Norwegian University
of Life Sciences

Master's Thesis 2018 60 ECTS
Faculty of chemistry, Biotechnology og Food Science

Identification of celiac disease related immunogenic peptides in Norwegian wheat types after *ex vivo* digestion

Rashida Rehman

ACKNOWLEDGMENTS

This thesis is performed as part of a two-year master program in Chemistry at the faculty of Chemistry, Biotechnology and Food Science (IKBM) at the Norwegian University of Life Sciences (NMBU). The majority of the laboratory work was performed at IKBM, while LC-MS analysis of the digested samples was performed at national institute of research in Avelino, Italy.

Firstly, I am greatly thankful to my main supervisor Gerd Vegarud and co-supervisors Anne Kjersti Uhlen and Tove G. Devold for their guidance along with helpful discussions and providing me an opportunity to travel to Italy for analysis of my samples. Secondly, I would like to thank Tora Asledottir for good instructions and company in the lab and the guidance for interpretation of results and help during thesis writing process. I would also like to thank Irene Comi for technical guidance in the lab. A special thanks to Gianfranco Mamone, for his help in interpretation of results. I am really thankful to Gianluca Picariello, Pasquale Ferranti, GioVanna and Laura for their help while staying in Italy.

Finally, I would like to thank my Husband Muhamad Tanveer for supporting me throughout my master thesis.

Norwegian University of Life Sciences

Ås

December 2018

_____

Rashida Rehman

ABSTRACT

Wheat has been part of human diet for centuries in the form of bread, pasta, cakes and pastries. There has been a lot of research to introduce better crop yield and disease resistant wheat varieties. Today's wheat varieties are more disease resistant and yield much more as compared to the old varieties. Modern wheat varieties have better baking qualities also, due to unique functional properties of the gluten. However, wheat is considered to cause many different food-born allergies and intolerances. Celiac disease is a food related enteropathy induced in genetically susceptible individuals by the ingestion of gluten proteins from wheat and gluten like proteins from barley and rye. It is due to the reason that the immunogenic peptides derived from gluten can induce autoimmune response in human body, resulting in villous atrophy. Until now the recommendation for the celiac disease is a lifetime gluten free diet.

The aim of the study was to investigate immunogenic peptides produced during *ex vivo* digestion of different wheat types including ancient wheat; einkorn, emmer, spelt and modern wheat varieties; Fram, Børsum, Bastian and Mirakel. All wheat types were digested according to a static *ex vivo* digestion model with human gastrointestinal (GI) juices. The Protein digestion pattern from all wheat types interpreted by SDS-PAGE and OPA showed degradation of some proteins after 120 min of gastric digestion and almost complete protein degradation after120 minutes of duodenal digestion. Mass spectrometric (LC-MS/MS) analyses of peptides released after 240 minutes of GI digestion revealed that the number of total peptide fragments and their amino acid sequences varied remarkably between different wheat types. In ancient wheat varieties fewer immunogenic peptides were identified compared to the modern varieties. 35, 49 and 71 immunogenic peptides were identified in einkorn, spelt and emmer, respectively, as compared to identification of 59, 78, 96 and 133 immunogenic peptides in Fram, Bastian, Børsum and Mirakel, respectively. The identified immunogenic peptides harboured different T-cell reactive epitopes. However, well-known 33mer LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF from α-gliadin was not observed in any of the wheat types. The same was observed for 26mer FLQPQQPFPQQPQQPYPQQPQQPFPQ and 25mer LGQQQPFPPQQPYPQPQPFPSQQPY. To conclude, all wheat types lead to the generation of immunogenic peptides, hence the analysed wheat types cannot be considered as safe for celiac patients. Further digestion of theses wheat samples by brush border enzymes is required to understand the complete behaviour of gluten in human GI system.

SAMMENDRAG

Hvete har vært en del av menneskers kosthold i århundrer i form av brød, pasta, kaker og andre bakverk. Det har vært mye forskning på hvete for å introdusere høyre avling- og sykdomsresistente hvetesorter. Dagens hvete sorter er mer sykdomsresistente og gir større avling enn gamle hvetesorter. Moderne hvetesorter har gode bakeegenskaper på grunn av de funksjonelle egenskapene i gluten. Hvete er også regnet som en årsak til matvareallergier og intoleranse. Cøliaki er en matrelatert tarmsykdom hos genetiske disponerte individer forårsaket av gluten proteiner fra hvete og gluten lignende proteiner fra bygg og rug. Immunogene peptider avledet fra gluten har vist evnen til å fremkalle auto immunrespons i menneskekroppen, noe som fører til ødeleggelse av tarmtotter. Den eneste foreslåtte kuren mot cøliaki er et livslangt glutenfritt kosthold.

Formålet med denne studien var å undersøke immunogene peptider produsert under *ex vivo* fordøyelse av forskjellige hvetetyper, inkludert gammel hvete; einkorn, emmer, spelt og moderne hvete varianter; Fram, Børsum, Bastian og Mirakel. Alle hvetetyper ble fordøyd i henhold til en statisk *ex vivo* fordøyelsesmodell med human mage- og tarmsaft (GI). Proteinfordøyelsen fra alle hvetetyper tolket ved hjelp av SDS-PAGE og OPA viste nedbrytning av noen proteiner etter 120 min magefordøyelse og nesten fullstendig protein nedbrytning etter 120 minutters tarmfordøyelse. Massespektrometriske (LC-MS/MS) analyser av peptider frigjort etter 240 minutters fordøyelse viste at antall totale peptidfragmenter og deres aminosyresekvenser varierte betydelig mellom hvetetyper. I de gamle hvetesortene ble færre immunogene peptider identifisert i forhold til de moderne varianter. 35, 49 og 71 immunogene peptider ble identifisert i henholdsvis i einkorn, spelt og emmer, sammenlignet med identifikasjon av 59, 78, 96 og 133 immunogene peptider i henholdsvis Fram, Bastian, Børsum og Mirakel. De identifiserte immunogene peptidene inneholdt forskjellige T-celle-reaktive epitoper. Imidlertid ble den velkjente 33mer LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF fra α-gliadin ikke observert i noen av de fordøyde hveteprøvene. Det samme ble observert for 26mer FLQPQQPFPQQPQQPYPQQPQQPFPQ og 25mer LGQQQPFPPQQPYPQPQPFPSQQPY. For å konkludere, alle hvetetyper førte til generering av immunogene peptider, og kan dermed ikke anses som sikre for cøliaki-pasienter. Videre fordøyelse av disse hveteprøvene med tarmtottenzymer er nødvendig for å forstå den fullstendige oppførselen til gluten i humant fordøyelsessystem.

ABBREVIATIONS

LMW-GS – Low molecular weight glutenins

HMW-GS – High molecular weight glutenins

MW – Molecular weight

GI – Gastrointestinal

HDJ – Human duodenal juices

HGJ – Human gastric juices

DH – Degree of hydrolysis

OPA – ortho-phthadialdehyde

ESI – Electrospray ionization

RP-HPLC – Reversed phase High performance liquid chromatography

MS – Mass spectrometer

SDS-PAGE – Sodium-Dodecyl Sulphate- Polyacrylamide Gel Electrophoresis

TAME – p-toluene-sulfonyl-L-arginine methyl ester

tTG2 – Tissue transglutaminase

CD – celiac disease

HLA-DQ – Human Leukocyte antigen

MHA – Major histocompatibility complex

APC – Antigen presenting cells

IL – Interleukin

IEL – Inter epithelial lymphocytes

APPENDEX

All the raw data of the LC-MS and detected peptides lists will be provided in USB flash drive on demand.

CONTENTS

INTRODUCTION

1.1 HISTORICAL EVOLUTION OF WHEAT

Wheat dominates worlds grain production along with maize and rice. These crops belong to grass family (Poaceae syn. Gramineae) and are primarily grown for their grains. Wheat is primarily grown in temperate regions of the world, but it is now widely adapted and grown in a variety of environments and altitudes. Being next to rice and maize, wheat is cultivated in more areas of the available land than any other crop (Gooding, 2009).

Pollen grain analyses show that these grasses were present 55 million years ago (MYA), in the Palaeocene, or maybe earlier (Kellogg, 2001). Genetic sequences (Huang et al., 2002) indicate that the emergence of wheat from barley and from rye is estimated to have occurred around 10-14 MYA and 7 MYA, respectively. All members of the Triticeae family have chromosomes in multiple set of seven (x = 7). The genome groups labelled as A, B, D, G, and S are relevant to wheat evolution (Kihara, 1929, Dewey, 1984). A schematic diagram that presents the linkage between ancient and modern wheat and their evolution has been described in **Figure 1**.

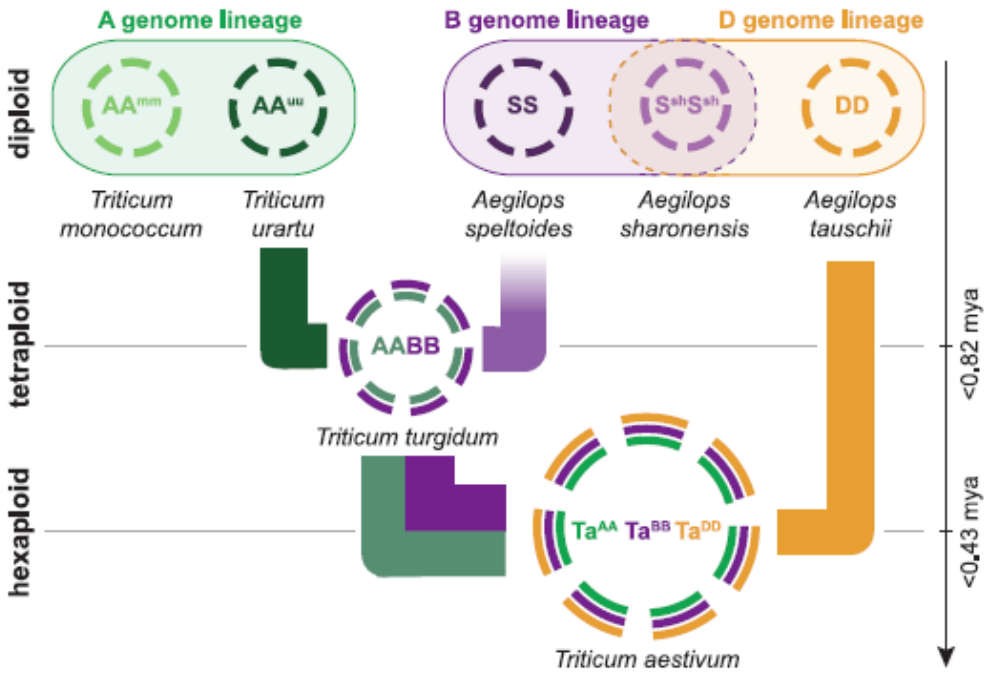


Figure 1: *Schematic diagram of the relationships between wheat genomes with history and genealogy. Names and nomenclature for the genomes are indicated within circles that provide a schematic representation of the chromosomal complement for each species. MYA, million years ago (International Wheat Genome Sequencing Consortium, 2014).*

Recently available wild diploid wheat species of einkorn are *Triticum. urartu* (genome $A^u A^u$) and *T. monococcum* subsp. *boeoticum* (A^bA^b). The most cultivated form of emmer *T. turgidum* subsp. *di-coccoides* genome $A^u A^ABB$ (tetraploid) is a hybrid of *T. urartu* and an ancestral goat grass (*Aegilpos speltoides*) (Huang et al., 2002, Brandolini et al., 2006) that appeared somewhere between 0.2 - 0.5 MYA. In these species, seeds are protected within spikelet, which makes it difficult for domestic use. For domestication, half-naked seeds are preferred which can easily be separated from the spikelet after maturation (Hillman and Davies, 1990). By the addition of the D Genome from wild goat grass *Aegilops. Tauschii (*DD) to *Triticum. turgidum* (AABB) developed to common wheat which is hexaploid (AABBDD) (Dvorak et al., 1998). These hexaploid wheat species are better adapted for domestication and have been distributed and utilized all over the globe (Nesbitt et al., 1996).

1.2 WHEAT BREEDING IN NORWAY

Norway is situated in western Scandinavia between latitudes 57°58' and 71°10'N. Only 1 mill ha (3%) of the total land area of about 324,000 km^2, is arable land (Lillemo and Dieseth, 2011). Norway is situated in a favourable position where the total amount of natural rainfall is enough to avoid water limitations problems (Grønlund et al., 2006). However, the year-to-year rainfall variation can be large. Pollen grain analysis from imprint of grain found in the remnants of burnt clay in Kråkerøya in Østfold provided evidence for the presence of emmer wheat (*Triticum dicoccum*) that dates back to 2500 BC (Mikkelsen, 1979). The main cereals cultivated in earlier times were naked barley (*Hordeum vulgare nudum*), with minor cultivation of einkorn (*T. monococcum*) and emmer wheat (*T. dicoccum*). During the late bronze age, cultivation of *T. spelta* and *T. aestivum* was observed, and naked barley was replaced by hulled barley (*H. vulgare*). Oats were introduced around 500 BC (Myhre, 2004). During the Viking and Medieval times, wheat was considered as a luxury grain only used by the elite class (Mikkelsen, 1979). Wild einkorn, emmer and spelt are cultivated in small amount nowadays for niche production and for preserving the genes (ostafjells.nlr.no). These ancient varieties are better suited for domestic use than mechanical processing used in today's baking industry (Uhlen, 1990).

In Norway, different spring and winter wheat varieties have been cultivated. Cultivation of spring wheat varieties is more favourable than winter wheat because of long and hard winter. Late harvesting of the previous crop and sometimes heavy rains in the planting season make the planting of winter wheat difficult (Strand, 1984). These uncertain conditions for winter

wheat make spring wheat yield more stable over the years (Statistics Norway). Most of the landraces cultivated here needed to develop good lodging resistance. Two landraces of spring wheat; Børsum, an awnless landrace from Ås and the awned landrace Østby from Tjølling in Jarlsberg; showed superior performance and were recommended for widespread cultivation before 1900. In early 1900, all races were susceptible to powdery mildew except Fram I and Fram II which had been developed by cross-breeding a mildew resistant type (J 03) identified from a landrace in Jåberg, with a resistant wheat type from USA (Vik, 1937). In 1952, the landmark variety Norrøna was developed by cross breeding Fram II with a Finnish variety Sopu (Bjaanes, 1962). This was superior in yield to all other varieties at that time and had good bread making quality. It was a landmark for the development of mildew resistant varieties with good baking qualities. Further in 1989, and later on, the spring wheat varieties named Bastian (Mjaerum, 1992) and then Mirakel was launched in 2015. These modern varieties have high crop yield, better resistance to diseases and good baking qualities (Uhlen et al., 2004). An overview of the Norwegian wheat types used for this study is presented in the **Table 1** below.

Table 1: *An overview of wheat types selected for the analysis along with their genome and origin.*

Wheat type	Accession	Variety	Breeding company/origin	Marked release
Einkorn	AA			
Emmer	AABB			
Spelt	AABBDD			
Common wheat	AABBDD	Fram	Norwegian landrace	before 1900
	AABBDD	Børsum	NLH	1936
	AABBDD	Bastian	Graminor, N	1989
	AABBDD	Mirakel	Graminor, N	2012

1.3 MORPHOLOGICAL STRUCTURE OF THE WHEAT GRAIN

The wheat grain (kernel) is composed of three parts; germ, endosperm, and an outer layer consisting of a seed coat and pericarp. These parts including a detailed overview of the coat are shown in **Figure 2**. The **wheat germ** is the embryo, present at one end of the seed. It would develop into the wheat plant when planted in soil. It is rich in certain vitamins, proteins and

lipids. The **endosperm** is the white part within the centre of the wheat grain and it contains starch and aleurone cells, with some content of protein and iron. The Endosperm stores nutrition for the developing germ when the seed is planted and makes almost 80 % of the grains content. The endosperm gets separated from the bran layers during the milling process. The hardness of the grain is determined by the ease of separating the endosperm from the bran layers. The **outer layers** are further divided into seed coat and pericarp. The Pericarp is the outer most layer of fruit, while seed coat present inside of the pericarp consists of Testa and nucellar layer. The outer layer is nutritionally rich in protein and minerals. (Hoseney, 1994)

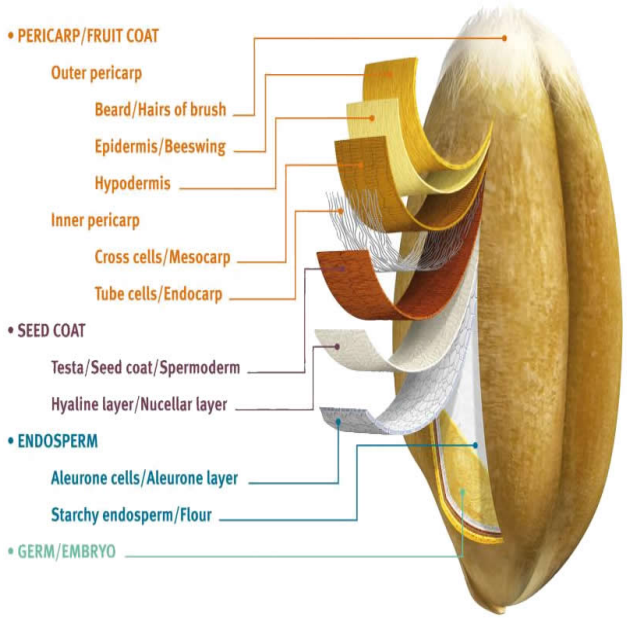


Figure 2: *The morphology of a wheat grain, cross section shows detailed description of the individual parts of wheat grain (Hoseney, 1994).*

1.4 COMPOSITION OF WHEAT GRAIN AND FLOUR

Wheat is mainly consumed as flour after milling the wheat grains. Wheat flour is composed of starch, proteins, oil, minerals, water and various vitamins. Some components are of great nutritious value for human beings e.g., proteins, starch, minerals, vitamins, and oils. Whereas protein and starch have great functional importance. The protein fraction plays a key role in bread making (Finney, 1948) by developing a strong gluten network that can retain air during baking. Starch is present in form of starch granules and has the ability to gelatinize when heated. Endosperm defines the hardness of a grain during milling which in turn affects absorption of water during dough formation (Williams, 2000). In the modern age, wheat is being used in a multitude of products and each product requires wheat flour with a specific

array of functional properties that are suitable for the product's processing and quality (Catterall and Cauvain, 2007). For optimum results during dough formation and baking, functional properties including kernel texture i.e, soft or hard wheat, starch content and gluten composition are important (Ross and Bettge, 2009).

1.5 WHEAT PROTEINS

Various animal and plant-derived food are used in the human diet as a source of protein. Protein is a key nutrient present in wheat along with starch. There can be variable amount of proteins in wheat. It varies between 8 and 16 % protein in most of the wheat varieties (Shewry et al., 2009). These cereal proteins are important from a nutritional point of view as well as for their functionality. Proteins are composed of amino acids linked together by peptide bonds, varies in molecular weight from a few thousand to several million. These peptide bonds are formed between the α-carboxyl and α-amino groups of neighbouring amino acids and form the backbone of the protein, also called the primary structure. Secondary structure is two dimensional and stabilized by the hydrogen bonds formed between the carbonyl oxygen and amide proton of the neighbouring peptide chains resulting in a-helix or b-sheets (Mathews et al., 2000). Sulphur containing cysteine residues on the peptide chains forms a disulphide bond (-S-S-) between two cysteine residues. Such linkages stabilize three-dimensional folding of the proteins providing tertiary or quaternary structure. These disulphide bonds can be made between two cysteine residues on the same protein chain (intramolecular bonds), making a loop, or they can be on different protein chains (intermolecular bonds), linking two polypeptide chains together resulting in tertiary globular structure of the wheat proteins (Wieser, 2007, Kasarda, 1999). During interaction with water, hydrophilic groups hydrogen-bond to water and/or to each other. Hydrophobic amino acid residues, tend to minimize their contact with water and hence associate together. (Hoseney, 1994). Disulphide bonds also play an important role in the quarternary structure of interacting polypeptides chains of different protein subunits (Shewry et al., 2003a).

1.5.1 CLASSIFICATION OF WHEAT PROTEINS

According to the work of **Thomas Burr Osborne** in the early 1900s, proteins have been classified into four types according to their solubility. **(1) Albumins** are soluble in water, and their solubility is not affected by reasonable (low) salt concentrations**. (2) Globulins** are

insoluble in pure water but soluble in dilute salt solutions and insoluble at high salt concentrations. **(3) Gliadins** are proteins soluble in aqueous 70% ethanol. **(4) Glutelins** are proteins soluble in dilute acids or alkali. The glutelin of wheat is named glutenin and that of barley is called hordein (**Figure 3**) (Osborne, 1907). The prolamin and glutelin are the storage proteins of the cereals and are collectively referred to as gluten, which is the fraction of interest in case of studying bread making and concerned with celiac disease (Inomata, 2009).

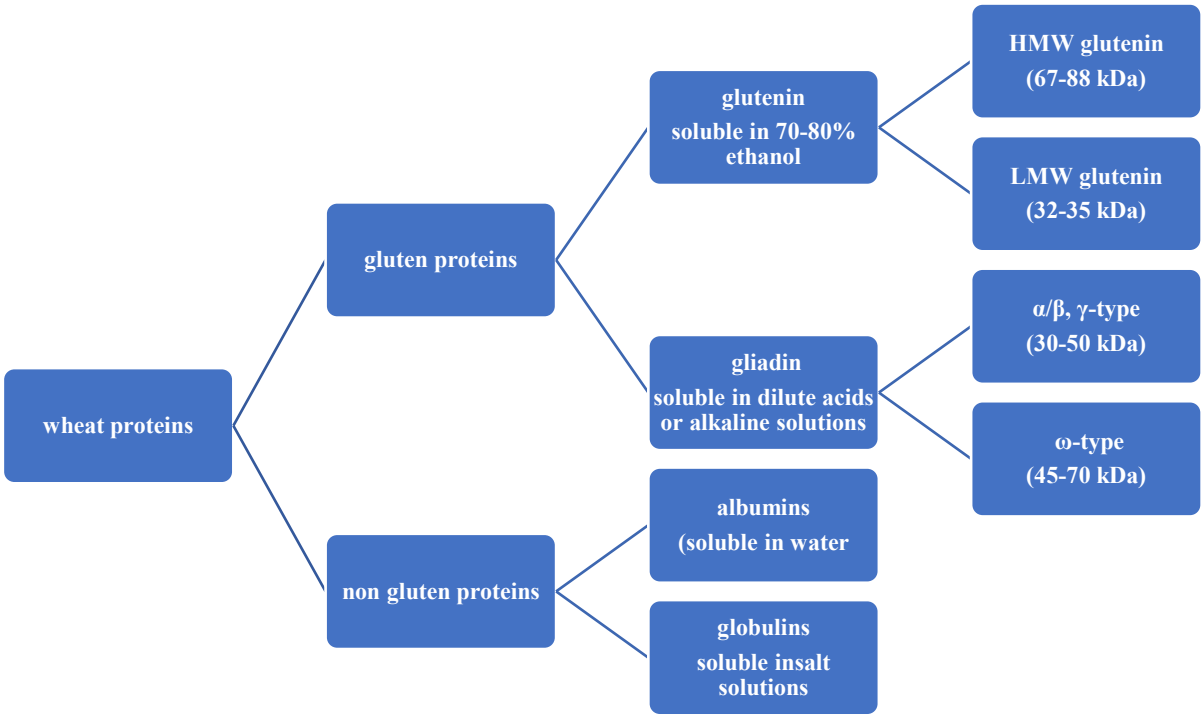


Figure 3: *The classification of wheat proteins based on their solubility based on Osborne extraction and gluten is further divided based on the molecular masses of the individual components adapted from Weiser et al (2007).*

1.5.2 NON-GLUTEN PROTEINS

Albumins and globulins are non-gluten proteins of wheat endosperm. It constitutes 20- 25% of total grain proteins (Belderok et al., 2000, Merlino et al., 2009). Albumins are soluble in water and globulins in salt solution. Albumins are mainly metabolic protein and perform an enzymatic activity in the breakdown of starch during germination, *e.g.* α-amylases, β-amylases, proteolytic enzymes, *etc*. (Matz, 1991). At the germination stage of the seedling, these enzymes make nutrients and provide energy by their hydrolytic and proteolytic activity (Stone and Savin, 1999).

