Volume 11 Number 10 1983 Nucleic Acids Research

The characterization of cDNA clones coding for wheat storage proteins

Dorothea Bartels and Richard D. Thompson

Cytogenetics Department, Plant Breeding Institute, Maris Lane, Trumpington, Cambridge CB2 2LQ, UK

Received 31 March 1983; Accepted 29 April 1983

ABSTRACT

Poly(A)$^+$ RNA isolated from the developing wheat endosperm var. Chinese Spring, has been used as template for the construction of a cDNA library. Within the library, clones have been identified by in vitro translation of hybrid-selected mRNA which encode α/β gliadin related sequences and γ-gliadin related sequences. The DNA sequence of one such clone has been determined and it shows homology with that of a clone encoding a barley storage protein, B-hordein. The sequence includes a tandem DNA repeat which is discussed in relation to the generation of diversity within the gliadins.

INTRODUCTION

Wheat flour is a major source of protein in the diet of humans and of livestock. The most abundant grain proteins are the storage proteins, termed gliadins and glutenins, which are deposited in the developing endosperm 2-3 weeks after fertilization (1). These proteins are important in determining the nutritive and baking properties of wheat flour. The gliadins are a complex, heterogeneous family of proteins, encoded by six, probably multigene loci on homoeologous group 1 and group 6 chromosomes (2). They are classified as prolamins on the basis of their solubility in alcohol. They resemble the prolamins of other Gramineae in their unusual amino acid composition, being rich in proline and glutamine but poor in charged amino acids, especially lysine and tryptophan.

Because of their low net charge the gliadins can be fractionated by electrophoresis at low pH. They have been subdivided into four groups, α, β, γ and ω, based on their mobility in aluminium lactate gels at pH 3.1 (2). The physical properties of gliadins make purification of individual protein components a difficult task. This factor and the unusual amino acid composition of the proteins has hindered the determination of protein sequence, which is mainly restricted to short amino terminal stretches (3).

This paper describes the production of a cDNA bank from the developing

endosperm of a homozygous, genetically characterized wheat variety (Chinese Spring). Clones from the bank have been shown to be complementary to mRNAs coding for α, β and γ gliadins. These clones are being used to further classify genes for gliadin polypeptides on the basis of sequence and genomic location.

MATERIALS AND METHODS

Plant material

Developing endosperms were isolated from ears of field-grown wheat, var. Chinese Spring, at 2-3 weeks post-anthesis, using a previously published method (4). Shoot RNA was isolated from 4-day old etiolated shoot tips of 'Chinese Spring' wheat. The material was stored at -80°C until required.

RNA extraction

40 g of endosperm tissue were ground to a fine powder under liquid nitrogen, incubated at 37°C for 15 min in 300 ml extraction buffer (0.1 M NaCl, 0.05 M Tris; HCl pH 9.0, 0.01 M EDTA, 2% SDS and 2 mg/10 ml Proteinase K) and extracted several times with phenol:chloroform (1:1) and finally with chloroform +0.1% isoamylalcohol only. The polyA$^+$ RNA fraction was isolated by chromatography on oligo dT cellulose (5) and stored under liquid nitrogen.

In vitro protein synthesis

Saturating amounts of poly(A)$^+$ RNA (30 µg/ml assay) were translated in a wheat germ, cell-free system for 90 min at 25°C in 20 µl assays using as a labelled amino acid L-[2,3,4,5-^{3}H] proline (100 Ci/mmol), L-[4,5-^{3}H]-leucine, (130 Ci/mmol), or L-[4,5-^{3}H] lysine (8 7 Ci/mmol).

Characterization of translation products

In vitro translation products were alkylated and incubated in SDS-mercaptoethanol-urea as described by Forde *et al.* (6) prior to analysis on 12.5% polyacrylamide gels. The loading buffer contained 2 M urea and 4 mM DTT (6). After electrophoresis the gel was fixed and stained in 0.02% Coomassie blue R (Sigma), 5% ethanol, and 6% TCA. The radioactive polypeptides were visualised by fluorography (7). In some cases the *in vitro* translation products were extracted with 55% aqueous isopropanol containing 0.1% DTT, alkylated with 5-vinylpyridine and then precipated with 2% (w/v) lithium chloride, using 10 µg wheat gliadins as carriers.

^{14}C-labelling of gliadins

1 mg of total gliadin extracted from var. Chinese Spring defatted flour (8) was labelled with 25 µCi ^{14}C-iodoacetic acid (Amersham) 40 mCi/mmol as described by (9). Ca. 70000 cpm were loaded on a gel for a 1-2 days exposure.

Construction of cDNA clones

First and second strand synthesis

The synthesis of double stranded cDNA followed the method described by (10). 10 µg endosperm polyA$^+$ RNA was incubated for 90 min at 42°C in 50 mM Tris-HCl pH 8.3, 140 mM KCl, 10 mM $MgCl_2$, 4 mM DTT, 500 µM of each deoxynucleotide triphosphate, 10 units/µg RNA avian myelobastosis virus reverse transcriptase (a generous gift of Dr. Y. Beard, Life Sciences, St. Petersburg, FL), 1 µg/µg RNA of oligo dT (Boehringer), and 1 unit/µl ribonuclease inhibitor RNasin (Biotec, Madison, USA). The second strand synthesis followed immediately by adding an equal volume of a buffer consisting of 200 mM Hepes pH 6.9, 500 µM of each deoxynucleoside triphosphate, 20 units of E. coli DNA polymerase I (Boehringer), and 15 µCi ^{32}P-dATP, in order to follow the reaction. This reaction was incubated for 3 hours at 22°C and then extracted with phenol and chloroform. The aqueous phase was desalted over a 1 ml Sephadex G-50-fine column in 20 mM NaCl, 10 mM Tris pH 8.5.

