EXPLORING THE CONDITIONS FOR THE EXPRESSION OF LOW MOLECULAR WEIGHT GLUTENIN 1D1 PROTEIN IN *ESCHERICHIA COLI (E. COLI)*

By

VENKATA HARINI GUDISEVA

Master of Science in Biotechnology

Andhra University, Visakhapatnam

India

2001

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
Degree of
MASTER OF SCIENCE
December 2006

EXPLORING THE CONDITIONS FOR THE

EXPRESSION OF LOW MOLECULAR WEIGHT

GLUTENIN 1D1 PROTEIN IN

ESCHERICHIA COLI (E. COLI)

Thesis Approved:

Dr. Patricia Rayas Duarte

Thesis Advisor

Dr. Robert L Matts

Dr Ramamurthy Mahalingam

A. Gordon Emslie

Dean of the Graduate College

Acknowledgements

I would like to express my sincere appreciation to my advisor Dr. Patricia Rayas Duarte for her constant guidance, encouragement and support through out my masters work. I would like to convey my gratitude to Dr. Carol Bender for kindly allowing me to use her laboratory facilities for my project. I am grateful to committee members Dr. Robert L Matts and Dr. Ramamurthy Mahalingam for their advice, support and suggestions through out my masters program.

My special thanks to Dr. Alejandro Penaloza for his guidance and help in my project. Also I would like to acknowledge other members of the lab Victor, Christy, Sarita for their friendship and making the lab a fun place to work.

I would like to thank my parents and sisters for their love, support and understanding throughout the whole process. Last but not least I would like to extend my heartfelt appreciation to my dearest friend Murthy Chavali for his help in enormous ways during the whole process and thesis writing.

TABLE OF CONTENTS

Chapter **Page**

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

Nomenclature

Da = Daltons

% = percentage

g = gram

hr = hour

L = liter

M = molar

min = minute

mg = milligram

ml = milliliter

°C = degrees centigrade

sec = second

μM = micromolar

μl = microliter

ER = Endoplasmic Reticulum

IPTG = Iso Propyl Thio Galactoside

α = alpha

β = beta

γ = gamma

δ = delta

ω = omega

HMW-GS = High Molecular Weight Glutenin Subunit

LMW-GS = Low Molecular Weight Glutenin Gubunit

Gli = Gliadin

MBP = Maltose Binding Protein

Chapter I

Literature Review

Literature Review

Introduction

Wheat is one of the largest cultivated crops belonging to the Poaceace family, the other two being barley and rice. The earliest known grains of wheat seem to originate approximately 5,000-6,000 B.C in the area known as the Fertile Crescent (consisting ancient Egypt, Mesopotamia and Syria). Varieties of wheat like Emmer and wild Eikorn were domesticated in Fertile Crescent (Hancock and James, 2004).

Wheat belongs to genus Triticacea and consists of species having diploid (n=14) e.g, *Triticum monococcum*, tetraploid (n=28), e.g *Triticum turgidum* or hexaploid (n=42) e.g. *Triticum aestivum* (bread wheat) sets of chromosomes. Within the species, wheat cultivars are classified further based on their growing season (spring or winter), by endosperm texture (hard or soft) or by color of grain (white, amber or red). Heterosis or hybrid vigor occurs in common wheat but cannot be produced on a commercial scale as wheat flowers are complete and self pollinating species. Commercially, hybrids have been generated either by chemical hybridizing agents or using plant growth hormones to interfere with pollen development (Bonjean and Angus, 2001).

Wheat kernels are made up of three major parts, the germ, bran and endosperm (Fig. 1). The germ is an embryonic plant and consists of embryo and scutellum. It has high concentrations of non-gluten proteins, lipids, vitamins and minerals (Posner, 2000). The bran is the outside covering and consists of several cellulose layers. It protects the kernel and is separated during the milling process. The endosperm is inside the kernel and is comprised mainly of starch and protein. Based on the endosperm texture wheat can be classified as vitreous (steely, flinty, grassy) or mealy (starchy, chalky) (Dowell et al., 2000). The endosperm is the food source for the germ plant and when milled becomes flour.

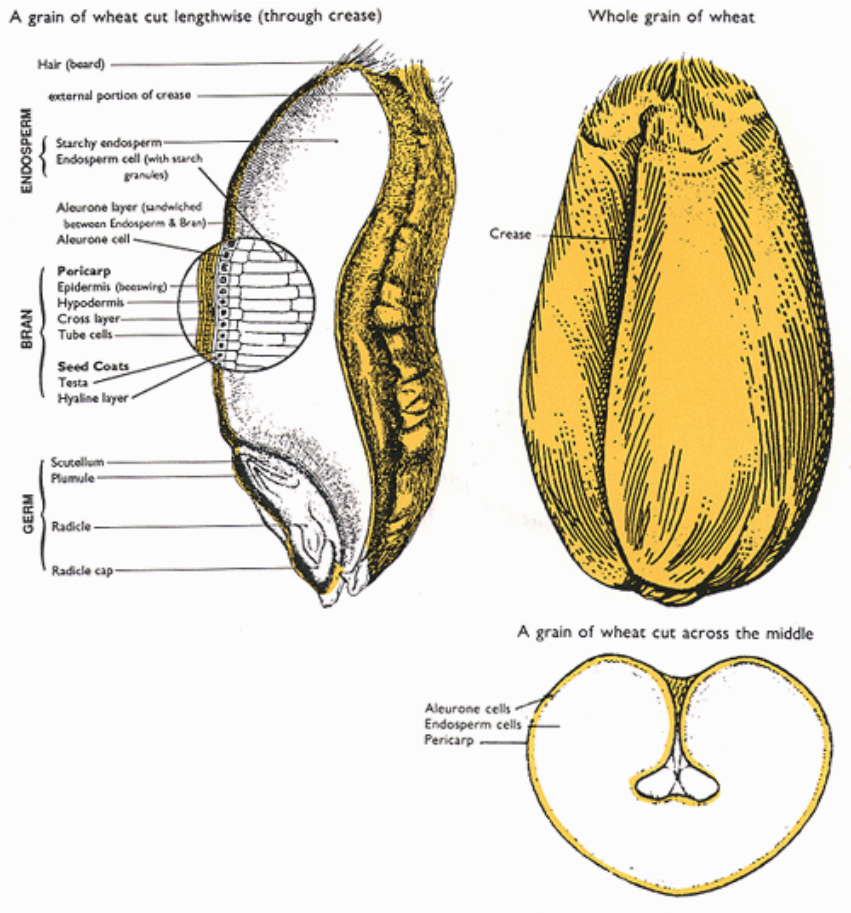


Figure 1: Image of whole wheat kernel and cross section of wheat kernel showing various layers**.**

Wheat Proteins

Wheat proteins are divided based on their solubility into gluten (gliadins and glutenin) which comprise 80-85% of total wheat protein and non gluten proteins (albumins and globulins) which are about 15-20% of total wheat protein. Albumins are water soluble while globulins are insoluble in water. Gliadins show high solubility in ethanol while glutenins dissolve in dilute acids or alkalis (Table 1) (Osborne, 1924).

Table 1. Solubility of wheat proteins (Osborne, 1924)

Proteins		Soluble in	Location in
Non-gluten proteins	Albumins	water	embryo (metabolic proteins) and endosperm cells (cytoplasmic proteins)
	Globulins	dilute salt solutions (O.5 M NaCl)	embryo and aleurone layer (storage proteins) and endosperm cells (cytoplasmic proteins)
Gluten proteins	Gliadins	70-80% ethanol	endosperm (storage proteins)
	Glutenins	dilute acids or alkalis solutions (0.05 M Acetic acid)	

The supply of nitrogen seems to influence the balance of protein fractions in the grain. The rate of accumulation also differs between these two classes of proteins. Non-gluten proteins (albumins and globulins) are present in high concentrations during the first 10 days of growth (Gupta, 1996) and steadily decline with kernel maturity. Gliadins are the first formed storage proteins that accumulate 5-10 days after anthesis. Glutenins accumulate last in the grain and are seen in significant quantities after about 20 days of anthesis (Stone and Nicolas, 1996). All of the wheat storage proteins (albumins, globulins and gluten proteins) are synthesized in the lumen of the ER. Most of the storage proteins seem to follow the secretory transport pathway from the ER to Golgi apparatus and finally into the storage vacuoles. Secretory proteins assume their folded conformation in the lumen of the ER by interactions of both intra and inter chain disulphide bonds and are assisted by molecular chaperones (Hsp 70/ Bip family) present in the ER. However,

some proteins accumulate in the ER and are incorporated in vacuole-like compartments that transform into protein bodies (Fig. 2) (Levanony et al., 1992). In wheat especially, the protein vacuoles are compressed between the starch granules but the precise mechanism of intracellular transport and deposition is still uncertain (Shewry and Tatham, 1995).

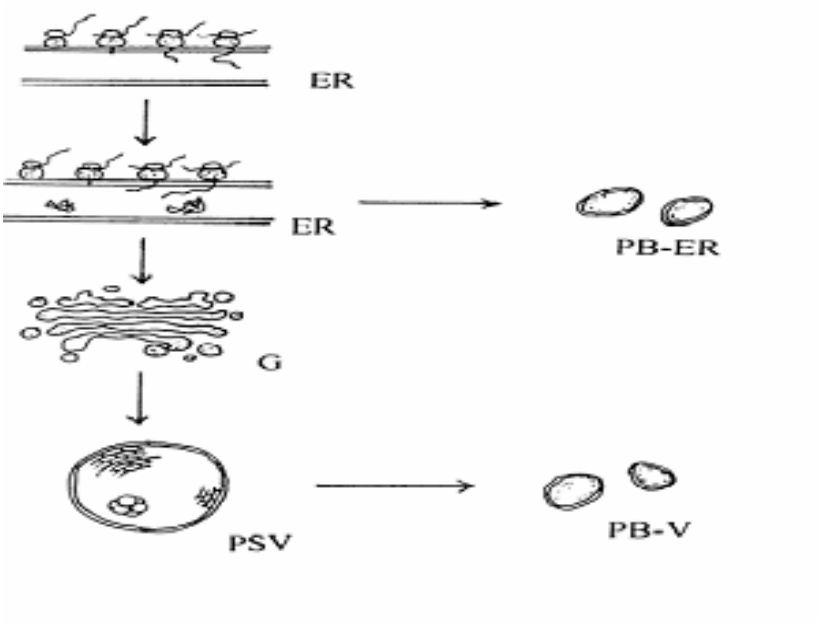


Figure 2. Various steps involved in storage protein processing and sorting in the Endoplasmic reticulum (Levanony et al., 1992); ER = Endoplasmic reticulum; G = Golgi; PSV= Protein sorting vesicle; PB-V –Protein body vesicle;

A. Non gluten proteins

The two kinds of non-gluten proteins albumins and globulins account for about 15-20% of the total wheat protein.

(i) ***Albumins***

Albumins are metabolic proteins and can be divided into six different classes with molecular weight between 17,000-28,000 Da. Albumins consist of mainly enzymes such as α-, β- amylases or proteases. High molecular weight albumin subunits are mainly β amylases (Gupta, 1996). All these enzymes provide nutrition to the growing embryo during germination by their hydrolytic and proteolytic actions.

(ii) ***Globulins***

Globulins are storage proteins that are located in protein bodies in the cells of the embryo and aleurone layer. They are classified into two groups, 7S and 11S globulins. Both these groups have very low concentration of sulphur containing amino acids (Shewry, 2000). The 7S globulins are comprised of three different subunits and have a total molecular weight of about 150,000-190,000 Da. They are located in the aleurone layer and their hydrolysis provides nutrition to the embryo during germination. The 7S globulins contain high amounts of glutamate, glutamine, arginine but lack cysteine and are thus unable to form disulphide bonds (Shewry, 2000).

The 11S globulin is also called triticin and is present in small quantities in the wheat endosperm cells. They form hexamers with a molecular weight of 60,000 Da and are polymeric storage proteins (Shewry, 1996). The S type globulins are the major storage proteins in legumes, oats and rice (Delcour, 2002). The large globulin subunits contain lysine residues and on proteolysis yield two types of subunits 20,000 Da (basic) and 40,000 Da (acidic) that are associated by intrachain disulphide bonds referred to as subunit pairs (Shewry, 1996).

The albumins and globulins together play only a minor role in the bread making potential of wheat flour. Low molecular weight cysteine rich proteins possibly influence the rheological properties of dough by sulphydryl/disulphide interchange reaction with gluten proteins (Kaufman, 1986). However, it has been shown that polymeric storage globulins are associated with low bread making performance (Veraverbeke, 2002). Most of these proteins are excluded from the flour during the milling process (Stone and Savin, 1999).

(B) Gluten proteins

Wheat gluten proteins are classified into two groups, gliadins and glutenins. Gluten comprises about 62-90% of the total wheat protein (Fig. 3). The gluten proteins are known as prolamins because of their high content of proline and glutamine amino acids and solubility in alcohol water mixtures (Shewry et al., 1995). These proteins form the major component of the white flour.