1.5.3 GLUTEN PROTEINS

Gluten proteins are the main storage proteins in cereals. Gluten proteins can be divided into two main fractions according to their solubility in aqueous alcohols: the soluble gliadins and the insoluble glutenin. Gliadins are monomeric proteins with molecular weights (MWs) around 28,000 – 55,000 Da and can be classified according to their relative mobility in the gel electrophoresis depending on their mass to charge ratio (m/z) into the α/β-, γ- and ω-type. The glutenin fraction comprises aggregated proteins linked by interchain disulphide bonds with sizes varying from 500,000 Da to more than 10 million Da (Wieser, 2007). Glutenin subunits have been divided into the high-molecular-weight (HMW) subunits (MW 67,000 – 88,000 Da) and low-molecular-weight (LMW) subunits (MW 32,000 – 35,000 Da). (Wrigley and Bietz, 1988). These two protein fractions contribute to the viscoelastic properties of dough made with wheat flour and water. Gliadins contribute mainly to the viscosity and extensibility of the dough system, while the glutenin fraction is responsible for dough strength and elasticity. These two fractions play a key role in determining the unique baking quality of wheat by conferring water absorption, cohesivity, viscosity and elasticity in the dough (Payne, 1987) (Hoseney, 1994). A simplified illustration of the viscoelastic properties of gluten and gluten network development during dough formation by hydration is presented in **Figure 4**.

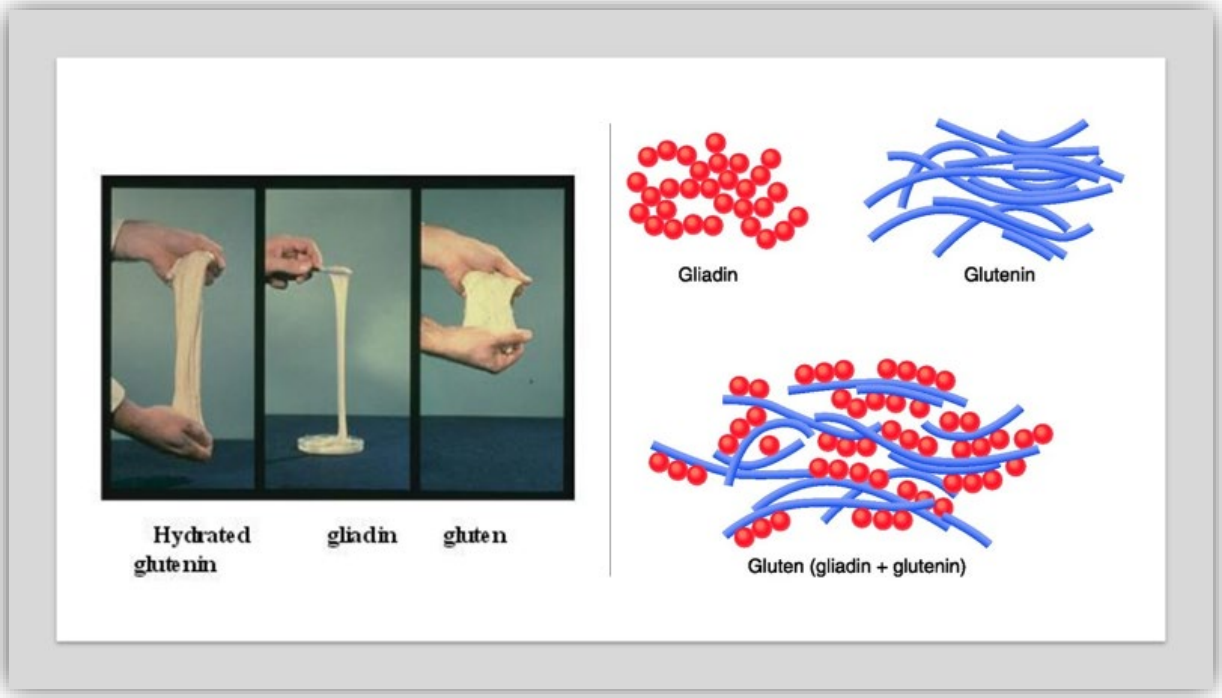


Figure 4: *The viscoelastic properties of gluten (Hoseney 1986) and arrangement of gliadin and glutenin during dough formation to form a viscoelastic dough.*

1.5.4 HMW-GLUTENIN SUBUNITS

Modern wheat is a hexaploid species with three genomes (called A, B and D) derived from related wild grass species as previously described. Single loci encoding HMW-glutenin subunits are present on the long arms of the group 1 chromosome (1A, 1B, 1D) and each locus comprising two genes encoding subunits, named *Glu-1-1* (x-type) and *Glu-1-2* (y-type). Although bread wheat could theoretically contain six HMW subunits (1Ax, 1Ay, 1Bx, 1By, 1Dx, 1Dy), the silencing of specific genes results in the presence of only three to five subunits, where 1Ay is always silent (Payne, 1987). A study of the genes coding HMW glutenin subunits from bread wheat and from related wild species have shown that the HMW glutenin subunits have conserved amino-acid sequences, comprising three distinct parts (**Figure 5**).

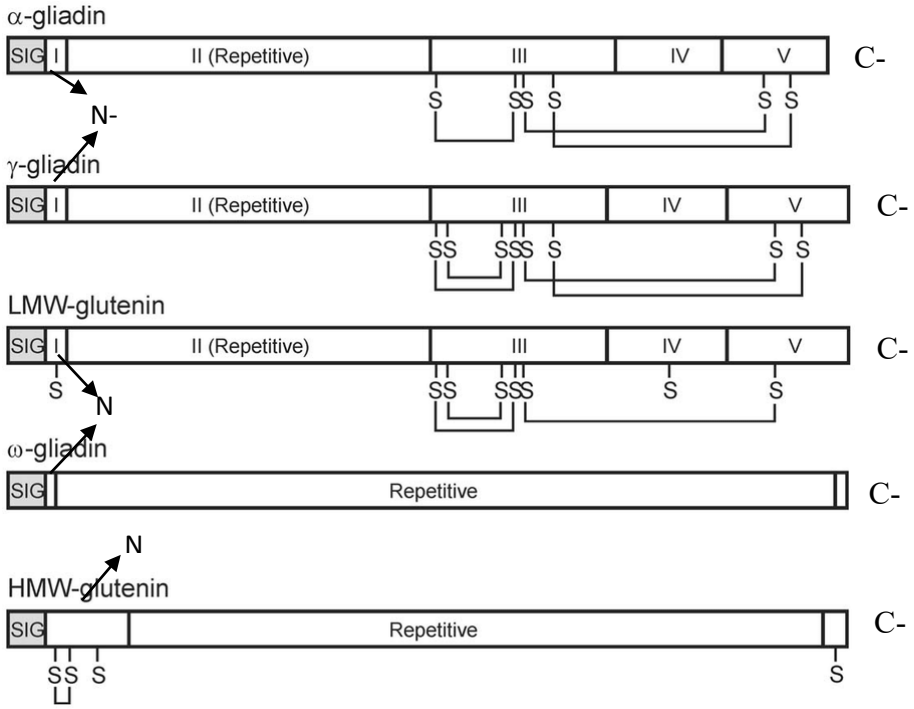


Figure 5: *The structure of HMW-, LMW-glutenin and α/β-, γ- and ω-gliadins subunits presenting different regions in each subunit. SIG represents the signal region, N represents N- terminal region, C- represents C-terminal region comprising III, IV, V subdivisions in α-, γ-gliadins and LMW glutenins whereas single region in HMW-glutenins and ω-gliadins and repetitive presents the region of repetitive motifs present in each subunit. Adapted from (Shewry et al., 2003a)*

It comprises the short non-repetitive N-terminal (of 81-89 residues in x-type subunits and 104 residues in y-type subunits) and the C-terminal (42 residues in all subunits). The central part consists of repeating peptide sequences that vary in length from about 272 to 872 residues, based on repeating hexapeptide and nonapeptide motifs in y-type. While in x-type, an

additional tripeptide is added (Shewry et al., 2003a). N- and C-terminals are abundant in cysteine residues as shown in **Table 2**.

Table 2: *Characterisation of gluten protein types adapted from Weiser (2000)*

Type	MWx10^{-3}	Proportions* (%)	Partial amino acid composition (%)				
			Gln	Pro	Phe	Tyr	Gly
ω5-Gliadins	49–55	3–6	56	20	9	1	1
ω1,2-Gliadins	39–44	4–7	44	26	8	1	1
α/β-Gliadins	28–35	28–33	37	16	4	3	2
γ-Gliadins	31–35	23–31	35	17	5	1	3
x-HMW-GS	83–88	4–9	37	13	0	6	19
y-HMW-GS	67–74	3–4	36	11	0	5	18
LMW-GS	32–39	19–25	38	13	4	1	3

*According to total gluten proteins.

The cysteine residues make both inter-and intramolecular disulphide bonds, which are the back bone of gluten proteins. Disulphide bonds play an important role in determining the structure and properties of gluten proteins (Shewry and Tatham, 1997). The study of repetitive domains indicate presence of β-reverse turns (Shewry et al., 2003b) and they also form a loose spiral which is responsible for gluten's elasticity as shown in **Figure 6**.

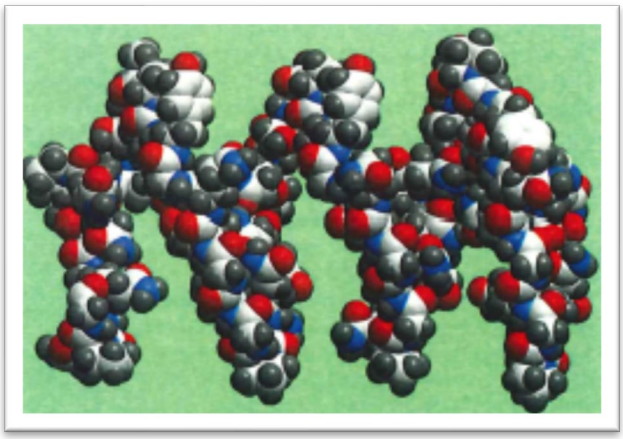


Figure 6: *β-Spiral model constructed for the consensus repeat peptides present in the HMW subunits of glutenin. Atoms are shown in white (carbon), blue (nitrogen), red (oxygen), and grey (hydrogen). (Parchment et al., 2001).*

1.5.5 LMW-GLUTENIN SUBUNITS

The LMW-glutenin subunits are more abundant than the HMW subunits, but they are more complex and difficult to study due to their similarity in their peptide sequences to some gliadin components. Unlike HMW glutenin, LMW subunits of glutenin are encoded by a clustered gene family comprises of 35-40 members which are located at the Glu-3 loci in hexaploid wheat (Sabelli and Shewry, 1991). This cluster of genes encoding LMW glutenin subunits makes it difficult to identify the proteins encoded by individual genes. The proteins encoded by these genes are comprised of four main regions: 1) a signal peptide of 20 amino acids; 2) a short N-terminal region of 13 amino acid residues, 3) a repetitive domain comprises of a single repeated motif of 5-8 residues makes about 70-186 AA. and 4) a C-terminal region of 180 AA-residues. The C-terminal can be subdivided into three parts based on its AA composition: a cysteine-rich region, a glutamine-rich region (**Table 2**), and a conserved C-terminal sequence as presented in **Figure 5** (Cassidy et al., 1998).

1.5.6 GLIADINS

ω- and γ-gliadins are controlled by clusters of tightly linked genes present at the Gli-1 loci on the short arms of the homologous group 1 chromosome (Gli-Al, Gli-Bl, Gli-Dl), whereas the α- and β-gliadins are encoded genes present at the Gli-2 loci (Gli-A2, Gli-B2, Gli-D2) present on the short arms chromosomes (Payne, 1987). Small structural differences are present among these types due to substitutions, deletions and insertions of single amino acid residues. Proteins encoded by these genes are comprised of four main parts; A Conserved signal peptide, a short N-terminal, a repetitive domain and C-terminal. The C-terminal can be subdivided into three parts based on its AA composition: a cysteine-rich region, a glutamine-rich region, and a conserved C-terminal sequence of 35-39 residues (**Figure 5**). C-terminal can vary within α/β-, ω-, and γ-gliadins and are specific for those proteins. α/β- and γ-gliadins are rich in cysteine and show the presence of three and four intrachain disulphide bonds respectively, whereas ω-gliadins are rich in glutamine, proline and phenylalanine but lack in cysteine (**Table 2**), so these are not capable of making disulphide crosslinks. α/β- and γ-gliadins have lower proportions of glutamine and proline than those of ω-gliadins (**Table 2**). α/β- and γ-gliadins differ significantly in the contents of tyrosine from the ω-gliadins. Studies on the secondary structure of gliadins and LMW-glutenins have shown similarities. The C-terminal domains of α/β-, γ-gliadins and LMW-glutenins exhibits α-helix and a globular structure which is stabilized by disulphide bonds (Shewry et al., 2009), while the repetitive domain exhibits non-globular

structure and exhibits considerable proportions of β-sheet structures (Cole et al., 1984, Shewry and Tatham, 1997).

1.5.7 FUNCTIONAL PROPERTIES OF GLUTEN

The wheat gluten proteins i.e., glutenins and gliadins are brought together to form a viscoelastic network shown in **Figure 4**, when flour is mixed with water to form a dough. Viscoelastic properties of wheat proteins have increased the span of its utilization for bread and other processed foods. HMW-subunits of glutenin, are particularly important in conferring high levels of elasticity (i.e. dough strength). Inter and intramolecular disulphide bonds formed at the beginning of protein synthesis, in the growing plant and ending in the baked products, are very important to understand the elasticity and stability of gluten network shown in **Figure 4**. These disulphide bonds are the known crosslink between LMW and HMW glutenins that form the 'elastic backbone' of gluten (Wieser, 2007). However, the glutamine-rich repetitive sequences that comprise the central parts of the HMW subunits also form extensive arrays of interchain hydrogen bonds that may contribute to the elastic properties via a 'loop and train' mechanism. Genetic variation in gluten quality is used in traditional breeding methods to manipulate the amount and composition of the HMW subunits to increase dough strength (Shewry et al., 2003b).

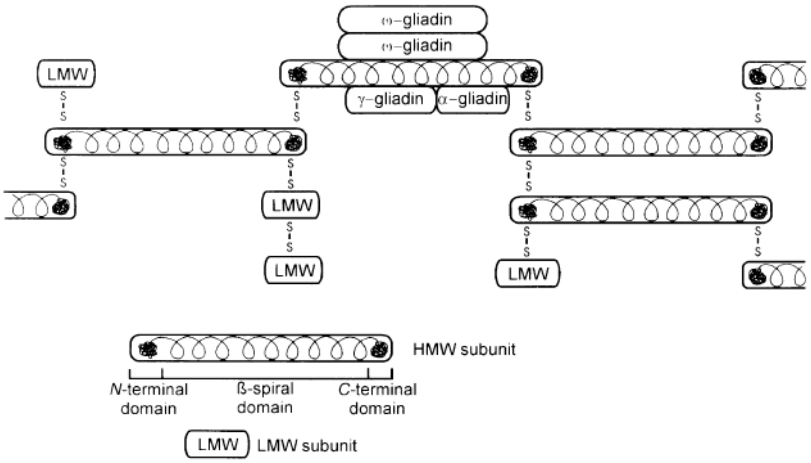


Figure 7: *A structural representation of disulphide bonds formed between LMW and HMW-glutenins and non-covalent bonds with gliadins that form backbone of elastic gluten network (Shewry et al., 2003a)*

1.6 CELIAC DISEASE (CD)

Wheat proteins are found to be triggering hypersensitivity reactions such as allergy (Inomata, 2009) or intolerance in susceptible individuals. Most allergic reactions belong to immediate type I responses caused by the production of IgE antibodies and are due to different environmental or genetic factors (Johansson et al., 2004). Among the different hypersensitivity reactions induced by wheat, celiac disease is prominent. **Celiac disease (CD)** is a non-IgE mediated enteropathy induced by ingestion of wheat gluten proteins and some other proteins similar in structure to gluten obtained from oat, rye and in rare cases from barley (Schalk et al., 2017). It is thought that 1-2 % of the total population in Europe and America is suffering from celiac disease (Lohi et al., 2007, Green, 2007). The inflammation caused by gluten protein can lead to small intestinal mucosal injury, villous atrophy and nutrient malabsorption. (Kagnoff, 2007). Several genetical and environmental factors are involved in CD pathogenesis. The presence of major histocompatibility complex (MCH) class II, HLA-DQ alleles are the main genetic factor associated with CD in humans (Sollid, 2002). As many as 99.5 % of the CD patients express the HLA-DQ alleles that encodes the heterodimer HLA-DQ2.5 on antigen presenting cells (APCs) in the lamina propria of the small intestine as shown in **Figure 8**

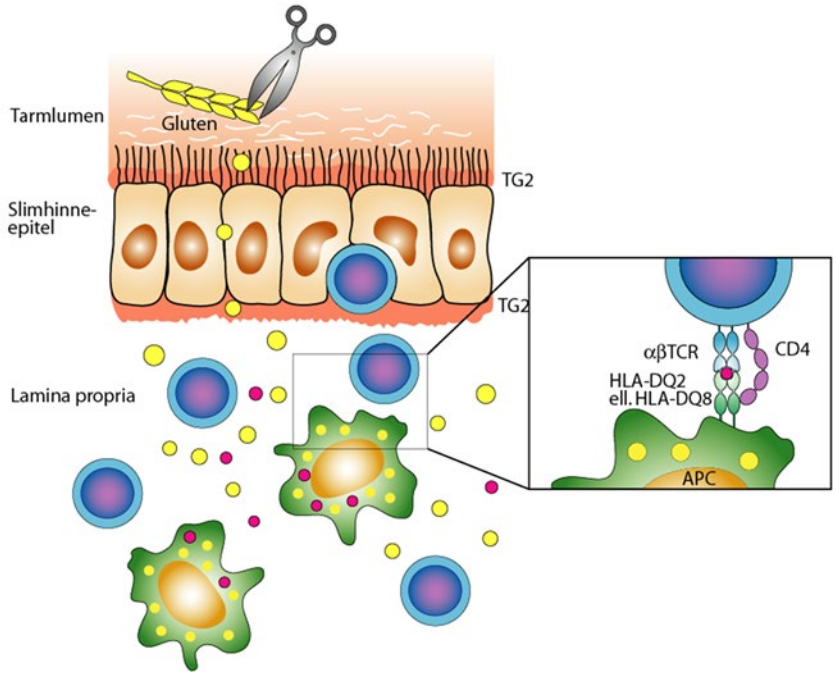


Figure 8: *HLA-DQ molecules located on the antigen presenting cells (APCs) presents gluten peptides for CD4+ T cells (Lea, 2018).*

The alleles may be present in cis- or trans-configuration depending on homo/heterozygosity of the individuals, carrying these haplotypes (Sollid et al., 1989). Most of the remaining patients

show association with HLA-DQ8 (Lundin et al., 1994), while a few show association with HLA-DQ2.2 or DQ7.5 (Karell et al., 2003) (**Table 3**). Both HLA-DQ2.5 or HLA-DQ8 are important for celiac disease development as gluten-reactive T-cells recognize peptides presented on these HLA molecules (Sollid, 2002). It is also important to note that HLA is not the only genetic factor in CD. 39 other loci have also been identified to be associated with disease development (Trynka et al., 2011).

Table 3: *Description and nomenclature of HLA-DQ molecules that are associated with CD and with antigen-presenting molecules for CD4+ T cells of celiac disease patients (adapted from Sollid et al, 2012).*

HLA-DQ molecule	**Encoded by**		**The risk for celiac disease**	**Expression in *cis* or *trans* position**	**Haplotype**
	DQA1*	**DQB1***			
HLA-DQ2.5	05	02	high	*cis, trans*	DR3DQ2
HLA-DQ2.2	02	02	Low	*cis, (trans)*	DR7DQ2
HLA-DQ7.5	05	03:01	Very low	*cis, (trans)*	DR5DQ7
HLA-DQ8	0.3	03:02	Low	*cis*	DR4DQ8

CD4+ T cells of CD patients (Molberg et al., 1997) recognize gluten peptides when presented by disease-associated HLA-DQ molecules as shown in **Figure 8**. The risk for disease is linked with the ability of HLA-DQ molecules to bind the gluten reactive T-cell epitopes. The selection of gluten reactive T-cell epitopes is influenced by at least three factors: (a) resistance of the polypeptide sequence to proteolytic degradation, (b) specificity of TG2 and (c) HLA binding specificity.

Table 4: *List of celiac disease relevant T-cell epitopes recognized by CD4+ T cells (Sollid et al, 2012)*

Previous name*	Recent name	Protein binding register **
		1 2 3 4 5 6 7 8 9
DQ2.5 binding epitopes		
DQ2.5-glia-α1a	DQ2-α-I, α9	P F P Q P **E** L P Y
DQ2.5-glia-α1b	DQ2-α-III	P Y P Q P **E** L P Y
DQ2.5-glia-α2	DQ2-α-II, α2	P Q P **E** L P Y P Q
DQ2.5-glia-α3	glia-α20	F R P **E** Q P Y P Q
DQ2.5-glia-γ1	DQ2-γ-I	P Q Q S F P **E** Q Q
DQ2.5-glia-γ2	DQ2-γ-II, γ30	I Q P **E** Q P A Q L
DQ2.5-glia-γ3	DQ2-γ-III	Q Q P **E** Q P Y P Q
DQ2.5-glia-γ4a	DQ2-γ-IV	S Q P **E** Q **E** F P Q
DQ2.5-glia-γ4b	DQ2-γ-VIIc	P Q P **E** Q E F P Q
DQ2.5-glia-γ4c	DQ2-γ-VIIa	Q Q P **E** Q P F P Q
DQ2.5-glia-γ4d	DQ2-γ-VIIb	P Q P **E** Q P F C Q
DQ2.5-glia-γ5	DQ2-γ-VI	Q Q P F P **E** Q P Q
DQ2.5-glia-ω1	DQ2-ω-I	P F P Q P **E** Q P F
DQ2.5-glia-ω2	DQ2-ω-II	P Q P **E** Q P F P W
DQ2.5-glut-L1	glutenin-17	P F S **E** Q **E** Q P V
DQ2.5-glut-L2	glutenin-156	F S Q Q Q **E** S P F
DQ2.5-hor-1	Hor-α9, Hα9	P F P Q P **E** Q P F
DQ2.5-hor-2	Hor-α2 Hα2	P Q P **E** Q P F P Q
DQ2.5-hor-3	hor-I-DQ2	P I P **E** Q P Q P Y
DQ2.5-sec-1	Sec-α9, Sα9	P F P Q P **E** Q P F
DQ2.5-sec-2	Sec-α2, Sα2	P Q P **E** Q P F P Q
DQ2.5-ave-1a	Av-α9A	P Y P E Q **E** E P F
DQ2.5-ave-1b	Av-α9B, 1490	P Y P E Q **E** Q P F
DQ2.2-glut-L1	glutenin-17	P F S **E** Q **E** Q P V
DQ8 binding epitopes		
DQ8-glia-α1	DQ8-α-I	**E** G S F Q P S Q **E**
DQ8-glia-γ1a	DQ8-γ-Ia	**E** Q P Q Q P F P Q
DQ8-glia-γ1b	DQ8-γ-Ib	**E** Q P Q Q P Y P **E**
DQ8-glut-H1	HMW-glutenin	Q G Y Y P T S P Q
DQ8.5 binding epitopes		
DQ8.5-glia-α1	DQ8-α-I	E G S F Q P S Q **E**
DQ8.5-glia-γ1		P Q Q S F P **E** Q **E**
DQ8.5-glut-H1	HMW-glutenin	Q G Y Y P T S P Q

*In the epitope names, these short terms are used to denote the type of proteins that the epitopes derive from: glia-α, glia-γ, glia-ω, glut-L, glut-H, hor, sec, ave denote α-gliadin, γ-gliadin, ω-gliadin, low molecular weight glutenin, high molecular weight glutenin, hordein, secalin and avenin respectively.