S_1 nuclease digestion

The excluded fractions from the Sephadex column were adjusted to S_1 buffer (300 mM NaCl, 30 mM Na-acetate, pH 4.5 and 3 mM $ZnCl_2$). Double stranded cDNA was digested with 1 unit/µl S_1 nuclease (Boehringer) for 30 min at 37°C. One unit catalyses the formation of 1 µg acid soluble deoxynucleotides after 30 min in incubation of denatured DNA at 37°C. At the end the reaction mixture was adjusted to pH 7.8 with 1 M Tris pH 8.0, extracted twice with phenol-chloroform (1:1) and chloroform and then ethanol precipitated.

Size fractionation of cDNAs

The double stranded cDNAs were size-fractionated on a 5-20% sucrose gradient in 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 10 mM NaCl) at 50,000 rpm. 4.5 ml gradients were run in a SW 65 rotor (Beckman) for 4 hrs at room temperature. The gradient was collected in 20 fractions. Small aliquots of each fraction were separated on a 5% polyacrylamide-urea gel in order to estimate the size of the cDNA. Fractions containing cDNA bigger than about 500 nucleotides were pooled and ethanol precipitated in the presence of 10 mM Mg-acetate.

Homopolymer Tailing of cDNA

5 µl of double-stranded cDNA was incubated in 100 µl of 0.1 M Cacodylic acid, 21 mM Tris base, 78 mM KOH, 0.15 mM dCTP and 25 ^{3}H-dCTP (spec. act. 12 $Ci.mmol^{-1}$) at 37°C. 1 µl 0.1 M $CoCl_2$ was added, and the reaction started by adding 15 units of terminal deoxynucleotidyl transferase (Bethesda Research Labs.). Samples were withdrawn after 10 and 20 minutes incubations to assay

for incorporation. The incubated samples were adjusted to 10 mM EDTA, 0.4 M Ammonium Acetate, 100 μg.ml^{-1} tRNA, extracted with phenol and chloroform and precipitated with 2.5 volumes ethanol overnight at -20°C. A similar reaction was performed to add deoxyguanosine residues to a DNA sample of Pst-digested pBR322.

Reannealing

75 ng pBr-dG was incubated with 5 μl cDNA-dC in 0.2 M NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA at 60°C for 1 hour. The waterbath was then switched off and allowed to cool to 4°C overnight. Recombinant plasmids were recovered by transformation into HB101 and subsequent screening for sensitivity to ampicillin (25 μg/ml^{-1}) and resistance to tetracycline (10 μg ml^{-1}).

Electrophoresis of RNA and transfer to nitrocellulose filters

Samples of RNA (up to 5 μg per slot) were treated with glyoxal in formamide and fractionated on 1.5% agarose gels containing 7 M urea (11). The gels were soaked in 10 mM sodium phosphate buffer pH 7.0 for 20 min prior to transfer to nitrocellulose filters (12). Alternatively the RNA was transferred to Ultrablot (Collaborative Res.) which was handled according to the manufacturer's instructions. Ultrablot filters were reused. The hybridisation was performed in 50% formamide, 5xSSC, 50 mM sodium phosphate pH 6.5, 0.2% SLS, 0.02% w/v each of bovine serum albumin, ficoll, and polyvinylpyrrolidone and 10 μg/ml poly rA. The filters were washed in 2xSSC 0.1% SLS at room temperature and in 0.2xSSC, 0.2% SLS at 50°C with two changes of each buffer.

Hybrid-selected translation

a) Hybrid-arrested translation. Selected cDNA fragments (see text for details) were transferred by sub-cloning to the M13 vector Mp9 (13). DNA preparations of the sub-clones were incubated essentially as described (14), at various ratios with 0.5 μg of polyA$^+$ RNA in 0.4 M NaCl, 10 mM PIPES pH 6.4. The samples were sealed in glass capillaries, heated at 100°C for 30 seconds and at 65°C for 1 hour. 5 μl of each sample was added to 100 μl H_2O, boiled for 2 minutes and frozen in liquid nitrogen prior to precipitation with 2.5 volumes ethanol. The remaining non-boiled portion was also ethanol-precipitated. All samples were resuspended in H_2O for *in vitro* translation.
b) Hybrid release translation was carried out as described (6); in some experiments, as indicated in the text, DNAs from M13 subclones were used instead of relaxed plasmid DNAs.

Preparation of ^{32}P-labelled probes

γ-^{32}P-ATP end-labelling of polyA$^+$ RNA. RNA samples in 25 μl of 50 mM Tris-

HCl pH 9.0, 10 mM $MgCl_2$, 5 mmM dithiothreitol and 5% v/v glycerol were partially hydrolysed by incubation at 100°C for 1 minute. The sample was diluted with 25 µl of the same buffer and incubated with 20 µCi γ-^{32}P-ATP (2000 Ci.mmol^{-1}, Amersham) and 4 units T4 polynucleotide kinase (Bethesda Research Labs) at 37°C for 30 minutes. Unincorporated ATP was removed by gel-filtration on G-100 Sephadex in 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.1% w/v SLS.

α^{32}P dATP labelling of plasmid inserts. Plasmid cDNA inserts were prepared by fractionating PstI digests on 6% polyacrylamide gels (15) and recovery according to Maxam and Gilbert (1977) or from low gelling temperature agarose gels (17). The DNA fragments were labelled with α-^{32}P-dATP (400 Ci.mmol^{-1}, Amersham) by 'nick-translation' (18).

Preparation of plasmid DNAs

Plasmids were prepared as single colony lysates (19) or by CsCl-ethidium bromide centrifugation of Triton lysates (20).

Filter hybridizations

Plasmid DNA samples were transferred to nitrocellulose according to (21). Hybridizations were performed at 65°C in 2xSSC, 10 x Denhardt's solution (22). Filters were washed in 2xSSC, 0.1% SLS at 65°C, unless otherwise indicated.

