Cereal chemistry, molecular biology, and toxicity in coeliac disease

R P Sturgess, H J Ellis, P J Ciclitira

The health of Dutch children with coeliac disease improved during the second world war when cereal foods were in short supply. Dicke, extending a prewar observation on the potential toxicity of cereal products,1 noted that the reintroduction of bread into the diet of these children produced a deterioration in their health which led to the conclusion that a gluten free diet could be used to treat the disorder.2 Wheat, rye, barley, and possibly oats were found to exacerbate the condition.3

Classification of cereal fractions
Studies have concentrated on wheat. Individual grains of wheat can be separated into the outer husk or bran, the germ or semolina, and the endosperm or flour (Fig 1), which in the United Kingdom represents 70–72% of the total grain by weight. The main constituents of wheat flour comprise starch (70–72%), protein (7–15%), lipids (1–2%), and water. Gluten is the product of a ball of wheat flour dough that has been exhaustively washed in tap water. Osborne4 classified wheat endosperm proteins into four solubility classes: albumins, which are soluble in water; globulins, which are soluble in salt solutions but are insoluble in water; gliadins, which are soluble in 50–90% ethanol; and glutenins, which are insoluble in neutral aqueous or saline solutions and ethanol (Fig 2). The baking qualities of wheat depend on the ability of glutenins to trap carbon dioxide in dough. The ethanol soluble products of cereals, collectively known as prolamins, are termed secalins from rye, hordeins from barley, avenins from oats, and zeins from maize, the latter of which is not toxic. These solubility characteristics are only relative; for example, the lower molecular weight gliadins are known to be soluble in ethanol. The gliadins are further subdivided into subfractions according to their relative electrophoretic mobility.5–8 The subfractions of gliadin are termed α, β, γ, andω with corresponding molecular weights ranging from 32 to 58 kDa. Unfractionated gliadin comprises approximately 30% α, 30% β, 30% γ, and 10% ω gliadin.

Early studies
Dicke and colleagues separated wheat into fractions including (i) gluten, (ii) a water soluble fraction that contained starch and some water soluble proteins including gliadins, (iii) glutenin, (iv) gliadin, (v) crude fibre, (vi) fat, and (vii) ash from heating the flour to 500°C.9 Clinical testing was performed on a 6 year old

Classification of wheat proteins

Figure 1: Diagrammatic representation of a grain of wheat.

Figure 2: Classification of wheat proteins.
Figure 3: Production of wheat grain endosperm cDNA.

Ground wheat grain endosperm

Poly A+ mRNA

5’

Phenol/chloroform extraction + oligo dt column purification

AAA 3’

Reverse transcriptase + deoxynucleotides

AAAA

Heat snap cool separation

TTT 5’

DNA polymerase + deoxynucleotides

TTT 5’

AAA 3’

S1 nuclease

TTT 5’

AAA 3’

Terminal transferase + dCTP

CCC

Homopolymer tailed cDNA

'Hairpin' loop

Double stranded cDNA

3’

5’

Figure 3: Production of wheat grain endosperm cDNA.

A patient who had previously been maintained on a gluten free diet. Histological relapse occurred in serial jejunal biopsies after an intraduodenal infusion of 7–5 g of α gliadin. Kendall and colleagues separated gliadin into twelve fractions by ion exchange chromatography, pooling them into three groups which were termed pre-α, α, and post-α-gliadins. They reported that only the α-fraction was toxic to coeliac patients. However, they failed to provide evidence for the purity of their fractions and used the urinary D-xylene test, an inaccurate method of assessment, to investigate their fractions. Clearly, to attempt to define the putative toxic peptide, further methods for the purification, characterisation and identification of cereal proteins needed to be developed.

Molecular biology

To assess any given cereal fraction the purity, method of separation, and starting material need to be defined. Methods for the purification of wheat albumins, globulins, and glutenins have been reviewed, and of importance for toxicity studies, methods have been developed for the consistently reproducible preparation, in sufficient quantities, of gliadin subfractions.

An alternative approach for the production of well characterised cereal proteins is molecular cloning of the genes for wheat gliadins, the methods for which will be considered in detail. Developing endosperm is the site of synthesis of major grain storage proteins. Both gliadins and glutenins are synthesised on membrane bound polysomes and deposited in protein bodies during maturation. They provide an important source of dietary proteins and together constitute the majority of cell free translation products of mRNA isolated from the developing endosperm of wheat grains. The proteins occur as complex, heterogeneous families of polypeptides which have similar physical properties and amino acid compositions. Structural studies have identified amino terminal sequence homologies between wheat and related species, including rye and barley. The homologous group of proteins from these related species may be the products of gene families derived from common ancestors. To assess the extent and distribution of sequence conservation in genes coding for different storage proteins and as an aid to assigning genes to sequence related groups, a library of wheat cDNA clones was prepared. Endosperm tissue was ground to a fine powder, suspended in extraction buffer, and extracted with phenol/chloroform. Shoot tip mRNA was extracted from four day old etiolated shoot tips, to serve as a control.