**Glutamate residues (E) formed by tTG2-mediated deamidation which are important for recognition by T cells are shown in bold. Additional glutamine residues (Q) also targeted by tTG2 are underlined.

Wheat (gluten) and products made with wheat is one of the environmental factor that cause an autoimmune reaction in the CD patients. Gluten proteins have specific amino acid sequence motifs, rich in glutamine and proline residues. Proline and glutamine-rich gluten proteins resist proteolytic degradation in the gastrointestinal lumen, resulting in long peptide fragments ranging from 15 to 50 residues (Shan et al., 2002). Class II-restricted T-cell receptors are sensitive to a minimum of nine amino acid core region as shown in **Table 4**. However, peptides surviving digestion can be much longer than these peptides (Sollid et al., 2012). An example is the 33mer (LQLQPFPQPQLPYPQPLPYPQPQLPYPQPQPF) from α-gliadin, that has been considered as one of the main culprits in celiac disease (Stepniak et al., 2005, Shan et al., 2002). This peptide fragment contains several overlapping T cell epitopes. Other long gluten peptides have also been recognized as being reactive to intestinal T cells of CD patients (Arentz-Hansen et al., 2000, Sjöström et al., 1998). A list of epitopes and their respective gluten peptides that have been identified by CD4+ T cells, is shown in **Table 4**. This includes peptide sequences that mainly derive from glutenins and gliadins in wheat. However, other epitopes with peptide sequences of hordeins, secalins and avenins from barley, rye and oats, respectively are also included in this list. After deamidation of glutamine (Q) to glutamate (E) by the enzyme tissue transglutaminase 2 (tTG2), these peptides bind effectively with the HLA-DQ molecules so the HLA-DQ2 and DQ8 restricted T cells of CD patients can react with them (Molberg et al., 1998); (van de Wal et al., 1998). Deamidation of glutamine to glutamate is described in **Figure 9**.

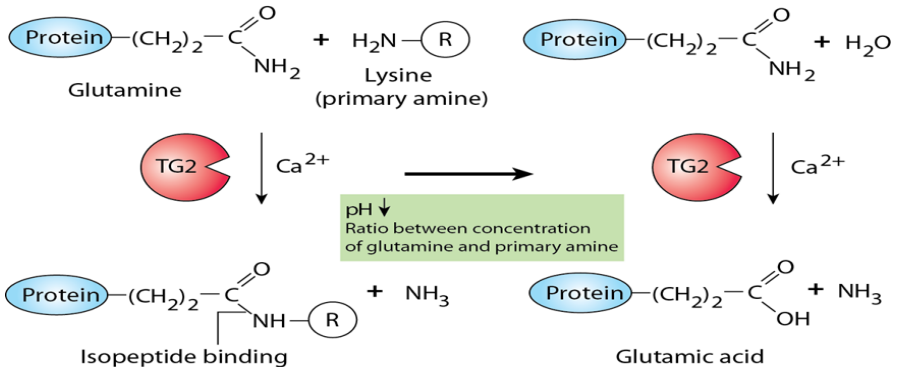


Figure 9: *Deamidation of glutamine to glutamate (right part) by transglutaminase 2 (tTG2) (Lea, 2018).*

Tissue transglutaminase (tTG2) recognizes glutamine residues in glutamine-X-proline sequences present in a large peptide fragment. The glutamate residues serve as anchor sites for binding of the peptides to HLA-DQ molecules. In HLA-DQ2.5 restricted epitopes, the glutamates introduced by TG2 are present at the positions P4, P6 and P7 and prolines at position P1, P6 and P8 of a nine-amino acid sequence peptide. HLA-DQ8 restricted epitopes alternatively presents glutamates at position P1, P4 and P9 while no specific sites for prolines (**Figure 10**) (Sollid et al., 2012). Deamidated gluten peptides bind with higher affinity to HLA-DQ, and a longer gluten peptide sequence forms more stable complexes with HLA-DQ (Fallang et al., 2009).

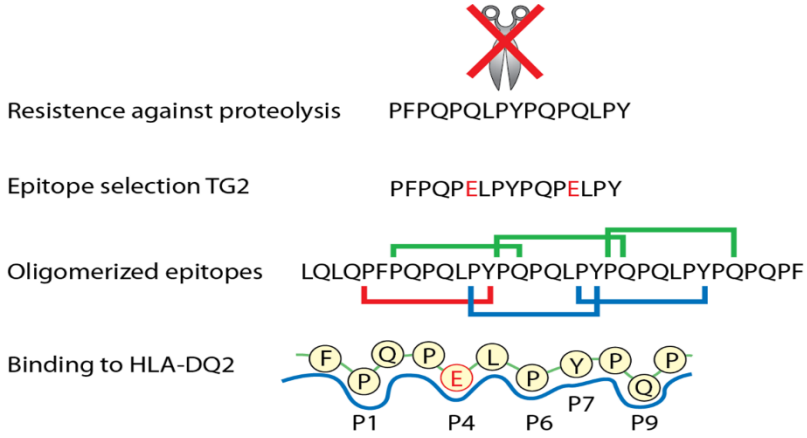


Figure 10: *Selective binding sites of amino acids to HLA-DQ to form a stable complex (Lea, 2018).*

1.6.1 MECHANISM OF CELIAC DISEASE PATHOGENESIS

In the pathogenesis of CD, activation of pro inflammatory CD4+ T cells plays a major role which subsequently lead to tissue damage. The native and tTG2-deamidated gluten peptides are presented to T-cell receptors after binding to the HLA-DQ molecules present on the surface of antigen presenting cells (APCs) in lamina propria of the intestinal wall. Deamidation is crucial in activation of T-cells as it converts the less immunogenic native gluten peptides to highly immunogenic antigens for CD4+ T cells. Deamidation increases the gluten peptide's binding affinity to disease-associated HLA dimers (Tye-Din et al., 2018). These activated T-cells are characterized by production of different cytokines such as interleukin (IL-15) and interferon (IFN-γ) (**Figure 11**). These T-cell cytokines increase the cytotoxicity of intraepithelial lymphocytes (IEL) and production of natural killer cells (NK cells) which leads to villous atrophy (Kagnoff, 2007). The T-cell cytokines stimulates the B-cell differentiation into plasma cells producing specific anti-gliadin and anti-tissue transglutaminase antibodies, which can be used in diagnostics of CD (Rescigno and Di Sabatino, 2009).

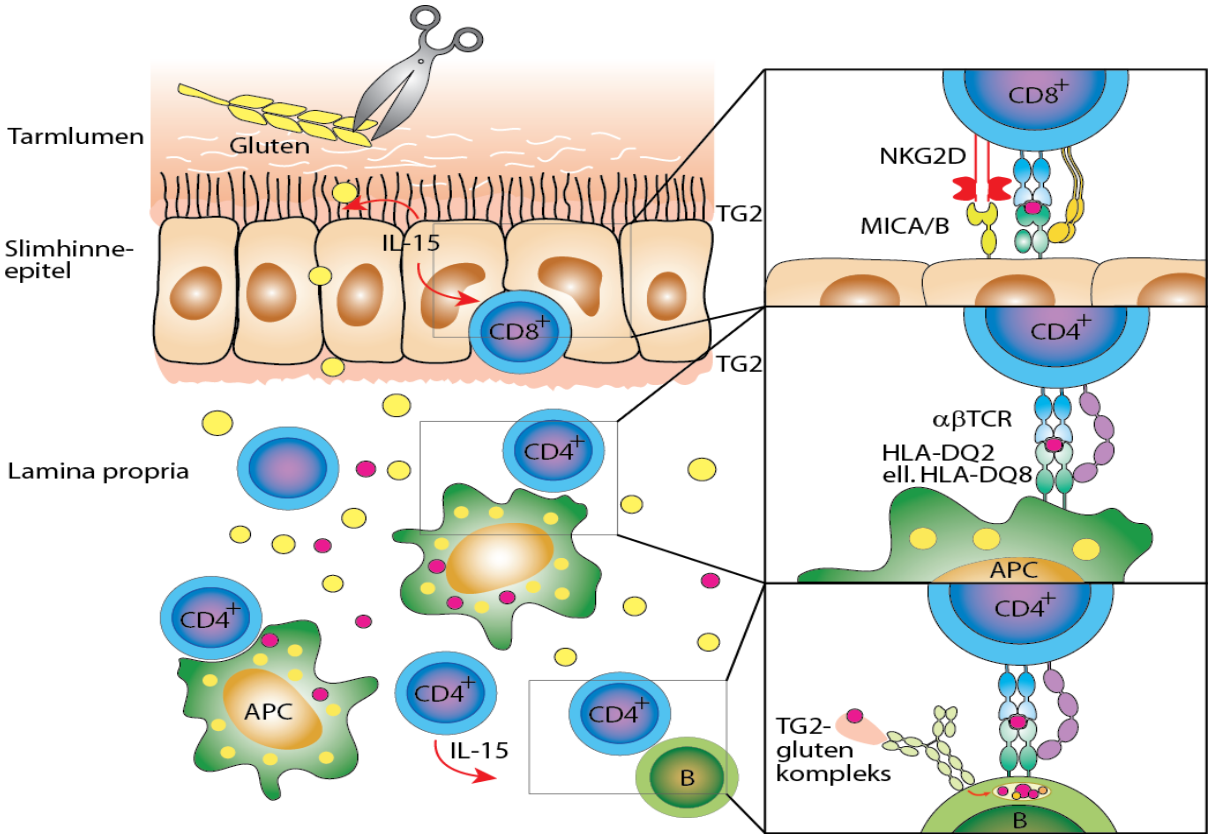


Figure 11: *Pathogenesis of CD: Gluten peptides cross the intestinal epithelium by different pathways. Tissue transglutaminase (tTG) deamidates gluten peptides which are then recognized by HLA-DQ2 or -DQ8 molecules of antigen presenting cell (APC). APC presents the toxic peptides to CD4+ T cells. Activated CD4+ T-cells produce high levels of pro-inflammatory cytokines. Some cytokines such as interleukin 15 (IL-15) induce the formation of natural killer cells (NK) resulting in epithelial damage (Lea, 2018).*

1.7 DIGESTION OF PROTEINS

The human body is dependent on proteins to function. The protein's role is vital for biochemical reactions, the immune system, metabolism and structure and strength of the cells (Mathews et al., 2000). Peptide sequences encrypted within the intact protein may exert physiological functions in the human body after enzymatic release during digestion (Foltz et al., 2008). It is important to keep in mind that protein digestion is not as simple as eating a slice of bread and magically getting the required amino acids. In the process of protein digestion, the protein molecules are broken down to peptide fragments and amino acids under physiological conditions, mainly in the stomach and in the intestine, then the degraded products are transported and absorbed from the gut through the brush border membrane to the blood. The human digestive system, called the gastrointestinal tract (GIT) is comprised of several organs, and extends from mouth to anus, illustrated in **Figure 12**. The GIT can be divided into

two tracts; upper and lower tract. The upper tract consists of oral cavity, oesophagus, stomach and duodenum. The lower tract consists of most of the small intestine i.e., jejunum and ileum and all parts of the large intestine., cecum, colon, rectum and anal canal.

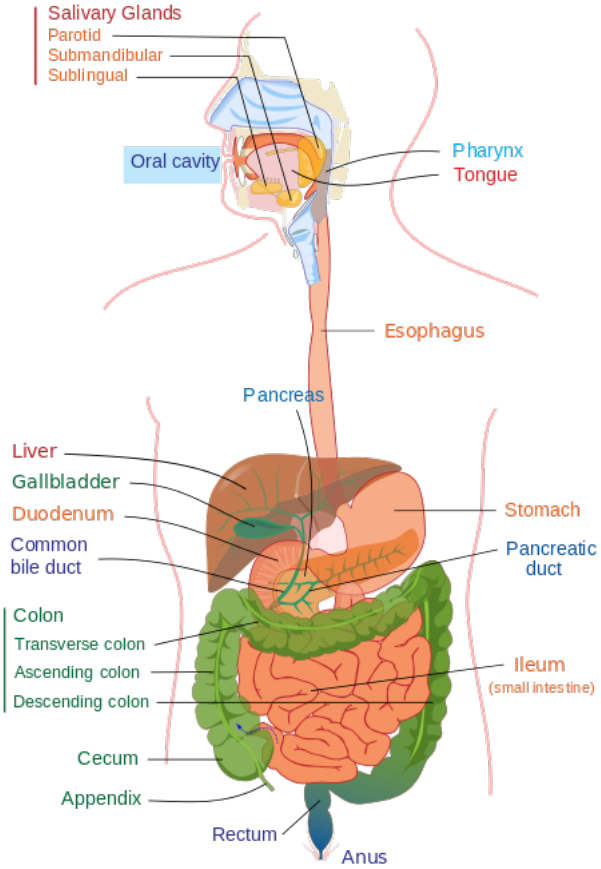


Figure 12: *The human digestive system (Verhoeckx et al., 2015). Copyrights to Wikipedia*

Digestion of food starts in the oral cavity where the food is masticated and secretion of saliva increases (Verhoeckx et al., 2015). The saliva contains amylase, which starts the digestion of starch. The mastication of food forms a bolus that is transported to the stomach via the oesophagus. G-cells of the stomach lining release mainly pepsin and hydrochloric acid which lowers the pH in the stomach to pH 1-2. When food enters the stomach pH rises to 3, this acid environment provides optimal pH for the gastric enzymes to digest the food (Saladin, 2013). In the stomach, pepsinogen is converted into active pepsin due to acidic pH. The digestion of protein starts in the stomach, where pepsin breaks some specific bonds of the proteins, preferably at sites of aromatic and hydrophobic amino acids, such as Tyr (Y), Phe (F) and Leu (L) (Rawlings et al., 2017). Pepsin exhibits high activity at pH 2.0 and is almost inactive at pH 6.5, which is attained in the duodenum. When the semi-digested food reaches the duodenum, a

more intensive hydrolysis of the proteins occurs. The intestinal enzymes are produced in the pancreas and are released together with bile salts from the gallbladder into the small intestine. The proteolytic enzymes trypsin, chymotrypsin, elastase, together with lipases, amylases are secreted from the pancreas. Trypsin cleaves the proteins C-terminally to Lys (K) and Arg (R), chymotrypsin cleaves C-terminally to aromatic and bulky amino acids, and elastase cleaves C-terminally to Ala (A), Valine (V) and to some extent Leucine (L) (Rawlings et al., 2017). Pancreatic juice and bile salts increase pH in the duodenum from 2 to 5-7.5. It inactivates the gastric enzymes and provides the optimal pH for the duodenal enzymes to work. The final stage of digestion of proteins occurs on the surface of intestinal enterocytes, by brush border membrane (BBM) enzymes, where peptides are further hydrolysed to mono-, di- and tripeptides (Verhoeckx et al., 2015). The nutrients are now ready to be absorbed by the enterocytes of the jejunum and ileum called villi and can be further degraded by intracellular proteases before entering the bloodstream. Absorption of water and minerals occurs in the large intestine. Fermentation of complex nutrients i.e., non-degradable polysaccharides and fibre also occur in the large intestine. Human digestive enzymes are not able to digest these complex molecules, these are then fermented by microbiota in the large intestine. Undigested food is excreted out of the body through the anal opening.

1.8 *EX-VIVO* DIGESTION MODEL

In-vitro digestion (IVD) studies are widely used to predict the behaviour of food components in the digestive tract. To mimic and simplify the human digestive system, a simulated model of digestion is used. *In-vitro* simulated digestion imitates the human digestive system by using commercial enzymes at some fixed parameters e.g., pH and temperature and enzyme quantity. It has its advantages for being rapid, easily workable and efficient, however, it cannot address the diversity of human digestive system. The human digestive system is very complex comprising different enzymes, salts and phospholipids etc. that work together to provide the human body with optimal nutrition and energy. Gastric and duodenal enzymes from the commercial origin can be prepared in the lab or isolated from natural sources e.g., porcine and bovine enzymes. When industrially prepared single enzyme is used, digestion result is different from the complex mixture of enzymes of the natural origin. In the simulated model, the substrate is incubated for a specific time with commercial or human oral, gastric and small intestinal digestive enzymes at a fixed temperature and pH. The COST Action INFOGEST protocol (Minekus et al., 2014) has standardized an international method of ex vivo digestion to get better and reproducible results. Inter-laboratory studies have optimized the IVD

INFOGEST method, and parameters concerning protein digestion have been compared (Egger et al., 2016). The static digestion model with fixed parameters, such as time and pH, are the disadvantages of such an *in vitro* method as compared to *in vivo* digestion studies. In this study, human enzymes have been used to mimic *in vivo* digestion in humans, called *ex vivo* digestion model. This model gives digestion results close to human digested system, as compared to commercial mix of enzymes from animal origin or a single enzyme.

1.9 PROTEIN DEGRADATION PROFILE

For post digestion analysis and identification of undigested proteins and peptides produced during digestion, different techniques are being used, based on structural and functional properties of these proteins and peptides. Some of these techniques are discussed below.

1.9.1 ELECTROPHORESIS

SDS-PAGE Gel electrophoresis is a method used in clinical chemistry, biochemistry and molecular biology for separation of macromolecules (DNA, RNA and proteins) and their fragments, based on their size and charge in an applied electric field. When an anionic detergent (sodium dodecyl sulphate SDS) is added to a protein sample, proteins and peptide fragments gets a net negative charge on the peptide chain in an amount proportional to its mass. Protein molecules become negatively charged and repel each other which breaks down the quaternary, non-disulphide-linked tertiary structure as well as the secondary structure of proteins. Dithiothreitol (DTT) is a reducing agent and is added to the protein samples for a complete breakdown of disulphide bonds. When an electric field is applied, these proteins/ peptides will then move towards their respective electrodes with speed depending on their mass to charge ratio (Mathews et al. 2013). The different sized molecules form distinct bands on the gel. These bands are further used to determine the molecular mass of the individual proteins by comparing it with the standard molecular weight ladder. These gels are scanned with the help of a scanner and a software is used to convert the protein band intensity into colour intensity for comparison of different proteins present in gel. Different stain-free and staining gels are being used depends on the proteins being analysed.

1.9.2 THE DEGREE OF HYDROLYSIS (DH)

During gastrointestinal digestion, the proteins are hydrolysed by gastric and pancreatic enzymes to individual peptides and thereby liberating their fri-amino groups. The degree of hydrolysis can give information about the kinetics of protein digestion and how rapid proteins are degraded into peptides. The degree of hydrolysis (DH) is proportional to peptide bonds broken in a protein, and is calculated as follows:

$$DH\ \% = \frac{h}{htot} x100\ \%$$

Where h is the amount of hydrolysed peptide bonds and h_{tot} is the total amount of peptide bonds present (Rutherfurd 2010). O-phthaldialdehyde (OPA) method is based on the determination of colour changes spectrophotometrically as a result of peptide bonds hydrolysed during digestion and absorbance is measured at 340nm. The determination of DH was done by the spectrophotometric assay described by Church et al., (1983). O-phathaldialdehyde (OPA) binds to the primary amino acids released after hydrolysis of the peptide bond as well as to the amino groups present on the side chains of peptides in the presence of β-mercaptoethanol. This reaction results in the formation of coloured complex; **1-alkylthio-2-alkylisoindole** which is expressed in **Figure 13**. Colour intensity varies with the number of peptide ponds hydrolysed and these colour changes are observed spectrophotometrically at 340 nm.

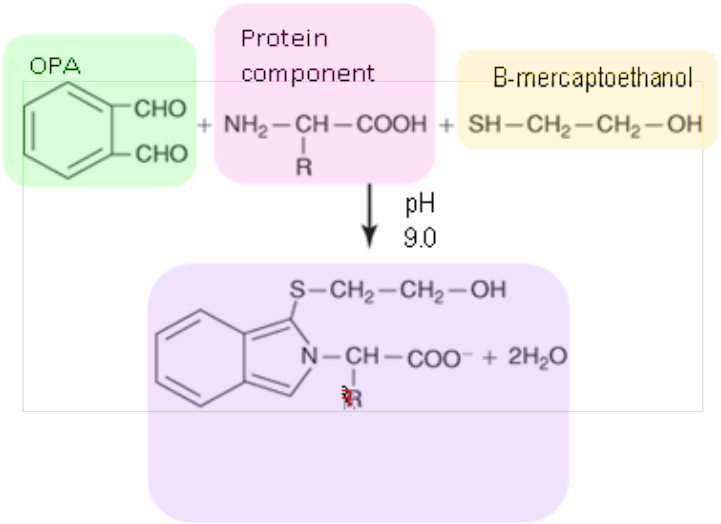


Figure 13: *The coloured complex formation by the reaction of OPA and β-mercaptoethanol with primary amino acids.*

1.10 PEPTIDE PROFILING

Several immunological and non-immunological methods are being used in the present-day proteomic analysis of immunogens present in the wheat. Immunological methods involve an antigen/antibody reaction to identify the immunogenic peptides present in wheat. Mass spectrometry coupled with liquid chromatography is the available nonimmunological method for proteomic analysis of wheat.