RESULTS

Isolation of mRNA and analysis of in vitro translation products

Before proceeding to the synthesis of a cDNA bank of clones it was necessary to isolate mRNA from endosperm tissue actively synthesizing storage proteins and to characterise the RNA.

PolyA$^+$ RNA isolated from membrane-bound polysomes (6) or extracted directly from wheat endosperm tissue stimulates the incorporation of [^{3}H]-proline or [^{3}H]-leucine into translation products in a wheat germ cell-free protein synthesizing system. PolyA$^+$ RNA obtained by both procedures and translated *in vitro* gave the same pattern of synthesized polypeptides (Fig. 1). We used the direct extraction method most often, because the yield was higher (10g endosperm gave 30 µg polysomal polyA$^+$ RNA or 50 µg total polyA$^+$ RNA) and the direct-extracted polyA$^+$ RNA showed a slightly higher template activity in the *in vitro* translation (2×10^6 cpm compared to 1.4×10^6 cpm per 0.5 µg RNA translated).

RNA was prepared routinely from endosperm tissue harvested about 20 days after anthesis. When polyA$^+$ RNA isolated from grains 10 days after anthesis

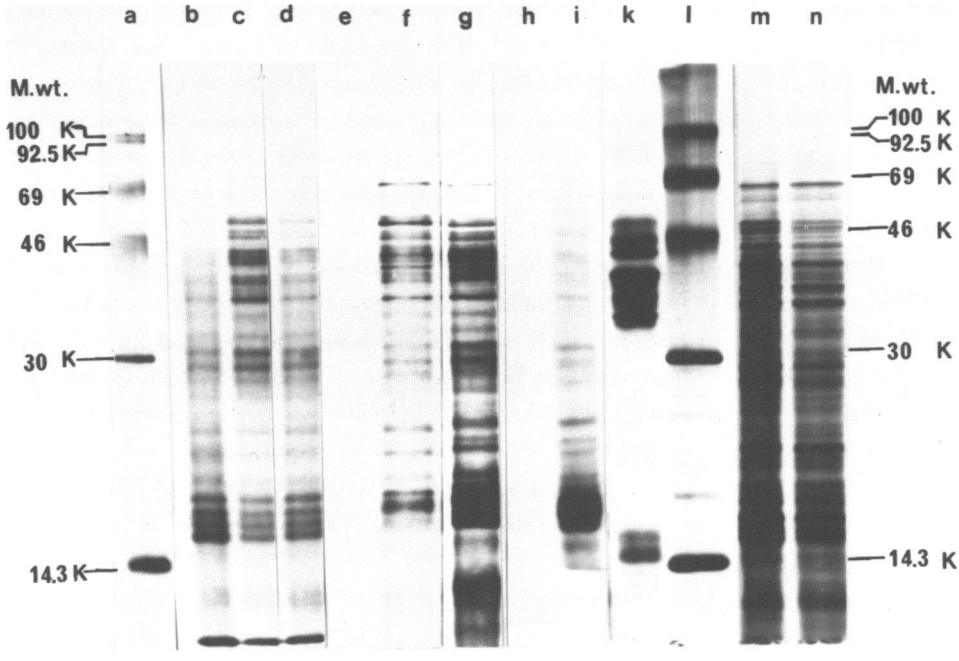


Fig. 1: Separation of in vitro translation products on a 12.5% SDS-urea-polyacrylamide gel. Different wheat endosperm poly(A) RNA preparations were used to direct protein synthesis in a wheat germ in vitro system. In lanes c, d, f, g, i, m, total endosperm poly(A) RNA was translated. In lane (b) polysomal endosperm poly(A) RNA was translated, and in lane (n) poly(A) RNA from 10 day old grains was translated. The products in lanes, b, c, e, f, g, m, n were labelled with ^{3}H-leucine and proline; the products in lane (d) were labelled with proline only and in lane (i) with lysine only. Lane (f) shows the isopropanol-soluble translation products. In lane (e) and (h) no RNA was used and labelled proline/leucine and lysine supplied respectively. Lane (k) shows ^{14}C-labelled authentic gliadins. ^{14}C-size marker proteins were separated in lanes (a) and (l).

was translated, the pattern of the in vitro translation products did not show significant differences (Fig. lm,n), indicating that the abundant mRNA populations present at that early developmental stage are similar to those extracted 20 days after anthesis.

The polypeptides synthesized in vitro ranged from 10-100 KD in size, with many prominent products (Fig. 1). Because of the difficulty of matching stained gels with autoradiographs, gliadins labelled in vitro with ^{14}C-iodoacetate were used as a reference to indicate the electrophoretic mobility of the in vitro gliadin translation products. Although not showing exactly the same pattern, gliadins labelled with ^{14}C-iodoacetate in vitro comigrate with

unlabelled gliadins (8) in a stained polyacrylamide gel. Proteins migrating similarly to the ^{14}C-gliadins (48-36 KD) were prominent amongst the in vitro synthesized proteins (Fig. 1). These in vitro translation products were analysed further for gliadin-like features. Gliadins are rich in proline (about 17 mol %) and leucine (about 7.3 mol %) but poor in lysine (0.5 mol %) (1), and a high proline or leucine to lysine ratio is therefore a characteristic of these proteins. When 1 µg polyA$^+$ RNA was translated in vitro using different labelled amino acids, 1.8 nmol ^{3}H-proline or 2.6 nmol ^{3}H-leucine were incorporated into trichloroacetic acid-insoluble material compared to 0.36 to 0.16 nmol ^{3}H-lysine. Most of the lysine was incorporated into proteins of low molecular weight (10-16 KD) (Fig. 1i), although after the gel had been exposed for two weeks, radioactivity could also be detected in higher molecular weight proteins comigrating with polypeptides which had incorporated proline and leucine.

