COELIAC DISEASE SAFE GLUTEN

The challenge to reduce toxicity while preserving wheat technological properties

Teun W.J.M. van Herpen

Promotoren:	**Prof. dr. R.J. Hamer** Hoogleraar in de Technologie van Graaneiwitten, leerstoelgroep levensmiddelenchemie, Wageningen Universiteit
	Prof. dr. H.J. Bosch Hoogleraar Enzymatische modificaties van secretoire eiwitten, Universiteit Utrecht
Co-promotoren:	**Dr. M.J.M. Smulders** Senior onderzoeker, Plant Research International, Wageningen UR
	Dr. L.J.W.J. Gilissen Senior onderzoeker, Plant Research International, Wageningen UR
Promotiecomissie:	**Prof. dr. ir. P.C. Struik** Wageningen Universiteit
	Prof. dr. H.F.J. Savelkoul Wageningen Universiteit
	Prof. dr. ir. M.A.J.S. van Boekel Wageningen Universiteit
	Prof. D. Lafiandra University of Tuscia, Italy

Dit onderzoek is uitgevoerd binnen de onderzoeksschool VLAG (Voeding, Levensmiddelentechnologie, Agrobiotechnologie en Gezondheid)

COELIAC DISEASE SAFE GLUTEN

The challenge to reduce toxicity while preserving wheat technological properties

Teun W.J.M. van Herpen

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof dr. M.J. Kropff
in het openbaar te verdedigen
op vrijdag 23 mei 2008
des namiddags te half twee in de Aula

Coeliac Disease Safe Gluten - The challenge to reduce toxicity while preserving wheat technological properties

Teun W.J.M. van Herpen

PhD thesis, Wageningen University, The Netherlands, 2008
With references – With summary in English and Dutch

ISBN: 978-90-8504-882-4

Contents

List of abbreviations

amiRNAs	artificial microRNAs
APC	Antigen Presenting Cells
bp	base pairs
BWPR	Band Width at Peak Resistance
CD	Celiac Disease
CS	Chinese Spring
CSLM	Confocal Laser Scanning Microscopy
DAF	Days After Flowering
DDT	Dough Development Time
ER	Endoplasmic Reticulum
FITC	Fluorescent Protein Label
FPLC	Fast Protein Liquid Chromatography
GMP	Glutenin MacroPolymer
GUS	Beta-glucuronidase
HMW	High Molecular Weight
HPLC	High Performance Liquid Chromatography
LMW	Low Molecular Weight
mAbs	Monoclonal antibodies
PAGE	PolyAcrylamide Gel Electrophoresis
PBs	Protein Bodies
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Saline with Tween
RNAi	RNA interference
SDS	Sodium Dodecyl Sulphate
SE	Size Exclusion
shRNA	Short Hairpin ribonucleic acid
siRNA	Short Interference ribonucleic acid

Chapter 1

General Introduction

Celiac disease (CD) is a frequently occurring food intolerance causing inflammation reactions in the small intestine and a range of related symptoms after ingestion of gluten protein from wheat and homologous proteins from rye and barley. The only treatment is a life-long gluten-free diet. Such a diet poses several social disadvantages for the patient which makes CD also a food problem. While gluten-free food products are available on the market in increasing numbers, the compliance to a strict gluten free diet is rather low among CD patients. Mariani *et al*. [1998] found a compliance rate of 53%. These authors also found that a strict gluten free diet may be a nutritional risk, because it leads to incorrect nutritional choices [Mariani *et al.*, 1998]. In addition, gluten-free bread is relatively expensive and not easily available. An other problem of a gluten-free diet is that gluten are very hard to avoid. Wheat gluten proteins have unique food-technological properties and are therefore applied in obvious products like bread, pasta and cookies, but also and more unexpectedly in a many unrelated food products such as battered snacks, many meat products, ketchups, bouillon cubes, and instant soups. Even medicines and alcoholic beverages like beers may contain gluten. Furthermore avoiding gluten is done after people are diagnosed to have CD.

As a consequence, the development of wheat varieties containing non-toxic or less toxic gluten while maintaining the unique technological properties would be highly useful. The aim of this thesis is to explore the feasibility of various ways to achieve this goal. Gluten genes from wheat have been studied extensively for their technological quality in the bread making process. However it is not known to what extent it is possible to remove toxicity while preserving technological properties.

In this introduction, first the mechanism of CD will be elaborated. Then an overview will be given of the existing knowledge on wheat with regard to its gluten genes and proteins. CD-toxic epitopes are identified in these gluten proteins and their relation to CD is discussed.

Following, we will introduce strategies to reduce these CD-toxic gluten proteins in order to develop CD-safe wheat. However, gluten proteins have essential technological characteristics that are important to be retained in strategies to develop CD-safe and technological interesting wheat. The unique technological properties of wheat gluten will be discussed in the fourth paragraph. In the last paragraph an outline of this thesis is given.

Celiac disease

In 1950, WK Dicke [1950] discovered that proteins especially from wheat, rye and barley are the main cause in the development of celiac disease. The responsible proteins in wheat were found to be the gluten proteins, in rye the secalins, and in barley the hordeins. Today estimations are that about 1% of the population suffers from celiac disease. Fasano and Catassi [2001] suggested that the ratio of known to undiagnosed cases of CD was 1 in 7. Major symptoms are chronic diarrhoea, osteoporosis, lymphoma and fatigue. After consumption of gluten proteins they are broken down into peptides, some of which have the capacity to invoke an immune response. The surface of a healthy small intestine is covered with villi that function in the uptake of nutrients (Figure 1). The villi contain enterocytes (a type of epithelial cells) that produce enzymes like lactase, trypsin and chymotrypsin that are important in the digestion. When a celiac patient consumes gluten, an immune-mediated response takes place against several of these peptide residues resulting in local inflammations leading to flattening of the mucosa of the small intestine. Generally, this flattening is reversible: after removal of gluten from the diet the villi will recover. Basically, the immune system can respond in two ways to gluten peptides, i.e. through the adaptive and through the innate immune system. Details of both systems will be given in the next two sub-paragraphs.

Until today, a complete and life long elimination of gluten from the daily diet is the only effective treatment strategy. However, this places a considerable burden upon patients because an increasing number of regular food products contain gluten and because many gluten-free food products on the market are not truly appreciated by the patients. Therefore, wheat with gluten that is low or not CD-toxic would be of great benefit to CD-patients. A cultivar low in CD-toxicity can

possibly be tolerated by CD patients. Oats for example, seem to be tolerated by most celiac disease patients [Janatuinen *et al.,* 2002] although oats contains few sequences that can be recognized by T-cells of certain CD patients [Vader *et al.,* 2003]. Since a relation exist between population incidence of CD patients and gluten exposure [Ivarsson *et al.,* 2000; Fasano, 2006; Ventura *et al.*, 1999] therefore prevention of CD based on a diet based on wheat cultivars with little CD-toxicity seems to be a promising option.

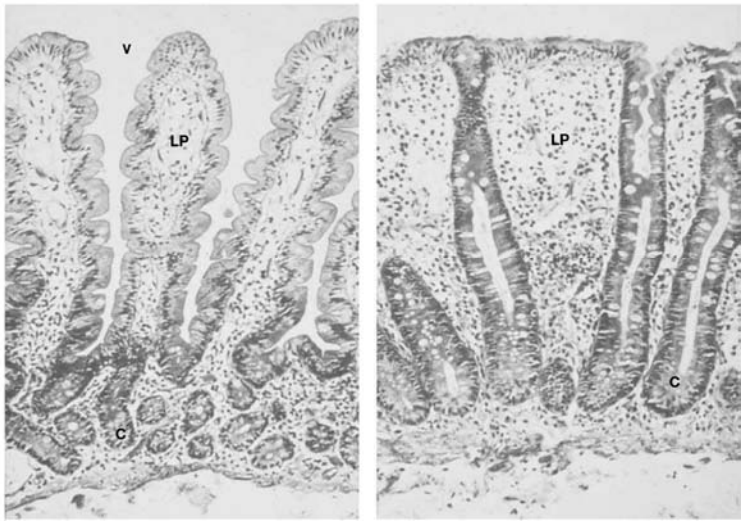

Figure 1 - The lefthand figure shows the mucosa of a healthy individual and the righthand figure shows a damaged flat and thickened mucosa of a celiac patient, showing villi (V), crypts (C) lamina propria (LP) [from Sollid 2002].

Adaptive immune response against gluten

In the small intestine some native gluten peptides can bind directly to particular receptors called HLA-DQ2 or HLA-DQ8 that are present on antigen presenting cells (APC). These APC cells can present the gluten peptide to gluten sensitive T-cells and thus activate [Koning, 2003; Vader *et al.,* 2002b].

When examining the T-cell reactive gluten peptides more closely, it was found that they are rich in the neutral amino acid glutamine. However, the MHC receptors of the APCs have a preference for negatively charged amino acids. Tissue transglutaminase (tTG) is an enzyme that is present in the intestinal wall and is normally involved in tissue repair. tTG can deamidate

glutamine into glutamate, which considerably increases the affinity of the peptide to the HLA-receptors (Figure 2). When the gluten peptides are bound to the APCs, gluten sensitive T-lymphocytes can be activated. The T-lymphocytes release certain cytokines after activation. This will cause an inflammation reaction leading to damage of the intestinal villi. The APC-gluten-peptide-T-lymphocyte complex simultaneously cause an increase in the immune response by attracting more gluten sensitive T-lymphocytes. This will result in an increased production of cytokines and tTG, subsequently followed by an amplification of the cascade due to increased deamidation of peptides resulting in further tissue damage. In this process, a glutamine from one protein is covalently bound to a lysine of another protein. However, after removal of all gluten from the small intestine, the tissue is able to recover and the vicious circle will stop.

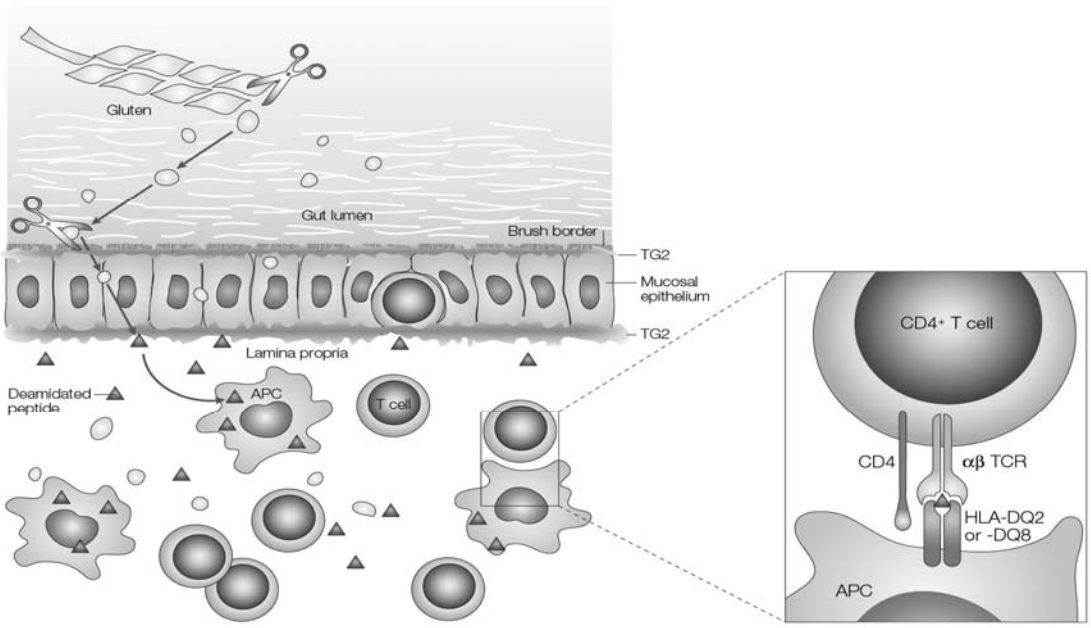

Figure 2 - The activation of the immune system by immuno-toxic gluten. 1. The gluten proteins arrive in the small intestine. 2 Proteolytic enzymes digest the gluten into smaller peptides which are taken up by the small intestine. Gluten resistant to proteolytic digestion can be taken up as polypeptide. 3 The tTG deaminates the gluten which can bind with high affinity to the HLA-DQ2 or DQ8 receptor of the APCs. 4 This complex is subsequently bound to the CD4+ T lymphocyte which results in activation of the immune system. Enlarged is the complex that is formed after binding of the APC-gluten with CD4+-T-lymphocyt [illustration from Sollid 2002].

The innate immune system in relation with gluten in Celiac disease

The development of an adaptive immune response is strongly controlled by innate immunity. Without molecular signals provided by intestinal dendritic cells, no gluten-specific T-cell responses will develop (Figure 2). It was shown that gliadin is capable of inducing the maturation of monocyte-derived dendritic cells [Palova-Jelinkova *et al.,* 2005]. Studies using tissue cultured cells showed that the α-gliadin-derived fragment p31–43 (see Table 2) can induce IL-15 secretion by activated intestinal dendritic cells and possibly other antigen presenting cells. This p31-43 peptide, in its native and deamidated form, is not binding to the APC cells. IL-15 is a potent stimulant of intraepithelial lymphocytes (IELs) [Ebert 1998]. In vitro, this has been shown to result in targeted cell killing [Meresse *et al.,* 2004; Hue *et al.,* 2004]. An increase in the number of IELs throughout the small-intestine as observed using biopsies of CD patients is one specific markers of this disease [Hoper *et al.,* 2006]. These new results now indicate that IL-15 secretion can lead to epithelial cell destruction by IELs.

Wheat

Cereals belong to the family of the *Gramineae,* also called *Poaceae* or grass family (Figure 3) in which various subfamilies and tribes are distinguished. Wheat (*Triticum* spp.), rye (*Secale cereale*) and barley (*Hordeum vulgare*) are closely related and are classified in the same tribe, called *Triticeae* or *Hordeae*. Other cereals like various species of millet (*Panicum spp.* and *Eleusine coracana*), sorghum (*Sorghum bicolor*), maize (*Zea mays*), and Job's Tears (*Coix lacryma-jobi*) belong to other subfamilies. Oat (*Avena* spp.), rice (*Oryza sativa*) and teff (*Eragrostis tef*) are classified in other tribes. Buckwheat, amaranth and quinoa are often mentioned regarding gluten-free diets, but are pseudocereal dicots and are much more distant to wheat.

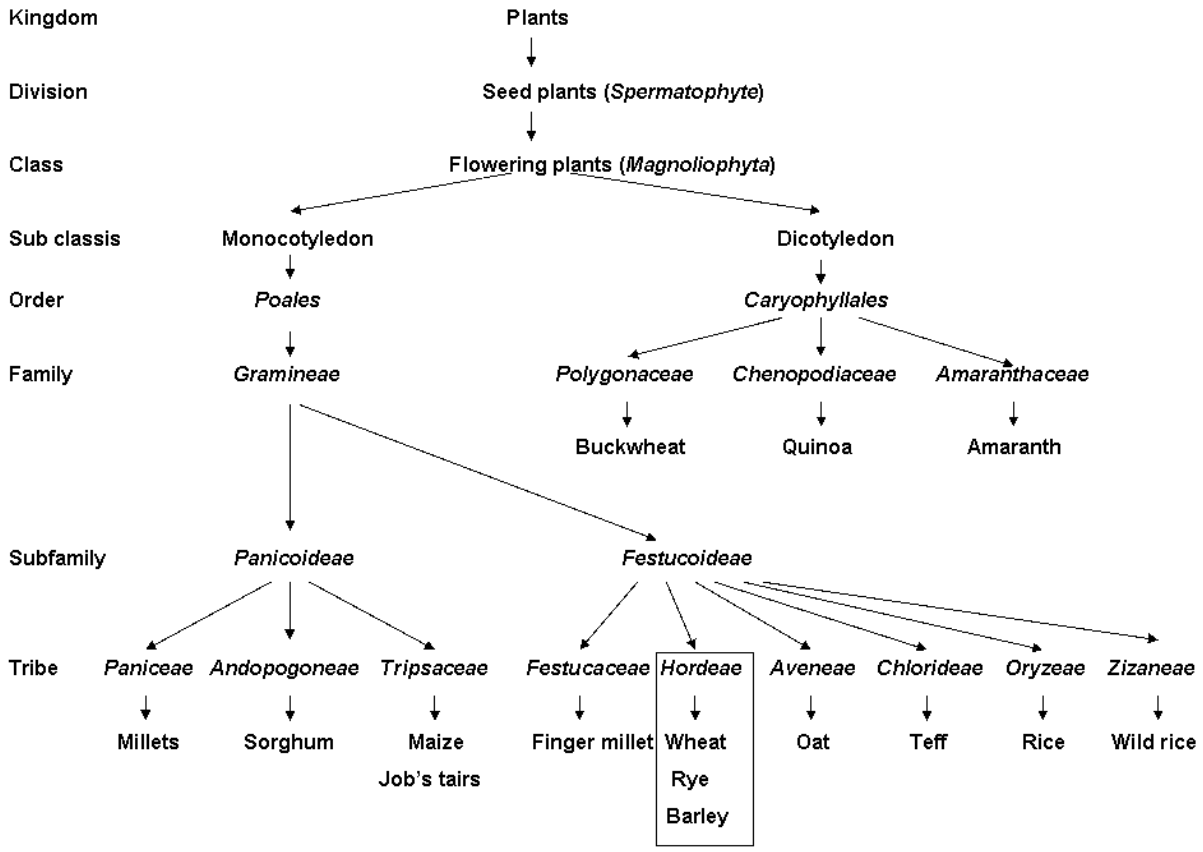

Figure 3 - Taxonomic relationships of grains and some other staple foods [after Thompson, 2000].

Hexaploid *Triticum aestivum* or bread wheat originated around 8,000 years ago from a hybridization of a tetraploid *Triticum* species with the diploid donor of the D genome *T. tauschii* [Feldman *et al.,* 1995] as depicted in figure 4. The A and B genomes were most likely provided by *T. dicoccoides* (AABB). Estimates for the age of the *T. dicoccoides* origin range from 250,000 to 1,300,000 years ago [Huang *et al.,* 2002; Mori *et al.,* 1995]. *T. dicoccoides* is formed from the wild diploid *T. monococcum* or *T. urartu* (A genome) and the donor of the B genome [Kilian *et al.,* 2007; Feldman *et al.,* 1995]. Morphological, geographical and cytological evidence suggests *T. speltoides* (S genome) (also known as *Aegilops speltoides*) or a closely related species as the B genome ancestor. Genetic research using AFLP markers also indicate an origin of the B genome from *T. speltoides* [Kilian *et al.,* 2007]. According to Isidore *et al.* [2005] polyploidization had

enabled intergenic hybridizations. In these hybrids, the genomes in the cellular nucleus remained separate, with no occurrence of recombinations between the homoeologous chromosomes.

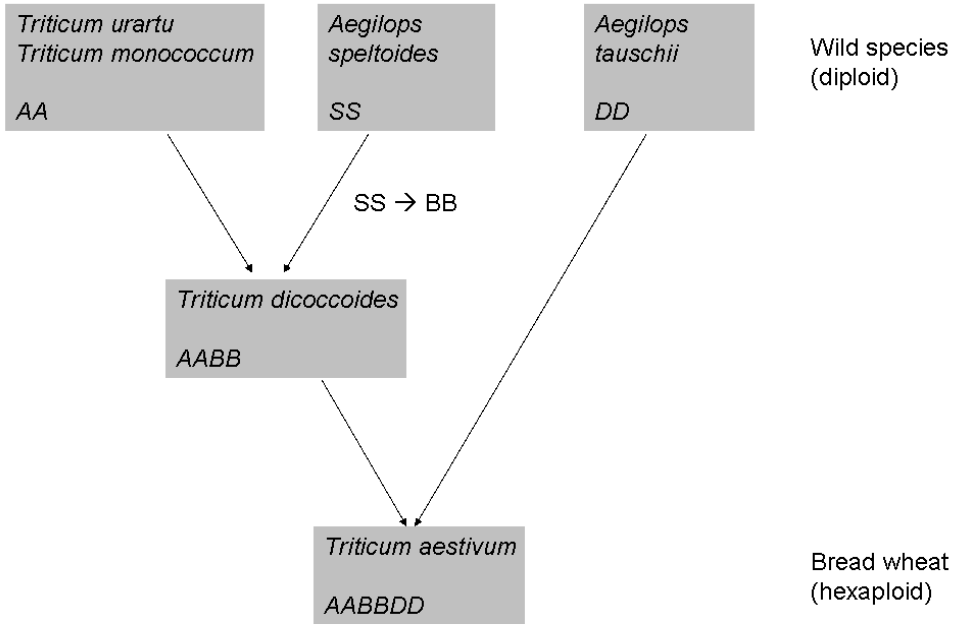

Figure 4 - Evolution of wheat

Wheat gluten proteins

Wheat is an important staple food because of its high nutritional characteristics, technological properties and long shelf life. In the Netherlands the per capita wheat consumption corresponds to a mean daily gluten intake of 13.1 g [Van Overbeek *et al.*, 1997]. The wheat kernel contains 8-17% of protein from which around 15% is albumin/globulin and around 85% is gluten. Wheat gluten can be classified into three large groups: sulphur-rich (S-rich, with a molecular weight [MW] of ~50 kD), sulphur-poor (S-poor, MW ~50 kD) and high molecular weight (HMW, MW ~100kD) glutenins, with a number of subgroups within the S-rich and S-poor group (Figure 5). This classification does not correspond directly to the polymeric and monomeric fractions in the wheat kernel (glutenins and gliadins, respectively). The S-poor group consists of ω-gliadins and the D-type LMW glutenin subunits (LMW-GS). These are encoded on the short arm of

chromosome 1A, 1B and 1D. For clarity, LMW-type does not refer to the chromosome location but is refer to structural differences between LMW-GS. The S-rich group consist of three major families: the B-type LMW-GS, the γ- and α-gliadins. These genes are located as multi gene families on the short arms of chromosomes 1 and 6 (A, B and D). The cysteine residues in the γ-gliadin and α-gliadins form intra-chain disulfide bonds (Table 1). The B-type LMW-GS also form intra-chain disulfide bonds and subsequently possess one or more additional cysteine residues, which may form inter-chain disulfide bonds (with cysteine residues present in other subunits) [Kohler *et al.*, 1993]. The C-type LMW-GS appear to comprise a mixture of α- and γ-gliadins [Shewry and Tatham, 1997] but posses one cysteine residue that can form inter-chain disulfide bonds. HMW-GS are located on the long chromosome 1A, 1B and 1D, and each of these chromosomes can encode for two different HMW-GS, one x-type and one y-type HMW-GS. These HMW-GS contain two or more cysteine residues that can form inter-chain bonds disulfide bonds [Table 1; Shewry and Tatham 1997].

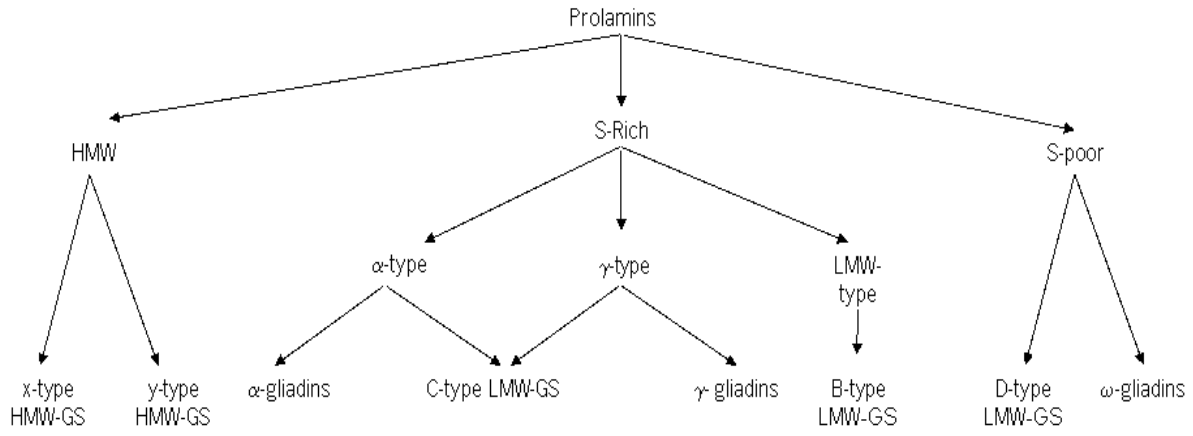

Figure 5 - The classification of wheat gluten proteins based on amino acid sequence [after Shewry and Lookhart, 2003]

Table 1 - Summary of the types and characteristics of wheat grain prolamins (gluten proteins)

Prolamin class	**Size (kD)**	**% of total**	**S-resi-dues**	**Intra-chain bonds**	**Inter-chain bonds**	**Chromosome location**
HMW glutenins x-type	65-90	6-10	2-5	0-1	2-3	Long arm of 1 (ABD)
HMW glutenins y-type	65-90		6-7	0-2	3>	Long arm of 1 (ABD)
α -gliadins	30-45	70-80	6	3	0	Short arm of 6 (ABD)
C-type LMW glutenins	30-45		7-8	3-4	1	Short arm of 1+6 (ABD)
γ -gliadins	30-45		8	4	0	Short arm of 1+6 (ABD)
B-type LMW glutenins	40-50		7-8	3	1-2	Short arm of 1 (ABD)
D-type LMW glutenins	30-75	10-20	1	0	1	Short arm of 1 (ABD)
ω -gliadins	30-75		0	0	0	Short arm of 1 (ABD)

Epitopes in relation to CD

Gluten are rich in the amino acids proline and glutamine. As a consequence gluten proteins have few trypsin cleavage sites, which limit the breakdown of these proteins. Multiple T cell epitope motifs have been identified in the α- and γ-gliadins as well as in the glutenins [Sollid, 2002; Arentz-Hansen *et al.,* 2000; Vader *et al.,* 2002b; Koning, 2003; see Table 2], the majority of which require deamidation for T cell recognition. Vader *et al.* [2003] identified 11 homologous T-cell stimulatory sequences in hordeins and secalins, located in similar regions of the proteins. Seven of these 11 peptides were recognized by gluten-specific T-cell lines from CD-patients. These results showed that the disease-inducing properties of barley and rye can, at least in part, be explained by T-cell cross-reactivity against gluten-, secalin-, and hordein-derived peptides. The results also show that oats contains sequences that can be recognized by T-cells of certain celiac disease patients. In that respect it is intriguing that oats can be tolerated by most celiac disease patients [Janatuinen *et al.,* 2002].

Peptides derived from α-gliadins are recognized by T cells from almost all celiac patients, whereas T-cell responses to γ-gliadins and glutenins are much less frequent [Arentz-Hansen *et al.,* 2000; 2002; Vader *et al.,* 2002b; 2003; Molberg *et al.,* 2003]. The α-gliadin proteins contain a stable 33-mer fragment that contains a cluster of these epitopes [Shan *et al.,* 2002]. This 33-mer

fragment is naturally formed by digestion with gastric and pancreatic enzymes. The 33-mer fragment contains the peptides glia-α2 and glia-α9 (Table 2).

Table 2 - Amino acid sequence of T cell stimulatory gluten peptides [after Koning, 2003]

Peptide	**Sequence**
Glia (206–217)	SGQGSFQPSQQN
Glt (723–735)	QQGYYPTSPQQSG
Glia-γ1 (138–153)	QPQQPQQSFPQQQRPF
Glia-α2 (62–75)	PQPQLPYPQPQLPY
Glia-α9 (57–68)	QLQPFPQPQLPY
Glia-α20 (93–106)	PFRPQQPYPQPQPQ
Glt-156 (40–59)	QPPFSQQQQSPFSQ
Glt-17 (46–60)	QQPPFSQQQQQPLPQ
Glu-5	QQQXPQQPQQF
Glia-γ30 (222–236)	VQGQGIIQPQQPAQL
33-mer (57–89)	LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF *Glia-α9* *Glia-α2*
p31-43	LGQQQPFPPQQPY

α-Gliadins

In the innate immune response as well as the adaptive immune response, the α-gliadins are considered most toxic. The α-gliadin derived p31–43 fragment was found to activate the innate immune system. And in the adaptive immune response T-cells against α-gliadin peptides are most frequently found. Therefore specifically the α-gliadin protein is described in more detail here.

The α-gliadins of hexaploid *Triticum aestivum* are encoded by the genes of the *Gli-2* locus on the short arm of the group 6 chromosomes of the A, B and D genome [Marino *et al.,* 1996]. Estimates for α-gliadin copy numbers range from 25-35 copies [Harberd *et al.,* 1985] to 100 [Okita *et al.,* 1985] or even 150 copies [Anderson *et al.,* 1997] per haploid (=single hexaploid) genome. Anderson and Greene [Anderson and Greene, 1997] compared the sequences of 27

known cDNA and genomic clones of α-gliadins and concluded that about half of the latter contained "in frame" stop codons and were presumably pseudogenes. The detailed constitution of the multi-gene locus is not known.

The schematic structure of an α-type gliadin protein is depicted in Figure 6. The protein consists of a short N-terminal signal peptide (S) followed by a repetitive domain (R) and two longer non-repetitive domains (NR1 and NR2), separated by two polyglutamine repeats (Q1 and Q2). In the non-repetitive domains, five conserved cysteine residues are present which are indicated with bold vertical lines. Gluten proteins can contain different T-cell epitopes in its structure (Figure 6). The currently known T-cell epitopes are shown and their approximate position is indicated by arrows.

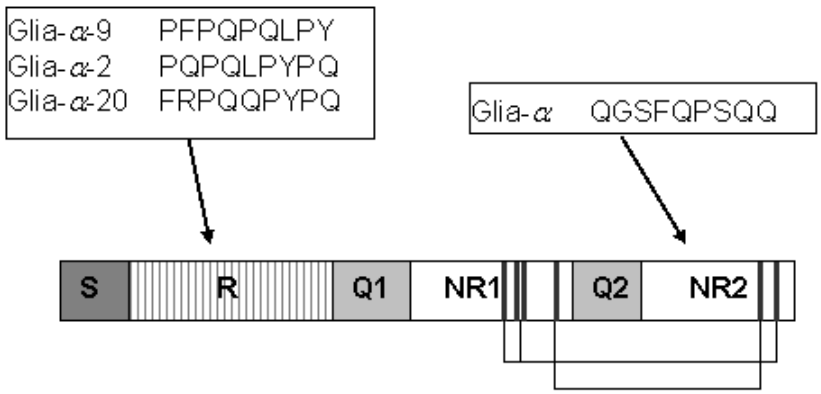

Figure 6 - Schematic representation of the structure of an α-gliadin protein. The protein consists of a short N-terminal signal peptide (S) followed by a repetitive domain (R) and a longer non-repetitive domain (NR1 and NR2), separated by two polyglutamine repeats (Q1 and Q2). In the non-repetitive domains six conserved cysteine residues are present which all form intra-chain disulphide bonds. The cysteine residues and bonds are indicated with bold lines. The T-cell epitopes are shown at their approximate position.

The development of low toxic gluten

The following three strategies to produce wheat varieties containing low-toxic gluten proteins can be considered. These strategies will be shortly introduced here and elaborated in the following sections of this paragraph. One strategy is classical breeding on existing hexaploid wheat varieties could focus on lowering the level of CD-toxic α-gliadins. Another strategy is to lower the expression of CD-toxic genes by influencing promoter activity, while the genes from less toxic genomes are enhanced in expression. Thirdly, using a RNA interference technique, we could down regulate the expression of the most CD-toxic proteins, like α-gliadins.

Breeding for low CD-toxic wheat

The validity of breeding or reconstructing bread wheat for low CD-toxicity depends on whether sufficient genetic diversity is present among varieties. Spaenij-Dekking *et al.* [2005] demonstrated that a large diversity appears to exist in the amount of T cell stimulatory epitopes present in α- and γ-gliadins, and in glutenins within and among different hexaploid *Triticum* varieties. In order to find one or more varieties with a naturally low level for all epitopes, if existing, a larger group of varieties need to be tested. Alternatively, reconstructions of hexaploid bread wheat may be used for selection of CD-safe varieties.

Promoter influencing

The D-genome of wheat contains most of the α-gliadin epitopes [Molberg *et al*., 2005; Spaenij-Dekking *et al.,* 2005]. These differences in toxicity are linked to specific differences in the α-gliadin sequences per genome of origin [Chapter 2]. Differences in expression between the three different genomes in the α-gliadin family have been observed which might point at differences in gene expression regulation mechanisms between the three genomes [Kawaura *et al.,* 2005]. By identifying the transcription factors responsible for these different expression patterns between the genomes a strategy can be developed to influence the expression of these different gene loci [Aharoni *et al.,* 2001]. Down-regulation of α-gliadin gene expression specifically of the D-genome might be a new strategy to reduce the CD-toxicity of wheat.

RNAi

Another strategy to eliminate the production of CD-toxic gluten proteins involves genetic modification. Classical RNAi has been shown to be effective in silencing all α-gliadins in bread wheat [Becher *et al.,* 2006]. This approach can, however, be carried out more specifically and be directed to only those genes carrying harmful epitope sequences to maintain the specific and unique technological properties as much as possible. Not all α-gliadins contain toxic T-cell epitopes [Chapter 2, this thesis].

In 1999, siRNAs were first discovered as part of post-transcriptional gene silencing in plants [Hamilton and Baulcombe, 1999]. Shortly thereafter, synthetic siRNAs were shown to be able to induce RNAi in mammalian cells [Elbashir *et al.,* 2001]. Miller *et al*. [2003] showed two years later that in mammalian cell models allele-specific silencing of disease genes was achieved by targeting a linked SNP using siRNA. The presence or absence of the DQ-8 Glia-α T-cell epitope is also linked to an allele-specific SNP in the α-gliadins gene sequence [Chapter 2, this thesis]. It is expected that a similar approach towards this specific α-gliadins epitope will result in specific silencing of those gliadin genes carrying that epitope. This strategy might reduce CD-toxicity while retaining other gliadins and relevant technological properties.

Technological properties

Gluten protein networks are among the most complex and large protein networks in nature due to numerous different components and different sizes, and due to variability caused by genotype, growing conditions and technological processes [Wieser, 2007].

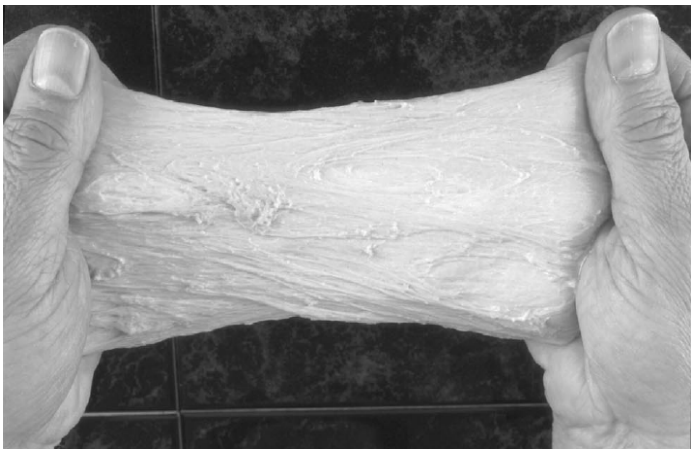

Figure 7 - Gluten can form a viscoelastic protein mass [after Shewry et al., 2002b]

Both the gliadins and the glutenins are important components with different functions in determining the rheological properties of dough (Figure 7). Purified hydrated gliadins have little elasticity and are less cohesive than glutenins. They mainly contribute to the viscosity and extensibility of the dough system. In contrast, hydrated glutenins are both cohesive and elastic, and are responsible for dough strength and elasticity. As a figure of speech, gluten can be seen as a "two-component glue", in which gliadins can be understood as a "plasticizer" or "solvent" for the glutenins. A correct mixture of the two is crucial to obtain the viscoelastic properties of the dough and the quality of the end product [Wieser *et al.*, 2007].

Synthesis and deposition of wheat seed storage proteins

Wheat seed storage proteins are produced in immature grains and deposited in spherical-shaped discrete protein bodies [Shewry and Halford 2002a]. After synthesis in the ER, some prolamins, principally gliadins, are transported via the Golgi to the protein storage vacuole, whereas other, principally glutenins, are retained within the ER to form ER-derived protein bodies [Kim *et al.*, 1988; Rubin *et al.*, 1992]. It is suggested that ER-derived protein bodies are subsequently taken up by protein storage vacuoles in a process analogous to autophagy [Shewry and Halford, 2002a; Galili, 1997; Shewry, 1999]. The protein bodies in developing grains are accompanied by smaller

dark-staining particles of the globulin storage protein triticin [Bechel *et al.*, 1991], which is presumably transported via the Golgi to the vascular protein bodies.

The hyper-aggregation model

In the hyper-aggregation model [Hamer and Van Vliet, 2000], the formation of the glutenin network is proposed as network in which covalent and non-covalent processes are involved (Figure 8). This model includes gluten proteins as well as starch and arabinoxylans.

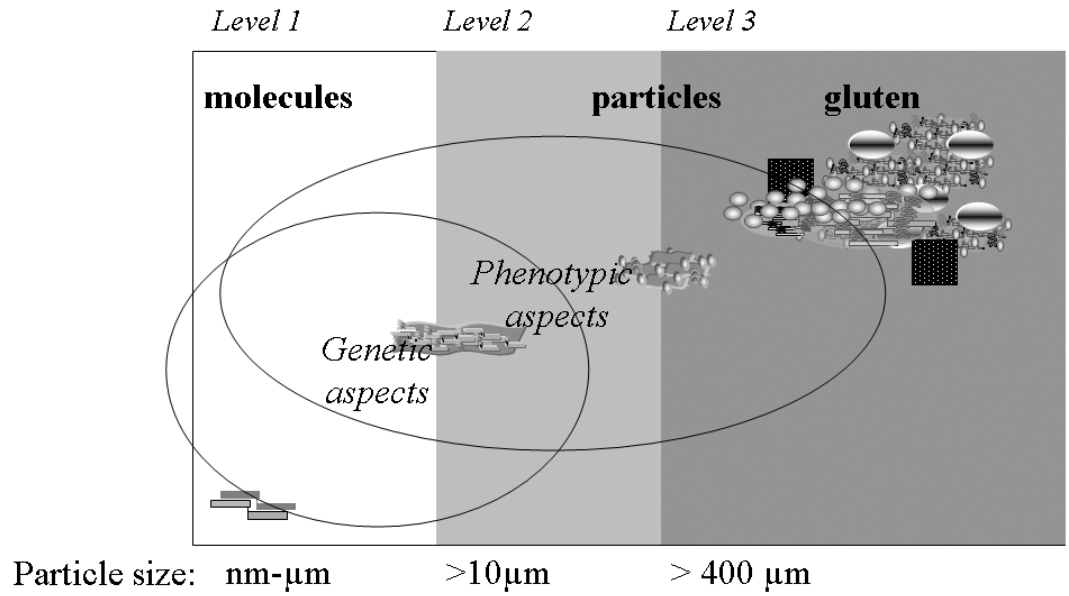

Figure 8 - A schematic representation of the hyper-aggregation model. The levels refer to different length scales. Genetic aspects refer to the molecular characteristics (amino acid sequence) of the participating gluten proteins and the formation of covalent bonds between these proteins. Phenotypic aspects (heat drought) refer to physico-chemical interactions including the formation of hydrogen bonds and electrochemical interactions.

At Level 1 of the hyper-aggregation model, HMW-glutenins and LMW-glutenins form covalently linked polymers (Figure 8). This step is determined by the present individual glutenin subunits and their ability to form and terminate the network. At Level 2, the covalently stabilized glutenin

polymers form larger aggregates by physical interactions: hydrogen bonds, and electrostatic and hydrophobic interactions. The glutenin subunit composition of the polymers formed at Level 1 will largely determine the size of the aggregates at Level 2. The aggregates at Level 2 will play a major role in the dough properties during further mixing and resting. At Level 3 further aggregation occurs by physical interactions only. Here interactions with non-protein constituents such as starch particles and arabinoxylans come into play as well.

Most modifications of wheat to obtain less toxic bread wheat are expected to have its influence on technological properties at Level 1 of the hyper-aggregation model.

Glutenin macropolymer (GMP)

The largest polymers in wheat flour are called “unextractable polymeric protein” (UPP) or “glutenin macropolymer” (GMP). GMP is the polymeric fraction which remains insoluble in various solvents (SDS or acetic acid) [Graveland *et al.*, 1982; Weegels *et al.*, 1996, 1997]. Differences in technological properties among flours parallel differences in the quantity of the fraction of GMP in flour. During dough mixing, the content of GMP decreases and after resting of the dough this content will increase again [Weegels *et al.*, 1997; Aussenac *et al.*, 2001]. It was shown that GMP consists of spherical glutenin particles [Don *et al.*, 2003a], which can vary in average size. Genetic background and growth conditions affect GMP quantity, GMP particle size and, consequently, flour quality [Don *et al.*, 2005a]. The quantity of GMP depends particular on the ratio of HMW-GS to LMW-GS [Don *et al.*, 2003c] and the types of individual glutenin subunits, with the HMW-GS are most important [Gupta *et al.*, 1993]. Studies using near-isogenic lines grown under different heat stress conditions have revealed that the HMW- to LMW-GS ratio is an important determinant of glutenin particle size, where lower ratios gave larger particles [Don *et al.*, 2005a].

The accumulation of wheat storage proteins is a continuous process commencing as early as 7 days after flowering (DAF) and only stopping at the desiccation phase, the last phase of grain maturation. During dessication, a close correlation was found between the accumulation of the GMP and the rapid loss of water [Carceller and Aussenac, 1999, 2001].

Disulphide bonds

As shown before disulphide bonds play a key role in determining the structure and properties of gluten protein [Grosch and Wieser, 1999]. These bonds are important in stabilising the conformation of proteins or protein aggregates and determine the size of the glutenin polymers. Disulphide bonds are formed between the sulphydryl group of the cysteine units, either within a single protein (intrachain) or with another protein (interchain). Some cysteine units remain as free thiol [Kasarda 1999]. Intrachain disulphide bond formation already starts after synthesis of protein within the lumen of the endoplasmatic reticulum as a part of protein folding [Shewry, 1999]. After residing in the protein bodies, glutenin undergo redox changes during the development and maturation of the grain. Free thiol groups become oxidised during the grain desiccation phase which coincides with the formation of high-MW polymers (GMP) [Carceller and Aussenac, 1999, 2001; Razi *et al.*, 2003].

Most α- and γ-gliadins show intra-chain cysteine bonds where specific residues within the protein bind together (Figure 9, Table 1). LMW glutenin has similar intra-chain cysteine bridges as the α- and γ-gliadins, and also has two free cysteine residues (Figure 9; Table 1). One residue can bind to other LMW subunits and the other can bind to other LMW subunits and y-type-HMW subunits and γ-gliadins. γ-Gliadins having an odd number of cysteine residues (also called C-type glutenins) might act as terminator of polymerisation, whereas the HMW-GS and B-type LMW-GS, with more than one free cysteine residue, can act as a chain extender. Besides γ-gliadins, also α-gliadins (also called C-type LMW glutenins) have been detected as terminators in a purified glutenin fraction [Lew *et al.*, 1992].

N-terminal sequencing of isolated GMP revealed that C-terminal parts of an x-type HMW-GS is linked to the N terminal domain of a y-type HMW-GS suggesting a specific head to tail orientation of HMW-glutenins [Tao *et al.*, 1992]. Three additional cysteine residues in the y-type HMW glutenins were found to form interchain cysteine bonds. Two of these cysteine residues were found to bind other y-type HMW glutenins and one cystein residue could bind to a LMW glutenin (Figure 9) [Grosch and Wieser, 1999].

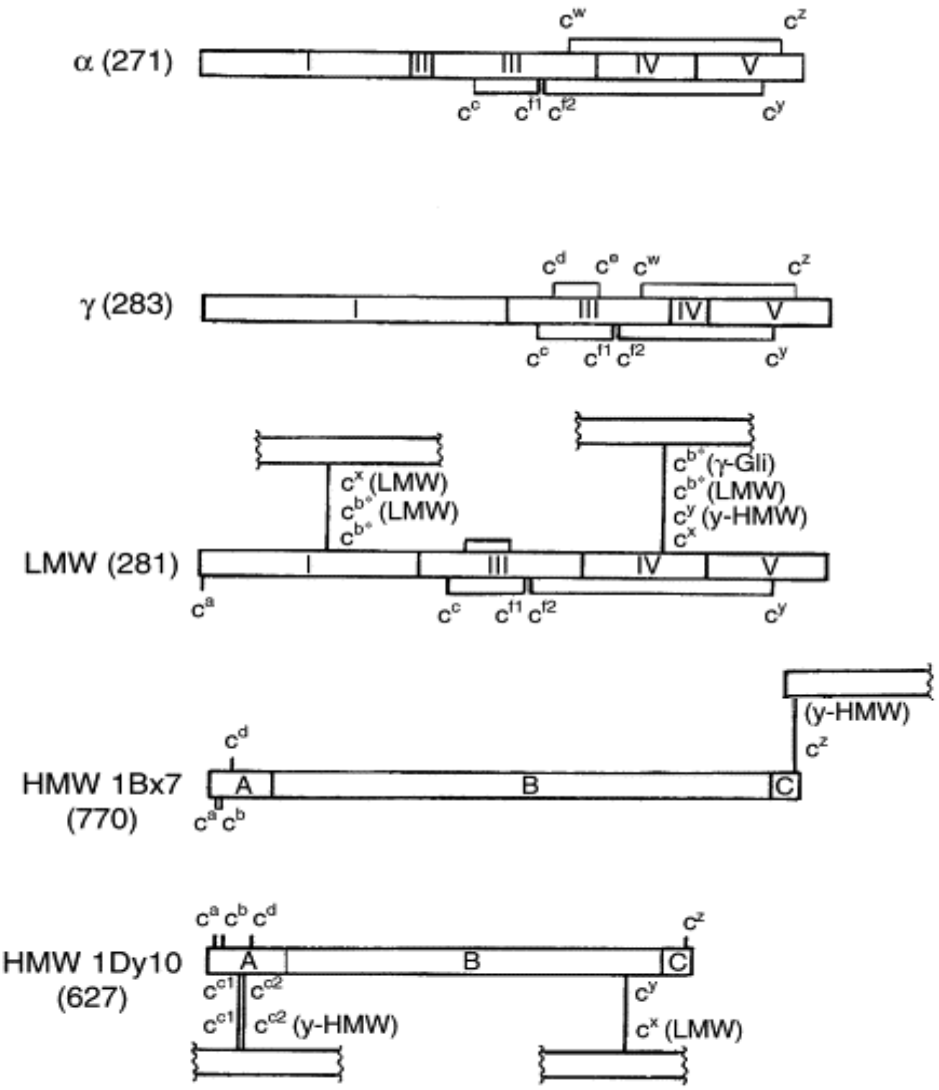

Figure 9 - Position and polymerization of cysteine residues' in α- and γ-gliadins, LMW, HMW 1Bx7 and 1Dy10-GS. Most α- and γ-gliadins show internal cysteine bonds. LMW glutenin has similar internal cysteine bonds but also has two free cysteine residues (Cb and Cx). The Cb residue was bound to other LMW subunits and the Cx residue could be bound to other LMW subunits, y-type-HMW subunits and γ-gliadins. The C-terminal parts of x-type HMW-GS was linked to the N terminal domain of y-type HMW-GS. Three cysteine subunits of the y-type HMW glutenin were bound to other y-type HMW glutenins (Ce1 and Cc2) or to LMW glutenin (Cy) [figure adopted from Grosch and Wieser, 1999].

A range of models have been developed to explain the structure of the glutenin network. Graveland *et al.* [1985] proposed a model in which x-type and y-type HMW glutenins form a linear head-to-tail backbone where the LMW-glutenins branch off from the y-type HMW-glutenins. Apparently, HMW-glutenins and LMW-glutenins polymerize separately. The mechanism of aggregation of HMW-glutenins and LMW-glutenins during synthesis, and bringing the separately polymerised polymers together at the appropriate time and place for glutenin formation during grain development is not yet known.

When manipulating the genetic background of bread wheat in an attempt to obtain less CD-toxic wheat, we can expect that changes in the HMW and LMW glutenins can change the ability to form a cohesive and elastic network. This should be considered when maintenance of the technological properties of bread wheat is aimed at.

However, the function and mechanisms involved in the "plasticizer" effect of the gliadins are still not well understood, because most research till now has focused on the characteristics of the disulphide bonds in the glutenin network which are active in the Level 1 of the hyper-aggregation model. We suggest that the "plasticizer" effect of the monomeric gliadins would be especially relevant at the Levels 2 and 3 of the hyper-aggregation model.

Outline of this thesis

Wheat varieties with low or not CD-toxic gluten would be of great benefit to CD-patients in toleration and maybe more importantly in prevention. Developing such a wheat variety is challenging, because firstly wheat contains a large number of individual gluten proteins, all differing in their contribution to CD-toxicity. Secondly removing specific CD-toxic gluten proteins can result in a loss of the unique technological properties of the gluten protein. The feasibility of different strategies to develop low CD-toxic wheat varieties will be assessed in this thesis. An overview of this thesis is given below.

Chapter 2: Alpha-gliadin genes from the A, B, and D genomes of wheat contain different sets of celiac disease epitopes

The α-gliadins are an important group of wheat storage proteins in relation to celiac disease. Epitopes derived from the α-gliadins are responsible for activating the innate- and adaptive immune response. α-Gliadins therefore are the main target when lowering the toxicity of bread wheat. The questions to be studied in chapter 2 is which α-gliadins are present in the different wheat genomes and whether sufficient genetic diversity is present among the α-gliadin family.

Chapter 3: Detailed analysis of the expression of an α-gliadin promoter and the deposition of α-gliadin in wheat grain development

One way of lowering the α-gliadins in wheat is by directly influencing the activity of α-gliadin promoter. More information about the expression can give a better prediction of the effects of lowering the α-gliadins on the wheat kernel and possibly on the technological properties. The central question in this chapter is how the α-gliadin proteins are expressed in the wheat kernel?

Chapter 4: The origin and early development of wheat glutenin particles

Changes in the technological properties of wheat resulting from genetic modifications are expected to be revealed in modifications in aggregation behaviour and the formation of small particles. These small particles called protein bodies are formed early in the development of the wheat kernel. This first question related to technological properties of wheat studied in chapter 4. The central question here is what is the relation between early development of wheat and the technological properties of the wheat gluten?

Chapter 5: The feasibility of decreasing CD toxicity while retaining technological properties: A study with Chinese Spring deletion lines

The feasibility to develop a strategy to reduce CD-toxicity in hexaploid bread wheat with a minimal effect on the technological properties of the gluten protein was studied. The following question is studied in chapter 6: What is the effect of the complete silencing of different sets of

gluten proteins on both technological properties and CD-toxicity? In combining the results of these two aspects a breeding strategy was developed.