1.10.1 LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY

Liquid chromatography-mass spectrometry (LC-MS) combines the two individual techniques. First by separating a sample into its individual components by liquid chromatography and then analysing these components by mass spectrometry. In this technique, proteins/peptides are separated according to their hydrophobicity, these separated peptides move through a vacuum chamber following a specific path based on their mass to charge ratio (m/z). In liquid chromatography, different stationary and mobile phases are used depending on the sample molecules affinity towards mobile or stationary phase. If the analyte has a higher affinity to the stationary phase, it will bind to the column and would move slower than compounds with higher affinity to mobile phase. A common method for separation of peptides is reversed-phase high-performance LC (RP-HPLC). Reversed-phase chromatography is based on a hydrophilic mobile phase and a hydrophobic stationary phase. Proteins being hydrophobic have a higher affinity to the stationary phase and are eluted through the column by the degree of their hydrophobicity. Now, these liquid molecules are transferred into the gas phase and ionised by passing through an ionization source before entering MS. Formation of gas-phase ions is achieved by different means i.e., laser, electron beam, UV light. Electrospray Ionization (ESI) and matrix-assisted laser desorption ionisation (MALDI) are frequently used ionization techniques in proteomics. In electrospray ionization chamber, the gaseous-phase ions are formed from proteins and peptides by spraying a dilute solution of the analyte through the tip of a fine stainless-steel capillary (Aebersold and Goodlett, 2001). Ionized samples are then fragmented and separated according to their m/z by application of electric and magnetic field i.e., by quadrupole or time of flight (TOF). Separated ions are then detected in the detector and their masses are recorded and presented as a mass spectrum. **Tandem mass spectrometry**, also known as **MS/MS** or **MS^2**, combines multiple steps of mass spectrometry, where fragmentation of the selected ions occurs in the later stage (De Hoffmann, 2000). In a tandem mass spectrometer, selection of separated ions in the first stage (precursor ions) of mass spectrometry (MS1) are fragmented (product ions) by using different fragmentation techniques

e.g., collision-induced dissociation, ion-molecule reaction, photodissociation, or another process (De Hoffmann, 2000). The resulting ions are then separated and detected in the second stage of mass spectrometry (MS2). A schematic diagram (**Figure 14**) showing the separation and fragmentation of ions in tandem mass spectrometry. In MS/MS all the peptides are fragmented and peptides with specific m/z are further proceeded to determine the sequence of the peptide. ESI-MS and MALDI-TOF-MS have become an important tool for gluten proteins characterization and providing a basis to classify wheat varieties at the molecular level. As well as for the identification of allergens and for trace detection of proteins in complex food matrices. (Mamone et al., 2011)

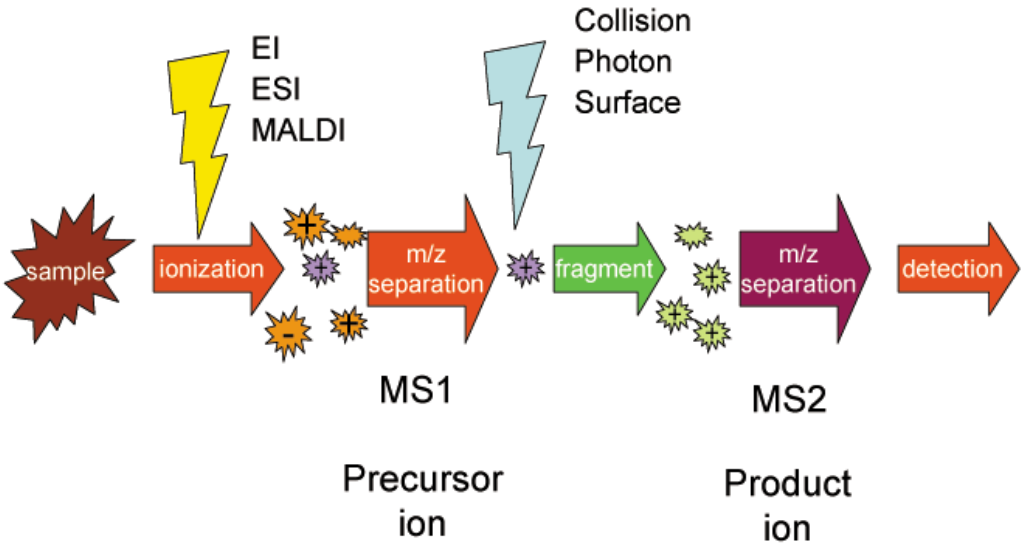


Figure 14: *Schematic representation of Tandem mass spectrometry(MS/MS). Peptides are first separated according to m/z and then fragmented to determine the peptides sequences. (Wikipedia)*

1.11 AIM OF THE STUDY

The aim of the study was

- To detect the differences in gluten protein profile of wheat types belonging to different ancient wheat types i.e., emmer, einkorn, spelt, and some common wheat varieties; Fram, Børsum, Bastian, Mirakel and their digestion with *ex vivo* model digestion with human gastrointestinal (GI) juices.
- To determine the difference in celiac disease related immunogenic peptides profile in different wheat types after digestion with normal/healthy human gastrointestinal (GI) juices.

2 MATERIALS AND METHODS

2.1 MATERIALS

The wheat material used in this study were collected from field trials grown in 2017 at Vollebekk Research Farm, Department of Plant Sciences, NMBU, Ås. They were all spring types, consisting of both diploid, tetraploid and hexaploid wheat species, as well as varieties of common wheat released in different time periods. These will hereafter be referred to in the text as "wheat types". A brief description of these wheat types is given below.

Einkorn (*Triticum. monococcum* subsp. *monococcom*). The accession originated from Gotland. Einkorn is a diploid specie, having the A genome designated AA. The plants have short straw, small spikes and small hulled grain with tough glumes that tightly enclose the grain. The grain remains intact in spikelet when threshed.

Emmer (*Triticum. turgidum* subsp. *dicoccom*) 'Gotland'. Emmer is a tetraploid species having both A and B genome, designated AABB. The plants have relatively tall straw with hulled grain. The grain remains intact in spikelet when threshed.

Spelt (*Triticum spelta*). 'Gotland'. Spelt is a hexaploid wheat specie having A-, B-, and D genome designated AABBDD. Spelt is an ancient specie which is a close relative to the common wheat. Spelt kernel has a large hulled grain, where kernels are intact in spikelet when threshed.

Common wheat (*Triticum aestivum subsp.aestivum*)

Old Wheat varieties

- Børsum
- Fram

Børsum was an old Norwegian landrace with origin from Ås, Akershus, in the periods around 1900. However, Fram was developed by crossbreeding with an American wheat variety in 1936. These varieties have plants with long straw, relatively small ears and naked seed. These varieties are susceptible to diseases and have weak gluten quality.

Modern wheat varieties

- Bastian
- Mirakel (GN06600)

Bastian was bred in Norway and released in 1989. Whereas Mirakel is a new variety, released in 2012, and is also bred in Norway. These varieties have plants with relatively longer straw

and naked seed. They show good resistance to diseases, in particular Mirakel. Both varieties have strong gluten and were classified as class 1 varieties according to gluten quality for baking in the Norwegian system.

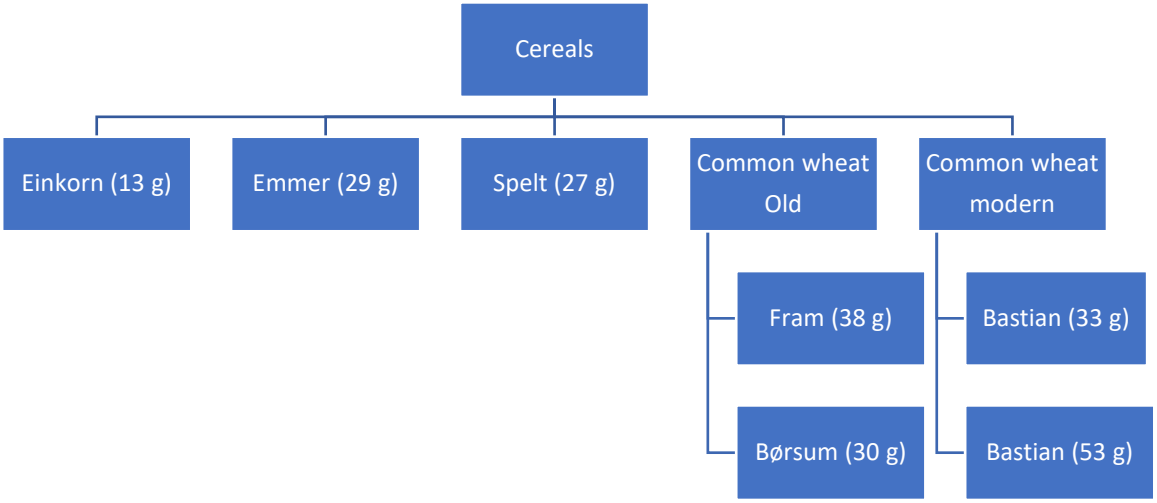


Figure 15: *An overview of the samples and amount for each sample in grams collected for analysis.*

2.2 SAMPLES COLLECTION

All species and varieties were grown in the same field trial in plots of 1.5m*3m. At maturity, 50 ears for each sample were harvested from each plot. For Einkorn, being small in ear and grain size, the double number of ears was collected. The harvested samples were dried at 30^0C for 3 days before threshing at Vollebekk Research Farm. Kernels were first threshed and cleaned by using threshing and cleaning machines (Perten Instruments AB, Sweden). Einkorn, emmer and spelt were dehulled manually after threshing and the weight of the samples after threshing and dehulling is presented in **Figure 15**. Then the samples were milled to whole meal flour by Falling Number Laboratory.

2.3 CHARACTERIZATION OF THE WHEAT TYPES

Following preliminary tests were performed to characterize the wheat types and determine the quality of grains and their flour used for digestion.

2.3.1 THOUSAND KERNELS WEIGHT

Weight per thousand kernels is used as a measure of quality of cereal crop. Seeds counted by Elmor C1 seed counter (Elmor Ltd., Switzerland, presented as weight in gram per thousand grains.

2.3.2 % MOISTURE

Water content of the grain was measured by weighing the kernel before and after drying in the oven for 24 hrs at 105ºC and then reduction in water was converted into % moisture.

2.3.3 FALLING NUMBER

Falling number is an indication of the damage that might occurred to starch due to late harvesting season. In late harvesting season alpha amylase activity starts due to preharvest germination in grains exposed to rain. Falling number was performed by Falling Number 1800 (Perten Instruments AB, Sweden) according to standard procedure (ICC) by using the flour/water mixture and then measuring viscosity of the flour/water mixture as resistance to the stirrer in terms of time (seconds) required for the stirrer to reach the bottom of glass tube.

2.3.4 PROTEIN CONTENT

Kjeldahl method was used to measure protein content. Kjeldahl method measures the amount of nitrogen present in the wheat flour. The amount of nitrogen is further used to calculate the protein content by using a numerical factor called Kjeldahl factor and Kjeldahl factor used for wheat is 5.7 (Eijsnik, 2017).

PROCEDURE

500mg of each sample was measured in glass tubes, in three parallels. 3ml concentrated sulphuric acid and a catalyst (Kjeldahl tablet) was added to all the tubes and the samples were heated to 420ºC for 60 minutes on a heating block. Sulphuric acid reacts with the N present in the organic sample and convert it into Ammonium sulphate ($(NH_4)_2SO_4$) (Eijsink et al., 2017). Afterwards, tubes were removed from the heating blocks and were cooled till room temperature. The tubes were placed in the Kjeltec™ 8400 one by one where water and 33% Sodium hydroxide (NaOH) were automatically added to the tubes. During this process, ammonium Sulphate is converted into Ammonia gas (NH_3) which correspond to the nitrogen present in the sample. This ammonia (NH_3) is automatically distilled into 200ml Erlenmeyer flask which contains boric acid and an indicator (Bromine cresol green). In this flask ammonium (NH_4^+) and borate ($H_2BO_3^-$) ions are generated. Amount of Borate ions ($H_2BO_3^-$) generated is equal to the amount of ammonium ion (NH_4^+) and it is determined by titrating Borat ion ($H_2BO_3^-$) against hydrochloric acid (0.05M HCl). Amount of HCl is used to calculate N present in the sample with the help of formula below.

$$\% \text{ N} = C\,\frac{mol}{1000mL} * \frac{volume\ of\ HCl(mL)}{weight\ of\ sample} * 14\left(\frac{g}{mol}\right) * 100$$

Where,

C is the concentration of HCl; 0.05 mol/L = 0.05 mol/1000ml.

14 is the atom number for Nitrogen.

Volume of HCl used for titration of Borate ions is automatically registered by the instrument. Then % protein is calculated by multiplying % N with 5.7, which is Kjeldahl factor for wheat proteins.

2.3.5 STARCH CONTENT

An enzymatic method is used to determine starch in the wheat flour samples. This method measures the concentration of starch spectrophotometrically after hydrolysis of starch to glucose (McCleary et al., 1994).

PROCEDURE

0.09 g of each flour sample was mixed with 0.2 ml 80% ethanol and 3 ml α-amylase with vigorous shaking. All the samples were prepared in two parallels. These parallels were incubated in the water bath at 95ºC for 6 minutes while shaking after every two minutes to gelatinize the starch. And α-amylase hydrolyzed α-(1-4) bonds of starch into its building blocks, amylose and amylopectin. Afterwards, 100 µl amyloglucosidase was added after sodium acetate buffer and incubated in water bath at 50ºC for 30 minutes to hydrolyze the amylose into glucose molecules. Samples were filtered to remove undissolved flour particles and transferred to 10 ml measuring flask. 0.1ml of the filtrate was taken into glass tubes in 2 parallels. After adding 3 ml GOPOD reagent, and incubating for about 20 minutes, 4-aminoantipyrine was formed which gave pink color to samples**.** After cooling down, the absorbance of the sample was measured with a spectrophotometer (SpectraMax M2, Molecular Devices) at 510 nm and the starch content of the sample was calculated by using standard curve of glucose (0 – 1.0 mg/ml). Absorbance is calculated by the following formula

$$\%\ \boldsymbol{starch} = \frac{\mathrm{X}\left(\frac{\mathrm{mg}}{\mathrm{ml}}\right).* \mathrm{F(ml)} * \ 0.9}{\text{weighed sample(mg)}} * 100$$

Where

X is glucose concentration (mg/ml) calculated by standard curve

F is dilution factor i.e., volume of the measuring flask used for dilution

Factor 0.9 corresponds to 162/180; adjustment for molecular mass of glucose (180 kDa) and polymerised glucose (162 kDa) in amylose and amylopectin

2.4 *EX VIVO* DIGESTION

2.4.1 HUMAN GASTRIC AND DUODENAL JUICES

Human gastric and duodenal juices were collected according to Ulleberg et al. (2011) by aspiration of some volunteers at Moss Hospital, Norway. The gastric and duodenal juices were aspirated simultaneously through three-lumen silicone tube, and the aspirates were stored at -20 °C, then at -80 °C. The pepsin and trypsin activities of the human gastrointestinal (GI) juices were assayed prior to the simulated digestion.

2.4.2 GASTROINTESTINAL DIGESTIVE ENZYME ACTIVITIES

The measurements and calculations of enzyme activities in the gastrointestinal juices were performed according to COST Action INFOGEST protocol (Minekus et al. 2014). **Pepsin Activity Assay** used 2% haemoglobin (w/v solution in HCl) at pH 2 as substrate. Haemoglobin solution was incubated along with gastric enzymes into a shaking incubator at 37°C for 10 minutes. TCA (5% w/v trichloroacetic acid) was added to stop the reaction prior to centrifugation at 6,000 rpm for 30 minutes to precipitate haemoglobin. TCA soluble peptides were removed as the supernatant and the absorbance was recorded at 280 nm by spectrophotometer (SpectraMax M2, Molecular Devices). For **Trypsin Activity Assay**, TAME (p-Toluene-Sulfonyl-L-arginine methyl ester) was used as substrate. TAME was mixed with Tris-HCl buffer (0.046 M Tris, 0.0115 M CaCl2 and 1 M HCl) at pH 8.1 and incubated in the spectrophotometer for 3 min to achieve temperature equilibrium. Duodenal juices were added to the mixture and the absorbance was recorded at 247 nm and 25^{o}C for 10 minutes with readings every 60 sec. (SpectraMax M2, Molecular Devices). One unit of enzyme activity was defined as the amount of duodenal juices (ml) that hydrolyses 1 μmole of TAME per min at 25 °C.

PREPARATION OF SAMPLE FOR DIGESTION

Porridge containing 10% protein was prepared by mixing 0.5 g wheat flours with 10 ml water and heated at 100^{o}C in a water bath for 10-15 min of each sample. The heat-treated samples were ultraturexed to obtain a homogenized sample mixture, cooled and stored at 4^{o} until further used. Due to water-loss during preparation of the porridge the protein content in the samples were recalculated as shown in the formula below.

$$Volume\ correction\ for\ protein\ = \frac{\%\ \text{protein in Mirakel wheat} * 500}{\%\ \text{protein in other wheat}}$$

Correction for water loss (WL) is done by the following formula

$$Concentration\ factor\ = \frac{\text{reduced volume of sample}}{\text{original volume}} + 1$$

Where 1 is the concentration ratio factor.

$$Volume\ correction\ for\ WL = \text{theoratical volume required}\ * \text{concentration factor}$$

% protein in Mirakel was used as standard for calculation other samples and volume of Porridge for each type used is shown in **Table 5**.

Table 5: *Volume per porridge sample used to obtain a fixed protein concentration during the digestion*

Sample	Einkorn	Emmer	Spelt	Fram	Børsum	Bastian	Mirakel
Volume (µl)	445.7	415.3	430.6	566.4	502.6	452.4	525.2

2.4.3 *EX VIVO* DIGESTION MODEL FOR WHEAT TYPES

Digestion with human GI enzymes was performed according to the COST Action INFOGEST protocol (Minekus et al. 2014), with some modifications. Enzyme activities in the human juices were assayed prior to *ex vivo* digestion, and juices corresponding to pepsin activity of 2000 U/ml and trypsin activity of 100 U/ml were added to the digestion phases. According to INFOGEST, 2000 U of pepsin activity and 100 U of trypsin activity are required per 1 ml of total volume of digestion. So, 614 µl of the gastric enzyme was required for 2 ml volume of gastric phase digestion and 1544 µl of the duodenal enzyme was required for 4 ml volume of duodenal phase digestion.

2.4.3.1 *EX VIVO* DIGESTION

Digestion was carried out by heat treated flour and water mixture (2:8 W/V) in two parallels and volume of sample was adjusted to 5 mg/ml protein per sample prior to digestion. Flour mixture was mixed 50:50 (v/v) with salivary fluid (SSF) followed by simulated gastric fluid

(SGF) (components listed in **Table 6**) and human gastric juices (HGJ) (2000 U/ml). pH was adjusted to 3.0 by adding 1M HCl with an electrode pH meter (827 pH lab, Metrohm). Final volume of gastric sample was made up to 2000 µL by adding dH_2O. The samples were incubated in a water bath (Julabo) at 37 °C with magnetic stirrer (RCT basic, Kika labortachnik) for 120 min. Gastric samples were taken out and enzyme activity was stopped by increasing pH >7 using (1M $NaHCO_3$) and samples were stored at -20 ° C immediately. Duodenal samples were further processed by adding simulated intestinal fluid (SIF; **Table 6**) and human duodenal juices (HDJ, 100 U/ml) 50:50 (V/V). PH was adjusted to 7.0 by addition of 1 M NaOH and volume of duodenal sample was made up to 4000 µL by adding dH_2O. The samples were incubated in water bath at 37 °C for another 120 min with magnetic stirring. Enzyme activity was stopped by adding 5mM Pefabloc® (76307) (Sigma Aldrich) to the duodenal samples and stored cold at -20 °C until further use.

Table 6: *Electrolyte stock solution prepared for oral, gastric and duodenal phase of digestion (Minekus et al. 2014)*

	Stock conc.		**SSF pH 7**		**SGF pH 3**		**SIF pH 7**	
			Volume of stock	Conc. In SSF	Volume of stock	Conc. In SGF	Volume of stock	Conc. In SIF
components	g/L	Mol/L	mmol/L	ml	mmol L^{-1}	ml	mmol L^{-1}	ml
KCl	37.3	0.5	15.1	0.5	6.9	6.9	6.8	6.8
KH2PO4	68	0.5	3.7	3.7	0.9	0.9	0.8	0.8
NaHCO3	84	1	6.8	6.8	13.6	13.6	12.5	12.5
NaCl	117	2	-	-	11.8	47.2	9.6	38.4
MgCl2(H2O)6	30.5	0.15	0.5	0.15	0.4	0.1	0.1	0.33
(NH4)2CO3	48	0.5	0.06	0.06	0.5	0.5	-	-
CaCl2(H2O)2	44.1	0.3	1.5	-	0.15	-	0.6	-

2.5 PROTEIN PROFILE BY SODIUM DODECYL SULPHATE POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE)

SDS-PAGE uses net charge and the size of protein molecules to separate proteins. In SDS-PAGE proteins are denatured and loaded with negative charge on their side chains by using a strong anionic detergent and the negatively charged proteins move towards the respective electrodes and their speed depends on their molecular mass and charge. The samples were analysed by SDS-PAGE to observe the enzymatic degradation of proteins during the digestion. Electrophoresis was done according to Morel et al (2002). All samples, undigested and digested, were made up to the same concentration by lyophilizing all the samples and then concentrating up to 5 mg/ml protein before mixing 1:1 with fresh SDS buffer containing (0.125M tris-HCl, 10 % SDS, 20 % glycerol and 200 mM dithiothreitol (DTT at pH 6.8)) and heated at 95 °C for 5 min to denature the proteins. 30 µl of all digested and undigested samples was added to the comb of a 12 % Mini-PROTEAN TGX Stain-Free Precast Gel (Bio-Rad) with 10 wells and covered with 10x running buffer containing (0.025M Tris, 0.192M glycine, 0.1 % SDS at pH 8.3), and run at constant voltage 200 V for approx. 37 min. Low molecular weight protein ladder was used as a standard. The proteins gels were transferred to dH_2O in a plastic box and scanned immediately by a UV-scanner to get a protein degradation profile. The electrophoresis of the samples was performed several times with varying protein concentrations to get optimum results.

2.6 DETERMINATION OF HYDROLYSIS OF PROTEINS BY SPECTROPHOTOMETRIC ASSAY USING O-PHTHALDIALDEHYDE (OPA)

Proteins are broken down to primary, secondary and tertiary amines by the action of digestive enzymes. This method was used to determine the degree of hydrolysis during digestion by measuring the number of primary amines present in digested samples and comparing them in each phase of digestion (Church et a, 1983). 200 µL of OPA buffer (containing 100 mM Na-tetra Borate, 10 % SDS, 40 mg OPA and 100 µL BME) and 50 µL sample were directly added to the microplates. Undigested and gastric samples were diluted 1:2 for gastric samples and 1:3 for duodenal samples with dH_2O prior to further dilution (1:2 in gastric and 1:4 in duodenal samples) directly on microplates. Absorbance was measured for plate spectrophotometrically at 340 nm. Absorbance of L-leucine standard (0 - 3 mM) and blank samples containing only gastric and duodenal juices were also measured along with sample. Church et al (1983) calculated degree of hydrolysis using amino acid sequences. However, the current calculations

were based on the Leucine equivalents per mg wheat proteins as the amino acid sequences for wheat proteins were not available.

2.7 LIQUID CHROMATOGRAPHY-MASS SPECTROSCOPY

Performed by Gianfranco Mamone at national institute of research, Avelino, Italy.

2.7.1 LC-HIGH RESOLUTION (HR)-MS/MS ANALYSIS

Mass spectrometry analysis was performed using a Q Exactive Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA), online coupled with an Ultimate 3000 ultra-high-performance liquid chromatography instrument (Thermo Scientific). Samples were diluted in 0.1 % (v/v) formic acid solution, loaded through a 5 mm long, 300 mm wide pre-column (LC Packings, USA) and separated by an EASY-Spray™ PepMap C18 column (2 mm,15 cm-75 mm) 3 mm particles, 100 Å pore size (Thermo Scientific). Eluent A was 0.1 % formic acid (v/v) in Milli-Q water; eluent B was 0.1 % formic acid (v/v) in acetonitrile. The column was equilibrated at 5% eluent B. Peptides were separated applying a 4 – 40 % gradient of eluent B over 60 min. The flow rate was 300 nL/min. The mass spectrometer operated in data-dependent mode and all MS1 spectra were acquired in the positive ionization mode with an m/z scan range of 350e1800. Up to 10 most intense ions in MS1 were selected for fragmentation in MS/MS mode. A resolving power of 70,000 full width at half maximum (FWHM), an automatic gain control (AGC) target of 1/106 ions and a maximum ion injection time (IT) of 256 ms were set to generate precursor spectra. MS/MS fragmentation spectra were obtained at a resolving power of 17,500 FWHM. To prevent repeated fragmentation of the most abundant ions, a dynamic exclusion of 10s was applied. Ions with one or more than six charges were excluded. Spectra were elaborated using the Xcalibur Software 3.1 version (Thermo Scientific).