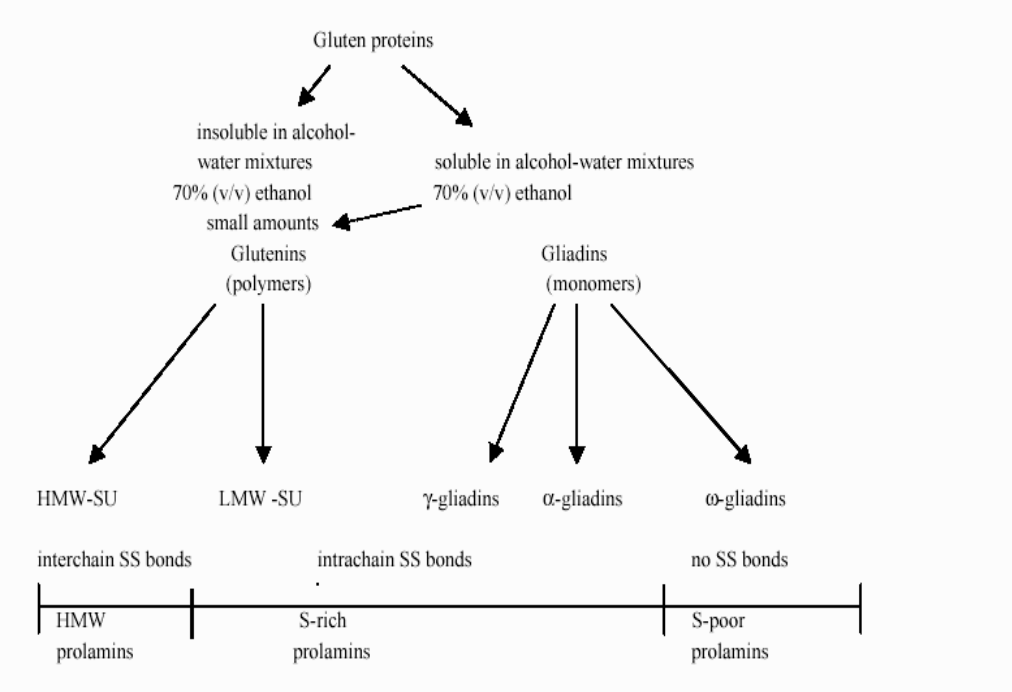


Figure 3. Classification and nomenclature of wheat gluten proteins (Shewry, 1996).

Bread wheat (*T. aestivum*) flour is widely used in the production of leavened and unleavened breads, Asian noodles, hard rolls and general purpose flour. The starch and gluten industries also employ wheat flour as raw material for their production. Wheat dough exhibits unique rheological (visco-elastic) properties that have been attributed to its storage proteins (glutenins and gliadins). Gluten proteins occur as simple monomers (gliadins) and polymers (glutenins) consisting of high molecular weight subunits (HMW-SU) and low molecular weight subunits (LMW-SU) linked by disulfide bonds (Fig. 3). Gliadins contribute to the viscous nature of the dough while the elastic properties are contributed by glutenins. Hence the ratio of gliadins and glutenins determine the balance between dough viscosity and elasticity thereby affecting its gluten protein content (Fig. 4). Variations in quality and quantity of gluten determine the dough properties that affect the bread making process (Shewry, 2000).

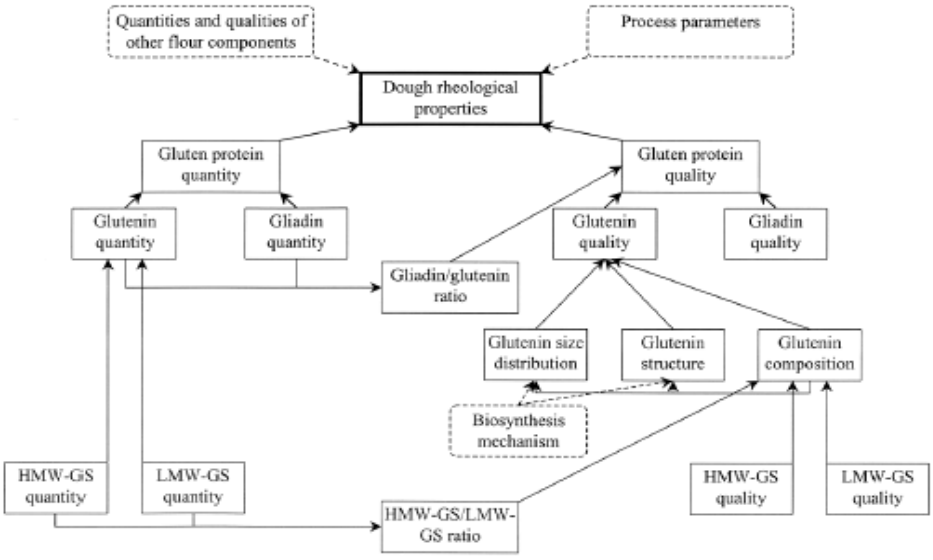


Figure 4. Factors governing wheat dough rheological properties. The high molecular weight glutenin subunits and low molecular weight glutenin subunits are represented as HMW-GS and LMW-GS, respectively (Veraverbeke et al., 2002).

When kneading or mixing flour with water, gluten proteins enable the formation of cohesive visco- elastic dough, capable of holding gas produced during fermentation and oven rise, resulting in the typical fixed open structure of bread after baking. The precise balance of viscosity and elasticity is important for quality of bread making. Insufficient elasticity of gluten leads to low bread volume while increased elasticity leads to higher loaf volume. Too elastic gluten impedes the expansion of gas cells, again leading to lower loaf volume. The elasticity of glutenin is believed to be mediated by covalent and noncovalent interactions between and within the glutenin chains. Gliadins act as plasticizers that weaken the interaction between the glutenin chains thereby increasing dough viscosity (Belton, 1999). The rheological properties of wheat flour can further be modified by the addition of oxidants, reducing agents, or proteases that directly change gluten proteins or by the addition of lipids/emulsifiers or hemicelluloses that can modify gluten protein interactions (Vereverbake, 2002).

(i) Gliadins

Gliadins are highly heterogeneous mixture of monomeric gluten proteins that are connected by intrachain polypeptide disulphide links. The cysteine residues in the polypeptide chains are involved in the formation of the disulphide bonds. Gliadins are low molecular weight gluten proteins of about 35,000 Da. The structure of gliadins is determined by the pH. They are arranged in a fibrillar form at pH greater than 5, and as partially unfolded monomers, at lower pH. Thus, the fibrillar structure seems to be enhanced by the increase in ionic strength (Shukla, 1975). Gliadins are classified into four groups (α-, β-, γ-, ω- gliadins) based on their amino acid sequence and electrophoresis mobility.

The α-, β-, γ- gliadins are similar in structure and have 250-300 amino acid residues (Wrigley, 1998) and are classified as S- rich prolamins due to their high content of amino acids, proline and glutamine (Fig. 5). The N-terminus of the protein is rich in proline while the C-terminus is rich in glutamine. Cysteine residues within the proteins help in stabilizing the protein structure (Shewry and Tatham, 1997). The non covalent protein-protein interactions of the gliadins are contributors to the viscosity and extensibility of gluten (Shewry, 2000). The ω- gliadins contain no disulphide bonds and have a different amino acid composition compared to other gliadins. They have small amounts of basic amino acids and are classified as S-poor prolamins. Hydrogen bonds within the ω- gliadins are responsible for their non covalent interactions in wheat dough (Shewry, 2000).

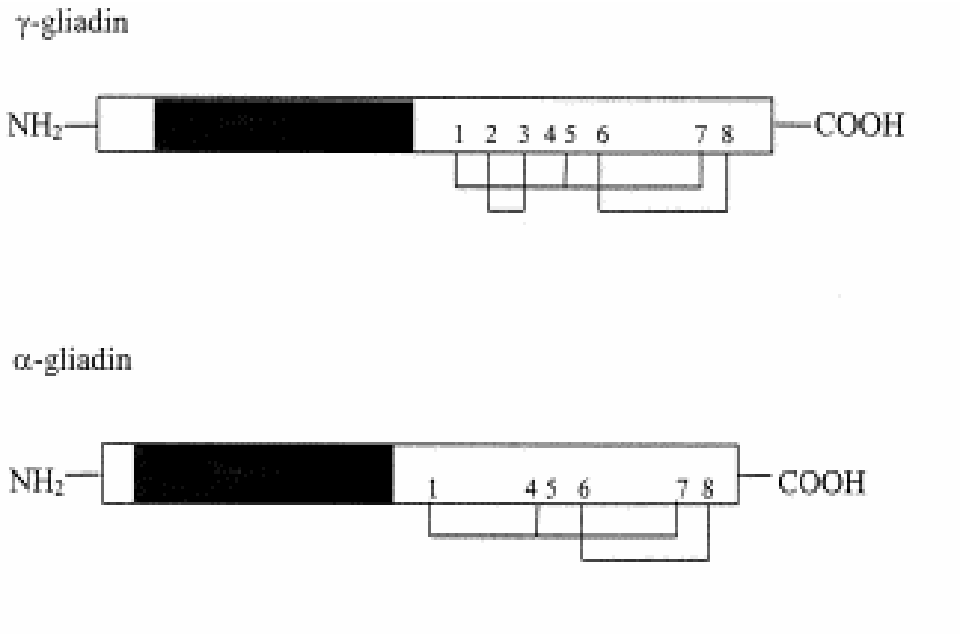


Figure 5. Structure of wheat gliadin protein. The numbers 1-8 represents the cysteine residues in gladins and the intra chain disulphide bonds formed by the cysteine residues have been represented by joining lines (Lindsay et al., 1999).

(ii) Glutenins

Glutenins consist of a mixture of polymeric proteins that are connected by interchain disulphide bonds (Shewry, 1997). The glutenins are insoluble in aqueous alcohol, but on reduction, the individual subunits dissolve in alcohol/water mixtures (Shewry, 1997). Glutenins are responsible for providing the elastic property to wheat dough. Glutenins can be separated into individual polypeptide subunits by reduction of their disulphide bonds. Based on their electrophoretic mobility in SDS-PAGE gels the subunits can be classified into two groups:

a) High molecular weight (HMW) glutenin subunits
b) Low molecular weight (LMW) glutenin subunits

(a) High molecular weight (HMW) glutenin subunits

The HMW glutenin subunits have a molecular weight ranging from 80-160,000 Da and contain low amounts of proline (Payne, 1997). They account for 5-10% of the total protein. SDS-PAGE analysis revealed about three to six different types of HMW glutenin subunits. The structure of HMW subunits has been studied by various spectroscopic (Circular Dichroism, Fourier Transform, Nuclear Magnetic Resonance) and microscopic methods. These studies have revealed that the individual subunit consists of 627-827 amino acid residues, with molecular weight values of 67,000- 88,000 Da. All the HMW subunits consist of central repetitive domain flanked by non repetitive domains at the N- and C- termini (Shewry, 1997). The N-terminal domain of the subunits consists of 81-104 amino acid residues while the C-terminus consists of 42 residues in all the subunit (Shewry, 1997). The length of the repetitive domain varies between 480-680 residues and

is responsible for the variation in the sizes of the proteins. Biophysical and computer generated structure predictions revealed that the repetitive domains form a β-spiral structure while the N and C termini exist as α-helices (Shewry, 1997).

The β- spiral structure of the repetitive domain of HMW subunits has been implicated in the biophysical and elastic properties of glutenin. It has been predicted that the sheet might act as a molecular spring that recoils back on stress (Fig. 6). Furthermore, it has a high content of glutamine residues that form strong hydrogen bonds with other subunits or proteins via their amide groups. Disruption of these bonds by shearing forces followed by reformation might also contribute to the elasticity of glutenin (Belton, 1994).

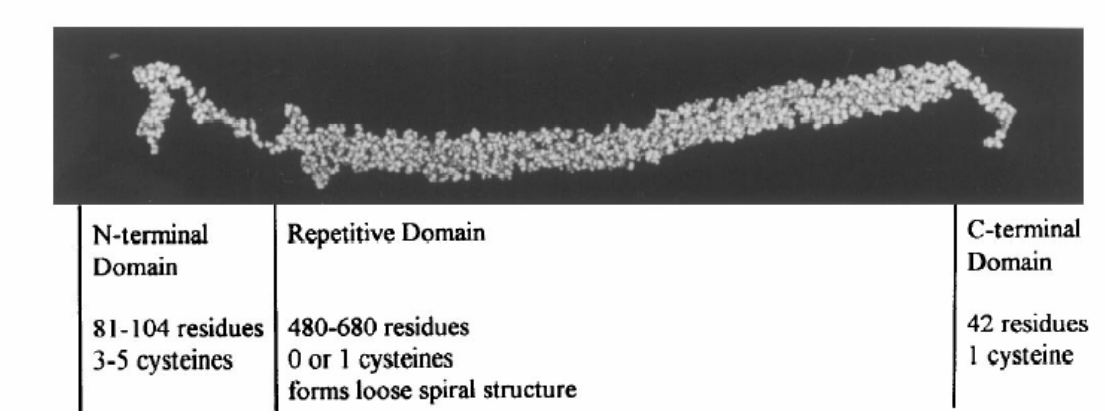


Figure 6. Computer predicted structure of wheat HMW glutenin subunit (Kasarda, 1994).

All types of hexaploid bread wheat cultivars have six HMW subunits of which only 3-5 are expressed. Two genes are present on the long arms of each of the group 1 chromosomes (1A, 1B, 1D). These encode two types of subunits X (1Ax, 1Bx, 1Dx) and Y type (1Ay, 1By, 1Dy) (Fig. 7). The allelic variation in the expression of the number of HMW leads to changes in the elastic property and quality of wheat flour. The polymers of HMW subunits are stabilized by formation of interchain disulphide bonds between the

cysteine residues. The number and the distribution of the bonds influence the size and biophysical properties of the subunits. HMW subunits amino acid sequence shows the presence of 3-7 cysteine residues. The C terminus of all HMW subunits has a single cysteine residue while the N-terminus contains either 3 (X type HMW subunits) or 5 (Y type HMW subunits) (Shewry, 1997). As there seems to be a dosage relationship between the number of expressed genes and the amount of HMW subunits, several labs are studying the improvements to quality of wheat flour, by addition of HMW subunits.