Since cereal prolamins are soluble in 55% isopropanol (23) in contrast

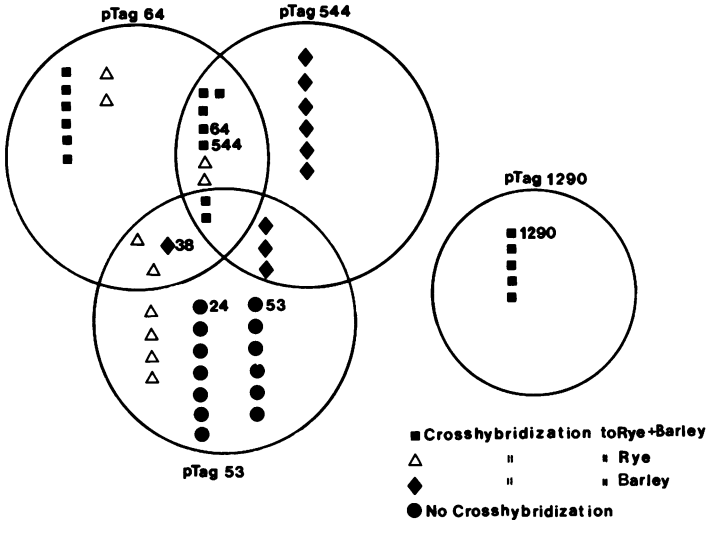
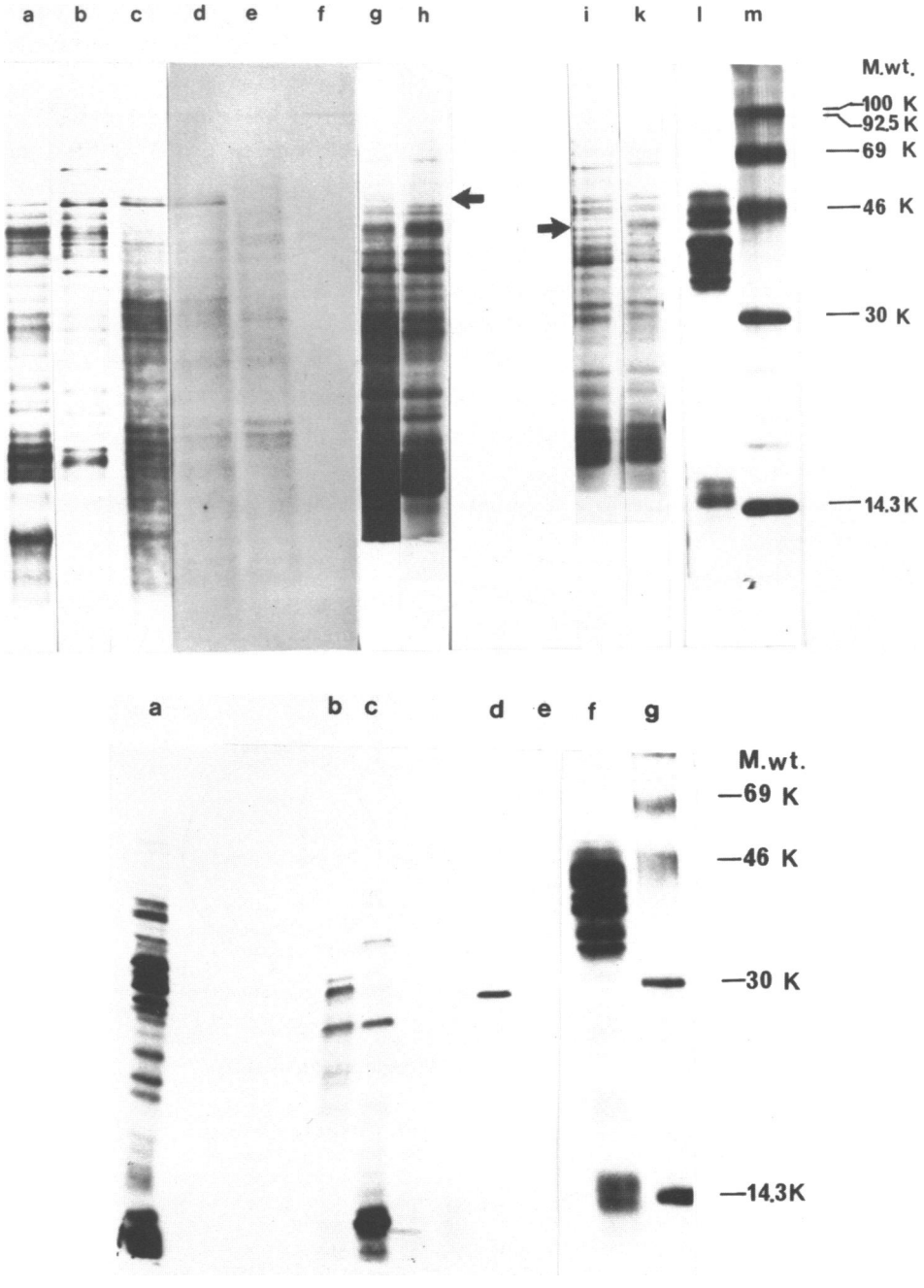


Fig. 2: Venn Diagram showing the cross hybridization between 50 selected cDNA clone, which all hybridized strongly to wheat endosperm poly(A) RNA. Pst I excisable inserts were prepared from four different clones. The inserts were labelled with ^{32}P by nick translation and hybridized to the selected clones to identify sets of related sequences. Clones within the overlaps hybridized to more than one probe. Only those clones which were examined further in the present study are indicated by their numbers, the remainder are identified by symbols representing the cross hybridization they show to endosperm poly(A) RNA from related cereals.


a b c d e f g h
i k l m
M.wt.
100 K
92.5 K
69 K
46 K
30 K
14.3 K
a b c d e f g
M.wt.
69 K
46 K
30 K
14.3K

to most other proteins, the *in vitro* translation products were extracted at 60°C, with 55% aqueous isopropanol containing 0.1% DTT. Many of the proteins synthesized *in vitro* were soluble in isopropanol (Fig. 1f).