Purified wheat endosperm mRNA was incubated with deoxynucleotides in the presence of reverse transcriptase (Fig 3). The mRNA was removed by heating the mixture to 100°C for two minutes and snap cooling in liquid nitrogen. This produced complementary DNA with a hairpin loop. After the addition of DNA polymerase I in the presence of deoxynucleotides, double stranded cDNA was produced. The hairpin loop was excised by the addition of S1 nuclease. Cohesive, homopolymer-tailed ends of
Cereal chemistry, molecular biology, and toxicity in coeliac disease

medium because of the presence of the tetracycline resistant gene in the transformed bacteria. Plasmids containing inserts at the PstI site were identified by their sensitivity to ampicillin on replica plates.

The mRNA was labelled with $^{32}$P, by incubation in the presence of a polynucleotide kinase. After replication of the bacterial clones on a millipore filter, they were lysed in situ with alkali and hybridised to the radioactively labelled RNA probe. This was undertaken with both the shoot-specific and endosperm-specific mRNA probes. The filter was exposed to x ray film after the unhybridised material had been washed away. Only the endosperm positive clones were selected for further analysis, on the basis that storage protein synthesis is restricted to endosperm tissue alone.

Secondary screening of the clones was undertaken by hybrid arrested translation (HART). The principle of the HART method relies on the observation that an mRNA-DNA hybrid is not translated in an in vitro translation system (Fig 5). Therefore, if most of the particular mRNA sequence in an mRNA population is hybridised to its complementary DNA and the mRNA population is translated, the translation which is encoded by the hybridising mRNA will be reduced or eliminated. Total wheat endosperm mRNA in the presence and absence of the cDNA clone under study was added to a reticulocyte lysate translation system. After incubation the labelled translation products were resolved in a sodium dodecyl sulphate polyacrylamide gel and detected by autoradiography. Identification of translation products which were not expressed in the presence of the cDNA clone under study indicated that the clone was complementary to RNAs for these products.

Further confirmation of the activity of the expressed proteins was undertaken by immunoprecipitation with specific antigliadin rabbit antisera. The immune precipitates were isolated with protein A-Sepharose and identified with one dimensional polyacrylamide gel electrophoresis.

These experiments permitted isolation of cDNA clones for several classes of wheat protein, including one $\alpha$ gliadin, and two $\gamma$ gliadin proteins as well as high molecular weight glutenin proteins. The clones were sequenced by the method of Maxam and Gilbert which allowed subsequent determination of their amino acid sequences. The results of sequencing clones encoding an $\alpha$ gliadin protein were largely in agreement with similarly produced amino acid sequences for these proteins reported by Kasarda and colleagues. These studies have permitted the full 266 amino acid sequence of $\alpha$-gliadin to be determined (Table) allowing the domain structure of this molecule to be defined (Fig 6). It should, however, not be forgotten that the amino acid sequences of individual $\alpha$ gliadin proteins are not all identical. Three gliadin N-terminal amino acid sequences are recognised, $\alpha$, $\gamma$, and $\omega$, whose nomenclature was derived from the electrophoretic classification of gliadin subfractions into four groups, $\alpha$, $\beta$, $\gamma$, and $\omega$.

The inserts of well characterised cDNA clones for gliadin sequences were integrated into the double stranded cDNA were generated by the addition of deoxyCTP in the presence of DNA terminal transferase.

The homopolymer-tailed cDNA clones were introduced into a plasmid, PBR 322. One advantage of PBR 322 is that it contains two antibiotic resistance markers, one for ampicillin and another for tetracycline, allowing for selection of the plasmid when one marker has been inactivated by an insertion. The plasmid was linearised with a restriction enzyme which cleaved the plasmid at the PstI site of PBR 322, allowing insertion into the ampicillin resistance gene (Fig 4). After this homopolymer-tailing was undertaken by enzymatic addition of deoxyGTP using terminal transferase. The deoxyCTP homopolymer-tailed double stranded cDNA was incubated with deoxyGTP complementary homopolymer-tailed PBR 322, permitting recircularisation of the plasmid after insertion of the double stranded DNA. Plasmid recombinants were recovered by transformation of competent Escherichia coli cells. The bacteria were plated out onto standard culture plates. Only those bacteria which were transformed with PBR 322 grew on tetracycline-containing medium because of the presence of the tetracycline resistant gene in the transformed bacteria. Plasmids containing inserts at the PstI site were identified by their sensitivity to ampicillin on replica plates.