2.7.2 MS ANALYSIS SPECTRA IDENTIFICATION

Mass spectra were elaborated using the Proteome Discoverer 2.1 software (Thermo Scientific), restricting the search to *Triticum* database extracted from the UniProt (downloaded in February 2018). Database searching parameters for peptide identification were the following: Met oxidation and pyroglutamic for N-terminus Gln as variable protein modifications; a mass tolerance value of 10 ppm for precursor ion and 0.01 Da for MS/MS fragments; no proteolytic enzyme selected. The false discovery rate and protein probabilities were calculated by Target Decoy PSM Validator working between 0.01 and 0.05 for strict and relaxed searches, respectively.

3 RESULTS

3.1 CHARACTERIZATION OF THE WHEAT TYPES

Several tests were performed to determine the quality of each wheat type including einkorn, emmer, spelt, some common wheat types; Fram and Børsum, Bastian and Mirakel. Moisture content (%) and thousand kernel weight (TKW) (g) of the dried grain samples of einkorn, emmer and spelt, was performed on harvested dehulled samples, while the common wheat types were free-threshing. Results showed that water content (%) of all the wheat types was approx. 10 % except for spelt which was 11 %. **Table 7** shows that the thousand-kernel weight (TKW) that varied between different wheat types. TKW was relatively high for spelt 41.1g and the for the common wheat type Mirakel 38.5 g, while for other wheat types between 30 - 33 g per thousand kernels. Falling number and starch content presented in **Table 7**, indicated the starch content and falling number differs between wheat types. Especially the old winter wheat types e.g., Fram and Børsum had very low falling number 150 and 89 respectively. All the other types had falling number almost 200 or above. Starch content was in the standard range (50-70 %) for all wheat types. Protein content of different wheat types varied between 8.2 and 11%.

Table 7: *Variations in thousand kernel weight (TKW) (g), falling number, starch (%) and proteins (%) among the different wheat types; einkorn, emmer, spelt, common wheat types i.e., Fram, Børsum, Bastian, and Mirakel*.

Wheat types	Thousand kernel weight (g)	Protein %	Falling number	Starch %
Einkorn	30.5	10.26	253	66.53
Emmer	31.7	10.96	255	59.96
Spelt	41.1	10.64	260	60.03
Fram	32.3	8.24	150	63.10
Børsum	31.9	9.21	89	55.61
Bastian	33.1	10.16	364	58.81
Mirakel	38.5	9.11	197	63.42

3.2 *EX VIVO* MODEL DIGESTION OF THE WHEAT PROTEINS

Digestion of different wheat types was performed with human gastrointestinal (GI) enzymes according to the COST Action INFOGEST protocol (Minekus et al. 2014), with some modifications. Pepsin activity in the human gastric juice (HGJ) was measured to be 651.9 U/ml and trypsin activity in human duodenal juice (HDJ) was 25.9 U/ml. All wheat porridge samples were digested with gastric (HGJ) and duodenal juices (HDJ) in parallels, then aliquoted and kept frozen at -20 ºC to stop enzyme activity until further used for lyophilization.

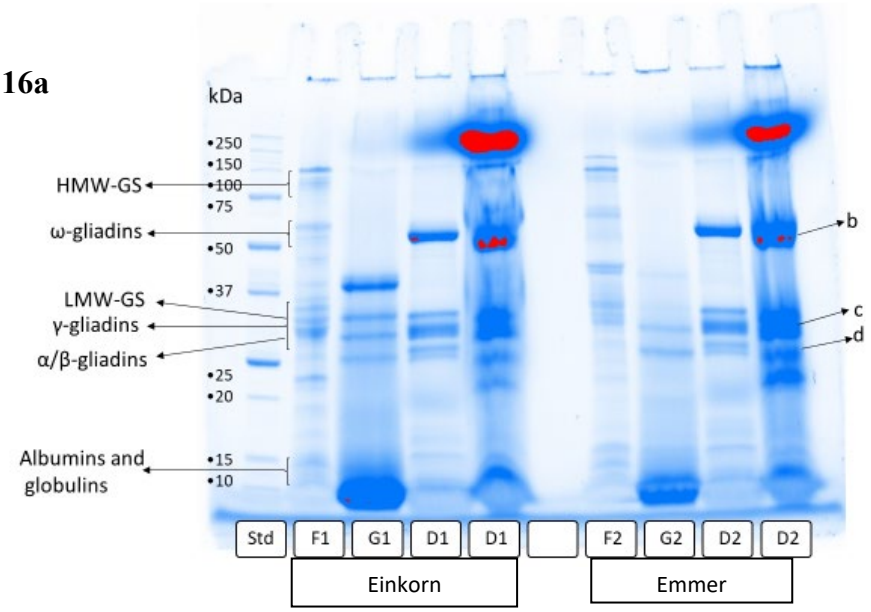
3.2.1 DEGRADATION PROFILE OF THE PROTEINS BY SDS-PAGE

Protein and degraded protein profile for the digested and undigested wheat types were analysed by **SDS-PAGE**. A trial run for the degradation profile with SDS-PAGE showed very low protein concentration in the digested samples, with almost invisible bands on the gels. These samples were then concentrated by lyophilization. The lyophilized samples were resuspended in water and used to get a clear protein degradation profile in all wheat types. The resulting bands of undigested samples showed presence of expected bands i.e., α-, β-, γ-, ω-gliadins and high molecular weight (HMW-) and low molecular weight (LMW-) glutenins in all wheat types (**Figure 16**). Globulins and albumins were also detected in the low molecular weight region of the gels. Bands in undigested einkorn exhibited fewer bands in the HMW-glutenin region as compared to other undigested wheat types. All the wheat types presented bands after gastric digetion in the molecular weight (MW) region (25-40 kDa). This region corresponded to gliadins which remained partially undigested during the gastric phase, however, later digested by HDJ. While HMW-glutenins were almost completely digested by HGJ, which is indicated by the absence of distinct protein bands in the respective region in all the wheat types. A new indistinct band appeared in the low molecular weight (LMW) fraction (MW<20 kDa), at the bottom of the gels in gastric digested samples, which disappeared again in duodenal digested samples. The lack of bands in the duodenal phase in all wheat types as shown in **Figure 16** a, b, c provided evidence for complete digestion of the gluten proteins, however, some bands observed belonged to the digestive enzymes in HGJ and in HDJ, previously confirmed by Devle et al., (2014).

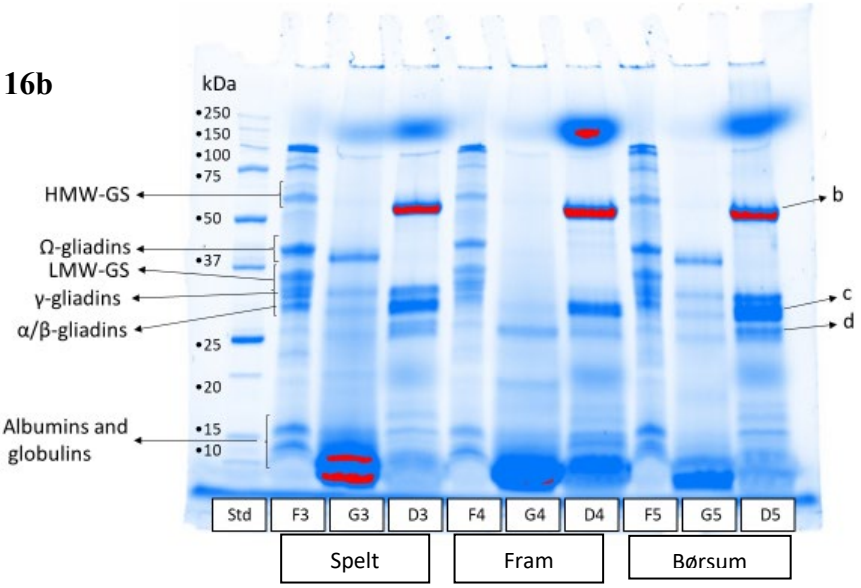
Emmer, Fram, and Mirakel seemed to contain more intact gliadins than einkorn, spelt and Børsum after gastric digestion. Results after the duodenal phase showed a very disinct ambigous band in the high molecular weight region of the gel for all the duodenal phase samples. All the doudenal samples showed a distict band at MW 55 kDa, some other distinct

bands in MW range (20-35) kDa and few indistinct bands in lower MW-region. Intensity of these observed bands varied among different wheat types e.g., Bastian and Mirakel showed higher intensity band as compared to spelt, Fram and Børsum. Wheras in case of einkorn and emmer, two different duodenal sample concentrations were applied to detrmine the optimun sampling concentration and the lower sampling showed optimum results.

16a



16b



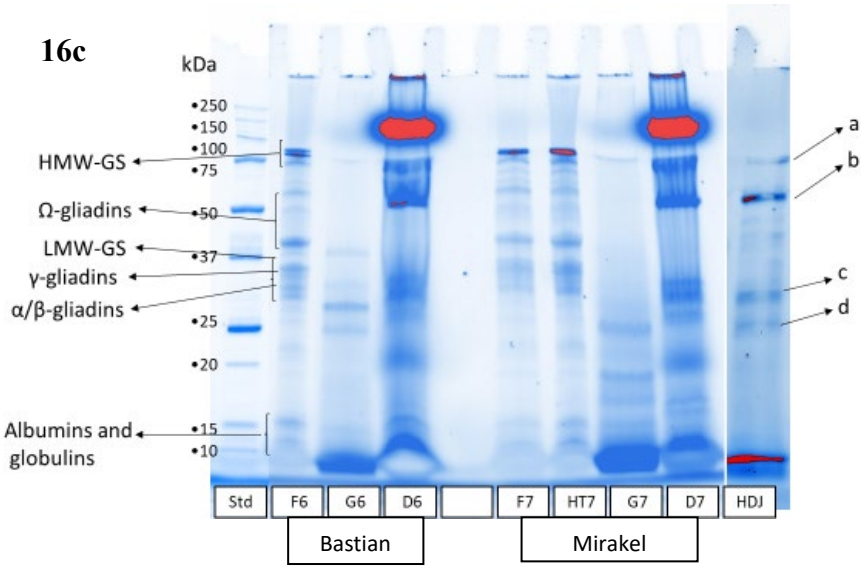


Figure 16 a, b, c): *SDS-PAGE of different wheat types showing digestion of proteins in gastric and duodenal phases and undigested wheat types. Std) denote 'standard MW ladder', F) denotes 'undigested flour', G) denotes 'gastric digested sample', D) denotes 'duodenal digested sample', HT) denotes 'heat treated undigested sample', HDJ) denotes 'human duodenal juices', HMW-GS) denotes 'high molecular weight glutenins' and LMW-GS) denotes 'low molecular weight glutenins'. Molecular weight in the standard ladder is expressed in kDa and D1 and D2 are two different concentrations of the same sample. Intestinal digestive enzymes in HDJ have been identified as a, amylase; b, carboxypeptidase, chymotrypsin, elastase, lipase, gastricsin and amylase; c, carboxypeptidase, elastase, lipase, trypsin and amylase; d, carboxypeptidase, chymotrypsin, elastase, lipase, trypsin and amylase enzymes (Devle et al., 2014).*

3.2.2 DEGREE OF HYDROLYSIS (DH) OF THE WHEAT PROTEINS

The degree of hydrolysis (DH) of the proteins was measured according to o-phthaldialdehyde OPA-assay (Church et al., 1983). Analysis measured the presence of primary amines produces during the hydrolysis of long peptide chain. And the coloured complexes formed with OPA-mercaptoethanol depends on the number of peptide bonds hydrolysed and measured spectrophotometrically. The concentration of primary amino-groups in all samples was measured and the degree of hydrolysis was defined as L-leucine equivalents per mg wheat proteins.

The DH obtained based on the L-Leucine equivalent per mg wheat protein is presented in **Figure 17**. Mirakel and Bastian showed only a small increase in Leucine equivalents of hydrolyzed proteins after digestion with HGJ. The other samples decreased in Leucine equivalents concentration (mM) after gastric phase digestion where einkorn decreased least and Fram decreased most as compared to Leucine equivalents (mM) in undigested samples. However, when digestion proceeded with HDJ, high increase in Leucine equivalents

concentration (mM) of all wheat types observed. Duodenal phase analysis represented a great variation in Leucine equivalents concentration (mM) among different wheat types. Analysis of duodenal phase in mirakel showed highest Leucine equivalents concentration (mM) whereas Børsum showed lowest Leucine equivalents concentration (mM). All other wheat types showed moderate Leucine equivalents concentration (mM).

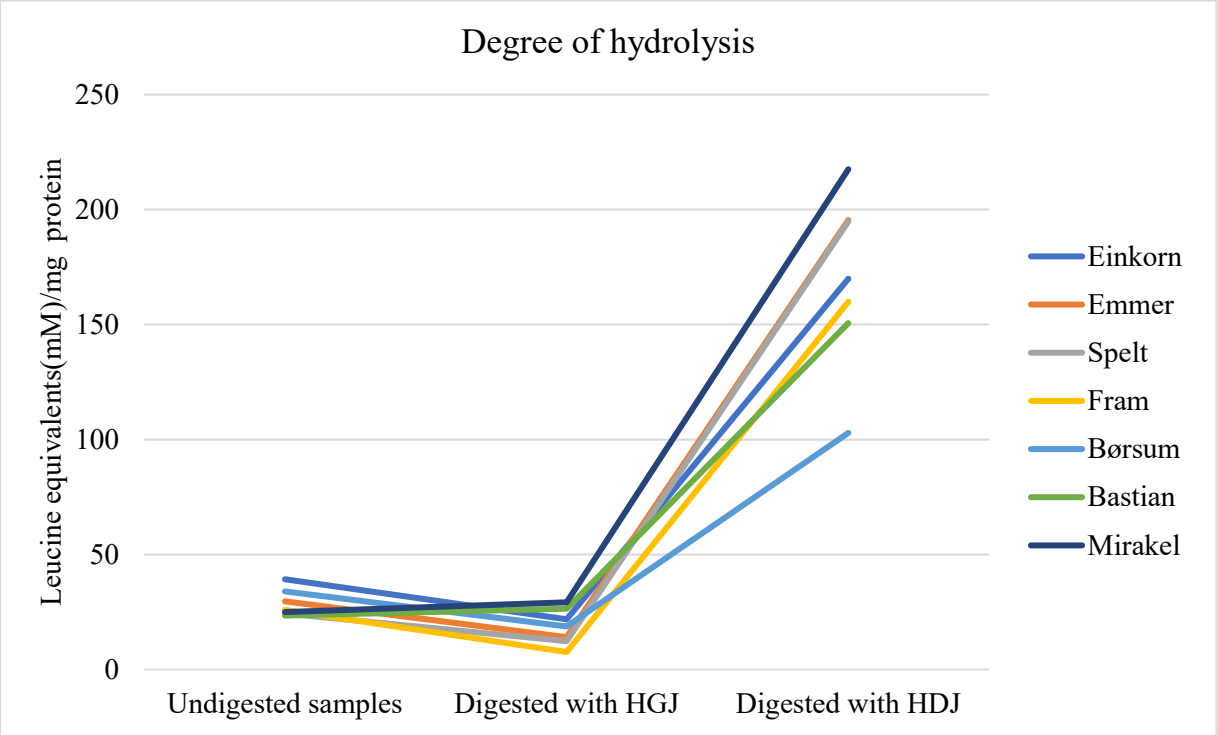


Figure 17: *Leucine equivalents concentration (mM) per mg wheat protein in undigested and digested wheat types; einkorn, emmer, spelt, Fram, Børsum, Bastian and Mirakel with human gastric juice (HGJ) and human duodenal juice (HDJ)*

3.2.3 IDENTIFICATION OF TOTAL PEPTIDES AND IMMUNOGENIC PEPTIDES AFTER *EX VIVO* DIGESTION OF THE WHEAT TYPES

The wheat types after digestion with human GI juices resulted in many peptide fragments. These peptide fragments were characterised by using LC-ESI-MS, where mass spectra were recorded for each peptide and fragmentation of these peptides by MS/MS gave a possible amino acid sequence for each peptide. The data obtained by MS/MS was elaborated by the software (Xcalibur) and to determine a complete peptide profile UniProt database was consulted. The database search was restricted to the proteome of *Triticum* and known sequences enabled identification of peptides and their respective mother proteins. The peptide profiles of digested samples contained a total of 1000 – 2686 peptide fragments each (**Table 8**) with variations in different wheat types. By manually processing, redundancies in the peptide

sequences were removed and an assorted list of unique peptides with their specific access number, m/z ratio, and name of mother protein was obtained. These assorted peptides were used as raw data for identification of immunogenic peptide sequences.

Immunogenic peptides were detected by comparing the assorted peptides list to the reference list of T-cell reactive epitopes presented by Sollid et al., (2012) (**Table 4**). The reference peptide sequences corresponded to specific T-cell reactive epitopes, which are relevant for celiac disease. The results presented in **Table 8** showed a great variation in the total number of peptides identified, the lowest in spelt (among hexaploid wheat varieties) and gradually increasing from Fram to Mirakel. However, digested einkorn and emmer showed a moderate number of total peptides with little variation. Similar behaviour was observed in assorted peptides for hexaploid wheat, whereas einkorn being diploid wheat had lower number of assorted peptides than emmer which is tetraploid. Digestion of Mirakel resulted in the highest number of identified immunogenic peptides (**Table 8**) with 133 unique peptides harbouring different T-cell reactive epitope sequences, whereas, einkorn produced the lowest number of identified immunogenic peptides 35. Emmer produced relatively high 71 immunogenic peptides while Bastian produced relatively low 78 immunogenic peptides which was unexpected. **Figure 18** shows the percentage of immunogenic peptides present and non-immunogenic peptides among the total assorted peptides in each wheat type. These results showed only little variation in the content of immunogenic peptides in different wheat types and the similar results of non-immunogenic content in each wheat type.

Table 8: *Number of detected peptides in each wheat type, shown as total number of peptides identified after LC-MS, assorted number of peptides by removing redundancy in data and number of immunogenic peptides according to the reference list of Sollid (2012), which provided the peptide sequences |associated with T-cell reactive epitopes*

	Total numbers of identified peptides	**Number of assorted peptides**	**Number of immunogenic peptides**
Einkorn	1586	215	35
Emmer	1551	264	71
Spelt	1051	178	49
Fram	1335	236	59
Børsum	1641	291	96
Bastian	2472	424	78
Mirakel	2686	433	133

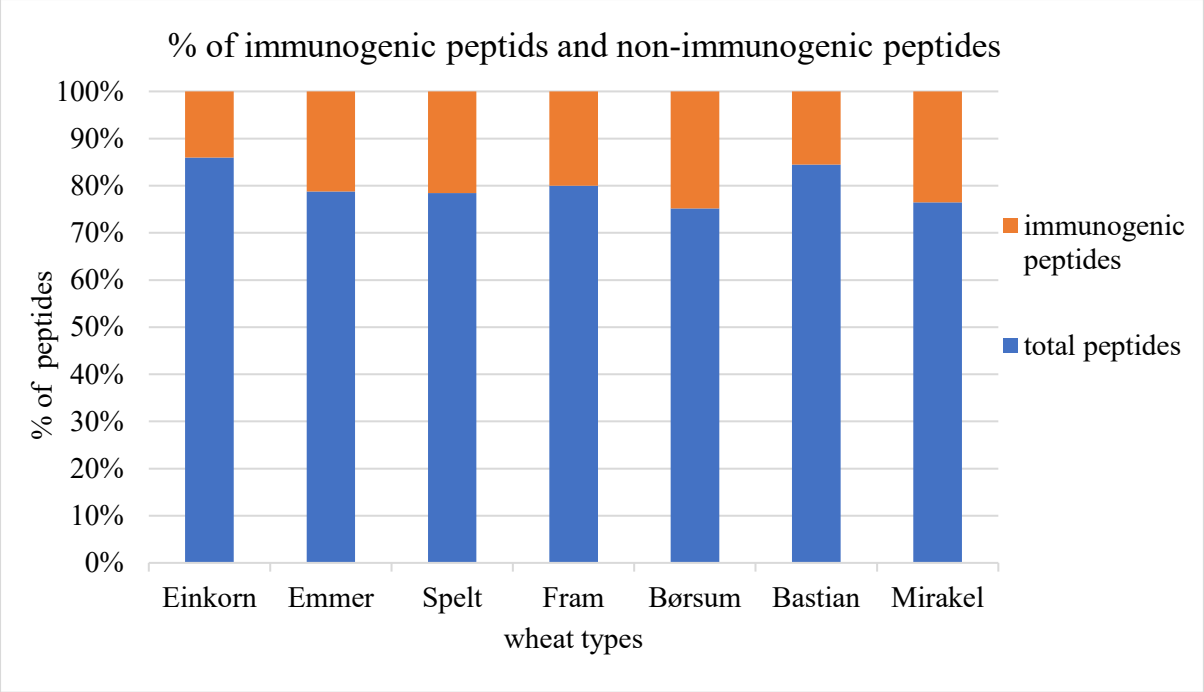


Figure 18: *Relative % of immunogenic and non-immunogenic peptides detected among the total assorted peptides in all wheat types.*

By comparing the reference peptide sequence list (Sollid et al., 2012) with the assorted peptides in wheat types, few completely matching peptides sequences were found. However, many larger peptide fragments harbouring immunogenic peptides sequences were identified. Figures 19,20 presents number of the immunogenic peptides identified in all wheat types corresponding to different T-cell reactive epitope sequences in the reference list. Number of peptide fragments corresponding to each reference epitope present in various wheat types are presented by different colours. Bastian and Mirakel showed large number of immunogenic peptides and a large variety of T-cell reactive epitopes, as compared to ancient wheat types e.g. einkorn, emmer and spelt. QQPEQPFPQ (DQ2-γ-VIIa) epitope was present in all wheat types, while PQPELPYPQ (DQ2-α-II, α2) epitope was only present in hexaploid wheat types e.g., Fram, Børsum, Bastian and Mirakel (*Triticum aestivum*). Einkorn was clearly distinguishable among all wheat types by carrying fewer immunogenic peptides and a small variety of T-cell reactive epitopes corresponding to these peptide sequences, PQP**E**Q**E**FPQ (DQ2-γ-VIIc), QQPEQPFPQ (DQ2-γ-VIIa), QQPFP**E**QPQ (DQ2-γ-VI), PFPQPEQPF (DQ2-ω-I), PFSEQEQPV (glutenin-17), and PQPEQPFPQ (Hor-α2, Hα2), followed by spelt which contained PFPQP**E**LPY (DQ2-α-I, α9), FRP**E**QPYPQ (glia-α20), SQP**E**Q**E**FPQ(DQ2-γ-IV), PQP**E**Q**E**FPQ (DQ2-γ-VIIc), QQPEQPFPQ (DQ2-γ-VIIa), QQPFP**E**QPQ (DQ2-γ-VI),

PFPQPEQPF (DQ2-ω-I), PQPEQPFPQ (Hor-α2, Hα2). In **Figure 20, 21** each T-cell reactive epitope is presented against its presence in different wheat types. In einkorn, only six of these 21 T-cell reactive epitopes were detected in peptide fragments harbouring immunogenic peptide sequences, while in spelt, Bastian and Mirakel 8, 12 and 14 different T-cell reactive epitopes were detected, respectively. Several peptide fragments were harbouring more than one T-cell reactive epitopes overlapping with each other and hence were counted several times with respect to immunogenic peptide sequence.