Chapter 6: Silencing epitope-specific alpha gliadin genes using siRNA on specific SNPs
A possible strategy to remove toxic α-gliadins proteins from wheat is by silencing the expression of α-gliadins with a specific T-cell-epitope. This strategy might reduce CD-toxicity while retaining other gliadins and relevant technological properties.

Chapter 7: General discussion
Finally, in the general discussion the results from all the experimental chapters will be discussed together. From this, we will consider the viability of the different strategies to develop low toxic and technologically attractive gluten.

Chapter 2

Alpha-gliadin genes from the A, B, and D genomes of wheat contain different sets of celiac disease epitopes

Teun WJM van Herpen, Svetlana V Goryunova, Johanna van der Schoot, Makedonka Mitreva, Elma Salentijn, Oscar Vorst, Martijn F Schenk, Peter A van Veelen, Frits Koning, Loek JM van Soest, Ben Vosman, Dirk Bosch, Rob J Hamer, Luud JWJ Gilissen, Marinus JM Smulders

Abstract

Bread wheat (*Triticum aestivum*) is an important staple food. However, wheat gluten proteins cause celiac disease (CD) in 0.5 to 1% of the general population. Among these proteins, the α-gliadins contain several peptides that are associated to the disease.

We obtained 230 distinct α-gliadin gene sequences from several diploid wheat species representing the ancestral A, B, and D genomes of the hexaploid bread wheat. The large majority of these sequences (87%) contained an internal stop codon. All α-gliadin sequences could be distinguished according to the genome of origin on the basis of sequence similarity, of the average length of the polyglutamine repeats, and of the differences in the presence of four peptides that have been identified as T cell stimulatory epitopes in CD patients through binding to HLA-DQ2/8. By sequence similarity, α-gliadins from the public database of hexaploid *T. aestivum* could be assigned directly to chromosome 6A, 6B, or 6D. *T. monococcum* (A genome) sequences, as well as those from chromosome 6A of bread wheat, almost invariably contained epitope glia-α9 and glia-α20, but never the intact epitopes glia-α and glia-α2. A number of sequences from *T. speltoides*, as well as a number of sequences from chromosome 6B of bread wheat, did not contain any of the four T cell epitopes screened for. The sequences from *T. tauschii* (D genome), as well as those from chromosome 6D of bread wheat, were found to contain all of

these T cell epitopes in variable combinations per gene. The differences in epitope composition resulted mainly from point mutations. These substitutions appeared to be genome specific.

Our analysis shows that α-gliadin sequences from the three genomes of bread wheat form distinct groups. The four known T cell stimulatory epitopes are distributed non-randomly across the sequences, indicating that the three genomes contribute differently to epitope content. A systematic analysis of all known epitopes in gliadins and glutenins will lead to better understanding of the differences in toxicity among wheat varieties. On the basis of such insight, breeding strategies can be designed to generate less toxic varieties of wheat which may be tolerated by at least part of the CD patient population.

Introduction

Wheat is an important staple food because of its characteristics of high nutritional value, technical properties and the long shelf life of the kernels. The wheat endosperm contains 8-15% protein, of which 80% is gliadins and glutenines. Hexaploid *Triticum aestivum* or bread wheat originated around 8,000 years ago from a hybridization of a tetraploid *Triticum* species with the diploid donor of the D genome *T. tauschii* [Feldman et al., 1995]. The A and B genomes were most likely provided by *T. turgidum*, itself presumably formed from the wild diploid *T. monococcum* (A genome) and the donor of the B genome, a species which has so far defied conclusive identification [Feldman *et al.,* 1995]. Morphological, geographical and cytological evidence suggests *T. speltoides* (S genome) or a closely related species as the B genome ancestor. Cytogenetic research showed that the B genome is actually an altered S genome arisen by an exchange of chromosomal segments with other diploids and amphiploids, such as *Aegilops bicornis* (S^b genome) or *T. longissima* (S^l genome) [von Buren *et al.,* 2001]. According to Isidore *et al.* [Isidore *et al.,* 2005] polyploidization had a strong effect on intergenic sequences but the gene space was conserved.

The α-type gliadins of hexaploid *Triticum aestivum* are encoded by the *Gli-2* locus on the short arm of the three different group 6 chromosomes [Marino *et al.,* 1996]. Estimates for α-

gliadin copy number range from 25-35 copies [Harberd *et al.*, 1985] to 100 [Okita *et al.*, 1985] or even 150 copies [Anderson *et al.*, 1997] per haploid genome. Anderson and Greene [1997] compared the sequence of 27 known cDNA and genomic clones of α-type gliadins and concluded that about half of the latter contained "in frame" stop codons and were presumably pseudogenes. The detailed constitution of the multi-gene locus is not known.

Celiac disease (CD) is caused by inflammatory, gluten-specific T cell responses in the small intestine. Specific native gluten peptides can bind to HLA-DQ2/8 and induce lamina propria CD4 T cell responses causing damage of the small intestine mucosa [Vader *et al.*, 2002a; 2002b]. Tissue damage initiates secretion of the enzyme tissue transglutaminase (tTG) for wound healing. However, this enzyme also deamidates gluten peptides, resulting in high affinity HLA-DQ2/8 binding peptides that can further increase T cell responses. Multiple T cell epitope motifs have been identified in α- and γ-gliadins as well as in glutenines [Arentz-Hansen *et al.*, 2000; Koning, 2003; Vader *et al.*, 2003; Van de Wal *et al.*, 1998], the majority of which show enhanced T cell recognition after deamidation. It also became clear that patients are generally sensitive to more than one gluten peptide. Although the DQ2/8 interaction represents the most significant association with CD so far defined, it is becoming clear that non-immunogenic gluten peptides also have an impact on the innate immunity system [Koning *et al.*, 2005; Sturgess *et al.*, 1994; Maiuri *et al.*, 2003]. Clearly, the gluten peptide repertoire involved in CD is not yet complete.

Molberg et al. [2005] and Spaenij-Dekking et al. [2005] used T cell and antibody-based assays to demonstrate that a large variation exists in the amount of CD4 T cell stimulatory peptides present in α- and γ-gliadins and glutenines among diploid, tetraploid, and hexaploid wheat accessions. If this is the result of genetic differences in gluten proteins with toxic epitopes, then this would allow to design strategies for selection and breeding of wheat varieties suitable for consumption by CD patients.

In this study we first determine whether the α-gliadin genes present in the A-, B- and D-genome ancestral species are sufficiently different to attribute the ancestral genomic origin of the α-gliadin genes in hexaploid bread wheat. Secondly, we aim at understanding the diversity of CD epitopes in the α-gliadin gene family in diploid and hexaploid wheat.

Experimental

[GenBank: DQ002569- DQ002798]

DNA extraction from wheat kernels

Accessions (Table 1) were obtained from VIR, St. Petersburg, Russia (*T. longissima*) and CGN, Wageningen, the Netherlands (*T. tauschii*, *T. monococcum* and *T. speltoides*). We followed the taxonomy of *Triticum* of Morris & Sears [1967]. Wheat kernels (250 mg) were grinded in liquid nitrogen and subsequently 5 ml of 65°C preheated extraction buffer (0.1 M Tris-HCl, pH 8.0; 0.5 M NaCl; 50 mM Na_2EDTA; 1.25 % (w/v) SDS; 3.8 g/l $NaHSO_4$) was added to the powder and was incubated at 65°C for 45 minutes. Then, 8 ml of chloroform/isoamylalcohol (24:1 v/v) was added. The mixture was shaken and centrifuged for 15 min at 3000 rpm. The supernatant was discarded and 8 ml ice-cold ethanol 96 % (v/v) was added. The tubes were shaken and consequently centrifuged for 10 min at 3000 rpm. The pellet was washed 2 times with 4 ml of 70 % ethanol and subsequently centrifuged for 10 min at 3000 rpm. The pellet was air-dried and dissolved in 500 µl of TE (10 mM Tris-HCl, pH 7.5 and 1 mM EDTA) + 10 µg/ml RNaseA. The solution was finally heated for 10 min at 60 °C and carefully shaken.

Amplification of α-gliadin genomic sequences

Primers to amplify α-gliadin genes from genomic DNA using PCR were designed on the conserved sequences at the 5' and 3' end of the coding region of the α-gliadin gene sequences obtained from the public database (forward primer, 1F: 5'-ATG AAG ACC TTT CTC ATC C-3', and reverse primer, 5R: 5'-GTT AGT ACC GAA GAT GCC-3'). Amplification was performed in a 25 µl reaction volume, containing 0.2 µM reverse and 0.2 µM forward primer, dNTP mix (0.25 mM each), 1 x Pfu buffer (Stratagene), 20 ng chromosomal DNA and a mixture of (1/4 v/v) Pfu DNA polymerase (Stratagene) (2.5 U/µl) and Goldstar DNA polymerase (Eurogentec) (5 U/µl). The PCR amplification utilized 3 min at 94˚C followed by 25 cycles consisting of 94˚C for 1 min, 55˚C for 1 min and 72˚C for 2 min with a final extension at 72˚C for 10 min.

Cloning and sequencing

The PCR products (lengths ranging from 900 to 1100 bp) were ligated into the pCRII-TOPO vector (Invitrogen) and subsequently used for the transformation of *E. coli*-XL1-blue cells (Stratagene). Recombinants were identified using blue-white color selection. Positive colonies were picked and grown overnight at 37˚C in freeze media (36 mM K_2HPO_4, 13.2 mM KH_2PO_4, 1.7 mM trisodium citrate, 0.4 mM $MgSO_4$, 6.8 mM $(NH_4)2SO_4$, 4.4 % v/v glycerol, 100 μg/ml ampiciline, 10 g/l tryptone, 5 g/l yeast extract and 5 g/l NaCl). The cloned insert was amplified directly from the culture in a PCR reaction using the M13 forward primer (5'-CGC CAG GGT TTT CCC AGT CAC GAC-3') and the M13 reverse primer (5'-AGC GGA TAA CAA TTT CAC ACA GGA-3') in 20 μl reaction volume containing 2 μl of culture. The reaction mixture consisted of the same components as well as concentrations, and utilized the same PCR program as described before. The amplified product was used in a sequencing reaction using 1F and 5R primers. Additional primers were designed on two other conserved regions of the α-gliadin gene to sequence the insert: one internal forward primer (designed on pos. 292-309), Fi1: 5'-CAA CCA TAT CCA CAA CCG-3', and one internal reverse primer (designed on position 599-615), Ri1: 5'-CA(C/T) TGT GG(A/C) TGG CTT GGC-3'. The sequence data were manually checked using the computer program Seqman from the DNAstar package. The obtained sequences were deposited in GenBank (accession numbers in Table 1).

Phylogenetic analyses of the obtained α-gliadin clones

The deduced amino acid sequences were aligned using Clustal X (version 1.81). Phylogenetic trees were inferred by neighbour-joining (Clustal X) and parsimony (PHYLIP version 3.57c; DNAPARS) [Felsenstein, 1985] and subsequently viewed using TreeView (version 1.6.6). The phylogenetic trees from the neighbour-joining (Figure 2) and parsimony analysis (not shown) were nearly identical and differed only in the organization of branches that were supported by low bootstrap values.

The deduced amino acid sequence of the full-ORF clones were analyzed without the targeting sequence (first 17 amino acids were removed) up to the former last conserved cystein residue (lengths range from 244 to 271 amino acids). In this way both primer regions were

omitted. The first repetitive domain (R) (Figure 1) was analyzed from the amino acid residue on position 18 (targeting sequence was removed) until the start of the first polyglutamine repeat (length of first domain 93-105 amino acids). The first non-repetitive domain (NR1) starts with the first amino acid after the first polyglutamine repeat and ends one amino acid before the second glutamine repeat (length 68-73 amino acids). The third domain starts with the first amino acid after the second polyglutamine repeat until the former last conserved cystein residue (length 57 or 58 amino acid residues). The glutamine repeats were analyzed using the number of amino acid residues located between the beginning and the end of the polyglutamine repeat.

Analysis on synonymous and non-synonymous substitution

The obtained nucleotide sequences were aligned codon-by-codon using Clustal W. We analysed general selection patterns at the molecular level using DnaSp 4.00 [Rozas *et al.*, 2003]. Insertions or deletions that cause a frame-shift were treated as non-synonymous substitutions. The number of synonymous (K_s) and non-synonymous substitutions (K_a) per site were calculated from pair wise comparisons with incorporation of the Jukes-Cantor correction, as described by Nei and Gojobori [1986]. Pair wise comparisons with fewer than seven non-synonymous mutations refer to closely related sequences and contain no useful information on substitution rates. This concerned 2528 out of 9243 pair wise comparisons, which were excluded from the analyses.

Epitope screening

All α-gliadin DNA sequences obtained in this study were translated to protein sequences and converted into FASTA format. In addition, public domain gliadin and glutenin sequences from bread wheat were extracted in FASTA-format from the Uniprot database (www.uniprot.org) with the following conditions: *Triticum aestivum* and (gliadin or glutenin). The program PeptideSearch [Mann and Wilm, 1994] was used for matching the predicted epitopes from α-gliadin with the databases described above. Only perfect matches were considered in the scoring.

Results

Analysis of the genomic α-gliadin genes from diploid species that represent the ancestral genomes of bread wheat

The typical structure of the α-gliadin is depicted in Figure 1. The fact that the sequences at the 5' end (signal peptide) and 3' end of the genes are highly conserved within the α-gliadin gene family enables to obtain different members of the gene family by a PCR-based method on genomic DNA of various wheat species (Table 1). Accessions used were *Triticum monococcum*, which represents the A genome; *T. speltoides* (two accessions) and *T. longissima* that represent relatives to the B genome, and *T. tauschii* as representative of the D genome of wheat. We included these two species to represent the B genome, since these are thought to be related to the as yet unknown ancestor. This yielded 230 unique DNA clones with high similarity to known α-gliadin genes (Table 1) that were not present in the public databases. Only 31 of these sequences contained a non-interrupted full open reading frame (full ORF) α-gliadin gene. The great majority of the obtained sequences contained one or more internal stop codons or (rarely) a frameshift mutation (Table 1). We refer to the latter sequences as pseudogenes. Remarkably, no full-ORF genes but only pseudogenes from *T. longissima* were found.

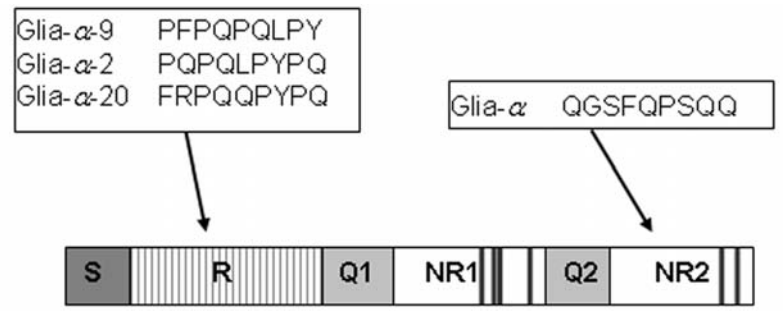

Figure 1 - Schematic structure of an α-type gliadin protein. The protein consists of a short N-terminal signal peptide (S) followed by a repetitive domain (R) and a longer non-repetitive domain (NR1 and NR2), separated by two polyglutamine repeats (Q1 and Q2). In the non-repetitive domains five conserved cystein residues are present which are indicated with vertical lines. The T cell epitopes are shown and their approximate position is indicated.

Table 1 - Number of obtained unique full open reading frame (full-ORF) and sequences with one or more stop codons (pseudogenes) from various diploid Triticum species. Accession numbers are given between brackets.

Genome	Species, Accession	Full-ORF	Pseudogenes	Total
A	*T. monococcum*, CGN 06602	15 (DQ002569 - DQ002583)	39 (DQ002600 - DQ002638)	54
B	*T. speltoides*, CGN 10682[1]	2 (DQ002584 - DQ002585)	23 (DQ002639 - DQ002661)	25
	T. speltoides, CGN 10684[1]	3 (DQ002586 - DQ002588)	9 (DQ002662 - DQ002670)	12
	T. longissima[1]	0	66 (DQ002671 - DQ002736)	66
D	*T. tauschii*	11 (DQ002589 - DQ002599)	62 (DQ002737 - DQ002798)	73
Total		31	199	230

[1] The correct annotation is S genome for *T. speltoides* and S^1 genome for *T. longissima* [Feldman *et al.*, 1995], but as they are here taken as closest representatives of the B genome, we will, for clarity, refer to them as B genome.

A phylogenetic analysis of the deduced amino acid sequence of the full-ORF α-gliadin genes demonstrated a clear clustering of the sequences according to their genome of origin (Figure 2). The sequences derived from the A genome (*T. monococcum*) as well as the sequences from the D genome (*T. tauschii*) each formed a separate cluster of relatively closely related genes in the phylogenetic tree. The sequences originated from the two *T. speltoides* accessions (B genome) formed a relatively diverse cluster. All five sequences derived from the two different accessions of *T. speltoides* differed from each other. Accordingly, the fact that the B genome sequences were more diverse is not an artifact from the use of more than one representative accession.

To investigate whether the observed clustering of the sequences can be related to specific domains of the α-gliadin gene (Figure 1), the first repetitive domain (R), the first (NR1) and the second non-repetitive domain (NR2) were used separately in a phylogenetic analysis (not shown). In all cases the sequences clustered according to their genome of origin and again the A (*T. monococcum*) and D genome (*T. tauschii*) sequences clustered separately in two groups with closely related sequences whereas the sequences originating from the B genome (*T. speltoides*) formed a more diverse group with nodes of high bootstrap values. Only when using domain NR2 no significant bootstrap values were attached to the nodes within this group.

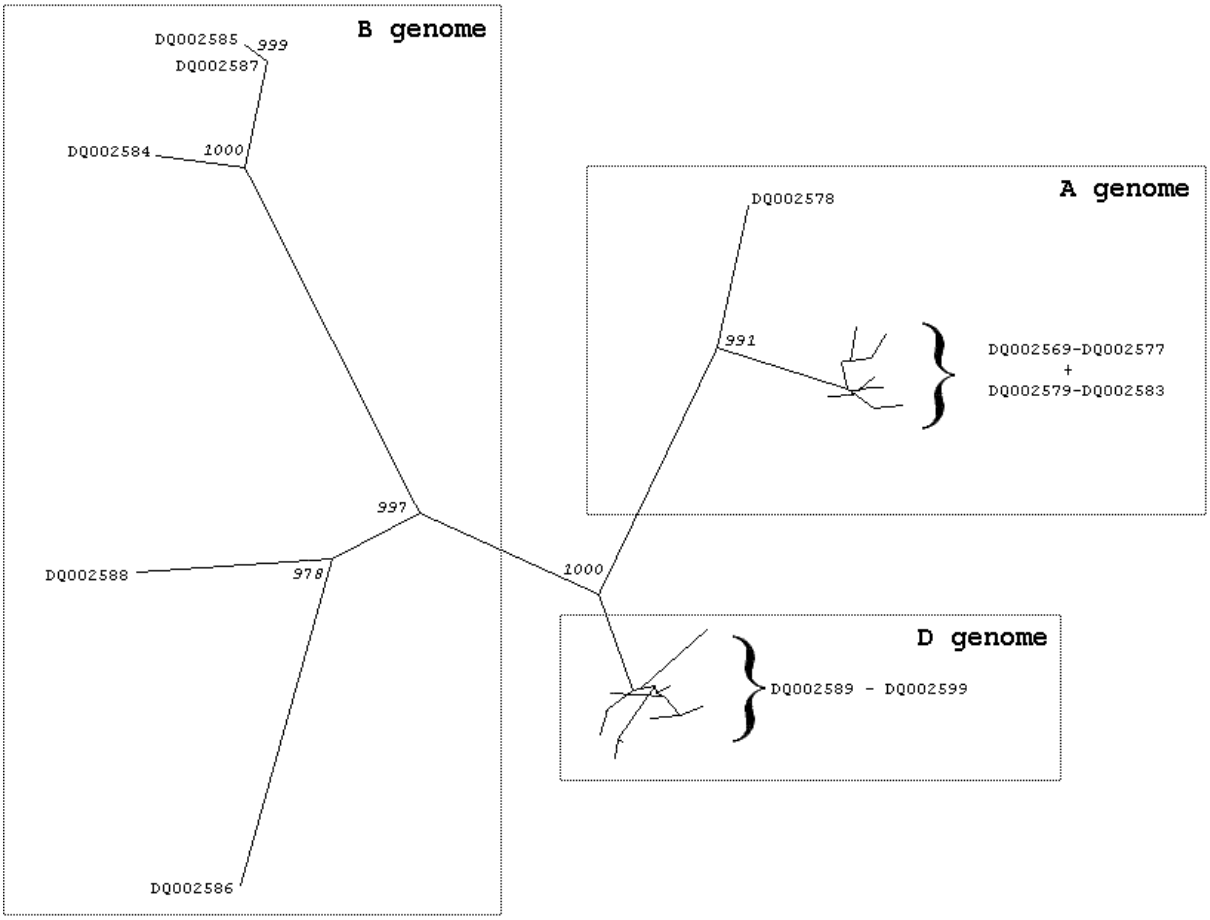

Figure 2 - Dendrogram of a ClustalX alignment of the obtained full-ORF α-gliadin deduced proteins, which are indicated by their accession numbers (see Table 1). A PAM350 matrix and the neighbor joining method were used. Bootstrap values (of 1000 replications) are given for nodes only if they were 950 or higher.

The two polyglutamine repeat domains were analyzed for differences in the average number of glutamine residues. Figure 3 shows large and also significant differences between the average lengths of the polyglutamine repeats depending on the genome of origin. The A genome (*T. monococcum*) coded for a significantly larger average number of glutamine residues in the first polyglutamine repeat than the B and D genomes. In the second polyglutamine repeat, the B genome showed a significantly larger number of glutamine residues than those of the other two genomes (Figure 3). The analysis of the repeat domains indicates that nearly all α-gliadin

sequences can be assigned to one of the three genomes using only the combination of both repeat lengths.

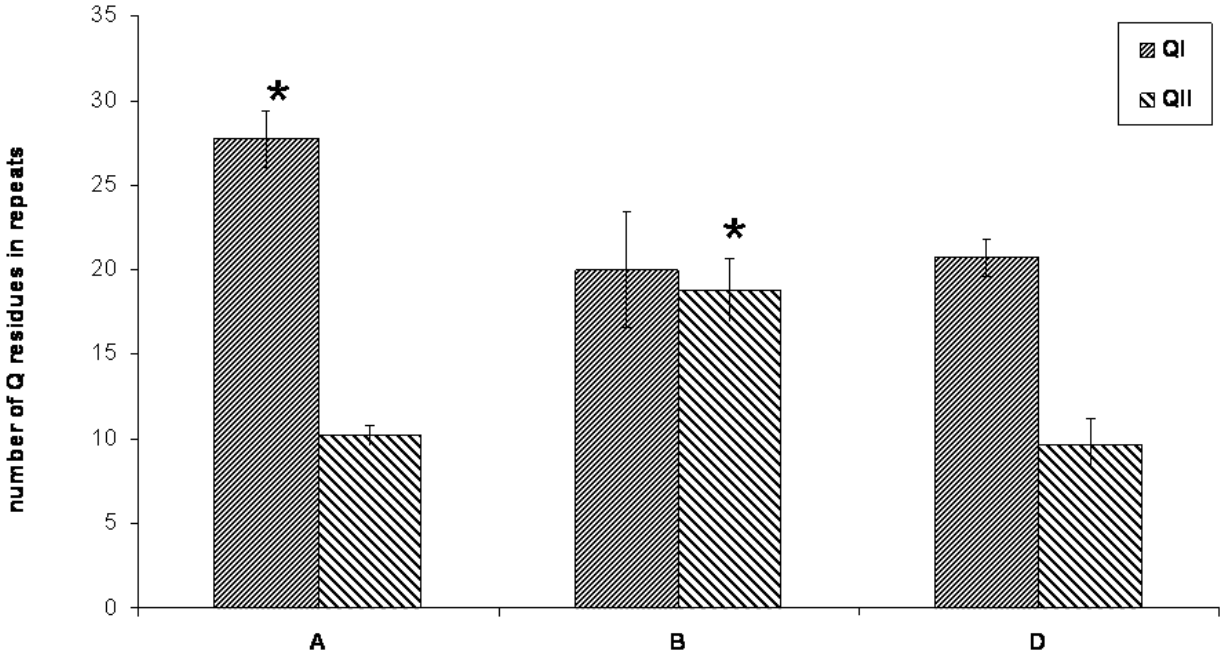

Figure 3 - Analysis of the two glutamine repeats in the 31 obtained full-ORF α-gliadin proteins from diploid wheat species, according to the genome of origin. The average number of the glutamine residues in the first (Q1) and second repeat (Q2) are shown according to the genome of origin. The A genome (T. monococcum) sequences possessed a significantly higher average number of glutamine residues in the first glutamine repeat (27.7 +/- 1.7) than the B (20.0 +/- 3.4) and D (20.7 +/- 1.1) genomes did. For the second glutamine repeat, the B genome sequences demonstrated a significantly higher number of glutamine residues (18.8 +/- 1.9) than those of the other two genomes (A, 10.2 +/- 0.6; D, 9.7 +/- 1.4).

Analysis of the pseudogenes

The great majority of the gliadin genes contained one or more internal stop codons. We refer to them as pseudogenes, although we cannot predict from the genomic data whether a subset is being expressed or not. A question is how and when these pseudogenes did evolve. Therefore, we determined their position in the clustering of the three genomes, and the relationship with intact

ORFs in the same loci. These pseudogenes are structurally similar to the full-ORF genes. The stop codons were nearly always located at positions where the full-ORF genes contained a glutamine residue codon. A stop codon was the result of a C to T change in 77.2% of the cases when compared with the full-ORF genes, altering a CAG or CAA codon for glutamine into a TAG or TAA stop codon. In addition, we observed that 15.5% of the stop codons were caused by T to A change, altering the codon for leucine (TTG) into a stop codon (TAG). Beside these major occurring substitutions we observed some C to A, C to G, G to T, and G to A changes. Twenty of the 199 pseudogenes contained a frameshift mutation (two were obtained from *T. monococcum* (A genome), two from *T. tauschii* (D genome) and 16 from *T. longissima* and the two *T. speltoides* accessions (B genome)).

The changes into stop codons were not distributed randomly across the amino acid residue positions in the sequences, and they were not distributed evenly across the various diploid species. A high percentage of stop codons occurred jointly in one pseudogene, and many pseudogenes from one species contained the same set of stop codons, suggesting that they have been duplicated after the mutations created the stop codons (Figure 4). A dendrogram of the deduced amino acid sequence of the great majority of non-frameshift pseudogenes, including the deduced amino acids downstream of the internal stop codon, closely resembled that of the full-ORF sequences. Only eleven percent of all pseudogene sequences clustered separately from the rest of the sequences of the same genome of origin.

To study the selection pressure on the obtained sequences the number of synonymous (K_s) and non-synonymous (K_a) substitutions per site were calculated from pair wise comparisons among the obtained full-ORF gene sequences and the pseudogene sequences (Figure 5). The trendlines indicated a relative excess of synonymous substitutions compared to non-synonymous substitutions and showed a stronger excess for the full-ORF genes. Consequently, the mean K_a/K_s ratio for the genes was significantly lower than that of the pseudogenes (*t* test; $P<0.0005$), indicating the occurrence of selection.

Since the first stop codons occur in various positions in the pseudogenes, it was not feasible to select a large number of sequences of sufficient and similar length to compare the selection pressure of the sequences up to the first stop codon with that of the sequences beyond it.

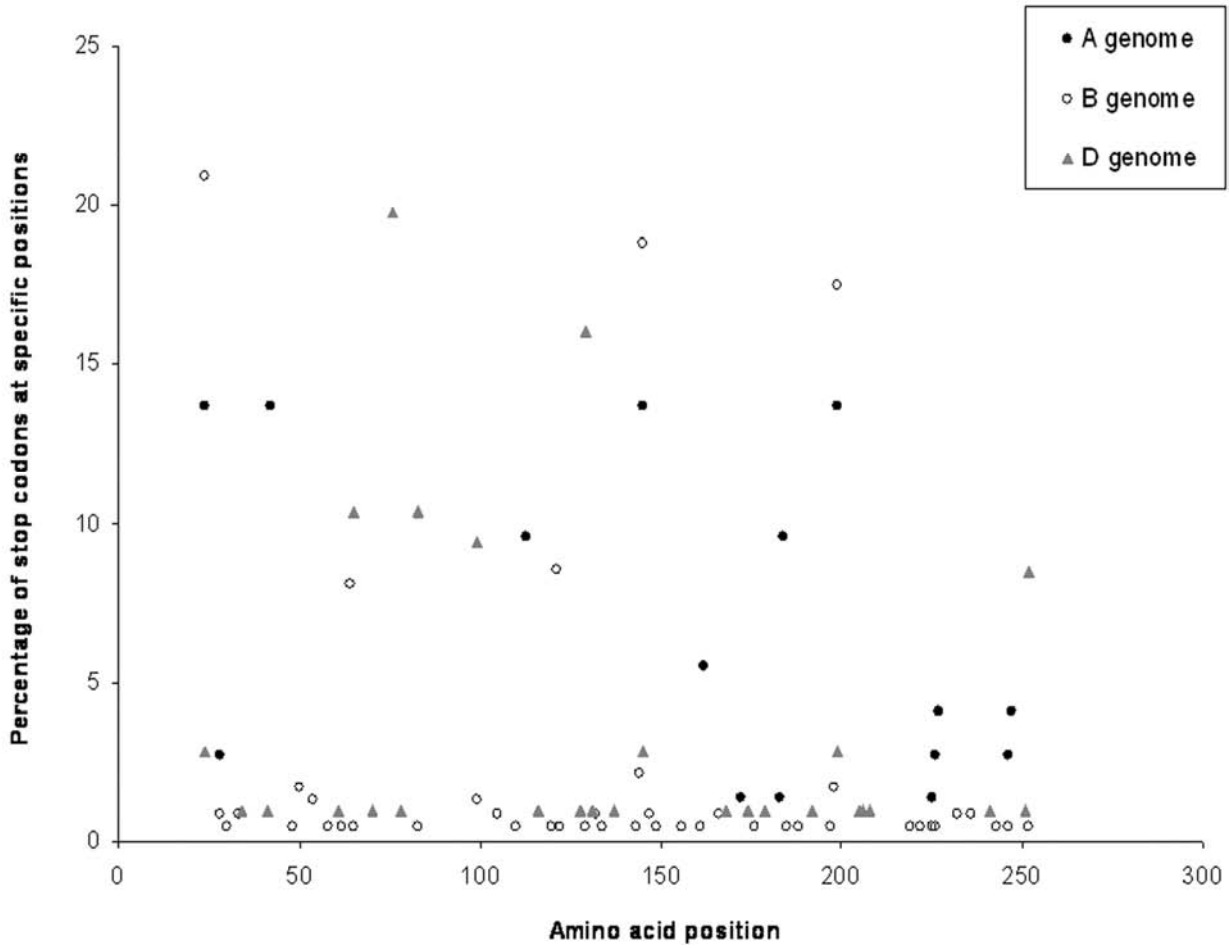

Figure 4 - Distribution of stop codons in the pseudogenes according to the amino acid position in the sequences. The positions of the stop codons are not distributed evenly across the various diploid species. The A genome sequences have a high percentage of stop codons at positions 24, 42, 145, 199 and these four stop codons may occur jointly in one pseudogene sequence. The B genome sequences also contain a high percentage of the jointly occurring stop codons at position 24, 145 and 199 but do not contain the stop codon at position 42. The jointly occurring stop codons 24, 145 and 199 are present in a few pseudogenes originating from the D genome. Pseudogenes from the A genome may contain another pair of jointly occurring stop codons at position 113 and 184 whereas the pair at positions 64 and 121 occurs in B genome pseudogenes, and pairs of stop codons at positions 65 and 83 and at the positions 99 and 252 occur in D genome pseudogenes.

Analysis of sequences from hexaploid bread wheat

If the features described above that distinguish the α-gliadin genes from different diploid genomes, are present in hexaploid wheat in the same way, this would make it possible to assign the sequences as well as the known T cell stimulatory epitopes of α-gliadins from hexaploid wheat to one of the three loci, on chromosome 6A, 6B, or 6D. Since many hexapoid sequences are present in the public database of EMBL/Genbank/DDBJ, we tested this using the deduced amino acid sequence of these 56 full-ORF genes to build a phylogenetic tree (accession numbers are given in Table 2). The sequences of hexaploid wheat clustered into three different groups (data not shown), as did the obtained sequences from this study, separated by a very high bootstrap value (998/1000). Joint analysis together with our full-ORF sequences from diploid species showed that the three groups coincide, and this allowed us to assign each of the genes of database sequences to one of the three *Gli*-2 loci (Table 2).

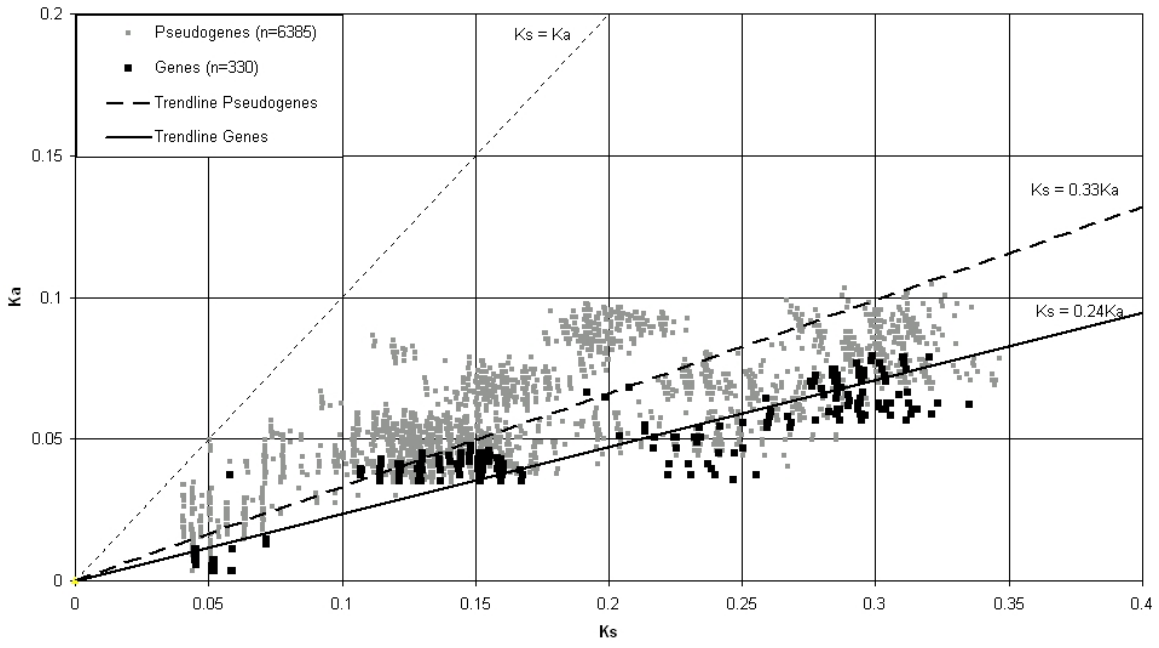

Figure 5 - The relation of the relative numbers of synonymous substitutions (K_a) and non-synonymous substitutions (K_s) per site for pairwise comparisons among full-ORF α-gliadins and pseudogene sequences. The dotted line represents a K_a/K_s ratio of 1. Linear trendlines with the intercept set to zero are shown both for full-ORF sequences and pseudogene sequences.

Analysis of CD-toxic epitopes

Our phylogenetic analyses show that the α-gliadin genes are distinct in their sequence conservation depending on the genomic origin. Are these patterns also being reflected in the occurrence of T cell stimulatory epitopes in the genes depending on their genomic origin? Table 3 shows the number of perfect matches in the obtained full-ORF genes and in the pseudogenes to the four epitopes studied. The results demonstrate that the set of epitopes is indeed distinct for each genome. Firstly, in the A genome (*T. monococcum*) sequences, the epitopes glia-α9 and glia-α20 were present in all 17 different full-ORF genes and in 39 (glia-α9) and in 38 (glia-α20) of the 44 pseudogenes. However, the epitopes glia-α and glia-α2 were absent. Also among the database sequences from hexaploid *T. aestivum* the sequences assigned to chromosome 6A showed the same trend in epitope occurrence (Table 2). Secondly, in the five obtained full-ORF sequences from the B genome species epitopes were completely absent except for two genes which contained the epitope glia-α only. Correspondingly, only four out of the 20 hexaploid wheat database sequences that were assigned to chromosome 6B contained epitope glia-α, whereas all others were without epitopes. Of the pseudogenes we obtained from the B genome species, 17% contained the glia-α epitope and only 3% the glia-α2 epitope, but these pseudogenes did contain the epitopes glia-α9 and glia-α20 at frequencies of 53% and 55%, respectively. Finally, in the 11 full-ORF sequences and the 64 pseudogenes obtained from the D genome, a frequent occurrence of all four different epitopes was found. This also applied to the five hexaploid wheat database sequences assigned to chromosome 6D.

Each epitope had its own position in the α-gliadin protein. Glia-α was in all cases present in the second non-repetitive domain (NR2), whereas glia-α2, glia-α9 and glia-α20 were all found in the first repetitive domain (R). A closer look at these sequences revealed that a single nucleotide polymorphism (SNP), which resulted in an amino acid change in a particular epitope, was present in most or all genes originating from one of the three genomes. For example, Figure 6 shows that the glia-α epitope in all of the full-ORF genes derived from the A genome were disrupted at the fifth amino acid of the epitope by the presence of an arginine (R) instead of a glutamine (Q). In three B genome sequences the glia-α epitope was disrupted at the second amino acid of the epitope by the presence of valine (V) instead of a glycine (G). A detailed overview of

presence of the epitopes glia-α2, glia-α9 and glia-α20 in the obtained full-ORF sequences are shown in Figure 7.

Table 3 - Number of T cell stimulatory toxic epitopes present in full-ORF genes (upper panel) and pseudogenes (lower panel). N is the total number of genes used in the analyses.

Genome	Full-ORF genes from	glia-α	glia-α2	glia-α9	glia-α20	N
A	*T. monococcum*	0	0	15	15	15
B	*T. speltoides*	2	0	0	0	5
	T. longissima	-	-	-	-	-
D	*T. tauschii*	10	8	11	10	11
Genome	**Pseudogenes from**	**glia-α**	**glia-α2**	**glia-α9**	**glia-α20**	**N**
A	*T. monococcum*	0	0	34	34	39
B	*T. speltoides*	7	3	18	19	32
	T. longissima	10	0	34	35	66
D	*T. tauschii*	45	23	28	49	62

Here we show for the first time that large differences exist in the content of predicted T cell epitopes (glia-α, glia-α2, glia-α9, glia-α20) in full-ORF genes and pseudogenes from the diploid species. This phenomenon was also in hexaploid wheat. None of the diploid A genome sequences and none of the sequences from chromosome 6A in the hexaploid bread wheat contained glia-α and glia-α2 epitopes (Table 2 and 3). In contrast, the sequences from the D genome contained all four epitopes at high frequencies, both in the diploid species and in the hexaploid bread wheat. For the B genome, the five diploid and 20 hexaploid full-ORF sequences rarely contained the epitope glia-α and did not contain one of the other three epitopes. Based on this analysis, we predict that among the α-gliadin proteinss, those coded by the B genome are the least likely to stimulate CD4 T cells. Remarkably, the pseudogenes revealed the presence of all the epitopes. In these analyses we have assumed that a single amino acid substitution is sufficient to prevent such peptides from stimulating the T cells, especially since the substitutions often concern a glutamine residue. Glutamine residues can be deamidated to glutamic acid by tTG in the human gut providing. The negative charges necessary to enhance binding in the DQ2 groove [Vader *et al* ., 2002a; 2002b].

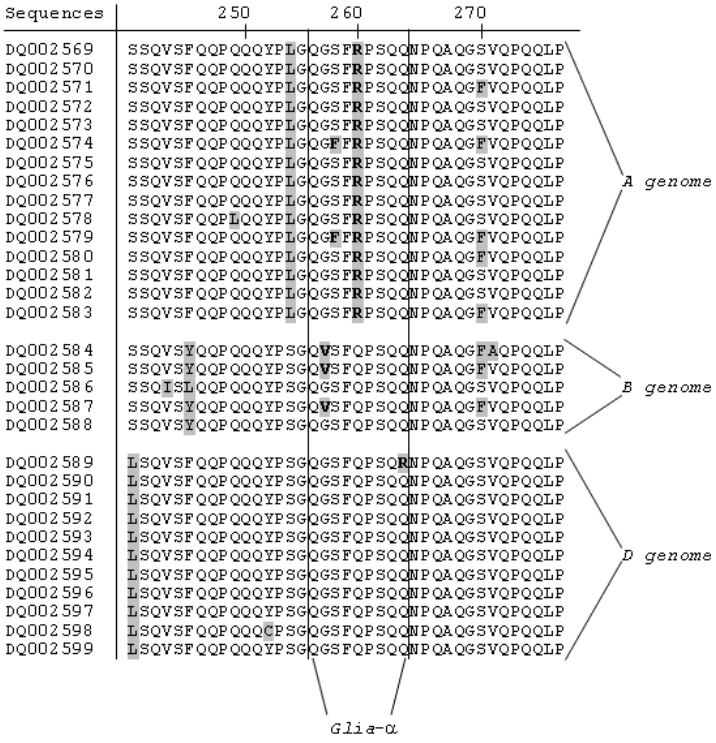

Figure 6 - Partial detailed alignment of the obtained full-ORF α-gliadin proteins. The figure shows the disruption of epitope glia-α (QGSFQPSQQ) by a single amino acid change in all T. monococcum (A genome) sequences and three of the T. speltoides (B genome) sequences.

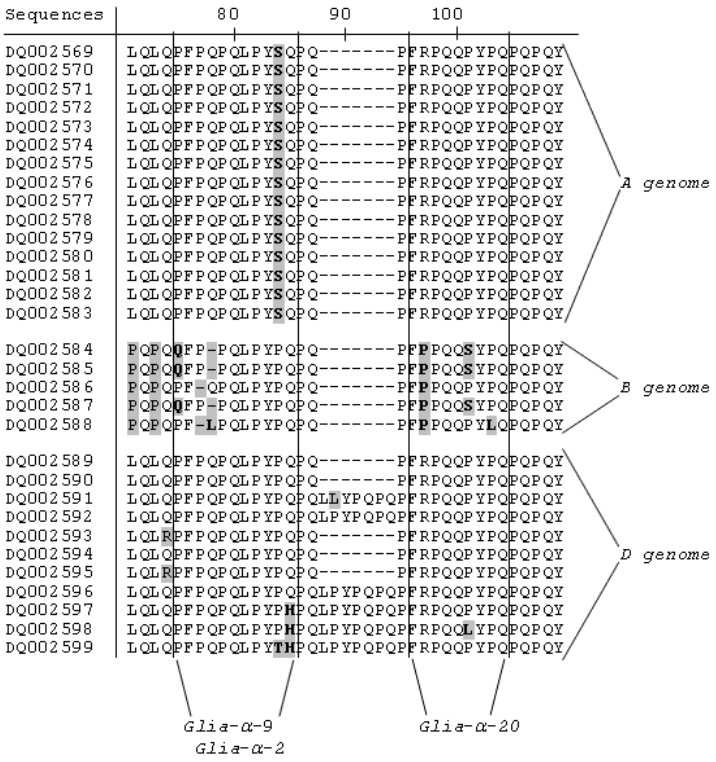

Figure 7 - Partial detailed alignment of the obtained full-ORF α-gliadin proteins, showing the disruption of epitope glia-α2 (PQPQLPYPQ) in all T. speltoides (B genome), T. monococcum (A genome) and three T. tauschii (D genome) sequences. Secondly the figure shows the disruption of epitope glia-α9 (PFPQPQ LPY) in the T. speltoides (B genome) sequences and finally the disruption of epitope glia-α20 (FRPQQPYPQ) in all T. speltoides (B genome) and in one T. tauschii (D genome) sequence.

Table 2 - Number of T cell stimulatory epitopes present in full-ORF α-gliadin genes originating from T. aestivum according to the deduced genome of origin. Sequences are obtained from the public databases. The α-gliadin locus is on chromosome 6, but the genome (i.e., chromosome 6A, 6B, or 6D) is deduced from clustering together with sequences from the diploid species representing the ancestral genomes.

Accession number	Deduced chromosome	glia-α	glia-α2	glia-α9	glia-α20
AAA17741	6A				1
AAA34280				1	1
AAA34281					1
AAA96276				1	1
AAA96523					1
AAA96524				1	1
AAA96525				1	1
B22364					1
BAA12318				1	1
CAA10257				1	1
CAA25593				1	1
CAA26384				1	1
CAB76955				1	1
CAB76958					1
CAB76959					
CAB76960					1
CAB76961				1	1
CAB76962				1	1
CAB76963				1	1
P02863				1	1
P04721					1
S07923				1	1
T06282					1
A22364	6B				
A27319					
AAA34275					
AAA34277					
AAA34278					
AAA34279		1			
AAA34283					
AAA96522					
CAA26383					
CAA26385		1			
CAB76954					
CAB76957					
E22364					
P04723					
P04725					
P04726					
P04727		1			
S07361					
S07924		1			
T06504					
AAA34276	6D		1	1	1
AAA34282		1	1	1	
C22364					1
CAA35238		1	2	1	1
CAB76956					1
CAB76964		1	2	1	1
D22364		1	1		1
P04722			1	1	1
P04724		1	1	1	
P18573		1	2	1	1
S10015		1	2	1	1
T06498			1	1	1
T06500		1	1	1	

Discussion

Gene copy number and complexity

The diploid wheat species used in this study contain a large number of α-gliadin copies in their genome. The sequences we obtained show that the fraction of genes with in-frame stop codons is very high, ranging from 72% in the A genome species to 95% in the B genome species (Table 1). Our *in silico* comparison shows a similar situation in hexaploid wheat. The fraction of these pseudogenes appears to be higher than previously found by Anderson and Greene [1997]. Analysis of the synonymous (K_s) and non-synonymous (K_a) substitutions in the obtained full-ORF genes and pseudogenes revealed that the pseudogenes contain more non-synonymous substitutions than the full-ORF genes. This is consistent with a reduced selection pressure on the pseudogenes. These results suggest that the majority of these sequences are not expressed (or only expressed up to the first stop codon).

Evolution

The obtained full-ORF genes cluster together according to their genome of origin in a phylogenetic analysis. The sequence differences in the various domains of the α-gliadin genes all contribute to this clustering. The differences consisted of point mutations leading to amino acid changes at specific positions. These amino acid changes are often genome specific, suggesting that most of the duplications of this gene family have taken place after the different diploid species separated from a common ancestor. From our data, the length differences in the two glutamine repeats of the gliadin genes, which were as observed by Anderson and Greene [1997], turned out to be related to the genomic origin of the genes as well. This may have occurred through the same mechanism as was found in the evolution of microsatellite repeats, where large-range mutations (duplication or deletion of a larger number of repeats through unequal crossing-over) occur infrequently, while small-step mutations (one repeat longer or shorter due to slippage) are frequent [Li *et al.,* 2002]. This would produce groups of similarly-sized repeats in the sequences from each genome, but the average length of each glutamine repeat could be quite different between different genomes. In addition, the large differences in the average lengths of

the two repeats in the same gene indicate that unequal crossing-over between the two repeats does not take place.

Interestingly, our results clearly indicate that at least 70% of the stop codons in the pseudogenes are position and genome specific. The occurrence of stop codons at identical positions in different sequences demonstrates that pseudogene duplication has occurred. The observation that three of the stop codon positions are shared between the A and the B genome implies that some pseudogene duplications must have taken place in the common ancestor.

Based on the structural similarities to other gliadin storage proteins like the γ- and ω-gliadins [Shewry and Tatham, 1990], the α-gliadin genes on chromosome 6 are suggested to have originated from a gliadin gene on chromosome 1 through a duplication and/or translocation event [22] after the separation of wheat from rye and barley [Shewry and Tatham, 1990]. We observed that the α-gliadin genes of *T. speltoides*, and of the corresponding B genome in hexaploid bread wheat as well, are more diverse than the α-gliadin genes on the A and D genome. One explanation for this phenomenon is chromosome exchange with other species during the formation of the ancestral B genome, which is also suggested by other authors [Feldman *et al.*, 1995]. The diversity of the pseudogenes obtained from *T. speltoides* and *T. longissima* also supports this assumption. In addition, the outbreeding character of these species may have further facilitated this recombination and maintenance of diversity.

T cell stimulatory epitopes in α-gliadin sequences

Our results indicate that, with respect to T cell toxicity as far as caused by α-gliadins, and based on currently known α-gliadin epitopes, the *Gli-2* locus on the D genome should be considered as the most relevant. This is in agreement with the results of Spaenij-Dekking and colleagues [2005] who found the highest presence of T cell-stimulatory epitopes (glia-α-2/9) in D genome species compared to A and B genome species. In addition Molberg et al. [2005] found that fragments identical or equivalent to a αG-33mer protein fragment appear to be encoded by α-gliadin genes on the wheat chromosome 6D and are absent from gluten of diploid Einkorn wheat (A genome) and even certain cultivars of the tetraploid pasta wheat (AB genome). If these predictions are confirmed in *in vivo* studies it may follow that breeding of bread wheat for low toxicity should

focus, as one of the targets, on lowering the α-gliadin proteins from the D genome. The D genome has contributed significantly to many characteristics of hexaploid wheat, including baking quality, through HMW glutenins on chromosome 1D, but there is no evidence for a specific contribution of the *Gli-2* locus on chromosome 6D to baking quality.

Our study focused on α-gliadin genes present in the genome, and did not consider possible differences in expression among the multiple copies of α-gliadin genes. Spaenij-Dekking and colleagues [2005] found large differences in T cell stimulatory epitopes in protein from different hexaploid and tetraploid accessions. Combined with our results, this may imply large differences in expression of toxic D-genome α-gliadin genes, possibly through interaction with the homologous loci on other chromosomes, as was found for ω-gliadins [Islam *et al.*, 2003]. In that case, the mRNA pool of α-gliadins would not perfectly match the genomic composition. Alternatively, genetic differences do exist in α-gliadin sequences among hexaploid wheat cultivars.

