in 65% formamide (v/v)(twice recrystallised), 10 mM Pipes pH 6.4, 0.4 M NaCl and 85  $\mu$ g/ml poly A. Poly A<sup>+</sup> RNA (20-30  $\mu$ g) was incubated with the immobilized DNA with gentle shaking for 4 h at 52°C in 120  $\mu$ l of prehybridization buffer without poly A. The filters were then washed 10 times at 60°C with SSC, 0.5% SDS and 3 times with 1 ml 2 mM EDTA pH 7.9 (2 x 5 min at room temperature, 5 min at 60°C)(25) and the hybridized RNA was eluted by boiling for 70-80 s in 220  $\mu$ l sterile distilled water. The filters were removed and the released RNA was quickly cooled in liquid N<sub>2</sub>. The RNA was ethanol precipitated using 10  $\mu$ g of calf thymus tRNA as carrier. Translation of selected mRNA's was carried out in the wheat germ cell-free system (8) using L-[4,5-<sup>3</sup>H]leucine, (197 Ci/mmol), L-[2,3,4,5-<sup>3</sup>H]proline (100 Ci/mmol) or L-[4,5-<sup>3</sup>H]lysine (80 Ci/mmol). Before electrophoresis the in vitro translation products were alkylated (8) and then incubated at 60°C for 5 min in 4 M urea, 2% SDS, 2% 2-mercaptoethanol and 25 mM Tris-HCl pH 6.8. The proteins were analysed in an SDS-polyacrylamide gel system based on that of Laemmli (26) with the addition of 4 M urea and 4 mM DTT to both stacking and separating gels. After electrophoresis the gel was treated with sodium salicylate, dried and fluorographed (27).

#### Nucleotide sequencing

Sequencing was carried out by the dideoxy method (28,29) using restriction fragments from hybrid plasmids which had been sub-cloned in E. coli JM101 using the phage M13mp7 as vector (30).

#### Physical containment

Recombinant DNA experiments were carried out under Category II containment conditions as defined by GMAG.

### RESULTS AND DISCUSSION

#### Cloning of poly A<sup>+</sup> RNA sequences from barley endosperm

Double-stranded cDNA was synthesized enzymatically using poly A<sup>+</sup> RNA from the membrane-bound polysomes of barley endosperm as the template. This RNA fraction has been shown to be highly enriched for mRNA which specifies storage proteins (8,31). In the first transformation experiment the double-stranded cDNA was inserted into the Hind III site of pPH207 by A-T tailing. Successful transformants were selected by their resistance to

chloramphenicol (*cmp<sup>r</sup>*) and those transformants with inserted cDNA were selected by their sensitivity to tetracycline (*tet<sup>S</sup>*). Only 5 *cmp<sup>r</sup>*, *tet<sup>S</sup>* colonies were obtained and these have been designated pHvE-c901 to pHvE-c905 (abbreviated to pc901, pc902 etc). In subsequent cloning experiments the cDNA was inserted into the *Pst* I site of pBR322 by G-C tailing and 273 *tet<sup>r</sup>* transformants (pc1-73 and pc101-300) were obtained.

Endosperm-specific cDNA clones were detected by comparison of their hybridization to poly A<sup>+</sup> RNA probes from endosperm and etiolated shoots. Translation of these two RNA preparations in a wheat germ system and subsequent gel electrophoresis showed, as expected, that the major translation products differed markedly and that there was no detectable synthesis of B or C hordein polypeptides by poly A<sup>+</sup> RNA from etiolated shoots. The results of hybridizing the two <sup>32</sup>P-labelled RNA probes to the colonies confirm that at least the abundant RNA sequences in the preparations share very little homology (Fig. 1). The RNA probe from etiolated shoots hybridized to only 7 of the 273 *tet<sup>r</sup>* colonies, while the endosperm probe hybridized to 133 colonies in addition to these 7. The 7 clones to which both probes hybridized have been found to carry sequences related to barley rRNA (see below). Cloning of rRNA sequences in this way

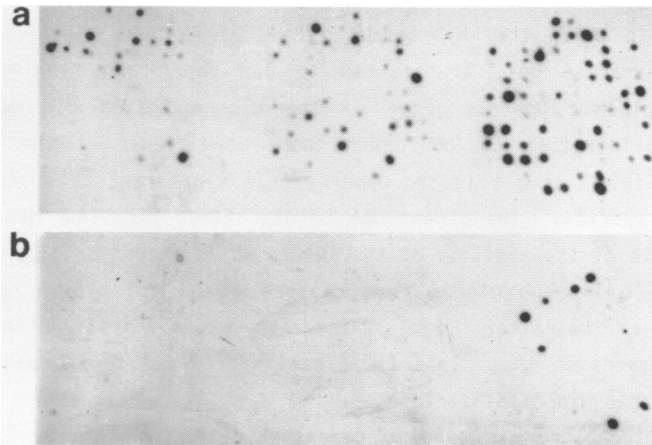


Fig. 1. Autoradiograph showing hybridization of colonies of <sup>32</sup>P-labelled poly A<sup>+</sup> RNA from (a) barley endosperm (b) etiolated barley shoots. Poly A<sup>+</sup> RNA from the membrane-bound polysomes of each tissue was <sup>32</sup>P-labelled and incubated under hybridizing conditions with a set of three nitrocellulose filters carrying a total of 237 *tet<sup>r</sup>* transformants. Conditions of hybridization were as described in Methods.

has been observed previously (32,33).