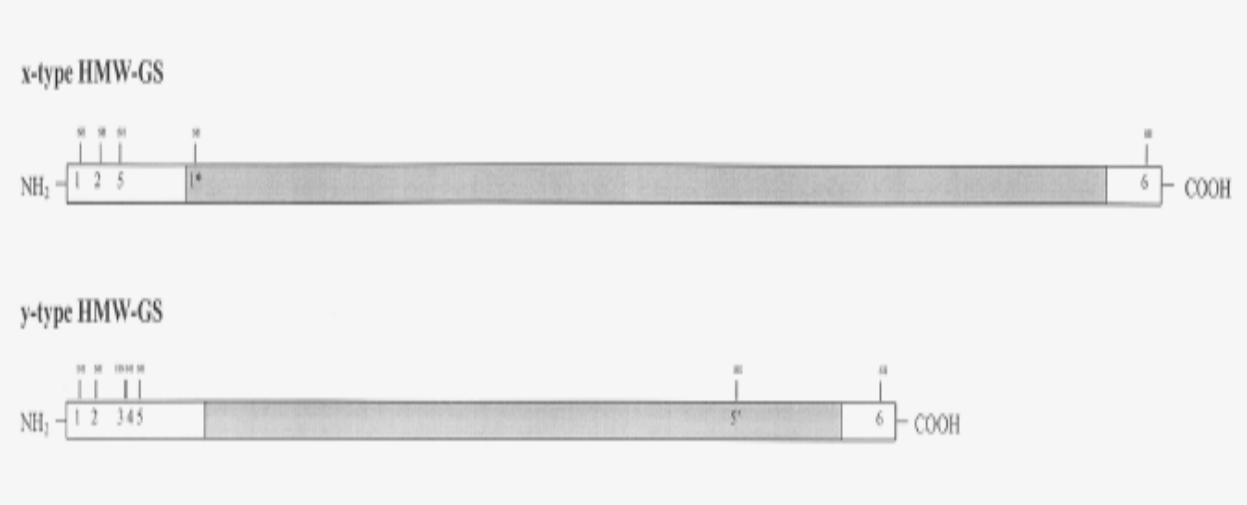


Figure 7. Schematic representation of structure of the HMW subunits. The numbers 1-6 represents the position of cysteine residues (Veraverbeke et al., 2002).

(b) Low molecular weight subunits (LMW subunits)

The low molecular weight subunits (LMW subunits) of wheat glutenin are polymorphic proteins that resemble in sequence to α-type gliadins. Based on their electrophoretic mobility in SDS-PAGE gels they have been classified into three separate groups A, B and C (Jackson et al., 1983). A and B type LMW subunits have a molecular

weight ranging from 40,000 Da - 50,000 Da, while C type LMW subunits range from 30,000 Da - 40,000 Da (Fig. 8). The LMW subunits belong to the B group and based on their N terminal amino acid sequence are divided into three sub groups called LMW-s, LMW-m and LMW-i according to the first amino acid residue of the mature protein serine, methionine or isoleucine, respectively (Kasarda, 1989). The LMW glutenin subunits have the ability to form intermolecular disulphide bonds with each other and also with HMW glutenin subunits. This property of the LMW glutenin is important for the elasticity of the gluten polymers and thus affects dough quality. LMW glutenins are also associated with dough resistance as evidenced by the reheological measurements (Veraverbeke et al., 2002). Different allelic forms of LMW glutenins seem to play different roles in determining the different quality parameters of wheat dough.

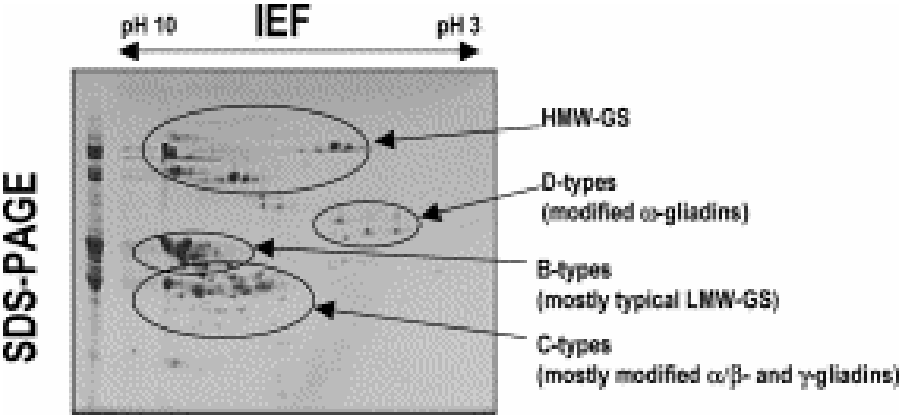


Figure 8. Analysis of wheat glutenin subunits by 2D gel electrophoresis (D' Ovidio, 2004).

The structure of LMW glutenin subunits has only been partially elucidated. The N terminal region of the protein forms beta turns whereas the C-terminus predominantly has alpha helices. (Thompson et al., 1987). Various modeling techniques have shown that the N-terminus seems to be highly repetitive and flexible while the C-

terminus is more compact. The repetitive subunit contains two cysteine residues that are involved in intermolecular disulphide bonds (Fig. 10) (Masci et al., 1998). These residues are predicted to be surrounded by regions of high flexibility, which might be a mechanism that facilitates polymerization (Fig. 9) (Ovidio et al., 2004). The remaining six residues form intra chain disulphide bonds and thus are responsible for stabilizing the structure of the LMW- GS (Orsi et al., 2000). Several site specific mutagenesis experiments of the first and the seventh cysteine have shown the importance of these cysteine residues in polymerization of gluten (Patacchini et al., 2003). Subunits of the B type LMW glutenins has been shown to form more polymers due to the presence of two cysteine residues in contrast to C-type subunits that have only one cysteine residue (Veraverbeke et al., 2000).

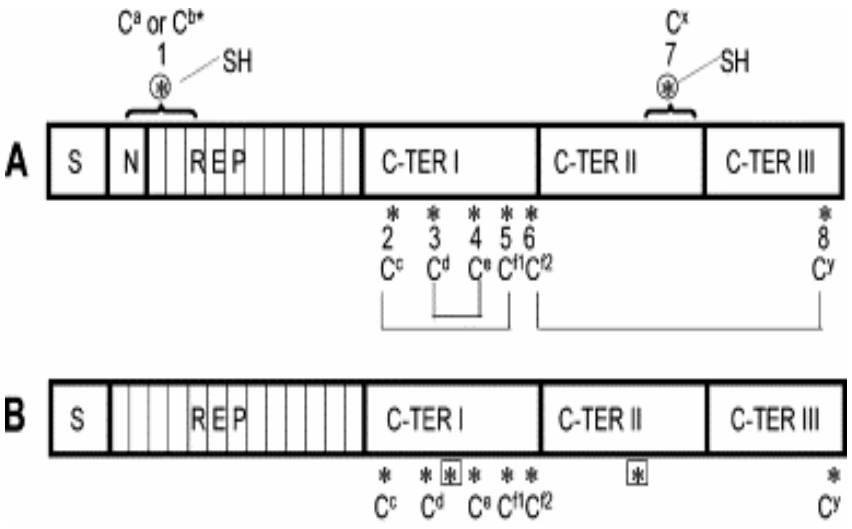


Figure 9. Diagrams representing the structures of typical LMW-GS as deduced from their encoding genes. (A) LMW-m and LMW-s type (B) LMW-i type. S: Signal peptide; N: N-terminal region; REP: repetitive domain (small boxes indicate repeat motifs); C-TER I, II, III: C-terminal regions. Cysteine residues are indicated by asterisks with sequential numbers. The asterisks encircled in (A) indicate the cysteines that are most likely to be

involved in inter-molecular disulphide bonds. Their position is also variable (as indicated by the brackets). Boxed asterisks in (B) indicate the two extra cysteine residues reported in the C-terminal domain of LMW-i type LMW-GS (D' Ovidio, 2004).

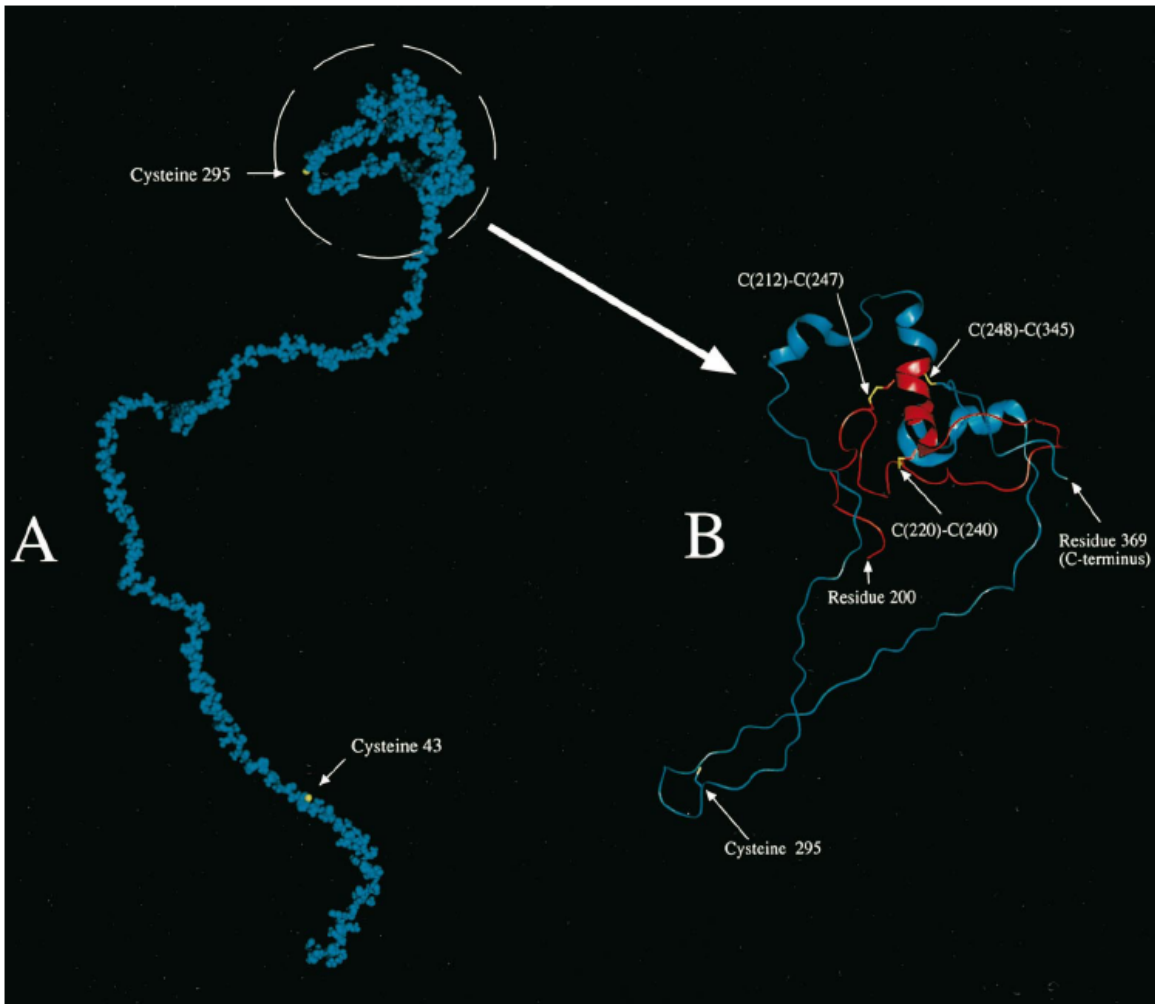


Figure 10. Computer generated model of the 42 K LMW glutenin sub unit space-filling form with all atoms shown in blue, except the sulfur atoms of Cys or cystine side chains, which are shown in yellow. B, Simplified model of the region containing the intramolecular disulfide linkages and Cys-295, which presumably forms one of two intermolecular disulfide cross-linkages. Residues 200 to 369 of the 42K subunit are shown in protein cartoon format, which displays the α-helical structure as a helical ribbon. The main polypeptide chain is shown in red for residues 200 to 248 and in blue

from there on to the C-terminal end at residue 369. The numbers of connected (intramolecular) Cys residues are shown (Masci et al., 1998).

The genes coding for the LMW subunits of wheat occur on the short arms of the group 1 chromosomes, at the *Gli-1* loci. This loci also contains the genes that code for the ω and, γ gliadins. (Payne et al., 1984). Genes for the B type LMW proteins undergo low levels of recombination with genes coding for gliadins resulting in designation of a separate locus for 1B-coded LMW glutenin subunits called, *Glu-B3*. The results from various labs have shown that the *Gli-1* locus occur at the same relative positions on the three homeologous chromosomes (Fig. 11). Southern blot analysis has shown that the gene copy number of LMW glutenin genes to be around 10-15 to 30-35 in hexaploid wheat (Cassidy et al., 1998).

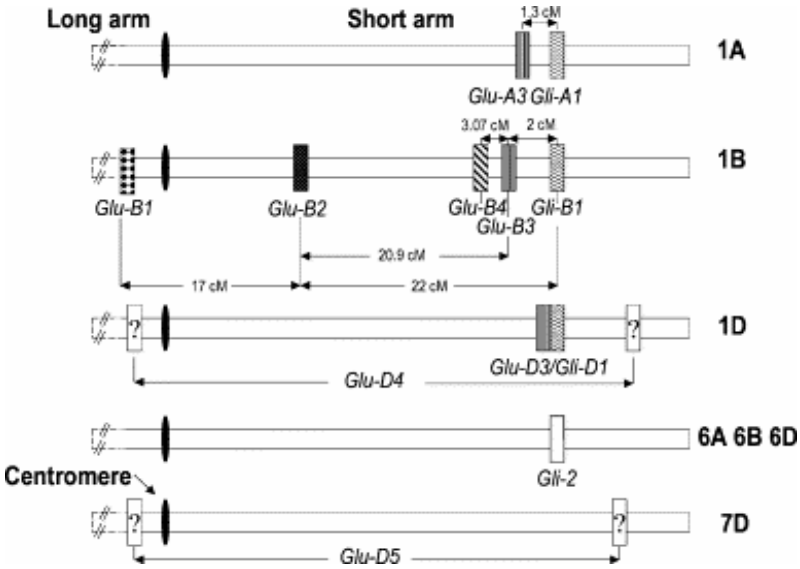


Figure 11. Chromosomal localization of genes encoding LMW-GS. Relative positions of loci are not proportional to their actual genetic distances, but are only indicative. The question mark indicates that the positions of the *Glu-D4* and *Glu-D5* loci are not known. (D' Ovidio, 2004)

Comparison between the coding regions of LMW glutenin genes has shown that they share more than 80% identity at both nucleotide and protein levels. LMW gluten proteins are synthesized on the rough endoplasmic reticulum and undergo folding in the lumen of ER, assisted by molecular chaperones. Expression of the LMW genes seems to be controlled at the level of transcription. All prolamin genes are regulated by trans-acting factors that bind to the cis-elements called the prolamin box or the endosperm box (300 bp) (Bartels et al., 1986). Environmental conditions like fertilizers seem to determine the gluten protein synthesis. Fertilizers rich in sulphur enhance the LMW glutenin subunits as they tend to be rich in sulphur containing amino acids while sulphur deficiency leads to negative effect (Shewry et al., 2001).