Conclusions

We have shown for the first time that α-gliadins from diploid *Triticum* species form distinct groups. This is reflected in large differences in the content of four T cell stimulatory epitopes (glia-α, glia-α2, glia-α9, glia-α20) in full-ORF α-gliadin genes and pseudogenes from these diploid species. Similar differences were shown to exist between the three genomes of hexaploid bread wheat. The sequence information we obtained forms a useful prerequisite for study of expressed α-gliadin mRNA and determination of both their genome of origin and their epitope content. Besides, the genetic composition of the α-gliadin loci needs to be compared across a large series of hexaploid bread wheat cultivars. As there may be more, still unknown, T cell stimulatory epitopes in all types of gluten proteins, and given that the role of the innate immune system is only beginning to be understood, it may be premature to start breeding of non-toxic wheat varieties. However, our results indicate that (re)construction of hexaploid wheat using a non-toxic D genome donor would reduce the overall T cell stimulation in CD patients.

Chapter 3

Detailed analysis of the expression of an α-gliadin promoter and the deposition of α-gliadin protein during wheat grain development

Teun WJM Van Herpen, Mandy Riley, Caroline Sparks, Huw D Jones, Cristina S Gritsch, Liesbeth Dekking, Rob J Hamer, Dirk Bosch, Elma MJ Salentijn, Marinus JM Smulders, Peter R Shewry, Luud JWJ Gilissen

Abstract

We report detailed studies of the pattern of expression of the promoter of a B genome-encoded α-gliadin gene and deposition of α-gliadin protein in developing wheat grains. A 592 bp fragment from an α-gliadin promoter drove expression of a GUS reporter gene primarily in the cells of the starchy endosperm, notably in the subaleurone but also in the aleurone layer. The α-gliadin promoter was active from 11 DAF (days after flowering) until maturity showing a pattern of expression similar to that of a 326 bp LMW subunit gene promoter reported previously. An α-gliadin-specific antibody detected α-gliadin protein in protein bodies in the starchy endosperm and in the subaleurone layer but, in contrast to the promoter activity, not in the aleurone cell layer. This suggests that the 592 bp promoter fragment does not include all of the elements required in cell-specific expression, and that additional regulatory elements upstream of the region used are required to repress expression in the aleurone layer. Sequence comparison showed differences in regulatory elements between the promoters of α-gliadin genes originating from different genomes (A and B) of bread wheat both in the region used here and upstream.

Introduction

Wheat is an important staple food over much of the world, being consumed after processing into a range of food products including breads, noodles and pasta. These processing properties depend primarily on the gluten proteins, which account for about 80% of the total proteins in white flour.

Much of the research on wheat is focused on understanding and improving its properties in bread making, including understanding the relative contributions of genetic and environmental factors to variation in these properties. However, the consumption of products containing wheat can have serious consequences for the health of those suffering from allergy or intolerance to wheat proteins, with celiac disease (CD) in particular, affecting about 0.5 to 1% of the human population in Western Europe [Koning *et al*., 2005]. The α-gliadins of gluten appear to be the most active group of wheat proteins in triggering CD [Arentz-Hansen *et al.*, 2000; 2002; Vader *et al.*, 2002b; 2003; Molberg *et al.*, 2003]. These proteins are encoded by *Gli-2* loci on the shorts arms of the group 6 chromosomes of bread wheat, with estimates of the numbers of individual α-gliadin genes ranging from 25-35 copies [Harberd *et al.*, 1985] to 150 copies [Anderson *et al.*, 1997] per haploid genome. A considerable amount of data on the sequences of epitopes that are recognised by T-cells of celiac disease patients (CD toxic epitopes) and their distributions within the α-gliadin family of proteins is available [Chapter 2]. This shows that they are non-randomly distributed between the sequences encoded by the different genomes: four different epitopes have been found among the D genome-encoded α-gliadins, two of these were also in A genome-encoded genes, but hardly any in the B genome-encoded genes. Antibodies against these epitopes have been developed recently, enabling the detection of specific α-gliadins [Mitea *et al.*, 2008].

Modern biotechnology offer opportunities to remove the proteins containing the most toxic epitopes or even the epitopes themselves, using transgenesis or *in vivo* mutagenesis. However, such modifications may also lead to altered technological properties, as the gluten proteins are the major determinants of the functional properties. For example, wheat has been genetically engineered to add additional genes encoding high molecular weight (HMW) glutenin subunits using their own endosperm-specific promoters [Blechl and Anderson, 1996; Altpeter *et*

al., 1996, Barro *et al.*, 1997], leading to both positive and negative effects on the properties of the dough for mixing and bread-making [Popineau *et al.*, 2001].

Becker *et al.* [2006] demonstrated that inhibition of the expression of the complete α-gliadin family can be achieved by using RNA interference. This drastic modification resulted in little effect on dough resistance and extensibility but in an increase in dough strength and a small decrease in loaf volume [Wieser *et al.*, 2007]. A less drastic approach would be to reduce the expression of only A and D genome-encoded α-gliadin genes, which contain most of the known CD toxic epitopes [Chapter 2] while retaining or enhancing the expression of α-gliadin genes from the B genome, which do not contain CD epitopes. However, to achieve this we require a more detailed understanding of the structure and regulation of expression of the α-gliadin gene promoters.

Previous studies revealed three important conserved *cis*-motifs in endosperm-specific promoters from wheat, barley and rice, namely the GCN4-like motif, the prolamin box and the AACA/TA motif [Dong *et al.*, 2007]. Additionally, the RY repeat is described to have a key role in seed-specific gene regulation of seed storage proteins [Fujiwara and Beachy, 1994]. The CGN4-like motif plays a central role in controlling endosperm-specific expression [Onodera *et al.*, 2001] and is the target of basic leucine zipper transcription factor (bZIP) proteins that belong to the Opaque2 subfamily [Albani *et al.*, 1997; Onate *et al.*, 1999; Wu *et al.*, 2000]. Deletion and point mutation experiments revealed that the prolamin box is important for the regulation of expression of endosperm-specific genes [Mena *et al.*, 1998; Diaz *et al.*, 2005]. The AACA motif was found to be conserved in the rice glutelin gene promoter and is involved in endosperm-specific expression [Wu *et al.*, 2000; Takaiwa *et al.*, 1996].

Stoger *et al.* [1999, 2001] studied an endosperm-specific LMW subunit gene promoter, showing specific expression in the outer subaleurone cells of the endosperm of transgenic bread wheat (*Triticum aestivum*). These cells are known to be rich in protein [Evers, 1970; Kent, 1966]. Lamacchia *et al.* [2001] similarly characterised an endosperm-specific HMW subunit gene promoter in transgenic durum wheat (*Triticum durum*), showing that expression was higher in the central, starchy endosperm cells with no expression in the aleurone cells.

Reeves and Okita [1987] described a gene encoding an α-gliadin isolated from the bread wheat cv Yamhill and identified a promoter region with various regulatory elements. Aryan *et al.* [1991] also showed that a segment of the α-gliadin gene promoter from -151 to -75 was required for optimum expression in a heterogeneous tobacco protoplast system. Six nuclear proteins from developing wheat kernels were found to interact with the first 165 bp upstream of the transcriptional start and this region was therefore suggested to have a role in the transcription of α-gliadin synthesis [Vellanoweth and Okita, 1993]. However, to our knowledge, expression of a functional gene under control of an α-gliadin promoter in wheat has not been reported previously.

To determine the pattern of α-gliadin expression in various tissues of wheat during kernel development we studied the expression of a GUS reporter gene under control of a 592 bp α-gliadin promoter fragment derived from the B-genome in stably transformed bread wheat. Using immunogold labelling and tissue printing we determined the deposition of α-gliadin protein in developing and mature wheat kernels, and compared the results to the deposition of HMW glutenin subunit in developing wheat kernels.

Experimental

Sequence similarity analysis of the α-gliadin sequence

The clone of Reeves and Okita [1987] (accession number M16496) contains the coding region of an α-gliadin gene including 1814 bp of 5' upstream sequence. The coding region was translated into an amino acid sequence and aligned from the *N*-terminus of the mature protein (omitting the first 17 amino acids comprising a signal peptide) to the last conserved cysteine residue (lengths ranging from 244 to 271 amino acids) with the same of set accessions of α-gliadin genes from the three genomes of hexaploid wheat and the diploid α-gliadin genes using Clustal X 1.81, as in Van Herpen *et al.* [Chapter 2]. A neighbour-joining tree was subsequently produced in TreeView 1.6.6.

Epitope screening of the α-gliadin sequence

The α-gliadin protein sequence was searched for known α-gliadin epitopes (Glia-α, Glia-α2, Glia-α9 or Glia-α20) [Mitea *et al.,* 2008] with only full identity matches being considered in the scoring [Chapter 2]. The antibody recognition motif, QPFPQPQL was then used for full identity matching, re-analysing the results from Van Herpen *et al.* [Chapter 2] (DQ002569-DQ002599) and the available database sequences assigned to chromosomes 6A, 6B and 6D.

Regulatory motif screening of database α-gliadin promoter sequence

The known α-gliadin promoter sequences were extracted from the NCBI (http://www.ncbi.nlm.nih.gov) database. This gave 30 promoter sequences from putative α-gliadin genes. Nine of these promoter sequences formed part of pseudogenes and seven were not accompanied by an open reading frame, so that it was not possible to determine their genomic origin based on gliadin sequence homology. The 14 remaining promoter sequences were accompanied by full α-gliadin open reading frames (Table 1) and were assigned to chromosomes 6A, 6B or 6D as in Van Herpen *et al.* [Chapter 2]. A pattern search on the promoter regions of the 14 promoter sequences allowed us to identify the presence of various regulatory sequences including the GCN4-like motif (TGAGTCA) [Onodera *et al.,* 2001], the prolamin box (TGT/CAAAG) [Vicente-Carbajosa *et al.,* 1997], AACA/TA motif [Wu *et al.*, 2000; Takaiwa *et al.,* 1996] and the RY repeat (CATGCAC) [Fujiwara and Beachy, 1994]. The percentage identities of the different promoter sequence regions were determined using MegAlign 6.1 from the DNASTAR package using Clustal V method. A blastn search was performed using the ncbi blast service at http://www.ncbi.nlm.nih.gov/BLAST/ using the nucleotide collection (nr/nt) database and optimised for somewhat similar sequences (blastn).

Constructs

A construct comprising 593 bp of the α-gliadin promoter, GUS reporter gene and nos (nopaline synthase) terminator was made. The α-gliadin-promoter was obtained using PCR from the pCR-1 construct [provided by Okita, Washington State University] using forward 5'-GCCGGAATTCAAGCTTGTCTAGTTACAGTAA-CAAC-3' and reverse 5'-

GCCGGGATCCGGTGGATTTGTATTGACCACTGC-3' primers, introducing Eco RI at the 5' side and Bam HI restriction sites at the 3' side. The PCR product was introduced in a PUC-based (PUC-M) vector using EcoRI and BamHI sites with the nos terminator at the BamHI side of the PCR product. The GUS gene was obtained using PCR on the pMDC162 construct [Curtis and Grossniklaus, 2003] using forward 5'-GCCGGGATCC*ATG*TTACGTCCTGTAGAAACCCC-3' and reverse 5'-GCCGGGATCCTCATTGTTTGCCTCCCTGCTG-3' primers, introducing BamHI sites to 3' and 5' side of the gene. The GUS gene was finally introduced in the BamHI side of the previously obtained PUC-based vector.

The pAHC20 selectable marker construct [Christensen and Quail, 1996], which contains the bar gene encoding phosphinothricin acetyltransferase (PAT), was used to confer resistance to the herbicide phosphinothricin (PPT).

Transformation

Immature embryos of bread wheat var. Cadenza were used as targets for transformation by particle bombardment using the protocol of Sparks and Jones [2004]. The following modifications were made: regeneration media contained 0.05 mM $CuSO_4$ instead of 10 mg/l $AgNO_3$; selection media did not contain 2,4D; bombardments were carried out at a pressure of 900 psi. The presence of the transgenes in putative transgenic plants was confirmed by PCR (Forward primer: 5'-AGTGTACGTATCACCGTTTGTGTGAAC-3' Reversed primer: 5'-ATCGCCGC-TTTGGACATACCATCCGTA-3'). Transformed plants were grown to maturity in a containment glasshouse and T_1 kernels harvested.

Histochemical GUS assay

Transverse sections approximately 0.5 mm thick were cut manually with a razor blade from developing or mature kernels of transgenic lines and controls. Expression of the uidA reporter gene was detected by incubating the sections in X-Gluc (5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid) buffer containing 1 mM X-Gluc, 100 mM sodium phosphate buffer, pH 7, 0,5 mM potassium ferricyanide, 0,5 mM potassium ferrocyanide and 0,1 % (v/v) Triton X-100 at 37°C for 1 hour for developing kernels or 24 hours for mature kernels. Staining was visually

assessed using a Zeiss stemi SVII stereomicroscope for mature kernels and a Leica MZ FLIII stereomicroscope for developing kernels.

Transformant progeny lines with homozygous segregations

A total of 62 independent T_0 lines were initially shown by PCR amplification to contain the uidA gene. Histochemical staining was carried out on 24 individual mature T_1 half kernels of each of 60 different PCR-positive T_0 lines (Figure 1) and on six and 12 half kernels of two other PCR-positive T_0 lines, identifying fifty-four lines which expressed GUS in the endosperm. The other halves of the kernels were kept for sowing. Thirty-one of these fifty-four lines that survived and set kernel showed segregation ratios for the expression of GUS consistent with 3:1 segregation (χ^2 test with p-values < 0.05). Six lines out of these thirty-one lines that showed the highest p-values (all higher then 0.5) were selected for segregation analysis in the T_1 generation. Eight individual GUS-positive T_1 half-kernels of each of these six lines were planted and eight T_2 kernels from each plant were tested for histochemical GUS staining. Five out of the six selected lines again showed segregation ratios in the eight progeny that were consistent with the expected segregation ratios for a single insertion site (2:1 of homozygote:hetrozygote segregation, χ^2 test with p-value=0.05). One progeny of each of these five lines that appeared to be homozygous in the T_2 kernels (with all eight kernels stained) was selected and again appeared homozygous in the T_3 kernels (all eight kernels again being positive for GUS staining). The T_3 kernels of these five lines were used for detailed analysis.

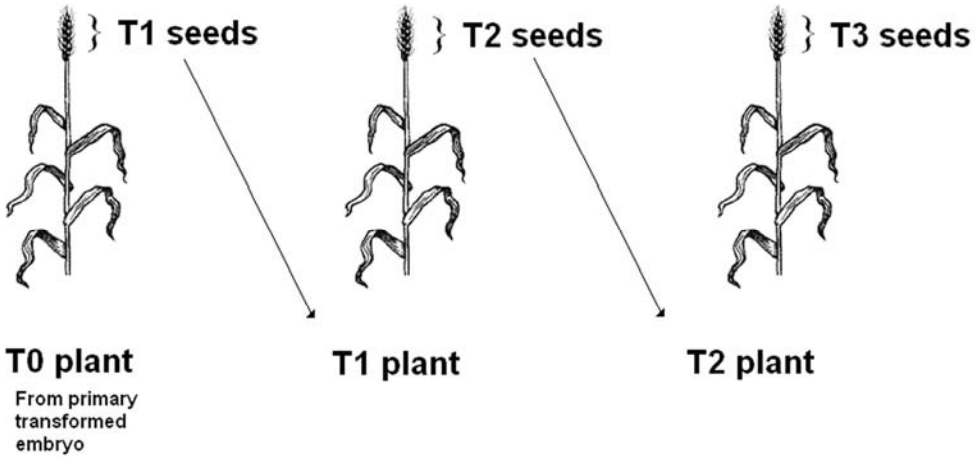

Figure 1 - Terminology of transgenic lines and kernels.

Cryostat sectioning (GUS stained)

Previously GUS stained mature kernels were frozen in cryostat embedding medium (Tissue-Tek, OCT Compound, Sakura) on a cryostat holder inside of the cryostat microtome (MicroM HM500 O). Sectioning was carried out at a temperature of -17°C. Sections of 200 μm were cut, placed on a microscope slide and visually assessed using a Zeiss Axiophot light microscope at 100 x magnification.

Cryostat Tissue printing

The method was based on that of Conley and Hanson [1997]. Whole mature kernels (approximately 30-35 DAF) were imbibed in order to soften the grain by placing the kernels onto moistened filter paper in Petri dishes sealed with Parafilm for 42 hours at 4ºC. A small transverse portion of the kernel was removed on the embryo side and the kernel was attached cut side down onto a cryostat holder with tissue freezing medium (Jung, Leica Microsystems) and quickly frozen in liquid nitrogen. Transverse sections of 15-20 μm thickness were cut with a Leica CM1850 cryostat at -17°C and attached onto nitrocellulose coated slides by gentle pressure (Oncyte Film-slides, Grace Bio-Labs Inc.). The slides were dried overnight at room temperature before proceeding with the immunolocalisation.

Immunolocalisation (Tissue prints)

The nitrocellulose coated slides with tissue prints were rinsed in washing buffer (PBS, pH 7.2, 0.3% (v/v) Tween 20) twice 15 minutes each with agitation and blocked for 1h with 5% (w/v) skimmed milk powder in washing buffer (tissue remnants were carefully removed at this stage) and then washed twice for 10 minutes each with washing buffer. The slides were then incubated for 2h in either of the two primary antibodies: the mouse monoclonal anti-α-gliadin specific antibody (Mitea *et al*. [2008]; Glia-α9 antibody obtained from Koning, Leiden University Medical Center) diluted 1:7000, or the rabbit polyclonal antibody raised to oat 8S globulin (Yupsanis *et al.*, 1990) diluted 1:5000 in washing buffer containing 0.5% BSA (bovine serum albumin). Following incubation, the slides were washed four times, 8 minutes each, with PBS and incubated for 1 hour in the secondary antibody (goat anti-mouse or goat anti-rabbit IgG alkaline phosphatase

conjugate (Sigma A-3688 and A-9919) diluted 1:3000 in washing buffer containing 0.5% BSA. The prints were washed twice, 5 minutes each of the following solutions: washing buffer+0.5% BSA, washing buffer and finally in washing buffer containing 0.05% SDS before developing using the Sigma BCIP/NBT Purple liquid solution for membranes (Sigma B3679). The reaction was stopped after a red/purple colour was observed by rinsing slides with distilled water. Tissue prints were also stained with Ponceau S (Sigma P7767) for total protein after examination of the immuno reactions. The slides were examined with a Leica MZ8 stereomicroscope and photographed with a Leica DFC 300FX digital camera. The Glia-α9 T-cell epitope is absent from most sequences originating from the B genome [Chapter 2 and Salentijn *et al.,* in prep.]. However, the determined minimal recognition motif of the Glia-α9 antibody (QPFPQPQ) is smaller then the minimal recognition motif of the Glia-α9 T-cell clone (PFPQPQLPY) [Mitea *et al.,* 2008]. By re-analysing the results of Chapter 2 we found that the minimal recognition motif of this Glia-α9 antibody is found in 2 out of 5 B-genome diploid sequences and in 15 out of 20 hexaploid 6B database sequences (also found in M16496). This indicates that this antibody recognises α-gliadin proteins originating from all three different genomes.

Fixation and embedding in LR White for immuno assay

Approximately 0.5 mm thick sections of developing wheat kernels 11, 18, 25 and 32 DAF were manually cut with a razor blade in 0.05 M phosphate buffer. The sections were fixed for 5 hours at room temperature in 4% (v/v) paraformaldehyde, 2.5% (v/v) glutaraldehyde in 0.05 M phosphate buffer pH 7.2. The fixed sections were washed three times for 30 minutes each in 0.05 M phosphate buffer pH 7.2, dehydrated in an ethanol series of 10 steps from 10% (v/v) to 100% (v/v) and infiltrated in increasing concentrations of LR white resin (TAAB Lab. Equipment Ltd.). Finally, the samples were embedded at 50°C in an oxygen-free environment for 24 hours. Sections of 1 μm thickness were cut using a Reichert-Jung Ultracut (Reichert, Vienna, Austria) ultramicrotome and placed on Poly-l-lysine (Sigma P-1399) coated multi-well slides for immunogold labelling.

Immunogold labelling and silver enhancement

The samples were blocked with blocking buffer containing 3% (w/v) BSA in 0.1 M PBST at room temperature for 30 minutes and then washed for 30 minutes with 1% (w/v) BSA in PBST. They were then incubated for 2 hours with a monoclonal α-gliadin specific (Glia-α9, 9-68, QPFPQPQ) and HMW specific (HMW 30-29 GYYPTS) antibody (obtained from Koning, Leiden University Medical Center) diluted 1:100 in 1% (w/v) BSA in PBST and washed three times for 5 minutes with 1% (w/v) BSA in PBST. They were then incubated for 1 hour using a secondary 10nm gold conjugated anti-mouse antibody diluted 1:50 in 1% (w/v) BSA in PBST, washed three times for 5 minutes with 1% (w/v) BSA in PBST, then three times with PBST and finally three times with distilled water. The samples were developed using a silver enhancement kit (British Biocell International, UK) and observed using a Zeiss Axiophot light microscope.

Pre-fixation, GUS staining, fixation and embedding in LR white

Approximately 0.5 mm thick sections of developing wheat kernels at 11, 18, 25 and 32 DAA were manually cut with a razor blade in 0.01 M phosphate buffer. The sections were pre-fixed for 30 minutes in 0.1% (v/v) glutaraldehyde in 0.05 M phosphate buffer pH 7 and then washed five times for 10 minutes in 0.05 M phosphate buffer pH 7. Histochemical GUS staining was then performed as described above and the samples washed again in 0.05 M phosphate buffer pH 7, re-fixed in 3% (v/v) glutaraldehyde in 0.05 M phosphate buffer pH 7 for 2 hours and washed again four times for 10 minutes each. The samples were then dehydrated, embedded and sectioned as described above before examination using a Zeiss Axiophot light microscope

Results

The expression pattern of an α-gliadin gene promoter in developing grains of wheat was determined by transformation of bread wheat with an B-genome α-gliadin promoter-GUS reporter construct. The accumulation of α-gliadin protein was also determined in the same tissues and stages of development by immunolocalisation using a specific monoclonal antibody. Finally, the

sequence of the α-gliadin promoter that was used was compared to those of other α-gliadin genes originating from the different genomes of hexaploid bread wheat in order to relate differences in sequence and genome origin to expression patterns.

Phylogenetic analyses, epitope- and regulatory motif screening

Alignment the α-gliadin gene (M16496) accompanying the promoter sequence as reported by Reeves and Okita [1987] showed a close relationship to α-gliadin genes encoded by chromosome 6B (bootstrap value 951/1000). The predicted α-gliadin protein encoded by the gene (M16496) was also screened for the presence of toxic celiac disease (CD) epitopes. None of the four identified celiac disease epitopes (called Glia-α, Glia-α2, Glia-α9 or Glia-α20) [Chapter 2] were detected, which is consistent with its origin from chromosome 6B [Chapter 2].

Other α-gliadin promoter sequences in the database were derived from chromosomes 6A and 6B (Table 1), but no sequences from the D genome were found. As shown previously [Chapter 2; Molberg *et al.,* 2005], CD epitopes are non-randomly distributed, indicating that the three genomes make different contributions to the ability of wheat to trigger CD. To determine if there are also sequence differences in the promoter sequences of the α-gliadin genes encoded by different genomes, we compared database sequences for the presence of several regulatory motifs [Wu *et al.,* 2000; Takaiwa *et al.,* 1996] (Table 1).

The thirteen promoter sequences of α-gliadin genes show a high sequence identity (<80%; Table 1) to the 593 bp promoter region used in this study (M16496). The GCN4 like motif was present twice in the 593 bp region of M16496 (at positions -300 and -492). The -300 GCN4-like motif was present in all of the sequences from the A and B genome, but the -492 bp GCN4-like motif was primarily present in the genes originating from the B genome (Table 1). The prolamin box was not present in the 593 bp promoter region of M16496 but is present in some other α-gliadin promoters (Table 1). The AACA/TA motif was found five times (-107, -252, -380, -421, -574) in the 593 bp region of M16496. The AACA/TA motifs at -252, -421 and -574 were present in all gliadin gene promoters from the A and B genomes, but the motifs at -107 and -380 were present in some α-gliadin promoters (Table 1). Finally, the RY repeat at position -294 was present in seven out of the eight A genome sequences and in none of the B genome sequences.

Table 1 - The presence of RY-repeat, CGN4-like motif, prolamin box and AACA/TA motif, and the percentage of identity for the 593 bp promoter region used in this study and the region further upstream of the used 593 promoter region. Positions indicated refer to sequence M16496. Genome of origin was determined based on homology of the associated gliadin gene sequence to sequences from diploid and hexaploid Triticum species [Chapter 2]. The accession number of the DNA sequence is given in the first column (the protein database number is given between brackets). n/a = not available

		Region not used		*Region used in the transformation*					
Database accession number	Genome of origin	identity % with M16496 (-594 to -895)	RY repeat (-774)	identity % with M16496 (0 to -593)	GCN4-like motif (-492)	AACA/TA motif (-380)	Prolamin box (-315)	RY repeat (-294)	AACA/TA motif (-107)
K03076 (AAA34280)	A	*n/a*	*n/a*	90	-	-	+	+	+
U08287 (AAA17741)	A	43	-	89	-	+	+	+	-
U50984 (AAA96276)	A	*n/a*	*n/a*	86	-	+	+	+	+
U51304 (AAA96523)	A	*n/a*	*n/a*	86	-	+	+	-	-
U51306 (AAA96524)	A	*n/a*	*n/a*	90	-	+	+	+	+
U51307 (AAA96525)	A	*n/a*	*n/a*	90	-	-	+	+	+
X01130 (CAA25593)	A	43	-	89	-	-	+	+	+
X02539 (CAA26384)	A	*n/a*	*n/a*	90	-	+	+	+	+
M16496 (AAA34290)	B	100	+	100	+	+	-	-	+
K03074 (AAA34277)	B	*n/a*	*n/a*	99	+	+	-	-	-
X02538 (CAA26383)	B	*n/a*	*n/a*	99	+	+	-	-	-
X02540 (CAA26385)	B	*n/a*	*n/a*	89	+	-	-	-	-
K03075 (AAA34279)	B	85	-	80	+	-	+	-	-
U51303 (AAA96522)	B	92	-	81	+	-	+	-	-
Not available in database	D	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*

Of four out of the thirteen promoter fragments, sequences are available upstream -593 bp (Table 1). Analysis of the region upstream -593 bp failed to identify a CGN-4-like motif or prolamin box. A RY repeat was present at -774 bp in the original promoter of which a fragment was used in this study (M16496, Table 1). When analyzing for sequence identity of the region -593 to -895 bp the two B genome sequences were highly similar to each other and to M16496, whereas the two A genome sequences were very similar to each other (not shown) but not to the B genome sequences. From these data we observe a general difference between A and B genome promoter sequences. The M16496 sequence that was used in this study also for the promoter sequence shows consistency with its origin from chromosome 6B.

Segregation

Segregation analysis of GUS expression in the individual T_1 kernels of each GUS expressing line showed that 31 of the 54 lines gave ratios that were consistent with a single insertion being inherited in a 3:1 ratio (Table 2). Segregation ratios consistent with 15:1, suggesting two transgenic insertion sites, were found in seven of the 54 lines. Finally, 16 out of 54 lines showed segregation ratios that were significantly different from 3:1 and 15:1 ratios (χ^2 test with p-values < 0.05; ranging from 1 to 12 GUS expressing kernels out of 24). The reason for this non-Mendelian pattern of expression is not known but it is possible that transgene silencing may have occurred in some lines thus underestimating the number of positive segregants.

GUS expression and phenotypes in different transgenic lines

All 54 lines expressed GUS in the endosperm as shown by histochemical GUS staining of T_1 kernels. A total of 1,266 kernels from the 54 lines were tested, of which 742 kernels expressed GUS in the endosperm (Table 2, column “Total”). Although considerable variation in the intensity of the GUS staining was observed, there was a consistent pattern in most lines. In 554 out of the 742 GUS expressing kernels (about 75%) the GUS staining was strongest in the outer part of the kernel and this pattern was therefore called “normal staining”. Of these, 234 kernels showed strong staining only in the outer part (Figure 2b), another 189 kernels had a strong staining in the outer part and a weak staining in the inner part of the endosperm (Figure 2c) and 131 kernels

showed weak staining only in the outer part of the endosperm (Figure 2a). In 145 out of the 742 GUS expressing kernels (about 20%) showed strong staining across the whole endosperm (Figure 2d), with no apparent spatial differences in expression. However, this may be due to limitations with the assay in strongly-expressing lines resulting in loss of discrimination between peripheral and central endosperm. The remaining 43 kernels (about 6%) showed partial, weak, and irregular staining.

Table 2 - Analysis of segregation and phenotypes of the T_0 transgenic lines. Percentages were calculated as the fraction of the total number of T_1 kernels where GUS staining was observed (row 4, numbers in bold). The different T_0 lines showed different segregation ratios and were classified into three segregation classes (1:3, 1:15 and non-Mendelian). The T_1 kernels showed different patterns of staining and were classified into five groups, where we considered the first three groups as"normal" staining.

	Segregation class			
	15:1	3:1	Non-Mendelian	*Total*
Numbers of T_0 lines	7	31	16	*54*
Numbers of T_1 seeds tested	168	714	384	*1266*
Numbers of T_1 seeds with no GUS staining observed	10	217	297	*524*
Numbers of T_1 seeds with GUS staining observed	*158*	*497*	*87*	*742*
Numbers of T_1 seeds with weak outer staining and no inner staining (Figure 2a)	9	76	46	*131*
Numbers of T_1 seeds with strong outer and no inner staining (Figure 2b)	29	193	12	*234*
Numbers of T_1 seeds with strong outer and low inner staining (Figure 2c)	48	137	4	*189*
Numbers of T_1 seeds with strong staining of the complete kernel (Figure 2d)	67	75	3	*145*
Numbers of T_1 seeds with partial, weak, and irregular staining (a-typical)	5	16	22	*43*
Numbers of "normal" staining (Figure 2a, 2b and 2c) of all T_1 seeds with GUS staining observed	86	406	62	*554*

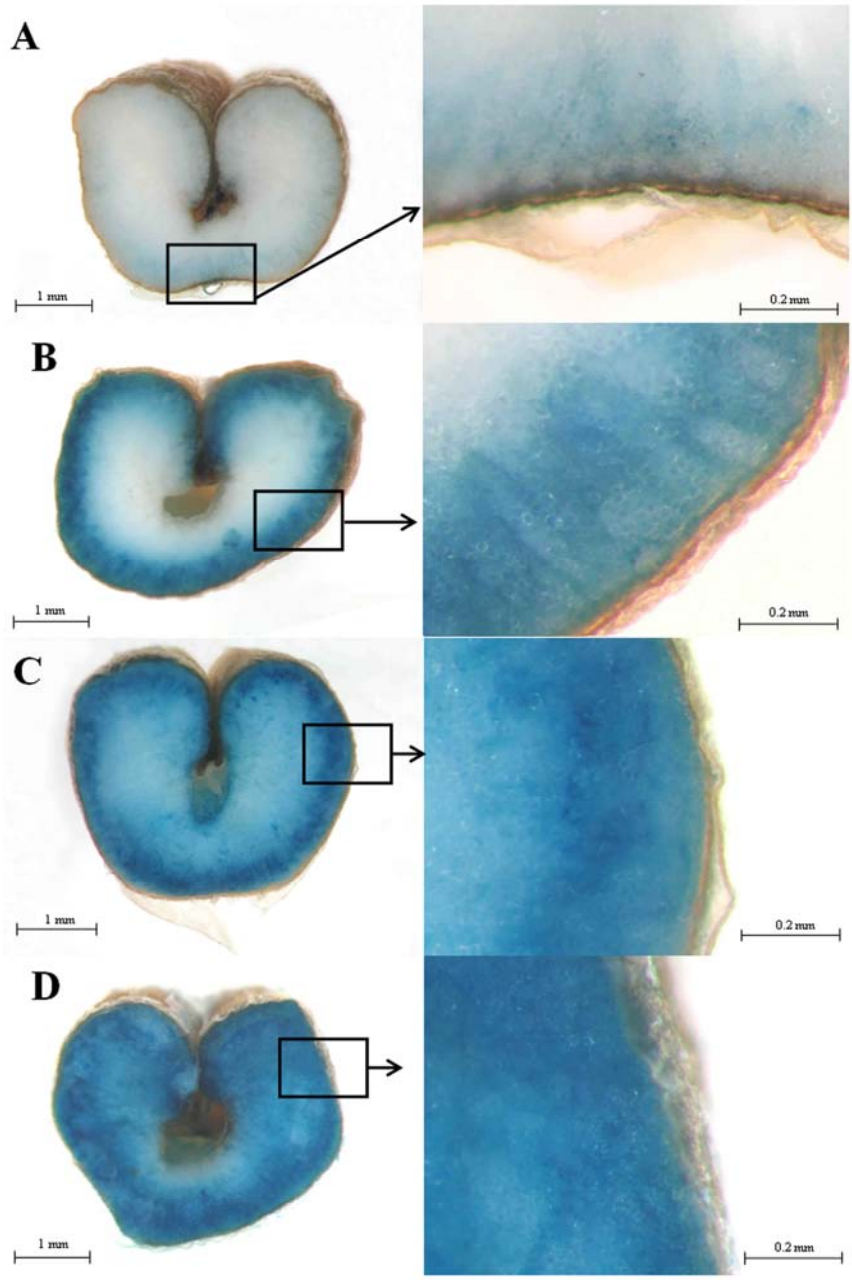

Figure 2 - Representative examples of four different phenotype classes based on differences in GUS expression levels in the mature T_1 kernels of different transgenic lines showing α-gliadin promoter activity. Left panel shows an overview of the kernel, right panel shows a five times magnification of the outermost cell layers of the grain. All samples were histochemically stained for 24 hours as described in material and methods. A: Low expression only on the outer part of the endosperm (18% of the kernels, picture of line B2219R2P5). B: Strong expression only in the outer part of the endosperm (31% of the kernels, picture of line B2219R1P12F) C: Strong expression in the outer part and a weak expression in the inner part of the endosperm (25% of the kernels, picture of line B2252R5P1) D: Strong staining across the whole endosperm (20% of the kernels, picture of line B2219R9P13C).

Correlation between phenotype and segregation

There appeared to be a relationship between the segregation ratio and intensity of expression (Table 2). Strong staining of the whole kernel (Figure 2d) was observed in only 3 of the 87 GUS expressing kernels in the non-Mendelian group (about 3%) but in 75 of the 497 GUS expressing kernels segregating in a 3:1 ratio (about 15%), and in 67 of the 158 GUS expressing kernels in the 15:1 group (42%). In contrast, weak and irregular staining was observed in 5 out of the 158 GUS expressing kernels in the 1:15 group (3%), in 16 out of the 497 GUS expressing kernels in the 3:1 group (3%), but in 22 out of the 87 GUS expressing kernels in the non-Mendelian group (25%). This indicates that expression pattern and intensity were related to differences in insertion patterns with multiple insertions (15:1 ratio) giving more intense staining and non-Mendelian events giving weak, irregular staining.

Detailed analysis of expression patterns

To determine tissue specificity, six T_3 lines were tested for GUS staining in different plant tissues. One of the lines (B2219R4P5.6.1) showed staining of the embryo, flower, pollen and at the connection point of the grain to the lemma in addition to the endosperm itself, but no GUS expression was observed in root, leaf and stem tissues. This expression pattern was only observed in the one line which was therefore excluded from further detailed analysis. The remaining five lines showed GUS staining only in endosperm tissue with no staining being observed in other tissues (stem, leaf, root, flower, pollen, embryo, lemma or testa).

To determine GUS expression during kernel development, developing grains were harvested from the central part of the head between 11 and 28 DAF (days after flowering). GUS staining was observed in half kernels from 11 DAF until maturity. The pattern was similar in all five lines and is illustrated by Figure 3 which shows developing grains at 14, 21 and 28 DAF at low magnification and by Figure 4 which shows a developing grain at 11 DAF and a mature grain at higher magnification.

Although these analyses show expression across the endosperm, the staining is less intense in the central than in the outer cells. The latter include the protein-rich subaleurone cells but also the cells of the aleurone layer, shown clearly in Figure 4.

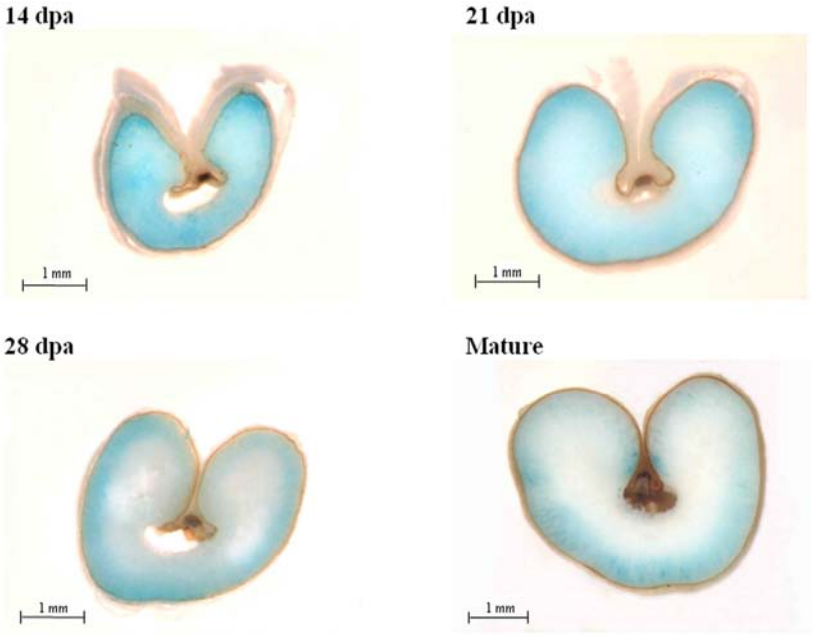

Figure 3 - GUS staining in developing kernels (18, 21, 28 DAF) and mature kernel of a representative homozygous T_3 transgenic line (B2219R1P10e.2.2).

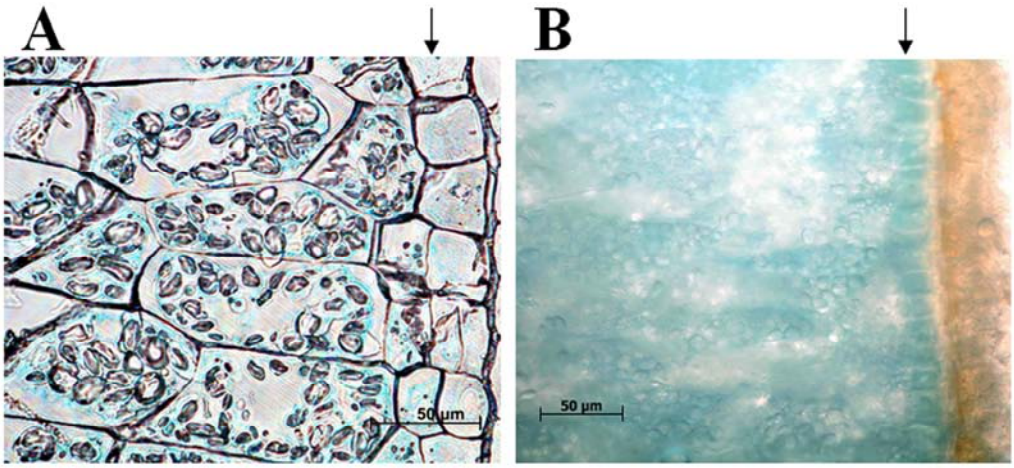

Figure 4 - GUS staining visualised at the cellular level on two different ages during development of homozygous T_3 kernels of 11 DAF and in the mature status. A: B2255R9P7, 11 DAF. Microscopic observation of 1 µm fixed and embedded section of a GUS stained kernels showing GUS staining in the subaleurone and aleurone cell layer of the wheat kernel.

B: B2219R1P10e, mature stage. Microscopic observation of a 200 µm-tick cryostat section of a mature GUS stained kernel showing clear GUS staining in the subaleurone and aleurone cell layer of the wheat grain. The arrow indicates the aleurone cell layer.

α-Gliadin labelling on thin (1 μm) sections and tissue prints

Although the promoter studies showed clear expression in aleurone cells of GUS driven by the α-gliadin gene promoter, this is not consistent with other reported studies which show that native gluten proteins are not present in this tissue [Fincher 1989]. To confirm the latter, we carried out immunolocalisation studies using the α-gliadin-specific Glia-α9 epitope antibody [Mitea *et al.*, 2008] and two different experimental approaches. Firstly, 1 μm sections of developing grains of 11, 18 and 25 DAF were fixed for conventional immunolocalisation using light microscopy (Figure 5). Secondly, proteins were transferred to nitrocellulose membrane by tissue printing (Figure 6). As a control, a monoclonal antibody specific for HMW subunits of glutenin [Mitea *et al.*, 2008] was also used on the thin sections of developing grains. Figure 5 compares immunostaining of thin sections with the α-gliadin-specific antibody (left) and HMW gluten antibody (right). This shows that α-gliadin and HMW subunits are present throughout the endosperm of the wheat kernel, including the subaleurone and central cells, and both are located in the protein bodies. However, in contrast to the GUS staining, no staining was observed in the aleurone cells. Similarly, immunostaining of tissue prints of mature grain (Figure 6) showed the location of α-gliadin in the subaleurone and starchy endosperm cells but not in the aleurone cell layer. Post-staining with Ponceau S revealed the aleurone cells that were not labelled with the anti-α-gliadin antibody (Figure 6c and d). To further confirm this, an antibody to 8S globulin, a protein present only in the aleurone cells and embryo of the grain (Figure 6) was used. This stained only the outer layer of the endosperm corresponding to the aleurone (Figure 6e) which was shown by double labelling to be outside the layer stained by the α-gliadin antibody (Figure 6f) [Wiley *et al.*, 2007].

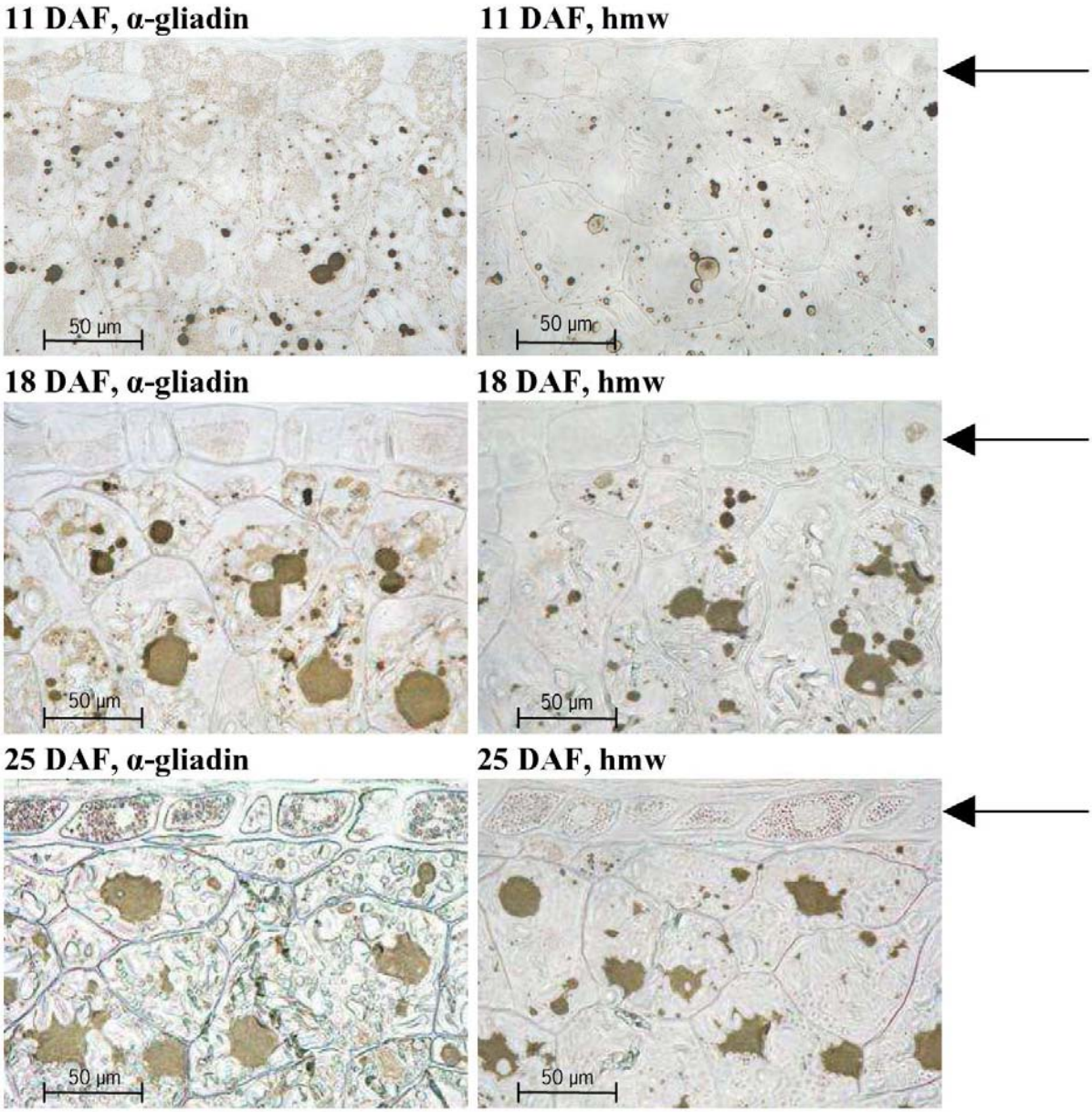

Figure 5 - Expression of α-gliadins and HMW glutenins in developing non-transformed wheat grains at 11, 18 and 25 DAF. Immunogold labelling was carried out on 1 µm sections using the anti α-gliadin and HMW glutenin specific antibodies. The figures on the left show immunogold labelling with an anti-α-gliadin antibody. The figures on the right show immunogold labelling with an anti-HMW glutenin antibody. At all stages only the protein bodies in the subaleurone cell layer are labelled and no staining in the aleurone cells was observed (arrows indicate aleurone).

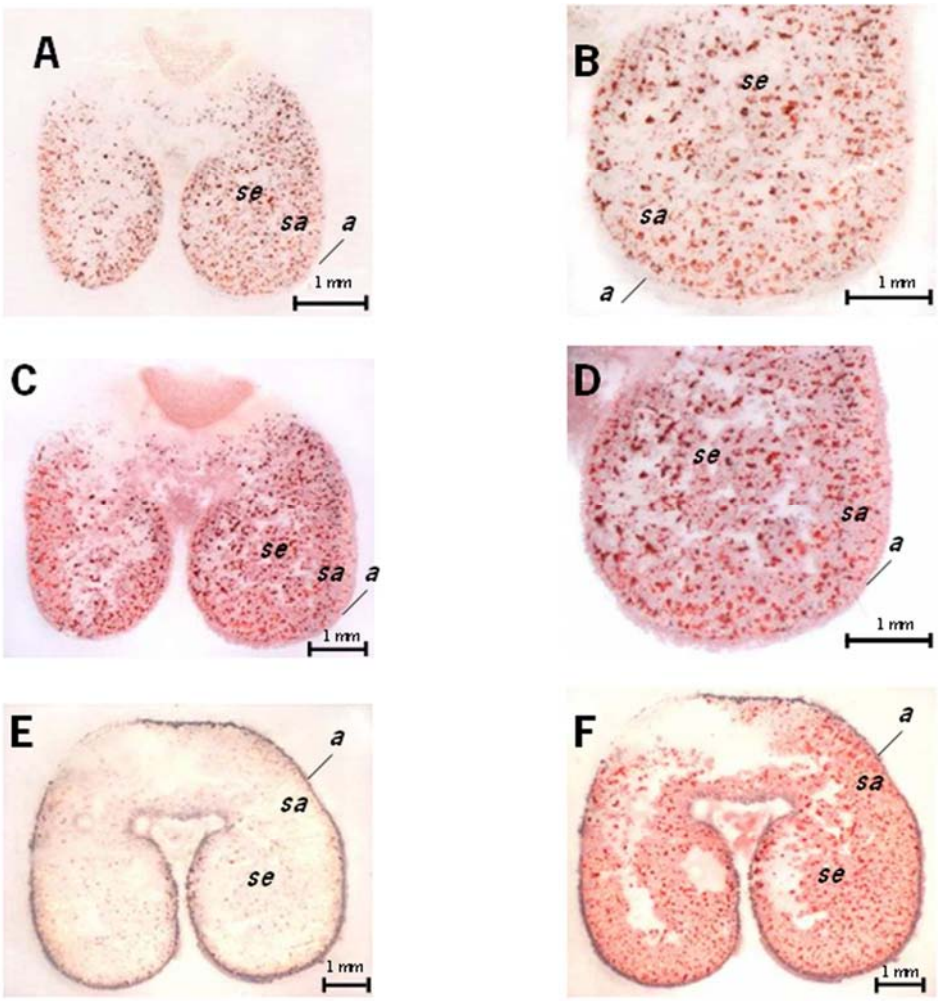

Figure 6 - Immunodetection of α-gliadin protein containing the Glia-α9 epitope and 8S globulin in tissue prints of mature wheat grains. A: Immunodetection of α-gliadin shown by the red stain in the starchy endosperm labelling B: Detail of figure 6A. Note the lack of labelling in the aleurone layer C: Same tissue print as in A labelled with the anti-gliadin antibody, and also post-stained with Ponceau S which stains the protein-rich aleurone pink and changes the colour of the starchy endosperm cells. D: Detail of figure 6C. Note the purple/pink stain in the aleurone layer not previously labelled in figures 6A and B. E: Immunodetection of 8S globulin shown by the clear purple labelling of the aleurone layer. F: The same print as in figure 6E, but also post-stained with Ponceau S, showing the presence of protein in the endosperm not previously labelled with the anti-8S globulin antibody. a=aleurone, sa=sub-aleurone; se=starchy endosperm. Bars: 1mm.

Discussion

Genome specificity and epitope content

The promoter sequence used in this study was obtained by Reeves and Okita in 1987. Based on sequence homology, we concluded that this is not the promoter sequence of an A genome gene as suggested by Reeves and Okita [1987] but of a gene from the B genome (the *Gli-2* locus on chromosome 6BS). It is not known if the promoters of α-gliadin genes originating from the A, B and D genomes differ in their expression patterns. Van Herpen *et al.* [Chapter 2] and Molberg *et al*. [2005] showed that a number of α-gliadins from the B genome do not contain any of the four identified α-gliadin T-cell toxic CD-epitopes (Glia-α, Glia-α2, Glia-α9 or Glia-α20) and these epitopes were also absent from the α-gliadin encoded by the gene used in this study.