Before further analysis we attempted to reduce the number of clones to be screened by dividing them into groups of related sequences. Plasmid DNA was prepared from a number of randomly selected clones and tested for excisability of the cDNA insert by Pst I digestion. The Pst I-excised cDNA inserts were then purified by preparative acrylamide gel electrophoresis from 10 of these plasmids. The sizes of these inserts ranged from 230-900 base pairs. The inserts were <sup>32</sup>P-labelled by nick translation and hybridized separately to the tetr<sup>r</sup> clones by colony hybridization. The probes from plasmids pc16, pc59, pc105, pc114, pc154, pc169, pc179, pc194, pc218 and pc287 hybridized to a total of 110 colonies (14 of which had not hybridized detectably to poly A<sup>+</sup> RNA). The inserts from pc169 and pc287, although of different sizes, hybridized to the same group of 7 clones. Many clones hybridized to two or more of the probes so that the pattern of cross-hybridization was complex, as illustrated in Fig. 2 using a Venn diagram. Also indicated in Fig. 2 is the separate group of 7 clones which hybridized to the EcoR1-excised insert of pHV29, a barley rDNA sequence (34). Fig. 2 has been used as a guide for more effective analysis of the cDNA clones as described below.

#### Hybridization of cDNA clones to size-fractionated RNA

Twenty eight cDNA clones were characterized with respect to the size class of RNA to which they hybridized. Poly A<sup>+</sup> RNA was electrophoresed in a CH<sub>3</sub>HgOH-agarose gel, transferred to DBM paper and hybridized to <sup>32</sup>P-labelled plasmid DNA's. Five of the clones tested did not show detectable hybridization to RNA under these conditions but the results presented in Fig. 3 show that the other clones hybridized to a number of different RNA species. The clones can be classified into four groups (I-IV) on the basis of the mobility of the RNA to which they hybridize (Fig. 3 and Table I). Although clones from classes I and III hybridized to a similar region of the filter (Fig. 3) they were clearly distinguishable on the autoradiographs in that Class III clones hybridized to a broader size range of RNAs than did Class I clones.

The same technique was used to demonstrate that one of the 7 clones which hybridized to both endosperm and shoot RNA (Fig. 1) carries a sequence related to barley 26S rRNA. When tested against equal amounts of endosperm poly A<sup>-</sup> and poly A<sup>+</sup> RNA, pc222 hybridized much more strongly to the poly A<sup>-</sup> than to the poly A<sup>+</sup> fraction (Fig. 3D lanes a and b). Hybridization to poly A<sup>-</sup> RNA was seen as a major band corresponding in



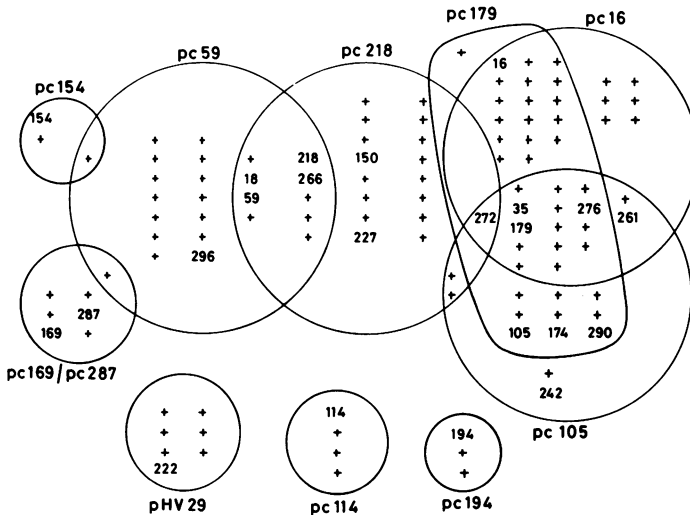


Fig. 2. Venn diagram illustrating the complex pattern of cross-hybridization between cDNA clones. The cDNA inserts from 10 hybrid plasmids were each excised with *Pst* I, purified by gel electrophoresis,  $^{32}\text{P}$ -labelled by nick translation and incubated under hybridizing conditions with all 237 *tet<sup>r</sup>* colonies on nitrocellulose filters. In addition, the DNA insert was excised from a barley ribosomal DNA clone pHV29 (34) by digestion with *Eco* R1, and similarly purified, labelled and tested for hybridization to the colonies. The cDNA clones to which each probe hybridized were detected by autoradiography and are indicated within the respective circles of the diagram. Many clones hybridized to two or more of the probes, as shown by the overlapping circles. Only those clones which were examined further in the present study are indicated by their numbers; the remainder are designated by + signs. A total of 107 of the 237 *tet<sup>r</sup>* colonies hybridized to one or more of the probes.