Structure of Gluten

Various models have been proposed on the structure of gluten polymer in dough. Earlier work has shown that only intra chain disulphide bonds are formed in glutenins. These bonds force the glutenin subunits to aggregate and form long chains of gluten polymer. Another model was proposed in which the polypeptide chains of glutenin were connected in a linear fashion with each other with a single disulphide bond (Ewart, 1979). Recent studies have shown that the gluten polymers consist of a mixture of dimers, polymers that are joined by both intrachain and interchain disulphide bonds. They have a molecular weight ranging from 50,000-100,000 Da. The size of the gluten polymers makes them insoluble in water (Fig. 12) (Wrigley, 1996). The amount and size of the polymeric protein is determined by a parameter referred to as % UPP (unsoluble

polymeric protein content). A high percentage of UPP offers a greater resistance to elasticity and mixing of dough (Gupta et al., 1993).

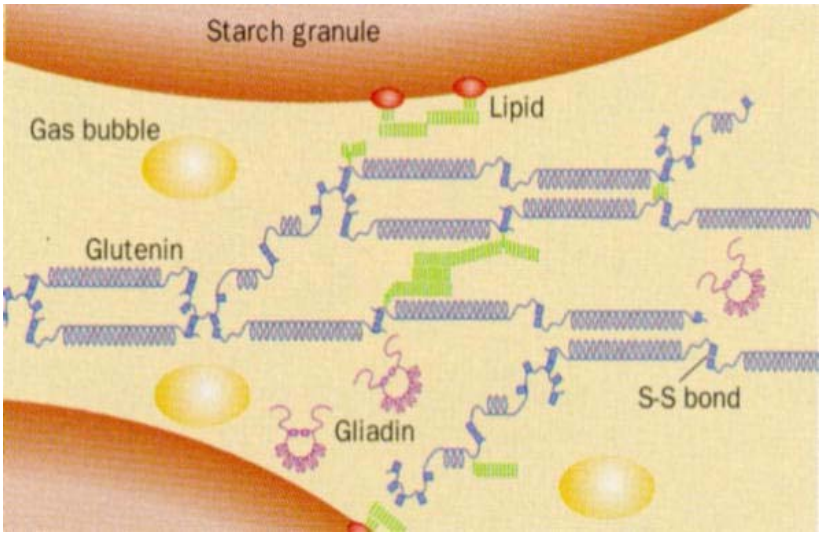


Figure 12. Model depicting the structure of gluten polymers in wheat dough (Wrigley, 1996). The disulphide bonded polymeric structure of wheat glutenin forming protein matrix between starch granules and gas bubbles in dough.

In the past decade more focus has been given to HMW-GS and its contribution to bread making quality and scientists have been successful in studying the genetic basis, expression and modification of the genes coding for these subunits. Only recently investigations have focused on LMW-GS, as it is attributed to the dough resistance and extensibility and it has been shown to have quantitative effects in many varieties of wheat.

The aim of the present study is to express a single type of LMW-GS protein in *E. coli* and purify the protein. The purified protein will be used to study its rheological and mixographic properties which are important parameters for determining the quality of

dough in the bread making industry. The gene encoding the Low Molecular Weight Glutenin 1D1 was synthesized by Integrated DNA Technologies (Coralville, IA) from the published sequence X13306 (Colot et al., 1989). This gene was then cloned in various bacterial vectors to achieve expression of the protein in large quantities for purification and studying the rheological properties of wheat dough.

Chapter II

Materials and Methods

Materials and Methods

LMWG1D1 Construct:

The gene encoding the Low Molecular Weight Glutenin 1D1 from *T. aestivum* was synthesized by IDT from the published sequence X13306 (Colot et al., 1989).

GAGGAAGCGGAAGGCGAGAGTAGGGAACTGCCAGGCATCAAACTAAGCAGAAGGCCCC
TGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTCTGTGTTGTAAAACGACGGCCAGT
CTTAAGCTCGGGCCTCAAATAATGATTTTAGATTAAgagctcgaacaacaacaacaataacaataacaacaac
ctcgggatcgagggaaggatgaaaacgttcctggtctttgcgctgctggcagttgcggccacctccgcaatcgcgcagatggaaacccgctgtatt
cctggcctggaacgtccgtggcagcagcaacctctgccgccgcagcagacctttccgcaacagccgctgttcagccaacagcagcaacaacagc
tgtttccgcagcagccgtctttctctcagcagcaaccgccgttttggcaacaacaaccgccattctctcagcaacagccaattctgccgcagcagcc
gccgttttcccaacagcaacagctggttctgccacagcaaccgcctttcagccaacaacagcaaccggttctgccgcctcagcagtctccttttcca
caacagcagcagcagcaccagcaactggtacaacaacagatcccagttgtgcagccaagcattctgcaacaactgaatccttgtaaagtgtttct
gcagcagcagtgtagcccggtggctatgccgcagcgtctggcgcgctcccaaatgctgcaacaatcttcttgccatgtcatgcaacagcaatgttgt
caacagctgccgcagatccctcagcaaagccgttatgaggctattcgcgcgattatttattccattattctgcaggaacagcagcaggtacaaggtt
ccatccagagccagcagcaacagccgcaacaactgggccaatgcgtgagccagcctcagcaacaatctcagcagcagctgggccagcaaccgc
aacaacagcagctggcccagggtacttttctgcagccgcatcagatcgctcagctggaagtaatgacctctattgcgctgcgcattctgccaactat
gtgttctgtgaatgtgcctctgtatcgtacgacgaccagcgtgccgtttggtgttggcaccggcgtgggcgcatactaaaagcttTTAATCGC
CATCCAGCTGATATTCCCTATAGTGCATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCG
TGTCTCAAAATCTCTGATGTTACATTGTACAAGATAAAATAATATCATCATGAACAATAA
AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAA
ACGTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGC
TCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTGTATGGGAAGCCCGATG
CGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAG
ATGGTCAGACTAAACTGGCTGACGGAATTTATGCCACTTCCGACCATCAAGCATTTTATC
CGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGAAAAACAGCGTTCCA
GGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCT
GCGCCGGTTGCACTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCG
CCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATG
ACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAACTTTTGCCA
TTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGAC
GAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCA
GGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCT
TTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTC
GATGAGTTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTG
ACTTGACGGGACGGCGCAAGCTCATGACCAAAATCCCTTAACGTGAGTTACGCGCGCGT
CGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTT
TTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTT
TGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAG
ATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGCCCACCACTTCAAGAACTCTGTA
GCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGT
CGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGA
ACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGG

CGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCC
AGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGC
GTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCG
GCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTAT
CCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGC
AGCCGAACGACCGAGCGCAGCGAGTCAGTGAGC.

Figure 13. The synthesized gene from IDT cloned in pJ plasmid. The nucleotide sequence in small letters represents the GLUIDI inset and the remaining sequence from the plasmid.

The gene was supplied in pJ plasmid in XL1Blue strain of *E. coli* which is isolated and then digested with *Sac*I and *Hind*III. The insert is further cloned into pMAL vector series, pET 21a and pQE 30 vectors.

Reagents, Chemicals, Vectors and Enzymes

Methanol, ethyl alcohol, acetic acid, EDTA, di-Sodium hydrogen phosphate, potassium-di-hydrogen phosphate, hydrochloric acid, sodium chloride, glycerol, ethidium bromide, β-mercaptoethanol, agarose, acrylamide, bis-acrylamide, triton X, ammonium per sulphate (APS), kanamycin, choromphenicol, ampicillin were purchased from Sigma Chemical Company.(St Louis, MO). Restriction enzymes, T4 DNA ligase and Taq DNA polymerase, were purchased from Recombinant DNA and protein core facility at the Biochemistry Department and Molecular Biology, Invitrogen (Carlsbad, CA) and Promega (Madison, WI). Plasmid miniprep kit, PCR purification kit, nucleic acid extraction kits were purchased from Qiagen (Valencia, CA). pMALc2X vectors were purchased from New England Biolabs, pET21a vector from Novagen (San diego, CA) and pQE30 vector from Qiagen, respectively.

IPTG & X-gal preparation

IPTG stock solution was prepared by dissolving 2 gm in 8 ml of water and adjusted to 10 ml and then filtered through a 0.22 micron disposable filter and stored at -20°C. To induce the cultures 0.5mM of IPTG was used. Stock solution of X-Gal was prepared by dissolving X-Gal in dimethylformamide (DMF) to make 20 mg/ml. The stock solution was then stored in a glass or polypropylene tube with aluminum foil and stored at -20°C. For 250 ml of LB agar media 300 µl of IPTG was used.

Preparation of bacterial growth media

Bacteria were cultured in Terrific broth, LB broth and LB media. To make 1 L of Terrific broth, 12 gm of Bacto-tryptone, 24 gm of Bacto-yeast extract, 4 ml of Glycerol were dissolved in water. The solution was made to 900 ml and pH was adjusted to 8.0 and autoclaved. Salt solution was prepared by dissolving 2.31 gm of KH_2PO_4 and 12.54 gm of K_2HPO_4 in 100 ml of water. Autoclaved media and salt solution were mixed prior to inoculating the media.

To make 1 L of LB media, 10 mg of Bacto tryptone, 15 mg of Bacto yeast, and 10 mg of NaCl were dissolved in water to make up to 1000ml. The pH was adjusted to 8.0 and then 15 gm of agar was added and autoclaved.

Preparation of Antibiotic stocks

Antibiotics were prepared as follows: 25 mg of ampicillin was dissolved in sterile water to a make a stock solution with a concentration of 25 mg/ml, and filtered through a 0.22 micron filter using disposable syringe. For 250 ml of LB media 250 µl of the ampicillin stock was added.

Chloramphenicol stock solution was prepared by dissolving 35 mg of chloramphenicol (Sigma) in ethanol to a final concentration of 35 mg/ml. For 250 ml of LB media 350 µl of chloramphenical was added.

Stock solution of kanamycin was prepared by dissolving 25 mg of kanamycin in sterile water to a final concentration of 25 mg/ml. The solution was then filtered through a 0.22 µM filter using a disposable syringe.

Small scale plasmid DNA isolation (miniprep) from recombinant colonies

Several recombinant bacterial colonies, which appeared white on LB plates, were picked randomly and grown in 5 ml Terrific broth (containing 100 µg/ml ampicillin) at 37°C overnight with vigorous shaking. Small scale preparation of plasmid DNA was done using Qiagen spin miniprep kit as per the protocol supplied by the manufacturer.

Briefly, transformed recombinant bacteria (5 ml) were harvested by centrifugation at 10,000 g for 2 min and resuspended in 200 µl of P1 buffer containing RNase A (supplied with the kit). Bacteria were then lysed and DNA was extracted by mixing gently with 250µl of P2 buffer followed by neutralization with 350 µl of N3 buffer. The genomic DNA/protein was then precipitated by centrifugation at 10,000g for 10 min and the supernatant containing the plasmid DNA was loaded into a Qiagen mini spin column. The spin column was centrifuged at 13200 rpm (Eppendorf centrifuge 5415D). Bound plasmid DNA was washed with 750 µl of buffer PE (supplied in the kit) by centrifugation. The spin column was then placed in a 1.5 ml microcentrifuge tube and bound DNA was eluted from the membrane by adding 50 µl of Buffer EB, elution buffer (10 mM Tris.Cl, pH 8.5) and centrifuging it at 10,000g for 1 min. The eluted DNA was

stored at 4°C. The amount of plasmid DNA was quantified spectrophotometrically, and the presence of insert within the plasmid was determined by digesting 2 µg of plasmid DNA with restriction enzyme at 37°C for 2 hrs. Undigested and restriction enzyme digested plasmid DNA were analyzed on a 1% agarose gel to identify the recombinant plasmid.

S.No.	Primer designation	Primers	Length (bp)
1.	pMAL Fwd pMAL Rv	5′-GAG**GAATTC**ATGAAAACGTTCCTG-3′ 5′-GAT**AAGCTT**TTAGTATGCGCCCACGCC-3′	24 27
2.	pNDE	5′-ATAG**CATATG**AAAACTGAAG-3′	20
3.	pET Fwd pET Rv	5′-GAG**GAATTC**ATGAAAACGTTCCTG-3′ 5′-GAT**AAGCTT**TTAGTATGCGCCCACGCC-3′	24 27
4.	pQE Fwd pQE Rv	5′-CGGATAACAATTTCACACAG-3′ 5′-GTTCTGAGGTCATTACTGG-3′	20 19

Table 2. Primers designed to amplify the insert and to sequences the recombinant clones. The restriction enzyme sites are represented in bold.

Amplification of insert DNA by PCR

The DNA insert was amplified with primers (Table 2) by PCR using a *Taq polymerase* (NEB, Promega). For 50 µl of PCR reaction, 2 µl of DNA template, 50 µl of each primer, 5 U of *taq* polymerase, 5 µl of 10X PCR buffer, 25 mM each dNTPs, were added and subjected to 30 cycles of PCR. The denaturation of plasmid DNA was

done at 95 ^{o}C for 2.0 min followed by annealing at 50^{o}C for 30 s and elongation at 72^{o}C for 1.0 min in an Eppendorf Mastercycler. An aliquot (10 µl) of PCR product was checked by 1% agarose gel electrophoresis for the size of the amplified products. The DNA bands of interest were excised and eluted by using the QIAquick gel Extraction kit. The eluted DNA was stored at 4°C. The amount of plasmid DNA was quantified spectrophotometrically, and the presence of insert within the plasmid was determined by digesting 2 µg of plasmid DNA with restriction enzyme at 37°C for 2 hrs. Undigested and restriction enzyme digested plasmid DNA were analyzed on a 1% agarose gel to identify the recombinant plasmid.