Location of α-gliadin promoter activity during wheat kernel development

Expression of the GUS reporter gene driven by the α-gliadin promoter in transgenic wheat grains was observed as early as 11 DAF. This is in agreement with the previous report that gliadin transcripts began to accumulate in the endosperm by 9 DAF [Drea *et al.,* 2005]. The expression became more prominent in the outer cells of the kernel as maturity progressed. The most frequently observed phenotype in the mature T_1 kernels (75% of the GUS-positive kernels) was that GUS staining was strongest in the outer part of the kernel, and weaker in the central starch-rich endosperm cells. Previous studies by Evers and Kent [Evers, 1970; Kent, 1966] showed that all starchy endosperm cells appear to contain similar total amounts of protein but that the accumulation of starch occurs mainly in the central starchy endosperm cells. As a consequence, protein may represent up to half of the total cell mass in the subaleurone cells, but only 10% of the total mass of the starch-rich central endosperm cells [Evers, 1970; Kent, 1966]. The most frequently observed expression pattern of the α-gliadin-promoter is consistent with a relatively greater accumulation in the outer layers as a consequence of dilution of the protein in the central cell by starch granules. This is similar to the reported expression pattern of a LMW glutenin promoter [Stoger *et al.,* 2001].

Correlation between phenotype and segregation

A correlation was observed between the segregation ratios and the phenotype, with a higher proportion of the T_1 kernels showing strong expression in lines segregating in a 15:1 ratio compared to 3:1 or non-Mendelian ratios. This is consistent with two transgene insertion sites and possibly a higher number of transgene copies, although transgene copy number was not measured here. Some atypical expression patterns, often with weak staining, were also observed in T_1 kernels from a few lines, particularly those showing non-Mendelian segregation ratios. The observation of non-Mendelian segregation may be due to expression levels being under the detection limit of the GUS staining procedure leading to an underestimation of true segregation ratios in these lines and patchy light staining. Alternatively, rearrangements of the transgene cassette during biolistic transformation may have resulted in truncation or rearrangements of functional promoter elements resulting in loss of the typical expression pattern seen in the majority of lines.

Cellular location of the α-gliadin and HMW glutenin protein

The demonstration of GUS expression in the aleurone cells of the transgenic plants expressing the α-gliadin promoter-GUS construct was unexpected as gluten protein gene expression has not previously been reported in this tissue. Lamacchia *et al.* [2001] and Stoger *et al.* [2001] reported similar studies, in which the GUS reporter gene was fused to glutenin subunit promoters and expressed in transgenic wheat: the HMW subunit 1Dx5 promoter in durum wheat [Lamacchia *et al.*, 2001] and a LMW subunit promoter in bread wheat [Stoger *et al.*, 2001]. Both promoters drove GUS expression in the starchy endosperm cells, including the subaleurone, but not in the aleurone layer. Similarly, the *in situ* hybridisation study of Drea *et al.* [2005] also showed α-gliadin transcripts in the subaleurone and central starchy endosperm but not in the aleurone.

To confirm our results we therefore used a specific monoclonal antibody to the Glia-α9 epitope of α-gliadin for immunolocalisation studies, using a specific monoclonal antibody to HMW glutenin subunits as a control [Mitea *et al.*, 2008]. These antibodies showed similar labelling to the starchy endosperm cells including the subaleurone but not to the aleurone cells.

This is consistent with the generally accepted view that aleurone cells do not contain gluten proteins, as reviewed by Fincher [1989].

The aleurone and starchy endosperm cells arise from the same cell lineage with some of the starchy endosperm cells being derived from cell divisions in the aleurone which acts as a meristem. It is therefore possible that (1) additional regulatory elements are present in the native α-gliadin promoter configuration that prevent expression in the aleurone cell layer, and (2) the lack of these elements in the 593 bp promoter used here resulted in aleurone mis-expression. Of thirteen promoter sequences from the NCBI database (Table 1), nine start at -593 bp, probably because of the presence of a Pst 1 restriction site that was used while cloning these genes. Analysis of the four remaining sequences of the region upstream -593 bp failed to identify a CGN-4-like motif or prolamin box. A RY repeat (CATGCAC) [Fujiwara and Beachy, 1994] was present at -774 bp in the original promoter (M16496), but not used in the 593 promoter region of this study. RY repeats were described to play a key role in the seed-specific regulation of seed storage protein genes in leguminous species. A deletion of part of the RY repeat in the *Vicia faba* legumin promoter [Baumlein *et al.,* 1992] resulted in expression in leaves of stably transformed tobacco plants whereas a legumin promoter with an intact RY repeat was primary expressed in seed tissue. This indicates that this *cis*-acting element represses expression in tissues other than seed tissue. The RY-element could therefore also repress expression of the α-gliadin gene in the aleurone cells. The RY element was also present in a number of the other α-gliadin promoter sequences, but located much closer to the transcription start site (position -294) (Table 1).

Regulatory motifs in α-gliadin promoters from different genomes

We also observed an association between the genome of origin and the presence of endosperm-specific promoter motifs in the 593 bp promoter region. The GCN4-like motif at -492 bp was only present in the genes originating from the B genome and not in the sequences originating from the A genome (Table 1). Aryan *et al.* [1991] described a 20% increase in promoter strength in the -592 to -448 region and the authors suggested the presence of an activating element in this region. The CGN4-like motif present at -492 is a candidate for such an activating element. These results indicate that differences in regulation of expression may exist between different members of the

α-gliadin gene family with the genome of origin being important. The RY repeat at position -294 was present in seven out of the eight A genome sequences and in none of the B genome sequences. The RY repeats have been proposed to have a key role in seed specific gene regulation [Baumlein *et al.,* 1992]. When analysing for sequence identity of the region upstream -593 bp we observed an association between genome of origin and sequence identity.

The differences in regulator motifs between sequences originating from the A or B genome suggest that the regulation of expression between α-gliadin genes present on chromosomes 6A and 6B may be different. Kawaura *et al.* [2005] observed that early α-gliadins gene expression were preferentially from the D genome, and late expression from the A genome. However, further research is required to determine whether this can be exploited to manipulate the proportion of α-gliadin encoded by the *Gli 2* loci on the different genomes of wheat to decrease the toxicity in celiac disease.

Conclusions

We have studied the expression of a specific B genome-encoded α-gliadin gene promoter fragment and the deposition of α-gliadin proteins. The promoter was active in the cells of the starchy endosperm, the subaleurone and the aleurone cell layer. No expression of GUS was found in other tissues. However, α-gliadin protein was detected together with HMW subunit protein in the protein bodies of the starchy endosperm and subaleurone cells but was not detected in the aleurone cells. The α-gliadin promoter was active from 11 DAF until maturity and showed a similar pattern of expression to the LMW subunit promoter. A large number of different transgenic plants with the same construct revealed differences in the observed phenotypes of expression. An association between segregation ratios of the transgene and the different phenotypes of expression was observed. These results emphasise the importance of testing a number of different transgenic lines when studying the effects of transgenes. Our results indicate that additional regulator elements upstream of the promoter region used may specifically repress expression in the aleurone cell layer. Observed differences in expression regulator motifs between

the α-gliadin genes on the different genomes (A and B) of bread wheat open the possibility of manipulating the balance between the α-gliadins from the different genomes in bread-wheat.

Acknowledgements

This research was partially funded by the Celiac Disease Consortium, an Innovative Cluster approved by the Netherlands Genomics Initiative and partially funded by the Dutch Government (BSIK03009). Microscopy sample preparation and part of the imaging was undertaken at the Rothamsted Research Bioimaging Unit, Harpenden, UK. We are grateful to Dr Paola Tosi at Rothamsted Research for her help during discussions on this work to Prof. Dr. Thomas W Okita (Washington State University) for providing the pCR1 construct and for further advice and discussion.

Chapter 4

The origin and early development of wheat glutenin particles

Teun WJM van Herpen, Jan HG Cordewener, HJ Klok, Jacky Freeman, Antoine HP America, Dirk Bosch, Marinus JM Smulders, Luud JWJ Gilissen, PR Shewry, Rob J Hamer

Abstract

Breadmaking quality is strongly related to the glutenin macropolymer (GMP) fraction. Don and co-workers [2003a] showed that GMP consists of spherical glutenin particles and suggested that these originate from the protein bodies (PBs) observed in developing grain. We have tested this hypothesis by systematically comparing SDS-soluble and SDS-insoluble protein fractions from both PBs and flour. These preparations were analysed at the molecular, oligomer, particle and microscopic levels. Comparison of PBs isolated from immature seeds with glutenin particles isolated from mature seeds revealed strong similarities in protein composition and the presence of large glutenin oligomers. However, the glutenin particles from mature wheat were significantly larger than PBs. We suggest that PBs are the building blocks for the formation of much larger glutenin particles which are formed during the desiccation phase of kernel development.

Introduction

Gluten polymers from wheat are among the largest and most complex protein networks in nature with MWs of more than 10 million [Wieser, 2007]. The differences in breadmaking quality between flours of different wheat varieties also parallel differences in a gluten protein fraction called glutenin macropolymer (GMP) [Graveland *et al.,* 1982, 1984; Weegels *et al.,* 1996, 1997]. GMP is the glutenin fraction which is insoluble in various solvents (SDS or acetic acid) [Weegels *et al.,* 1996, 1997] and consists of spherical glutenin particles [Don *et al.,* 2003a]. These glutenin particles also vary in size, and this variation correlates with a key technological quality parameter: dough mixing properties [Don *et al.,* 2005b]. The genetic background and growing conditions also affect the quality of GMP, the glutenin particle size and, consequently, flour quality, but the mechanism behind this effect is not known [Don *et al.,* 2005a]. However, Don and co-workers [Don *et al.,* 2003a] suggested that the glutenin particles in GMP are related in origin to the PBs observed in developing wheat endosperm.

The wheat gluten proteins can be classified into three groups on the basis of their structural and evolutionary relationships. These are sulphur-rich (S-rich), sulphur-poor (S-poor) and high molecular weight (HMW) groups, with a number of subgroups within the S-rich and S-poor groups. However, this classification does not correspond directly to the polymeric and monomeric fractions in the wheat kernel (glutenins and gliadins, respectively) [Kohler *et al.,* 1993], with the HMW subunits and the S-rich LMW subunits being the main contributors to glutenin particles.

Wheat seed storage proteins are synthesised in developing grain and are initially deposited in discrete protein bodies (PBs) [Shewry and Halford 2002]. All gluten proteins are initially synthesised on the rough ER and co-translationally translocated into the lumen. However, it appears that some components, probably mainly gliadins, are transported via the Golgi to the protein storage vacuole whereas others, probably mainly glutenins, are deposited within the lumen of the ER to form a second population of ER-derived protein bodies [Kim *et al.,* 1988; Rubin *et al.,* 1992]. It has also been suggested that these ER-derived protein bodies are subsequently absorbed by protein storage vacuoles in a process analogous to autophagy [Galili 1997; Shewry

and Halford, 2002; Shewry, 1999]. This leads to the formation of an apparently continuous gluten protein matrix in the cells of the mature grain [Parker 1980]. The PBs in developing grains also contain dark-staining inclusions of the globulin storage protein triticin [Bechtel *et al.*, 1991].

The accumulation of wheat seed storage protein commences as early as seven days after flowering (DAF) and ends at the beginning of the desiccation phase. In contrast, a close correlation was found between the accumulation of the GMP and the rapid loss of water during desiccation [Carceller and Aussenac 1999, 2001]. The percentage of SDS-insoluble polymers as a proportion of total polymers can increase from less than 10% at the end of kernel ripening to 50% in as little as 10 days, although the synthesis of gluten proteins is completed before this occurs [Carceller and Aussenac 1999]. Since glutenin insolubility is related to aggregate size, this indicates a lower degree of aggregation within PBs from immature wheat grain compared to within the gluten matrix of the mature wheat grain.

It has been recently suggested that the glutenin particles isolated from mature wheat flour and observed as spherical shaped particles using Confocal Laser Scanning Microscopy (CLSM) originate from the individual PBs as observed in immature wheat endosperm [Don *et al.*, 2003a]. If this is true, the PBs and glutenin particles should exhibit similarities at different levels: subunit composition, polymer composition, polymer size and polymer network properties.

In this study we have tested this hypothesis by comparing PBs from immature wheat with flour from mature grain of the same variety. The preparations from mature and immature grains were also compared for their ability to form a physical gel and for their composition of protein subunits and oligomers. Finally, their particle sizes were determined by laser diffraction and CLSM was used to study their compositions using specific stains for protein and free sulfhydryl groups.

Our results support the hypothesis that glutenin particles originate from the PBs. Additionally we suggest that further aggregation of individual PBs is required to generate the large particles observed in GMP.

Experimental

Wheat material used for protein body isolation was *Triticum aestivum* cv Cadenza of approximately 15 DAF (grown under controlled environment conditions at Rothamsted Research, Harpenden,UK). Flour was obtained by grinding mature dry grains from the same cultivar grown under the same conditions with an A11 Basic IKA Analytical Mill (IKA -WERKE GmbH & Co KG).

Protein body and protein body gel isolation

Endosperms were removed from all of the individual developing grain from 10 complete immature wheat spikes of *Triticum aestivum* cv Cadenza at approximately 15 DAF and chopped with a razor blade in 20 ml buffer 1 (20 mM, HEPES, pH7.6, 100 mM NaOAc, 5 mM $MgCl_2$). The homogenate was filtered through four layers of cheese cloth, pre-wetted with buffer 1. The residue was washed with 10 ml buffer 1 and the filtered homogenate was layered on top of a 1.75 M sucrose cushion in buffer 1 and centrifuged at 500 g for 2 minutes at 10°C. The material from the top of the cushion was collected and resuspended in 20 mM HEPES, pH7.6, 100 mM NaOAc, 5 mM EDTA, and 0.25 M sucrose, to a total volume of 10 ml. A two-step Percoll density gradient was prepared: a 1.13 g/ml and a 1.08 g/ml. The suspension was layered onto the two-step Percoll gradients and centrifuged at 7100 g at 10°C for 60 minutes. The PBs were collected in 0.5 ml volume from the surface of the 1.13 g/ml Percoll layer. The same volume of demineralised water was added to the PBs and the mixture vortexed briefly and then centrifuged for 2 minutes at 16,000 g. The supernatant was removed and the pellet was resuspended in a total of 1 ml water. This was vortexed and centrifuged as before. Finally, the PB pellet was resuspended in an appropriate volume of demineralised water for immediate use. Further details are provided elsewhere [Davy *et al.*, 2000]

SDS-insoluble protein was isolated by dissolving the PB pellet in 1.5% (w/v) SDS and centrifugation for 10 minutes at 16,000 g at room temperature. The supernatant containing SDS-soluble proteins was decanted and kept for later analysis while the SDS-insoluble protein was recovered as a gel (PB-gel).

Isolation of SDS-insoluble (GMP) and SDS-soluble wheat protein from mature grain

0.5 g flour of cv Cadenza was dissolved in 10 ml 1.5% (w/v) SDS. Ultracentrifugation was performed at 78,000 g at 20°C for 30 minutes (Centrikon T-2060, Kontron Instruments, USA). After ultracentrifugation, GMP was observed as a gel on top of a starch layer. The supernatant containing SDS-soluble protein was decanted and retained while the gel layer (GMP) was collected.

SE-HPLC analysis of SDS-soluble proteins (Profible method)

10 mg of PB pellet and 0.8 mg of flour was suspended in 1 ml 1% (w/v) SDS, 0.1 *M* sodium phosphate buffer (pH 6.9). Controlled ultrasonication with a sonicator (Vibra Cell 72434, Bioblock, Illkirch, France) delivering ultrasonic vibrations at 20 kHz and equipped with a 3 mm diameter tip probe was performed for 2 minutes without interruption (no overheating was observed), prior to separation by size-exclusion HPLC using a TSKgel G 4000 SW column (7.5 mm x 30 cm resolving column, Sigma Cat No 805790 and a 7.5 mm x 7.5 cm guard column, Sigma Cat No 805371). Further details are provided elsewhere [Dachkevitch and Autran, 1989, Millar, 2003; Morel *et al.*, 2000].

SE-FPLC chromatography of SDS-soluble proteins

Size-exclusion FPLC was performed using a Superose 6 HR 10/30 column (GE Healthcare, UK) pre-equilibrated with 1% (w/v) SDS in 0.1 M sodium phosphate buffer (pH 7). Aliquots (0.1 ml) of SDS-soluble protein from flour and from PBs were loaded on the column without prior sonication and fractionated using a flow rate of 0.3 ml per minute. Eluate fractions of 0.5 ml were collected for SDS-PAGE analysis. The elution pattern was monitored at 280 nm.

SDS-PAGE gel electrophoresis

Fractions from SE-FPLC were analysed on a Multiphor II horizontal SDS electrophoresis system using ExcelGel SDS gradient 8-18 gels (GE Healthcare, UK). Proteins were precipitated using 10% (w/v) trichloroacetic acid (TCA) and after washing with ice-cold acetone, the air dried pellets were dissolved in sample buffer (2% (w/v) SDS, 62.5 mM Tris-HCl pH 6.2, 10% (w/v)

glycerol, 0.01 mg/ml bromophenol blue) with and without 100 mM DTT as a reducing agent. Proteins were visualized using silver staining [Rabilloud *et al.*, 1988]. Protein body gel and fresh GMP fractions were also analysed on an ExcelGel SDS gradient 8-18 (GE Healthcare, UK) without prior precipitation.

Protein identification using nano-liquid chromatography – mass spectrometry

The bands of interest were excised the SDS-PAGE gel and sliced into 1 mm^3 pieces. Gel pieces were processed essentially according to Shevchenko [1996]. Tryptic digests were then analysed by one-dimensional LC-MS using an Ettan^{TM} MDLC system (GE Healthcare, UK) in a high-throughput configuration directly connected to a Q-TOF-2 Mass Spectrometer (Waters Corporation, Manchester, UK). Samples (5 µl) were loaded on 5 mm x 300 µm ID Zorbax^{TM} 300 SB C18 trap columns (Agilent Technologies, USA), and the peptides were separated on a 15 cm x 100 µm ID Chromolith CapRod monolithic C18 capillary columns at a flow rate of approximately 1 µl/min. Solvent A contained an aqueous 0.1% FA solution and solvent B contained 84% ACN in 0.1% FA. The gradient consisted of 5% B for 10 min, a linear gradient to 30% B over 40 min, a linear gradient to 100% B over 10 min, and then a linear gradient back to 5% B over 5 min.

MS analyses were performed in positive mode using ESI with a NanoLockSpray source. As lock mass, [Glu^1]fibrinopeptide B (1 pmol/µl) (Sigma, USA) was delivered from a syringe pump (Harvard Apparatus, USA) to the reference sprayer of the NanoLockSpray source at a flow rate of 1 µl/min. The lock mass channel was sampled every 10 seconds. Data dependent acquisition (DDA) was performed with the Q-TOF-2 operating in MS/MS mode. MS/MS data collected during a LC-MS/MS run were submitted to database search using the Protein Lynx Global Server V2.2.5 (Waters Corporation, USA), taking fixed (carbamidomethyl) and variable (oxidation of methionine) modifications into account. MS/MS spectra that led to no identification in the NCBI non-redundant protein database, were compared to translated wheat triticin EST sequences obtained from the NCBI database.

Particle size distribution

Freshly prepared wet GMP from flour was dispersed in 1.5% (w/v) SDS (10 ml) by rotating overnight. The PB-gel was also dispersed in 1.5% (w/v) SDS (2 ml) by rotating overnight. Intact PBs were dispersed in water and used directly. Particle size distributions were measured using a Mastersizer 2000 (Malvern Instruments, UK). The laser diffraction pattern obtained with the instrument can be transformed into particle size distribution using the Fraunhofer theory assuming a spherical particle shape. The range of the instrument is 0.02–2000µm. Dispersions were transferred to the suspension water in the sample vessel. The obscuration signal for diffraction was approximately 10%. Data were exported to Excel to perform calculations as described by Wang and co-workers and by Don and co-workers [Wang *et al.,* 2005, Don *et al.,* 2005a].

Confocal scanning laser microscopy (CLSM)

Freshly prepared wet GMP from flour was dispersed in 1.5% (w/v) SDS (10 ml) by rotating overnight. Protein body gel was also dispersed in 1.5% (w/v) SDS (2 ml). Intact PBs were dispersed in water and used directly. Samples were stained for protein with FITC (a fluorescent label, specific for proteins) and with Nile Red for hydrophobic molecules. Free SH groups were stained with the fluorescent marker Alexa Fluor 568 C5 maleimide (Invitrogen, USA, cat nr A20341). In addition, the protein bodies were extracted with a detergent (in this case 0.02% (w/v) SDS or 1% (v/v) Triton X-100) after labelling with FITC and Alexa Fluor 568. Samples were observed using a Leica TCS NT confocal laser scanning microscope (model Leica DM IRBE).

Results

Breadmaking quality relates to the capacity of GMP formation which consists of spherical glutenin particles [Don *et al.,* 2003b]. It was further suggested that these glutenin particles originate from PBs. We have tested this hypothesis by systematically comparing the properties and compositions of PBs from developing grain and GMP from flour.

GMP was isolated from mature wheat (cv Cadenza) flour using detergent solution (1.5% (w/v) SDS) and ultracentrifugation while PBs were isolated from immature (15 DAF) grains of the same wheat variety and purified using Percoll density gradient centrifugation. Extraction of the protein body preparation with SDS followed by centrifugation led to the formation of a gel which had a similar appearance to the GMP gel isolated from mature wheat (Figure 1). SDS emulsified the lipid membrane surrounding the protein body but left discrete bodies. Wellner *et al*. [2005] performed a similar experiment using chloroform/methanol to remove the lipid membrane but in contrast, this resulted in the fusion of the contents to form a cohesive, elastic protein mass.

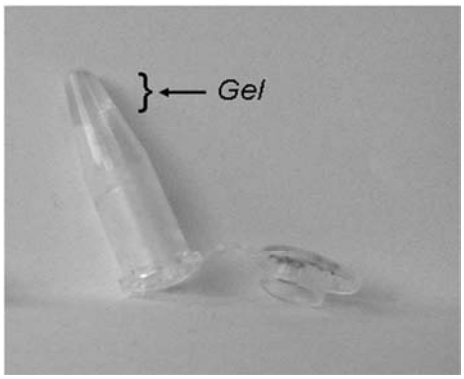

Figure 1 - The SDS-insoluble protein from protein bodies form a gel-like pellet after centrifugation.

SDS-soluble proteins in PB from developing grain and in flour from mature grain

The ability to form a SDS insoluble gel indicates the presence of polymers in the PB and GMP samples. More information about proteins in these preparations was obtained using SE-HPLC. This method uses sonication to shear large polymers and render them soluble. Consequently, although the profiles are reproducible and relate to quality parameters, they do not represent the precise molecular mass distributions of the unextracted polymers. The separation profiles obtained using the SDS-soluble proteins in PBs from developing grain (Figure 2A) and the SDS-soluble proteins in flour from mature grain (Figure 2B) were essentially similar. Five fractions could be distinguished, which represent HMW glutenin polymers (F1), LMW glutenin polymers (F2), gliadins (F3 and F4) and albumins and globulins (F5), respectively. Apart from differences relating to the amount of protein applied to the column, the profile obtained with the protein body

sample is similar to the mature wheat grain sample except that the proportion of the F4 fraction is lower for the protein body proteins. This indicates that the PB sample contains relatively less monomeric gliadin proteins and suggests that it is enriched in protein bodies derived from the ER which are enriched in glutenins [Kim *et al.,* 1988; Rubin *et al.,* 1992]. In contrast, the GMP fraction from flour corresponds to the matrix proteins which comprise both ER and vacuolar-derived PBs.

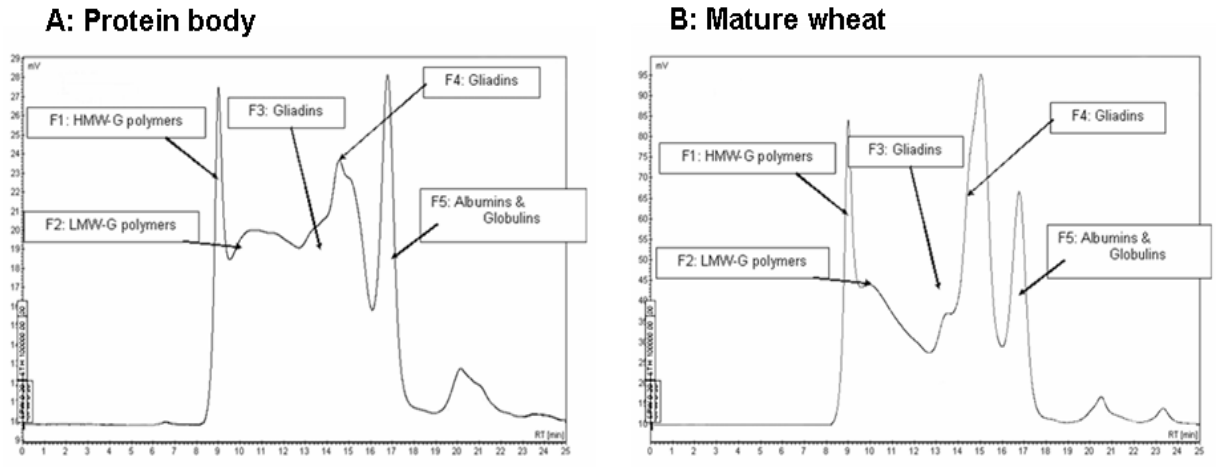

Figure 2 - HPLC profiles of SDS-soluble proteins prepared from protein bodies (A) and mature wheat grain (B). Five fractions can be distinguished. Fraction 1 (F1) eluting at approximately 9 minutes was identified as HMW-glutenin polymers, F2 eluting from 10 till 13 minutes was LMW-glutenin polymers, F3 from 13 till 14 minutes and F4 from 14 till 18 minutes were gliadins and F5 from 16 till 18 minutes was albumins and globulins. The fraction from protein bodies is characterized by a relatively low proportion of gliadin proteins. (note: voltage range (mV) is 3 times lower in protein body sample).

Composition of SDS-soluble polymeric and SDS-insoluble gel fractions

The polymeric and oligomeric glutenin fractions (F1 and F2) would be expected to contribute mainly to gel formation observed in both PB and GMP. These fractions were therefore analysed in more detail with respect to their protein composition. SE-FPLC was used to prepare comparable fractions as obtained with SE-HPLC (Figure 2). The F1 and F2 fractions were

analyzed with SDS-PAGE under reducing and non-reducing conditions. Under non-reducing conditions a smear was visible in all lanes, whereby fractions eluting later from the size-exclusion column showed the highest mobility in the gel (Figure 3A, 3C). The reduced fractions (Figure 3B, 3D) show the presence of both HMW glutenin subunits (between 150 and 75 kD) and LMW glutenin subunits/gliadins. A typical difference between the protein patterns of the SDS-soluble preparations from mature and immature grain is the presence in the latter of some low molecular weight bands at approximately 20 kD. This agrees with Wellner *et al.* [2005] who also observed comparable low molecular weight bands in the PB preparations but not in the sample prepared from wheat flour.

To compare the proteins present in the SDS-insoluble gel fractions from mature and immature grain, SDS-insoluble gel was boiled in SDS sample buffer in the presence and absence of the reducing agent, DTT. Under non-reducing conditions the solubilised polymers were not able to enter the gel except for a small fraction of the gel from immature grain, indicating that at least some of the polymeric components in immature grain were smaller than those in mature grain. However, the two SDS-insoluble fractions show similar patterns of subunits to the SDS-soluble polymers after reduction. It is of interest that the protein band of approximately 20 kD was not observed in the gel fractions from either the immature (Figure 3E) or the mature (Figure 3F) samples.

Identification of Triticin in the SDS-soluble fraction of PBs of immature wheat

Three protein bands at an estimated molecular mass of 20 kD in the SDS-PAGE gel, (indicated by an arrow in Figure 3B) were cut out of the gel and after tryptic digestion analyzed by LC-MS/MS. All three bands were identified as the small basic subunit of triticin, yielding between 8 and 11 peptides that could be assigned to the EST sequence CJ634623 (Figure 4). In addition to pure tryptic peptides (tryptic recognition at K/R-X, with X is not P), two half tryptic peptides were detected due to aspecific cleavage, i.e. VHLYQNAIISPLWNINAH.. and …VIANAYGISR. Furthermore, apart from the N-terminal peptide GLEENFCDHK, a truncated form was observed missing the first two amino acids G and L. As no clear differences in peptide composition of the tryptic digests of the three individual bands were found, the reason for the differences in mobility

on SDS-PAGE could not be explained by differences in amino acid sequence. Overall we found a sequence coverage of 67% (131 out of 195 amino acids) of the predicted EST protein sequence (Figure 4).

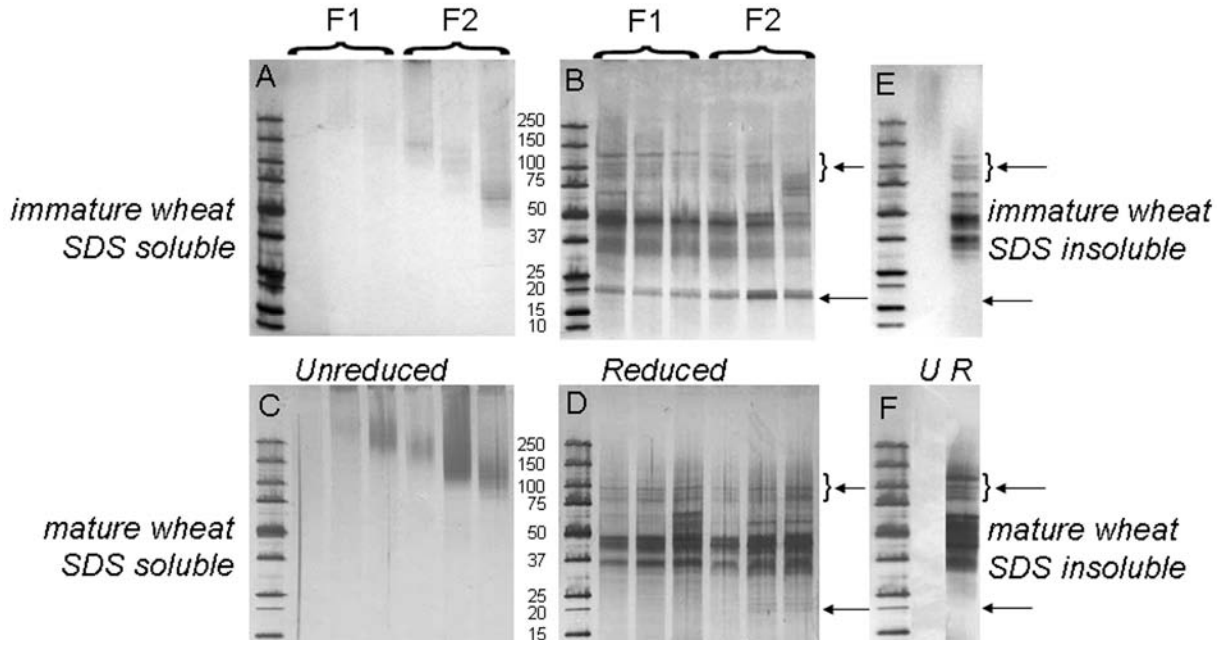

Figure 3 - SDS-PAGE of SDS-soluble and SDS-insoluble protein fractions from wheat. The SDS-soluble fractions collected after SE-FPLC fractionation were analyzed on 8-18 % gradient gels under non-reducing (A, C) and reducing (B, D) conditions. Under non-reducing conditions the protein complexes showed a gradual decrease in molecular mass when eluting later from the size-exclusion column. Under reducing conditions the different subunits of the polymers could be visualized, showing the presence of triticin in all fractions obtained from immature wheat (the lower arrow in B). The higher arrows indicate the HMW glutenin subunits. A) SDS-soluble FPLC fractions from immature wheat under unreduced conditions, and B) reduced conditions. C) SDS-soluble FPLC fractions from mature wheat under unreduced conditions, and D) reduced conditions. E) SDS-insoluble proteins from immature wheat under unreduced and reduced conditions. F) SDS-insoluble protein from mature wheat under unreduced and reduced conditions.

GLEENFCDHK/**LSVNIDDPSR**/**ADIYNPR**/AGTITR/**LNSQTF**
PILNIVQMSATR/**VHLYQNAIISPLWNINAH**SVMYMIQGHIWV
QVVNDHGR/**NVFNDLLSPGQLLIIPQNYVVLK**/K/AQR/DGSK
/**YIEFK**/**TNANSMVSHIAGK**/NSILGALPVD**VIANAYGISR**/T
EAR/SLK/FSR/**EEELGVFAPK**/**FSQSIFR**/SFPNGEEESS

Figure 4 - Tryptic peptides of wheat triticin identified by LC-MS/MS. After tryptic digestion of the 20 kDa protein bands indicated in Fig. 3, the resulting peptides were fractionated by nano-LC and on-line sequenced by Q–TOF MS/MS. The amino acid sequences given in bold were observed in the tryptic digests, and covered 67% of the predicted protein sequence of the wheat EST, CJ634623.

Morphological appearance of the PBs and the effect SDS

The PB preparations were analysed by CLSM microscopy, using specific fluorescent labels for protein, hydrophobic groups and free sulfhydryl groups. In addition, the labelled PBs were extracted with a detergent to observe effects on protein solubility and free SH groups under conditions similar to those used for GMP extraction.

Double staining to show protein and protein hydrophobic groups (Figure 5A) revealed filled spheres with a size around 10 μm. The protein spheres were accompanied by inclusion bodies which stained with Nile Red indicating hydrophobic properties. Double staining for protein and free SH groups (Figure 5D) showed the presence of free SH groups throughout the protein body. The hydrophobic inclusion bodies were rapidly dissolved when the preparations were treated with SDS, either by passive diffusion (Figure 5B) or with shaking to form an emulsion (Figure 5C). This occurred within a few seconds and could be observed in real time under the CLSM microscope (supplementary data). After staining for protein and hydrophobic groups, SDS was added and the samples were shaken (inputting energy), resulting in dispersion (Figure 5C). Similar results were obtained when Triton X-100 was used instead of SDS, and indicate that the PBs observed in Figure 5B are stabilised mainly by non-covalent interactions. The same procedure was applied to PBs which were stained for free SH groups before adding SDS and shaking (Figure 5E). In this case particles that were stable in a SDS solution were observed, and these appeared to consist of aggregates of several individual PBs.

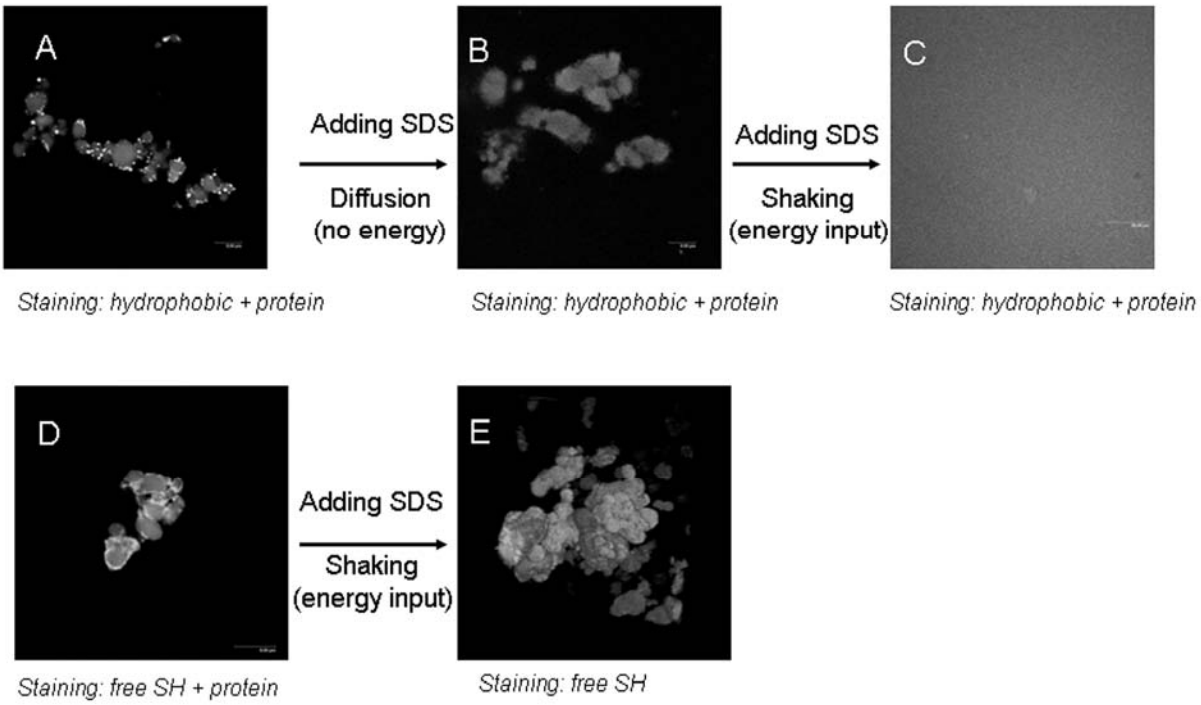

Figure 5 - CSLM images of protein body preparations after different treatments. A) Double stained (FITC and Nile Red) protein bodies; protein is grey, hydrophobic areas are white. B) Double stained (FITC and Nile Red) protein bodies after treating with SDS by diffusion; protein is grey, hydrophobic areas are white. C) Double stained protein bodies (FITC and Nile Red) after adding SDS and input of energy (shaking), protein is grey , hydrophobic areas are white;. D) Protein bodies stained with Alexa Fluor (which binds to free SH groups) and FITC, protein is grey while the white areas are stained with Alexa Fluor. E) Protein bodies pre-stained with Alexa Fluor and FITC and addition of SDS and input of energy (shaking).

Size distribution of the PBs and GMP particles

The particle size distribution of the protein bodies observed by CLSM was determined using laser diffraction (Figure 6). This showed that the weighted mean diameter (D) [4,3] of the untreated preparations was 11.7 +/- 2.0 μm, while the value for the insoluble fraction isolated from protein bodies (Protein body gel) was 12.3 +/- 0.2 μm. In contrast, the particle size of GMP from mature

seeds was consistently larger with a volume weighted mean of 76.5 +/- 2.2 μm. This size difference with GMP from mature seeds being larger than PB and PB-gel from immature wheat is in agreement with the results described in paragraph 3.4 and shown in figures 3E and 3F. The particle size distributions of the PB and PB-gel preparations show almost perfect Gaussian distributions while the GMP preparation from mature seeds shows the presence of fractions of smaller sizes than the main peak. If all of the particles are considered as spheres (as suggested by the CLSM analysis), the average volume of a protein body can be calculated as 0.839×10^{-9} ml, of a glutenin particle from PBs as 0.972×10^{-9} ml and of a glutenin particle from mature endosperm as 0.234×10^{-6} ml. Thus, the average volumes of glutenin particles from immature wheat and dry wheat grain differ by a factor of 241.

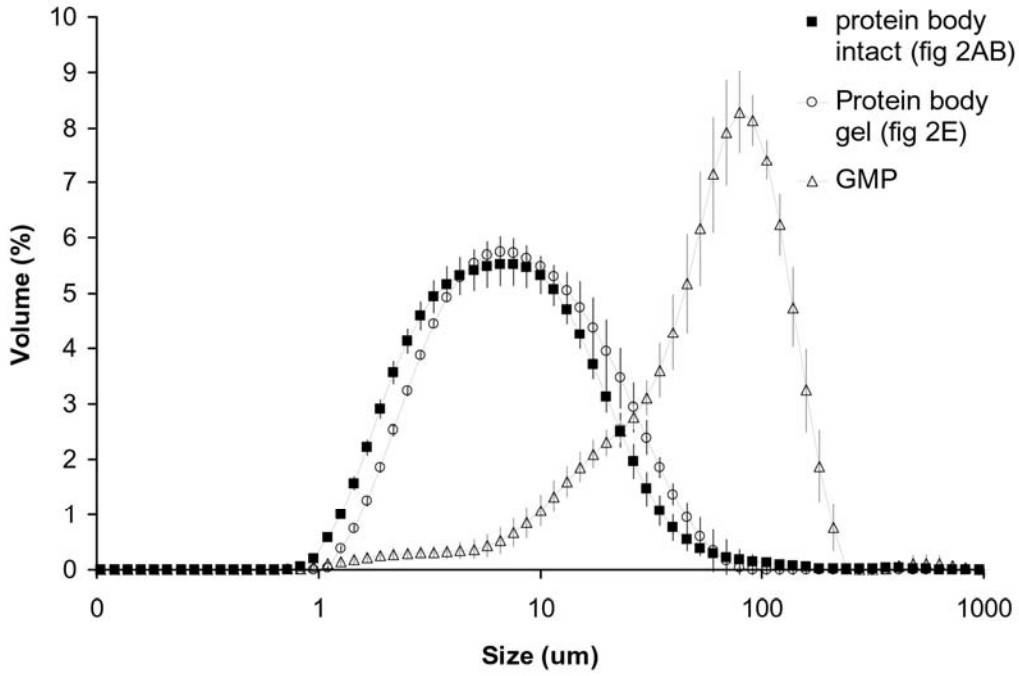

Figure 6 - Volume percentage distributions of protein bodies and gel protein fractions. Intact protein bodies are shown as squares, SDS-insoluble gel protein from isolated protein bodies as circles and SDS insoluble gel protein isolated from mature wheat as triangles. The gel protein from mature wheat is clearly much larger compared to the gel protein isolated from protein bodies.

Discussion

Differences in breadmaking quality parallel differences in the GMP fraction, which consists of spherical glutenin particles whose size correlates with a key functional parameter: dough mixing properties [Don *et al.*, 2005b]. It has also been suggested that there is a developmental relationship between the protein bodies deposited in developing grain and the GMP particles prepared from mature grain [Don *et al.*, 2003a]. Such a relationship may also offer an explanation for the established effects of environmental and genetic factors on GMP properties and processing quality [Don *et al.*, 2005a].

A systematic comparison between PBs isolated from immature wheat seeds and GMP isolated from mature seeds revealed strong similarities in protein composition including the presence of large glutenin oligomers. Both preparations also gave an SDS-insoluble gel protein fraction. However, the proportion of protein that was insoluble in SDS was much greater in the GMP than in the PBs, as shown previously by Carceller and Aussenac (1999, 2001). There were also large (more than 200-fold) differences in the sizes of the particles present in the SDS-insoluble protein from mature wheat as compared to immature wheat despite the fact that their subunit compositions were similar as demonstrated by SDS-PAGE after reduction. These results indicate that the SDS-insoluble networks in immature and in mature wheat kernels consist of the same subunits, but that they differ in their degree of polymerisation. The SDS-soluble glutenin polymers (F1 and F2), from immature and mature wheat also had similar subunit compositions to the SDS-insoluble fractions, except that triticin was present in a polymeric state in the SDS-soluble polymers of immature wheat, but not in the SDS-insoluble fractions of mature or immature grains.

Triticin

Triticin is related to the 11S storage globulins present in many seeds and has been suggested to affect the breadmaking properties of wheat [Sievert *et al.*, 1991; Singh *et al.*, 1993]. Two subunits of triticin, the larger acidic subunit and the smaller basic subunit, are coded by a single gene and are produced by post-transcriptional cleavage of a single protriticin [Singh *et al.*, 1993]. Singh *et*

al. [1991] showed that triticin forms tetrameric oligomers that consist of two acidic subunits of 52 and 58 kDa, and two basic subunits of 22 and 23 kDa as observed after reduction of the disulphide bonds. Triticin has free SH groups in its structure which allow it to polymerise [Singh *et al*., 1991], however the basic subunit contains only one cysteine residue. We identified three bands of approximately 20 kDa as the basic subunit of triticin. These bands were found in the SDS-soluble protein fractions. The absence of the 20 kDa bands in the SDS-insoluble fractions suggests a function of triticin in the degree of solubilization/polymerization. In the SDS-soluble fraction of mature wheat only a faint band was visible at the expected mass of the basic triticin subunit. This corresponds to the results of Singh *et al.* [1991] who showed that triticin protein accumulation occurs mainly between 10-21 DAF with no significant increase in accumulation beyond 21 DAF. In our analysis the basic subunit, and not the acidic subunit of triticin, was identified. One reason for this is that the 20 kDa basic subunit was well separated from the other proteins in the gel, while in the mass range where we can expect the acidic subunit (around 55 kDa) several other proteins were present. It is possible that the triticin observed in the present study was either covalently bound to the SDS-soluble gluten polymers or was itself present in a range of polymeric forms. However, it was clearly not present in the SDS-insoluble gel fraction. This is in agreement with the results obtained by CLSM (Figure 5B), as Bechtel *et al.* (1991) showed that triticin was present in inclusion bodies which were similar in appearance to those that were dissolved by treatment with SDS in the present study.

Particle size

It has been shown that GMP isolated from mature flour consists of spherical glutenin particles [Don *et al*., 2003a]. Despite the similarity in the subunit compositions of the PB and GMP fractions, the particles present in the GMP fraction from mature wheat were significantly larger (241-fold) than those in the SDS-insoluble fraction from PBs (average diameter of 76 +/- 2 μm compared with 12.3 +/- 0.2 μm). Furthermore, the mean volume of the particles present in the SDS-insoluble fraction from PBs (0.972×10^{-9} ml) is similar to that of the protein bodies themselves (0.839×10^{-9} ml): the fact that the PBs are 14% smaller may be explained by swelling of the protein in the presence of SDS.

The average volume of the inner starchy endosperm cells is $0.994x10^{-6}$ ml per cell whereas the average volume of the peripheral (subaleurone) cells is $0.217x10^{-6}$ ml per cell [Evers, 1970; Kent and Jones, 1952]. All wheat endosperm cells also contain approximately the same weight of protein [Evers, 1970] with the differences in volume relating to variation in starch content. In particular, the subaleurone cells are densely packed with protein compared with the central cells which contain mainly starch. If we assume that a glutenin particle with an average diameter of 76.49 μm is a perfect sphere then it would have an average volume of $0.234x10^{-6}$ ml, which is equivalent to the average volume of a subaleurone cell at the end of the desiccation phase. In addition, the SDS-insoluble GMP fraction of mature wheat also contains fractions of smaller sizes than the main peak. Hence, it is possible that some of the GMP particles correspond to the matrix proteins present in single subaleurone cells and others to smaller deposits in the central endosperm cells.

From protein body to glutenin particle

It has been suggested that a glutenin particle could result from the swelling of a single protein body [Don, 2005b]. The results presented in this paper show that a 241-fold increase in volume would be required, which is in agreement with the calculations of Don [2005b]. Furthermore, we observed that PBs from immature wheat can form a physical gel in SDS solution which also consists of particles of similar size (with only a 14% increase in volume) to the original protein body. This suggests a different mechanism for the formation of GMP than simple swelling of a protein body, with the GMP particles originating from the fused protein body matrix present in mature cells rather than individual PBs in developing cells.

Although the accumulation of polymeric protein is a continuous process starting as early as 7 days after flowering (DAF), a close correlation was found between the accumulation of the GMP and the rapid loss of water, which occurs during the late stages of grain development [Carceller and Aussenac, 1999, 2001]. During this stage the percentage of SDS-insoluble polymers as a proportion of total polymers can increase from 10% to 50%, whereas the synthesis of protein had already stopped at the end of the kernel ripening [Carceller and Aussenac, 1999]. The fusion of ER-derived PBs with protein storage vacuoles has already occurred before the

desiccation phase and it is possible to observe ER-derived PBs inside the vacuole and surrounded by one or two membranes in addition to the vacuolar membrane [Galili, 1997]. Therefore, the absorption of PBs into protein storage vacuoles alone cannot explain the rapid increase in SDS-insoluble protein during the desiccation phase and a second mechanism must operate during the desiccation phase. Carceller and Aussenac (2001) showed that premature desiccation of the grain induces SDS-insoluble polymer formation while Razi *et al*. [2003] suggested that glutathione may play a role in controlling the degree of polymerisation.

Another mechanism which may be relevant to the formation of GMP is the loss of membrane integrity due to programmed cell death at the end of seed development. Young and Gallie [1999] demonstrated the loss of membrane integrity using Evans Blue staining of wheat endosperm cells during seed development, which increased from 20 DAF until the entire endosperm was stained at 30 DAF. The loss of membrane integrity therefore could be associated with the desiccation phase and the formation of GMP which could make the free SH groups present in the protein bodies and protein storage vacuoles accessible for aggregation and cross linking.

Our results support the proposed relationship between PBs and glutenin particles, but suggest that further aggregation of individual PBs is required to generate the large particles observed in GMP.

Acknowledgements

This research was partially funded by the Celiac Disease Consortium and the Netherlands Proteomics Centre, both an Innovative Cluster of the Netherlands Genomics Initiative and partially funded by the Dutch Government (BSIK03009). This research was partially funded by the Allergy Consortium Wageningen (ACW) and the Top Institute Food and Nutrition Wageningen (TIFN).

Chapter 5

The feasibility of decreasing CD toxicity while retaining technological properties: A study with Chinese Spring deletion lines

Teun WJM van Herpen, Hetty C van den Broeck, Cees Schuit, Elma Salentijn, Liesbeth Dekking, Dirk Bosch, Marinus JM Smulders, Luud JWJ Gilissen, Ingrid M van der Meer, Rob J Hamer

Abstract

This study investigates the feasibility of eliminating coeliac disease (CD)-toxic gluten proteins from bread wheat while retaining the technological properties. The gluten proteins are mainly encoded by loci present on chromosomes 1 and 6 on the three different genomes of bread wheat. Not all these proteins contain CD-toxic epitopes. The effects of deleting different gluten loci was analysed using a set of deletion lines of *Triticum aestivum* cv. Chinese Spring. The effect of these deletions on CD-toxicity was analysed using monoclonal antibodies that recognize different CD T-cell epitopes of specific gluten proteins. The same deletion lines were technologically tested (mixing properties and dough rheology). Our results show that removing the gluten loci on the short arm of chromosome 1D and the short arm of chromosome 6D resulted in a large decrease in the antibody response. Removing the locus on 1D showed a minimal effect on technological properties. The deletion of the locus on 6DS resulted in a significant change in technological parameters. However these can be corrected for.

Introduction

Coeliac disease

Coeliac disease (CD) is a disorder, which is characterised by a permanent intolerance to certain wheat gluten proteins. Over 0.5% of the Western population suffers from CD, which presents itself by chronic diarrhoea, osteoporosis, lymphoma and several other clinical symptoms after gluten consumption. Until now, a complete and life long elimination of gluten from the daily diet is the only option to combat the symptoms [Koning *et al.*, 2005].