mobility to barley 26S rRNA ( $M_r 1.3 \times 10^6$ ) and as a series of minor lower molecular weight bands. After longer exposure of the autoradiographs similar bands were seen with poly A<sup>+</sup> RNA (Fig. 3D lane d), indicating some contamination of this fraction with rRNA. There was no detectable hybridization of pc222 to *K. aerogenes* rRNA (fig. 3D lane c). Although it is thought that the other 6 clones which hybridize to shoot RNA are derived from rRNA (see Fig. 2) we have not determined whether they are related to the 26S rRNA.

#### Translation of mRNAs selected by hybridization

To study the relationship between cDNA sequences and polypeptides we chose 20 cDNA clones for further characterization by hybrid-release

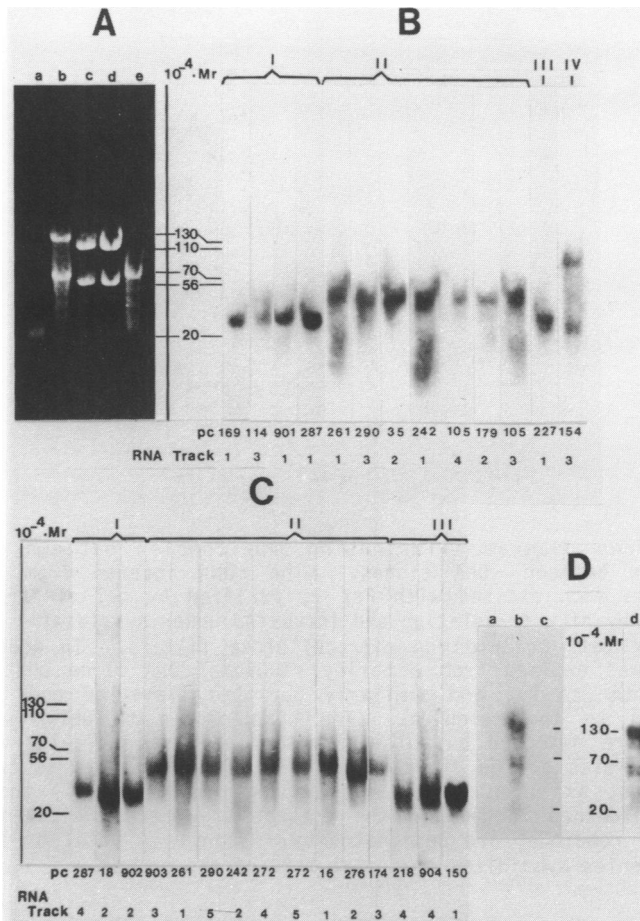


Fig. 3. 'Northern' hybridizations of cDNA clones to RNA fractionated on CH<sub>2</sub>HgOH-agarose gels. Plasmid DNA (ca. 0.1µg) was <sup>32</sup>P-labelled by nick translation to a specific activity of about 10<sup>7</sup> cpm µg<sup>-1</sup> and hybridized to the covalently bound RNA. Following hybridization the filters were washed and autoradiographed.

A: Gel stained with ethidium bromide to show molecular weight markers. Lane a) rabbit globin mRNA; b) polysomal poly A<sup>-</sup> RNA from barley endosperm; c) and d) *Klebsiella aerogenes* rRNA; e) polysomal poly A<sup>+</sup> RNA from barley endosperm. B and C: Autoradiographs obtained after hybridization of various <sup>32</sup>P-labelled plasmid DNAs to polysomal poly A<sup>+</sup> RNA from barley endosperm; in B the RNA had been fractionated on the same gel as in A, while C is from another but similar gel. D: Autoradiographs obtained after hybridization of <sup>32</sup>P-labelled pc222 to endosperm poly A<sup>+</sup> RNA (lanes a and d), endosperm poly A<sup>-</sup> RNA (lane b), *K. aerogenes* rRNA (lane c). Lanes a, b and c were autoradiographed for 1 day, lane d for 5 days.

Table 1. Classification of clones according to the size(s) of RNA to which they hybridize

Class	cDNA clones	<sup>a</sup> Molecular weight of hybridizing RNA ( $\times 10^{-3}$ )	Estimated size of RNA (nucleotides)
I	pc18, pc114, pc169, pc287, pc901, pc902	280 ( $\pm 10$ )	875
II	pc16, pc35, pc105, pc174, pc179, pc242, pc261, pc272, pc276, pc290, pc903	420 ( $\pm 20$ )	1300
III	pc150, pc218, pc227, pc904	250-330	780-1050
IV	pc154	240 ( $\pm 5$ ) and 800 ( $\pm 20$ )	750 and 2500
V	pc222	<sup>b</sup> 1300 ( $\pm 50$ ) (also minor bands: 230, 300, 520, 600 and 1000)	4000
VI	pc59, pc145, pc266, pc296, pc905	No detectable hybridization	

<sup>a</sup> Molecular weight estimates are based on the results of at least four hybridizations of each clone to different RNA tracks, usually including tracks from more than one gel.