The mRNA was labelled with $^{32}$P, by incubation in the presence of a polynucleotide kinase. After replication of the bacterial clones on a millipore filter, they were lysed in situ with alkali and hybridised to the radioactively labelled RNA probe. This was undertaken with both the shoot-specific and endosperm-specific mRNA probes. The filter was exposed to x ray film after the unhybridised material had been washed away. Only the endosperm positive clones were selected for further analysis, on the basis that storage protein synthesis is restricted to endosperm tissue alone.

Secondary screening of the clones was undertaken by hybrid arrested translation (HART). The principle of the HART method relies on the observation that an mRNA-DNA hybrid is not translated in an in vitro translation system (Fig 5). Therefore, if most of the particular mRNA sequence in an mRNA population is hybridised to its complementary DNA and the mRNA population is translated, the translation which is encoded by the hybridising mRNA will be reduced or eliminated. Total wheat endosperm mRNA in the presence and absence of the cDNA clone under study was added to a reticulocyte lysate translation system. After incubation the labelled translation products were resolved in a sodium dodecyl sulphate polyacrylamide gel and detected by autoradiography. Identification of translation products which were not expressed in the presence of the cDNA clone under study indicated that the clone was complementary to RNAs for these products.

Further confirmation of the activity of the expressed proteins was undertaken by immunoprecipitation with specific antigliadin rabbit antisera. The immune precipitates were isolated with protein A-Sepharose and identified with one dimensional polyacrylamide gel electrophoresis.

These experiments permitted isolation of cDNA clones for several classes of wheat protein, including one $\alpha$ gliadin, and two $\gamma$ gliadin proteins as well as high molecular weight glutenin proteins. The clones were sequenced by the method of Maxam and Gilbert which allowed subsequent determination of their amino acid sequences. The results of sequencing clones encoding an $\alpha$ gliadin protein were largely in agreement with similarly produced amino acid sequences for these proteins reported by Kasarda and colleagues. These studies have permitted the full 266 amino acid sequence of $\alpha$-gliadin to be determined (Table) allowing the domain structure of this molecule to be defined (Fig 6). It should, however, not be forgotten that the amino acid sequences of individual $\alpha$ gliadin proteins are not all identical. Three gliadin N-terminal amino acid sequences are recognised, $\alpha$, $\gamma$, and $\omega$, whose nomenclature was derived from the electrophoretic classification of gliadin subfractions into four groups, $\alpha$, $\beta$, $\gamma$, and $\omega$.

The inserts of well characterised cDNA clones for gliadin sequences were integrated into the double stranded cDNA were generated by the addition of deoxyCTP in the presence of DNA terminal transferase.

The homopolymer-tailed cDNA clones were introduced into a plasmid, PBR 322. One advantage of PBR 322 is that it contains two antibiotic resistance markers, one for ampicillin and another for tetracycline, allowing for selection of the plasmid when one marker has been inactivated by an insertion. The plasmid was linearised with a restriction enzyme which cleaved the plasmid at the PstI site of PBR 322, allowing insertion into the ampicillin resistance gene (Fig 4). After this homopolymer-tailing was undertaken by enzymatic addition of deoxyGTP using terminal transferase. The deoxyCTP homopolymer-tailed double stranded cDNA was incubated with deoxyGTP complementary homopolymer-tailed PBR 322, permitting recircularisation of the plasmid after insertion of the double stranded DNA. Plasmid recombinants were recovered by transformation of competent Escherichia coli cells. The bacteria were plated out onto standard culture plates. Only those bacteria which were transformed with PBR 322 grew on tetracycline-containing medium because of the presence of the tetracycline resistant gene in the transformed bacteria. Plasmids containing inserts at the PstI site were identified by their sensitivity to ampicillin on replica plates.

The mRNA was labelled with $^{32}$P, by incubation in the presence of a polynucleotide kinase. After replication of the bacterial clones on a millipore filter, they were lysed in situ with alkali and hybridised to the radioactively labelled RNA probe. This was undertaken with both the shoot-specific and endosperm-specific mRNA probes. The filter was exposed to x ray film after the unhybridised material had been washed away. Only the endosperm positive clones were selected for further analysis, on the basis that storage protein synthesis is restricted to endosperm tissue alone.