Relation between gluten proteins and coeliac disease: CD-toxic epitopes

Gluten proteins can be divided into three main groups: high molecular weight glutenin subunits (HMW-GS), low molecular weight glutenin subunits (LMW-GS), and gliadins. Gliadins can be subdivided into α/β-, γ- and ω-type-gliadins [Shewry and Halford, 2002].

CD is caused by an immune system mediated response in the small intestine against certain gluten peptides. In the small intestine, some native gluten peptides can bind directly to specific HLA-DQ2/8 receptors on antigen presenting cells (APCs). After deamidation by tTG, the affinity for these HLA-receptors is strongly increased. APCs can present the gluten peptides to gluten-sensitive T-cell lymphocytes leading to the release of cytokines, which will cause inflammation reactions and result in damaged intestinal villi [Vader *et al*., 2002b]. Multiple T-cell activating gluten peptides were mainly found in the α-gliadins but also in γ-gliadins and glutenins [Sollid 2002; Arentz-Hansen *et al*., 2000; Vader *et al*., 2002b; Koning, 2003]. Peptides derived from α-gliadins are recognized by T-cells from most CD patients, while T-cell responses to γ-gliadins and glutenins are less frequently found [Arentz-Hansen *et al*., 2000; 2002; Vader *et al*., 2002b; 2003; Molberg *et al*., 2003]. The development of an adaptive immune response is strongly controlled by innate immunity. Only α-gliadin derived peptides are reported to induce this innate response [Palova-Jelinkova *et al.*, 2005].

Development of CD-safe wheat

Wheat with much lower amounts of epitopes that trigger this T-cell response (in short: low CD-toxic or CD-safe wheat) can be favourable in two areas. Firstly, some CD-patients can possibly tolerate these cultivars low in T-cell CD-toxic epitopes. Indications for this can be found in studies regarding oats [Janatuinen *et al.*, 2002]. Oats, which contain few sequences that can be recognized by T-cells of certain CD-patients, appear to be tolerated by most CD patients [Vader *et al.*, 2003]. Secondly, a diet based on wheat cultivars low in CD-toxic epitopes may help in the prevention of CD. Researchers observed that the amount of gluten exposure is correlated to the population incidence of CD [Fasano, 2006; Ivarsson *et al.*, 2000; Ventura *et al.*, 1999]. However, progress in breeding for bread wheat with less CD-toxic gluten may at the same time result in an unwanted loss of gluten technological properties.

Gluten and technological properties

Technological properties of wheat gluten are related to interactions between different flour components e.g. proteins, starch and arabinoxylans [Hamer and Van Vliet, 2000]. Total protein content was found to affect dough and baking properties [Uthayakumaran *et al.,* 1999], but just quantity is not sufficient to explain differences in flour or dough properties [Wieser and Kieffer, 2001]. Proteins important for dough properties are both glutenins and gliadins. Hydrated glutenins are both cohesive and elastic, and are responsible for dough strength and elasticity. In contrast, gliadins contribute mainly to the viscosity and extensibility of the dough. Purified hydrated gliadins have little elasticity and are less cohesive than glutenins. A correct mixture of the two is crucial to obtain optimal viscoelastic dough [Wieser, 2007].

The polymerisation behaviour of the glutenins has been studied extensively. Both the quantity and the size distribution of the polymeric gluten proteins are important factors in dough properties [Gupta *et al.*, 1993]. The largest polymers in wheat flour are the glutenin macropolymers (GMP). Differences in technological quality correlate with differences in the characteristics of GMP [Graveland *et al.*, 1982, 1985; Weegels *et al.*, 1996]. GMP is characterised by high levels of cross-linked HMW- and LMW-glutenins [Don *et al.*, 2003a]. The relative sizes of GMP depend upon its composition, particularly the ratio of HMW-GS to LMW-GS and the

types of individual glutenin subunits [Gupta *et al.*, 1993]. HMW-GS composition affects the quantity of GMP and at the same time affects flour-mixing time [Shewry *et al.*, 1989; MacRitchie, 1999]. Mainly during the kernel desiccation phase [Carceller and Aussenac, 1999] GMP is accumulated as spherical particles (glutenin particles) in the wheat grain [Don *et al*., 2003a]. The size of these glutenin particles correlates with dough mixing properties [Don *et al.*, 2005b].

Gluten Loci

Elimination of gluten containing CD-toxic epitopes from wheat is challenging because bread wheat contains a large number of copies encoding gluten proteins and these genes share a high sequence homology [Chapter 2]. Wheat gluten genes are located on three different genomes, A, B and D, which originate from three individual ancestors. A few (for HMW-GS) to several hundred (for α-gliadins) of gene copies are present in hexaploid wheat but not all of these gluten genes are expressed. If expressed they can have different expression patterns and levels [Shewry and Halford, 2002]. Sequences of individual copies of genes from the same gluten family, such as the α-gliadins, are very similar and can contain multiple and different CD-toxic peptides [Chapter 2].

Genetic analysis demonstrates that most gluten proteins of hexaploid *Triticum aestivum* are encoded by different sets of loci on group 1 and 6 chromosomes [Shewry *et al.*, 2003; Sabelli and Shewry, 1991]. The HMW-GS are encoded by loci on the long arm of group 1 chromosomes (*Glu-A1, Glu-B1, and Glu-D1*) [Harberd *et al.*, 1986]. Each of these loci encodes two subunits (one x-type and one y-type subunit). The LMW-GS are mainly encoded by the *Glu-3* loci located at the short arms of group 1 chromosomes [Singh and Shepherd, 1988] and are tightly linked to the loci encoding the γ-gliadin (*Gli-1*) and ω-gliadins (*Gli-1, Gli-3*). Most α/β-gliadins are encoded by loci on the short arms of group 6 chromosomes (*Gli-2*) [Marino *et al*., 1996]. As shown before, the contributions to CD-epitopes by these gluten genes can show large differences according to the genome of origin [Chapter 2; Molberg *et al.*, 2005].

Objective of this study

In this study, we explore the feasibility of a strategy to reduce CD-toxicity in hexaploid bread wheats with a minimal effect on the technological properties of gluten proteins. For this, various

deletion lines of *Triticum aestivum* cv. Chinese Spring were selected. A large series of deletion lines have been developed in Chinese Spring. Deletion lines and nullisomic lines form an important cytogenetic resource.

The deletion lines selected here generally lack one locus of gluten genes from one of the three homologous chromosomes. Such deletions result in a complete removal of the proteins encoded by the deleted chromosome parts. These lines give us the unique opportunity to study the effect of removal of any chosen set of gluten genes on both the reduction of CD-toxicity and on the degree to which the technological properties of flour are affected.

The questions to be answered in this study are: 1) What is the contribution of the different gluten loci originating from the three homologous chromosomes to CD-toxicity? 2) What is the importance of these gluten loci for the technological properties of wheat flour? By combining the answers to these two questions, we explore the feasibility of a strategy to reduce CD-toxicity in hexaploid bread wheats with a minimal effect on the technological properties of gluten.

Experimental

Monoclonal Abs that recognize CD-toxic T-cell epitopes have been raised in mice [Mitea *et al.*, 2008]. The availability of these mAbs enables the screening of gluten proteins for the occurrence of immune responsive epitopes. In previous research, screening for CD toxicity on gluten protein extracts was carried out using T-cell proliferation assays [Molberg *et al.*, 2005; Spaenij-Dekking *et al.*, 2005]. Because most gluten proteins are poorly soluble in aqueous solvents, they have to be digested when analyzed in T-cell assays. Using the mAbs, screening can be performed on native gluten without previous digestion.

Wheat materials

Twenty-six *T. aestivum* Chinese Spring deletion lines were selected (obtained from Wheat Genetics Resource Centre, Kansas State University) having partial deletions of the long and short arm of chromosomes 1 and 6 (Figure 1B and 4B). Only lines with homologous deletions were

selected. One line had deletions of both chromosome 1 and chromosome 6 (1BS-19/6DS-4, Figures 1B). All deletion lines were grown in containment glasshouses. No morphological differences between the different lines were observed. Seeds were harvested from mature wheat plants.

Database search on the specificity of the mAb compared to T-cell epitopes

The minimal recognition sequence determined for the mAbs is not identical to the minimal recognition sequence of the corresponding human T-cell clones [Mitea *et al.*, 2008]. The mAb sequences are shifted a few amino acids to the left or to the right. To examine to what extent different sequences are recognized by the mAbs compared to the sequences recognized by the T-cell clones, the occurrence of the sequences was analysed. From the NCBI protein database (http://www.ncbi.nlm.nih.gov/) five different groups of gluten protein sequences were extracted, and subsequently converted into FASTA formats, using the following search queries: “alpha gliadin”, “gamma gliadin”, “omega gliadin”, “LMW glutenin” and “HMW glutenin”. All non-*Triticum*, non-*Aegilops* entries and sequences containing <100 amino acids were removed. The obtained protein sequences were aligned using ClustalW to validate if the correct groups were assigned to the sequences. In the “gamma gliadin” group, four sequences (AAA34286, P04729, P04730, and AAA34285) were more similar to LMW glutenin proteins as observed from ClustalW and BLASTP (http://www.ncbi.nlm.nih.gov/) results. These four sequences were transferred to the “LMW glutenin” group. In the “omega gliadin” group, one sequence (ABI20696) was specific for the “alpha gliadin” group, observed from ClustalW and BLASTP results, and was transferred to the “alpha gliadin” group.

The sequences in the five established groups were analysed for the different minimal recognition sequence of the mAbs and T-cells [Mitea *et al.*, 2008]. No mismatches were allowed in the scoring. Scores were expressed as the number of sequences and as the percentage of the sequences in the established group that contained one or more recognition sequences. The T-cell minimal recognition sequences used in the analyses were: Glia-α9 (PFPQPQLPY), Glia-α20 (FRPQQPYPQ), Glt-156 LMW (PFSQQQQSPF), HMW-glt (QQGYYPTS) and mAb minimal

recognition sequences used were: Glia-α9 (QPFPQPQ), Glia-α20 (RPQQPYP), Glt-156 LMW-1 (PPFSQQ), LMW-2 (QSPFS), HMW-glt (GYYPTS) [Mitea *et al.*, 2008; Mitea *et al.*, in prep].

Protein isolation

Prolamins were extracted from wheat grains according to Van den Broeck *et al*., [2008]. Grains were ground in an analytical mill (A 11 Basic, IKA-Werke) and sieved through mesh (0.5 mm). Prolamins were extracted from 50 mg wheat flour by addition of 0.5 ml of 50% (v/v) iso-propanol with continuous mixing (MS1 Minishaker, IKA Works, Inc.) at 1000 rpm for 30 min at room temperature, followed by centrifugation at 10,000 rpm for 10 min at room temperature. The residue was re-extracted twice with 50% (v/v) iso-propanol/1% (w/v) DTT/50 mM Tris-HCl, pH 7.5, for 30 min at 60ºC with mixing every 5 to 10 min followed by centrifugation at 10,000 rpm for 10 min at room temperature. After addition of each next extraction solution, the residue was resuspended by shaking in a Fastprep® FP220A Instrument for 10 sec at 6.5 m/sec followed by sonication for 10 min in an ultrasonic bath (Branson 3510, Branson Ultrasonics Corporation). The three obtained supernatants were combined and considered the prolamin extract. The protein content was quantified using the Biorad Protein Assay (Biorad Laboratories), based on the Bradford dye-binding procedure, according to manufacturer's instruction with BSA as a standard.

SDS-PAGE

Proteins were separated on SDS PAGE gels (10%) using a SE260 mighty small II system (GE Healthcare, UK). SDS-PAGE was followed by immunoblotting or by silver staining according to Rabilloud *et al*., [1988] with some modifications. Gels were fixed in 50% (v/v) ethanol/10% (v/v) acetic acid in water for 30 min. Then, gels were washed in 5% (v/v) ethanol/1% (v/v) acetic acid in water for 10 min, followed by three times washing for 5 min in MilliQ water. Gels were sensitized in 0.02% (w/v) sodium thiosulfate for 1 min and again washed three times for 30 sec in MilliQ water. Gels were incubated in 0.1% (w/v) silver nitrate for at least 20 min. After this incubation, gels were rinsed 2 times for 5 sec in MilliQ water and developed in 6% (w/v) sodium carbonate containing 0.05% (v/v) formaldehyde (37%)/0.4‰ (w/v) sodium thiosulfate.

Immunoblotting

Proteins were blotted onto nitrocellulose (0.2 µm, Biorad Laboratories), in buffer omitting methanol, using a Mini Trans-Blot Cell (Biorad Laboratories) at 100 V for 1 hour. Blots were incubated and visualised as described by Cordewener *et al*., [1995] using mAbs specific for stimulatory T-cell epitopes against Glia-α9, Glia-α20, GLT-156 (LMW-1 and LMW-2), HMW-glt [Mitea *et al*., 2008]. Monoclonal Ab binding was visualised by staining for alkaline phosphatase, using Nitro Blue tetrazolium (NBT) and 5-Bromo-4-chloro-3-indolyl phosphate (BCIP) (Sigma).

Quadrumat milling

To obtain white wheat flour, wheat kernels (total weight ranging 7.6-36 g) were milled using a Quadrumat JR (Brabender, Germany). Kernel moisture was adjusted to 16.5%. Bran was separated from endosperm flour by sieving through mesh (150 µm). After sieving the average yield was 50% (w/w), noting that samples 6AS-1 and 6DS-2 had a typically higher flour yield of 64% and 60%, the other samples ranged from 43% to 51%.

DUMAS protein determination

Flour protein content (N×5.7) was determined by the Dumas method [Sebecic and Balenovic, 2001] using an NA2100 Nitrogen and Protein Analyzer (ThermoQuest-CE Instruments, Rodeno, Italy). The Dumas method is based on the measurement of total nitrogen in the sample. Methionine was used as a standard.

GMP isolation and glutenin particle size determination

Dough strength is correlated to the amount of GMP and to the size of glutenin particles and was therefore determined in this study. GMP was isolated by dispersing flour in 1.5% (w/v) SDS followed by ultracentrifugation as described earlier by Graveland *et al*., [1982]. GMP volume (µl/mg flour protein) was determined and used for comparison. Fresh GMP from flour was dispersed in 1.5% (w/v) SDS (10 ml) and left rotating overnight at room temperature. Particle size distributions were measured using a Mastersizer 2000 (Malvern Instruments, UK). The laser

diffraction pattern obtained with the instrument was correlated to the particle size distribution based on Fraunhofer theory, assuming a spherical particle shape. The range of the instrument was 0.02–2000µm. GMP dispersions were transferred to the water filled sample vessel at an obscuration of approximately 8%. The surface area mean [D3.2] was taken from the particle size distribution data for comparisons. Further details of this method are described elsewhere [Wang *et al*., 2005; Don *et al*., 2005b].

Mixing

Dough strength is determined using a micro-Mixograph. A 2 g Mixograph (National Manufacturing Co., USA) pin-mixer was used to analyse the mixing properties of the different flour samples. Mixing was performed at 20°C. Water was added according to the Plastograph method [ICC 115/1 (ICC, 1992); Shewry and Lookhart, 2003], which corrects for differences in flour protein contents. Dough contained 2% (w/w) sodium chloride (Merck, Germany). Bandwidth at peak resistance (BWPR) (%) and dough development time (DDT) in minutes were taken from the midline analysis for comparison.

Flow-relaxation measurements

Relaxation tests were performed to study dough elasticity. Longer relaxation half times indicate more elastic dough behaviour [Lásztity, 1996; Lichtendonk *et al*., 2000]. Dough was mixed to peak in the 2 g Mixograph pin-mixer, carefully taken from the mixer and transferred into the Bohlin VOR rheometer (Bohlin Instruments, Sweden). Flow-relaxation measurements were performed using aluminium grooved plate geometry with a cross-section of 30 mm and a gap of 1 mm [Lichtendonk *et al.*, 2000]. Moisture loss from the dough piece was prevented using paraffin oil. The actual measurement was made after an equilibration time of 30 min. to allow appropriate release of dough stress. The measuring temperature was 20°C. During measurement the sample was deformed to a strain of 100% at a shear rate of 0.0208 s^{-1}. The strain was kept constant and the subsequent decrease of stress of the dough was recorded as a function of time. The time necessary for the dough to relax to a stress of 50% of the initial stress, recorded directly after stopping deformation, was taken as the flow-relaxation half time ($T_{1/2}$).

Results

Validity of used mAbs: specificity of minimal recognition sequences in different classes of gluten database sequences

Table 1 shows the numbers (and percentages) of sequences that contain the different minimal recognition sequences of the mAbs and corresponding T-cell clones. From this, we observed that the mAb and T-cell minimal sequences were specific for the epitopes in each of the expected groups, with the exception of the mAb Glia-α9, which also was present in γ- and ω-gliadin protein sequences, even though the T-cell recognition sequence is not. The minimal recognition sequences for the anti-LMW-1 and LMW-2 mAbs (both have a smaller minimal recognition sequence than the corresponding T-cell related sequence) were more frequently found in the LMW-GS group than the corresponding T-cell recognition sequence. The predicted HMW recognition was identical for T-cells and Ab.

Epitope	α/β-gliadin	γ-gliadin	ω-gliadin	LMW-GS	HMW-GS
T-cell Glia-α9	70 (67%)	-	-	-	-
T-cell Glia-α20	72 (69%)	-	-	-	-
T-cell LMW-1/2	-	-	-	12 (12%)	-
T-cell HMW	-	-	-	-	147 (54%)
mAb Glia-α9	85 (82%)	42 (48%)	3 (60%)	-	-
mAb Glia-α20	72 (69%)	-	-	-	-
mAb LMW-1	-	-	-	90 (91%)	-
mAb LMW-2	-	-	-	43 (43%)	-
mAb HMW	-	-	-	-	147 (54%)
Total number of sequences	104	87	5	99	273

Table 1 - Scoring of the minimal recognition sequences of the T-cell clones and mAbs in the different gluten protein classes. The table contains for each epitope tested for the numbers (and percentages of total sequences tested) of minimal recognition sequences from both T-cell stimulatory epitopes and mAbs found in different classes of gluten sequences from the NCBI database.

–:no recognition sequence was found in that class of gluten sequences.

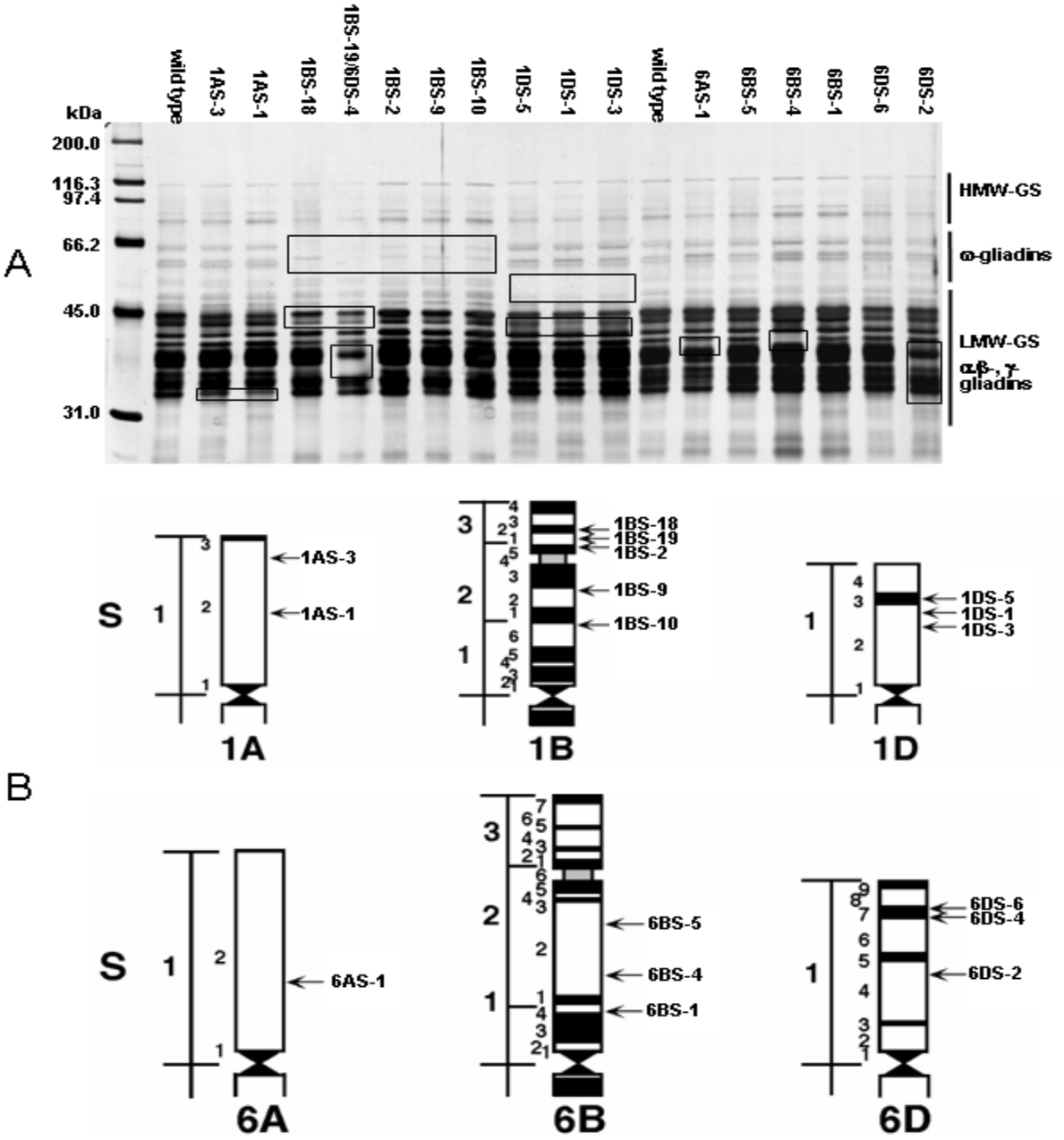

Figure 1 - A: Chinese Spring deletion lines analysed by SDS-PAGE (10%) followed by silver staining. Samples are indicated above each lane. Positions of different gluten proteins are indicated. Boxes indicate major differences compared to wild type. B: Group-1 and 6 idiogram of Chinese Spring deletion lines for the short arm of chromosome 1 (Wheat Genetics Resource Centre, Kansas State University) indicating the approximate location of the deletions.

Protein composition Chinese Spring deletion lines: SDS-PAGE analysis

Prolamins were extracted from Chinese Spring deletion lines and analysed by SDS-PAGE followed by silver staining (Figure 1A). Because most HMW-GS do not stain well using silver staining [Van den Broeck *et al*., 2008], these deletion lines were not analysed at this level. Major differences compared to wild type are indicated in Figure 1A. Differences in protein content compared to wild type were observed in the ω-gliadin region for lines with deletions of the short arms of chromosomes 1B and 1D, which lack some bands. Most of the deletion lines showed differences in the LMW-GS/α/β-, γ-gliadin region as expected. The double deletion line, 1BS-19/6DS-4 (Figure 1A, Lane 5), was lacking the largest number of bands. Deletion line 6BS-4 (Lane 15) lacks a protein band that is present in the other deletion lines of chromosome 6B, even though these have been reported (Wheat Genetics Resource Centre; Figure 1B) to have larger deletions. This deletion line also contains the 5BS-2 deletion, but, to our knowledge, no gluten protein locus has been identified to be located on the short arm of chromosome 5B.

Epitopes in the gluten proteins of Chinese Spring

Various mAbs recognizing different T-cell epitopes were used for analysing Chinese Spring deletion lines (as depicted on idiograms in figure 1B, 4B) for their contribution to CD-toxicity. In figure 2A, immunoblot results are presented using the mAbs Glia-α9 and Glia-α20 for deletion lines of the short arm of chromosomes 1 and 6. Major differences, compared to wild type, are indicated with arrowheads. Deletion lines 1AS-3 (Lane 2) and 1AS-1 (Lane 3) lack two protein bands using mAb Glia-α9 and no protein bands by using mAb Glia-α20. This suggests that these proteins only contain the sequence recognized by mAb Glia-α9. All five deletion lines of the short arm of 1B (Lanes 4-8) lacked two protein bands by using mAb Glia-α9 and no protein bands by using mAb Glia-α20, except in case of the double deletion line 1BS-19/6DS-4 (Lane 5), which lacked two extra bands by using mAb Glia-α9 and three by using mAb Glia-α20. One protein band was recognised by both mAbs. All 1BS deletion lines lacked the same protein bands, including the ones reported to have a very small deletion, which may suggest that the locus encoding these proteins is close to the end of the chromosome arm. All three deletion lines of the short arm of chromosome 1D (Lanes 9-11) lacked four protein bands by using mAb Glia-α9 and

two protein bands by using mAb Glia-α20. Deletion lines of 1DS lacked two protein bands by using mAb Glia-α9 and one protein band by using mAb Glia-α20. These protein bands were also absent from the prolamin pattern on SDS-PAGE using silver staining (Figure 1A, lane 9-11). Deletion line 6AS-1 (Figure 2A, lane 13) lacked one protein band by using mAb Glia-α9, and two bands by using mAb Glia-α20. Deletion line 6BS-4 (Figure 2B, lane 15) lacked one protein band by using mAb Glia-α9, but this was not the case for the other two 6BS deletion lines (6BS-1 and 6BS-5). A band of a similar size was also absent from the prolamin pattern on SDS-PAGE using silver staining. In the 6BS deletion lines, no changes were observed, compared with CS wild type, in protein bands by using mAb Glia-α20. From the 6DS deletion lines, the line with the largest deletion (6DS-2, Figure 2 lane 18) lacked three protein bands by using mAb Glia-α9 and four bands by using Glia-α20.

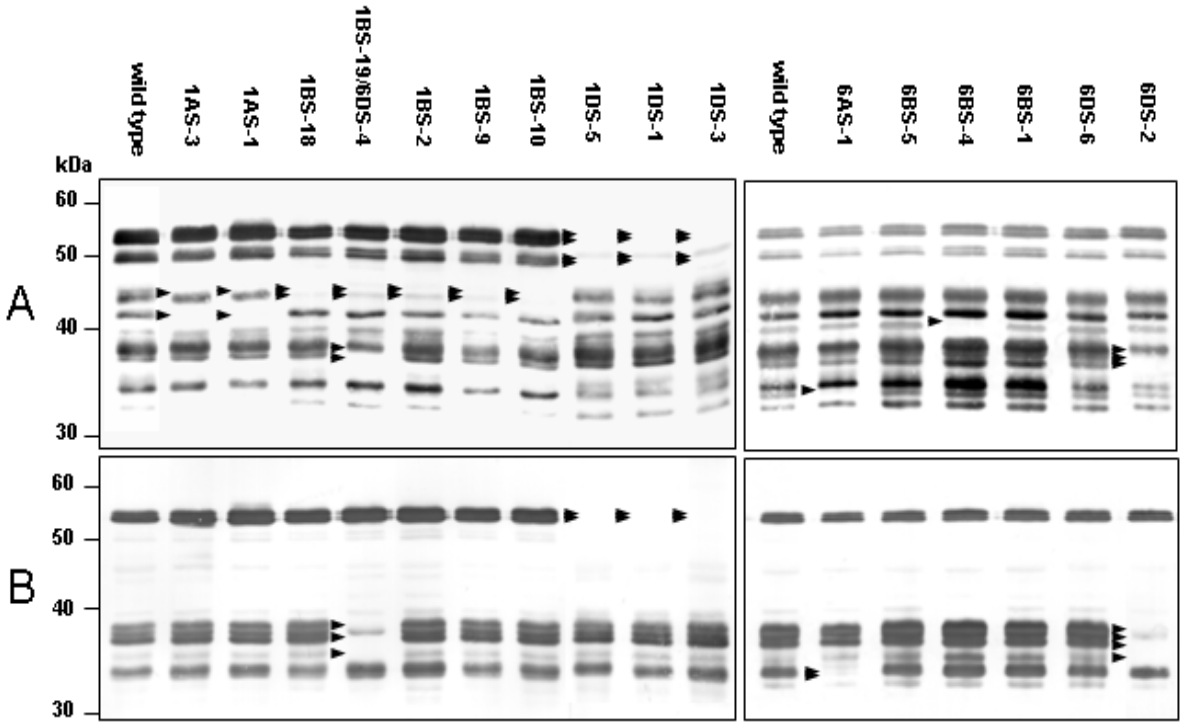

Figure 2 - Immunoblot analysis of deletion lines of the short arm of chromosome 1 and 6 using, mAbs recognizing T-cell epitopes Glia-α9 (A), Glia-α20 (B). Names of the different deletion lines are indicated above each lane. Arrowheads indicate major differences, compared to CS wild type.

Figure 3 shows the immunoblot results using mAb LMW-2 for the various deletion lines of the short arm of chromosome 1. One band was observed in all the deletion lines and in CS wild type without any significant differences. Immunoblot results using mAb LMW-1 showed similar patterns (results not shown).

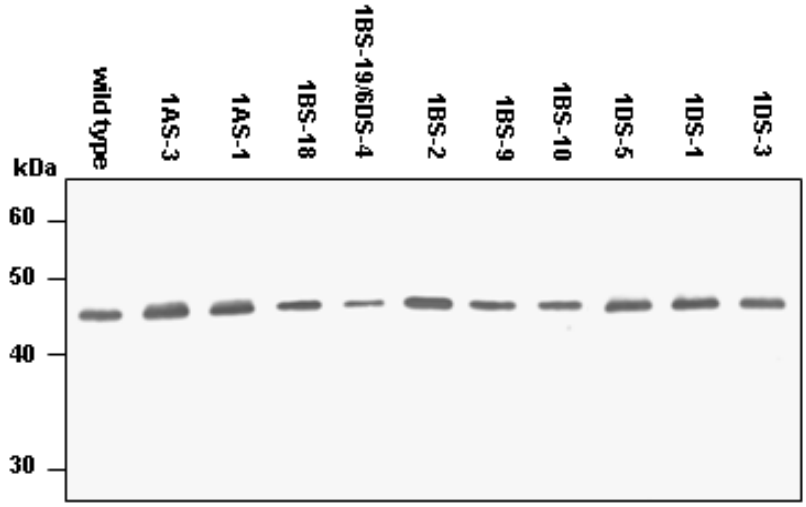

Figure 3 - Immunoblot analyses of deletion lines with mAbs recognizing T-cell epitopes LMW-2.
Names of the different deletion lines are indicated above each lane. Molecular sizes are indicated at the left.

Figure 4 shows the immunoblot results for the deletion lines of the long arm of chromosome 1 using the mAb recognising HMW-glt. In CS wild type all four HMW gluten subunits were detected. No contribution to HMW-GS was observed for the long arm of chromosome 1A, as expected for a silent locus. Two HMW-GS (HMW-GS 1Bx7 and HMW-GS 1By8) were absent in the deletion lines 1BL-1 and 1BL-6. The two other HMW-GS (2 and 12) were absent in deletion line 1DL-4.

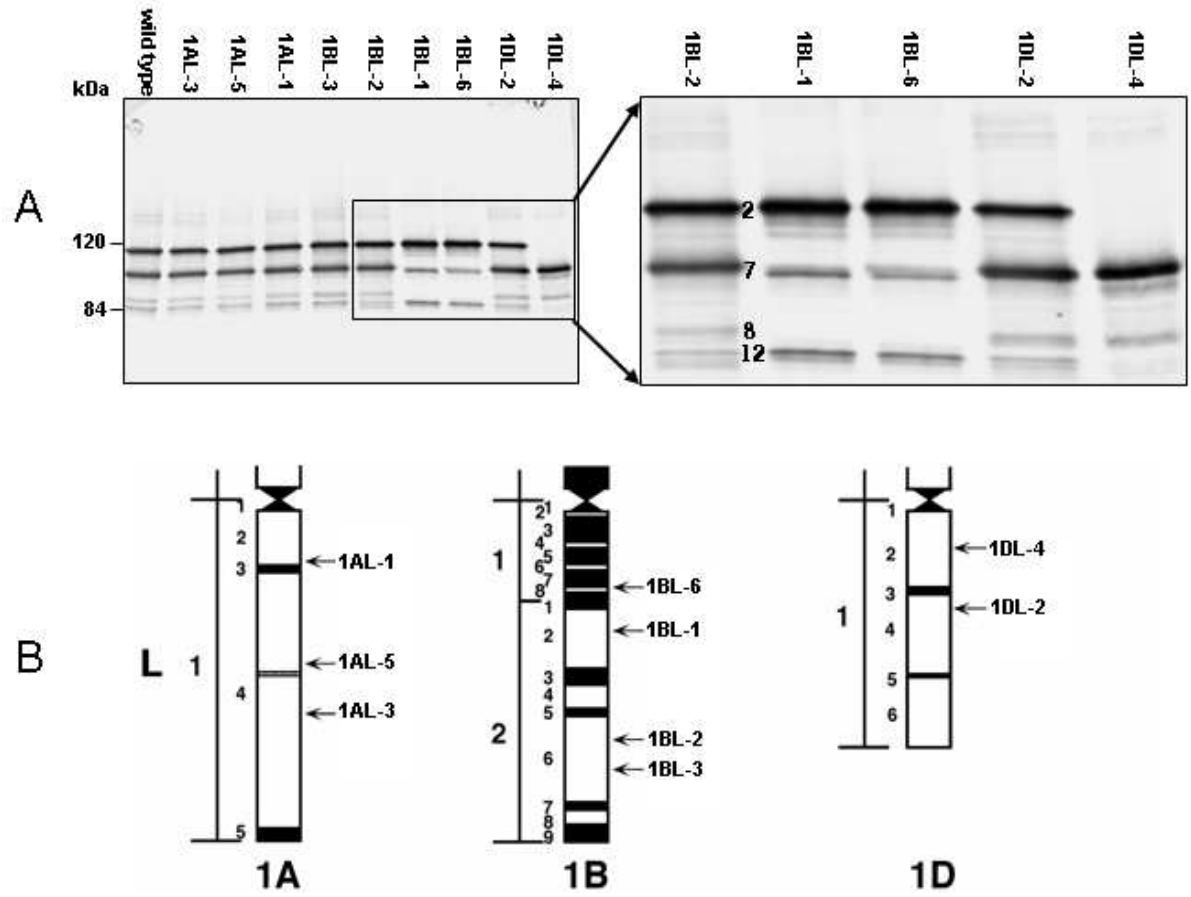

Figure 4 - A: Immunoblot analysis of deletion lines of the long arm of chromosome 1 using the mAb recognizing a T-cell epitope in the HMW-glt. Names of the different deletion lines are indicated above each lane. The last five lanes were magnified to show further detail. The positions of the different HMW-GS are indicated. B: Group-1 idiogram of Chinese Spring deletion lines for the long arm of chromosome 1 (Wheat Genetics Resource Centre, Kansas State University).

Technological parameters of Chinese Spring deletion lines

The lines with the largest deletions from chromosomes 1 and 6 were used for technological testing (Table 2). Parameters among flours of different deletion lines are represented in Table 2. Total protein content in flour (% w/w) of each deletion line was higher than in wild type flour (Table 2). Especially protein content in flour of line 6AS-1 was high (20.5%), followed by protein content in flour of deletion line 1BS-19/6DS-4 (18.6%). This does not reflect actual kernel protein content

and is presumably because of the low extraction achieved by the Quadrumat mill. For technological tests, corrections were made for these differences in protein content (described in the experimental section).

Table 2 - Different technological parameters as tested for different deletion lines. All technological measurements are performed in duplicate, except the relaxation test for the deletion lines 1DL-4, 6AS-1, 6DS-2 and 6DS-4/1BS-19. Protein content of flour was higher for all deletion lines compared to the wild type. Corrections for protein content were applied for the different technological tests.

Line	Deletions in	% protein in flour	DDT (min)	BWPR (%)	$T_{1/2}$ (sec)	GMP volume (µl/mg)	Particle surface area, D3.2 (µm)
Wild type	-	11.7 ± 0.2	3.1 ± 0.3	18.3 ± 0.5	45 ± 3	10.2 ± 0.2	12.2 ± 0.2
1AL-1	HMW (silent)	14.7 ± 0.4	3.6 ± 0.2	27.3 ± 0.2	48.7 ± 0.2	12.2 ± 0.5	15.9 ± 0.3
1BL-1	HMW	12.9 ± 0.7	2.1 ± 0.1	19.1 ± 0.8	30.9 ± 0.5	1.4 ± 0.2	n/a
1DL-4	HMW	17 ± 1	1.6 ± 0.1	17.6 ± 0.0	17.5	0.3 ± 0.1	n/a
1AS-1	LMW/ γ	15.2 ± 0.5	4.0 ± 0.4	23 ± 2	67.2 ± 0.5	11.7 ± 0.0	11.9 ± 0.1
1BS-10	LMW/ γ	13 ± 2	2.8 ± 0.2	22.4 ± 0.6	42.2 ± 0.5	7.3 ± 0.1	11.2 ± 0.2
1DS-1	LMW/ γ	13 ± 2	3.0 ± 0.1	24.9 ± 0.1	45.8 ± 0.3	12.6 ± 0.0	12.4 ± 0.2
6AS-1	α	20.5 ± 0.8	3.0 ± 0.0	42.9 ± 0.9	41.7	12 ± 1	18.2 ± 0.1
6BS-1	α	14.3 ± 0.4	4.7 ± 0.1	23.2 ± 0.2	36.8 ± 0.5	10.8 ± 0.1	10.9 ± 0.1
6DS-2	α	14.2 ± 0.8	4.1 ± 0.3	27.3 ± 0.7	33.9	8.7 ± 0.7	10.5 ± 0.1
1BS-19/6DS-4	α / γ /LMW	18.6 ± 0.3	3.7 ± 0.1	33.0 ± 0.6	55.4	11.6 ± 0.2	11.4 ± 0.1

GMP quantity expressed as volume per mg protein was decreased in deletion line 1BL-1 and was almost absent in deletion line 1DL-4 (Figure 5A; Table 2). This will lead to a decrease in dough strength [Graveland *et al.*, 1982, 1985; Weegels *et al.*, 1996]. Flours of two other deletion lines, 1BS-10 and 6DS-2, showed smaller but significant decreases in GMP quantity. For all other deletion lines, GMP volume was increased. Because of the low amount of GMP present in flour of the deletion lines 1BL-1 and 1DL-4, it was impossible to measure glutenin particle sizes for these lines. Average glutenin particle size was increased in flour of deletion line 1AL-1, and even more in flour of deletion line 6AS-1 (Figure 5B; Table 2). In deletion lines 6DS-2, 6BS-1, 1BS-10, 1BS-19/6DS-4 and 1AS-1 average particle size was decreased compared to CS wild type (Figure 5B; Table 2).

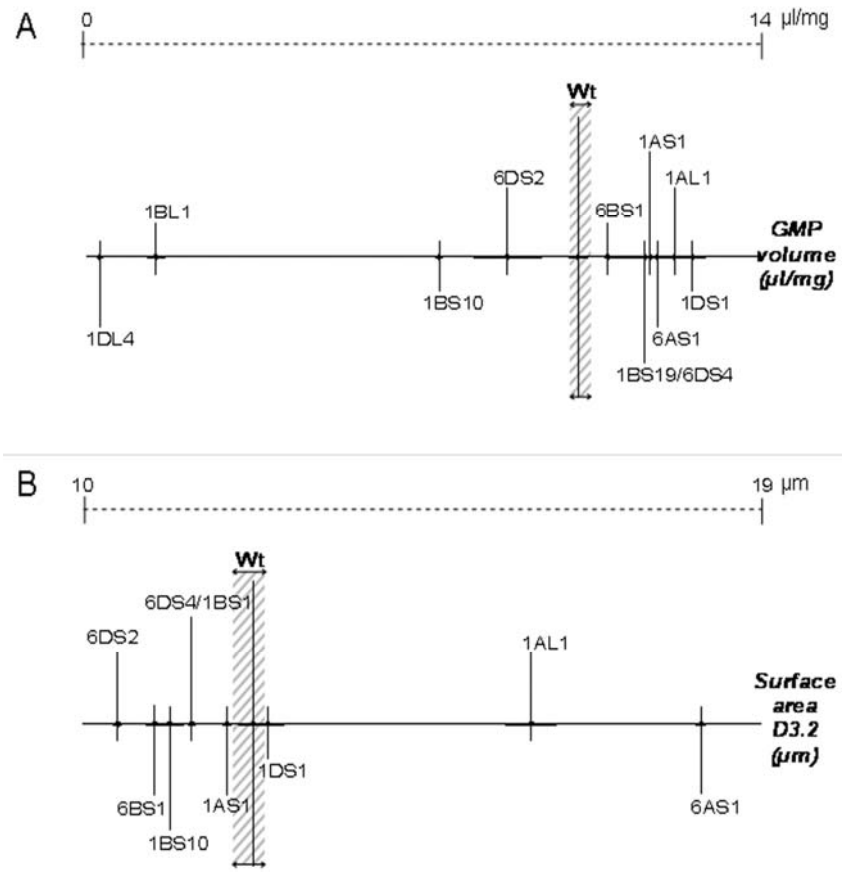

Figure 5AB - GMP properties of different deletion lines and wild type of Chinese Spring wheat. GMP volume as µl per mg protein (A) and glutenin particle sizes as D3.2 (B) are expressed on a linear bar. Thick lines on the bar accompanying each value indicate the standard error of each sample, while the standard error of the wild type is indicated with shaded boxes. Please note that in figure 5B, 1BL1 and 1DL4 are absent because the amount of GMP was too low.

Dough made from the two deletion lines missing HMW-GS (1BL-1 and 1DL-4) showed a significant decrease in DTT (Figure 6A; Table 2). Dough made from other deletion lines, especially the lines with deletions of the short arm of chromosome 6 (6AS-1, 1BS-19/6DS-4, 6DS-2 and 6BS-1) and 1AS-1, showed an increased DDT. Deletions of the *Gli-2* loci seem to have a strong effect on increasing DDT. BWPR was slightly decreased in deletion line 1DL-4 and was increased in all other deletion lines compared to CS wild type dough (Figure 6B, Table 2). BWPR was especially high in the deletion lines 6AS-1 and 1BS-19/6DS-4. It is relevant to note that these are the same deletion lines having the highest protein content in flour. Dough elasticity expressed by $T_{1/2}$ was mostly decreased in flours of the deletion lines 1BL-1 and 1DL-4 that lack

HMW-GS (Figure 6C; Table 2). Dough elasticity was also decreased in the deletion lines 6BS-1 and 6DS-2. In contrast, line 1BS-19/6DS-4, and even more line 1AS-1, showed an increase in relaxation half time ($T_{1/2}$), indicating a more elastic dough [Lásztity, 1996; Lichtendonk *et al.*, 2000].

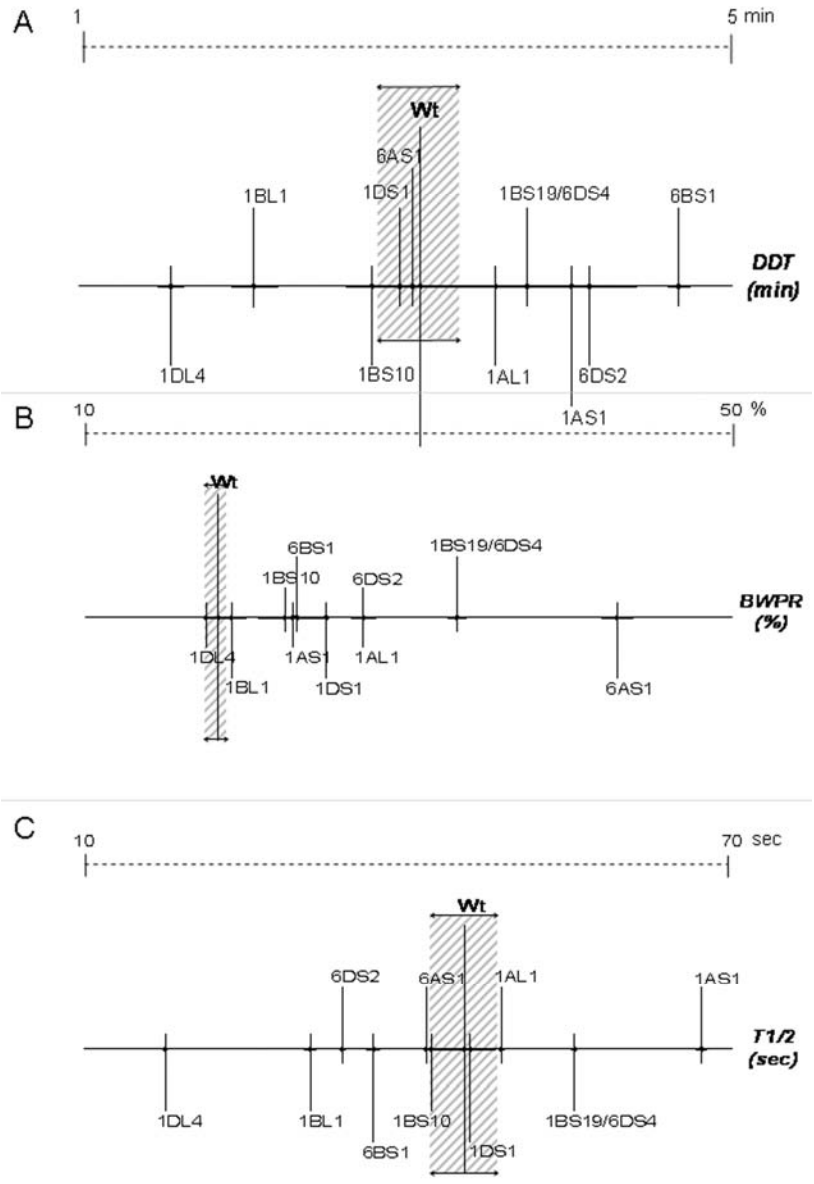

Figure 6ABC - Dough technological parameters of different deletion lines and wild type of Chinese Spring wheat. Four was mixed in a pin-mixer described in the experimental part. Corrections for differences in protein content were applied. Technological parameters obtained were DTT (A) and BWPR (B) and $T_{1/2}$ (C). These parameters are expressed on a linear bar. Thick lines on the bar accompanying each value indicate the standard error of each sample, while the standard error of the wild type is indicated with shaded boxes.

Discussion

In this study, we investigated the possibilities to develop a hexaploid bread wheat variety, which would combine reduced CD-toxicity with good technological properties. We attempted to answer this using a set of Chinese Spring deletion lines, which lack different gluten loci from the group 1 and 6 chromosomes of wheat. The different deletions were studied on both CD-toxicity and technological properties. The combination of these results can tell whether removal of certain gluten genes as a strategy to reduce CD-toxicity can be carried out with minimal effect on the technological properties of bread wheat.

Relevance of mAbs used

In this study, we used five mAbs specific for T-cell epitopes: Glia-α9, Glia-α20, LMW-1, LMW-2 and HMW-glt [Mitea *et al.*, 2008]. The antibody recognition sites are not identical to those of the T-cells (Mitea *et al.*, in prep). An in silico prediction, using Genbank sequences, indicated that the number of database sequences containing the minimal recognition sequence of the mAbs Glia-α20 and HMW-glt were the same as for the corresponding T-cell clones. The mAbs LMW-1 and LMW-2 had a much smaller minimal recognition sequence than the corresponding T-cell clones. This resulted in a higher number of sequences containing the LMW-1 and LMW-2 mAb recognition sequences than the LMW T-cell recognition sequence. However the predicted number of proteins detectable using these mAbs is large, we could only identify a single prolamin band in the CS wild type and deletion lines (Figure 3). Analysis of the minimal recognition sequence of the mAb Glia-α9 indicated that besides the α-gliadin proteins this mAb would also recognise γ- and ω-gliadins. Indeed, the Glia-α9 mAb recognised protein bands that disappeared in deletion lines of the short arm of chromosome 1A, 1B and 1D (where γ- and ω-gliadin encoding genes are located). It is shown that ω-gliadins may have epitopes that are involved in gluten-sensitive response of CD patients [Ensari *et al.*, 1998]. The sequence of the mAb Glia-α20 was only found in the α-gliadin sequences from the database, but the mAb Glia-α20 stained two bands that were encoded by the 1DS chromosome. One possibility to explain this is that on chromosome 1DS an α-gliadin or another gliadin gene containing the Glia-α20 sequence is present. This is in

agreement with reports suggesting that α-gliadin genes on chromosome 6 have originated from a homologous gliadin gene from chromosome 1 [Gu *et al.,* 2004; Shewry and Tatham, 1990]. This ancestral α-gliadin gene could be located on the short arm of chromosome 1D of Chinese Spring. Alternatively, the epitope could be rarely present in γ-gliadins, but was not detected in the set of sequences studied here. A more extensive analysis of database sequences could solve this, perhaps by mining large EST libraries.

Overall, among the three loci of the short arm of chromosome 1, the γ-gliadin locus on 1DS contributed more to CD-toxicity compared to 1AS and 1BS. For the short arm of chromosome 6, the α-gliadin locus on 6DS contributed the most to CD-toxicity, followed by 6A and 6B, respectively. We did not find differences in CD- toxicity between the two HMW-GS loci, because both the loci on 1BL and 1DL contained HMW-glt T-cell epitopes.

Relevance of the used deletion lines of Chinese Spring wheat

Based on the knowledge that gluten genes are mainly located on chromosomes 1 and 6 [Shewry *et al.,* 2003; Sabelli and Shewry, 1991] various deletion lines missing parts of these chromosomes were selected. Deletions of a complete gluten locus (as is the case of these deletion lines) results in the complete absence of the gluten genes encoded by this locus, and of several other genes near this locus. Chinese Spring is extensively used as a model for hexaploid wheats in both genetic and proteomic studies [Payne *et al.,* 1987b; Islam *et al.,* 2003]. In spite of this, the results obtained from this study cannot automatically be extrapolated to all wheat varieties. Various studies have shown that different wheat varieties contain different HMW-GS and that these have different effects on technological parameters [Don *et al.,* 2003b; Payne *et al.,* 1987a; Lafiandra *et al.,* 1993; Popineau *et al.,* 1994]. Using acid-PAGE analysis differences in the LMW and gliadin loci were shown among cultivars [Metakovsky *et al.,* 1993]. Differences in contribution to technological parameters were observed among flours with different LMW-GS loci [Eagles *et al.,* 2004; Branlard *et al.,* 2001] and flours with different gliadin contents [Van Lonkhuijsen *et al.,* 1992].