<sup>b</sup> 18S and 26S barley rRNA have molecular weights of  $700 \times 10^3$  and  $1300 \times 10^3$  daltons respectively.

translation. Plasmid DNA from each clone was bound to nitrocellulose filters and then hybridized to poly A<sup>+</sup> RNA. RNA which hybridized was eluted, translated in a wheat germ system and the *in vitro* translation products were analysed by electrophoresis on SDS-urea polyacrylamide gels. To test the specificity of the hybridization, endosperm poly A<sup>+</sup> RNA was hybridized to immobilized DNA of pBR322 and pc222 (a rRNA-related clone, see Figs. 2 and 3). As expected neither of these plasmid DNAs selected RNA capable of directing *in vitro* translation (Fig.4, lanes i and j). When one of the cDNA clones, pc179, was hybridized with endosperm poly A<sup>+</sup> RNA it selected RNA which directed synthesis of a number of polypeptides (Fig. 4 lane m). However when RNA from etiolated shoots was used in the

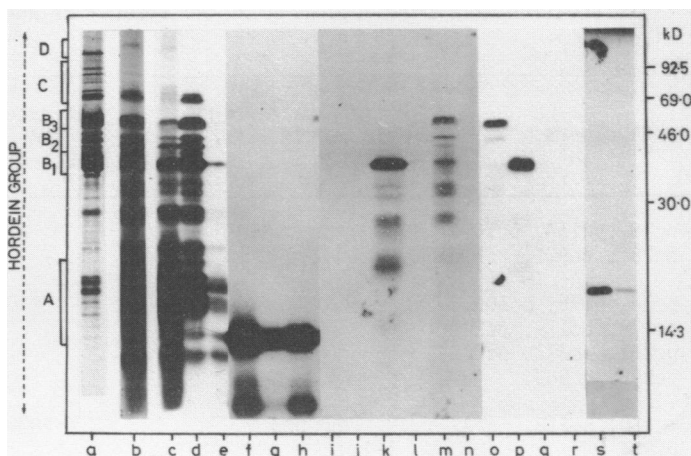


Fig. 4 Characteristics of translation products specified by different mRNAs Alcohol-soluble barley storage proteins (lane a) and *in vitro* <sup>3</sup>H-labelled translation products (lanes b to t) were separated on a 13% SDS-urea-polyacrylamide slab gel. The mRNAs translated in the wheat germ were: no RNA (n); endosperm poly A<sup>+</sup> RNA (b to e); rabbit reticulocyte poly A<sup>+</sup> RNA (f to h); RNA selected from endosperm poly A<sup>+</sup> RNA by hybridization to plasmid DNAs: pc222 (i), pBR322 (j), pc290 (k), pc179 (m), pc179 (o), pc290 (p), pc179 (q), pc290 (r), pc114 (s), pc114 (t). RNA selected from shoot poly A<sup>+</sup> RNA by hybridization to pc179 (l). Products in lanes b, i, j, k, l, m, n were labelled with [<sup>3</sup>H]leucine and [<sup>3</sup>H]proline; in lanes c and f with [<sup>3</sup>H]leucine, in lanes d, g, o, p, s with [<sup>3</sup>H]proline and in lanes e, h, q, r, t with [<sup>3</sup>H]lysine.

hybridization, pc179 failed to select any translatable RNA (Fig. 4 lane l), confirming the endosperm specificity of this cDNA clone (Fig. 1) and further demonstrating the specificity of the procedure.

The variety of results obtained when hybrid-release translations were carried out on 20 hybrid plasmids is shown in Fig. 5. Although a few cDNA clones (e.g. pc114 lane i) gave a single polypeptide, the majority selected mRNAs which directed the synthesis of several polypeptides, sometimes of widely differing molecular weights, in spite of stringent washing (see Methods). One group of clones which falls into the latter category is of particular interest. The 11 clones which hybridized to RNA of 1300 nucleotides on Northern transfers (Class II, Table 1) each gave hybrid-release translation products which correspond in electrophoretic mobility to one or more of the nascent B hordein polypeptides (Fig. 5). For example, pc16 (lane q), pc179 (lane d), pc242 (lane w) and pc903 (lane t)