Secondary screening of the clones was undertaken by hybrid arrested translation (HART). The principle of the HART method relies on the observation that an mRNA-DNA hybrid is not translated in an in vitro translation system (Fig 5). Therefore, if most of the particular mRNA sequence in an mRNA population is hybridised to its complementary DNA and the mRNA population is translated, the translation which is encoded by the hybridising mRNA will be reduced or eliminated. Total wheat endosperm mRNA in the presence and absence of the cDNA clone under study was added to a reticulocyte lysate translation system. After incubation the labelled translation products were resolved in a sodium dodecyl sulphate polyacrylamide gel and detected by autoradiography. Identification of translation products which were not expressed in the presence of the cDNA clone under study indicated that the clone was complementary to RNAs for these products.

Further confirmation of the activity of the expressed proteins was undertaken by immunoprecipitation with specific antigliadin rabbit antisera. The immune precipitates were isolated with protein A-Sepharose and identified with one dimensional polyacrylamide gel electrophoresis.

These experiments permitted isolation of cDNA clones for several classes of wheat protein, including one $\alpha$ gliadin, and two $\gamma$ gliadin proteins as well as high molecular weight glutenin proteins. The clones were sequenced by the method of Maxam and Gilbert which allowed subsequent determination of their amino acid sequences. The results of sequencing clones encoding an $\alpha$ gliadin protein were largely in agreement with similarly produced amino acid sequences for these proteins reported by Kasarda and colleagues. These studies have permitted the full 266 amino acid sequence of $\alpha$-gliadin to be determined (Table) allowing the domain structure of this molecule to be defined (Fig 6). It should, however, not be forgotten that the amino acid sequences of individual $\alpha$ gliadin proteins are not all identical. Three gliadin N-terminal amino acid sequences are recognised, $\alpha$, $\gamma$, and $\omega$, whose nomenclature was derived from the electrophoretic classification of gliadin subfractions into four groups, $\alpha$, $\beta$, $\gamma$, and $\omega$.

The inserts of well characterised cDNA clones for gliadin sequences were integrated into...
specialised expression vectors. The procedure fused the cDNA insert into the operon of a bacterial gene just downstream from a strong promoter. One of these vectors was introduced into E.coli to obtain the synthesis of the encoded protein.\(^8\) Regulatory sequence blocks between storage protein genes that determine endosperm specific expression can be detected in transgenic tobacco plants.\(^9\) This finding opens prospects of gene manipulation and transfer between divergent plant species for agricultural uses.

Detection and assessment of toxicity of cereal fractions
Biochemical and immunological methods to detect cereal proteins have been reviewed.\(^11\)\(^13\) Sensitive assays to detect cereal proteins become increasingly important since strict adherence to a gluten free diet seems to protect from the development of malignancy.\(^12\) Certain foods previously thought to be gluten free may need to be reassessed. Monoclonal and polyclonal antibody based assays have recently shown that beer and malt contain immunoreactive barley hordein, implying that patients with coeliac disease should not ingest beer or foods that contain malt.\(^13\)

Evidence implicates T cell mechanisms in the pathogenesis of coeliac disease. It is important to remember that the epitopes recognised by monoclonal antibodies raised against whole gliadin may differ from those that are potentially recognised by small intestinal T cells, after processing and presentation of cereal peptide antigens, in vivo. As Howdle and Losowsky emphasise,\(^1\) there is a need to be able to detect accurately the putative toxic cereal peptide and not just immunoreactive gliadin. The production of such assay systems is dependent on the identification of that putative peptide toxic.

In vitro organ culture, leucocyte migration inhibition factor assays, and in vivo challenge studies have been used to assess the coeliac toxicity of cereal fractions. In vivo jejunal challenge remains the ‘gold standard’ test of toxicity, but Loft and colleagues have recently extended studies on rectal challenge, which has the potential to be a simpler and less invasive means of studying cereal toxicity in coeliac disease.\(^1\)\(^1\)\(^4\)

Leigh et al.\(^16\) studied the effects of an oral challenge of 100–1500 mg of Frazer’s Fraction III (FFIII) in patients with treated coeliac disease. There were appreciable changes in the jejunal epithelial lymphocyte population but no changes in morphology, with 500, 1000, and 1500 mg of FFIII. We investigated dose and time responses with unfraccionated gliadin, given intraduodenally, and found that 10 mg produced no change, 100 mg minimal change, 500 mg moderate change, and 1000 mg extensive damage to small intestinal morphology, which began at two hours. We extended these studies to assess the toxicity of 1000 mg quantities of α, β, γ, and ω gliadin subfractions and found all to be toxic.\(^1\)\(^5\)