Contribution to toxicity and technological properties of different gluten loci: Possibilities of lowering CD-toxicity with a minimal negative effect on technological properties

We observed that all HMW-GS present in Chinese Spring contain the HMW-glt epitope. Encoded by chromosome 1B: HMW-GS 7 (x-type) and 8 (y-type) and by chromosome 1D: HMW-GS 2 (x-type) and 12 (y-type) [Harberd *et al.,* 1986]. We observed that technological properties were strongly affected by the removal of the different HMW-GS (Table 2) with the strongest effect in deletion line 1DL-4. Dough strength (DDT, GMP volume) and dough elasticity ($T_{1/2}$) were both highly decreased which is in agreement with published results [Don *et al.,* 2006; Wieser and Kieffer, 2001; Wieser *et al.,* 2007]. Deletion of the locus on 1AL resulted in some increase in dough strength (DDT and GMP volume) and elasticity ($T_{1/2}$). In line 1AL-1, also glutenin particle sizes were significantly increased. Both the x-type and y-type genes of Chinese Spring at *Glu-A1* are silent [Harberd *et al.,* 1986]. In most studies the silent locus at *Glu-A1* was not found to be important to determine dough strength compared to non-silent loci [Eagels *et al.,* 2004; Branlard *et al.,* 2001].

Based on obtained results we consider the *Glu-1* loci of Chinese Spring inappropriate as a focus to breed for less CD toxic wheat if technological properties are to be preserved. HMW-GS without T-cell epitopes do exist, as indicated by our database search. Our database search indicates that 46% of the database HMW-GS sequences did not contain the HMW-glt T-cell recognition site (Table 2). Varieties containing such HMW-GS could be appropriate in a low CD-toxic variety, regarding that technological criteria are met.

On the short arm of chromosome 1 the gluten loci mainly encoding the LMW-GS, γ- and ω-gliadins are located. We observed that the genes on the short arm of chromosome 1DS had a larger contribution to CD-toxicity than genes on 1AS and 1BS. These differences are mainly due to differences in the γ-gliadin proteins because the Glia-α9 mAb can recognise γ-gliadins but no LMW-GS, while the ω-gliadins migrate at higher levels on a SDS-PAGE gel. Technological measurements revealed that deleting part of the short arm of chromosome 1A led to an increase in DDT and GMP volume. A decrease in LMW-GS or gliadins results in a relative increase in HMW-GS/LMW-GS or glutenin/gliadin ratios. Such a change is suggested to increased dough strength [Gupta *et al.,* 1993]. The removal of the locus on chromosome 1AS also resulted in

increased dough elasticity. In the deletion lines 1AS-1 and 1DS-1 an increased GMP volume was observed, while in line 1BS-10 a decreased GMP volume was found together with a decreased DDT. This suggests that a LMW-GS encoded by the *Glu-B3* locus is important for the formation of the glutenin polymer. This is in agreement with Masci *et al.* [1998] who characterised a LMW-GS from hexaploid bread wheat that is a major component of the glutenin polymer. The authors indicate that this LMW-GS is similar to LMW-GS genes encoded by the *Glu-B3* locus. Subsequently, the increase in GMP volume in the deletion lines 1AS-1 and 1DS-1 can be explained by a higher expression of *Glu-B3* encoding LMW-GS because of compensation for lost *Glu-A3* or *Glu-D3* encoding LMW-GS. Compensation behaviour in wheat because of down regulation of storage proteins after inhibition of the expression of α-gliadins by RNAi was observed by Wieser *et al.* [2006]. However, line 1BS-19/6DS-4 did not show a decreased GMP volume. The LMW-GS locus (*Glu-B3*) is located more proximal to the centromer than the γ-gliadin locus (*Gli-1*) [Shewry *et al.,* 2003b]. The deletion in line 1BS-10 is more proximal to the centromer than the deletion in line 1BS-19/6DS-4 (Figure 1B). Therefore, the *Glu-B3* locus could be deleted in 1BS-10 and not in 1BS-19/6DS-4. We did not observe any LMW-GS disappearing from the immunoblot using the LMW-1 and 2 mABs, which is possible if this LMW-GS has no or very few LMW epitopes. To summarise the results for the deletion on the short arm of chromosome 1, we consider lowering expression from the short arm of chromosome 1D as a strategy to lower CD-toxicity which would not affect technological properties too much.

On the short arm of the group 6 chromosomes the gluten loci that encode for α-gliadin proteins are located. In the adaptive immune response as well as the innate immune response, the α-gliadins are considered the most toxic [Arentz-Hansen *et al.,* 2002; Vader *et al.,* 2002b; Molberg *et al.,* 2003; Palova-Jelinkova *et al.,* 2005]. We observed that the locus on chromosome 6DS, located between 6DS-6 and 6DS-2, is responsible for the most CD-toxic proteins. The locus located on 6BS had the least contribution to toxicity. These results are in agreement with the results obtained by Molberg *et al.* [2005]. These authors showed no decrease in T-cell responses for 6DS-6 proteins but a significant decrease in T-cell responses for 6DS-2 proteins. With respect to technological properties, line 6AS-1 showed an increase in GMP volume and a strong increase in average glutenin particle size. In contrast, deletions of the loci on 6B and 6D showed a decrease

in average glutenin particle size. It was shown that α-gliadins from the three different genomes are structurally distinct [Chapter 2] and are differently expressed at different times during kernel development [Kawaura *et al.*, 2005]. Lew *et al.* [1992] detected an α-gliadin (also called C-type LMW glutenin) in a purified glutenin fraction. Such an α-gliadin protein with one free cysteine residue can function as a chain terminator, while LMW-GS and HMW-GS with two or more free cysteine residues can function as chain extenders of the gluten polymer. We suggest that the short arm of chromosome 6A of CS is encoding a chain terminating α-gliadin. The decrease of chain terminators could cause the gluten polymer to grow larger and therefore explain the larger glutenin particle size as observed in line 6AS-1. Because of compensation, deletions of 6BS and 6DS can lead to an increased expression of chain terminating α-gliadins encoded by chromosome 6AS and result in the observed smaller particle sizes.

The deletions of 6BS and 6DS resulted in stronger dough (increased DDT). This effect on dough strength is expected because a decrease in α-gliadins results in a relative increase in the glutenin/gliadin ratio. This ratio is reported to correlate to dough strength [Wieser *et al.*, 2007]. Deletions of 6BS and 6DS also led to decreased dough elasticity (Figure 5C). GMP volume of 6DS-2 flour was decreased (indicating a weaker dough), while DDT was increased (indicating a stronger dough). These results indicate that in line 6DS-2 the decrease in GMP volume resulted in decreased elasticity rather then decreased dough strength. Breeding for decreased CD-toxicity by focusing on to decrease the expression of the proteins encoded by the locus on 6DS may have a significant decrease in CD-toxicity but will most likely result in a wheat variety with decreased dough elasticity. The decreased elasticity could be the result of an increased glutenin/gliadin ratio. This ratio can be corrected for by adding non-CD-toxic monomeric proteins to the flour, for instance avenins from oats. Progeny of crossings using the two deletion lines 1DS-1 and 6DS-2 can result in a line with decreased CD-toxicity without affecting technological properties too much. The decreased technological properties could possibly be corrected using monomeric proteins.

In summary, deletions targeted to remove HMW-GS on the long arm of chromosome B and D lead to a loss of technological functionality and a moderate loss of CD toxic epitopes.

Deletions targeted to remove LMW-GS and γ- and ω-gliadins leads to moderate changes in technological properties. The 1BS-10 deletion is an exception to this, presumably because of a loss of an essential LMW-GS, since the 1BS-19/6DS-4 (a shorter deletion on the same chromosome) does not lead to significant losses in technological functionality. With respect to toxicity, the 1DS-1 deletion is the most effective.

Deletions on the short arm of chromosome 6A, 6B and 6D are meant to remove α-gliadins. In general, this leads to stronger and stiffer dough. This is likely to be caused by a change in glutenin/gliadin ratio, and can be corrected for. With respect to loss of CD-toxicity, deletions of 6D are the most effective.

Conclusions

Our results show that a strategy to breed for low CD-toxic wheat while retaining technological properties is feasible if focused on lowering the gluten proteins encoded by the short arm of chromosome 1D and the short arm of chromosome 6D. This will lead to an extensive removal of CD-toxic epitopes. The changes in most technological properties could be corrected for by adding monomeric proteins like for example avenins from oats.

Acknowledgements

This research was partially funded by the Celiac Disease Consortium, an Innovative Cluster approved by the Netherlands Genomics Initiative and partially funded by the Dutch Government (BSIK03009). This research was partially funded by the Allergy Consortium Wageningen (ACW).

We are grateful to Wim Lichtendonk for his help during discussions on the technological part of this work and for his help during the technological experiments.

Chapter 6

Silencing epitope-specific alpha gliadin genes using siRNA on specific SNPs

Teun WJM van Herpen, Hetty C van den Broeck, Elma Salentijn, Thomas Bastien, Dirk Bosch, Ingrid M van der Meer, Peter A van Veelen, Frits Koning, Mandy Riley, Caronline Sparks, Huw D Jones, Peter R Shewry, Rob J Hamer Luud JWJ Gilissen, Marinus JM Smulders

Abstract

Celiac disease (CD) is an immune system-mediated disorder to wheat gluten that occurs in approximately 1% of the Caucasian population. Among the gluten proteins, the α-gliadins are most often associated with CD, as epitopes derived from α-gliadins are recognized by T cells from almost all CD patients. Not all α-gliadins contain all epitopes in an intact form, as specific SNPs in the gene sequence lead to amino acid differences, notable between α-gliadins from the three homologous loci in hexaploid wheat. Miller *et al.* [2003] showed in mammalian cell models SNP specific silencing by using siRNA. Following the same strategy, we attempted to specifically down-regulate epitope-containing α-gliadins, while non-epitope α-gliadins would be retained. We predicted that this would lead to less CD-toxic cultivars with good technological properties. We designed our siRNA to specifically down-regulate α-gliadins containing the DQ8-α-Glia epitope.

As a result of our transformations, we observed a different storage protein expression pattern as shown by SDS-PAGE compared to the control line. 2D analyses revealed several down-regulated spots in the gliadin region. Sequencing of these spots revealed that these were gamma gliadins. Expression analysis showed that in three lines the total α-gliadin mRNA was up-regulated and in one line this was down-regulated. An altered mRNA balance was found in two out of four transgenic lines. These two lines showed opposite results. One line had relatively less

epitope-containing α-gliadin mRNA and the other line had relatively more. Different transgenic events seemed to influence this siRNA approach. These results could indicate that our approach resulted in off-targeting and/or transitivity.

Artificial micro-RNAs (amiRNAs) can provide an alternative design to down-regulate genes with less off-targeting and transitivity. The amiRNA approach could allow silencing of individual isoforms. It was shown that one SNP at the cleavage site strongly influenced silencing of a transcript and could therefore be used to study the individual activity on two alleles. However, as yet, an approach that can distinguish between a single SNP has not been accomplished in plants.

Introduction

Celiac disease (CD) is an immune system mediated disorder with symptoms including chronic diarrhoea, osteoporosis and lymphoma. Up to 1% of the western population suffers from CD, although the disease is estimated to be highly undiagnosed [Fasano and Catassi, 2001; Fasano, 2006]. Until today, a complete and life long elimination of gluten from daily diet is the only option to combat the symptoms. Among CD patients the compliance to a strict gluten-free diet is rather low. Mariani *et al.* [1999] found a compliance rate of 53%. These authors also found that a strict gluten free diet may be a nutritional risk because it may lead to incorrect nutritional choices [Mariani *et al.*, 1999]. Taking the example of oats, low toxic gluten could provide an alternative to a gluten-free diet. It could increase the compliance rate and nutritional aspects of the diet, as a cultivar low in T-cell epitopes may possibly be tolerated by some CD patients. Since population incidence is related to gluten exposure [Ivarsson *et al.*, 2000; Fasano, 2006], a diet with low CD-toxic gluten would imply a much reduced exposure to epitopes which could also be of importance for the prevention of CD in susceptible individuals or could postpone clinical complaints in susceptible individuals.

Wheat gluten is composed of gliadins and glutenins with a number of subgroups within each group. Among these proteins the α-gliadin proteins are most often associated with CD as

peptides derived from α-gliadins are recognised by T-cells from almost all CD patients whereas T-cell responses to other gluten proteins are much less frequent [Arentz-Hansen *et al.,* 2002; Vader *et al.,* 2002b; 2003; Molberg *et al.,* 2003]. In addition, an α-gliadin-peptide (p31-43) was implicated in the innate immune system response in CD patients [Maiuri *et al.,* 2003].

The α-gliadins of wheat (*Triticum aestivum*) are encoded by three multicopy loci. Estimates for copy number range from 25-35 copies [Harberd *et al.,* 1985] to 150 copies [Anderson *et al.,* 1997]. A considerable amount of sequence data on CD-toxic epitopes amongst the α-gliadin family is available [Chapter 2]. Epitopes are non-randomly distributed over the gene sequences [Chapter 2; Molberg *et al.,* 2005; Spaenij-Dekking *et al.,* 2005], the absence of epitopes being caused by amino acid replacements due to locus-specific SNPs in the α-gliadin gene sequences [Chapter 2].

Lowering CD-toxicity in bread wheat can be achieved by decreasing the production of gluten proteins with T cell stimulatory epitopes using genetic modification. Becker *et al.* [2006] demonstrated that inhibition of the expression of the complete α-gliadin family can be achieved with RNA interference (RNAi). The effect of this drastic experiment on baking quality revealed that gluten strength was strongly increased and bread volume was slightly lowered [Wieser *et al.,* 2006]. A less drastic approach would be to only lower the expression of α-gliadin genes which contain CD-toxic epitopes. This would hardly affect technological properties (baking quality).

In 1999, siRNAs were first discovered as part of post-transcriptional gene silencing in plants [Hamilton and Baulcombe, 1999]. Shortly thereafter, synthetic siRNAs were shown to be able to induce RNAi in mammalian cells [Elbashir *et al.,* 2001]. Typically, siRNAs are 19 to 24 nucleotides in length and produced from fold-back precursors that are transcribed from imperfect inverted repeats in the genome. Sequential processing of the siRNA precursor by the double-strand specific RNases in plants [Kurihara and Watanabe, 2004] produces a stable small RNA that is incorporated into the silencing complex. Perfectly complementary siRNAs are widely used in animals as a tool to down-regulate RNA expression of genes of interest [Hannon and Rossi, 2004]. Miller *et al.* [2003] showed that in mammalian cell models allele-specific silencing of disease genes could be achieved by targeting a linked SNP using the siRNA. Designing siRNA that distinguish between genes that differ by only a single nucleotide is a powerful tool to reduce

the expression of a single gene in human cells [Schwartz *et al.*, 2006]. Following the same design strategy, targeting SNPs in the α-gliadins related to the presence or absence of perfect CD-toxic epitopes could result in less CD-toxic cultivars with good technological properties.

After having obtained more information about the α-gliadins family [Chapter 2], we observed that one SNP can result in two isoforms of the α-gliadin protein. One isoform contains the perfect DQ8-α-Glia (Glia-α) epitope QGSFQPSQQ and another that contains QGSFRPSQQ. The glutamine (Q) residue at position five is important because it can enhance T-cell recognition after deamidation [Koning, 2003]. A substitution of a Q to R is therefore expected to decrease T-cell recognition. Down-regulation of these α-gliadins while retaining the α–gliadins without this epitope could result in decreased toxicity of the gluten with a minimal effect on technological parameters. Therefore, in this study we designed a siRNA molecule targeted at α-gliadins containing the DQ8-α-Glia-(206-217) epitope [Koning, 2003]. This sequence is mainly present in α-gliadins encoded by the loci on chromosome 6B and 6D (i.e., the B and D genome of bread wheat) [Chapter 2]. The α-gliadin genes on the D genome are considered to exert the most influence on the immune response compared to A and B genome α-gliadins [Molberg *et al.*, 2005, Spaenij-Dekking *et al.*, 2005]. The transgenic wheat plants containing the siRNA construct were analysed for total and relative expression levels and for gluten proteins expression.

Experimental

Constructs

The ubiquitin promoter was amplified from the construct pAHC25 using the following primers: 5'-GCC GGG ATC CTG CAG AAG TAA CAC CAA ACA ACA GGG-3' and 5'-GCC GGG ATC CGT GCA GCG TGA CCC GGT CGT-3'. This introduced two BamHI restriction sites (underlined in the primer sequence) on both sides of the promoter. The PCR product was subsequently cut with BamHI and ligated into a BamHI and BglII digested pFLAP10 construct. This resulted in a plasmid with an ubiquitin promoter and Nos Terminator with one BamHI

restriction site in between. After the ligation of the ubiquitin promoter in the pFlap10 plasmid, the correct orientation was checked using digestion and the construct finally verified by sequencing.

The active shRNA (short hairpin RNA) molecule contained a 19 bp double-stranded region targeting the DQ8-α-Glia-(206-217) epitope and a loop of nine nucleotides (Figure 1). A loop of nine nucleotides was found to be most effective in down-regulating gene expression by Brummelkamp *et al.* [2002]. At position nine the SNP directed to select for epitope or non-epitope was positioned [as in Miler *et al.,* 2003; Brummelkamp *et al.,* 2002].

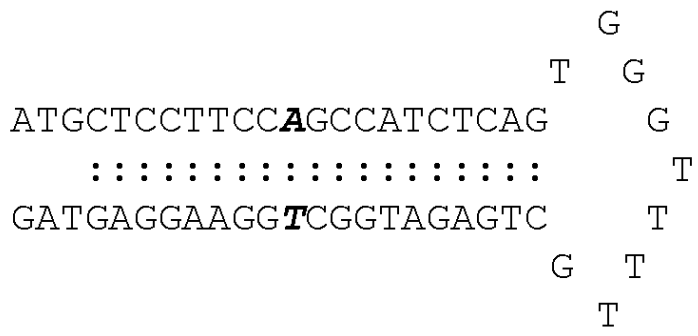

Figure 1 - Predicted configuration of the designed shRNA molecule. The molecule contains a 19 bp double stranded targeting α-gliadin sequence, and a loop of nine nucleotides to target specifically the DQ8-α-glia-(206-217) epitope. At position nine the SNP directed to select for epitope or non-epitope was positioned.

To generate the shRNA part two partially complementary 59 bp DNA sequences were designed: forward strand 5'-GA TCC ATG CTC CTT CCA GCC ATC TCA GTG GGT TTT GCT GAG ATG GCT GGA AGG AGT AGG-3' and reverse strand 5'-GAT CCC TAC TCC TTC CAG CCA TCT CAG CAA AAC CCA CTG AGA TGG CTG GAA GGA GCA TG-3' (the loop sequence of 9 bp is underlined). After annealing, these two primers gave a BamHI overhang which was, after kinase treatment, ligated into the BamHI digested ubiquitin promoter-pFlap10 construct. The new construct was verified by sequencing.

The pAHC25 selectable marker construct [Christensen and Quail, 1996], which contains the bar gene encoding phosphinothricin acetyltransferase (PAT), was used to give resistance to the herbicide phosphinothricin (PPT).

Transformation

Immature embryos of bread wheat cv Cadenza were used as targets for transformation by particle bombardment using the protocol of Sparks and Jones [2004]. The following modifications were made: regeneration media contained 0.05 mM $CuSO_4$ instead of 10 mg/l $AgNO_3$; selection media did not contain 2,4D; bombardments were carried out at a pressure of 900 psi. The presence of the transgenes in putative transgenic plants was confirmed by PCR. Control lines were bombarded with gold only and went through the same regeneration process. Transformed plants of several generations were grown in containment glasshouses. No phenotypic changes were observed.

Transformed progeny lines

A total of fifteen independent T_0 lines comprising the siRNA genotype were initially shown by PCR amplification to contain the shRNA construct. Fourteen of these lines gave seeds. Twelve individual T_1 seeds from these fourteen transgenic lines and one control line were grown for T_1 plant generation. PCR amplification identified that four of these fourteen putative transgenic lines inherited the construct in the T_1 generation, as tested on 12 different T_1 progeny seeds. From these four lines, we selected one progeny plant each that was PCR positive (control line: R1P1.2, transgenic lines: R3P2a.3, R3P4.7, R3P8b.3 and R11P1.6). These plants were grown as T_2 lines which produced the T_3 seeds that were analysed in detail.

DNA isolation, RNA isolation and cDNA synthesis

Genomic DNA was isolated from the different transgenic lines, from young leaves of different T_2 progeny seedlings and from T_3 half seeds harvested at 30 DAF. Total RNA was isolated from grains harvested at 30 DAF. Both isolations were performed according to the method described by Doyle and Doyle [1990] but with 1% (w/v) poly-(vinylpyrrolidone)-10 in the extraction buffer. For the isolation of first strand cDNA 1 µg of total RNA was treated with DNAse I (Invitrogen,

UK) followed by RT-PCR (Invitrogen™, UK, SuperScript™ III First-Strand Synthesis System for RT-PCR) using oligo (dT_{16-20}) primers. Samples without SuperScript™ III reverse transcriptase (minus RT-PCR) were included as controls for the DNAse I treatment.

Transgene detection

To specifically detect the transgene in the progeny plants and seeds of different transgenic lines, two construct-specific PCR primers were designed. Amplifications were performed in 50 µl reaction volume, containing 0.4 µM forward primer (5'-ATT GCC AAA TGT TTG AAC GA-3'), and 0.4 µM reverse primer (5'-TTT AGC CCT GCC TTC ATA CG-3'), dNTP mix (0.2 mM each), 1.5 mM of $MgCl_2$, 1x Goldstar buffer (Eurogentec Belgium), and 1U Goldstar DNA polymerase (Eurogentec Belgium) (5 U/µl). Primers were based on the Ubiquitin promoter (forward primer) and T-nos terminator (reversed primer). An annealing temperature of 57°C was used.

Pyrosequencing

The pyrosequence protocol is described by Salentijn *et al.* [in prep]. In the pyrosequence analysis cDNA of grains, previously determined to be PCR positive (PCR performed on half of the seed) were used. For line R1P1.2.1 (control) four grains, for line R3P2a.3.1 six grains, for line R3P4.7 three grains, for R3P8b.3.1 three grains and for line R11P1.6.1 four different grains were studied.

Amplifications were performed for pyrosequencing from 2 µl cDNA sample. Amplifications were performed in 50 µl reaction volume, containing 0.4 µM reverse (α-3prime-R1: 5'-Biotin-TGG AGG GAT *R*TA BAC ATT GC-3'), and 0.4 µM forward primer (α-3prime-F1: 5'-CAG *Y*CT C*WR* *B*A*R* CAA TAT CC-3'), dNTP mix (0.2 mM each), 1.5 mM of $MgCl_2$, 1x Goldstar buffer (Eurogentec, Belgium), and 1U Goldstar DNA polymerase (Eurogentec, Belgium) (5 U/µl). PCR program used was: 94°C for 5 min followed by 50 cycli of 94°C for 30 sec, 51°C for 1 min and 72°C for 2 minutes.

The PCR product was linked to streptavidine sepharose HP beads (Amersham Biosciences, UK) by incubation for 10 min at room temperature while shaking. Subsequently, the biotinylated PCR product was isolated using a Vacuum prep tool and used as template for

pyrosequencing. For this, the PCR product was washed in 70% (v/v) ethanol; denatured in 0.2M NaOH and neutralized in 10 mM TRIS-acetate, pH 7.5. Next, the biotinylated strand was transferred to 45 µl primer solution (0.3 µM primer in 20 mM TRIS-acetat, 2 mM $MgAc_2$). Pyrosequencing was done with a degenerated primer that consisted of the primers α3'-SQ1a to α3'-SQ1e mixed in equally amounts (0.06 µM each): α3'-SQ1a 5'-CTC TGC AAC AAT ATC CAT-3';α 3'-SQ1b 5'-CTC AGC AAC AAT ATC CAT-3'; α 3'SQ1c 5'-CTCA GGA ACA ATA TCC AT-3'; α 3'-SQ1d 5'-CTC AGC AGC AAT ATC CAT-3'; α 3'-SQ1e 5'-CTC AAC AAC AAT ATC CAT-3'. The pyrosequencing reaction was performed with the following nucleotide dispensation order: GTCTGAGTCAGTACTCGTCGAGTCATCTCAGCTAG.
Comparison of samples was done on each individual peak using Multivariate comparisons (ANOVA test performed using the statistical computer program GenStat Release 9.2 with a 95% significance level).

Real Time PCR

RNA expression levels of total α-gliadin expression in the individual seeds at 30 DAF from each line were tested using real time PCR on a MyIQTM Single Color Real Time PCR detection system (Bio Rad Laboratories, Germany). All reactions were performed in duplicate using two different primer sets. Reaction mixes consisted of 45 µl containing: 22.5 µl SYBR green PCR mastermix (Applied Biosystems, CA), 4.5 µl forward primer (3 µM), 4.5 µl reverse primer (3 µM), 9 µl MQ, and 4.5 µl cDNA (0.02 µg/µl). Amplification program utilised was as followed: 3 min at 95°C, 15 sec at 95°C, 1 min at 60°C utilising 40 cycles. Specific primer sets used were 1) the same set as used for pyrosequencing and an Alpha-3 primer set (forward primer: 5'-AGC CAG GTC TCC TTC CAA CAG-3'; reverse primer was: 5'- CTG TAG GGC TAG GTT CCT TAT TTC C-3'). As an internal control for amount of RNA Cytochrome b was amplified using the following primers: forward: 5'-TTT CGT GGT CTA TAT CAT GCG AGT TA-3'; reversed: 5'-AAA GCT CAT CTG ACC CCA AGG T-3'.

DUMAS protein determination

For each line, three mature T_3 grains were harvested from three different T_2 progeny. For line R3P8.3 only two seeds from two progeny were tested. First the bran was removed from the kernel using a razor blade. Subsequently the kernels were ground using mortar and pestle. Approximately 15 mg of flour was used to analyse the protein contents by the Dumas method [Sebecic and Balenovic, 2001] using an NA2100 Nitrogen and Protein Analyzer (ThermoQuest-CE Instruments, Italy). The Dumas method is based on the measurement of total nitrogen in the sample. Methionine was used as a standard.

Protein isolation

Individual, mature, T_3 grains from two or three different progenies of the different transgenic lines were harvested. Prolamins were extracted according to Van den Broeck *et al.*, [2008]. Grains were ground in an analytical mill (A 11 Basic, IKA-Werke) and sieved through mesh (0.5 mm). Prolamins were extracted from 50 mg wheat flour by addition of 0.5 ml of 50% (v/v) iso-propanol with continuous mixing (MS1 Minishaker, IKA Works, Inc.) at 1000 rpm for 30 min at room temperature, followed by centrifugation at 10,000 rpm for 10 min at room temperature. The residue was re-extracted twice with 50% (v/v) iso-propanol/1% (w/v) DTT/50 mM Tris-HCl, pH 7.5, for 30 min at 60ºC with mixing every 5 to 10 min followed by centrifugation at 10,000 rpm for 10 min at room temperature. After addition of each next extraction solution, the residue was resuspended by shaking in a Fastprep FP220A Instrument for 10 sec at speed 6.5 m/sec followed by sonication for 10 min in an ultrasonic bath (Branson 3510, Branson Ultrasonics Corporation). The three supernatants were combined and considered as the prolamin extract. The protein content was quantified using Biorad Protein Assay (Biorad Laboratories), based on the Bradford dye-binding procedure, according to manufacturer's instructions.

SDS PAGE

Proteins were separated on SDS PAGE gels (10%) as described by using a SE260 mighty small II system (GE Healthcare, UK). Gels were fixed in 40% (v/v) ethanol/10% (v/v) acetic acid in water for 30 min. Subsequently, gels were stained with Sypro Ruby (Invitrogen, UK).

2D-electrophoresis

In 2D-electrophoresis, proteins were separated in the first dimension on Immobiline Drystrip pH 3-10 of 24 cm (GE Healthcare, UK). For overnight rehydration of IPG strips, protein (100 µg) was combined to a final volume of 450 µl in rehydration buffer (6 M urea, 2 M thio-urea, 2% (w/v) CHAPS, 20 mM DTT) complemented with 0.5% IPG buffer pH 3-10 (GE Healthcare, UK) according to manufacturer's instructions. The rehydrated strips were focused on an IPGphor (GE Healthcare, UK) at 50,000 Vh. Prior to the second dimension, strips were equilibrated for 15 min in equilibration buffer (6 M urea, 30% (w/v) glycerol, 2% (w/v) SDS, 50 mM Tris-HCl, pH 8.8) containing 1% (w/v) DTT, followed by 15 min in equilibration buffer containing 2.5% (w/v) iodoacetamide. Separation in the second dimension was performed on SDS-PAGE gels (10% acrylamide) using the Hoefer ISO-Dalt System (GE Healthcare, UK). After electrophoresis gels were stained with Sypro Ruby (Invitrogen, UK), and post-stained with PageBlue (Fermentas, USA) for spot-isolation.

Protein sequencing

In-gel digestions were performed with the use of the Proteineer DP digestion robot (Bruker, Germany). The plate holder was adapted in-house to accommodate a hydrophilic 96-well filter plate (Millipore MSRL N04). Selected gel spots, isolated from a PageBlue stained gel, were placed in the 96 well filter plate and washed with a solution of a 1/1 (v/v) 50 mM ammonium bicarbonate pH 8.3/CAN (acetonitrile), followed by washing with 100% water. Next, a reduction step was performed with 10 mM DTT in 50 mM ammonium bicarbonate pH 8.3 at 56°C, followed by alkylation step using 55 mM iodo acetamide in 50 mM ammonium bicarbonate pH 8.3 at 22°C. Subsequently, the gel pieces were thoroughly washed with 50 mM ammonium bicarbonate pH 8.3 at 22°C, and shrunk in 100% ACN. These latter steps were performed 3 times. A solution of 12 ng/µl chymotrypsin in 50 mM ammonium bicarbonate pH 8.3 was added and the gel pieces were allowed to swell for 45 min at 10°C, after which excess enzyme solution was removed. After swelling the gel pieces were transferred to polypropylene eppendorf tubes and 30 µl 50 mM ammonium bicarbonate pH 8.3 was added. Digestion took place overnight and

extraction was done by addition of 50/50/0.1 water/ACN/formic acid and 50/50/1 water/ACN/formic acid. Extracts were pooled and freeze dried.

For mass spectrometry, peptides were dissolved in 95/3/0.1 (v/v/v) water/ACN/formic acid and subsequently sequenced by tandem mass spectrometry. Peptides were analyzed by nanoflow liquid chromatography using an Agilent 1100 HPLC system (Agilent Technologies, USA) coupled on line to a 7-tesla LTQ-FT mass spectrometer (Thermo Electron, Germany). The chromatographic system consisted of the following components. ReproSil-Pur C18-AQ, 3 µm (Dr. Maisch GmbH, Ammerbuch, Germany) was used as a resin for the analytical nano column and AQUA-C18 5 µm was used as a resin for the trapping column. Peptides were trapped at 5 µl/min on a 1 cm column (100 µm internal diameter, packed in house) and eluted to a 15 cm column (50 µm internal diameter, packed in house) at 150 nl/min in a 60 min. gradient from 0 to 50% ACN in 0.1% formic acid. The eluent was sprayed via emitter tips (made in house) butt-connected to the analytical column. The mass spectrometer was operated in data-dependent mode, automatically switching between MS and MS/MS acquisition. Full scan MS spectra were acquired in FT-ICR with a resolution of 25,000 at a target value of 5,000,000. The two most intense ions were then isolated for accurate mass measurements by a selected ion monitoring scan in FT-ICR with a resolution of 50,000 at a target accumulation value of 50,000. These ions were then fragmented in the linear ion trap using collision-induced dissociation at a target value of 10,000. In a post analysis process, raw data were converted to peak lists using Bioworks Browser software, Version 3.1. For protein identification, MS/MS data were submitted to the Swiss Prot database using Mascot Version 2.1 (Matrix Science) with the following settings: 1–5 ppm and 0.8-Da deviation for precursor and fragment masses, respectively; no enzyme specification. All reported hits were assessed manually, and peptides with MASCOT scores lower than 40 were discarded.

Results

Transgene detection

PCR analysis on four individual T_2 progeny of the four PCR positive T_1 lines showed that the line R11P1.6 segregated for the transgene (2 out of 4 plants). For the three other lines all four progeny plants were positive. In R11P1.6.1, individual T_3 grains segregated for the construct (12 out of 21 grains) as well as in grains from R3P2a.3.1 (18 out of 21 grains). The two other lines did not show segregation among the grains, namely line R3P4.7 (18 out of 18) and R3P8b.3 (21 out of 21). These results may indicate a higher copy number present in the lines R3P4.7 and R3P8b.3 compared to the two other lines.

Total α-gliadin mRNA expression

The results from the Real Time PCR using α-gliadin specific primers show that the total α-gliadin expression levels were increased, rather than decreased in the lines R3P2a.3.1, R3P4.7.1 and R11P1.6.1 (Figure 2). The higher levels of mRNA in these lines may be due to compensation behaviour as observed on protein level by Wieser *et al.* [2006]. Line R3P8b.3.1 in contrast showed a decreased level of total α-gliadin mRNA. This line (together with line R3P4.7.1) also showed the least degree of segregation indicating a higher copy number in this line. A higher copy number is expected to result in a higher expression level (Chapter 3), which may result in a higher degree of down-regulating [Miki *et al.* 2005]. Miki *et al.* suggested that genes with higher levels of mRNA expression may be silenced more effectively than those with lower levels of expression. A higher expression of the siRNA construct could therefore also increase the down-regulation of the α-gliadin.

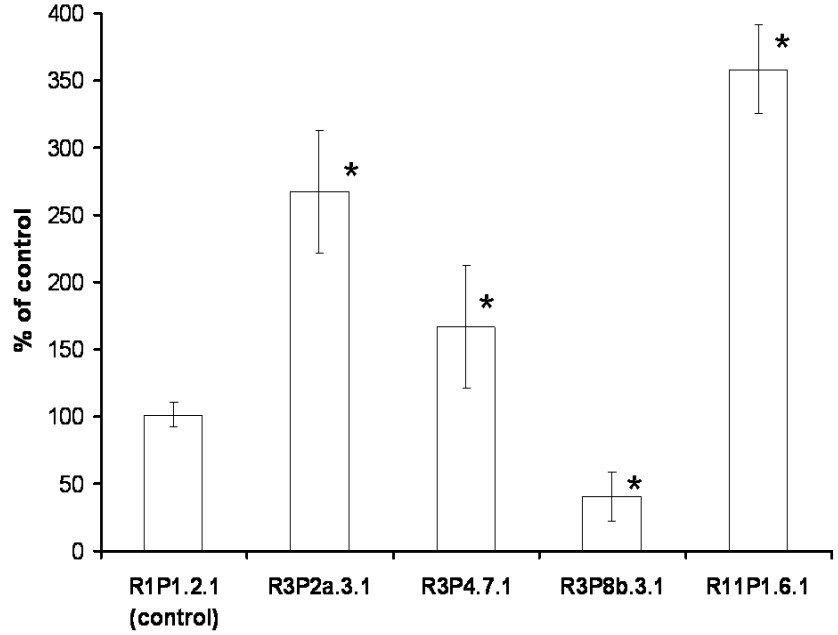

*Figure 2 - Total α-gliadin mRNA levels at 30 DAF. The graph shows the percentage of expression compared to the control for the different transgenic lines and the control line, *= a significant increase or decrease compared to control.*

Transcription balance of α-gliadin isoforms

Our siRNA strategy was intended to down-regulate the isoform of the α-gliadins containing the epitope Glia-α while retaining the other isoforms. To detect altered ratios between expressed α-gliadin mRNAs a pyrosequencing experiment was performed. Differences are expected in the relative amounts of nucleotides order number 20 (SNP G) and 21 (SNP A), as shown in figure 3. The SNPs represented by nucleotides order number 20 and 21 are linked to SNP T at number 2 (for 20) and SNP C at number 3 (for 21), figure 3 and Salentijn *et al.*, in prep.

Epitope-containing α-gliadins are D- or B-genome sequences and are represented by peak 21 (Figure 3). However, some B-genome sequences show the SNP at position 21 but have additional SNPs at other positions in the targeted sequences. If the targeted isoform represented by peak 21 is down-regulated the isoform represented by peak 20 will be up-regulated. This pyrosequencing study was done on several biological replications of cDNA from grains of 30 DAF (see experimental) and results were analysed using an ANOVA statistical test.

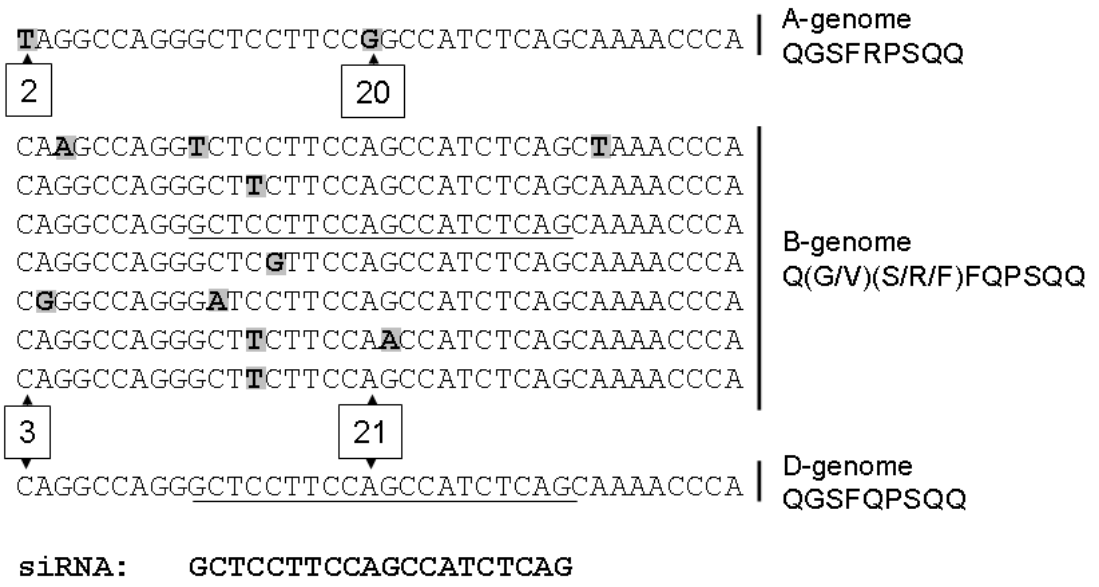

Figure 3 - SNPs in A, B and D genome EST sequences. In the figure the different SNP variations are shown in the region of targeting among the EST sequences [Salentijn et al., in prep]. A and D genome sequences show only one variation whereas for the B genome sequences seven variations were found. The amino acid sequence is also indicated. The expected siRNA molecule is depicted below these sequences. SNPs are shown in bold and gray boxes. Pyrosequence nucleotide order numbers 2, 3, 20 and 21 are indicated. Note that the designed siRNA is perfectly complementary to D-genome sequences and one B-genome sequence.

The data show a significant difference in nucleotides order number 20 (G) for the lines R3P4.7.1 and R11P1.6.1 (Figure 4). However, these lines show opposite results. Compared to control lines, line R3P4.7.1 has a higher level of non-epitope α-gliadin mRNAs, which will lead to a lower level of epitope-containing α-gliadins (expected results). Line R11P1.6.1, in contrast, shows the opposite effect (not expected). Changes in nucleotides order number 21 (A) were not statistically significant. For the nucleotides order number 2 (T) and 3 (C) only number 3 in line R11P1.6.1 was significantly higher compared to the control (data not shown). This correlates well with the effect observed at peak 20. Peak 2 was slightly higher in R3P4.7.1 and slightly lower in R11P1.6.1, but the differences were not statistically significant.

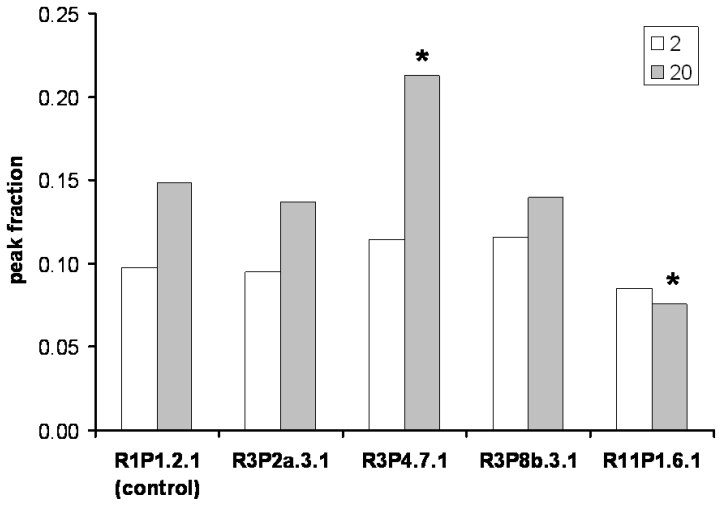

Figure 4 - Relative transcription levels of α-gliadin mRNA. The graph shows the pyrosequencing results from several biological repetitions at 30 DAF of nucleotide 2 (T) and 20 (G). Stars indicate a significantly different expression compared to the control line analysed using a multivariate comparisons (ANOVA, 95% significance level).

Flour protein content

Total flour protein content was higher in all transgenic lines compared to control lines (Figure 5). For line R3P2a.3 these results were not significant. For the other lines the protein content in flour was significantly higher than the control line. Apparently protein expression is increased in the modified wheat lines. We also observed an increase in total α-gliadin expression in three transgenic lines (Figure 2).

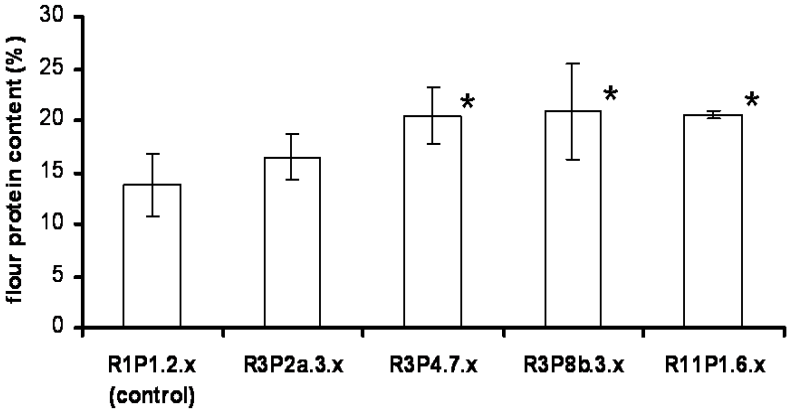

*Figure 5 - Protein content in endosperm. Dumas results were shown in this graph. Results are average values from three T2 progeny; "x" represents 3 progeny. On the y-axis flour protein content is represented. All transgenic lines show a higher level of protein content; *= significantly different from control.*

Different prolamin patterns in transgenic lines

Various mature seeds from individual transgenic lines were analysed on SDS-PAGE. The plants from which we harvested these seeds were previously determined to contain the transgene, as determined by PCR. However, individual grains can still segregate (see paragraph 3.1). In contrast to the pyrosequencing assay, we did not PCR checked each grain prior to SDS-PAGE analysis.

The prolamins show at least three distinct patterns on SDS-PAGE (Figure 6). One pattern is representing the control line (lanes 1-3). Another pattern shows an increased pattern in the LMW-GS/γ-/α/β-gliadin region. Lane 10, representing one individual seed from one T_2 progeny from line R3P8b.3 showed what appears to be a complete absence of α-gliadins. It also shows down-regulation of a large set of prolamins, including some proteins in the omega gliadin region. These results may indicate off-targeting, but the experiment must be repeated for other seeds as well before this can be concluded (Figure 2).

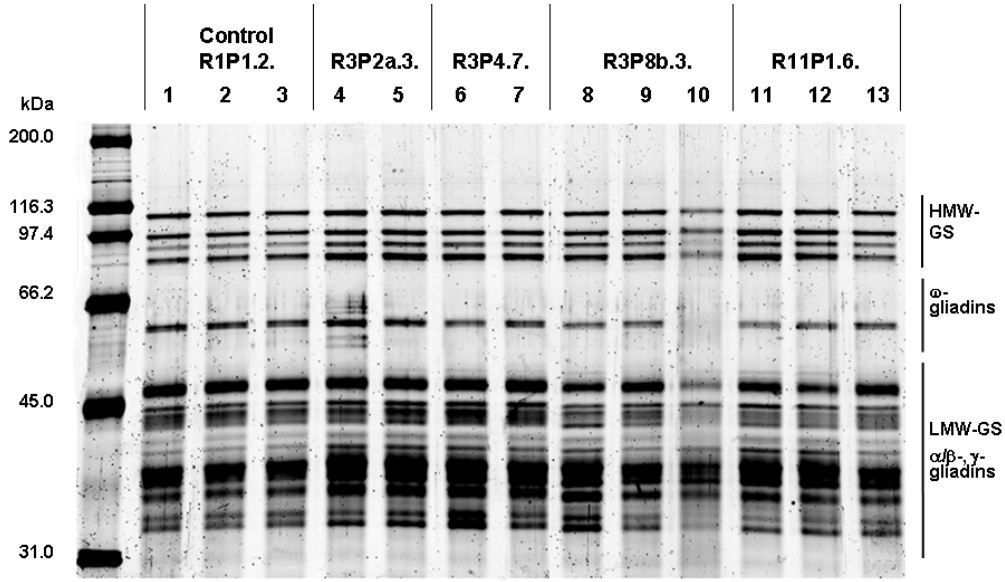

Figure 6 - Prolamin pattern analysis. Individual seeds from different control and transgenic lines analysed by SDS-PAGE (10%) followed by Sypro Ruby staining. Lanes: 1, 2 and 3: Control lines R1P1.2.1, R1P1.2.2, R1P1.2.3; 4 and 5: line R3P2a.3.1 and R3P2a.3.2; 6 and 7: line R3P4.7.1 and R3P4.7.2; 8, 9 and 10: R3P8b.3.1, R3P8b.3.2 and R3P8b.3.3; 11, 12 and 13: line R11P1.6.1, R11P1.6.2 and R11P1.6.3

Specific differences of individual protein in transgenic line

To analyse for specific differences in individual proteins we performed a 2D protein analysis on one control line (R1P1.2.1) and one transgenic line R11P1.6.1 (Figure 7). A protein extract from one seed from transgenic line R11P1.6.1 was used. The same sample as analysed on SDS-PAGE (Lane 11 in figure 6) was used. On 2D analysis we observed specific down-regulation of several protein spots in the LMW-GS/γ-/α/β-gliadin region. Three protein spots (encircled in figure 7) were cut out of the gel for protein sequencing.

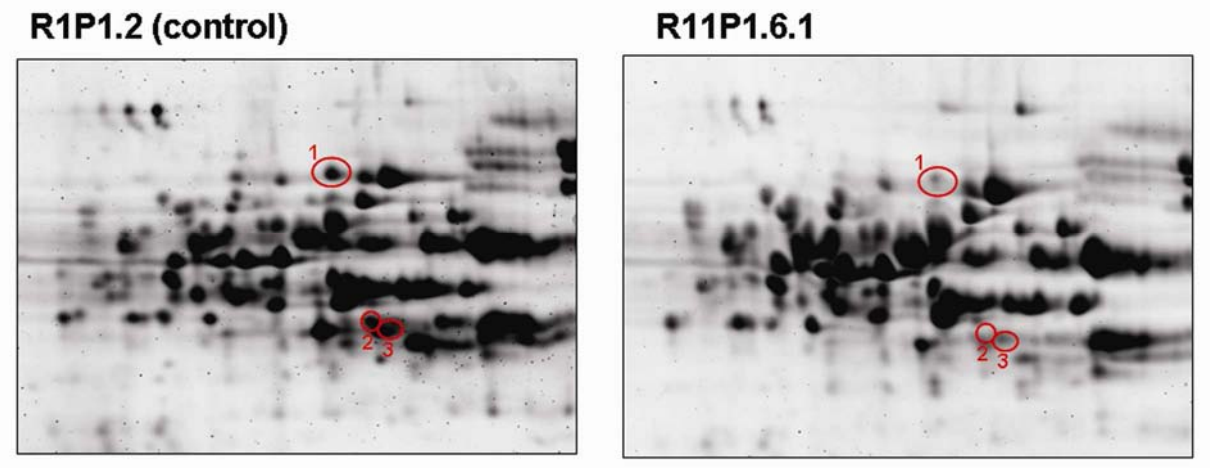

Figure 7 - 2D protein analysis. First dimensions were performed on a pH range of 3-10. The second dimension was performed on 10% SDS-PAGE gels. One seed from the control line (R1P1.2, also analysed and shown in figure 6, lane 1) was compared to transgenic line R11P1.6.1 that showed an interesting pattern on SDS-PAGE (Figure 6, lane 11). The lefthand figure shows the control line and the righthand figure shows line R11P1.6.1. Encircled spots were down-regulated in the transgenic line and were subsequently sequenced.

Identification of down-regulated proteins

Specific spots encircled in figure 7 were sequenced. These specific spots were confirmed to be gliadins, but from the γ-gliadin family. Spot 1 was sequenced and resulted in a coverage of 18% (58 out of 327 amino acids in 14 unique peptides) to the gamma-gliadin precursor (P08453). For spot 2 a coverage of 21% (61 out of 291 amino acids in 5 unique peptides) to a gamma-gliadin B precursor (P06659) was found. Sequencing of spot 3 obtained a coverage of 31% (91/291 amino

acids in 7 unique peptides) to a gamma-gliadin B precursor (P06659). Similar to the results obtained in the pyrosequencing analysis, these results may indicate off-targeting of the siRNA. An increased expression of total α-gliadin was also observed for this line.

Discussion

Until today, an approach to down-regulate genes that can distinguish between targets that differ by a single SNP has not been accomplished in plants [Ossowski *et al.,* 2008]. In this study we obtained several transgenic lines which inherited a siRNA construct. These transgenic lines had overall a normal phenotype. Although, the transgenic lines used for detailed analysis may still segregate for the presence of one or more constructs, we did observe differences in storage protein expression and deposition in the transgenic lines compared to control lines. Results varied between the different progeny seeds of one transformed plants probably as a result of segregation. However, results varied also between the different transgenic lines. Different transformation events are known to result in different expression levels, or other, so called, positional effects and this may explain differences seen among the different lines.

One line (R11P1.6) showed unexpected changes in the balance of expression of α-gliadin mRNAs targeted compared to non-targeted mRNAs, as the target sequence increased compared to the alternative sequence in the mRNA pool of developing seeds of this line. The seeds also lacked some protein spots compared to the untransformed control, but these turned out to be γ-gliadins, indicating off-targeting.

Another line (R3P4.7) did show the expected results in the pyrosequencing experiment. 2D-analysis followed by protein sequencing can confirm if this changed expression balance is also reflected on α-gliadin protein levels and in the presence of specific isoforms.