End filling of digested plasmid with Nde1

End filling of the 5' plasmid over hangs after digestion with *Nde1* was done by using using the Klenow fragment (Invitrogen) which catalyses and removes the 3' and fills in at the 5' to form blunt ends. The reaction mixture was set up using 5 µl of DNA (2 µg), 1 µl of each dNTP (100 mM), 3 µl of REACT 2 buffer, 2 µl of Klenow Polymerase and 16 µl of distilled water was added to bring the final volume of the reaction to 30 µl.

Agarose gel electrophoresis

All the plasmids were detected by agarose gel electrophoresis using a 1 % agarose gel in TAE buffer. The gels were stained in ethidium bromide for 2-5 mins and washed and analyzed in a UV transilluminator (Gel Doc) system. The 1 kb ladder from Invitrogen was dissolved in loading buffer and used as a marker.

SDS gel electrophoresis

Protein analysis was done by SDS-12 % PAGE. Acrylamide gels were run in Tris-Glycine buffer (Sambrook et al., 1989). The protein samples were prepared by adding sample buffer and boiled at 95^{o} C for 5 mins and then loaded on the gel. The gels were stained using Coomassie stain (0.5 mg of Coomassie in Methanol and acetic acid). The gel was then destained in a solution of methanol, acetic acid and water in a ratio of 4:1:5(v/v/v). The protein marker was procured from BioRad (Hercules, CA) and dissolved in sample buffer in 1:20 (w/v) ratio

Expression and purification of recombinant protein

The selected clones were grown in TB media till an O.D (optical density) of 0.6 was reached. The cells were then induced with 0.5 mM IPTG and grown over night at 18^{o} C. The cells were centrifuged at 4500 rpm. The cell pellet was dissolved in column buffer (NEB) (20 mM Tris-HCl, 200 mM NaCl, 1mM EDTA and 10mM β - Mercaptoethanol for 1 L of the buffer). This buffer was used throughout the purification process. To get the desired protein in soluble form the column buffer was replaced with GUS extraction buffer (Bender's lab protocol) consisting of 1M $NaHPO_4$ pH 7.0 [Na_2HPO_4 – 57.7 ml + NaH_2PO_4 – 42.3 ml = 1 M of stock] , 0.70 ml β – Mercaptoethanol, 0.5 M NaEDTA pH 8.0 30% Sarkosyl, 10.0mL Triton X 100 and 916 ml Distilled water. After eluting from the resin the protein was concentrated using Millipore concentrator and then cleaved with the specific protease in the cleavage buffer (NEB).

Chapter III

Results and Discussion

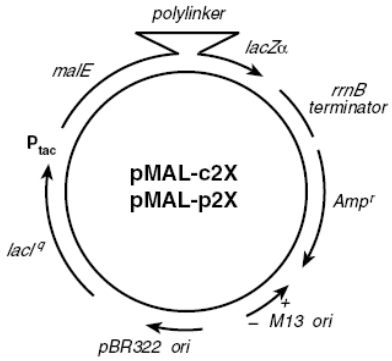
Results and Discussion

In order to get significant amounts of low molecular weight gluten needed for studying its rheological properties we cloned the LMWG1D1 gene in different cloning vectors and checked for the expression of low molecular weight glutenin protein 1D1. The protein yield and protein expression were analyzed.

(a) Cloning and expression of LMWG1D1 protein in pMALc2X vector

The 975 bp LMWG1D1 gene was cloned at the *Sac*I and *Hind*III sites of the pMALc2X vector (Fig. 14a) multicloning site and transformed into *E. coli* DH5α competent cells. The transformed cells were grown on LB plates with Ampicillin (100 μg/ml and X-gal (40 mg/ml) and the recombinant white colonies were picked and the plasmids were isolated. The plasmids were screened for the presence of insert DNA by digesting with *Sac*I and *Hind*III restriction enzymes. The plasmids containing the insert of desired size were further confirmed by sequence analysis at the Recombinant DNA and protein core facility of the Biochemistry and Molecular Biology department.

14 (a)



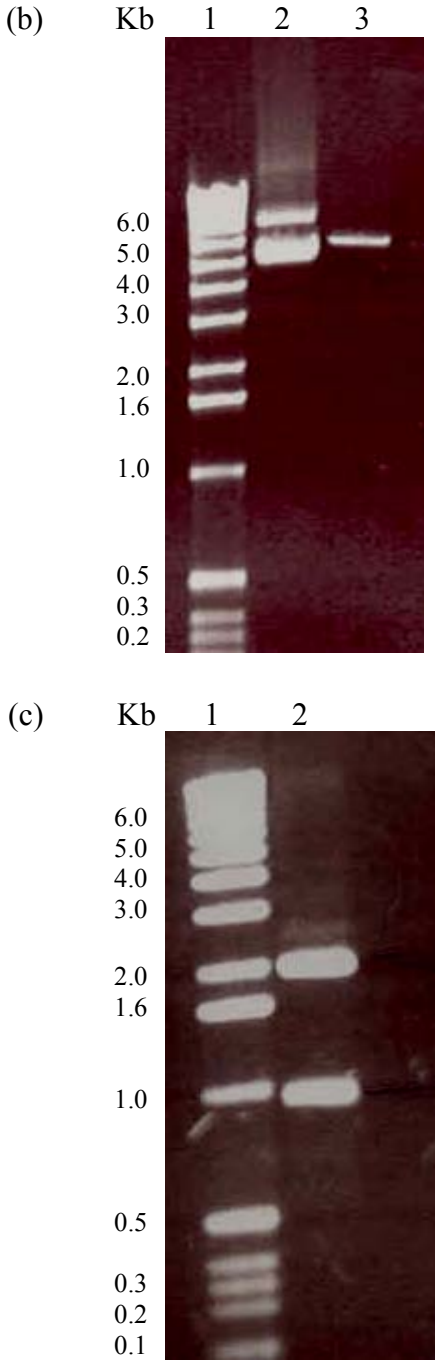


Figure 14 (a) Restriction map of pMal-c2X vector. (b) Analysis of pMaLc2X vector by restriction digestion using *Sac*I and *Hind*III enzymes. lane 1, 1 Kb DNA ladder; lane 2, undigested vector; lane 3, double digested vector. (c) Double digestion of pJ plasmid with insert. lane 1, 1 Kb DNA ladder; and lane 2, Double digested vector with insert (~1 kb).

For expression of the LMWG1D1 protein, the recombinant plasmids were transformed in *E. coli* Rosetta cells and grown in LB broth containing 25 µg/ml ampicillin and 35 µg/ml of chloramphenicol to an OD of 0.6. The cells were induced with 0.5 mM IPTG and incubated overnight at 18°C and harvested the next day. LMWG1D1 was expressed as a fusion protein with the N- terminus maltose binding protein (MBP) tag. The harvested cells were sonicated and the crude extract was loaded on to an amylose column. Proteins fused with the MBP-tag were bound to the resin and later eluted twice with buffer solution containing 10 mM maltose. The eluted protein samples were analyzed using 12% acrylamide SDS-PAGE (Fig. 15). The concentration of the fusion protein was determined using Nanodrop spectrophotometer, Nanodrop Technologies (Wilmington, DE). To obtain pure LMWG1D1 protein, the MBP-fusion tag was cleaved from the extracted fusion protein by using Factor Xa protease (Fig. 16).

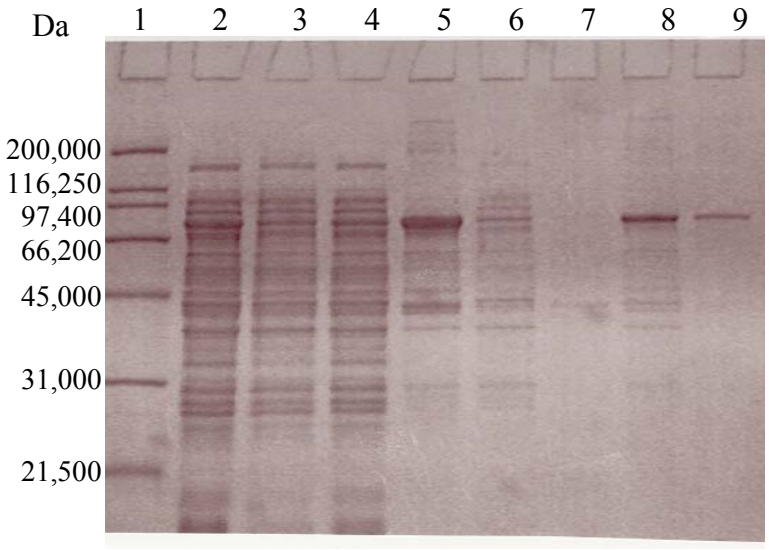
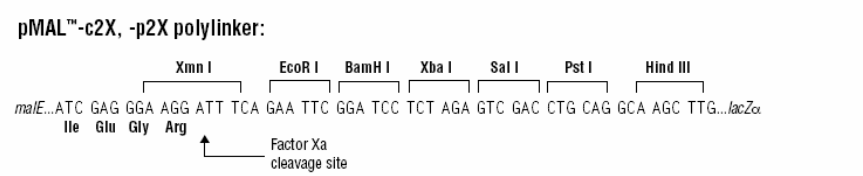


Figure 15. Analysis of LMWG1D1 expression in *E. coli* Rosetta cells using 12% acrylamide SDS- PAGE: lane 1, protein marker; lane 2, induced *E. coli* total cell lysate; lane 3, induced *E. coli* cell pellet dissolved in buffer; lane 4, induced *E. coli* cell supernatant; lane 5, MBP-LMWG1D1 protein bound to amylose column; lane 6,

supernatant after binding; lane 7, wash; lane 8, amylose resin after protein elution; lane 9 eluted LMWG1D1 fusion protein.

16 (a)



(b)

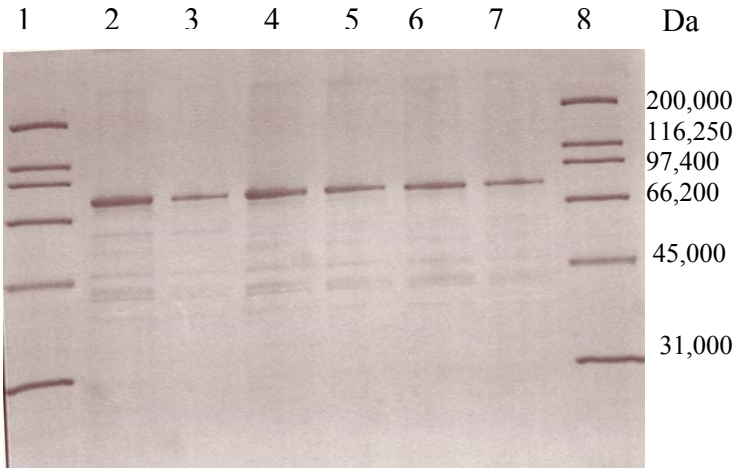


Figure 16. (a) Polylinker and Factor Xa cleavage site of pMAL-c2X vector. (b) Analysis of Factor Xa protease digested MBP-LMWG1D1 fusion protein, using 12% acrylamide SDS-PAGE: lane 1&8, Marker; lane 2, cleaved fusion at RT (50% enzyme); lane 3, cleaved fusion at 37^{o}C; lane 4, cleaved fusion at RT (100% enzyme); lane 5, cleaved fusion at 37^{o}C; lane 6, fusion in buffer at RT (control); lane 7, fusion protein in buffer at 37^{o}C.

Though we were able to obtain a good level of the protein expression (500 ng/ml of fusion protein from 20 ml of culture), cleavage of the MBP-tag from the fusion protein was not detected. Varying the cleavage conditions like incubation time and temperature

also could not improve the cleavage of MBP from the fusion protein as evident from (Fig. 16b). The above approach failed to cleave the MBP-tag from the fusion protein, as the folding of the fusion protein might have made the MBP-cleavage site inaccessible to protease. Cloning was done with the protease site incorporated in the insert during synthesis of the insert gene by IDT. It was hence hypothesized that a change in the fusion tag cleavage site and location might increase the chances of protein release. Hence the expression of LMWG1D1 was attempted by cloning in pMal-c2E and pMal-c2G vectors (NEB) with different fusion tag cleavage sites in vector for Enterokinase and Genenase respectively. The insert was modified with different restriction sites by PCR using primers 1 (Table 2).

(b) Cloning and expression in pMal-c2E and pMal-c2G vectors

The LMWG1D1 gene was PCR amplified and cloned at *EcoR*I and *Hind*III in pMal-c2E and pMal-c2G vectors. The vectors include a sequence coding for the recognition site of a specific protease. This allows the protein of interest to be cleaved from MBP after purification, without adding any vector-derived residues to the protein the polylinker includes a restriction site superimposed on the sequence coding for the site of the specific protease. This is where the gene of interest is inserted using primers1 (Table 2). *E.coli* DH5α competent cells were transformed using these vectors and recombinant plasmids were isolated and digested with *EcoR*I/*Hind*III. The digested plasmids were analyzed by 1% agarose gel electrophoresis.