Using in vitro jejunal biopsy organ culture, two groups\(^18\) independently concluded that α, β, γ, and ω-fractions of wheat gliadin exacerbate coeliac disease. Weiser et al.\(^1\) suggested from the results of leucocyte migration inhibition factor assays and organ culture that their fraction B3142, which is a peptide corresponding to the N-terminal amino acids 3–56 of A-gliadin, an α-type sequenced gliadin, was coeliac toxic, although they did not include a control peptide in their assays.\(^1\)\(^6\)

Specific peptides of known amino sequence have been produced from A-gliadin by cleavage of the protein with cyanogen bromide and chymotrypsin with subsequent purification of the resultant peptides.\(^1\) Three peptides derived from cyanogen bromide cleavage spanned the complete 266 amino acid residues of A-gliadin; four peptides derived from chymotryptic digestion covered the N-terminal sequence through residue 68. Five out of the seven peptides, assessed for their coeliac toxicity using the in vitro organ culture technique, showed evidence of toxicity. The most frequently occurring
residues among the toxic peptides were pro-ser-gln-gln and gln-gln-gln-pro, although again an alternative protein control was not included in all the experiments.

The similar N-terminal amino acid sequence present in wheat, rye, and barley prolamins suggests that a certain amino acid sequence may constitute the toxic determinant. Detoxification of wheat gluten proteins by deamidation implies that glutamine, which constitutes more than 35% of the amino acids in gluten, may be involved, although the concentrated acid used may have altered other amino acid residues.

Kagnost and colleagues reported that a sequence with A-gliadin shares eight amino acids in a span of 12 and an identical pentapeptide with the 54 kDa E1b protein of human adenovirus 12 (Ad12), an adenosivirus usually isolated from the gastrointestinal tract. An antibody that reacted with the Ad12 E1b protein cross reacted with A-gliadin, a 119 amino acid cyanoen bromide peptide fragment of A-gliadin and a synthetic heptapeptide of A-gliadin from within the region of sequence homology. They suggested that, in genetically susceptible people, an encounter of the immune system with this antigenic determinant, associated with an intestinal viral infection, might be an important initiating step in the pathogenesis of coeliac disease. Cell mediated immunity to this same 12 amino acid peptide was studied by an indirect leucocyte migration inhibition assay and peripheral blood transformation assay. Patients with coeliac disease on a gluten free diet had a significantly greater response than healthy subjects with the leucocyte migration inhibition test. The cellular immune response was dependent on antigen concentration and was not present in untreated coeliac patients, but a control peptide was not included. Subsequent studies have been unable to find a correlation between coeliac disease and antibody titres to the E1b protein of Ad12 in serum from patients with untreated coeliac disease or evidence of persistent adenosivirus infection. Jewell and colleagues have recently undertaken feeding studies in treated coeliac patients using this peptide, produced by solid phase synthesis, and found damage to the small intestinal villous architecture which did not occur in control subjects (D Jewell, personal communication).

The adenosivirus hypothesis is most attractive and has obvious parallels in theories regarding the pathogenesis of autoimmune disease. Until the results of toxicity studies using purified amino acid sequenced gliadin peptides are available, however, the verdict must remain that of 'unproved'.

Conclusion

It is 40 years since Dickey reported the toxicity of wheat gluten to patients with coeliac disease. Subsequent investigators have purified wheat fractions, enabling them to show the toxicity of gliadin subfractions. Molecular biology has facilitated the sequencing of these proteins. Full characterisation of the toxic fraction is now possible, opening up the exciting prospect of developing cereals which are non-toxic to coeliac patients, but which retain the baking qualities of wheat.

PJG is a Wellcome Senior Research Fellow in Clinical Science. The authors wish to thank the Wellcome Trust, UK Coeliac Trust, Ministry of Agriculture, Fisheries and Food, St Thomas’s Hospital Research (Endowments) Committee, the American Institute for Cancer Research, Nurioca UK, and AB Juvel, Stockholm, for financial support.

25. Bartels D, Thompson RD. The characterisation of cDNA


28 Kasarda DD, Okita TW, Bernardin JE. Nucleic acid (cDNA) and amino acid sequences of alpha type gliadins from wheat (*Triticum aestivum*). *Proc Natl Acad Sci USA* 1984; 87: 4712–6.