A possible explanation for the off-targeting observed

Perfectly complementary siRNAs are widely used in animals as a tool to down-regulate RNA expression of genes of interest [Hannon and Rossi, 2004]. Designing siRNA that distinguish between genes that differ by a single nucleotide is suggested to be a powerful tool to reduce the

expression of a single gene in human cells [Schwartz *et al.*, 2006]. The designed siRNA from this study was based on previous research in animal cells [Miler *et al.*, 2003; Brummelkamp *et al.*, 2002]. The protein machinery involved in RNAi in plants is, however, somewhat different from animals. siRNAs can reduce protein levels through two mechanisms: translational repression and transcript cleavage. Translational inhibition is mainly found in animals and transcript cleavage appears to be the predominant mechanism in plants [Schwab *et al.*, 2005].

siRNAs can also affect mRNAs that are not perfectly complementary, generally termed 'off- or mis-targets' [Jackson *et al.*, 2003]. Naturally found plant siRNAs (also called miRNAs) have targets with up to five mismatches [Schwab *et al.*, 2005]. In our case the unexpected results could be due to the effect of off-targeting. Another problem observed when using siRNA approaches is called transitivity. Upon binding to target transcripts, siRNAs can serve as primers for RNA-dependent RNA polymerases. These extend the local RNA double strands and generate templates for production of secondary siRNAs by Dicer action [Voinnet, 2005]. These secondary siRNAs, which are unrelated in sequence to the initial trigger, can in turn affect other genes. Because of the high homology among gliadin genes [Shewry and Lookhart, 2003] transitivity could result in the down-regulation of different groups of gliadin genes. Because some γ-gliadin spots were found to be down-regulated in 2D analysis we can not exclude that transitivity has played a role in our approach. The down-regulated proteins in the omega-gliadin region seen in line R3P8b3.3 on 1D, would, if confirmed on 2D gels, indicate a high degree of transitivity.

Silencing of more than one gene was performed in rice by Miki *et al.* [2005]. These authors targeted multiple members in a gene family of rice using siRNA and showed that with one siRNA construct more genes can be down-regulated. However, the authors suggested that genes with higher levels of mRNA expression may in fact be better silenced than those with lower levels of expression. α-Gliadins can have different expression patterns and levels [Shewry and Halford, 2002], therefore the effect suggested by Miki *et al.* [2005] can also play an important role in our approach in down-regulating the α-gliadins from hexaploid wheat.

The preliminary analysis on the transgenic lines as reported here must be extended by the analysis of several seeds per transgenic line on 2D gels followed by sequencing of differential protein spots. This is particularly true for line R3P4.7.1. It would also be very useful to produce

subsequent generations of the four transgenic lines that are homozygous for the inserted constructs.

Even though the reported results are preliminary, and line R3P4.7.1 may turn out to be a very successful transgenic event with the results aimed for, progress in literature since the start of these experiments already has indicate that the design of our siRNA constructs was, by hindsight, suboptimal. Below we therefore indicate possible improvements for future siRNA constructs aimed to make wheat safe for CD patients.

Suggestions to reduce off-targeting

The use of a larger RNAi molecule as in Becher *et al.* [2006] is not expected to produce a very high specificity. Becher *et al.* [2006] showed down-regulation of a great number of α-gliadins by expressing a long RNAi molecule in bread wheat. Long double-stranded precursors generate a multitude of siRNAs with varying ends [Schwab *et al.*, 2006], which makes the prediction of off-targets particularly difficult. In fact it is possible that their approach has also silenced some gamma-gliadins.

As described by Schwab *et al.* [2005; 2006] artificial microRNAs (amiRNAs) can be a better option to down-regulate genes in plants than other siRNA techniques. Short hairpin RNAs, as used in our study, are intended to target perfectly complementary mRNAs, while amiRNAs preferentially avoid perfectly complementary targets in order to minimise problems caused by transitivity. amiRNA precursors generally generate only a single effective small RNA of known sequence. By contrast, several siRNAs with undefined ends are produced as a silencing trigger from hairpin constructs. In this way, potential off-targets of amiRNAs can be more accurately predicted than other hairpin constructs. The amiRNA approach could allow silencing of individual isoforms [Ossowski *et al.*, 2008]. It was shown that one SNP at the cleavage site strongly influenced silencing of a transcript, and could therefore be used to study the individual activity of two alleles [Schwab *et al.*, 2006]. However, an approach to down-regulate genes that can distinguish between a single SNP has not yet been accomplished in plants [Ossowski *et al.*, 2008].

Chapter 7

General Discussion

Coeliac disease (CD) is a disorder which is characterised in patients by a permanent intolerance to gluten proteins. CD presents itself by chronic diarrhoea, osteoporosis, lymphoma and several other clinical symptoms. Wheat of low CD-toxicity (with gluten proteins containing significantly lower levels of CD-toxic epitopes) would be beneficial for CD patients and in prevention of CD. Such low toxic wheat should still retain its good technological properties. The α-gliadins are considered a very important factor in causing CD because of their role in provoking innate as well as adaptive immune responses. By removing α-gliadins, a large amount of CD-toxicity could be eliminated. The effect of removing α-gliadins on technological properties was not well studied up to now. Gliadins have been described to mainly contribute to the viscosity and extensibility of the dough system. Several strategies aiming at the development of low CD-toxic hexaploid bread wheat (with A, B and D genome) will be discussed in this chapter. Here, we compare the findings as reported in this thesis with data from literature.

Strategies investigated

In this thesis, three strategies were investigated to reduce the presence of CD-toxic epitopes:

1. Identification of gluten gene loci in A, B and D wheat genomes with regard to CD-toxic epitope content;
2. Use of deletion lines lacking specific sets of gluten gene containing loci;
3. Application of the siRNA approach to down-regulate specific CD-toxic α-gliadin genes.

The effect of reduced CD-toxicity on industrial quality must be limited in order to maintain technological properties. This aspect was studied with regard to:

1. The relation between kernel development and technological properties;
2. Time and place of α-gliadin production during grain development;
3. Influence of elimination of gluten gene loci on industrial quality.

We will discuss these strategies according to their effectiveness, limitations and feasibility in time.

Reducing toxicity

To develop a strategy to reduce the number of CD-toxic α-gliadin genes, the distribution of CD-epitopes in the genes of the α-gliadin family were studied (Chapter 2). The main findings in this chapter are the large number of homologous α-gliadin sequences in this family with CD-toxic epitopes. Therefore, reduction of this number through conventional breeding methods will be a very complicated task. Also in this gene family, a large amount of pseudogenes were found: We estimated that 87% of these genes were pseudogenes, which is more than found in literature [Anderson and Greene, 1997]. In addition, it appeared that the epitope sequences, including those of the pseudogenes, showed a genome-specific distribution. The D-genome was found to contain the largest number of α-gliadin genes with CD toxic epitopes. Furthermore, in the majority of CD patients the α-gliadin 33-mer fragment p31-43 (LGQQQPFPPQQPY) is able to provoke an innate immune response [Shan et al 2004]. This fragment was found in one cloned *T. monococcum* (A genome) sequence (out of 17 sequences) and in 22 database sequences (out of 23 database sequences assigned to chromosome 6A based on sequence homology), by re-analysing the results from Chapter 2. P31-43 was not found in B- or D-genome sequences. In these genomes, different forms of p31-43 were present. In the D genome, **P**GQQQPFPPQQPY, and in the B genome, **P**GQQQ**Q**FPPQQPY were the most abundantly observed p31-43 variants. It is not yet known if these sequences can also provoke an innate response. In contrast to the T-cell recognition sequences, not much is known on the importance of specific individual amino acids in inducing IL-15 secreted by dendritic cells, which is crucial in the innate response.

Based on these results, we decided to focus on the α-gliadins encoded on chromosome 6D. The choice for this strategy depends partly on available data in public databases. We assumed

these to be representative for the *in-planta* situation, but this need not be correct. To give an example: ω-gliadins appear to be difficult to clone. As a result, only five ω-gliadin sequences are present in the database. In the case of α-gliadins, however, sequences were obtained from various sources, i.e. cloned from genomic DNA, from EST sequences (Salentijn *et al.,* in prep), and from cDNA (ongoing experiments). These sources produce very similar results with regard to the differences between sequences from the homologous loci, and with regard to the presence of intact and mutated epitopes.

One approach to reduce the α-gliadins expressed by genes on chromosome 6D was to focus on influencing gene expression directly at the promoter level. Another approach was to use specific deletion lines to study the effects of complete absence of the different gluten gene loci on the epitope content. A third approach was to apply RNA interference aiming at down-regulation of only those gliadin genes with CD-toxic epitopes. These approaches will be elaborate below.

Influencing gene expression directly at the promoter level

Down-regulation of the expression of D-genome α-gliadins while increasing the expression of the B-genome α-gliadins may retain the technological properties. A and D genome sequences were shown to be expressed differently during kernel development [Kawaura *et al.,* 2005]. In Chapter 3 we observed differences among promoter sequences from different genomes. These differences were also located at sites important for transcription factor binding (for example the sequence suggested for bZIP binding). Specific transcription factors might therefore be involved in the regulation of an expression balance between the different genomes. If this is true, the expression balance could possibly be artificially changed directly at the promoter sequence. In this way, down-regulation of the D genome while increasing B genome α-gliadins might enable reducing CD-toxicity while retaining technological properties. However, in this thesis only differences in sequence conservation between A and B genome promoter sequences have been analysed. No D-genome α-gliadin promoter sequence was found in the databases. The distinct sequence conservations of the α-gliadin coding sequences originating from the A, B and D genomes are probably the result of differently evolved ancestral genomes (Chapter 2). We therefore hypothesise that differences may also exist with D genome promoter sequences. Also the

expression of D genome α-gliadin genes is different from A genome genes during kernel development [Kawaura *et al.,* 2005] which indicates differences in promoter activity and thus in its sequence. A serious problem in attempting to influence the expression of the α-gliadin promoter using transcription factors is that these factors are expected to also influence other protein families. This can lead to unwanted effects. The LMW-GS promoters, for example, also contain CGN motifs as do gliadin gene promoters. Possibly, other storage protein genes may have this motif as well [Gobaa *et al.,* 2007]. In this regard it was observed that over-expressing a Dof-transcription factor in Arabidopsis, which is a homologue of endosperm-specific transcription factors in cereals, resulted in developmental changes of the plant [Kang and Singh, 2000]. This may indicate some risk of unwanted effects when applying this approach.

The effects of deletions of various gluten gene loci on the epitope content

In Chapter 5 we studied the epitope content immunologically (using epitope-specific antibodies) in the different classes of gluten proteins from the A, B and D genome. In summary, deletions in the hexaploid wheat cultivar Chinese Spring selected for removal of HMW-GS on the long arm of chromosome 1B and 1D led to a moderate loss of CD-toxic epitopes. Deletions targeted at the removal of LMW-GS and γ-gliadins result in a reasonable loss of CD toxic epitopes, of which deletions of 1DS-1 were most effective. Deletions on the short arms of chromosome 6A, 6B and 6D were targeted at the removal of α-gliadins. This led to a considerable loss of epitopes with the deletions of the loci on 6D being most effective. One unknown factor when applying such an approach more generally is that other hexaploid wheat varieties may have different CD-toxic gluten genes than Chinese Spring, although the results do correspond well with the distribution of epitopes in the diploid accessions sequenced in Chapter 2. Designing crosses between specific deletion lines might further open ways to reduce CD-toxicity. The deletions of part of the chromosomes result in the loss of specific gluten gene loci but this approach is quite crude, as genes located nearby are lost as well. A more delicate approach could be accomplished using RNAi.

RNA interference aiming at down-regulation of gliadin genes with CD-toxic epitopes only

Using siRNA, we attempted to specifically down-regulate a set of α-gliadin genes containing an α-gliadin epitope while retaining the ones that do not contain this epitope. This could limit negative effects on technological properties while obtaining a decrease in toxic epitopes comparable to a complete deletion of part of the chromosome. Four independent transgenic plants were obtained. The results on gene expression varied between the transgenic lines. One line showed a significant down-regulation in the transcription of the targeted gene, while another line showed an increase of the targeted α-gliadin gene. Seeds of the plant line with increased α-gliadin gene expression showed on a 2D gel down-regulated γ-gliadin gene expression. This suggests off-targeting by the siRNA construct (see below). Further analysis of the plant line showing a decrease in α-gliadin gene transcription, using detailed 2D analysis and protein sequencing of down-regulated spots, is on-going and will hopefully demonstrate whether our approach was successful in a least one of the four transgenic lines.

siRNAs can affect mRNAs that are not perfectly complementary, generally considered off-targets [Jackson *et al.*, 2003]. Natural plant siRNAs have targets with up to five mismatches [Schwab *et al.*, 2005]. In our case, the unexpected silencing of γ-gliadins could be due to off-targeting. The α-gliadin gene family contains several hundreds of very homologous gene copies (Chapter 2). Silencing of more than one gene was performed in rice by Miki *et al.* [2005]. These authors showed that with one siRNA construct multiple genes could be down regulated. However, the authors also indicate that genes with higher levels of mRNA expression might be better silenced than those with lower levels of expression. Our wheat transgenic lines showed different levels of expression of the transgene (Chapter 3). Because we saw many differences between the transgenic lines, the level of expression of the siRNA, perhaps due to copy number differences, may have had severe effects on the downregulation of various other gliadin genes.

Applying an RNAi molecule with a much larger complementing sequence, as used by Becher *et al.* [2006], is not expected to result in a higher specificity. Becher *et al.* [2006] showed down-regulation of most α-gliadin genes by transgenic expressing of a long RNAi molecule in a bread wheat cultivar. Long double-stranded precursors generate a multitude of siRNAs with varying ends, which make the prediction of off-targeting particularly difficult. As described by

Schwab *et al.* [2005; 2006], artificial microRNAs (amiRNAs) may be a better option to specifically down regulate genes in plants. Short hairpin RNAs, as used in this study, are intended to target perfectly complementary mRNAs, whereas amiRNAs preferentially avoid perfectly complementary targets in order to minimise problems caused by transitivity. amiRNA precursors generally generate only a single effective small RNA of known sequence. Potential off-targets of amiRNAs can be more accurately predicted than short hairpin RNAs [Schwab *et al.*, 2006]. The amiRNA approach achieved silencing of individual isoforms [Ossowski *et al.*, 2008]. It was shown that one SNP at the cleavage site strongly influenced silencing of a transcript, and could therefore be used to study the individual activity on two alleles [Schwab *et al.*, 2006]. However, an approach to down-regulate genes that can distinguish between a single SNP has not yet been accomplished in plant [Ossowski *et al.*, 2008].

Effects on technological properties

Relations between wheat kernel development and technological characteristics are well documented. A considerable body of literature exist on the effect of growing conditions (including soil composition, fertiliser, and light) on the formation of wheat prolamins. Less is known of the mechanism by which this impacts technological properties. Don *et al.* [2005a] studied the effect of heat stress and proposed a mechanism for this phenomenon related to the process of particle formation. Don *et al.* [2003a] suggested that the glutenin particles in GMP originate from the PBs observed in developing wheat endosperm. The results described in Chapter 4 further confirm that protein bodies are the origin for the larger glutenin particles observed in mature wheat flour. Based on these results we have adopted the model for glutenin particle formation (Figure 1). The cell wall of the endosperm cells acts as a physical barrier for the growing glutenin particle. The influence of non-gluten components on the formation of these glutenin particles like starch, arabinoxylans and triticin, can also be explained by this model. These non-gluten components are also suggested to be important for technological properties as explained in the hyperaggregation model of Hamer and Van Vliet [2000]. The inclusion bodies filled with triticin could inhibit this process by preventing the loss of membrane integrity. Starch granules are expected to negatively influence the size of the glutenin particles, by competing with glutenin particles for cell space.

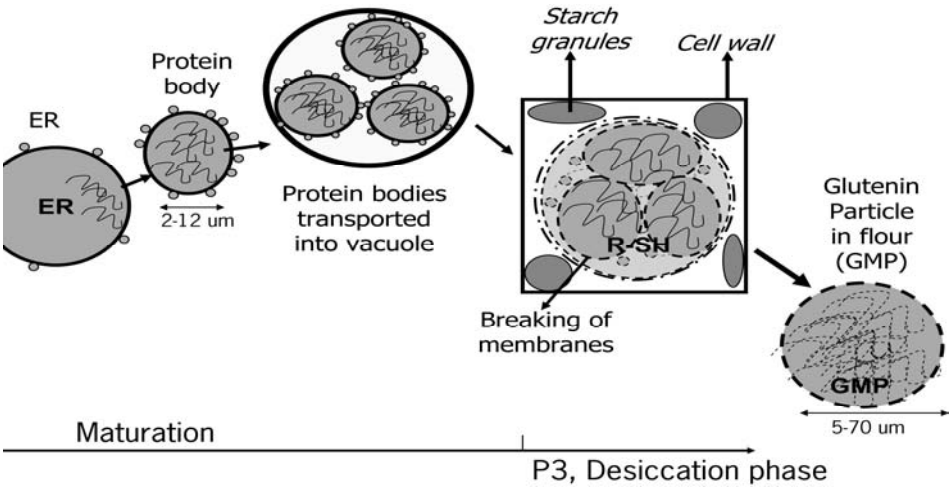

Figure 1 - Model of glutenin particle formation during different steps of kernel development. Until the desiccation phase the endoplasmatic reticulums (ERs) in the cells are filled with protein to form distinct protein bodies (PBs). These protein bodies are subsequently transported to the vacuole. During the desiccation phase membrane integrity is lost and protein bodies can fuse to form the glutenin particles.

The results presented in Chapter 4 accompanied with results from literature support this model. In this chapter, two time points were used: approximately 15 DAF and maturity. We did not observe what actually happened during the desiccation process. Additional experiments are required to further support this model of glutenin particle formation during kernel ripening. For this thesis, this model is relevant because of possible down-regulation of the α-gliadin genes which could affect protein body formation and therefore influence technological properties of the wheat flour.

When we want to predict the effects on technological properties, more knowledge is required on the expression and deposition of the α-gliadins. This is especially relevant when assuming the proposed model of glutenin particle formation. In Chapter 3 we observed that the α-gliadin protein was found in the protein bodies of the starchy endosperm cells and in the sub-aleurone cells. The α-gliadin promoter was active from 11 DAF until maturity. Since α-gliadins are deposited in the protein bodies during wheat during kernel development, a lower expression of

α-gliadins could disturb protein body formation and affect technological properties as proposed before.

A limitation in the design of this study is that only one fragment (of 592 bp) of the complete α-gliadin promoter sequence was used. We found a discrepancy between the observed cell specificity of promoter fragment activity and the actual deposition of the α-gliadin protein. This promoter fragment was active in the aleurone cells, while the α-gliadin protein was not found in these cells. This indicates a non-specific expression of the promoter fragment compared to the natural situation.

In Chapter 5 we tested the feasibility of a strategy to reduce CD-toxicity while affecting technological properties to a minimum. We analysed the contribution of the different classes of gluten proteins to the technological properties. Dough mixing should not be too short (risk of over-mixing), and dough needs to be sufficiently visco-elastic to be handled, and to be able to retain gas. In summary, deletions targeted at removing HMW-GS on the long arm of chromosome 1B and 1D led to a severe loss of technological functionality. Deletions targeted at the removal of LMW-GS and γ-gliadins led to moderate changes in technological properties. The 1BS-10 deletion is an exception to this, presumably due to a loss of an essential LMW-GS, since line 1BS-19/6DS-4 (a shorter deletion on the same chromosome) did not show significant losses in technological functionality. Deletions on the short arm of chromosome 6A, 6B and 6D were targeted at the removal of α-gliadins. In general this leads to stronger and stiffer dough. This is likely to be caused by a change in the glutenin/gliadin ratio, and hence can be corrected for. With respect to particle size, line 6AS-1 showed an increase in average glutenin particle size, while deletions of the loci on 6B and 6D showed a decrease in average glutenin particle size.

One limitation in this study is that other hexaploid wheat varieties may have different gluten loci compared to Chinese Spring and hence removal of loci may show somewhat different technological properties. It is also important to realise that real baking experiments should still be performed to confirm our assessment of technological quality.

Did we reach our objective?

How far do these strategies bring us towards a reduction of CD toxicity while at the same time retaining technological properties? Low-CD-toxic gluten can be beneficial in two ways: 1) CD-patients could benefit the increased tolerance from such a variety; 2) Such a variety may have an effect in prevention of CD in the general public: low-CD-toxic wheat reduces the exposure to epitopes, which may postpone the onset of CD.

Were we right to remove α-gliadins?

α-Gliadins, mainly from the 6D genome, seem the most toxic. Therefore lowering the α-gliadins from the 6D genome of bread wheat seems to be a logical option to obtain a less toxic bread wheat variant. Will this lead to 'safe' varieties or merely to varieties with reduced toxicity? In the following we will discuss the value of such a variety for CD patients on one hand and for prevention in the general public on the other hand will be discussed.

CD patients should use a gluten free diet. The Codex Alimentarius indicates for gluten free products a threshold of 0.02% gluten (20 mg/kg). A wheat variety lacking α-gliadins will therefore not be labelled 'gluten free'. Most T-cell clones were found to react to α-gliadin sequences, but also T-cells clones reacting to γ-gliadins and HMW and LMW glutenins were found in CD patients [Sollid, 2002; Arentz-Hansen *et al.*, 2000; Vader *et al.*, 2002b; Koning, 2003]. Based on these results, varieties lacking the chromosome 6D encoded α-gliadins will not be regarded 'safe' for CD patients.

α-Gliadins are the major contributor in both activating the adaptive immune response and the innate immune response. A variety low in CD-toxic α-gliadins could be beneficial for patients who do not yet know they have CD. These patients could develop fewer complaints. Several researchers observed a relation between population incidence of CD and gluten exposure [Ivarsson *et al.*, 2000; Fasano, 2006; Ventura *et al.*, 1999]. A variety low in chromosome 6D α-gliadins could therefore help to prevent or delay the development of the disease. Diploid and tetraploid wheat (i.e. *T. monococcum* and *T. durum*) do not contain the D genome. These wheats are hence already devoid of the α-gliadins considered here. In literature several reports have

appeared that wheats lacking the D genome are indeed less toxic. Auricchio *et al.* [1982] reported that peptic/tryptic digests of durum proteins exhibited much less adverse effects in a human pathologic tissue culture system. Molberg *et al.* [2005] reported that several wheats lacking the D-genome also lack the immunodominant 33-mer. Molberg *et al.* [2005] did not extend his conclusion to all durum wheat. This is probably due to the fact that, as we have observed, durum wheat samples are often contaminated with *T. aestivum* wheat. These reports lend further support to our strategy to eliminate 6D-encoded α-gliadins.

For prevention purposes in the general public such a variety must be cultivated on a large scale to be consumed by most people. If this can be achieved even CD patients could benefit. It is reported that keeping to a strict gluten free diet is difficult. Mariani *et al.* [1998] have reported that the compliance rate to a strict gluten free diet by CD patients was only approximately 50%. We expect that wheat low in 6D α-gliadins are a better option for patients when breaking their diet.

Two factors are important for these varieties to be cultivated on a large scale:

1) Yields and agronomical properties have to be sufficient to be economically interesting for farmers;
2) Technological properties have to be sufficient to be interesting for industrial purposes.

Decreased technological properties as a result of the removal of 6D α-gliadins can in principle be corrected for. Therefore it is expected that varieties low in 6D α-gliadin will still be interesting for industrial purposes. Agronomical properties of such a variety will still have to be determined. They will have to meet modern criteria for cultivation and for usability in order to be selected for cultivation on a large scale.

Best approaches

Based on new insights in the field of RNAi research the amiRNA technique is expected to be more specific than the siRNA technique. The advantage of this approach is that this technique can be applied to a cultivar that already possesses good technological properties and good yields. It will take several years from the design of the amiRNA construct to an actual low CD-toxic

transgenic bread wheat variety. To obtain genetically stable (homozygous) transgenic wheat line it already takes several years, as each generation of wheat will take at least 6 months and several generations are needed to obtain a homozygous line. After this a sufficient amount of kernels have to be obtained for technological and immunological tests and human intervention studies, which will again take several years.

RNAi is expected to down-regulate genes targeted for. Toxic proteins can be lowered using this technique but will not be completely (100%) absent. Such a variety will be considered GM wheat and this can cause problems in consumer acceptance. However, Schenk *et al.* [2008] found that beneficial "medical-GM" has an increased consumer acceptance over non medical-GM. We suggest that the value of amiRNA in transgenic wheat to specifically down-regulate CD-toxic α-gliadins should be investigated.

We are capable of eliminating the 33-mer and some other epitopes by deleting the α-gliadins from the D-genome locus. Practically, this can be obtained using deletion lines (Chapter 5). Line 6DS-2, in which this was achieved, did not show a decrease of dough rheological properties that cannot be corrected for. Specific crosses of deletion lines (for example 1DS-1 with 6DS-2) can be made to further decrease CD-toxic-epitopes. The effects on technological properties should be tested again for such a line followed by real baking experiments. The effects are expected to be minimal, because the two individual deletion lines did not show strong effects on technological parameters. Finally human intervention studies have to show that these lines can be tolerated well by CD-patients. Again agronomical properties are important to consider. Yield assessments of such a deletion line should also be performed.

Another approach is to include markers which detect the toxic α-gliadins from chromosome 6D in breeding programs and select for commercial varieties with lower 6D α-gliadins. This can be performed if there is sufficient genetic variation among in the α-gliadin loci on chromosome 6D in commercial varieties. If there is insufficient variation among commercial varieties, diploid accessions, tetraploid varieties or synthetic wheat lines can be used to obtain a new, low toxic variety. This latter approach would take at least a 5 to 10 year period, and would not immediately lead to wheat varieties with a high level of production (yield) and with good technological properties.

Towards CD safe wheat

In this thesis, a large dataset on epitope content and distribution of CD-toxic epitopes in α-gliadins has been obtained. From this it became clear that these proteins contain a number of different epitopes against which T-cells have been found in CD patients. It was established clearly that the epitope content of α-gliadins from the three homologous loci was completely different. The deletion of the locus on chromosome 6D leads to by far the strongest reduction of epitope content. As industrial quality was not affected very much, it may be possible to also delete the locus on chromosome 6A, after which nearly all toxic epitopes from α-gliadins would be removed, including the innate epitope, which is only present in chromosome 6A-encoded α-gliadins. Whether such a line is viable, whether expression of the 6B-encoded α-gliadins is increased, and whether industrial quality is not affected too much, remains to be determined. An alternative approach could be to combine a deletion of the locus on chromosome 6D with selection of cultivars with low levels of epitopes in the α-gliadins from chromosome 6A. Tests for differences in toxicity among cultivars and germplasm at hexaploid and tetraploid level are currently being carried out in a parallel PhD project [Hetty van den Broeck].

Furthermore, we have obtained evidence that suggests that the approach of deleting the most toxic locus may also be useful for the γ-gliadins, which are encoded by homologous loci on chromosome 1A, 1B and 1D. The antibody assays on γ-gliadins from deletion lines suggest the existence of differences in epitope content between the proteins from these three loci. Deletion line 1BS-19/6DS-4 lacking one α-gliadin and one γ-gliadin locus still had a reasonable technological properties, even though the glutenin particles in this line were somewhat decreased. A minor reduction in technological properties may be compensated in the baking process by addition of safe proteins from other sources, e.g. from oat or rice. Epitope content of γ-gliadins from the different chromosomes is currently being studied in detail.

The approach of deleting the most toxic loci can most likely not be applied on the glutenins, as industrial quality is very much dependent on the three-dimensional particle structure generated by these proteins. Indeed, LMW and HMW deletion lines showed by far the greatest decrease in several parameters of industrial quality, beyond the point of repair. In addition, the antibody assays of deletion lines do not suggest that the epitope content of LMW glutenins

encoded by the homologous loci is grossly different. As LMW genes are currently being studied intensively by various research groups, it is expected that there will be sufficient sequence information available soon to verify this.

The clinical relevance of α-gliadin epitopes is clear, as all patients have T-cells that recognise at least some of the epitopes, so that these T-cells are probably abundant. The relevance of glutenin epitopes is much less clear. For that reason we believe that a wheat cultivar with strongly reduced content of gliadin epitopes is already a significant achievement. Whether existing CD patients will benefit from it remains uncertain, as they will have T-cells that recognise the remaining gliadin epitopes, as well as some glutenin epitopes. Such a cultivar may, however, be a significant improvement with regard to preventing the onset of the disease. The significant reduction in epitope content would lead to a much reduced exposure to gliadin epitopes, at least delaying the developing intolerance to gluten. If so, we should start to define, to measure and to express the presence of wheat or wheat derivatives in food not as gluten content but rather as CD-epitope content. Clinical studies have to provide more clarity about the differences in toxicity.

References

- Aharoni A, Vos CHR, Wein M, Sun Z, Greco R, Kroon A, Mol JNM and O'Connell AP (2001) The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. Plant J, 28:319-332.
- Albani D, Hammond-Kosack MC, Smith C, Conlan S, Colot V, Holdsworth M and Bevan MW (1997) The wheat transcriptional activator SPA: a seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. Plant Cell, 9:171-84.
- Altpeter F, Vasil V, Srivastava V and Vasil IK (1996) Integration and expression of the high-molecular-weight glutenin subunit 1Ax1 gene into wheat. Nat Biotechnol, 14:1155-1159.
- Anderson OD and Greene FC (1997) The α-gliadin gene family. II DNA and protein sequence variation, subfamily structure and origins of pseudogenes. Theor Appl Genet, 95:59-65.
- Anderson OD, Litts JC and Greene FC (1997) The α-gliadin gene family. I. Characterization of ten new wheat α-gliadin genomic clones, evidence for limited sequence conservation of flanking DNA, and southern analysis of the gene family. Theor Appl Genet, 95:50-58.
- Arentz-Hansen EH, Körner R, Molberg Ø, Quarsten H, Vader W, Kooy Y, Lundin KEA, Koning F, Roepstorff P, Sollid LM and McAdam S (2000) The intestinal T-cell response to α-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. J Exp Med, 191:603-612.
- Arentz-Hansen EH, McAdam S, Molberg Ø, Fleckenstein B, Lundin KEA, Jørgensen TJD, Jung G, Roepstorff P and Sollid LM (2002) Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. Gastroenterology, 123:803-809.
- Aryan AP, An G and Okita TW (1991) Structural and functional analysis of promoter from gliadin, an endosperm-specific storage protein gene of *Triticum aestivum* L. Mol Gen Genet, 225:65-71.

- Auricchio S, De Ritis G, De Vincenzi M, Occorsio P and Silano V (1982) Effects of gliadin-derived peptides from bread and durum wheats on small intestine cultures from rat fetus and coeliac children. Pediatr Res, 16:1004-1010.
- Aussenac T, Carceller JL and Kleiber D (2001) Changes in SDS solubility of glutenin polymers during dough mixing and resting. Cereal Chem, 78:39-45.
- Barro F, Rooke L, Bekes F, Gras P, Tatham AS, Fido R, Lazzeri PA, Shewry PR and Barcelo P (1997) Transformation of wheat with high molecular weight subunit genes results in improved functional properties. Nat Biotechnol, 15:1295-1299.
- Baumlein H, Nagyt I, Villarroel R, Inz D and Wobus U (1992) Cis-analysis of a seed protein gene promoter: the conservative RY repeat CATGCATG within the legumin box is essential for tissue-specific expression of a legumin gene. Plant J, 2:233-239.
- Bechtel DB, Wilson JD and Shewry PR (1991) Immunocytochemical Localization of the Wheat Storage Protein Triticin in Developing Endosperm Tissue. Cereal Chem, 68:573-577.
- Becker D, Folck A and Wieser H (2006) Inhibierung der α-Gliadin-Genexpression in hexaploidem Brotweizen. Getreidetechnologie, 60:153-156.
- Blechl AE and Anderson OD (1996) Expression of a novel high-molecular-weight glutenin subunit gene in transgenic wheat. Nat Biotechnol, 14:875-879.
- Branlard G, Dardevet M, Saccomano R, Lagoutte F and Gourdon J (2001) Genetic diversity of wheat storage proteins and bread wheat quality. Euphytica, 119:59-67.
- Brummelkamp TR, Bernards R and Agami R (2002) A system for stable expression of short interfering RNAs in mammalian cells. Science, 19:550-553.
- Carceller JL and Aussenac T (1999) Accumulation and changes in molecular size distribution of polymeric proteins in developing grains of hexaploid wheats: role of the desiccation phase. Aust J Plant Physiol, 26:301-310.
- Carceller JL and Aussenac T (2001) Size characterisation of glutenin polymers by HPSEC-MALLS. J Cereal Sci, 33:131-142.

- Carceller, JL and Aussenac T (1999) Accumulation and changes in molecular size distribution of polymeric proteins in developing grains of hexaploid wheats: role of the desiccation phase. Aust J Plant Physiol, 26:301–310.
- Christensen, AH and Quail PH (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res, 5:213-218.
- Conley CA and Hanson MR (1997) Cryostat tissue printing: an improved method for histochemical and immunocytochemical localization in soft tissues. Biotechniques, 22:491-496.
- Cordewener JHG, Hause G, Görgen E, Busink R, Hause B, Dons HJM Van Lammeren AAM, Van Lookeren Campagne MM and Pechan P (1995) Changes in synthesis and localization of members of the 70-kDa class of heat-shock proteins accompany the induction of embryogenesis in *Brassica napus* L. microspores. Planta, 196:747-755.
- Curtis MD and Grossniklaus U (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol, 133:462-469.
- Dachkevitch T and Autran JC (1989) Prediction of baking quality of bread wheats in breeding programs by size-exclusion high-performance liquid chromatography. Cereal Chem, 66:448-456.
- Davy A, Blom MS, Svendsen I, Cameron-Mills V and Simpson DJ (2000) Prediction of Protein Cleavage Sites by the Barley Cysteine Endoproteases EP-A and EP-B Based on the Kinetics of Synthetic Peptide Hydrolysis. Plant Physiol, 122:137-146.
- Diaz I, Martinez M, Isabel-LaMoneda I, Rubio-Somoza I and Carbonero P (2005) The DOF protein, SAD, interacts with GAMYB in plant nuclei and activates transcription of endosperm-specific genes during barley seed development. Plant J, 42:652-62.
- Dicke WK (1950) Coeliakie. PhD thesis. MD Thesis, Utrecht.
- Don C, Lichtendonk WJ, Plijter JJ and Hamer RJ (2003a) Glutenin Macropolymer: a gel formed by particles. J Cereal Sci, 37:1-7.

- Don C, Lichtendonk WJ, Plijter JJ and Hamer RJ (2003b) Understanding the link between GMP and dough: from glutenin particles in flour towards developed dough. J Cereal Sci, 38:157-165.
- Don C, Mann G, Bekes F and Hamer RJ (2003c) Linking glutenin particles to HMW subunit composition. In: Lafiandra D, Masci S and D'Ovidio R, Editors, The Gluten Proteins-Proceedings of the 8th International Workshop Gluten 2003, Royal Society of Chemistry, Cambridge, 188-192.
- Don C, Lookhart G, Naeem H, MacRitchie F and Hamer RJ (2005a) Heat stress and genotype affect the glutenin particles of the glutenin macropolymer-gel fraction. J Cereal Sci, 42:69-80.
- Don C, Lichtendonk WJ, Plijter JJ, Van Vliet T and Hamer RJ (2005b) The effect of mixing on glutenin particle properties: aggregation factors that affect gluten function in dough. J Cereal Sci, 41:69-83.
- Don C, Mann G, Bekes F and Hamer RJ (2006) HMW-GS affect the properties of glutenin particles in GMP and thus flour quality. J Cereal Sci, 44:127-136.
- Doyle JJ and Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus, 12:13-15.
- Drea S, Leader DJ, Arnold BC, Shaw P, Dolan L and Doonan JH (2005) Systematic spatial analysis of gene expression during wheat caryopsis development. Plant Cell, 17:2172-2185.
- Eagles HA, Eastwood RF, Hollamby GJ, Martin EM and Cornish GB (2004) Revision of the estimates of glutenin gene effects at the Glu-B1 locus from southern Australian wheat breeding programs. Aust J Agric Res, 55:1093-1096.
- Ebert EC (1998) Interleukin 15 is a potent stimulant of intraepithelial lymphocytes. Gastroenterology, 115:1439-1445.
- Elbashir S, Harborth J, Lendeckel W, Yalcin A, Weber K and Tuschl T (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature, 411:494-498.
- Ensari A, Marsh MN, Moriarty KJ, Moore CM, Fido RJ, Tatham AS (1998) Studies in vivo of omega-gliadins in gluten sensitivity (coeliac sprue disease). Clin Sci (Lond), 95:419-24.

- Evers AD (1970) Development of the endosperm of wheat (Triticum aestivum L.). Ann Bot (Lond), 34:547-555.
- Fasano A (2006) Systemic autoimmune disorders in celiac disease. Curr Opin Gastroenterol, 22:674-679.
- Fasano A and Catassi C (2001) Current approaches to diagnosis and treatment of celiac disease: an evolving spectrum, Gastroenterology, 120:636-651.
- Feldman M, Lupton FGH and Miller TE (1995) Evolution of crop plant. Edited by Smartt J, Simmonds NW. Harlow Essex: Longman Scientific & technical:184-192.
- Felix I, Martinant JP, Bernard M, Bernard S and Branlard G (1996) Genetic characterization of storage proteins in a set of F1-derived haploid lines in bread wheat. Theor Appl Genet, 92:340-346.
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution, 39:783-791.
- Fincher GB (1989) Molecular and cellular biology associated with endosperm mobilization in germinating cereal grains. Annu Rev Plant Physiol Plant Mol Biol, 40:305-346.
- Fujiwara T and Beachy RN (1994) Tissue-specific and temporal regulation of a beta-conglycinin gene: roles of the RY repeat and other cis-acting elements. Plant Mol Biol, 24:261-272.
- Galili G (1997) The prolamin storage proteins of wheat and its relatives. In: Larkins BA, Vasil IK, eds. Cellular and molecular biology of plant seed development. The Netherlands: Kluwer Academic Publishers 221-256.
- Gobaa S, Kleijer G and Stamp P (2007) A new high molecular weight glutenin subunit coded by Glu-A1: its predicted structure and its impact on bread-making quality. Plant Breeding, 126:1-4.
- Graveland A (1984) In: Graveland A and Moonen JHE, Editors, Proceedings of the 2nd International Gluten Workshop on Gluten Proteins, Pudoc, Wageningen 59-67.

- Graveland A, Bosveld AP, Lichtendonk WJ, Marseille JP, Moonen JHE and Scheepstra JA (1985) A model for the molecular structure of the glutenins from wheat flour. J Cereal Sci, 3:1-16.
- Graveland A, Bosveld P, Lichtendonk WJ and Moonen JHE (1982) Extraction and fractionation of wheat flour proteins. J Science Food Agri, 33:1117–1128.
- Greffeuille V, Abecassis J, Bar L'Helgouac'h C and Lullien-Pellerin V (2005) Differences in the aleurone layer fate between hard and soft common wheats at grain milling. Cereal Chem, 82:138-143.
- Grosch W and Wieser H (1999) Redox reactions in wheat dough as affected by ascorbic acid. J Cereal Sci, 29:1-16.
- Gu YQ, Crossman C, Kong X, Luo M, You FM, Coleman-Derr D, Dubcovsky J and Anderson OD (2004) Genomic organization of the complex alpha-gliadin gene locus in wheat. Theor Appl Genet, 109:648-657.
- Gupta RB, Khan K and Macritchie F (1993) Biochemical basis of flour properties in bread wheats. I. Effects of variation in the quantity and size distribution of polymeric protein. J Cereal Sci, 18:23-41.
- Hamer RJ and Van Vliet T (2000) Understanding the structure and properties of gluten: an overview. In: Shewry PR and Tatham AS, Editors, 2000. Wheat Gluten-Proceedings of the Seventh International Gluten Workshop, Royal Society of Chemistry, Cambridge, 125-131.
- Hamilton A and Baulcombe D (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. Science, 286:950-952.
- Hannon GJ and Rossi JJ (2004) Unlocking the potential of the human genome with RNA interference. Nature, 431:371-378.
- Harberd NP, Bartels D and Thompson RD (1985) Analysis of the gliadin multigene locus in bread wheat using nullisomic-tetrasomic lines. Mol Gen Genet, 198:234-242.
- Harberd NP, Bartels D and Thompson RD (1986) DNA restriction fragment variation in gene family encoding high molecular weight (HMW) glutenin subunits of wheat. Biochem Genet, 24:579-596.

- Hopper AD, Hurlstone DP, Leeds JS, McAlindon ME, Dube AK, Stephenson TJ and Sanders DS (2006). The occurrence of terminal ileal histological abnormalities in patients with coeliac disease. Dig Liver Dis, 38:815-819.
- Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R and Gornicki P (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploidy wheat. Proc Natl Acad Sci USA, 99:8133-8138.
- Hue S, Mention JJ, Monteiro RC, Zhang S, Cellier C, Schmitz J, Verkarre V, Fodil N, Bahram S, Cerf-Bensussan N and Caillat-Zucman S (2004) A direct role for NKG2D/ MICA interaction in villous atrophy during celiac disease. Immunity, 21:367-377.
- Isidore E, Scherrer B, Chalhoub B, Feuillet C and Keller B (2005) Ancient haplotypes resulting from extensive molecular rearrangements in the wheat A genome have been maintained in species of three different ploidy levels. Genome Res, 15:526-536.
- Islam N, Tsujimoto H and Hirano H (2003) Wheat proteomics: relationship between fine chromosome deletion and protein expression. Proteomics, 3:307-316.
- Ivarsson A, Persson LA, Nystrom L, Ascher H, Cavell B, Danielsson L, Dannæus A, Lindberg T, Lindquist B, Stenhammar L and Hernell O (2000) Epidemic of coeliac disease in Swedish children. Acta Paediatr, 89:165-171.
- Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G and Linsley PS (2003) Expression profiling reveals off-target gene regulation by RNAi. Nat Biotechnol, 21:635–637.
- Janatuinen EK, Kemppainen TA, Julkunen RJ, Kosma VM, Maki M, Heikkinen M and Uusitupa MI (2002) No harm from five-year ingestion of oats in coeliac disease. Gut, 50:332-335.
- Kang HG and Singh KB (2000) Characterization of salicylic acid-responsive, Arabidopsis Dof domain proteins: overexpression of OBP3 leads to growth defects. Plant J, 21:329-339.
- Kasarda DD (1999) Glutenin polymers: the in vitro to in vivo transition. Cereal Foods World, 44:566-571.

- Kawaura K, Mochida K and Ogihara Y (2005) Expression profile of two storage-protein gene families in hexaploid wheat revealed by large-scale analysis of expressed sequence tags. Plant Physiol, 139:1870-1880.
- Kent NL and Jones CR (1952) The cellular structure of wheat flour. Cereal Chem, 29:383-398.
- Kilian, B, Özkan H, Deusch O, Effgen, S Brandolini A, Kohl J, Martin W and Salamini F (2007) Independent Wheat B and G Genome Origins in Outcrossing Aegilops Progenitor Haplotypes. Mol Biol Evol, 24:217-227.
- Kim WT, Franceschi VR, Krishman HB and Okita TW (1988) Formation of wheat protein bodies: involvement of the Golgi apparatus in gliadin transport. Planta, 176:173-182
- Kiyosaki T, Matsumoto I, Asakura T, Funaki J, Kuroda M, Misaka T, Arai S and Abe K (2007) Gliadain, a gibberellin-inducible cysteine proteinase occurring in germinating seeds of wheat, *Triticum aestivum* L., specifically digests gliadin and is regulated by intrinsic cystatins. FEBS J, 274:1908-1917.
- Köhler P, Belitz HD and Wieser H (1993) Disulphide bonds in wheat gluten: further cystine peptides from high molecular weight (HMW) and low molecular weight (LMW) subunits of glutenin and from γ-gliadins. Z Lebensm Unters Forsch, 196:239-247.
- Koning F (2003) The molecular basis of celiac disease. J Mol Recognit, 16:333-336.
- Koning F, Gilissen L and Wijmenga C (2005) Gluten: a two-edged sword. Immunopathogenesis of celiac disease. Springer Semin Immunopathol, 27:217-232.
- Koning F, Schuppan D, Cerf-Bensussan N and Sollid L (2005) Pathomechanisms in celiac disease. Best Pract Res Clin Gastroenterol, 19:373-387.
- Kurihara Y and Watanabe Y (2004) Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. PNAS, 24:12753-12758.
- Lafiandra D, D'Ovidio R, Porceddu E, Margiotta B and Colaprico G (1993) New data supporting high Mr glutenin subunit 5 as the determinant of quality differences among the pairs 5+10 vs. 2+12. J Cereal Sci, 18:197-205.

- Lamacchia C, Shewry PR, Di Fonzo N, Forsyth JL, Harris N, Lazzeri PA, Napier JA, Halford NG and Barcelo P (2001) Endosperm-specific activity of a storage protein gene promoter in transgenic wheat seed. J Exp Bot, 52:243-250.
- Lásztity R (1996) The chemistry of cereal proteins (second ed.), CRC Press. Boca Raton, FL 78-79.
- Lending CR and Larkins BA (1989) Changes in the zein composition of protein bodies during maize endosperm development. Plant Cell, 1:1011-1023.
- Lew EJL, Kuzmicky DD and Kasarda DD (1992) Characterization of low molecular weight glutenin subunits by reversed-phase high-performance liquid chromatography, sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and N-terminal amino acid sequencing. Cereal Chem, 69:508-515.
- Li YC, Korol AB, Fahima T, Beiles A and Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol, 11:2453-2465.
- Lichtendonk WJ, Kelfkens M, Orsel R, Bekkers ACAPA and Plijter JJ (2000) The impact of water-soluble pentosans on dough properties In: Shewry PR and Tatham AS, Editors, Wheat Gluten-Proceedings of the Seventh International Workshop Gluten, Royal Society of Chemistry, Cambridge, 512-518.
- MacRitchie F (1999) Wheat proteins: characterization and role in flour functionality. Cereal Foods World, 44:188-193.
- Maiuri L, Ciacci C, Ricciardelli I, Vacca L, Raia V, Auricchio S, Picard J, Osman M, Quaratino S and Londei M (2003) Association between innate response to gliadin and activation of pathogenic T cells in coeliac disease. Lancet, 362:30-37.
- Mann M and Wilm M (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal Chem, 66:4390-4399.
- Mariani P, Viti MG, Montouri M, La Vecchia A, Cipolletta E, Calvani L and Bonamico M (1998) The Gluten-Free Diet: A Nutritional Risk Factor for Adolescents with Celiac Disease? J Pediatr Gastroenterol Nutr, 27:519-523.

- Marino CL, Nelson JC, Lu YH, Sorrells ME, Leroy P, Tuleen NA, Lopes CR and Hart GE (1996) Molecular genetic maps of the group 6 chromosomes of hexaploid wheat (Triticum aestivum L. em. Thell). Genome, 39:359-366.
- Masci S, D'Ovidio R, Lafiandra D and Kasarda DD (1998) Characterization of a low-molecular-weight glutenin subunit gene from bread wheat and the corresponding protein that represents a major subunit of the glutenin polymer. Plant Physiol, 118:1147-1158.
- Mena M, Vicente-Carbajosa J, Schmidt RJ and Carbonero P (1998) An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. Plant J, 16:53-62.
- Meresse B, Chen Z, Ciszewski C, Tretiakova M, Bhagat G, Krausz TN, Raulet DH, Lanier LL, Groh V, Spies T, Ebert EC, Green PH and Jabri B (2004) Coordinated induction by IL15 of a TCR-independent NKG2D signalling pathway converts CTL into lymphokine-activated killer cells in celiac disease. Immunity, 21:357-366.
- Metakovsky EV, Chernakov VM, Pogna NE and Bushuk W (1993) Gliadin alleles in Canadian western red spring wheat cultivars: use of two different procedures of acid polyacrylamide gel electrophoresis for gliadin separation. Genome, 36:743-749.
- Miki D, Itoh R and Shimamoto K (2005) RNA silencing of single and multiple members in a gene family of rice. Plant Physiol, 138:1903-1913.
- Millar SJ (2003) The development of near infrared (NIR) spectroscopy calibrations for the prediction for wheat and flour quality. Project Report No 310, Home Grown Cereals Authority, London.
- Miller VM, Xia H, Marrs GL, Gouvion CM, Lee G, Davidson BL and Paulson HL (2003) Allele-specific silencing of dominant disease genes. Proc Natl Acad Sci, 100:7195-7200.
- Mitea C, Havenaar R, Drijfhout JW, Edens L, Dekking L and Koning F (2008) Efficient degradation of gluten by a prolyl endoprotease in a gastrointestinal model: implications for coeliac disease. Gut, 57:25-32.

- Molberg Ø, Solheim FN, Jensen T, Lundin KEA, Arentz- Hansen EH, Anderson OD, Kjersti UA and Sollid LM (2003) Intestinal T-cell responses to high-molecular-weight glutenins in celiac disease. Gastroenterology, 125:337-344.
- Molberg Ø, Solheim FN, Jensen T, Lundin KEA, Arentz- Hansen EH, Anderson OD, Kjersti UA and Sollid LM (2003) Intestinal T-cell responses to high-molecular-weight gluteninsin celiac disease. Gastroenterology, 125:337-344.
- Molberg O, Uhlen AK, Jensen T, Flaete NS, Fleckenstein B, Arentz-Hansen H, Raki M, Lundin KE and Sollid LM (2005) Mapping of gluten T cell epitopes in the bread wheat ancestors: implications for celiac disease. Gastroenterology, 128:393-401.
- Morel M-H, Dehlon P, Autran JC, Leygue JP and Bar-L'Helgouac'h C (2000) Effects of temperature, sonication time, power settings on size distribution and extractability of total wheat flour proteins as determined by size-exclusion high-performance liquid chromatography. Cereal Chem, 77:685-691.
- Mori N, Liu YG and Tsunewaki K (1995) Wheat phylogeny determined by RFLP analysis of nuclear DNA. 2. Wild tetraploid wheats. Theor Appl Genet, 90:129-134.
- Morris R and Sears ER (1967) The cytogenetics of wheat and its relatives. In: Wheat and wheat improvement, Quisenberry KS and Reitz LP eds. American Society of Agronomy, Madison WI:19-87.
- Nei M and Gojobori T (1986) Simple Methods for Estimating the Numbers of Synonymous and Nonsynonymous Nucleotide Substitutions. Mol Biol Evol, 3:418-426.
- Okita TW, Cheesbrough V and Reeves CD (1985) Evolution and heterogeneity of the alpha/beta type and gamma-type gliadin DNA sequences. J Biol Chem, 260:8203-8213.
- Oñate L, Vicente-Carbajosa J, Lara P, Díaz I and Carbonero P (1999) Barley BLZ2, a seed-specific bZIP protein that interacts with BLZ1 in vivo and activates transcription from the GCN4-like motif of B-hordein promoters in barley endosperm. J Biol Chem, 274:9175-9182.
- Onodera Y, Suzuki A, Wu CY, Washida H and Takaiwa F (2001) A rice functional transcriptional activator, RISBZ1, responsible for endosperm-specific expression of storage protein genes through GCN4 motif. J Biol Chem, 276:14139-14152.