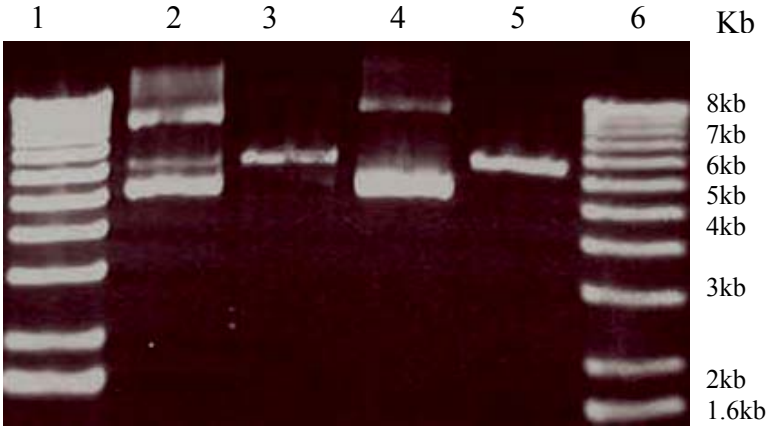
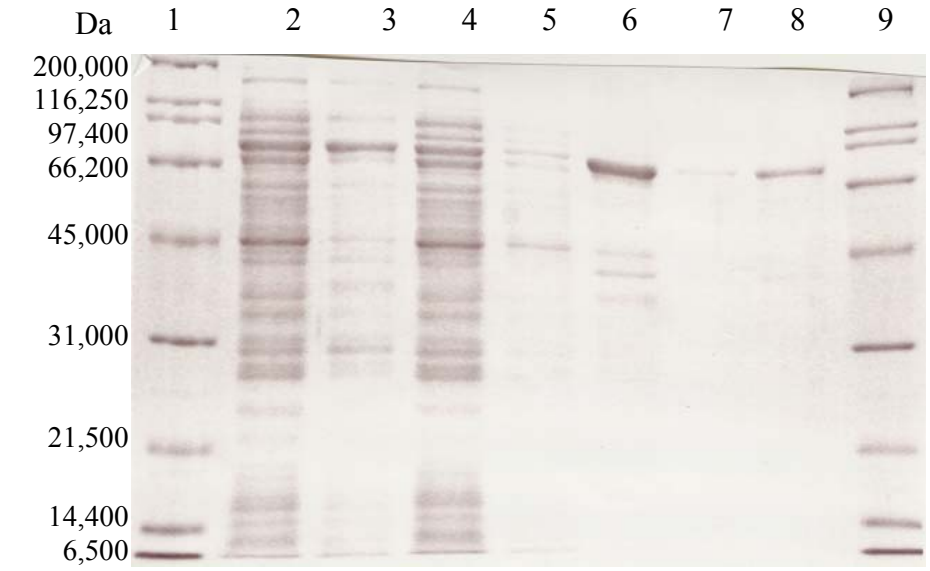


Figure 17. Restriction digestion of Pmal-c2E and pMAL-c2G vectors using *EcoR*I and *Hind*III enzymes. lane1 and 6, 1 Kb ladder; lane 2, undigested pMAL-c2E; lane 3, double digested vector with *EcoR*I and *Hind*III; lane 4, undigested pMAL-c2G; lane5, double digested vector with *EcoR*I and *Hind*III.

The selected recombinant clones containing the desired insert were then transformed into competent *E. coli* Rossetta cells and expression of the LMWG1D1 protein was conducted using the method described earlier. The eluted fractions were then run on 12% acrylamide SDS-PAGE to check for the expression and recovery of the fusion tag protein (Fig. 18).

18 (a)



18 (b)

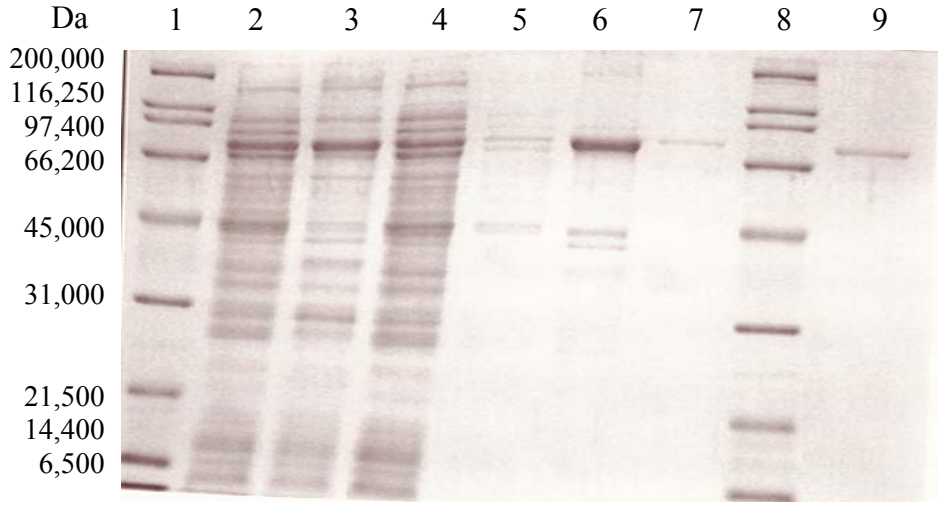


Figure 18. (a) Analysis of expression of LMWG1D1 in pMal-c2E vector using 12% acrylamide SDS-PAGE. Lane1 and 9, protein markers; lane 2, *E. coli* total cell lysate; Lane 3, *E. coli* cell pellet dissolved in buffer; lane 4, supernatant after binding; lane 5, wash fraction; lane 6 resin bound to MBP; lane 7, resin after elution; lane 8, eluted fusion protein. (b) Analysis of expression of LMWG1D1 in pMal-c2G vector using 12% acrylamide SDS-PAGE. lane1 and 8, protein markers; lane 2, crude extract; lane 3, pellet; lane 4, supernatant after binding; lane 5, wash; lane 6, resin bound to MBP; lane 7, resin after elution; lane 9, eluted fusion protein.

Figures 18 a and b show that we were able to detect the expression of the fusion protein. The concentration of the protein was estimated to be 2.2 mg/ml for pMAL-c2E vector and 2.4 mg/ml for pMAL-c2G vector from 100 ml of culture. The expressed fusion tag protein was then cleaved using Enterokinase and Genenase cleavage sites in pMal-c2E and pMal-c2G. Enterokinase is a specific protease that cleaves after lysine at its cleavage site Asp-Asp-Asp-Asp-Lys. Genenase I cleave His-Tyr-Glu and His-Tyr-Asp slowly, but will not cleave His-Tyr-Pro or His-Tyr-Ile. These samples were then analyzed by 12% acrylamide SDS-PAGE (Fig. 19).

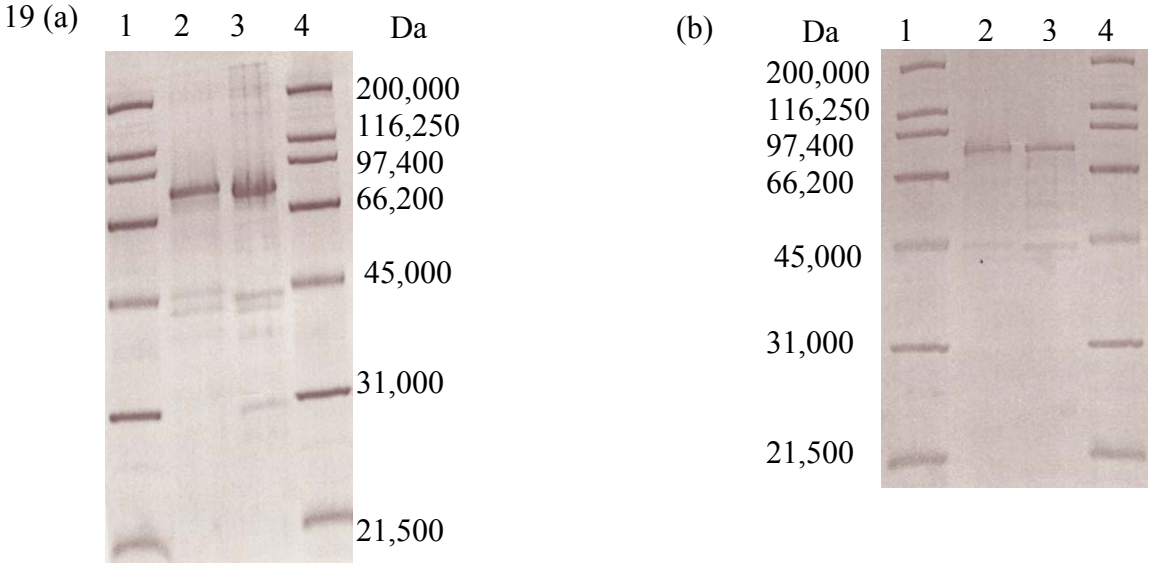


Figure 19. (a) Analysis of LMWG1D1-MBP fusion protein cleavage with Enterokinase. lane 1 and 4, Protein marker; lane 2, Cleaved fusion protein at RT; lane 3, Cleaved fusion protein at 37°C. (b) Analysis of LMWG1D1-MBP fusion protein cleavage with Genenase. lane 1 and 4, protein marker; lane 2, cleaved fusion protein at RT; lane 3, Cleaved fusion protein at 37°C.

Analysis of the Fig. 19 shows that the cleavage of the MBP-tag from the fusion protein did not occur. Hence this strategy also was not successful in obtaining a homogenous low molecular weight gluten protein. The MBP protein might be helping in solubulization of fusion protein which might form aggregates and hence cleavage of fusion tag was not achieved.

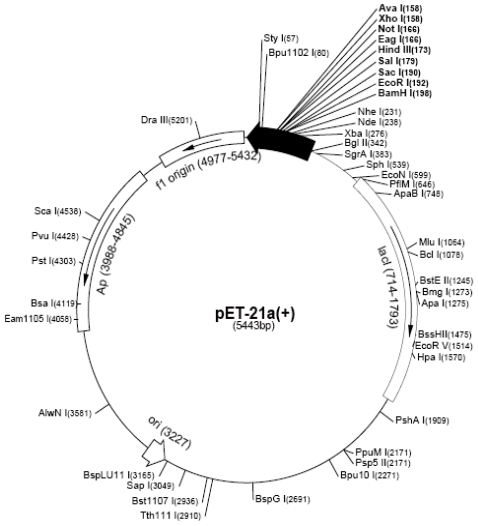
Since both the above strategies to cleave the MBP- tag did not work we decided to express the protein with an N-terminus His – tag under the control of a T7 promoter. In order to achieve this we cloned the LMWG1D1 gene in a pET21a vector.

(c) Expression and cloning in pET21a vector

The pET21a vector was restriction digested with *EcoR*I and *Hind*III enzymes and LMWG1D1 insert was cloned at these sites (Fig. 20a). The cloned vector was then transformed into *E. coli* DH5α cells and the recombinants were selected by their growth on LB ampicillin and X-gal plates. The plasmids obtained from the colonies were then checked for the presence of desired insert by double digestion with *EcoR*I/*Hind*III enzymes and then analyzed by 1% agarose gel electrophoresis (Fig. 20 b).

Expression of the protein was done by transforming the recombinants in *E. coli* BL21 competent cells. BL21 cells are high performance host cells that provide higher mRNA stability and tighter expression of the protein.

20 (a)



(b)

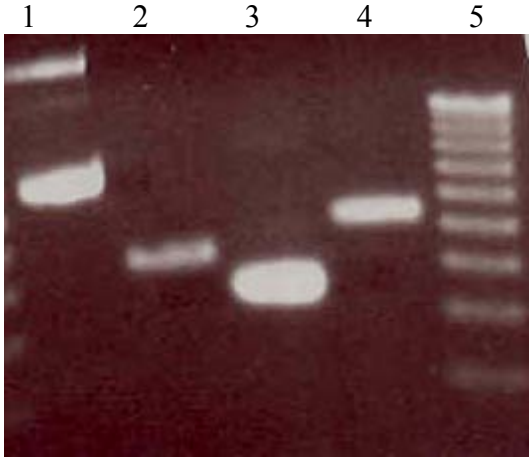


Figure 20. (a) Restriction map of pET21a vector. (b) Restriction digestion pET21 a using *EcoR*1 and *Hind*III enzymes. lane 1, undigested pET21a; lane 2, *Hind*III digested pET vector; lane 3, undigested clone in BL21 competent cells; lane 4, clone digested with *Hind*III; (vector showing ~1kb difference) lane 5, 1 Kb ladder.

Induction of the tagged protein was done with 0.5 mM of IPTG and incubated overnight at 18°C as described as before. The cells were then lysed and the extract was purified using a nickel-NTA column (NEB) following the manufacturers instruction. The column bound proteins were then eluted by a column buffer and analyzed by 12% acrylamide SDS-PAGE (Fig. 21).

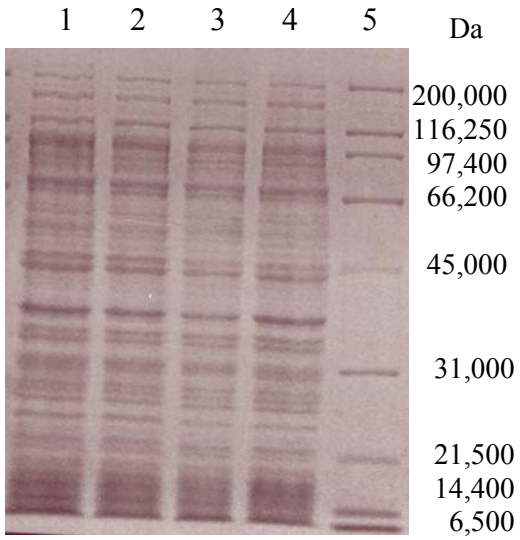


Figure 21. Expression analysis of LWMG1D1 protein in pET21a vector. lane 1, uninduced *E. coli* cells; lane 2, induced *E. coli* cells with 0.5mM IPTG; lane 3, uninduced *E. coli* cells; lane 4, induced *E. coli* cells with 1mM IPTG; lane 5, protein marker

Gel analysis showed similar pattern of expression in all the samples eluted from the column. As we could not detect any expressed protein of 32,000 Da size, we concluded that the LMWG1D1 protein was not able to be expressed in this vector. This experiment was repeated by varying the IPTG concentration (0.1, 0.5, 1.0, 1.5 and 2.0 mM) and temperature (18°C, 26°C and 37°C) using suitable controls, but we were unable to detect the expression LMWG1D1 protein.