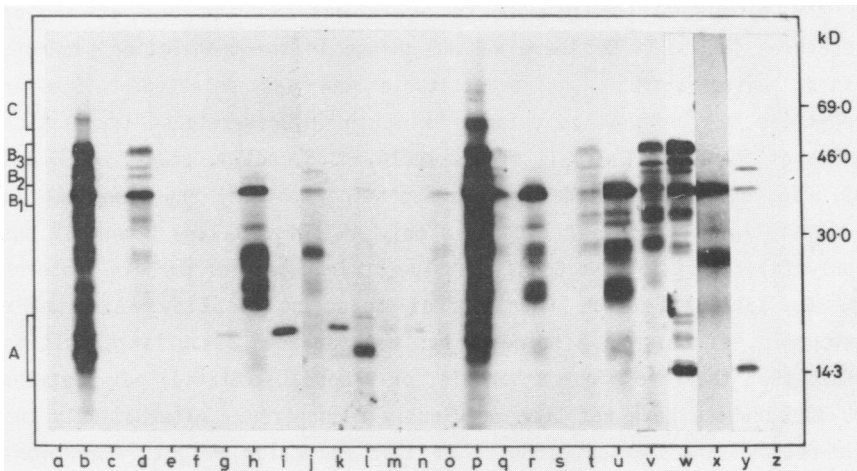


Fig. 5 Translation products of plasmid-selected mRNA. Lanes a and z: no RNA added to wheat germ; lanes b and p: translation products of endosperm poly A<sup>+</sup> RNA. All other lanes show the translation products of mRNA selected by plasmid DNAs: pBR322 (c), pc179 (d), pc59 (e), pc266 (f), pc169 (g), pc276 (h), pc114 (i), pc272 (j), pc287 (k), pc154 (l), pc218 (m), pc150 (n), pc105 (o), pc16 (q), pc35 (r), pBR322 (s), pc903 (t), pc290 (u), pc179 (v), pc242 (w), pc174 (x), pc261 (y). The translation products were analysed as for Fig. 4.

gave rise to a group of polypeptides of 33-47 kd which clearly resembles the group of major B hordeins (B1, B2 and B3) seen in the translation products of polysomal poly A<sup>+</sup> RNA (Fig. 4 lanes b, c, d and ref.8). Since each hordein band seen on SDS-PAGE can be resolved into a number of individual polypeptides by isoelectric focusing (6) it does not follow that pc16, pc179, pc242 and pc903 necessarily hybridize to the same highly conserved mRNA sequence. Preliminary evidence from isoelectric focusing of the hybrid-release translation products indeed suggests that the four clones are not identical.

The 7 other clones which fall into Class II on Northern hybridizations (Table 1) also give at least one hybrid-release translation product which co-migrates with a B hordein polypeptide. In the case of pc35 (lane r) and pc290 (lane u) the major translation product corresponds to a B1 hordein polypeptide (Fig. 5). However, in the case of pc105 (lane o), pc174 (lane x), pc261 (lane y), pc272 (lane j) and pc276 (lane h) the major translation products include, in addition to polypeptides with the mobility of B hordeins, a number of polypeptides of lower molecular weight (20-30 kd).

When the hybridization of RNA to pc276 was carried out at higher temperatures (up to 60°C) there was no change in the translation products obtained, while at 65°C no translatable RNA was selected. Similar polypeptides can be seen as components of the hybrid-release translation products of the other Class II clones (pc16, pc35, pc179, pc242, pc290 and pc903, Fig. 5). Polypeptides of corresponding mobility are also seen in the translation products of endosperm poly A<sup>+</sup> (Fig. 5 lane b and p) but are only faintly visible on Coomassie-blue stained gels of hordein proteins (Fig. 4, lane a). The identity of these polypeptides and their relationship, if any, to B hordeins has not yet been established. It is possible that they are minor B hordein, or B hordein-related, polypeptides of 20-30 kd which have not been previously recognized. Alternatively the hybrid-release translation products in this molecular weight range might result from premature termination. This interpretation cannot be completely excluded but seems unlikely since there is a considerable variation in the intensities of these bands between the different hybrid-selected mRNAs translated in the same system.

To assist in the identification of the hybrid-release translation products of Class II clones we have compared the ratio of incorporation of [<sup>3</sup>H]proline to [<sup>3</sup>H]lysine. It is known that the B hordeins are rich in proline (20.6 mol %) and poor in lysine (0.5 mol %) (2) and a high proline : lysine ratio is therefore a characteristic of these proteins. Because of the low levels of [<sup>3</sup>H]lysine incorporation into the hybrid-release translation products of the three plasmids tested, accurate estimates of the proline : lysine ratio were difficult to make from measurements of trichloroacetic acid-insoluble radioactivity. Approximate values obtained with two Class II clones were 6 for pc179 and 12 for pc290, as compared with ratios of 5.5 for the translation products of endosperm poly A<sup>+</sup> RNA and 0.5 for those of rabbit reticulocyte mRNA. Examination of the autoradiographs (Fig. 4) confirms that globin was readily labelled by both [<sup>3</sup>H]proline (lane g) and [<sup>3</sup>H]lysine (lane h). However, the translation products of RNA selected by pc179 (Fig. 4 lanes o and q) and pc290 (lanes p and r), like the majority of those of endosperm poly A<sup>+</sup> RNA (lanes d and e), were not detectable after [<sup>3</sup>H]lysine incorporation, although many were readily labelled by [<sup>3</sup>H]proline. Thus it is clear that the hybrid-release translation products of these two clones share with the translation products of endosperm mRNA, and with the B hordeins, the property of a high proline : lysine ratio. Our tentative identification of pc179, pc290 and

the other class II clones (Table 1) as B hordein cDNA clones is therefore supported by these observations.