- Ossowski S, Schwab R and Weigel D (2008) Gene silencing in plants using artificial microRNAs and other small RNAs. Plant J, 53:674-690.
- Palova-Jelinkova L, Rozkova D, Pecharova B, Bartova J, Sediva A, Tlaskalova-Hogenova H, Spisek R and Tuckova L (2005) Gliadin fragments induce phenotypic and functional maturation of human dendritic cells. J Immunol, 175:7038-7045.
- Parker ML (1980) Protein Body Inclusions in Developing Wheat Endosperm. Ann Bot (Lond), 46:29-36.
- Payne PI, Nightingale MA, Krattiger AF and Holt LM (1987a) The relationship between HMW glutenin subunit composition and the bread-making quality of British-grown wheat varieties. J Sci Food Agri, 40:51–65.
- Payne PI, Seekings JA, Worland AJ, Jarvis MG and Holt LM (1987b) Allelic variation of glutenin subunits and gliadins and its effect on breadmaking quality in wheat: Analysis of F5 progeny from Chinese Spring×Chinese Spring (Hope 1A). J Cereal Sci, 6:103-118.
- Popineau Y, Cornec M, Lefebvre J and Marchylo B (1994) Influence of high M[r] glutenin subunits on glutenin polymers and rheological properties of glutens and gluten subfractions of near-isogenic lines of wheat Sicco. J Cereal Sci, 19:231.
- Popineau Y, Deshayes G, Lefebvre J, Fido R, Tatham AS and Shewry PR (2001) Prolamin aggregation, gluten viscoelasticity and mixing properties of transgenic wheat lines expressing 1Ax and 1Dx high molecular weight glutenin subunit transgenes. J Agric Food Chem. 49:395-401.
- Rabilloud T, Carpentier G and Tarroux P (1988) Improvement and simplification of low-background silver staining of proteins by using sodium dithionite. Electrophoresis, 9:288-291.
- Reeves CD and Okita TW (1987) Analyses of alpha/beta-type gliadin genes from diploid and hexaploid wheats. Gene. 52:257-266.
- Rhazi L, Cazalis R, Lemelin E and Aussenac T (2003) Changes in the glutathione thiol-disulfide status during wheat grain development. Plant Physiol Biochem, 41:895-902.

- Rozas J, Sanchez-DelBarrio JC, Messeguer X and Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics, 19:2496-2497.
- Rubin R, Levanony H and Galili G (1992) Characterization of two types of protein bodies in developing wheat endosperm. Plant Physiol, 99:718–724.
- Sabelli PA and Shewry PR (1991) Characterization and organization of gene families at the *Gli-1* loci of bread and durum wheats by restriction fragment analysis Theor Appl Genet, 83:209-216.
- Schenk MF, Fischer ARH, Frewer LJ, Gilissen LJWJ, Jacobsen E and Smulders MJM (2008) Attitude towards GM application as for prevention of hey fever and food allergy by patients and non-patients. Health, Risk & Society (submitted).
- Schwab R, Ossowski S, Riester M, Warthmann N and Weigel D (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. Plant Cell, 18:1121-1133.
- Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M and Weigel D (2005) Specific effects of microRNAs on the plant transcriptome. Dev Cell, 8:517-527.
- Schwarz DS, Ding H, Kennington L, Moore JT, Schelter J, Burchard J, Linsley PS, Aronin N, Xu Z and Zamore PD (2006) Designing siRNA that distinguish between genes that differ by a single nucleotide. PLoS Genet, 8:e140.
- Sebecic B and Balenovic J (2001) Rapid ecologically acceptable method for wheat protein content determination-comparison of methods. Deutsche Lebensmittel Rundschau, 97:221-225.
- Shan L, Marti T, Sollid LM, Gray GM and Khosla C (2004) Comparative biochemical analysis of three bacterial prolyl endopeptidases: implications for coeliac sprue. Biochem J, 383:311-318.
- Shevchenko A, Wilm M, Vorm O and Mann M (1996) Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. Anal Chem, 68:850-858.

- Shewry PR, Halford NG and Tatham AS (1989) The high molecular weight subunits of wheat, barley and rye: genetics, molecular biology, chemistry and role in wheat structure and functionality, Oxford Surveys on Plant Molecular Cell Biology, 6:173-219.
- Shewry PR and Tatham AS (1990) The prolamin storage proteins of cereal seeds: structure and evolution. Biochem J, 267:1-12.
- Shewry PR and Tatham AS (1997) Disulphide bonds in wheat gluten proteins. J Cereal Sci, 25:207-227.
- Shewry PR (1999) The synthesis, processing, and deposition of gluten proteins in the developing grain, Cereal Foods World, 44:587-589.
- Shewry PR and Halford NG (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. J Exp Botany, 53:947-958.
- Shewry PR, Halford NG, Belton PS and Tatham AS (2002) The structure and properties of gluten: an elastic protein from wheat grain. Philos Trans R Soc Lond B Biol Sci. 357:133-142.
- Shewry PR and Lookhart GL (2003) Wheat gluten protein analysis. American Association of Cereal Chemists, St. Paul, Minnesota.
- Shewry PR, Halford NG and Lafiandra D (2003a) Genetics of wheat gluten proteins. Adv Genet, 49:111-184.
- Shewry PR, Halford NG, Tatham AS, Popineau Y, Lafiandra D and Belton PS (2003b) The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. Adv Food Nutr Res, 45:219-302.
- Sievert D, Sapirstein HD and Bushuk W (1991) Changes in electrophoretic patterns of acetic acid-insoluble wheat flour proteins during dough mixing. J Cereal Sci, 14:243-256.
- Singh NK and Shepherd KW (1988) Linkage mapping of genes controlling endosperm storage proteins in wheat. Theor Appl Genet, 75:628-641.
- Singh NK, Donovan GR, Carpenter HC, Skerritt JH and Langridge P (1993) Isolation and characterization of wheat triticin cDNA revealing a unique lysine-rich repetitive domain. Plant Mol Biol, 22:227-237.

- Singh NK, Shepherd KW, Langridge P and Gruen LC (1991) Purification and biochemical characterization of triticin, a legumin-like protein in wheat endosperm. J Cereal Sci, 13:207-219.
- Sollid LM (2002) Coeliac disease: dissecting a complex inflammatory disorder. Nat Rev Immunol, 2:647-655.
- Spaenij-Dekking EH, Kooy-Winkelaar EM, Nieuwenhuizen WF, Drijfhout JW and Koning F (2004) A novel and sensitive method for the detection of T cell stimulatory epitopes of alpha/beta- and gamma gliadin. Gut, 53:1267-1273.
- Spaenij-Dekking L, Kooy-Winkelaar Y, Van Veelen P, Drijfhout JW, Jonker H, Van Soest L, Smulders MJM, Bosch D, Gilissen LJWJ and Koning F (2005) Natural variation in toxicity of wheat potential for selection of non-toxic varieties for celiac disease patients, Gastroenterology, 129:797-806.
- Sparks CA and Jones HD (2004) Transformation of wheat by biolistics. In: Curtis IP (ed) Transgenic crops of the world-essential protocols. Kluwer Academic Publishers, Dordrecht, 19-34.
- Stoger E, Parker M, Christou P and Casey R (2001) Pea legumin overexpressed in wheat endosperm assembles into an ordered paracrystalline matrix. Physiol, 125:1732-1742.
- Stoger E, Williams S, Keen D and Cristou P (1999) Constitutive versus seed specific expression in transgenic wheat: temporal and spatial control. Transgen Res, 8:73-82.
- Sturgess R, Day P, Ellis HJ, Lundin KE, Gjertsen HA, Kontakou M and Ciclitira PJ (1994) Wheat peptide challenge in coeliac disease. Lancet, 343:758-761.
- Takaiwa F, Yamanouchi U, Yoshihara T, Washida H, Tanabe F, Kato A and Yamada K (1996) Characterization of common cis-regulatory elements responsible for the endosperm-specific expression of members of the rice glutelin multigene family. Plant Mol Biol. 30:1207-1221.
- Tao HP, Adalsteins AE and Kasarda DD (1992) Intermolecular disulfide bonds link specific high-molecular-weight glutenin subunits in wheat endosperm. Biochim Biophys Acta, 1159:13–21.

- Thompson T (2000) Questionable foods and the gluten-free diet: Survey of current 4 recommendations. J Am Diet Assoc, 100:463-465.
- Tosi P, D'Ovidio R, Napier JA, Bekes F and Shewry PR (2004) Expression of epitope-tagged LMW glutenin subunits in the starchy endosperm of transgenic wheat and their incorporation into glutenin polymers. Theor Appl Genet, 108:468-476.
- Uthayakumaran S, Gras PW, Stoddard FL and Bekes F (1999) Effect of varying protein content and glutenin-to-gliadin ratio on the functional properties of wheat dough. Cereal Chem, 76:389-394.
- Vader LW, de Ru A, Van Der Wal Y, Kooy YM, Benckhuijsen W, Mearin ML, Drijfhout JW, Van Veelen P and Koning F (2002a) Specificity of tissue transglutaminase explains cereal toxicity in celiac disease. J Exp Med, 195:643-649.
- Vader W, Kooy Y, Van Veelen P, De Ru A, Harris D, Benckhuijsen W, Peña S, Mearin L, Drijfhout JW and Koning F (2002b) The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides. Gastroenterology, 122:1729-1737.
- Vader LW, Stepniak DT, Bunnik EM, Kooy YM, De Haan W, Drijfhout JW, Van Veelen PA and Koning F (2003) Characterization of cereal toxicity for celiac disease patients based on protein homology in grains. Gastroenterology, 125:1105-1113.
- Van De Wal Y, Kooy Y, Van Veelen P, Pena S, Mearin L, Molberg O, Lundin L, Mutis T, Benckhuijsen W, Drijfhout JW and Koning F (1998) Small intestinal T cells of celiac disease patients recognize a natural pepsin fragment of gliadin. Proc Natl Acad Sci USA, 95:10050-10054.
- Van Den Broeck HC, America AHP, Smulders MJM, Gilissen LJWJ and Van Der Meer IM (2008) Staining efficiency of specific proteins depends on the staining method used: Wheat gluten proteins. Proteomics (in press).
- Van Lonkhuijsen HJ, Hamer RJ and Schreuder C (1992) Influence of specific gliadins on the breadmaking quality of wheat. Cereal Chem, 69:174-177.
- Van Overbeek FM, Uil-Dieterman IG, Mol IW, Köhler-Brands L, Heymans HS and Mulder CJ (1997) The daily gluten intake in relatives of patients with coeliac disease compared with that of the general Dutch population. Eur J Gastroenterol Hepatol, 9:1097-1099.

- Vellanoweth RL and Okita TW (1993) Analysis of nuclear proteins interacting with a wheat α/β-gliadin seed storage protein gene. Plant Mol Biol, 22:25-41.
- Ventura A, Magazzu G and Greco L (1999) Duration of exposure to gluten and risk for autoimmune disorders in patients with celiac disease. SIGEP Study Group for Autoimmune Disorders in Celiac Disease. Gastroenterology, 117:297-303.
- Vicente-Carbajosa J, Moose SP, Parsons RL and Schmidt RJ (1997) A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. Proc Natl Acad Sci USA, 94:7685-7690.
- Voinnet O (2005) Induction and suppression of RNA silencing: Insights from viral infections. Nat Rev Genet, 6:206-220.
- Von Buren M (2001) Polymorphisms in two homeologous gamma-gliadin genes and the evolution of cultivated wheat. Genet Res Crop Evol, 48:205-220.
- Wang M, Van Vliet T and Hamer RJ (2005) Interaction of water unextractable solids and xylanase with gluten protein: effect of wheat cultivar. J Cereal Sci, 41:251-258.
- Weegels PL, Hamer RJ and Schofield JD (1995) RP-HPLC and capillary electrophoresis of subunits from glutenin isolated by SDS and Osborne fractionation. J Cereal Sci, 22:211-224.
- Weegels PL, Hamer RJ and Schofield JD (1997) Depolymerisation and re-polymerisation of wheat glutenin during dough processing. II. Changes in composition. J Cereal Sci, 25:155-163.
- Weegels PL, Van De Pijpekamp AM, Graveland A, Hamer RJ and Schofield JD (1996) Depolymerisation and repolymerisation of wheat gluten during dough processing 1. Relationships between GMP content and quality parameters. J Cereal Sci, 23:103-111.
- Wellner N, Mills EN, Brownsey G, Wilson RH, Brown N, Freeman J, Halford NG, Shewry PR and Belton PS (2005) Changes in protein secondary structure during gluten deformation studied by dynamic Fourier transform infrared spectroscopy. Biomacromolecules, 6:255-61.
- Wieser H (2007) Chemistry of gluten proteins. Food Microbiol, 24:115-119.

- Wieser H and Kieffer R (2001) Correlations of the amount of gluten protein types to the technological properties of wheat flours determined on a micro-scale. J Cereal Sci, 34:19-27.
- Wieser H, Koehler P, Folck A and Becker D (2006) Characterization of Wheat with Strongly Reduced α-Gliadin Content. Proceedings of the 9th International Gluten Workshop, San Francisco, September 14-16.
- Wiley PR, Tosi P, Evrard A, Lovegrove A, Jones HD and Shewry PR (2007) Promoter analysis and immunolocalisation show that puroindoline genes are exclusively expressed in starchy endosperm cells of wheat grain. Plant Mol Biol, 64:125-136.
- Wu C, Washida H, Onodera Y, Harada K and Takaiwa F (2000) Quantitative nature of the prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal cis-element requirements for endosperm-specific gene expression. Plant J, 23:415-421.
- Young TE and Gallie GR (1999) Analysis of programmed cell death in wheat endosperm reveals differences in endosperm development between cereals. Plant Mol Biol, 39:915-926.

Summary

Coeliac disease (CD) is a disorder which is characterised in patients by a permanent intolerance to gluten proteins. CD presents itself by chronic diarrhoea, osteoporosis, lymphoma and several other clinical symptoms. Wheat of low CD-toxicity (with gluten proteins containing significantly lower levels of CD-toxic epitopes) would be beneficial, both for existing CD patients and for the prevention of CD in general. However, commercial considerations dictate that any reductions in CD-toxicity of wheat must not significantly compromise its good technological properties. The α-gliadins are considered an important factor in causing CD because of their role in provoking innate, as well as adaptive, immune responses. In this thesis, we have discussed several approaches that could be taken to develop low CD-toxic hexaploid bread wheat (with A, B and D genomes).

In Chapter 1, CD and the relationship between CD and gluten are described, as well as the biochemical and genetic background of gluten.

Chapter 2 outlines the screens performed to find T-cell stimulatory (CD-toxic) epitopes in the α-gliadins from the three wheat genomes. This resulted in the discovery that the D genome α-gliadins in particular encode multiple CD-toxic epitopes.The homologous genes in the A and B genomes possessed several SNPs leading to modified amino acid sequences and potentially reduced CD-toxicity. We also found a large number of pseudogenes in the corresponding loci on each of the genomes, especially in those of the B-genome.

Chapter 3 describes the expression of α-gliadin genes which was studied at the promoter level using chimaeric α-gliadin::GUS constructs in transgenic wheat. GUS was expressed from 11 days after flowering (DAF) to maturity in the endosperm and in the aleurone layer. Using a specific α-gliadin antibody, we could detect α-gliadin deposition throughout the endosperm. No α-gliadins were found in the aleurone layer, which may have been caused by the lack of promoter elements naturally present upstream to the promoter fragment used. It should be noted that gliadins were mainly deposited in the protein bodies.

We further studied the relation between protein bodies in developing wheat endosperm and glutenin particles in mature wheat flour (Chapter 4). Our results confirm the hypothesis of

Don *et al.* [2003a] that glutenin particles originate from protein bodies. We were able to isolate glutenin particles from immature wheat and observe that these particles are of the same size as protein bodies. In contrast, glutenin particles from mature wheat were much larger. This suggests that during the desiccation phase individual protein bodies can fuse to form the glutenin particles as found in wheat flour.

Chapter 5 describes a set of Chinese Spring deletion lines that were selected to obtain a set of wheat lines lacking HMW-GS, LMW-GS or gliadins. All lines were characterized using a panel of antibodies against specific CD-toxic epitope peptides (one HMW-GS peptide, two LMW-GS peptides and two α-gliadin peptides). Lines were also characterised for protein composition and technological properties. We found that deleting the HMW-GS led to an absolute loss of technological properties along with a limited reduction of CD-toxicity. The loss of chromosome 6D α-gliadins led to a large decrease in CD toxicity whereas technological properties were less affected. Our results indicate that a strategy of removing chromosome 6D α-gliadins in wheat breeding would result in a large decrease of CD toxicity with a limited loss of technological properties. It was concluded that such loss could be technologically corrected for, making the removal of 6D α-gliadins a feasible option to reduce CD-toxicity while retaining technological properties.

RNA interference (RNAi) methodology was utilised to specifically silence the expression of CD-toxic proteins. The results from this preliminary study are described in Chapter 6. RNAi had been used before by other researchers to down-regulate the complete α-gliadin family however, our specific (siRNA) approach was new and not yet established in plants. Wheat plants were transformed with a siRNA construct targeted at a specific CD-toxic epitope sequence. Transformed plants in the T_2 generation were not yet homozygous and T_3 seeds showed segregation for the transgene which made phenotypic and molecular analysis more difficult. Analyses of the transcriptome of single seeds using pyrosequencing showed the expression balance between the different α-gliadin mRNAs. This revealed that the seeds of one line had a lower expression and another line had a higher expression of the targeted α-gliadin genes. Analysis on the protein level in 2D gels showed that the expression of several proteins was down-regulated. Sequencing of several of these protein spots showed that these were γ-gliadins. This

may be due to off-targeting of the siRNA. Detailed analysis is still on-going at the time of completion of this thesis.

The final Chapter consists of a general discussion. First we considered whether it was correct to focus on the down-regulation of chromosome 6D α-gliadins. From a technological point of view, the removal of 6D α-gliadins resulted in increased dough mixing times and decreased rheology ($T_{1/2}$). 6D α-gliadin removal also affected particle formation; gutenin particles from these flours were smaller than those from wild type flour. The effects on technological properties could be corrected for by blending such flours with CD-safe flour from, e.g., oats or maize. Removal of α-gliadins encoded on chromosome 6D would result in the largest decrease in T-cell epitopes compared to other gliadins and glutenins. *T. monococcum* and *T. durum* lacking the D-genomes are reported less CD-toxic than *T. aestivum*. In addition, the number of other CD epitopes should be lowered as well, through selective breeding strategies, to further decrease CD-toxicity. Finally, we discussed who may benefit most from CD-low toxic wheat. For sensitive patients, wheat cultivars lacking the 6D encoded α-gliadins may still be CD-toxic if also not all other epitopes have been removed. But such cultivars might be beneficial for those individuals who are susceptible to CD and have not yet acquired the disease or have not yet developed acute clinical symptoms, provided they will be exposed to significantly lower levels of CD-toxic epitopes. This is of particular important since it has been reported that the ratio of known to undiagnosed cases of CD is 1:7. However, for preventive use of low CD-toxic wheat, it is important that it is grown on a large scale. This requires that low-CD toxic wheat meets agronomic criteria with regard to yield, stability, plant disease resistance and technological criteria. Developing commercially valuable CD-low-toxic wheat will take at least 5-10 years.

Samenvatting

Coeliakie is wereldwijd één van de meest voorkomende voedselintoleranties. Ongeveer 1% van de Westerse bevolking heeft er last van. De ziekte wordt gekenmerkt door een chronische ontstekingsreactie in de dunne darm ten gevolge van het eten van voedingsproducten waarin gluteneiwitten aanwezig zijn. Deze eiwitten komen voort uit de granen tarwe, gerst en rogge. Vooral gluten uit tarwe is een veelvuldig toegepast eiwit in de voedingsindustrie wegens de unieke visco-elastische eigenschappen. Uiteraard is het aanwezig in (tarwe)brood en veel andere bakkerijproducten, maar het eiwit wordt ook gebruikt in diverse levensmiddelen waarin het niet direct verwacht wordt, zoals in diverse vleesproducten (worsten), diverse soepen en sauzen, snoepgoed, bier, enz. Ten gevolge van de ontstekingsreactie bij gevoelige personen vlakt het darmepitheel sterk af waardoor een aanzienlijke verkleining van het darm oppervlakte plaatsvindt en de opnamecapaciteit voor nutriënten ernstig verminderd wordt. Symptomen van coeliakie zijn chronische buikpijn, diarree, en vermoeidheid. Probleem is dat naar schatting slechts 15% van de patiënten goed gediagnosticeerd is. Onbehandelde patiënten hebben kans op het ontwikkelen van groeiachterstand, en op latere leeftijd van osteoporose, onvruchtbaarheid en zelfs lymfomen. De enige remedie voor deze patiënten is een levenslang strikt glutenvrij dieet.

Door de grote en toenemende hoeveelheid glutenbevattende voedingsmiddelen is het volgen van een glutenvrij dieet geen gemakkelijke taak. Daarom zal de ontwikkeling van tarwe met lage coeliakie-toxiciteit een belangrijke bijdrage kunnen leveren aan de kwaliteit van leven, niet alleen van de personen die weten dat ze de ziekte hebben, maar vooral ook van hen die het niet weten en al lange tijd met veelal vage medische klachten rondlopen. In dit proefschrift wordt beschreven of we dergelijke tarwes kunnen maken.

Tarwe (*Triticum*) is een complex geslacht met een groot aantal soorten. De meeste tarwe soorten zijn diploid. Deze planten hebben in hun cellen een dubbele set van het erfelijk materiaal. Sommige van deze diploide soorten zijn met elkaar gehybridiseerd op zo'n manier dat het erfelijk materiaal bij elkaar opgeteld is. We noemen deze tetraploid. Zo kennen we bijvoorbeeld *Triticum dicoccoides (T. turgidum),* die een hybride is van *T. urartu* (AA genoom) en *T. speltoides* (BB genoom). *T dicoccoides* heeft als gevolg het AABB genoom. Later heeft een verdere hybridisatie

met *T. tauschii* (DD genoom) geleid tot het ontstaan van een nieuwe, hexaploide soort, *T. aestivum* (AABBDD), onze broodtarwe. Van deze diploide, tetraploide en hexaploide soorten zijn weer heel veel variëteiten bekend, die door selectie en verdere veredeling ontstaan zijn. Deze soort hybridisaties hebben ook gevolgen gehad voor de complexiteit van het gluteneiwit in de zaden. Dit eiwit is in de diploide soorten al complex en opgebouwd uit verschillende typen: α-, β-, γ- en ω-gliadinen en LMW- en HMW-gluteninen; in de tetra- en hexaploide tarwesoorten is dit nog complexer geworden. De vraag is nu of al deze gluteneiwitten in al die verschillende soorten en variëteiten verschillen laten zien in de mate van coeliakie-toxiciteit. We weten inmiddels dat de α-gliadinen beschouwd worden als één van de belangrijke gluteneiwitten bij het veroorzaken van coeliakie. Deze α-gliadinen bevatten het meest vaak bepaalde kleine *epitopen* die door de cellen van het immuunsysteem herkend worden, deze cellen vervolgens activeren en aanzetten tot de het ontwikkelen van ontstekingsreacties. Wellicht zouden op basis van verschillen in de coeliakie-toxiciteit van α-gliadinen strategieën kunnen worden ontwikkeld, gericht op de productie van coeliakie-veilige(r) tarwevariëteiten met weinig of geen epitopen. Zoals hierboven vermeld, zijn de visco-elastische eigenschappen van gluten belangrijk voor de voedingsindustrie en vooral in de bakkerij-industrie. Deze eigenschappen moeten uiteraard behouden blijven in deze nieuw te ontwikkelen rassen. In dit proefschrift worden verschillende benaderingen besproken voor het ontwikkelen van laag-coeliakie-toxisch tarwe met behoud van de technologische eigenschappen.

In Hoofdstuk 1 (algemene inleiding) wordt in detail uitgelegd wat de ziekte coeliakie inhoudt en wordt dieper ingegaan op de relatie tussen coeliakie en de gluten eiwitten. Hier komen de epitopen aan de orde en wordt de stand van zaken van de wetenschap hierover toegelicht. Ook wordt de genetische en moleculaire achtergrond van de gluteneiwitten beschreven. Vervolgens wordt ingegaan op de strategieën die in dit proefschrift ontwikkeld zijn voor de ontwikkeling van laag-coeliakie-toxische tarwe. Deze strategieën omvatten selectie aspecten, verandering van genexpressie, gebruik van deletielijnen, en toepassing van RNA-interferentie. Tot slot wordt de biochemische achtergrond van de industriële kwaliteit aan de hand van het hyper-aggregatiemodel toegelicht.

Hoofdstuk 2 beschrijft de isolatie en epitoopkarakterisatie van enkele honderden α-gliadine genen afkomstig uit de drie verschillende tarwesoorten (met AA, BB of DD genoom) die

uiteindelijk in broodtarwe met het AABBDD genoom gehybridiseerd zijn. Het bleek dat de α-gliadine genen naar het oorspronkelijke genoomtype herleid konden worden. Ook werd ontdekt dat er vooral op het D-genoom α-gliadinen gecodeerd zijn die verschillende coeliakie toxische epitopen bevatten. In de andere genomen (A en B) bleken specifieke veranderingen (puntmutaties) in het DNA de epitopen zodanig te wijzigen dat daardoor de coeliakie-toxiciteit naar verwachting verlaagd is. We hebben ook een grote hoeveelheid α-gliadine-pseudogenen gevonden in de verschillende tarwe genomen. Deze pseudogenen komen waarschijnlijk niet tot expressie en zorgen daardoor niet voor coeliakie-toxiciteit.

In Hoofdstuk 3 hebben we de expressie van de α-gliadine genen onderzocht op promotor niveau door een GUS gen achter een α-gliadine promotor te plaatsen en dit in de tarweplant tot expressie te laten komen. De α-gliadine expressie was, met behulp van een chemische reactie, zichtbaar in de graankorrel als een blauwe kleur. De promotor bleek in het endosperm (kiemwit) inclusief de aleuronlaag actief te zijn. Dit werd waargenomen vanaf 11 dagen na bloei tot en met in het rijpe zaad. Met behulp van een specifiek α-gliadine antilichaam werd het α-gliadine eiwit ook in het endosperm gedetecteerd, maar, in tegenstelling tot de promotor activiteit, niet in de aleuronlaag. De oorzaak van dit verschil kan liggen in het feit dat er verderop, voorbij het einde van het promotorfragment dat we in deze studie gebruikt hebben, andere factoren aanwezig zijn die de expressie in de aleuronlaag regelen. We hebben het α-gliadine-eiwit vooral waargenomen in bolvormige eiwitlichaampjes (protein bodies).

We hebben in Hoofdstuk 4 de relatie verder onderzocht van de protien bodies in onrijpe tarwekorrels met de technologisch belangrijke gluteninedeeltjes in rijpe tarwekorrels (en in tarwebloem). Onze resultaten versterken de theorie van Don en collega's, die in 2003 veronderstelden dat de gluteninedeeltjes in bloem afkomstig zijn van de eiwitlichaampjes in onrijpe tarwekorrels. We hebben gluteninedeeltjes uit onrijp zaad geïsoleerd en zagen dat deze deeltjes even groot zijn als de protien bodies en eenzelfde eiwitsamenstelling en oligomeeropbouw hadden. De gluteninedeeltjes in bloem bleken echter veel groter te zijn. Dit is zeer waarschijnlijk het gevolg van het feit dat tijdens de uitdrogingsfase van de zaadrijping individuele protien bodies fuseren tot grotere gluteninedeeltjes.

In Hoofdstuk 5 zijn deletielijnen van het ras Chinese Spring bestudeerd. Deze waren zodanig geselecteerd dat ze een set aan planten opleverde die ieder verschillende groepen gluteneiwitten (HMW-GS, LMW-GS of gliadinen) misten. De glutengenen van tarwe zijn gelegen op de chromosomen 1 en 6. Dit geldt voor alle drie de typen genomen A, B en D. De geselecteerde lijnen zijn getest met verschillende antilichamen (één gevoelig voor een HMW-GS peptide, twee voor een LMW-GS peptide en twee antilichamen voor α-gliadine peptiden). De zaden van de planten werden daarnaast gekarakteriseerd op hun eiwitsamenstelling en hun technologische eigenschappen. Het bleek dat het verwijderen van de HMW-GS leidt tot een definitief verlies van de technologische eigenschappen, gapaard gaand met slechts een beperkt verlies van coeliakie-toxiciteit. Het verwijderen van LMW-GS liet geen meetbaar verlies van coeliakie-toxiciteit zien, terwijl het verwijderen van de LMW-GS genen op chromosoom 1B de technologische eigenschappen negatief beïnvloedde. Het verwijderen van de α-gliadine genen op chromosoom 6D leidde daarentegen tot een groot verlies aan coeliakie–toxiciteit, terwijl de technologische eigenschappen veel minder werden beïnvloed. Deze resultaten laten zien dat een veredelingsstrategie gericht op het verwijderen van de op chromosoom 6D gecodeerde α-gliadinen een sterk verlagend effect zal hebben op de coeliakie-toxiciteit terwijl de technologische eigenschappen behouden blijven. Voor het eventuele verlies aan technologische eigenschappen kan waarschijnlijk voldoende worden gecorrigeerd.

In Hoofdstuk 6 is een nieuwe en specifieke RNA interferentie (siRNA) techniek onderzocht, die erop is gericht om specifiek de expressie van coeliakie-toxische eiwitgenen uit te schakelen. RNA interferentie is eerder gebruikt door andere onderzoekers om succesvol de complete α-gliadine familie uit te schakelen. Onze specifieke siRNA aanpak is nog niet eerder toegepast in planten. Wij hebben tarwe planten getransformeerd met een siRNA construct dat is gericht op het uitschakelen van een specifiek coeliakie-toxisch epitoop. Getransformeerde planten in de T_3 generatie lieten uitsplitsing zien. Dit duidt op nog aanwezige heterogeniteit. Analyse van het transcriptoom van jonge, individuele zaden liet de expressie balans zien tussen de verschillende α-gliadine mRNAs. Hieruit bleek dat één van de transformante lijnen een lager expressieniveau van de beoogde epitoop-bevattende α-gliadine genen vertoonde, terwijl een andere lijn juist een hogere expressie liet zien. Analyse op eiwitniveau in 2D gelen gaf aan dat de

expressie van verschillende eiwitten omlaag gebracht was. Op basis van aminozuur sequenties van het eiwit in de desbetreffende eiwitspots van deze gelen bleek het hier γ-gliadinen te betreffen. De verlaging van de expressie van deze niet beoogde genen kan het gevolg zijn van aspecifieke gerichtheid van het siRNA molecuul. Op het moment van voltooiing van dit proefschrift is de gedetailleerde analysen van de transgene planten nog niet afgerond en wordt nog verder voortgezet.

Hoofdstuk 7 bevat een algemene discussie van de resultaten van dit proefschrift. Eerst bediscussiëren we de juistheid van de strategie om te focussen op het verlagen van de α-gliadinen afkomstig van chromosoom 6D. Vanuit een technologisch oogpunt resulteerde het verwijderen van 6D α-gliadinen in een verhoogde mixtijd en een snellere deegrelaxatie ($T_{1/2}$) van het deeg. Verder leidde het verwijderen van 6D α-gliadinen tot kleinere gluteninedeeltjes. We verwachten dat de afwijkende technologische eigenschappen gecorrigeerd kunnen worden door het inmengen van coeliakie-veilige bloem van bijvoorbeeld haver of maïs. Het verwijderen van chromosoom 6D gecodeerde α-gliadinen zal leiden tot de grootste eliminatie van T-cel epitoop eiwitfragmenten vergeleken met de verwijdering van gliadine- of glutenine-genen van andere chromosomen van de andere genomen. In de diploide tarwesoort eenkoorn (*T. monococcum*) en de tetraploide pastatarwe (*T. dicoccoides/turgidum*) ontbreekt het D-genoom. In de literatuur zijn deze twee soorten beschreven als minder coeliakie-toxisch dan de gewone hexaploide broodtarwe (*T. aestivum*). Hierin zouden de andere T-cel epitopen ook verwijderd moeten worden om de coeliakie-toxiciteit verder te verlagen. We bediscussiëren in dit laatste hoofdstuk ook wie het meeste baat heeft bij deze laag-coeliakie-toxische tarwevariëteiten. Voor gevoelige coeliakie patiënten zullen tarwevariëteiten zonder 6D gecodeerde α-gliadinen waarschijnlijk nog steeds toxisch zijn omdat hierin nog niet alle andere epitopen verwijderd zijn. Een dergelijke variëteit is echter wel nuttig in de preventie van de ontwikkeling van coeliakie bij mensen die vatbaar zijn voor de ziekte of voor patiënten die nog geen klinische symptomen hebben ontwikkeld. Consumptie van een dergelijke variëteit zal leiden tot een verminderde blootstelling aan coeliakie-toxische epitopen. Dit is vooral belangrijk omdat uit studies is gebleken dat de meerderheid van de patiënten (85%) nog niet of niet goed gediagnosticeerd is. Voor preventief gebruik van laag-coeliakie-toxisch tarwe is het belangrijk dat het op grote schaal kan worden geconsumeerd en dat

het daarom ook op grote schaal moet worden verbouwd. Deze laag-toxische tarwevariëteiten moeten dan ook zeer goede agronomische (opbrengst, stabiliteit en ziekteresistentie) en goede technologische eigenschappen hebben. Naar verwachting zal de ontwikkeling van dergelijke coeliakie-veilige tarwevariëteiten tot grootschalig commercieel niveau nog minstens 5 tot 10 jaar in beslag nemen.

Dankwoord / Acknowledgements

Het dankwoord is waarschijnlijk wel het meest gelezen hoofdstuk van het proefschrift. Hierin probeer ik de mensen te bedanken die me geholpen hebben vooral tijdens de afgelopen 4 jaar. Hierbij dank ik jullie voor alle hulp die ik heb mogen ontvangen van iedereen. Ik kan helaas niet iedereen bij naam noemen. Voor de mensen die ik hier niet genoemd heb: wel bedankt en jullie hulp is niet vergeten en zeker gewaardeerd!

Ten eerste wil ik mijn begeleiders bedanken: Rob, Dirk, Luud en René. Rob, je hebt me heel goed geholpen met dit onderzoek. De buitenlandse congressen via de AACC waarvoor je me stimuleerde om naar toe te gaan waren erg leuk en leerzaam. Je betrokkenheid bij het project was fijn en motiverend, ook het organiseren van de meetings van de glutenclub. René en Luud, jullie hebben me als dagelijkse begeleiders erg goed geholpen hier op PRI. De tweewekelijkse meetings waren motiverend en structurerend. Luud, jij wist alles goed te regelen en bent een soort visionair die een goede helikopterview heeft. René, jij was altijd erg betrokken en lekker snel met dingen nakijken wat me erg goed heeft geholpen. Ook bedankt voor het meegaan naar de UK.

Tijdens dit project heb ik hulp gehad van verschillende studenten. Dank je wel Cees, je was een gezellige en goede student en hebt belangrijk voorwerk gedaan voor het tot stand komen van het vierde hoofdstuk. Esther, bedankt voor het samen opstarten van het project waarbij jij je thesis schreef en ik aan het inlezen was. Noor, bedankt voor je werk en interesse in het project. *Thomas, merci beaucoup pour ton travail dans le projet RNAi. Sans ton aide le cinquième chapitre n'aurait pas été possible*.

Alle collega's bij PRI van Bioscience, ik kan jullie niet allemaal bij naam noemen, maar stuk voor stuk dank jullie wel voor alle hulp en voor de interessante werkomgeving! Een aantal collega's hebben me geholpen bij het tot stand komen van de verschillende hoofdstukken. In het bijzonder wil ik hier Hetty noemen voor je expertise in het proteomics werk en je kritische blik op het werk, ik ben erg blij dat je met je PhD bent begonnen. Ik ben ook dank verschuldigd aan Froukje, Ingrid, Jan en Twan bij de verschillende werkzaamheden. Elma, dank je wel voor alle hulp tijdens deze vier jaar van onderzoek. Van TNO wil ik Wim Lichtendonk bedanken voor de

hulp met de deegproeven. En van het Nizo wil ik Jan Klok bedanken voor de hulp met het CSLM microscopie werk.

Dank je wel ook aan mijn kamergenoten Jules en Jos. Jullie waren gezellig en behulpzaam en ik heb me goed thuis gevoeld in ons mannenkantoor. Jules, en ook je familie, bedankt voor alles, zoals bijvoorbeeld de lekkere diners, borrels, hulp met verhuizen, gezelligheid, de wekelijkse spelletjes en nog veel meer. Nu ben je ook mijn paranimf samen met Ralph. Ralph, jij was mijn eerste begeleider op het lab van PRI en hebt me de fijne kneepjes van het moleculaire werk geleerd. Ik vond het erg gezellig met jou op het werk en daar buiten. We hebben ook een hele gezellige vakantie gehad samen in de USA!

Remko bedankt voor het ontwerpen van de cover en de uitnodiging en voor je vriendschap. Het was leuk mee te doen met de dansshow, walking gallery, en om te gaan zwemmen en te kletsen.

Het project CDC had een fijne multidisciplinaire samenstelling met partners in de medische, voeding- en industriële hoek. Van Leiden Medical Center, in het bijzonder dank je wel Frits, Liesbeth en Peter voor de hulp bij de verschillende hoofdstukken. Ook dank aan onze andere partners van het CDC zoals UMC Utrecht, TNO en NCV.

Dank je wel aan mijn collega's van de glutenclub waarin we interessante discussies hadden over de gluten eiwitten en waarmee we samen naar conferenties gingen. In het bijzonder bedank ik hier Clyde, Neleke, Hadi, Jendo, Cristina, Cathrine, Daan, Lieke en Eva. Eva bedankt voor je gezellige diners!

Het ACW clubje was een leuke groep collega AIO's. We waren in dezelfde tijd gestart en we waren allemaal "allergie AIO's". We hadden leuke borrels en konden zo van ideeën wisselen en over onze ervaringen tijdens het onderzoek praten. Dank je wel Evelien, Martijn, Mirco, Presilla, Margreet en Yvonne. Harry en Huub, bedankt voor de interesse en organisaties. Huub ook bedankt voor het lezen van het proefschrift. Verder iedereen bedankt van het ACW.

Bedankt, alle collega's van het laboratorium van Food Chemistry. Het was leuk dat ik al meteen in het begin mee mocht met de leerzame reis naar Japan, we hadden een leuke tijd daar, jammer dat ik de tweede keer verhinderd was. Ook bedankt voor de hulp die ik van verschillende mensen binnen Food Chemistry heb mogen ontvangen.

At PRI a lot of people came and went. Here I had the chance to meet a lot of international friends. We had a lot of nice dinners and drinks and this helped me at lot to keep up the good spirit. Thanks for these nice times we spent, I would like to put here all your names but I will limit myself to a few people in particular. Asaph, thanks for teaching me about the beauty of our research and your guidance and stimulation to start my PhD. Shital, you were there all the four years and you were very special to me. Tetty, and the rest of Indonesian community, thanks for the nice dinners and good times, especially during Ramadan. Salva, Ruud, Xinying and Raffa, thanks for the friendship and we had a nice time in Barcelona. Esra, thanks for showing me Istanbul. Sebastian, thanks for your friendship, man. Daniella, it was great that I could join your dancing show.

I also worked for 9 months in Rothamsted, UK. I would first like to thank Peter Shewry. When I started my PhD I just sent you an email to ask your opinion about some ideas that I had. At a conference in Wageningen we discussed about the different ideas and you were happy to give me the chance to work in your laboratory for as long as I wanted. Thanks for all your help Peter! For the transformation experiments especially Huw, Caroline and Mandy helped me a lot. Huw and Caroline, thank you very much for welcoming me in the lab and give me a comfortable stay. I could always come to ask for help which is much appreciated! Also thanks to all the members of the transformation lab and CPI division, in particular Mark, Wakako, Mariann, DP, Kexiu, Huixia, Javon, Angela, Lucia, Paul, Melloney, Smita, Sue and Helen. Mark, thanks a lot for the pleasant dinners and DVD evenings! Mandy, you were fabulous! You taught me everything about the transformation experiments and how to take care of our "babies". We had a lovely time talking about everything during our work in the laminar hoots. Often you let me enjoy dinners and delightful evenings at your house together with your lovely husband Rod. Jackie, thanks for your help and explanations and discussions on the work which was very helpful for the fourth chapter. Christina, thank you very much for teaching me everything about microscopy work and our nice talks we had during the work. Paula thanks for your help and discussions.

In England I lived in the Manor. In this house and also at work I got to know a lot of nice friends, which is especially important if you are away from home! We had a lot of parties and trips around the country and Europe. All my friends from Rothamsted, it was great to meet you all

and I will not forget one of you! Petra, thanks a lot for everything. Monica, Delia and Hiro, we had a nice time in Roma, I enjoyed it a lot.

From the plant breeding department I would like to thank in particular Krissana, Nicolas and Mariame, for the nice dinners and friendship! Also thanks to Estelle, Matieu, Antoine, Jeroen, Bjorn, Bjorn and all the others from Plant Breeding.

In Ede heb ik na het werk altijd lekker kunnen ontspannen met sporten samen met onder andere Alex, Marijn, Bart en Ruud. Bedankt voor de leuke tijd. Toch jammer dat ik niet geschikt was voor het patsergilde. Alex, jij in het bijzonder bedankt. We hebben veel tijd doorgebracht samen, wat erg gezellig was. We gingen veel sporten, 's morgens tijdens de zonsopgang in de ochtenddouw rennen in de uiterwaarden, bier zuipen, we hadden goede gesprekken en er is een mooie vriendschap ontstaan. Je hebt de tijd in Wageningen erg aangenaam gemaakt!

Nu is het ontspannend en ontstressend om weer Taekwondo te doen in een leuke groep. Margriet bedankt voor alles. En dank jullie wel alle sportvrienden van Sjabbens.

André en Dave, ik heb jullie tijdens deze periode minder gezien dan ik zou willen. Jullie zijn erg belangrijk voor me en bedankt voor de steun, motivatie en vriendschap.

Dank je wel aan al mijn familie en vrienden. Ons warme gezinnetje, Pap, Mam, Maarten en Melina, Wendy, Nathalie en Iris, bedankt voor alles, jullie waren er altijd, onvoorwaardelijk! Papa en Mama, dank jullie wel voor jullie onvoorwaardelijke liefde. Papa, je liet duidelijk merken trots te zijn dat nu je tweede zoon ook zijn doctors titel gaat halen. Dit geeft een goed gevoel en was ook een stimulans om dit te volbrengen. Mama, onze speciale band geeft me kracht en een sterke basis om alles te doen wat ik wil. Maarten, je bent een goed voorbeeld als grote broer, bedankt voor je hulp, steun en broederliefde. Melina, ik ben blij met jou als mijn schoonzus en bedankt voor alles. Wendy, Nathalie en Iris, van jullie heb ik liefde en steun mogen ontvangen en ik ben erg blij met jullie. *Anoma, merci beaucoup pour tout, ton amour, ton aide, ton écoute, ta compréhension et la stabilité que tu m'apportes! Je me réjouie de notre futur ensemble et de notre mariage cette année!*

Teun van Herpen

Curriculum Vitae

Teun van Herpen is geboren op 29 mei 1980 te Oss. In 1996 behaalde hij zijn MAVO diploma en is hij begonnen aan de beroepsgerichte opleiding laboratorium technieken met als specialisatie biochemie aan de ROC de Leijgraaf te Oss waarvoor hij in 2000 zijn diploma behaalde. Als afstudeer opdracht deed hij onderzoek naar de effecten van bioactieve stoffen op dikke darm kanker bij Numico Research in Wageningen. Hierna heeft hij zijn studie voortgezet aan de hogeschool voor laboratorium technieken met een dubbele afstudeer richting biochemie en biotechnologie. In 2003 behaalde hij hiervoor zijn diploma. Zijn tweede afstudeeropdracht vond plaats bij Plant Research International in de groep van Prof. dr. Bouwmeester aan het onderwerp terpenen en plantengezondheid. Dit onderzoek leverde ondermeer een publicatie op in het gezaghebbende tijdschrift Science. In januari 2004 begon hij zijn promotieonderzoek bij de Wageningen Universiteit (WUR) onder leiding van Prof. dr. Hamer, Dr. Smulders, Dr. Gilissen en Prof. dr. Bosch. Gedurende het tweede jaar heeft hij hiervoor onder begeleiding van Prof. Shewry gedurende 9 maanden gewerkt bij Rothamsted Research in Engeland aan de genetische modificatie van tarwe. De resultaten uit zijn promotieonderzoek heeft hij gepresenteerd op international congressen in de VS, Japan, Ierland, Oostenrijk en Turkije. Na vier jaar is hij gepromoveerd en de resultaten van dit onderzoek staan beschreven in dit proefschrift.

Teun van Herpen was born on the 29th of May 1980 in Oss. In 1996 he obtained his diploma for his secondary school and started the profession orientated education of laboratory techniques with the specialization biochemistry. In 2000 he obtained his diploma for this study. As a traineeship he studied the effects of bioactive compounds on colon cancer at Numico Research in

Wageningen. Hereafter he started his Bachelor study of laboratory techniques with a double specialization of biochemistry and biotechnology for which he obtained his diploma in 2003. His second traineeship was at Plant Research International in the group of Prof. dr. Bouwmeester on the subject terpenoids and plant health. As one of the highlights, this research resulted in a publication in the authoritative magazine Science. In January 2004 he started his PhD at Wageningen University (WUR) under supervision of Prof. dr. Hamer, Dr. Smulders, Dr. Gilissen and Prof. dr. Bosch. During his second year he worked for 9 months at Rothamsted Research in the UK under supervision of Prof. Shewry on the genetic modification of wheat. He presented the results from his research at international conferences in USA, Japan, Ireland, Austria and Turkey. After four years he promoted and the results of this research are described in this thesis.

List of publications

Related to this thesis

1. van Herpen TWJM, van den Broeck HC, Schuit C, Salentijn E, Koning F, Bosch D, Smulders MJM, Gilissen LJWJ, van der Meer IM, Hamer RJ. The feasibility of decreasing CD toxicity while retaining technological properties: A study with Chinese Spring deletion lines. **In prep for Biochimica et Biophysica Acta - Molecular Basis of Disease.**
2. Van Herpen TWJM, Riley M, Sparks C, Jones HD, Gritsch CS, Dekking L, Hamer RJ, Bosch D, Salentijn EMJ, Smulders MJM, Shewry PR, Gilissen LJWJ. Detailed analysis of the expression of an α-gliadin promoter and the deposition of α-gliadin protein during wheat grain development. **Submitted to Annals of Botany.**
3. van Herpen TWJM, Cordewener JHG, Klok HJ, Freeman J, America AHP, Bosch D, Smulders MJM, Gilissen LJWJ, Shewry PR, Hamer RJ. The origin and early development of wheat glutenin particles. **Submitted to Journal of Cereal Science.**
4. van Herpen TWJM, Goryunova SV, Salentijn E, Riley M, Sparks C, van Veelen PA, Bosch D, Gilissen LJWJ, Smulders MJM, Jones HD, Shewry PR, Hamer RJ. Alpha-gliadin genes from the A, B, and D genomes of wheat contain different sets of celiac disease epitopes. **Proceedings of the 9th International Gluten Workshop, San Francisco**, September 14-16, 2006, 321-325.
5. Hamer RJ, van Herpen TWJM. Studies on the relation between Glutenin particles and Protein bodies. **Proceedings of the 9th International Gluten Workshop**, San Francisco, September 14-16, 2006, 218-221.
6. van Herpen TWJM, Goryunova SV, van der Schoot J, Mitreva M, Salentijn E, Vorst O, Schenk MF, van Veelen PA, Koning F, van Soest LJ, Vosman B, Bosch D, Hamer RJ, Gilissen LJ, Smulders MJ. Alpha-gliadin genes from the A, B, and D genomes of wheat contain different sets of celiac disease epitopes. **BMC Genomics.** 2006, 10:7:1.

Other publications

7. Kappers IF, Aharoni A, van Herpen TWJM, Luckerhoff LL, Dicke M, Bouwmeester HJ. Genetic engineering of terpenoid metabolism attracts bodyguards to Arabidopsis. **Science**. 2005, 23:309:2070-2072.
8. van Herpen TWJM, Schwab W, Jongsma MA, Bouwmeester HJ, Aharoni A. The Biosafety of Metabolic Change: Altering Plant Terpenoid Metabolism. **Genomics for Biosafety in Plant Biotechnology** 2004, 35-46.

Overview of completed training activities

Discipline specific activities

- Bioinformation Technology-1, VLAG, 2004
- Allergy Matters! conference, Wageningen, 2004
- AACC 3rd European Young Cereal Scientist Workshop, Ireland, Dublin, 2004
- ALW Meeting Experimental Plant Sciences, Lunteren, 2004, 2005, 2007
- Meetings of 'learned societies' and CDC meetings, 2004-2008
- Training period at Rothamsted Research (Prof. Shewry), Harpenden, UK, 2005, 2006, 2007
- AACC 4th European Young Cereal Scientist Workshop, Austria, Vienna, 2005
- AACC 5th European Young Cereal Scientist Workshop, Gaziantep, Turkey, 2006
- AACC Gluten Workshop, USA, San Francisco, 2006

General courses

- VLAG PhD week, 2004
- Scientific Writing, CENTA, 2004
- NWO Talent Day, 2005
- Getting your message across – Media skills for scientific researchers, SENSE, 2006
- Intuitive Intelligence, Wageningen University, 2006
- Personal Efficacy, Wageningen Graduate School, 2006
- Organising and supervising MSc thesis project, Wageningen University, 2006
- Career Assessment, Meijaard en Meijaard, 2007

Optional courses and activities

- Preparation of PhD research proposal
- PhD trip to Japan, Food Chemistry, 2004
- Literature study program Bioscience group, 2004-2008, Wageningen

Cover Art: Remko Zijlstra 2008
Model: Madelon/ Wheat ear
ezel1@live.nl

Remko Zylstra created this cover with a strong symbolic and aesthetic black and white picture of a woman's body.

This creation revealed itself after reading the specific topics of the thesis and through the books theme.

The picture tells us the story about coeliac disease which is mostly taking place in the intestine area.

The hands symbolize the care taking of this body part with the cause (the wheat ear) in the centre of the body as a safe welcome now.