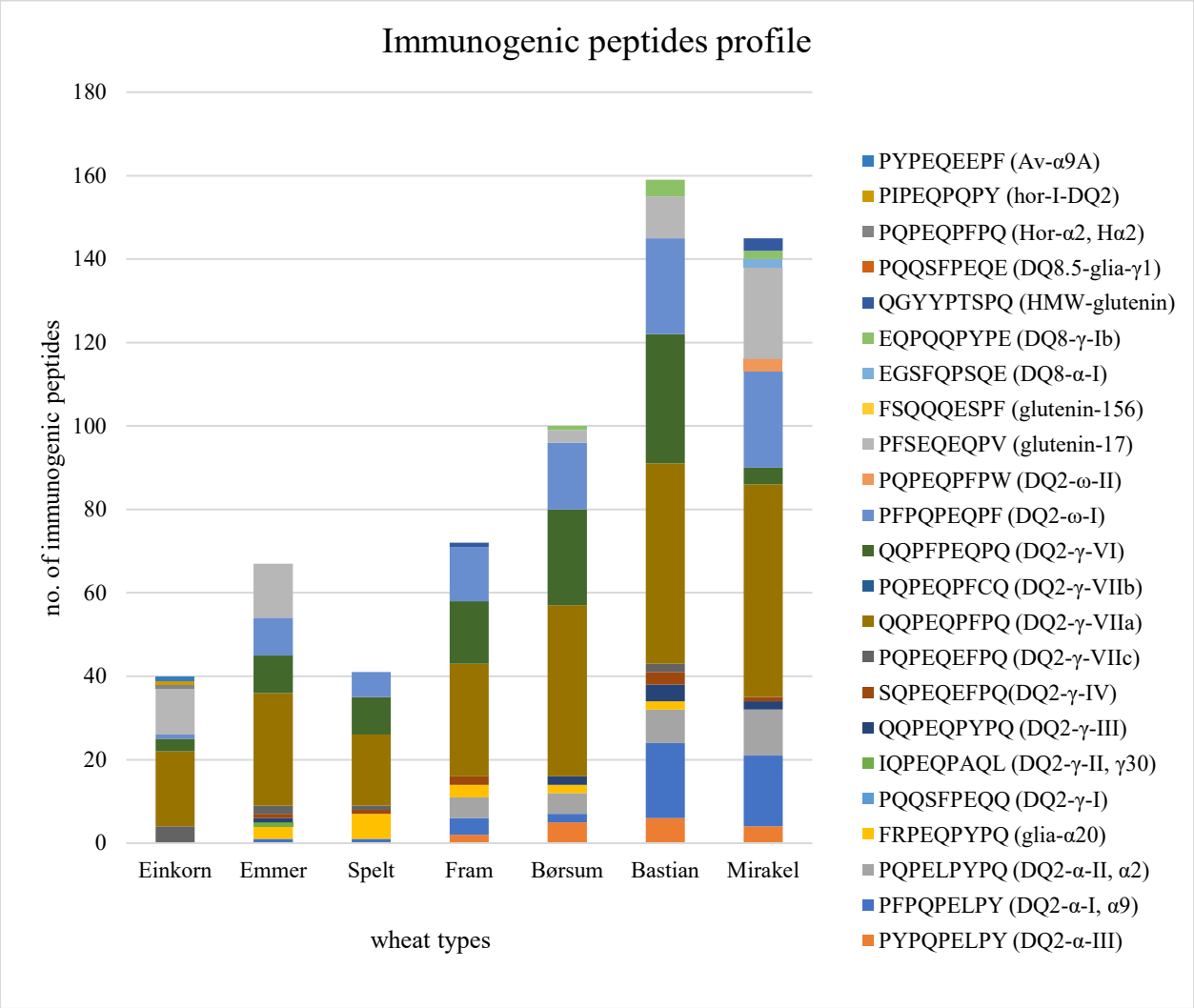


Figure 19: *Total number of identified immunogenic peptides present in all wheat types. Colours represent different T-cell reactive epitopes and the height of bar represents number of peptide fragments harbouring these epitopes previously identified by Sollid et al., 2012. In epitope names short terms are used to denote the type of proteins that epitopes are derived from: 'glia-α' denotes α-gliadin, 'glia-γ' denotes γ-gliadin, 'glia-ω' denotes ω-gliadin, 'glut-L' denotes low molecular weight glutenin, and 'glut-H' denotes high molecular weight glutenin.*

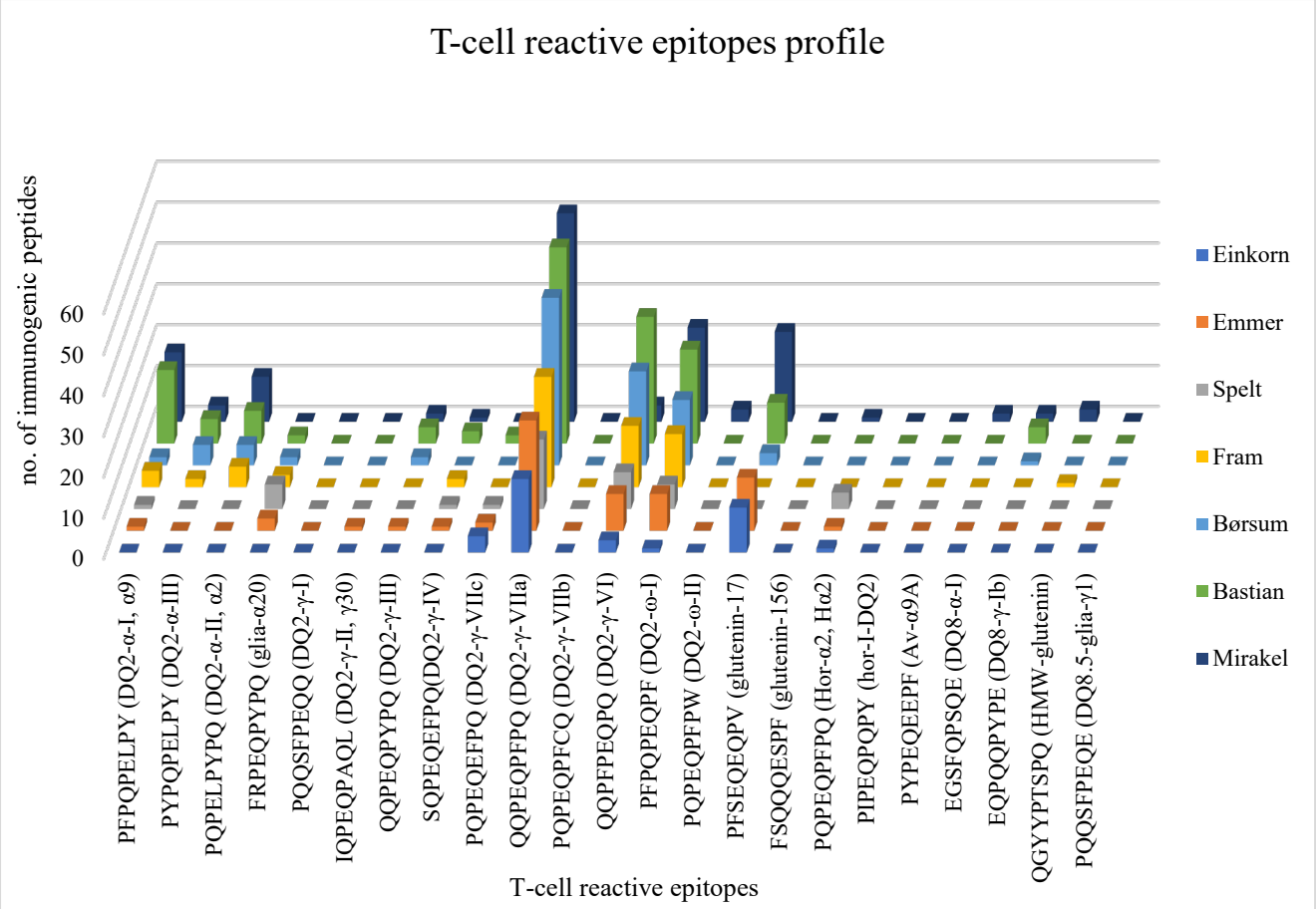


Figure 20: *A graphical representation of T-cell reactive epitopes and the number of peptide sequences harbouring these epitopes in each wheat type represented by different colours. Each bar represents a peptide sequence that is specific for T-cell reactive epitope e.g., PFPQPELPY is a peptide sequence that is specific to bind with T-cell reactive epitope DQ2.5-glia-α1a etc. (Sollid et. al, 2012). In epitope names short terms are used to denote the type of proteins that the epitopes are derived from: 'glia-α' denotes α-gliadin, 'glia-γ' denotes γ-gliadin, 'glia-ω' denotes ω-gliadin, 'glut-L' denotes low molecular weight glutenin, and 'glut-H' denotes high molecular weight glutenin.*

Information about the mother proteins that produced the immunogenic peptides was obtained from the database; the UniProt. These results were presented in **Figure 21**(a, b, c, d, e, f, g), where each sector of the pie diagram indicates respective mother protein and the % of immunogenic peptides which were derived from the mother proteins.

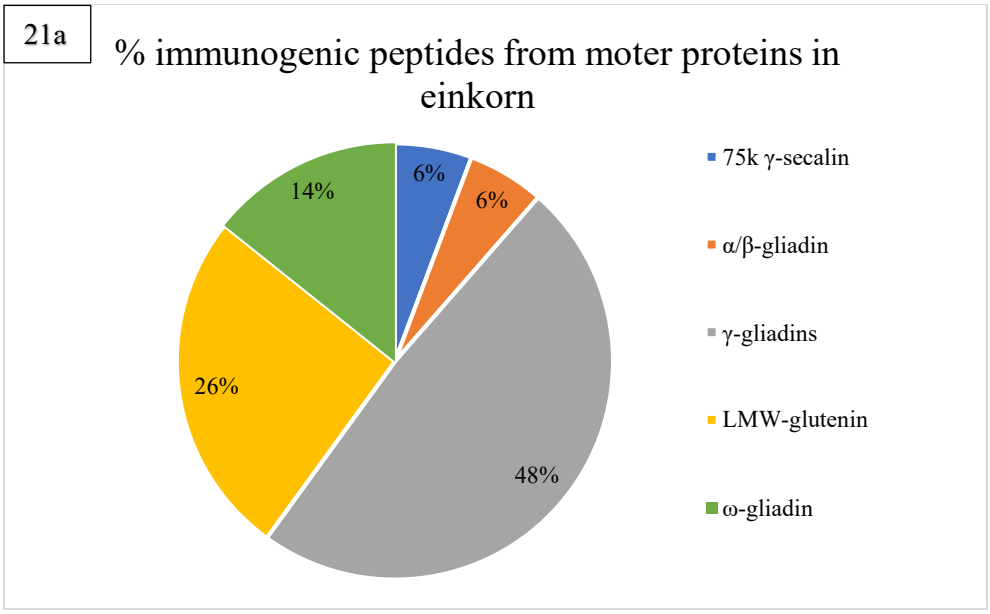


Figure 21a: *Diagrammatic representation of the % immunogenic peptides from each mother protein. γ-gliadins combined two different peptide fragments with access number M9TG60 and B6UKM7 in UniProt. LMW-glutenin represents low molecular weights glutenins and combined three peptide fragments A0A089VMD8, R4JB19 and F6M7E1.*

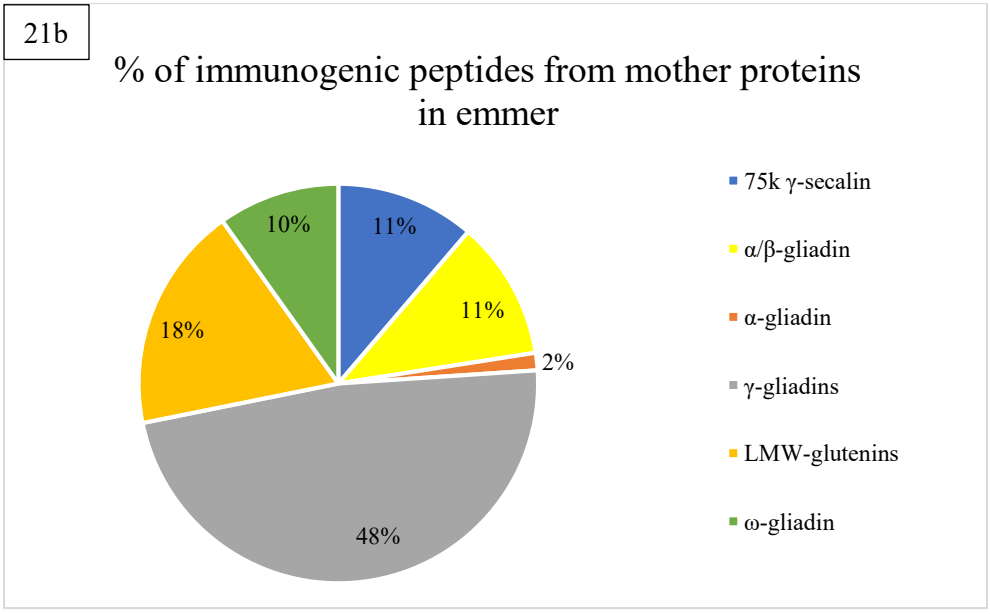


Figure 21b: *Diagrammatic representation of the % immunogenic peptides from each mother protein. γ-gliadins combined four different peptide fragments with access number B8XU42, B6UKS3, B6DQB2 and P08453 in UniProt. LMW-glutenin represents low molecular weights glutenins and combined four peptide fragments A0A089VMD8, D2DII7, B2Y2Q7 and F6M7E1.*

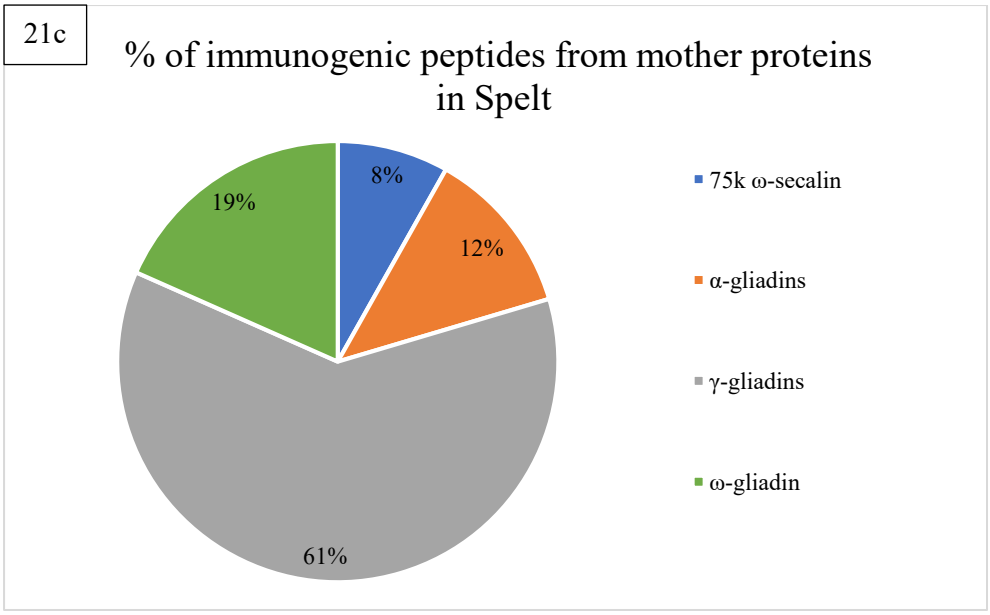


Figure 21c: Diagrammatic representation of the % immunogenic peptides from each mother protein. γ-gliadins combined four different peptide fragments with access number Q6EEW9, B6UKS3, D0EMA4 and R9XV74 in UniProt. A-gliadins combined two peptide fragments A0A0E3Z5A0, Q9M4M3.

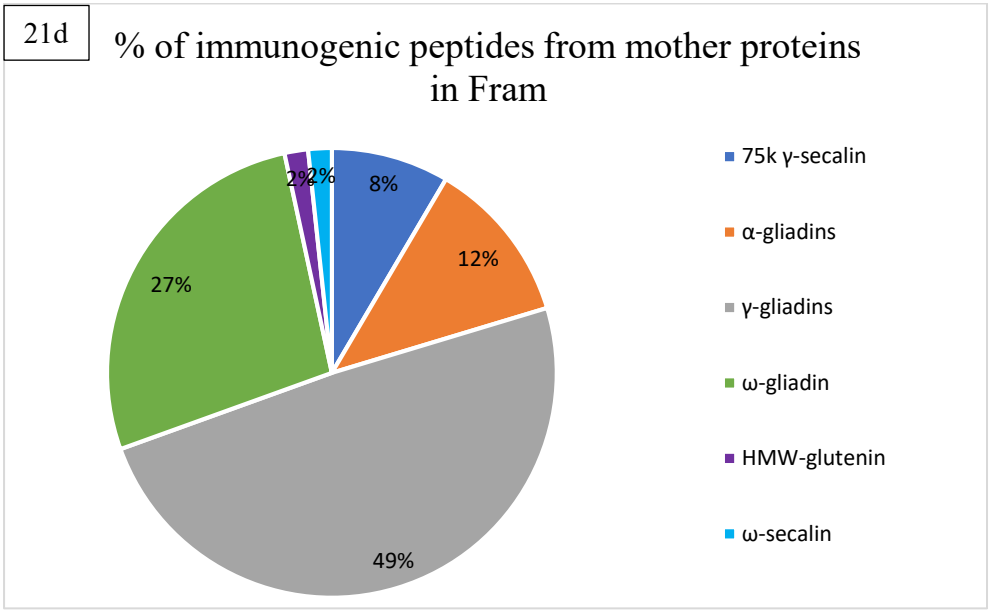


Figure 21d: Diagrammatic representation of the % immunogenic peptides from each mother protein. γ-gliadins combined three different peptide fragments with access number Q9M6P7, B6UKQ4, and P21292 in UniProt. α-gliadins combined two peptide fragments J7HT09 and X2KS61.

.

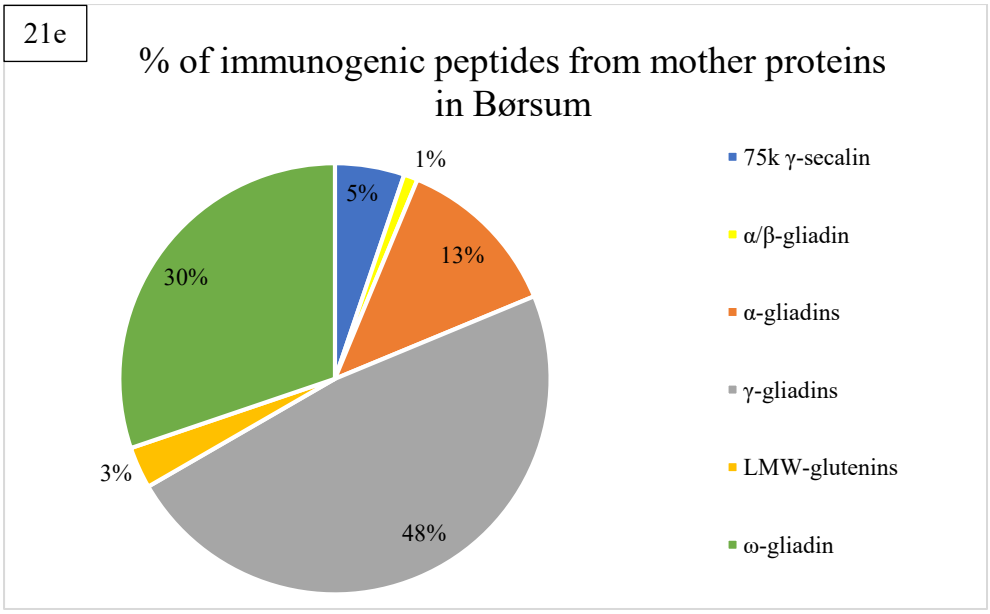


Figure 21e: Diagrammatic representation of the % immunogenic peptides from each mother protein. γ-gliadins combined three different peptide fragments with access number B6UKS3, B6UKN8, and M9TG60 in UniProt. α-gliadins combined four peptide fragments A5JSA7, J7HT09, K7WV37 and M4WH44. LMW-glutenins combine two peptide fragments B2Y2Q2 and C8KIL6.

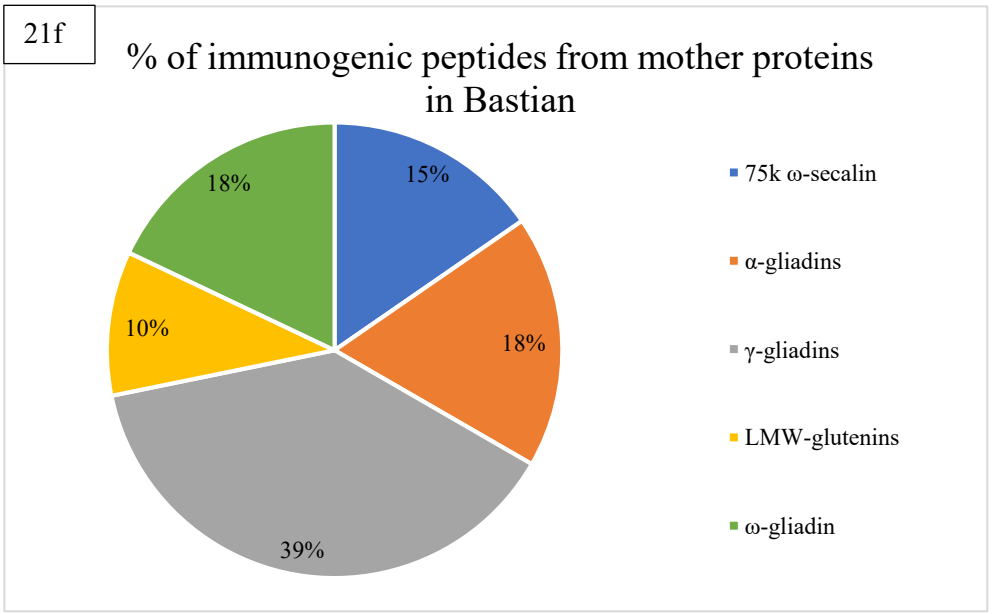


Figure 21f: Diagrammatic representation of the % immunogenic peptides from each mother protein. γ-gliadins combined sex different peptide fragments with access number B6UKM8, B6UKQ4, B6DQB2, B6UKL5, K7XEG8 and M9TG60 in UniProt. α-gliadins combined five peptide fragments A0A1K0IT12, A5JSA7, J7HT09, K7WV37 and M4WH44. LMW-glutenins combine four peptide fragments B2Y2Q7, P94021, A0A0A0R1F5 and B2Y2Q2.