In the case of pc114, a Class I clone (Table 1), the single low molecular weight polypeptide synthesized (ca. 20 kd) is detectable after [<sup>3</sup>H]lysine incorporation (Fig. 4 lane t) but less readily than after [<sup>3</sup>H]proline incorporation (lane s). The identity of this polypeptide with an intermediate proline : lysine ratio is not yet known.

An indication of the complexity of the family of mRNAs which code for B hordeins is given by the ability of several cross-hybridizing cDNA clones to select mRNA for different groups of B hordein polypeptides. For example, although pc179 hybridizes to both pc35 and pc272 (Fig. 2), the three plasmid DNAs give very different results on hybrid-release translation (Fig. 5 lanes d, r and j respectively). Similar results were obtained in an analysis of zein cDNA clones (24). Thus the B hordeins, like the maize storage proteins, seem to be composed of several families of polypeptides which are nevertheless related by sequence homologies in their mRNAs. However, the failure of any of the 11 B hordein cDNA clones to select C hordein mRNA (Fig. 5) suggests that there is little sequence homology between the B and C hordein mRNAs.

Only one cDNA clone which does not belong to Class II, pc154, selected mRNA specifying a polypeptide corresponding in mobility to a B hordein (Fig. 5 lane l). The major translation product related to pc154, however, was low molecular weight (ca. 15 kd) and in addition there were several translation products of intermediate molecular weight. Since pc154 does not hybridize to any of the putative B hordein clones (Fig. 2) it is unlikely that it is also derived from B hordein mRNA.

The remaining 5 cDNA clones tested (pc59, pc150, pc169, pc218, pc287 Fig. 5 lanes e, n, g, m, k) all gave hybrid-release translation products of low molecular weight (15-20 kd). As expected, pc169 and pc287, the cDNA inserts of which each hybridize to the same group of cDNA clones (Fig. 2) give comparable results on hybrid release translation (Fig. 5). The identity of these various low molecular weight polypeptides has not yet been established.

#### Nucleotide sequence determination

To confirm its relationship to the B hordeins, one of the cDNA clones pc16, was partially sequenced by the dideoxy chain termination method (28,29). The cDNA insert in pc16 is approximately 350 base pairs long and contains a single Alu 1 site (Fig. 6). One of the two cDNA-containing

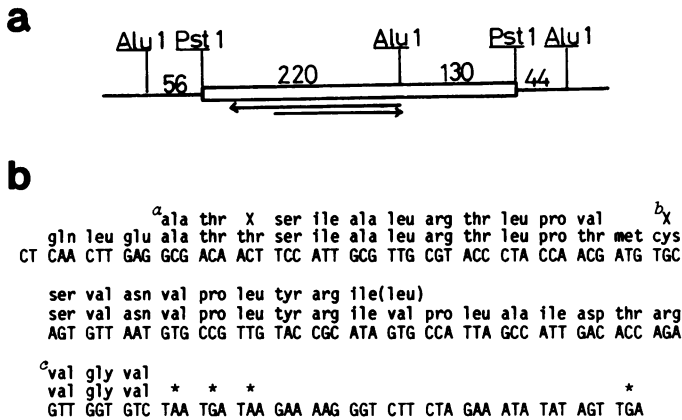


Fig. 6 Physical map and partial nucleotide sequence of pc16. a) A physical map of the region of pc16 which includes the cDNA insert (boxed segment). The numbers refer to fragment lengths in nucleotide base pairs. b) A partial nucleotide sequence of the cDNA insert of pc16 derived from the 280bp Alu 1 fragment sub-cloned into the Hinc II site of M13mp7 (30). The sequence was obtained by the dideoxy chain termination method. The predicted amino acid sequence is also shown and has been aligned with amino acid sequences from B1 hordein polypeptides. Asterisks indicate the positions of termination codons. <sup>a</sup>N-terminal sequence of SV8 peptide IV of B1 hordein (36). <sup>b</sup>N-terminal sequence of CNBr peptide V of B1 hordein (36). <sup>c</sup>C-terminal sequence of B1 hordein (37,38).