Having shown that the polyA$^+$ RNA extracted from endosperm tissue directs the *in vitro* synthesis of proteins with the properties of gliadins, this RNA was used as a template to construct cDNA clones.

Construction and characterization of cDNA clone bank

cDNA was prepared from 20 µg oligo dT-cellulose-purified RNA from immature endosperms of wheat var. 'Chinese Spring' as described in Materials and Methods. cDNA was inserted into the Pst site in pBR322 by homopolymer G-C tailing. The resulting Amp^S clones were screened by filter hybridizations (Materials and Methods) using ^{32}P-labelled polyA$^+$ RNA from shoot and endosperm tissues. Those clones hybridizing only to the endosperm polyA$^+$ RNA were used for further analyses. The cDNA insert varied in length from 350-1500 base pairs as determined by digestion with appropriate restriction enzymes.

In order to classify the cloned DNAs into sequence-related groups, the bank was also screened with ^{32}P-labelled polyA$^+$ RNA purified from the endosperm of cereal species related to wheat. These results are described in detail elsewhere (24). They enabled us to select three clones each representing a distinct sequence group. One clone hybridized to barley RNA in addition to wheat RNA, one to rye RNA in addition to wheat RNA and one to wheat RNA only. The bank was then probed with excised, ^{32}P-labelled inserts

Fig. 3a and 3b: SDS-Urea-polyacrylamide gel electrophoresis of hybrid-selected translation products.
Fig. 3a: lanes (a) and (b) translation products of wheat endosperm poly(A) RNA, lane (b) isopropanol-soluble products. These products are to be compared with the translation products of RNA selected by M13 544 Alu-3 (lane c) and isopropanol extracted products lane (d); lanes (e) and (f) show the controls obtained by using the vector only ((f) isopropanol extracted, (e) total products). Lane (h) presents the results from a hybrid arrested translation using M13 544 Alu-3, lane g shows the control experiment where the hybrid had been melted and the RNA subsequently translated. Lanes (i) and (k) show the same experiment for pTag53; in lane (i) the translation products after hybridisation of pTag53 to the RNA, and in lane k translation products from the melted sample. Lane (l) contains ^{14}C-labelled gliadin and lane (m) ^{14}C-size marker proteins.
Fig. 3b: lane (a) total translation products of wheat endosperm poly A RNA lane (b) translation products selected by pTag24, lane (c) translation products selected by pTag38, lane (d) translation products selected by pCAM 179 (identified as coding for a B-hordein, Forde *et al.*, 1981) by hybridizing to wheat endosperm poly A RNA. Lane f shows ^{14}C-labelled gliadins and lane g ^{14}C-marker proteins.

from these clones and others and the results are summarized in Fig. 2 (Venn Diagram). To examine the degree of heterogeneity within the hybridization groups we digested the cloned DNAs with restriction enzymes. Digestions of several cross-hybridizing clones with one restriction enzyme produced several different sized fragments arising from the cDNA inserts. The sum of the internal fragments was too great to be aligned on a single mRNA. Therefore we assume that sequence diversity exists within each hybridization group. The hybridization between members of the three groups was generally much less than within the group. Restriction enzyme maps were constructed for two clones, pTag64 and pTag544. These clones, which have 800 and 920 base pairs long inserts respectively, showed some cross-hybridization but had no common restriction fragments. The restriction enzyme map of pTag544 is shown in Fig. 5. The fragment indicated Alu-3 from pTag544 was subcloned into M13 mp9 for hybrid arrested (HART) and hybrid release translation (HRT) analyses.

Clones from the major hybridization groups in the Venn Diagram were used in hybridizations to mRNA, size-fractionated on agarose-urea gels and in hybrid release or arrest translation in order to determine the coding properties.

Identification of cDNA clones by hybrid-selected translation

cDNA clones representing different hybridisation groups were used to determine their coding potential. Three clones, pTag24, 38 and 544 were subjected to hybrid-selected translation (6). The results are shown in Fig. 3. pTag24 gave translation products in the α+β-gliadin region, with some evidence of smaller products, possibly artefacts. pTag38 gave translation products in the same region, but in addition there was strong hybridization to RNA encoding low molecular weight proteins. A sub-clone of pTag544 in M13mp8 containing Alu-3 (Fig. 5) was used for both hybrid-release and hybrid arrest experiments. In the hybrid arrest sample (Fig. 3A,h) two bands which comigrate with major polypeptides in the γ-gliadin region were absent; these polypeptides were synthesized when the hybrids were dissociated by boiling (Fig. 3A,g). These bands were also visible in the translation products of RNA released from this DNA in a hybrid- release sample (Fig. 3A,c). A hybrid-release sample obtained from plasmid pTag544 gave a similar result. The polypeptide products were further characterized by extraction with 55% isopropanol (Fig. 3A,d). This solvent extracts the high-molecular-weight products in the γ-gliadin region, but fails to extract the fainter bands of lower molecular weight. This may indicate that the signal in this region present in the unextracted sample is due to incomplete products of synthesis

which are not extracted from polysomes.