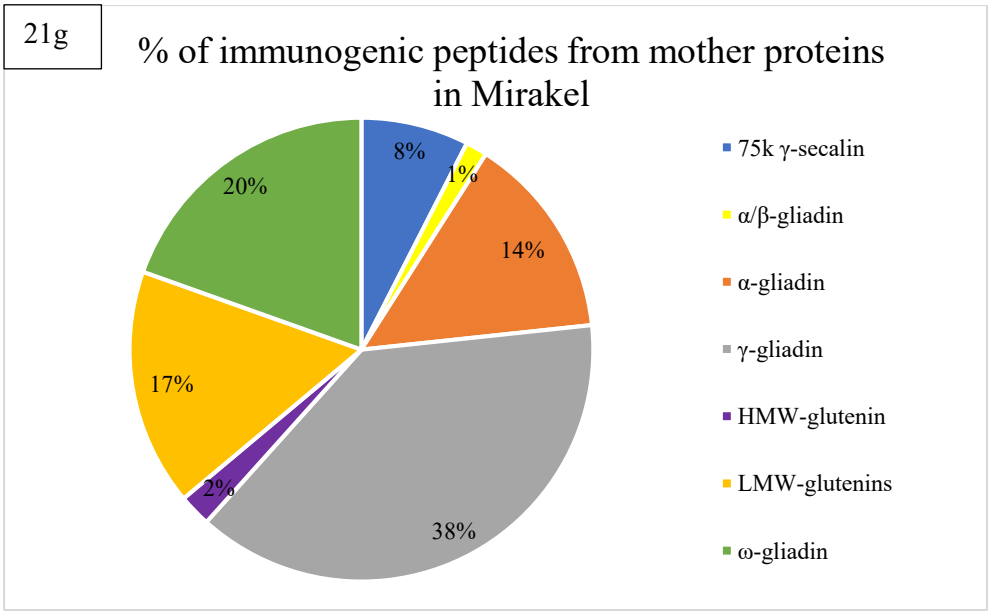


Figure 21g: Diagrammatic representation of the % immunogenic peptides from each mother protein. γ-gliadins combined seven different peptide fragments with access number B6UKS3, B6UKK2, B6UKK8, B6UKL1, P08453, R9XUR5 and U5U9Q1 in UniProt. α-gliadins combined four peptide fragments A5JSA7, J7HT09, K7WV37 and M4WH44. LMW-glutenins combine four peptide fragments B2Y2Q7, P94021, A0A068F9M7 and B8XU52. ω-gliadins combined A0A060N0S6 and R9XUR5.

Comparison of all the pie diagrams (21a-21g) gives a very good overview of the presence of different proteins in each sample. Pie diagram of each wheat type shows that the % of immunogenic peptides in main proteins varied greatly from einkorn to Mirakel. γ-gliadins were the main proteins that produced immunogenic peptides during digestion in all wheat types. Other proteins present in different wheat types are described as follows. LMW-glutenin contributed most towards immunogenic peptides % in einkorn and emmer, followed by ω-gliadins, while, α/β-gliadins contributed in a small amount. In spelt after γ-gliadin, α-, and ω-gliadins contributed significantly towards the immunogenic peptides production. ω-gliadins produced immunogenic peptides in significant amount in Fram, Børsum, Bastian and Mirakel. HMW-glutenins (only present in Fram and Mirakel) were the least contributors towards immunogenic peptides. 75k- gamma secalin was observed in all wheat types, generally this protein is found in rye.

Gluten proteins are complex macromolecules, when digested with GI enzymes, give rise to enormous amount of peptide fragments of different sizes and molecular mass. All the peptides presented in **Table 7** were originated from four main types of proteins; albumin, globulin, glutenin and gliadins. While immunogenic peptides were originated from glutenins and

gliadins. During digestion, the formation of peptides and their size is determined by the specificity of digestive enzymes e.g., pepsin, trypsin, chymotrypsin etc. **Table 9** presents the complexity of all these digested proteins. The proteins present in different wheat types were cleaved into peptide fragments of different sizes. **Table 9** shows a relationship between the immunogenic peptide fragments and mother proteins i.e., α-, β-, γ-, ω-gliadins and HMW and LMW glutenins. In **Table 9**, the column 1 listed T-cell reactive epitopes and their peptide binding register adapted from Sollid et al. 2012. Next columns represent the wheat types with mother protein that produced specific immunogenic peptides. Results indicated that different wheat types not only varied from each other in number of peptide fragments, but cleavage sites were also different that resulted in the varying length of the fragments. Einkorn showed the lowest number of immunogenic peptide fragments belonging to each T-cell reactive epitope. In einkorn, γ-gliadins produced peptides that harboured three different T-cell reactive epitope sequences. PFSEQEQPV associated with the epitope named (glutenin-17) was present in three different LMW-glutenin fragments. Emmer showed different set of proteins producing immunogenic peptides; i.e., γ gliadins produced peptides that were corresponded to three T-cell reactive epitopes (DQ2-γ-VIIc, DQ2-γ-VIIa, and DQ2-ω-I). Four different LMW-glutenin fragments corresponded to glutenin-17 epitope in emmer. In spelt γ -gliadins were the main protein corresponding to six different epitopes, with minor amount of ω-gliadin and α-gliadins. In Fram and Børsum, γ -gliadins corresponded to five and six T-cell reactive epitopes respectively, followed by ω-gliadins that corresponded to four epitopes in Fram and Børsum. In Fram HMW-glutenin also contributed towards T-cell reactive epitope. Bastian and Mirakel both had high in total amount of proteins and the number of the corresponding T-cell reactive epitopes was also highest among all wheat types. ω-gliadin was prominent in both samples that corresponded to six T-cell reactive epitopes in Bastian and four T-cell reactive epitopes in Mirakel. γ-gliadins were further divided into large fragments depending on their sequences and represented by different Accession code in UniProt of γ-gliadins. Repeated protein name or protein fragments present in **Table 9** represents different accession code.

***Table 9**: An overview of all the mother proteins that produced immunogenic peptide. These proteins have indirectly been identified by using a list of all the T-cell reactive epitopes adapted by Sollid et al., 2012. T-cell reactive epitopes derived from the respective main proteins are shown in different rows. For example, PFSEQEQPV is associated with epitope named (glutenin-17) was originated from the LMW-glutenin and identified in einkorn, emmer, Børsum, Bastian and Mirakel*

Peptide sequence	**Einkorn**	**Emmer**	**Spelt**	**Fram**	**Børsum**	**Bastian**	**Mirakel**
PFPQP**E**LPY (DQ2-α-I, α9)		Alpha/beta-gliadin	Omega-gliadin (Fragment)	Omega-gliadin Alpha-gliadin Alpha-gliadin protein	Omega-gliadin (Fragment)	Omega-gliadin (Fragment) Alpha-gliadin* Alpha-gliadin* Alpha-gliadin Alpha-gliadin	Omega-gliadin (Fragment) Alpha-gliadin Alpha-gliadin Alpha-/beta-gliadin storage protein
PYPQPELPY (DQ2-α-III)				Alpha-gliadin	Alpha-gliadin Alpha-gliadin	Alpha-gliadin Alpha-gliadin	Alpha-gliadin
PQPELPYPQ (DQ2-α-II, α2)				Alpha-gliadin Alpha-gliadin protein	Alpha-gliadin Alpha-gliadin Alpha/beta-gliadin	Alpha-gliadin Alpha-gliadin Alpha-gliadin Alpha-gliadin	Alpha-gliadin Alpha-gliadin Alpha-gliadin Alpha-/beta-gliadin storage protein
FRP**E**QPYPQ (glia-α20)		Alpha/beta-gliadin	75k gamma secalin Alpha-gliadin (fragment)	Alpha-gliadin	Alpha-gliadin	Alpha-gliadin	
PQQSFPEQQ (DQ2-γ-I)							
IQPEQPAQL (DQ2-γ-II, γ30)		Alpha/beta-gliadin					
QQPEQPYPQ (DQ2-γ-III)		75k gamma secalin			75k gamma secalin Alpha-gliadin	Gamma-gliadin 75k gamma secalin	75k gamma secalin
SQP**E**Q**E**FPQ (DQ2-γ-IV)		75k gamma secalin	Gamma gliadin	Gamma gliadin		Gamma-gliadin 1	Gamma-gliadin

PQP**E**QP**E**FPQ (DQ2-γ-VIIc)	Gamma-gliadin	Gamma-gliadin (Fragment)	Gamma gliadin			Gamma-gliadin (Fragment)	
QQPEQPFPQ (DQ2-γ-VIIa)	Omega-gliadin (Fragment) Omega-gliadin (Fragment) 75k gamma secalin Gamma-gliadin 1	Gamma-gliadin 75k gamma secalin 75k gamma secalin Alpha/beta-gliadin Omega-gliadin Omega-gliadin	Omega-gliadin (Fragment) Gamma gliadin 75k gamma secalin Gamma gliadin Gamma gliadin	Gamma gliadin Omega-gliadin 75k gamma secalin Gamma gliadin Gamma gliadin (Fragment)	Omega-gliadin (Fragment) Gamma gliadin 75k gamma secalin Gamma-gliadin 1	Omega-gliadin (Fragment) Gamma-gliadin Gamma gliadin 75k gamma secalin Gamma-gliadin Gamma-gliadin 1	Omega-gliadin (Fragment) Gamma-gliadin 75k gamma secalin Gamma-gliadin Gamma-gliadin Gamma-gliadin
PQP**E**QPFC**Q** (DQ2-γ-VIIb)							
QQPFP**E**QPQ (DQ2-γ-VI)	75k gamma secalin Gamma-gliadin 1	Alpha/beta-gliadin 75k gamma secalin Omega-gliadin	Omega-gliadin (Fragment) Gamma gliadin 75k gamma secalin Gamma gliadin	Omega secalin Omega-gliadin Gamma gliadin Gamma gliadin (Fragment) Gamma gliadin (Fragment)	Omega-gliadin (Fragment) Gamma-gliadin Gamma-gliadin Alpha-gliadin Gamma-gliadin 1 LMW-glutenin subunit	Omega-gliadin (Fragment) Gamma-gliadin Gamma-gliadin 75k gamma secalin Gamma-gliadin 1 LMW-glutenin subunit	75k gamma secalin Omega-gliadin (Fragment)
PFPQPEQPF (DQ2-ω-I)	75k gamma secalin	75k gamma secalin Gamma-gliadin Gamma-gliadin 75k gamma secalin	Gamma-gliadin Gamma gliadin Gamma gliadin	Omega-gliadin 75k gamma secalin Gamma gliadin	75k gamma secalin Omega-gliadin (Fragment) Gamma-gliadin Gamma-gliadin 1	75k gamma secalin Omega-gliadin (Fragment) Gamma-gliadin Gamma-gliadin Gamma-gliadin Gamma-gliadin 1	Omega-gliadin (Fragment) B6UKK8 Gamma-gliadin 75k gamma secalin Gamma-gliadin Gamma-gliadin

PQPEQPFPW (DQ2-ω-II)							Omega-gliadin (Fragment)
PFSEQEQPV (glutenin-17)	LMW-Glutenin subunit(Fragment) LMW-Glutenin subunit LMW-H6-5-2 LMW-Glutenin subunit (Fragment)	LMW-Glutenin subunit(Fragment) LMW-Glutenin LMW-Glutenin subunit A3-4 LMW-Glutenin subunit LMW-H6-5-2			LMW-glutenin subunit Gamma-gliadin 1	LMW-glutenin subunits LMW-glutenin subunit LMW glutenin 2 (fragment)	LMW-Glutenin LMW-Glutenin LMW-glutenin LMW-Glutenin 2 (Fragment)
FSQQQESPF (glutenin-156)							
PQPEQPFPQ (Hor-α2, Hα2)	75k gamma secalin	75k gamma secalin	Omega-gliadin (Fragment) Gamma gliadin 75k gamma secalin Gamma gliadin				75k gamma secalin
PIPEQPQPY (hor-I-DQ2)							
PYPEQEEPF (Av-α9A)							
EGSFQPSQE (DQ8-α-I)							Alpha-gliadin
EQPQQPYPE (DQ8-γ-Ib)					75k gamma secalin	Gamma-gliadin 75k gamma secalin	75k gamma secalin
QGYYPTSPQ (HMW-glutenin)				Glutenin,HMW subunit DX5			Glutenin,HMW subunit DX5
PQQSFPEQE							

* Repeated protein name or protein fragments represented different accession code for UniProt.

4 DISCUSSION

This study aimed to determine the type, sequence and quantity of immunogenic peptides related to Celiac disease (CD) after gastrointestinal digestion in different wheat types including einkorn, emmer, spelt, and four varieties of common wheat; Fram, Børsum, Bastian and Mirakel. Identification of the immunogenic peptides originated from these wheat types was done after *ex vivo model* digestion with human gastrointestinal enzymes (HGJ and HDJ) for 120 minutes each, respectively. Protein degradation profile, degree of hydrolysis and immunogenic peptide sequences related to CD were determined.

4.1 CHARACTERIZATION OF WHEAT TYPES

The grain samples were selected and collected at Vollebekk research field, Ås, Norway, and were grown there in the 2017 season. The main quality analyses used for food wheat were performed, as this overview of the sample material could be important for the interpretation of results from the main study dealing with ex vivo digestion and the identification of immunogenic peptides.

The tests included thousand kernel weight (TKW), moisture content, falling number, starch and protein content. Thousand kernels weight for einkorn, emmer and spelt was measured after dehulling the kernel while other wheat types were free-threshing. The einkorn ear seemed to be smaller in size than other wheat types, hence double sampling number for einkorn ears were used. Results for TKW of all the samples showed expected behaviour except einkorn. The einkorn sample had low TKW but compared to previous experience with einkorn grown in Norway, this sample had relatively large kernels (A.K.). Among the common wheat, the older varieties Fram and Børsum were small in kernel size, whereas the modern variety Mirakel has larger grains. This observation in grain size probably reflect the genetic variation between these varieties as described by McFall and Fowler (2009).

Starch content of all the samples was within the expected range i.e., 50 - 70 % (Kent, 1994). The Falling Number test indirectly shows alpha-amylase activity and damage to starch due to amylase activity. High amylase activity (expressed by lower falling number) results in less cohesive and sticky dough, hence, making this flour unsuitable for bread baking (Ross and Bettge, 2009). Results for falling numbers of these samples showed great variation within the samples. For a good quality baking, flour falling numbers values should be minimum 200 (Fk.no). Most of the samples showed desired values of falling numbers, however, Fram and

Børsum had lower falling number. In 2017, frequent precipitation during the maturation and harvesting period for spring wheat caused severe pre-harvest germination (A.K.). During pre-harvest germination, amylase activity started in kernels that might have resulted in starch degradation. The protein content of the flour samples was measured using Kjeldahl method and results showed that all the samples were generally low in protein content. The low protein content of these samples could be an important factor in the further analysis of peptides originated from these samples. The low protein content of all the samples could be explained by the fertilization regime for this field trial. Moderate amount of fertilizers was used during the growing season that resulted in lower content of nitrogen and ultimately reduced the amount of protein in all the samples.

4.2 PROTEIN DEGRADATION PROFILES

SDS-PAGE provides protein degradation profile after digestion with HGJ and HDJ. Comparison of undigested samples of different wheat types gives an overview of proteins present in them. The results showed the presence of many protein bands according to their MW size. HMW-glutenin in the range of 80-120 kDa, ω-gliadins at 37-55 kDa, LMW-glutenin at 3-39 kDa, α/β-gliadins at 28-35 kDa and γ-gliadins at 31-35 kDa were detected. These results were found in their respective regions as described by Weiser and Schalk (Wieser, 2007, Schalk et al., 2017). An overview of the protein composition in the individual wheat types was obtained from undigested sample bands. Comparison of these bands gave relative amount of individual proteins in each wheat type. Loci for HMW-glutenins are present on 1A, 1B, 1D and there should be theoretically 5 bands present in the HMW-glutenin region of hexaploid wheat as 1Ay is silenced (Shewry et al., 2003a). Similarly, there should be four bands for tetraploid and two for diploid wheat types for x and y- subunits. However, the HMW-glutenin bands present in einkorn were not very distinct as in the other samples. Einkorn being a diploid wheat (AA) lacks some genes encoding for HMW-glutenin that has resulted in absence of respective bands. Tetraploid wheat i.e., emmer showed comparatively two distinct and one diffused band, yet fewer as compared to hexaploid wheat varieties. It can also be explained by the gene loci for HMW-glutenin in the tetraploid wheat (AABB) as it lacks 1D locus for HMW-glutenin (Molberg et al., 2005). However, only three distinct bands were observed in hexaploid wheat types. It is suggested that some bands might be overlapping with each other. Interpretation of these bands was difficult as only single variety of each wheat type is used and

the results were not reproduced and confirmed by analysis of other varieties exhibiting similar genome.

All the wheat varieties were digested with HGJ and HDJ and the respective protein and peptide profile bands were analysed. In the gastric phase, all the proteins were not fully digested, particularly the gliadin proteins. The undigested gliadins were identified as strong bands (≈ 30-45 kDa) present in the gastric phase. It indicates that gliadins in emmer, Fram, Bastian were more resistant to gastric enzyme as compared to gliadins in einkorn, spelt, Børsum and Mirakel. The results are supported by the fact that only pepsin is present in the gastric phase, which cleaves specifically very few peptide bonds (Barrett et al., 2009). A strong band at the MW< 10 kDa was observed in all gastric samples that could be an indication of smaller peptide fragments produces during digestion. Similar results were observed by measuring the degree of hydrolysis, which showed that the few peptide bonds had been hydrolysed during the gastric phase in all the wheat types. However, einkorn and spelt proteins were somewhat less digested as compared to emmer and common wheat. This was an unexpected behaviour, as being ancient varieties, it was expected that einkorn and spelt proteins were digested more easily. This unexpected behaviour can be explained by the gliadin/glutenin ratio in einkorn, as α-gliadins are the major protein group while glutenins are the minor (Weiser, 1999). Gliadins are reported to be quite resistant to the gastric enzyme due to its structure and disulphide bonds (Mamone et al., 2007). However, this was not confirmed by LC-MS analysis, since gastric samples were not analysed further by this method. All wheat types showed no gluten protein bands after human duodenal digestion, however, the bands observed in this region were most probably duodenal enzyme bands reported previously by Devle et al., 2010. These results were confirmed by comparing the Leucine equivalents (mM) per mg wheat protein in gastric and duodenal phase. Value of Leucine equivalents in the duodenal phase was very high as compared to the gastric phase. It is suggested that high Leucine equivalents value in duodenal phase is due to almost complete digestion of the wheat proteins. The Leucine equivalents (mM) per mg wheat protein varied within the wheat types, Børsum was the lowest and Mirakel showed the highest Leucine equivalents value. It depends on the primary structure and amino acid composition of the gliadins and glutenins present in the wheat types (Church et al., 1983).

4.3 IDENTIFICATION OF IMMUNOGENIC PEPTIDES

The complexity of wheat proteins and peptide sequences and the importance of gluten for CD was already described in 2002 by Shan et al. Current study also showed similar results by

presenting the huge number of peptides after gastrointestinal digestion of seven Norwegian wheat types (**Table 8**). The focus in this study was on peptide sequences that are related to known immunogenic peptide sequences reported earlier by Ciccocioppo et al., (2005), Mamone et al., (2011) and Sollid et al. (2012). From these previous studies it was established that the 33mer fragment (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF) from α/β-gliadin (Ciccocioppo et al., 2005), 26mer (FLQPQQPFPQQPQQPYPQQPQQPFPQ) from γ-gliadin and 25mer (LGQQQPFPPQQPYPQPQPFPSQQPY) derived from α-gliadins (Mamone et al., 2007) were responsible for the immune reaction in CD patients (Gianfrani et al., 2007). Results from current study showed the absence of the 33mer, 26mer and 25mer fragments in all the digested wheat samples. Larger peptides seemed to be digested into smaller peptide fragments. This observation led to the conclusion that the 33mer was further cleaved into smaller peptide fragments by the action of human GI enzymes, as proposed earlier by Mamone et al. 2008. To verify this assumption, the data was further refined to detect peptide fragments harbouring immunogenic peptide sequences. Immunogenic peptides were identified using the peptide list of celiac disease related epitope/peptide-binding register given by Sollid et al., (2012). The results showed presence of many peptides, which harbour the immunogenic sequences. Many peptides contained more than one immunogenic sequence, overlapping with each other. The number of immunogenic peptides identified in each wheat type varied a lot. Einkorn carried the least number of immunogenic peptides followed by spelt as expected (**Table 8**), whereas Mirakel carried the highest number of immunogenic peptides. Comparison of einkorn and emmer with other hexaploid wheat types (**Figure 18**) in terms of the ratio for % of immunogenic peptides to the non-immunogenic peptides, did not show a major difference between diploid, tetraploid and hexaploid wheat types. These results contrast with a study performed on the similar wheat genome representative samples from different cultivars by Prandi et al., (2017). They suggested that the ancient wheat varieties produced more immunogenic peptides as compared to modern wheat varieties.

Detailed analysis of all the T-cell reactive epitopes present in digested peptides as shown in **Figure 19, 20** is comparable with a previous study done by Molberg (2005). Molberg (2005) suggested that the loci for the 33mer originated from the α-gliadins, is located on 6D (Molberg, 2005). Hence it is suggested that the 33mer is absent in einkorn (AA) and emmer (AABB), while present in spelt, Fram, Børsum, Bastian and Mirakel (AABBDD). Detailed analysis of the 33mer showed that it contained 6 overlapping immunogenic epitope sequences (PFPQPELPY, PYPQPELPY, PQPELPYPQ, repeated twice) and all these T-cell reactive

epitope sequences were not detected in einkorn and only one sequence (PFPQPELPY) was detected in emmer and spelt. In hexaploid wheat varieties, these T-cell reactive epitopes were present in many of the peptides fragments. These immunogenic peptides were cleaved at different positions, producing peptides of different length with similar epitope sequence or multiple epitope sequences overlapping with each other. Whereas 26mer contains two immunogenic sequences (QQPQQPYPQ, QQPQQPFPQ) overlapping with each other. The QQPQQPYPQ fragment was absent in einkorn and emmer while present in hexaploid wheat. The QQPQQPFPQ is a nine-amino-acid sequence which was present in all the samples in varying quantity. It is suggested that QQPQQPFPQ derived from the γ-gliadin and was a part of 26mer which was cleaved during digestion. These immunogenic sequences were present in low amount in old hexaploid wheat varieties and much higher in modern wheat varieties. The 25mer is not detected in all the samples. However, immunogenic peptide (LGQQQPFPPQQPY) derived from α-gliadin (Mamone et al., 2011)) was present in einkorn, emmer, Børsum, Bastian and Mirakel and absent in spelt and Fram and it is suggested that LGQQQPFPPQQPY is a cleaved part of the 25 mer. The 25mer did not contain any t-cell reactive epitopes from Sollid's T-cell reactive epitope list. It was considered immunogenic as it can activate innate immune reaction in CD patients (Maiuri et al., 2003). It is difficult to explain their absence in spelt and Fram, although these samples are hexaploid. It is suggested that the small fragments are still immunogenic as they correspond to T-cell reactive epitopes sequences.