Alu 1 fragments from pc16 was sub-cloned in M13mp7 (30) by blunt-end ligation into the Hinc II site. M13 clones were obtained with the Alu 1 fragment in both possible orientations. Part of the nucleotide sequence obtained from an M13 clone with the cDNA sequence (rather than the pBR322 sequence) adjacent to the region of primer annealing is presented in Fig. 6. This sequence has been confirmed by repeated sequencing reactions and electrophoretic analysis and by use of cytosine  $\beta$ -D-arabinose nucleoside triphosphate as an additional chain terminator (35) in some experiments. However, the sequence was not readily confirmed from a clone with the cDNA insert in the other orientation because of difficulty in obtaining a readable sequence beyond the oligo (dC) 'tail' which links the cDNA to the pBR322 fragment. Although the DNA polymerase reaction continued through the homopolymeric region and well beyond, the resulting DNA strands appeared as blurred bands on the autoradiograph. The explanation for this effect is not known. We have, however, been able partially to surmount this problem by annealing oligo (dG)<sub>12-18</sub> to the oligo (dC) sequence



and using this as a primer in the normal sequencing reaction (B. G. Forde and L. Blanco, manuscript in preparation). Because of the length heterogeneity of the primer and the multiple positions on the oligo (dC) tail to which it can anneal it was necessary to remove this heterogeneous 5' end after the reaction. As it was known that a Taq I site is located approximately 40 base pairs from the G-C tail, reaction products were digested with this enzyme before electrophoresis. It was then possible to read the DNA sequence from a point approximately 100 nucleotides from the G-C tail. Using this technique we were able to confirm the first 115 nucleotides of the sequence in Fig. 6, including the entire coding sequence.

The amino acid sequence predicted from one of the six possible reading frames for the nucleotide sequence in Fig. 6 is also shown. Although this reading frame contains 4 termination codons none of these occur in the first 111 nucleotides. The open reading frame is followed by three successive termination codons, giving a putative C-terminal sequence of three amino acids which are the same as those determined for the C-terminus of B1 hordein polypeptides (37,38). Furthermore there is very strong homology between two regions of the predicted amino acid sequence and the amino acid sequences obtained from two B1 hordein peptides (36). A total of 23 of the 37 amino acid residues in the predicted sequence can be matched with amino acid sequences determined from B1 polypeptides. We therefore conclude that pc16 is probably derived from mRNA which specifies a B1 hordein polypeptide.

#### CONCLUSIONS

A collection of about 270 endosperm-specific cDNA clones has been constructed and analysed by several techniques. Cross-hybridization analysis indicated that a number of families of related sequences had been cloned (Fig. 2). In particular, the 54 clones which hybridized to one or more of the probes pc16, pc105 and pc179, form a closely-knit group at the right-hand side of the Venn diagram (Fig. 2). The 11 clones from this group which were examined in more detail also constituted a distinct class (Table 1) that hybridized to RNA of 1300 nucleotides. The same 11 clones also selected mRNA which specified at least one B hordein-like polypeptide on hybrid-release translation (Fig. 5). For two of the clones, pc179 and pc290, the B hordein-like nature of the hybrid-release translation products was confirmed by their characteristically high proline:lysine ratio (Fig.

4). A nucleotide sequence derived from one of the putative B hordein clones, pc16, predicted amino acid sequences which correspond with those of the known C terminus and two peptides of B1 hordein (Fig. 6). On colony hybridization, pc16 itself hybridized to 39 other cDNA clones (Fig. 2). We therefore conclude that most, if not all of the clones on the right-hand side of the Venn diagram carry sequences related to the B hordeins.

A number of the results obtained point to the complexity of the products of the Hor 2 locus. Firstly, the complex nature of the cross-hybridization between pc16, pc105, pc179 and the cDNA library (Fig. 2) can only be fully explained if we assume that the 54 clones in this cross-hybridization group are derived from more than one mRNA species. Secondly, several cDNA clones which cross-hybridize are found to select mRNA for different sets of B hordein polypeptides (Fig. 5). This not only confirms that the B hordeins are specified by more than one mRNA (i.e. are not derived by post-translational modification of one gene product) but also demonstrates that there are sequence homologies between mRNAs coding for different polypeptides. The results therefore support the hypothesis (6) that the Hor 2 locus is complex and has evolved by the duplication and divergence of a single ancestral gene.

#### ACKNOWLEDGEMENTS

We are indebted to J. O. Bishop and M. Richardson for laboratory facilities and help with cDNA synthesis and cloning into pPH207. We are grateful to W. L. Gerlach and J. R. Bedbrook for use of pHV29; to J. Beard for the gift of AMV reverse transcriptase, to C. M. Bowman for Pst I and to J. D. Marriage for wheat germ. We also thank L. Blanco for help with sequencing, J. Ray and J. Wallace for technical assistance and P. R. Shewry for valuable discussion. D. Bartels received a fellowship by the Deutsche Forschungsgemeinschaft.