Size-determination of mRNAs by electrophoresis

cDNA clones representing different hybridisation groups were used as probes to identify the size-class of related mRNAs. Samples of 5 µg polyA$^+$ endosperm RNA were glyoxalated (11), separated on a denaturing urea-agarose gel and transferred to nitrocellulose or ultrablot paper. The filters were hybridized with ^{32}P-labelled excised inserts from selected cDNA clones (Fig. 4). Variation in complementary sequence concentration and in the size of the complementary RNAs were seen for different clones. pTag544 hybridized to RNA species of 1400-1550 nucleotides; pTag64 to RNA of 1400 nucleotides, pTag24 to RNA of 1300, and pTag38 to RNA of 1150-1250 nucleotides. The estimated minimum coding length, based on the apparent molecular weight of the largest

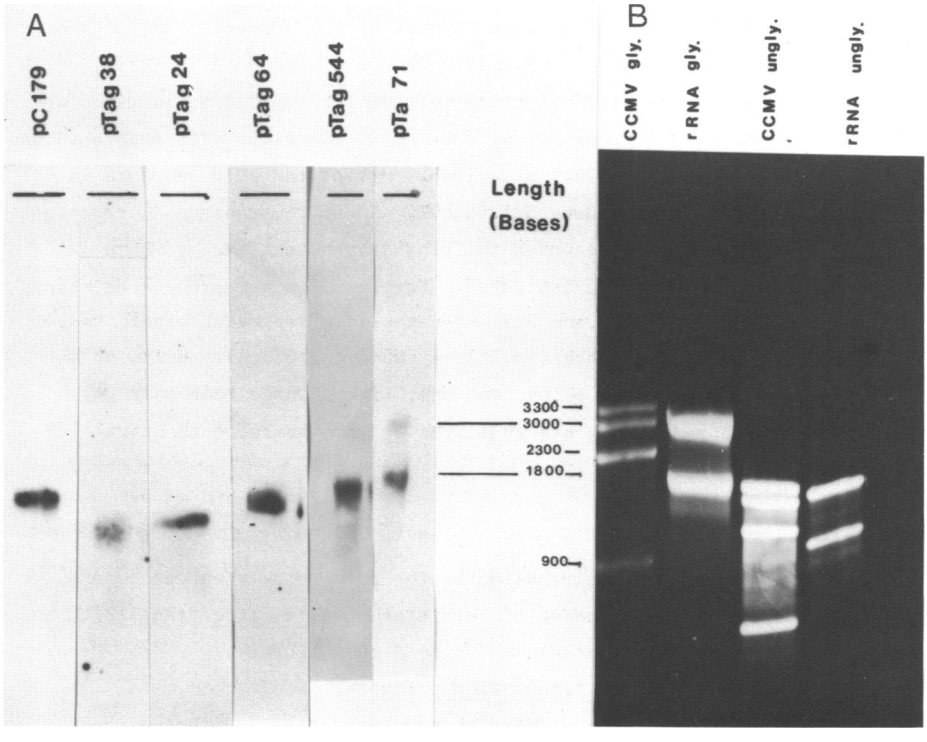


Fig. 4 (A): Filter hybridizations with pTag cDNA clones, pC179 (a B-hordein clone (6) and pTa71 (a wheat ribosomal RNA gene clone (32)) to wheat endosperm poly(A) RNA glyoxalated and fractionated on urea-agarose gels, before transfer.

(B) shows part of the gel stained with ethidium bromide, to visualize the size markers and monitor the effect of glyoxal treatment.

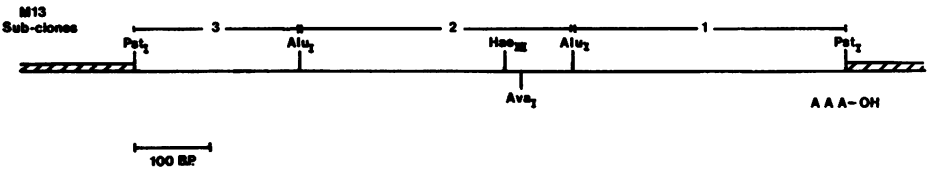


Fig. 5: Restriction enzyme cleavage site map of pTag544.

hybrid-release product, for pTag544 is 1400 nucleotides and for pTag24 is 1250 nucleotides. These estimates allow for 150-200 nucleotides of non-coding sequence. The largest hybrid-release product of pTag38 is apparently 40KD giving a minimum coding length of 1000 nucleotides. However the heterogeneity of the products encoded for in the hybrid selection translation is not reflected in different mRNA size classes.

The cDNA insert in pTag544 was sequenced by the chemical (16) method. The DNA sequence is presented in Fig. 6. The predicted reading frame, which gives an amino acid composition similar to that for γ-gliadin, is given above the DNA sequence. The sequence shows substantial homology with that obtained for barley clone pC16 (6). For comparison, part of the pC16 sequence is presented, aligned from residues 533 to 691. The pC16 sequence is represented by a solid horizontal line under the pTag544 sequence except where it differs, when the changes are indicated. Vertical lines represent deletions (positions 604-611 and 618), and bracketed residues represent single base insertions in the pC16 sequence (positions 666-7 and 673-4). Units of a tandem repeat two of 42 base pairs, and one, with a three base pair insertion, 45 base pairs long are indicated by arrowheads in the first 129 residues.

DISCUSSION

In order to study gene expression in the developing wheat endosperm, we have analysed the mRNA complement of this tissue by *in vitro* translation and nucleic acid hybridization techniques using a cDNA library.

Characterization of in vitro translation products

PolyA$^+$ RNA isolated from developing endosperm tissue directed the synthesis of proteins *in vitro* which were shown to be gliadin-like by the following criteria: (1) comigration on denaturing polyacrylamide gels with gliadin standards; (2) relative incorporation of ^{3}H-proline and ^{3}H-leucine versus ^{3}H-lysine in the *in vitro* translation products; (3) solubility of the proteins in 55% aqueous isopropanol after treatment with a reducing agent

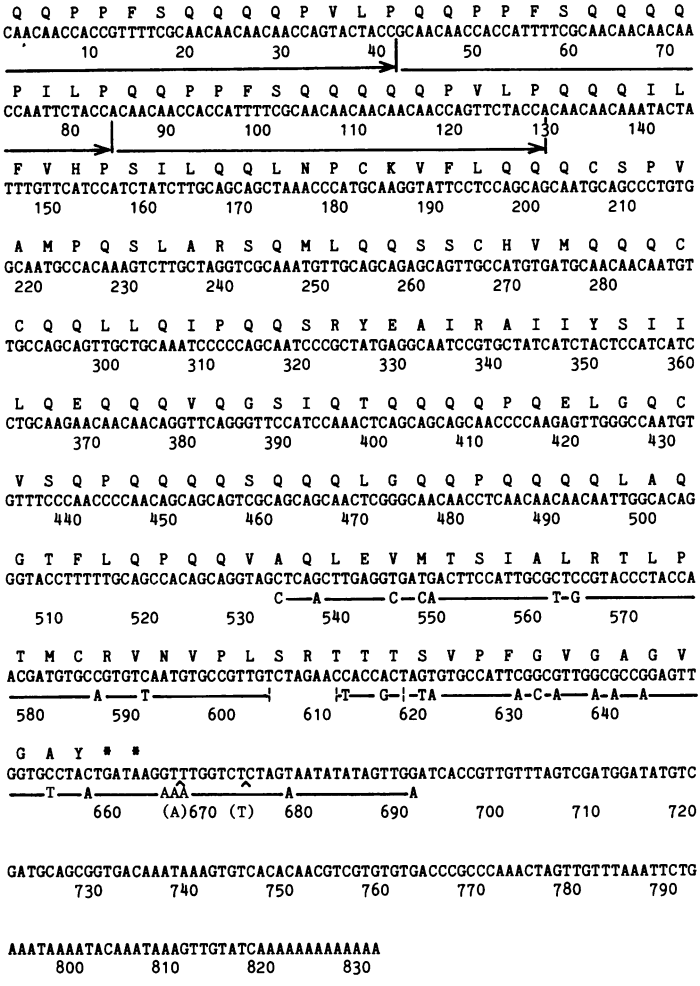


Fig. 6: Nucleotide sequence of pTag544 and comparison with part of pC16 sequence. Only pC16 residues differing from those in pTag544 DNA are shown, while insertions and deletions in pC16 relative to pTag544 are represented by Λ and respectively. Arrows in the first 129 residues denote position of tandem repeat units (see text).

(2-β-mercaptoethanol and dithiothreitol (DTT)).

These techniques have also been used to identify hordein polypeptides in the translation products of barley endosperm polyA$^+$ RNA (6, 26). We have also shown that these in vitro translation products can be immunoprecipitated with antibodies raised to gliadin fractions (Bartels, Thompson and Ciclitira, in preparation).

Other workers have reported the *in vitro* synthesis of gliadin-like proteins (27), (28). The sizes and properties of the products described here are very similar to these reports. We attribute slight differences between the size estimates to the use of different wheat varieties and different gel systems in fractionation of the proteins.

Okita and Greene (27) and Donovan *et al.* (28) found that the *in vitro* translation products are larger than the corresponding *in vivo* proteins and suggested this is due to the presence of short additional signal peptides not cleaved off in the wheatgerm translation system. We also do not see exact correspondence between *in vitro* synthesized polypeptides and those extracted from the grain. Because of the complexity of the product profile and the practical difficulty of aligning precisely the stained proteins with the translation products which are visualized by fluorography, we feel that further evidence for the presence of a signal sequence for each polypeptide is still necessary.

Characterization of cloned sequences

The properties of the cDNA clones obtained indicate that the gliadins are synthesized from families of related genes. Sequence homologies within the cDNA bank were detected by two experiments. Firstly the clones were hybridized with polyA$^+$ RNA from the endosperms of related cereals (24) and secondly selected clone inserts were hybridized to the clone bank. This data is summarized in the Venn Diagram (Fig. 2).

Both approaches indicated that a number of families of related sequences had been cloned. Three major sequence-related groups were identified, by clones 64, 53 and 544.

The 64 group hybridized strongly to rye polyA$^+$ endosperm RNA and some members weakly to barley, the 544 group hybridized strongly to barley polyA$^+$ endosperm RNA, whereas the 53 group only shared homology with the wheat diploid species *T. monococcum*. An additional discrete hybridization group was defined by pTag1290 insert. Clones in this group hybridized to rye and barley polyA$^+$ endosperm RNA.

All clones tested by hybrid-release translation were complementary to mRNAs coding for more than one polypeptide. pTag24, 38 and 53 were complementary to mRNA for polypeptides in the α/β-gliadin region. pTag544, which has sequence homology with barley hordein cDNA clone pC179 (6), was complementary to mRNA for polypeptides in the γ-gliadin region. Further analysis of the hybrid selected translation products is necessary in order to give a final alignment to *in vivo* synthesized proteins.

None of the clones examined selected mRNA for polypeptides in the ω-gliadin region.

The determined cDNA sequence of pTag544 extends 219 residues into the coding sequence. The corresponding mRNA was estimated by hybridization analysis to be approximately 1400 bases in length (Fig. 4), therefore the cDNA clone stops approximately 300 bases or 100 amino acid residues from the amino terminus. The amino acid composition of the region sequenced is similar to that determined for gliadins (1) although it is relatively rich in cysteine. The relatedness in sequence to B-hordein mRNA which was used to characterize the clone is confirmed by comparison to the DNA sequence of hordein cDNA clone pC16 (6). The homology extends into the 3' untranslated region of the two mRNAs.

At the 5' end of the cDNA clone there is a tandem DNA repeat with a periodicity of 14 amino acid residues. The DNA sequence is repeated with little degeneracy at the codon third base positions, suggesting that the sequence repetition is a comparatively recent event or has been homogenised recently, e.g. by intra gene conversion. Similar internal repetition may be responsible for generating some of the divergence seen within gliadin families which are of similar amino acid compositions andshare common amino and carboxy terminal sequences, but differ in size. As the repeat lies at the terminus of the cDNA clone, the extent and significance of it cannot be completely assessed. Within the repeat, hydrophobic tetraplets are separated by 2-5 glutamine residues in a similar arrangement to part of Z-19 zein (29). The repeated sequence observed here gives the clearest indication of the periodicity of a DNA repeat so far observed in prolamin genes. It seems likely that if this polypeptide has a structure consistent with the model of Argos et al. (30), based on a 20 amino acid repeat, then the wheat coding sequence is generated in part by repeats of DNA lengths other than the sixty-nucleotide repeat of zein. This observation is consistent with that made by Geraghty et al. (31) for the 19KD zein clone A20, which contains an apparent direct repeat of 108 nucleotides.