(a) Expression and cloning in pMal vector with deletion in mal E gene

All the above strategies involving cleavage of LMWG1D1 protein from the MBP fusion tag and His-tag purification with T7 promoter failed to over express the desired protein. We decided to express the protein without a tag under the control of *tac* promoter which is a high efficiency promoter. The *tac* promoter, a hybrid

of two strong promoters (trp and lac) is effective because it corresponds to the E. coli consensus -10 and -35 sequences, tataat and ttgaca. As we already achieved good amounts of expression in pMAL vector series we deleted the malE gene that codes for the maltose binding domain from the pMal vector and attempted to express the protein without a tag.

The malE gene was removed from the vector by restriction digestion using *Nde*I and *EcoR*I enzymes (Fig. 22). The cut vector was then ligated using a klenow polymerase. Glu1D1 DNA insert was then cloned in this vector and transformed into *E. coli* DH5α cells. The recombinants were transformed in to *E. coli* Rosetta cells and were analyzed for the presence of any expressed protein as per method described earlier.

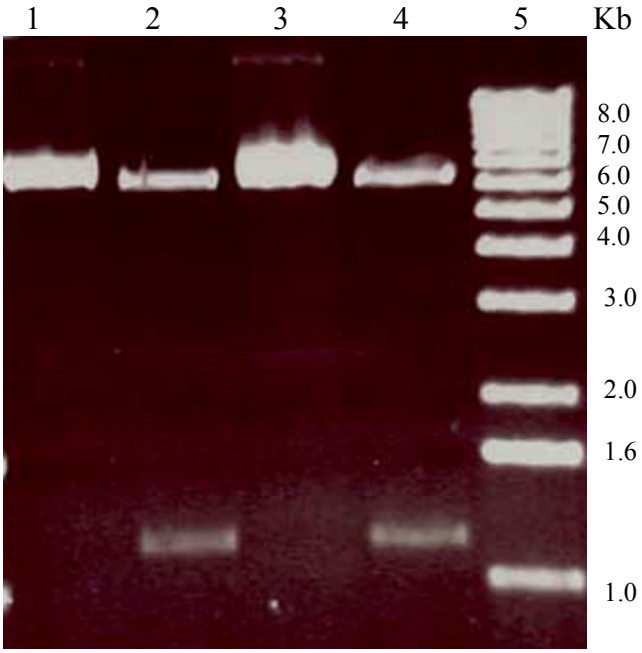


Figure 22. Double digestion of pMAL vectors to remove the malE gene: lane 1, undigested Pmal-c2E vector; lane 2, *Nde*I and *EcoR*I digested vector; lane 3, undigested pMAL-c2G vector; lane 4, *Nde*I and *EcoR*I digested vector; lane 5, 1 Kb ladder

The eluted samples were analyzed by 12% acrylamide SDS-PAGE and the desired protein (based on size 32,000 Da) was found in the pellet fraction of the elutant. As the protein was expressed in the insoluble fraction, it might have been expressed as an inclusion body in the bacterial cell.

The protein expressed as inclusion bodies was run on a 12% acrylamide SDS-PAGE and the overexpressed protein band of the desired size was cut from the gel and electro eluted (Fig. 23, lane 5) and subject to dialysis. After this the sample was again subjected to electrophoresis to confirm the size of the band (Fig. 23, lane 6).

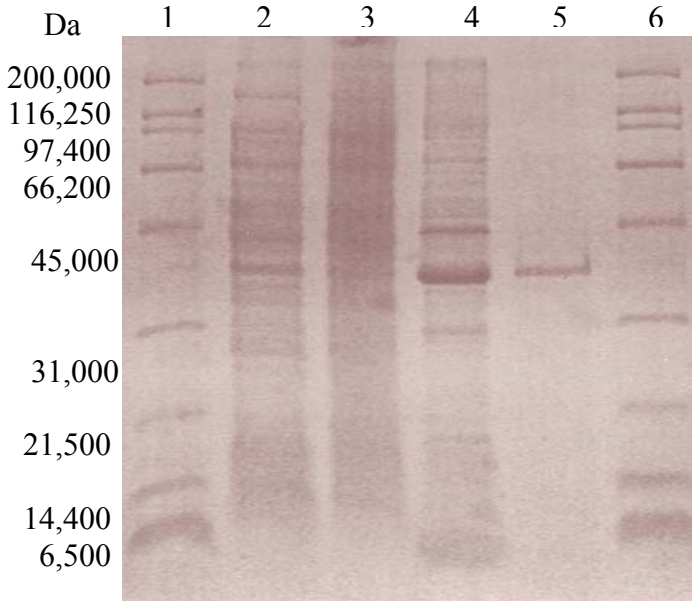


Figure 23. Expression of LMWG1D1 using the tac promoter: lane 1 and 6, markers; lane 2, uninduced *E. coli* cells; lane 3, Induced *E. coli* cells; lane 4, expressed protein in the pellet; lane 5, electro eluted protein from the gel.

The electro eluted band was analyzed by Matrix assisted laser desorption ionization mass spectrophotometer (MALDI – TOF MS) analysis to confirm the identity

of the protein (peptide finger printing was done at Recombinant DNA and protein core facility of Biochemistry and Molecular Biology department). Peptide mass fingerprinting analysis showed that the sequence of the eluted and the induced proteins were Chain A of Ompf Porin Mutant D74a, and not the LMWG1D1 protein. This suggests that MBP is aiding in the solubulization of the glutenin protein which forms aggregates upon expression and when glutenin is expressed without MBP-tag, the protein is expressed as inclusion bodies. The changes in vector restriction map might have caused frame shifts which might changed the reading frame of the expressed protein under *tac* promoter.

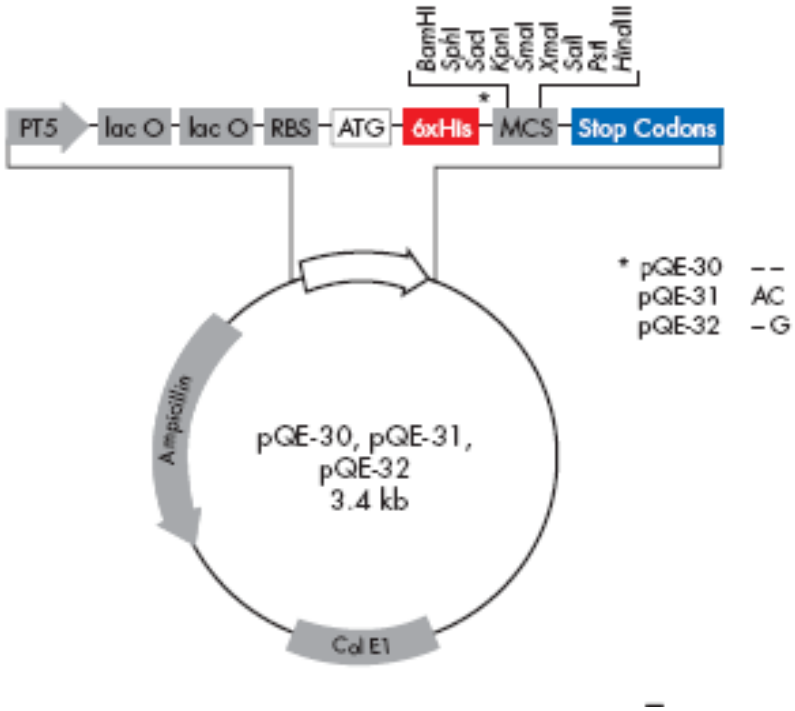
We also speculate that changes in osmotic pressure of the culture media and inside the cell may have resulted in over expression of Ompf Porin mutant protein as shown by peptide mass finger printing analysis. Porins form aqueous channels that aid the diffusion of small hydrophilic molecules across the outer membrane of Gram-negative bacteria. OmpF porin functions to regulate osmotic pressure between the cell and its surroundings and hence we speculate this might have reduced the expression LMW1D1. OmpF porin contains 340 amino acid residues similar to the expected LMWIDI protein which appeared as an artifact in SDS- 12% PAGE analysis.

(e) Expression and cloning in pQE vector

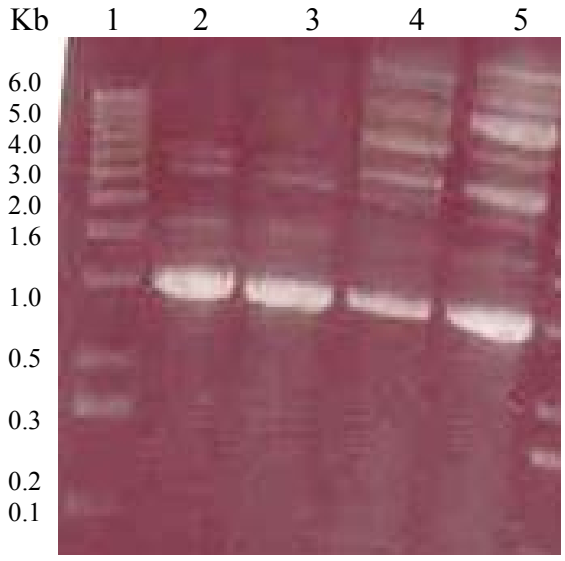
As an alternative approach for obtaining the homogenous low molecular weight gluten protein. The LMWG1D1 gene was cloned at the *Sac*I/*Hind*III sites of the pQE vector under the control of T5 promoter (Fig. 24). *E. coli* DH5α cells

were transformed using this vector and clones were selected on LB kanamycin and X-gal plates. Recombinants were subjected to expression analysis in *E. coli* M15 host cells as described earlier. The clones selected were sequenced at the Recombinant DNA and protein core facility at Department of Biochemistry using primer 4 (Table 2).

24 (a)



(b)



(c)

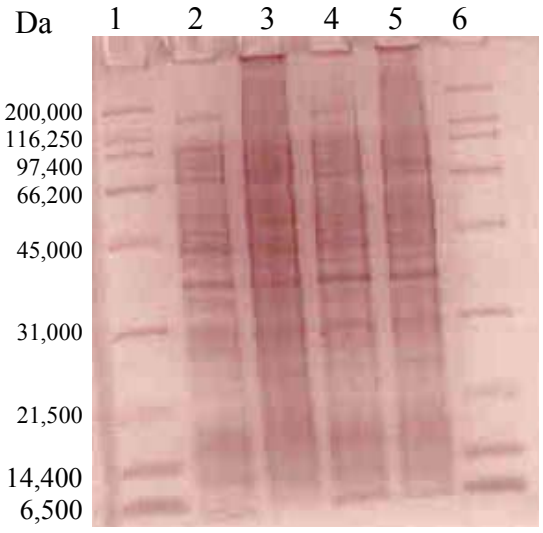


Figure 24. (a) Restriction map of pQE vector. (b) Cloning of pQE vectors in to DH5α and M15 cells: lane 1, 1 Kb ladder; lane 2, pQE30 vector in DH5α cells; lane 3 pQE32 vector in DH5α cells; lane 4, pQE30 vector in M15 cells; lane 5, pQE32 vector in M15 cells. (c) Expression of LMWG1D1 using the pQE vector: lane 1 and 6, markers; lane 2,

uninduced pQE 30 vector in *E. coli* cells; lane 3, induced pQE 30 vector in *E. coli* cells; lane 4, uninduced pQE 32 vector in *E. coli* cells; lane 5, induced pQE 32 vector protein from the gel.

The cells after expression were lysed and the extract was subjected to 12% acrylamide SDS-PAGE (Fig 24 c). There was no difference in the pattern of the protein bands between the induced and uninduced vector lanes. We concluded that there is no expression of the LMWG1D1 protein in pQE vector.

f) Conclusions and recommendations for future work

The results from the above experiments showed that the expression of the LMWG1D1 gene was not achieved using pET and pQE vector systems. The aim of this study was to express and purify the LMWG1D1 protein in large scale and in soluble form using various bacterial expression systems. One of the major difficulties encountered in our study was that the expression was not evident in other bacterial expression systems except pMAL. We concluded that MBP-tag is efficient in expressing and solubulizing the protein. The nature of the gluten protein to polymerize and form aggregates might be a reason for the improper cleavage of the fusion tag. Unpublished results from our lab also suggest similar problems associated with the expression of HMW glutenin subunits in the pMALc2x vector expression systems. The fusion protein was completely degraded during cleavage (Ana Romero personal communication). Studies to determine the reasons for the instability of the protein are undergoing in the lab. It was not possible for us to isolate and purify the soluble protein for its characterization. Western blotting of the

protein expressed in the bacterial system without the MBP-tag would give insight into the results we obtained from the mass spectrometry of that protein. Also it has been shown that bacterial systems tend to metabolize the abundant glutamine and proline amino acids of the gluten proteins. So we speculate that the use of yeast like eukaryotic systems could facilitate the expression and solubilization of the protein.

References

References

Barro, F., Rooke, L., Bekes, F., Gras, P., Tatham, A. S., Fido, R., Lazzeri, P. A., Shewry, P. R. and Barcelo, P. 1997. Transformation of wheat with high molecular weight subunit genes results in improved functional properties. Nat Biotechnol 15:1295-9.

Bartels, D., Altosaar, I., Harberd, N. P., Barker, R. F. and Thompson, R. D. 1986. Molecular Analysis of Gamma-Gliadin Gene Families at the Complex Gli-1 Locus of Bread Wheat (Triticum-Aestivum L). Theoretical and Applied Genetics 72:845-853.

Belton, P. S. 1999. On the elasticity of wheat gluten. Journal of Cereal Science 29:103-107.

Belton, P. S., Colquhoun, I. J., Field, J. M., Grant, A., Shewry, P. R. and Tatham, A. S. 1994. H-1 and H-2 Nmr Relaxation Studies of a High M(R) Subunit of Wheat Glutenin and Comparison with Elastin. Journal of Cereal Science 19:115-121.

Cassidy, B. G., Dvorak, J. and Anderson, O. D. 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.

Colot, V., Bartels, D., Thompson, R. and Flavell, R. 1989. Molecular characterization of an active wheat LMW glutenin gene and its relation to other wheat and barley prolamin Genes. Molecular & General Genetics 216:81-90.

Dobraszczyk, B. J. and Morgenstern, M. 2003. Rheology and the breadmaking process. Journal of Cereal Science 38:229-245.

D'Ovidio, R. and Masci, S. 2004. The low-molecular-weight glutenin subunits of wheat gluten. Journal of Cereal Science 39:321-339.

Dowell, F. E., Broce, A. B., Xie, F., Throne, J. E. and Baker, J. E. 2000. Detection of parasitised fly puparia using near infrared spectroscopy. Journal of Near Infrared Spectroscopy 8:259-265.

Ewart, J. 1979. Glutenin structure. Journal of the Science and Agriculture: 482-492.

Feeney, K. A., Wellner, N., Gilbert, S. M., Halford, N. G., Tatham, A. S., Shewry, P. R. and Belton, P. S. 2003.Molecular structures and interactions of repetitive peptides based on wheat glutenin subunits depend on chain length. Biopolymers 72:123-131.

Garg, M., Singh, H., Kaur, H. and Dhaliwal, H. S. 2006. Genetic control of high protein content and its association with bread-making quality in wheat. Journal of Plant Nutrition 29:1357-1369.

Gupta, R. B., Khan, K. and Macritchie, F. 1993. Biochemical basis of flour properties in bread wheat's 1. Effects of variation in the quantity and size distribution of polymeric protein. Journal of Cereal Science 18:23-41.

Gupta, R. B., Masci, S., Lafiandra, D., Bariana, H. S. and MacRitchie, F. 1996. Accumulation of protein subunits and their polymers in developing grains of hexaploid wheats. Journal of Experimental Botany 47:1377-1385.

Humphris, A. D. L., McMaster, T. J., Miles, M. J., Gilbert, S. M., Shewry, P. R. and Tatham, A. S. 2000. Atomic force microscopy (AFM) study of interactions of HMW subunits of wheat glutenin. Cereal Chemistry 77:107-110.

Jackson, E. A., Holt, L. M. and Payne, P. I. 1983. Characterization of High Molecular-weight gliadin and low-molecular-weight glutenin subunits of wheat endosperm by two-dimensional electrophoresis and the chromosomal localization of their controlling genes. Theoretical and Applied Genetics 66:29-37.

Jackson, E. A., Holt, L. M. and Payne, P. I. 1985. Glu-B2, A Storage Protein Locus controlling The D-Group of LMW glutenin subunits in bread Wheat (Triticum-Aestivum). Genetical Research 46:11-17.

Kasarda, D. D. 1999. Glutenin polymers: The in vitro to in vivo transition. Cereal Foods World 44:566-571.

Kasarda, D. D., Bernardin, J. E. and Qualset, C. O. 1976. Relationship of gliadin protein components to chromosomes in hexaploid wheat's (Triticum aestivum L.). Proc Natl Acad Sci U S A 73:3646-3650.

Kohler, P., Belitz, H. D. and Wieser, H. 1993. Disulfide bonds in wheat gluten - further cystine peptides from High-Molecular-Weight (HMW) And Low-Molecular-Weight (LMW) subunits of glutenin and from gamma-gliadins. Zeitschrift Fur Lebensmittel-Untersuchung Und-Forschung 196:239-247.

Kuktaite, R., Larsson, H. and Johansson, E. 2004. Variation in protein composition of wheat flour and its relationship to dough mixing behavior. Journal of Cereal Science 40:31-39.

Levanony, H., Rubin, R., Altschuler, Y. and Galili, G. 1992. Evidence for a novel route of wheat storage proteins to Vacuoles. Journal of Cell Biology 119:1117-1128.

Li, W., Dobraszczky, B. J., Dias, A. and Gil, A. M. 2006. Polymer conformation structure of wheat proteins and gluten sub fractions revealed by ATR-FTIR. Cereal Chemistry 83:407-410.

Lindsay, M. P. and Skerritt, J. H. 1999. The glutenin macropolymer of wheat flour dough's: structure-function perspectives. Trends in Food Science & Technology 10:247-253.

Lindsay, M. P., Tamas, L., Appels, R. and Skerritt, J. H. 2000. Direct evidence that the number and location of cysteine residues affect glutenin polymer structure. Journal Of Cereal Science 31:321-333.

Masci, S., D'Ovidio, R., Lafiandra, D. and Kasarda, D. D. 1998. Characterization of a low-molecular-weight glutenin subunit gene from bread wheat and the corresponding protein that represents a major subunit of the glutenin polymer. Plant Physiology 118:1147-1158.

Masci, S., Rovelli, L., Kasarda, D. D., Vensel, W. H. and Lafiandra, D. 2002. Characterization and chromosomal localization of C-type low-molecular-weight glutenin subunits in the bread wheat cultivar Chinese Spring. Theor Appl Genet 104:422-428.

McMaster, T. J., Miles, M. J., Kasarda, D. D., Shewry, P. R. and Tatham, A. S. 2000. Atomic force microscopy of A-gliadin fibrils and in situ degradation. Journal of Cereal Science 31:281-286.

Mecham, D. K., Kasarda, D. D. and Qualset, C. O. 1978. Genetic aspects of wheat gliadin proteins. Biochem Genet 16:831-53.

Muller, H. G. 1975. Rheology and conventional bread and biscuit-making process. Cereal Chemistry 52:R89-R105.

Muller, S., Vensel, W. H., Kasarda, D. D., Kohler, P. and Wieser, H. 1998. Disulphide bonds of adjacent cysteine residues in low molecular weight subunits of wheat glutenin. Journal Of Cereal Science 27:109-116.

Orsi, A., Sparvoli, F. and Ceriotti, A. 2001. Role of individual disulfide bonds in the structural maturation of a low molecular weight glutenin subunit. Journal of Biological Chemistry 276:32322-32329.

Osborne, T. 1924. The vegetable proteins. Longman's Green Co, London.

Patacchini, C., Masci, S., D'Ovidio, R. and Lafiandra, D. 2003. Heterologous expression and purification of native and mutated low molecular mass glutenin subunits from durum wheat. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences 786:215-220.

Payne, P. I. 1987. Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. Annual Review of Plant Physiology and Plant Molecular Biology 38:141-153.

Payne, P. I. and Corfield, K. G. 1979. Subunit composition of wheat glutenin proteins, isolated by gel-filtration in a dissociating medium. Planta 145:83-88.

Payne, P. I., Holt, L. M., Jackson, E. A. and Law, C. N. 1984. Wheat storage proteins - their genetics and their potential for manipulation by plant-breeding. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 304:359.

Payne, P. I., Holt, L. M., Lawrence, G. J. and Law, C. N. 1982. The genetics of gliadin and glutenin, the major storage proteins of the wheat endosperm. Qualitas Plantarum-Plant Foods for Human Nutrition 31:229-241.

Payne, P. I., Roberts, M. S. and Holt, L. M. 1986. Location of genes-controlling the D-group of LMW Glutenin subunits on chromosome-1d of bread Wheat. Genetical Research 47:175-179.

PJ Stone., R. S. 1999. Wheat: Ecology and physiology of yield determination in: Grain Quality and its physical determinants. Food Products Press, New York.

Posner, E. 2000. Handbook of cereal science and technology. Marcel Dekker Corp, New York.

Posner, E. S. 1988. Back to the issue of Wheat grading. Cereal Foods World 33:362-365.

Ram, S., Bhatia, V., Jain, V. and Mishra, B. 2006. Characterization of low molecular weight glutenin subunit gene representing Glu-B3 locus of Indian wheat variety NP4. Journal of Plant Biochemistry and Biotechnology 15:79-83.

Shewry, P. R. 1996. Cereal grain proteins. Pages 227-250 in: Cereal grain proteins, RJ Henry., P S Kettlewell R. Henry, ed. Chapman and Hall, London.

Shewry, P. R., Halford, N. G., Belton, P. S. and Tatham, A. S. 2002. The structure and properties of gluten: an elastic protein from wheat grain. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 357:133-142.

Shewry, P. R., Napier, J. A. and Tatham, A. S. 1995. Seed storage proteins - Structures and biosynthesis. Plant Cell 7:945-956.

Shewry, P. R., Popineau, Y., Lafiandra, D. and Belton, P. 2000. Wheat glutenin subunits and dough elasticity: findings of the EUROWHEAT project. Trends in Food Science & Technology 11:433-441.

Shewry, P. R., Sayanova, O., Tatham, A. S., Tamas, L., Turner, M., Richard, G., Hickman, D., Fido, R., Halford, N. G., Greenfield, J., Grimwade, B., Thomson, N., Miles, M., Freedman, R. and Napier, J. 1995. Structure, assembly and targeting of wheat storage proteins. Journal of Plant Physiology 145:620-625.

Shewry, P. R. and Tatham, A. S. 1997. Disulphide bonds in wheat gluten proteins. Journal of Cereal Science 25:207-227.

Shewry, P. R., Tatham, A. S., Barro, F., Barcelo, P. and Lazzeri, P. 1995. Biotechnology of breadmaking - unraveling and manipulating the multi-protein gluten complex. Bio-Technology 13:1185-1190.

Shewry, P. R., Tatham, A. S. and Lazzeri, P. 1997. Biotechnology of wheat quality. Journal of the Science of Food and Agriculture 73:397-406.

Shukla, T. P. 2001. Chemistry of dough development. Cereal Foods World 46:337-339.

Singh, M. and Khatkar, B. S. 2005. Structural and functional properties of wheat storage proteins: A review. Journal of Food Science and Technology-Mysore 42:455-471.

Sissons, M. J., Ames, N. P., Hare, R. A. and Clarke, J. M. 2005. Relationship between glutenin subunit composition and gluten strength measurements in durum wheat. Journal of the Science of Food and Agriculture 85:2445-2452.

Stone, B. 1996. Cereal grain carbohydrates. Chapmann and Hall, New York.

Thompson, S., Bishop, D. H. L., Madgwick, P., Tatham, A. S. and Shewry, P. R. 1994. High-Level expression of a Wheat LMW glutenin subunit using a Baculovirus system. Journal of Agricultural and Food Chemistry 42:426-431.

Tronsmo, K. M., Faergestad, E. M., Longva, A., Schofield, J. D. and Magnus, E. M. 2002. A study of how size distribution of gluten proteins, surface properties of gluten and dough mixing properties relate to baking properties of wheat flours. Journal of Cereal Science 35:201-214.

Tsiami, A. A., Bot, A., Agterof, W. G. M. and Groot, R. D. 1997. Rheological properties of glutenin sub fractions in relation to their molecular weight. Journal of Cereal Science 26:15-27.

Veraverbeke, W. S. and Delcour, J. A. 2002. Wheat protein composition and properties of wheat glutenin in relation to breadmaking functionality. Crit Rev Food Sci Nutr 42:179-208.

Weegels, P. L., Orsel, R., Vandepijpekamp, A. M., Lichtendonk, W. J., Hamer, R. J. and Schofield, J. D. 1995. Functional-Properties of LMW wheat Proteins 2. Effects on dough properties. Journal of Cereal Science 21:117-126.

Wrigley, C. W. 1996. Biopolymers - Giant proteins with flour power. Nature 381:738-739.

VITA

VENKATA HARINI GUDISEVA

Candidate for the Degree of

Master of Science

Thesis: EXPLORING THE CONDITIONS FOR THE EXPRESSION OF LOW MOLECULAR WEIGHT GLUTENIN 1D1 PROTEIN IN *ESCHERI COLI. (E.COLI)*

Major Field: Biochemistry and Molecular Biology

Biographical:

Born in Rajahmundry, Andhra Pradesh, India on August12th 1978, Daughter of Mr. Venkata Krishna Rao Gudiseva

Education: Received a Bachelor of Science in Genetics, Botany, Chemistry, P. B. Siddhartha College of arts and sciences, Vijayawada, India in 1999. Awarded Masters of Science in Biotechnology Andhra University, Visakhapatnam, India in 2001. Completed the requirements for the Master of Science degree in Biochemistry and Molecular Biology at Oklahoma State University in December, 2006.

Research Experience: Worked as an intern in Shantha Biotechnics, Hyderabad, India in 2001 Graduate research assistant, Department of Biochemistry and Molecular Biology Oklahoma State University, Stillwater, Oklahoma from 2004-2006. Member of GPSGA from 2005 to 2006.

Name: Venkata Harini Gudiseva Date of Degree: December 2006

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of the study: EXPRESSION OF LOW MOLECULAR WEIGHT GLUTENIN 1D1 PROTEIN IN *ESCHERICHIA COLI. (E. COLI)*

Pages in Study – 54 Candidate for the Degree of Master of Science

Major Field: Biochemistry and Molecular Biology

Scope of Study: Bread wheat (***Triticum aestivum***) is one of the major cereal crops used for making flour. Wheat flour has unique rheological properties which have been attributed to its protein content. Glutens are a major class of wheat proteins contributing to the elasticity and viscosity of wheat dough. Two classes of wheat gluten proteins have been identified: High Molecular Weight Glutenins (HMW) and Low Molecular Weight Glutenins (LMW). Different allelic forms of LMW seem to play different roles in determining the different quality parameters of wheat dough.

The goal of the present study is to express the gene encoding the Low Molecular Weight Glutenin 1D1 protein in a bacterial system and purify the protein. The purified protein would then be used to determine its rheological and mixographic properties which are important parameters for determining the quality of dough in the bread making industry.

To this end we have tried expressing the gene in various expression vectors and in different bacterial host strains but we were unable to express and purify the protein to its homogeneity. We could express the protein with a fusion tag which was difficult to cleave and hence, we hypothesize that the future studies to express and characterize Low Molecular Weight Glutenin 1D1 protein should utilize eukaryotic systems.

Advisor Dr. Patricia Rayas Duarte