#### REFERENCES

- 1 Mifflin, B.J. and Shewry, P.R. (1979) In: Seed Protein Improvement in Cereals and Grain Legumes, Vol. 1. pp.137-158, IAEA, Vienna
- 2 Shewry, P.R., Field, J.M., Kirkman, M.A., Faulks, A.J. and Mifflin, B.J. (1980) *J. exp. Bot.* 31, 393-407
- 3 Shewry, P.R., Faulks, A.J., Pickering, R.A., Jones, I.T., Finch, R.A., and Mifflin, B.J. (1980) *Heredity* 44, 383-389
- 4 Doll H. and Brown, A.D.H. (1979) *Canad. J. Genet. Cytol.* 21, 391-404
- 5 Jensen, J., Jorgensen, J.H., Jensen, H.P., Giese, H., and Doll, H. (1980) *Theor. Appl. Genet.* 58, 27-31
- 6 Faulks, A.J., Shewry, P.R., and Mifflin, B.J. (1981) *Biochem. Genet.* 19, 841-858

- 7 Brandt, A. (1979) *Carlsberg Res. Commun.* 44, 255-267
- 8 Matthews, J.A. and Mifflin, B.J. (1980) *Planta* 149, 262-268
- 9 Fox, J.E., Pratt, H.M., Shewry, P.R., and Mifflin, B.J. (1977) In *Nucleic Acids and Protein Synthesis in Plants*. pp.501-509, CNRS, Paris
- 10 Wienand, U. and Feix, G. (1978) *Eur. J. Biochem.* 92, 605-611
- 11 Bantle, J.A., Maxwell, I.H., and Hahn, W.E. (1976) *Anal. Biochem.* 72, 413-427
- 12 Efstratiadis, A., Maniatis, T., Kafatos, F.C., Jeffrey, A. and Vournakis, J.N. (1975) *Cell* 4, 367-378
- 13 Efstratiadis, A., Kafatos, F.C., Maxam, A.M., and Maniatis, T. (1976) *Cell* 7, 279-288
- 14 Bishop, J.O. and Davies, J.A. (1980) *Molec. Gen. Genet.* 179, 573-580
- 15 Cohen, S.N., Chang, C-Y, Hsu, L. (1972) *Proc. Natl. Acad. Sci. U.S.A.* 69, 2110-2114
- 16 Peacock, A. C. and Dingman, C. W. (1967) *Biochem.* 6, 1818-1827
- 17 Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 560-564
- 18 Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) *J. Mol. Biol.* 113, 237-251
- 19 Grunstein, M. and Hogness, D.S. (1975) *Proc. Natl. Acad. Sci. U.S.A.* 72, 3961-3965
- 20 Clewell, D.B. (1972) *J. Bacteriol.* 110, 667-676
- 21 Bailey, J.M. and Davidson, N. (1976) *Anal. Biochem.* 70, 75-85
- 22 Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5350-5354
- 23 Alwine, J.C., Kemp, D.J., Parker, B.A., Reiser, J., Renart, J., Stark, G.R. and Wahl, G.M. (1980) *Methods in Enzymology 68: Recombinant DNA*, R. Wu, ed, pp.220-242, Academic Press, New York
- 24 Park, W.D., Lewis, E.D. and Rubenstein, I. (1980) *Plant Physiol.* 65, 98-106.
- 25 Ricciardi, R.P., Miller, J.S. and Roberts, B.E. (1979) *Proc. Natl. Acad. Sci. U.S.A.*, 76, 4927-4931
- 26 Laemmli, V.K. (1970) *Nature* 227, 681-685
- 27 Chamberlain, J.P. (1979) *Anal. Biochem.* 98, 132-135
- 28 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463-5467
- 29 Sanger, F. and Coulson, A.R. (1978) *FEBS Lett.* 87, 107-110
- 30 Messing, J., Crea, R. and Seeburg, P.H. (1981) *Nucl. Acid Res.* 9, 309-321
- 31 Brandt, A. and Ingversen, J. (1978) *Carlsberg Res. Commun.* 43, 451-469
- 32 Ohno, T., Cozens, P.J., Cato, C.B. and Jost, J-P (1980) *Biochim. Biophys. Acta* 606, 34-46
- 33 Geiser, M., Doring, H-P., Wostmeyer, J., Behrens, U., Tillmann, E. and Starlinger, P. (1981) *Nucl. Acids Res.* 8, 6175-6188
- 34 Gerlach, W.L. and Bedbrook, J.R. (1979) *Nucl. Acids. Res.* 7, 1869-1885
- 35 Smith, A.J.H. (1980) In *Methods in Enzymology 65: Nucleic Acids Part 1*, L. Grossman and K. Moldave, eds. pp.560-580, Academic Press, New York
- 36 Schmitt, J.M. and Svendsen, I. (1980) *Carlsberg Res. Commun.* 45, 549-555
- 37 Schmitt, J.M. (1979) *Carlsberg Res. Commun.* 44, 431-438
- 38 Mifflin, B.J., Shewry, P.R., Faulks, A.J., Forde, B.G., Kreis, M., Matthews, J.A, Bahramian, M.B., Thompson, R.D., Bartels, D. and Flavell, R.B. (1982) In *Barley Genetics IV (Proc. 4th Int. Barley Genetics Symposium)*, Edinburgh University Press, Edinburgh (in press)