The 3' untranslated region of the clone contains two examples of the putative polyadenylation signal sequence AATAAA (33) at residues 974-799 and 805-810.

Analysis of the length of complementary mRNAs to selected clones by hybridization to size-fractionated, filter-bound RNA indicated the presence of size classes of mRNA complementary to pTag 544, 24, 64, and 38. In view of the results of the hybrid-selected translation experiments with these

clones, we conclude that more than one mRNA is present within one size class.

Six major genetic loci for the gliadins have been described using 2-dimensional electrophoresis of proteins from aneuploid lines (8). The translation products of mRNA from these lines are currently being used to assign the polypeptides synthesized in vitro to individual chromosomes, and to further classify the cloned sequences. This may provide more information about the homology between the gliadin loci and also about the homology between gliadin loci and genes in related cereals. The gliadin cDNA sequences identified also allow us to inspect gliadin gene number and organization in the wheat genome, and provide probes for the isolation of the DNA sequences which comprise the complex loci.

ACKNOWLEDGEMENTS

We thank Dr. R.B. Flavell for helpful discussions throughout this work and S. Martin and M. O'Dell for their help in harvesting the endosperm.

R.T. wishes to thank Dr. H. Sommer and Dr. C. Lazarus for instruction in sequencing techniques. D.B. was supported by the Deutsche Forschungsgemeinschaft.

REFERENCES

1 Kasarda, D.D., Bernardin, J.E., and Nimmo, C.C. (1976). In Advances in Cereal Science and Technology 1 (ed. Pomeranz Y.) pp.158-236 (American Association of Cereal Chemists, St. Paul, Minnesota, 1976).
2 Payne, P.I., Holt, L.M., Lawrence, G.J. and Law, C.N. (1982). 'The genetics of gliadin and glutenin, the major storage proteins of the wheat endosperm'. Qual. Plant. Plant Foods Hum. Nutr. 31, 3, pp.229-249.
3 Autran, J.C., Lew, E.J-L., Nimmo, C.C. and Kasarda, D.D. (1979). Nature 282, pp.527-529.
4 O'Dell, M., and Thompson, R.D. (1982). J. Sci. Fd. Agric. 33, pp.419-420.
5 Kloppstech, K. and Schweiger, H.G. (1976). Cytobiologie 13, pp.394-400.
6 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, pp.6689-6707.
7 Bonner, W.M. and Laskey, R.A. (1974). Eur. J. Biochem. 46, pp.83-88.
8 Payne, P.I., and Corfield, K.G. (1979). Planta 145, pp.83-88.
9 Weinryb, J. (1968). Arch. Biochem. Biophys. 124, pp.285-291.
10 Wickens, M.P., Buell, G.N. and Schimke, R.T. (1978). J. Biol. Chem. 253 (7) pp.2483-2495.
11 Koller, B., Delius, H. and Dyer, T.A. (1982). Eur. J. Biochem. 122, pp.
12 Thomas, P.S. (1980). Proc. Nat. Acad. Sci. U.S.A. 77, pp.5201-
13 Messing, J. (1981). In Proceedings of the Third Cleveland Symposium on Macro-molecules, ed. A.G. Walton, Elsevier Amsterdam, pp.143-153.
14 Paterson, P.M., Roberts, B.E., Kuff, E.L. (1977). Proc. Nat. Acad. Sci. U.S.A. 74 (10) pp.4370-4374.

15 Peacock, A.C. and Dingman, C.W. (1967). Biochem. 6, pp1818-1827.
16 Maxam, A.M. and Gilbert, W. (1977). Proc. Nat. Acad. Sci. U.S.A. 74, pp.560-564.
17 Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.G.H. and Roe, B.A. (1980). J. Mol. Biol. 143, pp.161-178.
18 Maniatis, T., Jeffrey, A. and Van De Sande, H. (1975). Biochemistry 14, pp.3787-3794.
19 Barnes, W.M. (1977). Science 195, pp.393-394.
20 Clewell, D.B. (1972). J. Bact. 110, pp.667-676.
21 Southern, E.M. (1975). J. Mol. Biol. 98, pp.503-517.
22 Denhardt, D.T. (1966). Biophys. Biochem. Res. Commun. 23, pp.641-646.
23 Larkins, B.A. (1981). Biochemistry of Plants 65, pp.40-489, Springer Verlag, Berlin.
24 Bartels, D. and Thompson, R.D. (1983). Theor. Appl. Genet. 64, pp.269-273.
25 Schmitt, J.M. and Svendsen, I. (1980). Carlsberg Res. Commun. 45, pp.549-555.
26 Matthews, J.A. and Miflin, B.J. (1980). Planta 149, pp.262-268.
27 Okita, T.W. and Greene, F.C. (1982). Plant Physiol. 69, pp.834-839.
28 Donovan, G.R., Lee, J.W. and Longhurst, T.J. (1982). Aust. J. Plant Physiol. 9, pp.59-68.
29 Pedersen, K., Devereux, J., Wilson, D.R., Sheldon, E. and Larkins, B.A. (1982). Cell 29, pp.1015-1026.
30 Argos, P., Pedersen, K., Marks, M.D. and Larkins, B.A. (1982). J. Biol. Chem. 257, pp.9984-9990.
31 Geraghity, D.E., Messing, J. and Rubenstein, I. (1982). EMBO Journal 1, pp.1329-1335.
32 Gerlach, W.L. and Bedbrook, J.R. (1979). Nucl. Acids. Res. 7, pp.1869-1885.
33 Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980). Nucl. Acids. Res. 8, pp.127-142.