A comparative study of the gliadin/glutenin ratio in hexaploid, tetraploid and diploid wheat varieties, suggested that hexaploid wheat had the highest amount of HMW-glutenins among all the varieties and the lowest gliadin/glutenin ratio, while diploid and tetraploid wheat had high gliadin/glutenin ratio (Wieser, 2000). This relationship can be used to explain the number of immunogenic peptides observed in this study (**Figures 19**, **20**). Immunogenic peptides derived from gliadins were dominant in all wheat types. However, HMW-glutenins (Payne, 1987, Cornish et al., 2006) seems to give rise to more immunogenic peptides in hexaploid wheat as compared to diploid and tetraploid wheat. These results (**Figures 19, 20**) are supported by the observation that T- cell reactive epitopes e.g. DQ2-γ-VIIa, were present in all wheat types, indicating that a gene encoded for γ-gliadins is expressed in all wheat types. While, HMW-glutenin was present only in hexaploid wheat, indicating its gene located on 1A, 1B, 1D (Payne, 1987).

Immunogenicity of the gluten protein using commercial digestive enzymes had previously been reported by Mamone et al (2011). However, for the first-time human gastrointestinal juices were used for the *ex-vivo* digestion of gluten proteins and the results should be considered closer to the digestion in humans. The results obtained by this study have some limitations since the digestion was performed with HGJ and HDJ and brush border enzymes were not included in the digestion prosess. current study lacked passage and further degradation of the peptides in the human intestinal cell. TG2 present in lamina propria of intestinal cells play a key role in the transformation of these native peptide sequences to immunogenic peptidse by deamidating glutamine (Q) to glutamate (E). These deamidated peptides presented by the HLA-DQ molecules on the surface of the antigen presenting cells (APCs) is the main factor for activation of CD4+ T-cell (Molberg et al., 1998) (van de Wal et al., 1998). Since brush border enzymes were not included in the model digestion, the results were determined by theoretic deamidation of Q into E, at particular positions as described by Sollid et al., 2012. These peptide profiles are obtained by only one representative duodenal sample and one representative sample was used for digestion of each wheat type. It limits the reproducibility of results and its application when comparing these samples to the other samples grown by different cultivars.

5 CONCLUSIONS AND FUTURE PERSPEKTIVE

This study aimed to determine the immunogenic peptide profile in different diploid, tetraploid and hexaploid wheat types. The representative of each type; einkorn (diploid), emmer (tetraploid) and spelt (hexaploid), some common wheat varieties (hexaploid) e.g., Fram, Børsum, Bastian and Mirakel were digestion with human GI juices and their peptides profile was analysed. Following conclusions can be drawn by the current study. The protein degradation profile of different wheat types showed almost complete protein digestion after the duodenal phase.

This study showed that all the wheat types despite their genome and cultivation era (ancient or modern) do not perform differently in the human GI tract and produce as much immunogenic peptides as their genome allows.

In all wheat types ratio of total peptides to immunogenic peptides showed very little variation among the wheat types.

The total number of immunogenic peptides originated from the mother proteins increased significantly from diploid towards hexaploid except spelt. Spelt presented lower count of total and immunogenic peptides as compared to the other hexaploid wheat types.

The 33mer, 26mer and 25mer fragments were not detected in our samples and it is suggested that these peptides are digested into smaller fragments. Smaller fragments are considered immunogenic too as they harbour T-cell reactive epitope sequences.

Although the conclusion is drawn by theoretical deamidation of glutamine (Q) into glutamate (E) in epithelial cells with the lack of brush border enzymes this study gave a good comparison of the immunogenicity of different wheat types. However, to understand it fully, quantification of these peptides will be very useful. With these aspects in mind, further work on quantification of peptides could be achieved by using proteomics.

Further digestion of these peptides with brush border enzymes will be useful to determine the actual behaviour of these peptides in the human intestine. Whether these large fragments will be degraded into mono, di or tri-peptides, ready for absorption into the bloodstream or some large fragment remain undigested to start an autoimmune reaction.

These wheat samples can also be digested by the GI juices from Celiac patients to understand the difference between the digestive enzymes of normal human beings and celiac patients and their effect on the digestion of different wheat types.

To understand the immunogenicity of these peptides, an immune assay using an antigen/antibody response can be used before and after digestion with the brush border enzymes.

6 REFERENCES

A.K., U. personal communication, 2018

AEBERSOLD, R. & GOODLETT, D. R. 2001. Mass spectrometry in proteomics. *Chemical reviews,* 101**,** 269-296.

ARENTZ-HANSEN, H., KÖRNER, R., MOLBERG, Ø., QUARSTEN, H., VADER, W., KOOY, Y. M., LUNDIN, K. E., KONING, F., ROEPSTORFF, P. & SOLLID, L. M. 2000. The intestinal T cell response to α-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *Journal of Experimental Medicine,* 191**,** 603-612.

BARRETT, K. E., BARMAN, S. M., BOITANO, S. & BROOKS, H. 2009. Ganong's review of medical physiology. 23. *NY: McGraw-Hill Medical*.

BELDEROK, B., MESDAG, J., DONNER, D. A. & MESDAG, H. 2000. *Bread-making quality of wheat: a century of breeding in Europe*, Springer Science & Business Media.

BJAANES 1962. Forsøk med vårhvete sorter i Hedmarkog Oppland 1953-1960. *Forskning og forsøk i landbruket,* 13**,** 22.

BRANDOLINI, A., VACCINO, P., BOGGINI, G., ÖZKAN, H., KILIAN, B. & SALAMINI, F. 2006. Quantification of genetic relationships among A genomes of wheats. *Genome,* 49**,** 297-305.

CASSIDY, B., DVORAK, J. & ANDERSON, O. 1998. The wheat low-molecular-weight glutenin genes: characterization of six new genes and progress in understanding gene family structure. *Theoretical and Applied Genetics,* 96**,** 743-750.

CATTERALL, P. & CAUVAIN, S. P. 2007. Flour milling. *Technology of breadmaking.* Springer.

CHURCH, F. C., SWAISGOOD, H. E., PORTER, D. H. & CATIGNANI, G. L. 1983. Spectrophotometric assay using o-phthaldialdehyde for determination of proteolysis in milk and isolated milk proteins1. *Journal of Dairy Science,* 66**,** 1219-1227.

CICCOCIOPPO, R., DI SABATINO, A. & CORAZZA, G. R. 2005. The immune recognition of gluten in coeliac disease. *Clinical & Experimental Immunology,* 140**,** 408-416.

COLE, E. W., KASARDA, D. D. & LAFIANDRA, D. 1984. The conformational structure of A-gliadin: intrinsic viscosities under conditions approaching the native state and under denaturing conditions. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology,* 787**,** 244-251.

CORNISH, G., BÉKÉS, F., EAGLES, H. & PAYNE, P. 2006. Prediction of dough properties for bread wheats. American Association of Cereal Chemists International.

DE HOFFMANN, E. 2000. Mass spectrometry. *Kirk-Othmer Encyclopedia of Chemical Technology*.

DEVLE, H., ULLEBERG, E. K., NAESS-ANDRESEN, C. F., RUKKE, E.-O., VEGARUD, G. & EKEBERG, D. 2014. Reciprocal interacting effects of proteins and lipids during ex vivo digestion of bovine milk. *International Dairy Journal,* 36**,** 6-13.

DEWEY, D. R. 1984. The genomic system of classification as a guide to intergeneric hybridization with the perennial Triticeae. *Gene manipulation in plant improvement.* Springer.

DVORAK, J., LUO, M.-C., YANG, Z.-L. & ZHANG, H.-B. 1998. The structure of the Aegilops tauschii genepool and the evolution of hexaploid wheat. *Theoretical and Applied Genetics,* 97**,** 657-670.

EGGER, L., MÉNARD, O., DELGADO-ANDRADE, C., ALVITO, P., ASSUNÇÃO, R., BALANCE, S., BARBERÁ, R., BRODKORB, A., CATTENOZ, T. & CLEMENTE,

A. 2016. The harmonized INFOGEST in vitro digestion method: From knowledge to action. *Food Research International,* 88**,** 217-225.
FALLANG, L.-E., BERGSENG, E., HOTTA, K., BERG-LARSEN, A., KIM, C.-Y. & SOLLID, L. M. 2009. Differences in the risk of celiac disease associated with HLA-DQ2. 5 or HLA-DQ2. 2 are related to sustained gluten antigen presentation. *Nature immunology,* 10**,** 1096.
FINNEY, K., F 1948. Loaf volume and protein content of hard winter and spring wheat. *Cereal Chem.,* 25**,** 291-312.
FOLTZ, M., CERSTIAENS, A., VAN MEENSEL, A., MOLS, R., VAN DER PIJL, P. C., DUCHATEAU, G. S. & AUGUSTIJNS, P. 2008. The angiotensin converting enzyme inhibitory tripeptides Ile-Pro-Pro and Val-Pro-Pro show increasing permeabilities with increasing physiological relevance of absorption models. *Peptides,* 29**,** 1312-1320.
GIANFRANI, C., SICILIANO, R. A., FACCHIANO, A. M., CAMARCA, A., MAZZEO, M. F., COSTANTINI, S., SALVATI, V. M., MAURANO, F., MAZZARELLA, G. & IAQUINTO, G. 2007. Transamidation of wheat flour inhibits the response to gliadin of intestinal T cells in celiac disease. *Gastroenterology,* 133**,** 780-789.
GOODING, M. J. 2009. The wheat crop. *Wheat: chemistry and technology***,** 19-49.
GREEN, P. 2007. R; Cellier, C. Celiac disease. *N Engl J Med,* 357**,** 1731-43.
GRØNLUND, A., NJØS, A., VESTGARDEN, L. S. & LYNGSTAD, I. 2006. Utilisation of agricultural databases for statistical evaluation of yields of barley and wheat in relation to soil variables and management practices. *Acta Agriculturae Scandinavica Section B-Soil and Plant Science,* 56**,** 1-8.
HILLMAN, G. C. & DAVIES, M. S. 1990. 6. Domestication rates in wild-type wheats and barley under primitive cultivation. *Biological Journal of the Linnean Society,* 39**,** 39-78.
HOSENEY, R. C. 1994. *Principles of cereal science and technology*, American Association of Cereal Chemists (AACC).
HUANG, S., SIRIKHACHORNKIT, A., SU, X., FARIS, J., GILL, B., HASELKORN, R. & GORNICKI, P. 2002. Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat. *Proceedings of the National Academy of Sciences,* 99**,** 8133-8138.
INOMATA, N. 2009. Wheat allergy. *Current opinion in allergy and clinical immunology,* 9**,** 238-243.
INTERNATIONAL WHEAT GENOME SEQUENCING CONSORTIUM 2014. A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science,* 345**,** 1251788.
JOHANSSON, S., BIEBER, T., DAHL, R., FRIEDMANN, P. S., LANIER, B. Q., LOCKEY, R. F., MOTALA, C., MARTELL, J. A. O., PLATTS-MILLS, T. A. & RING, J. 2004. Revised nomenclature for allergy for global use: Report of the Nomenclature Review Committee of the World Allergy Organization, October 2003. *Journal of allergy and clinical immunology,* 113**,** 832-836.
KAGNOFF, M. F. 2007. Celiac disease: pathogenesis of a model immunogenetic disease. *The Journal of clinical investigation,* 117**,** 41-49.
KARELL, K., LOUKA, A. S., MOODIE, S. J., ASCHER, H., CLOT, F., GRECO, L., CICLITIRA, P. J., SOLLID, L. M., PARTANEN, J. & DISEASE, E. G. C. O. C. 2003. HLA types in celiac disease patients not carrying the DQA1* 05-DQB1* 02 (DQ2) heterodimer: results from the European Genetics Cluster on Celiac Disease. *Human immunology,* 64**,** 469-477.

KASARDA, D. 1999. Glutenin polymers: the in vitro to in vivo transition. *Cereal Foods World,* 44**,** 566-571.
KELLOGG, E. A. 2001. Evolutionary history of the grasses. *Plant physiology,* 125**,** 1198-1205.
KENT, N. L. 1994. *Kent's Technology of Cereals: An introduction for students of food science and agriculture*, Elsevier.
KIHARA, H. 1929. Conjugation of homologous chromosomes in the genus hybrids Triticum× Aegilops and species hybrids of Aegilops. *Cytologia,* 1**,** 1-15.
LEA, T. E. 2018.
LILLEMO, M. & DIESETH, J. 2011. Wheat breeding in Norway. *AP Bonjean, WJ Angus and M. van Ginkel, editors, The World Wheat Book. A history of wheat breeding.,* 2**,** 45-79.
LOHI, S., MUSTALAHTI, K., KAUKINEN, K., LAURILA, K., COLLIN, P., RISSANEN, H., LOHI, O., BRAVI, E., GASPARIN, M. & REUNANEN, A. 2007. Increasing prevalence of coeliac disease over time. *Alimentary pharmacology & therapeutics,* 26**,** 1217-1225.
LUNDIN, K. E., SCOTT, H., FAUSA, O., THORSBY, E. & SOLLID, L. M. 1994. T cells from the small intestinal Mucosa of a DR4, DQ7/DR4. DQ8 celiac disease patient preferentially recognize gliadin when presented by DQ8. *Human immunology,* 41**,** 285-291.
MAIURI, L., CIACCI, C., RICCIARDELLI, I., VACCA, L., RAIA, V., AURICCHIO, S., PICARD, J., OSMAN, M., QUARATINO, S. & LONDEI, M. 2003. Association between innate response to gliadin and activation of pathogenic T cells in coeliac disease. *The Lancet,* 362**,** 30-37.
MAMONE, G., FERRANTI, P., ROSSI, M., ROEPSTORFF, P., FIERRO, O., MALORNI, A. & ADDEO, F. 2007. Identification of a peptide from α-gliadin resistant to digestive enzymes: Implications for celiac disease. *Journal of Chromatography B,* 855**,** 236-241.
MAMONE, G., PICARIELLO, G., ADDEO, F. & FERRANTI, P. 2011. Proteomic analysis in allergy and intolerance to wheat products. *Expert review of proteomics,* 8**,** 95-115.
MATHEWS, C., VAN HOLDE, K. & AHERN, K. 2000. *Biochemistry. 2000*.
MATZ, S. A. 1991. *Chemistry and technology of cereals as food and feed*, Springer Science & Business Media.
MCCLEARY, B., GIBSON, T., SOLAH, V. & MUGFORD, D. 1994. Total starch measurement in cereal products: interlaboratory evaluation of a rapid enzymic test procedure. *Cereal Chemistry,* 71**,** 501-504.
MCFALL, K. L. & FOWLER, M. E. 2009. Overview of wheat classification and trade. *Wheat science and trade***,** 437-454.
MERLINO, M., LEROY, P., CHAMBON, C. & BRANLARD, G. 2009. Mapping and proteomic analysis of albumin and globulin proteins in hexaploid wheat kernels (Triticum aestivum L.). *Theoretical and Applied Genetics,* 118**,** 1321-1337.
MIKKELSEN, E. 1979. *Kornets eldste historie i Norge*.
MINEKUS, M., ALMINGER, M., ALVITO, P., BALLANCE, S., BOHN, T., BOURLIEU, C., CARRIERE, F., BOUTROU, R., CORREDIG, M. & DUPONT, D. 2014. A standardised static in vitro digestion method suitable for food–an international consensus. *Food & function,* 5**,** 1113-1124.
MJAERUM, J. Development of quality properties in Norwegian wheat. 24. Nordic Cereal Congress, Stockholm (Sweden), 20-23 May 1990, 1992. Svaloef AB.
MOLBERG, Ø., KETT, K., SCOTT, H., THORSBY, E., SOLLID, L. & LUNDIN, K. 1997. Gliadin specific, HLA DQ2-restricted T cells are commonly found in small intestinal biopsies from coeliac disease patients, but not from controls. *Scandinavian journal of immunology,* 46**,** 103-109.

MOLBERG, Ø., MCADAM, S. N., KÖRNER, R., QUARSTEN, H., KRISTIANSEN, C., MADSEN, L., FUGGER, L., SCOTT, H., NORÉN, O. & ROEPSTORFF, P. 1998. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nature medicine,* 4**,** 713.

MOLBERG, Ø., UHLEN, A. K., SOLLID, L. M., JENSEN, T., FLÆTE, N. S., ARNTZ-HANSEN, H., LUNDIN, K. E. A., BURKHARD, F. & RAKI, M. 2005. Mapping of Gluten T-Cell Epitopes in the Bread Wheat Ancestors: implications on celiac disease. *gastoenterology*.

MYHRE, B. 2004. Agriculture, landscape and society ca. 4000 BC–AD 800. *Norwegian agricultural history***,** 12-77.

NESBITT, M., HILLMAN, G., PENA-CHOCARRO, L., SAMUEL, D. & SZABO, A. 1996. Checklist for recording the cultivation and uses of hulled wheats. *Hulled wheats. Promoting the Conservation and Use of Underutilised and Neglected Crops,* 4**,** 234-245.

OSBORNE, T. B. 1907. *The proteins of the wheat kernel*, Carnegie institution of Washington.

PARCHMENT, O., TATHAM, A., SHEWRY, P. & OSGUTHORPE, D. 2001. Molecular modelling of the central repetitive domains of the high molecular weight subunits of wheat. *Cereal Chem,* 78**,** 658-662.

PAYNE, P. The use of near isogenic lines with different HMW glutenin subunits in studying bread-making quality and glutenin structure. Proc. 3rd Int. Workshop Gluten Proteins. Budapest, Hungary, 1987, 1987. World Scientific, 216-226.

PRANDI, B., TEDESCHI, T., FOLLONI, S., GALAVERNA, G. & SFORZA, S. 2017. Peptides from gluten digestion: A comparison between old and modern wheat varieties. *Food research international,* 91**,** 92-102.

RAWLINGS, N. D., BARRETT, A. J., THOMAS, P. D., HUANG, X., BATEMAN, A. & FINN, R. D. 2017. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic acids research,* 46**,** D624-D632.

RESCIGNO, M. & DI SABATINO, A. 2009. Dendritic cells in intestinal homeostasis and disease. *The Journal of clinical investigation,* 119**,** 2441-2450.

ROSS, A. S. & BETTGE, A. D. 2009. Passing the test on wheat end-use quality. *Wheat Science and Trade***,** 455-493.

SABELLI, P. & SHEWRY, P. 1991. Characterization and organization of gene families at the Gli-1 loci of bread and durum wheats by restriction fragment analysis. *Theoretical and Applied Genetics,* 83**,** 209-216.

SALADIN, K. S. 2013. *Anatomía y fisiología: la unidad entre forma y función (6a*, McGraw Hill Mexico.

SCHALK, K., LEXHALLER, B., KOEHLER, P. & SCHERF, K. A. 2017. Isolation and characterization of gluten protein types from wheat, rye, barley and oats for use as reference materials. *PloS one,* 12**,** e0172819.

SHAN, L., MOLBERG, Ø., PARROT, I., HAUSCH, F., FILIZ, F., GRAY, G. M., SOLLID, L. M. & KHOSLA, C. 2002. Structural basis for gluten intolerance in celiac sprue. *Science,* 297**,** 2275-2279.

SHEWRY, P. & TATHAM, A. 1997. Disulphide bonds in wheat gluten proteins. *Journal of Cereal Science,* 25**,** 207-227.

SHEWRY, P. R., D'OVIDIO, R., LAFIANDRA, D., JENKINS, J. A., MILLS, E. C. & BÉKÉS, F. 2009. Wheat grain proteins. *Wheat: chemistry and technology***,** 223-298.

SHEWRY, P. R., HALFORD, N. G. & LAFIANDRA, D. 2003a. Genetics of wheat gluten proteins. *Advances in genetics.* Elsevier.

SHEWRY, P. R., HALFORD, N. G., TATHAM, A. S., POPINEAU, Y., LAFIANDRA, D. & BELTON, P. S. 2003b. The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties.

SJÖSTRÖM, H., LUNDIN, K., MOLBERG, O., KÖRNER, R., MCADAM, S., ANTHONSEN, D., QUARSTEN, H., NORÉN, O., ROEPSTORFF, P. & THORSBY, E. 1998. Identification of a gliadin T-cell epitope in coeliac disease: general importance of gliadin deamidation for intestinal T-cell recognition. *Scandinavian journal of immunology,* 48**,** 111-115.

SOLLID, L. M. 2002. Coeliac disease: dissecting a complex inflammatory disorder. *Nature Reviews Immunology,* 2**,** 647.

SOLLID, L. M., MARKUSSEN, G., EK, J., GJERDE, H., VARTDAL, F. & THORSBY, E. 1989. Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. *Journal of Experimental Medicine,* 169**,** 345-350.

SOLLID, L. M., QIAO, S.-W., ANDERSON, R. P., GIANFRANI, C. & KONING, F. 2012. Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics,* 64**,** 455-460.

STEPNIAK, D., VADER, L. W., KOOY, Y., VAN VEELEN, P. A., MOUSTAKAS, A., PAPANDREOU, N. A., ELIOPOULOS, E., DRIJFHOUT, J. W., PAPADOPOULOS, G. K. & KONING, F. 2005. T-cell recognition of HLA-DQ2-bound gluten peptides can be influenced by an N-terminal proline at p-1. *Immunogenetics,* 57**,** 8-15.

STONE, P. J. & SAVIN, R. 1999. Grain quality and its physiological determinants. *Wheat: ecology and physiology of yield determination***,** 85-120.

STRAND, E. 1984. *Korn og korndyrking*.

TRYNKA, G., HUNT, K. A., BOCKETT, N. A., ROMANOS, J., MISTRY, V., SZPERL, A., BAKKER, S. F., BARDELLA, M. T., BHAW-ROSUN, L. & CASTILLEJO, G. 2011. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nature genetics,* 43**,** 1193.

TYE-DIN, J. A., GALIPEAU, H. & AGARDH, D. 2018. Celiac disease: a review of current concepts in pathogenesis, prevention and novel therapies. *Frontiers in pediatrics,* 6**,** 350.

UHLEN, A. 1990. The composition of high molecular weight glutenin subunits in Norwegian wheats and their relation to bread-making quality. *Norwegian Journal of Agricultural Sciences,* 4**,** 1-17.

UHLEN, A. K., SAHLSTRØM, S., MAGNUS, E. M., FÆRGESTAD, E. M., DIESETH, J. A. & RINGLUND, K. 2004. Influence of genotype and protein content on the baking quality of hearth bread. *Journal of the Science of Food and Agriculture,* 84**,** 887-894.

VAN DE WAL, Y., KOOY, Y., VAN VEELEN, P., PEÑA, S., MEARIN, L., PAPADOPOULOS, G. & KONING, F. 1998. Cutting edge: selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *The Journal of Immunology,* 161**,** 1585-1588.

VERHOECKX, K., COTTER, P., LÓPEZ-EXPÓSITO, I., KLEIVELAND, C., LEA, T., MACKIE, A., REQUENA, T., SWIATECKA, D. & WICHERS, H. 2015. *The Impact of Food Bioactives on Health: in vitro and ex vivo models*, Springer.

VIK, K. 1937. Resistance to powdery mildew in spring wheat. *Sci Rep Agric Coll Nor,* 17**,** 435-495.

WIESER, H. 2000. Comparative investigations of gluten proteins from different wheat species I. Qualitative and quantitative composition of gluten protein types. *European Food Research and Technology,* 211**,** 262-268.

WIESER, H. 2007. Chemistry of gluten proteins. *Food microbiology,* 24**,** 115-119.

WILLIAMS, P. C. 2000. *Applications of the Perten SKCS 4100 in flour-milling*, Canadian Grain Commission, Grain Research Laboratory.
WRIGLEY, C. & BIETZ, J. 1988. Proteins and amino acids.

EIJSINK, V., DEVOLD, T.G. & SYNSTAD, B. 2017. *Anvendt og eksperimentell biokjemi laboratorietkurs*. In: IINSTITUTT FOR KJEMI , B. O. M. (ed). Universitet for miljø- og biovitenskap

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway