1 INTRODUCTION

Sorghum is a major subsistence cereal crop grown in Africa and Asia, due to its hardiness and drought-resistance, enabling its cultivation in semi-dry regions (FAO 1995, reviewed by Belton and Taylor, 2004). Hence, sorghum is a major food staple for the poorest people of Africa, and is the principal source of energy, protein, vitamins and minerals for these people (FAO, 2002; ICRISAT, 2009). Sorghum grain contains about 11% protein (ranging between 7 to 16%), which is similar to other major cereals (Serna-Saldivar and Rooney, 1995). However, the protein quality (specifically lysine protein content and protein digestibility) of sorghum is inferior to other cereals. The lysine content of its protein is substantially lower, by 35 to 90% compared to other cereals (reviewed by Henley, Taylor and Obukosia, 2010). The lower lysine content is due to the major proteins of sorghum, the kafirin prolamin storage proteins, being essentially lysine-free (reviewed by Belton, Delgadillo, Halford and Shewry, 2006). In addition, the digestibility of sorghum proteins is lower than that of maize, despite the fact that the proteins are very similar (reviewed by Duodu, Taylor, Belton and Hamaker, 2003). The low protein digestibility of sorghum is even further reduced after wet cooking, such as porridge making (Axtell, Kirleis, Hassen, D'Cros-Mason, Mertz, and Munck, 1981; Taylor and Taylor, 2002). The significance is that sorghum foods have much lower protein digestibility (protein quality) than foods made from other cereals (Mertz, Hassen, Cairns-Whittern, Kirleis, Tu and Axtell, 1984). These factors have a huge impact on communities that rely on sorghum as a major staple, where Protein Energy Malnutrition (PEM), insufficient food, lack of dietary diversity and disease outbreaks are common (Muller and Krawinkle, 2005).

Due to the above, efforts to improve the overall nutritional quality of sorghum are taking place. Nutrient fortification strategies for sorghum (and other cereals) include improved agronomic practices, breeding high nutrient yielding cultivars and nutrient targeted genetic modification (or biofortification) (reviewed by Poletti, Gruissem and Sautter, 2004). Biofortification can be defined as a process to increase the bioavailability and the concentration of nutrients in crops through both conventional plant breeding and recombinant DNA technology (genetic engineering) (Henley et

al., 2010). Effective biofortification of sorghum could contribute to improved nutritional quality of sorghum-based foods, which in turn could alleviate some of the macro- and micronutrient deficiencies of poorer communities. However, the accumulation of macro- and micronutrients in the grain is a limiting factor in health improvement, unless the bioavailability and retention of these nutrients in the processed food is improved.

Currently, nutritionally improved sorghum lines with reduced expression of different kafirin sub-classes, are being developed using genetic engineering techniques by the Africa Biofortified Sorghum (ABS) Project, under the Bill and Melinda Gates Foundation Grand Challenges (No. 9) in the Global Health Initiative (ABS Project, 2009). The Grand Challenge 9 (GC 9) intends to ‘‘Create a full range of optimal bioavailable nutrients in a single staple plant species’’ with the aim of improving nutrition to promote health. The specific aims of the ABS Project are to: 1) increase iron and zinc availability by 50%, 2) to increase provitamin A levels to up to 20 mg/kg, 3) to increase lysine content by 80–100%, 4) to increase tryptophan and threonine by 20%, 5) to concomitantly decrease leucine by 15%, and 6) to improve protein digestibility from its current value to approximately 60–80% (ABS Project, 2009; Henley et al., 2010). These improvements in sorghum nutrient content and availability are based on the average levels currently found in normal sorghum. In this work, only the improvements in lysine content and *in vitro* protein digestibility will be addressed.

2 LITERATURE REVIEW

In this review, focus is put on the development of high-lysine sorghum and maize mutants, with emphasis on protein nutritional quality (lysine content and protein digestibility) and grain endosperm quality. Research conducted on the sorghum protein composition (non-kafirins and kafirin proteins), the synthesis of sorghum and maize protein bodies, distribution of different prolamins in the protein bodies, and the role the different prolamins play in protein body structure and endosperm structure will be reviewed. Comparisons will be made between normal sorghum, normal maize, high-lysine and high-protein digestibility mutants, Quality Protein Maize and research conducted on transgenic maize. Biotechnology used to transform the protein nutritional quality of sorghum in the ABS Project will also be covered.

The nutritional quality of a protein source for humans (or animals) can be defined as the power of that protein source to cover the requirements for nitrogen and amino acids of these organisms (reviewed by Schaafsma, 2005). In this work the term '*protein nutritional quality'* will refer specifically to lysine content and protein digestibility.

2.1 Development of high-lysine maize and sorghum

Research to find cereals with improved protein quality has been ongoing since the early 1960's, beginning with reports of naturally occurring (spontaneous) high-lysine mutant maize genotypes, *opaque-2* (Mertz, Bates and Nelson, 1964) and *floury-2* (Nelson, Mertz and Bates, 1965). The lysine content of the high-lysine maize endosperm tissue was found to be 4 and 3 g lysine/100 g protein, respectively, compared to 2 g lysine/100 g protein of normal maize endosperm. The improved lysine content of these mutant grains was attributed to low levels of the lysine-poor seed storage proteins, the prolamins, which are called zeins in maize, with compensatory increases in the lysine- and tryptophan-rich non-zein seed proteins, as well as free lysine and tryptophan compared to normal maize (reviewed by Gibbon and Larkins, 2005). The improved protein nutritional value of *opaque*-2 maize was demonstrated in a series of growth experiments using rats, weaning pigs

and young children suffering from protein malnutrition (identified as kwashiorkor) (reviewed by Mertz, Axtell, Ejeta and Hamaker, 1993). Diets containing *opaque*-2 maize combined with minerals and vitamins were found to significantly increase the growth in rat and weaning pig feeding trials compared to normal maize, while in human feeding trials, diets based solely on *opaque*-2 flour was shown to cure children who suffered from kwashiorkor. The improved health in human children was attributed to the fact *opaque*-2 maize has 90% the protein nutritional value of milk protein (Mertz et al., 1993). Over the years, *opaque-2* has been used in a number of breeding trials to produce several high-lysine mutant maize genotypes with good corneous endosperm texture, referred to as Quality Protein Maize (QPM) (Gibbon and Larkins, 2005).

Success with *opaque*-2 maize encouraged extensive research to identify similar naturally occurring mutants in other cereals, including barley (Munck, Karlsson, Hagberg and Eggum, 1970) and sorghum (Singh and Axtell, 1973). The use of chemical mutagens to induce high-lysine mutants was also investigated (Mertz et al., 1993; Eggum, Brunsgaard and Jensen, 1995). In 1966, Purdue University, USA, initiated a comprehensive programme to systematically investigate the inheritance and improvement of protein content and protein quality in sorghum (reviewed by Wall and Paulis, 1978). Two naturally occurring high-lysine mutants (IS11167 and IS11758, of Ethiopian origin) were identified from 62 varieties, all with floury endosperm phenotype, obtained from the world sorghum collection (9 000 varieties) (Singh and Axtell, 1973). These high-lysine mutants were very high in protein (15.70% and 17.20%) and lysine content (3.33 and 3.13 g lysine /100 g protein), respectively, compared to the mean protein (12.7%) and lysine (2.05%) content of the remaining 60 normal varieties analysed.

Another approach used during this period was to induce high-lysine mutations in sorghums using the chemical mutagen, diethylsulphate (Mertz et al., 1993). The mutation resulted in at least a 60% increase in protein lysine content. High-lysine sorghum selection, P721 opaque (P721Q), contained similar protein content (12.9%) as the normal parent (P721N), but had 3.1 g lysine/100 g protein as opposed to 2.1 g lysine/100 g protein in the normal parent.

Inheritance studies on native high-lysine sorghums (Singh and Axtell, 1973) and the chemically induced high-lysine P721Q mutant, suggested that the increased lysine concentration is controlled by a single recessive gene that is simply inherited in both cases (Wall and Paulis, 1978). The high-lysine gene is therefore easily transferred by standard plant breeding procedures. Thus, both native high-lysine sorghums (Jambunathan, Mertz and Axtell, 1975) and the chemically induced high-lysine mutant P721Q (Weaver, Hamaker and Axtell, 1998; Tesso, Ejeta, Chandrashekar, Huang, Tandjung, Lewamy, Axtell and Hamaker, 2006; Nyannor, Adedokun, Hamaker, Ejeta and Adeola, 2007; Tesso, Hamaker and Ejeta, 2008) have been used in plant breeding programmes to develop high-lysine populations derived from crosses with normal commercially important sorghums.

2.2 Amino acid profile and protein composition of high-lysine mutants

In high-lysine sorghums, as is the case with other high-lysine cereals such as *opaque*-2 maize, the high-lysine gene alters the amino acid pattern in the floury endosperm tissue relative to normal vitreous endosperm tissue (Singh and Axtell, 1973). The major changes observed were increased lysine, arginine, aspartic acid, glycine and tryptophan concentration and decreased amounts of glutamic acid (glutamine), proline, alanine and leucine in the floury endosperm tissue. Differences in the distribution of soluble protein fractions between high-lysine and normal sorghums is considered to be the primary reason for the improved lysine content of high-lysine sorghums. Jambunathan et al. (1975) compared the distribution of soluble protein fractions in native high-lysine sorghum (IS11758), high-lysine hybrid (F_2) sorghums and normal sorghums with similar genetic background. It was shown by these authors that the native (IS11758) and hybrid high-lysine (F_2) kernels had a decrease in the concentration of lysine-poor kafirin proteins, with a compensatory increase in the lysine-rich non-kafirin proteins (albumins, globulins and glutelins, including free amino acids) compared to the normal sorghums. The reduced synthesis of kafirin proteins and compensatory increased synthesis in non-kafirin proteins no doubt accounts for the increased lysine content in high-lysine kernels. This was later confirmed in a study using high-lysine mutants (IS11167 and P721Q) and normal sorghums (Guiragossian, Chibber, Van Scoyoc, Jambunathan, Mertz

and Axtell, 1978). However, these authors showed that the modified distribution of soluble protein fractions and increased lysine content was endosperm specific. It was suggested that in high-lysine mutants it is necessary to look at the distribution of soluble proteins specifically in the endosperm tissue to fully understand the marked changes in amino acid patterns observed in the whole grain. This is due to the fact that lysine-rich albumin and globulin proteins are found primarily in the germ (and cells of the aleurone layer) (Wall and Paulis, 1978; Taylor and Schüssler, 1986). Sorghum kernels with a large germ and small endosperm portion, or floury endosperm can in fact show high lysine content. However, these grains generally have low protein content. Even so, differences in total protein soluble in 60% tertiary butanol with a reducing agent, 2-mercaptoethanol (2-ME) has been used to distinguish normal and high-lysine sorghums (Wall and Paulis, 1978).

In maize, a similar protein content was also found to occur between different high-lysine opaque mutants (ranging from 9.3 to 12.8%), QPM (ranging from 9.7 to 10.8%) and normal maize genotypes (ranging from 9.2 to 10.8%) (Gibbon and Larkins 2005). It would appear that the protein content of different sorghums (and maize) is relatively stable regardless of the genotype. Nitrogen fertilisation has been found to be a major factor in elevating the protein content in normal sorghum (Walls and Paulis, 1978). In these cases, a slight decline in lysine content is reported, in normal sorghum, which is thought to be due to increased synthesis of the lysine-poor storage proteins (kafirins) in the grain endosperm, thereby reducing the protein nutritional quality of the grain.

Nevertheless, the increased lysine content reported for high-lysine mutant sorghum and maize is still below that recommended by the World Health Organization (WHO) (4.8 g lysine/100 g protein for 4-18 year olds) for foodstuffs to meet basic protein requirements (WHO/FAO/UNU Expert Consultation, 2007). It is not known if genetic transformation can increase the lysine content of transformed cereals to meet the lysine requirements recommended by the WHO.

2.3 Classification of sorghum proteins

Sorghum proteins can be classified into five major classes namely: albumins (water-soluble), globulins (salt-soluble), prolamins (kafirins) (aqueous alcohol-soluble), crosslinked kafirins (aqueous alcohol + reducing agent-soluble) and glutelins (detergent + reducing agents + alkaline pH-soluble) (reviewed by Hamaker and Bugusu, 2003). This solubility-based classification scheme was proposed by Landry and Moureaux (1970) for the proteins of maize and was modified by Jambunathan et al. (1975) for sorghum. Later studies revealed that the kafirins and crosslinked kafirins, after reduction with a reducing agent like 2-ME, are essentially identical. Thus, sorghum proteins can be classified simply as kafirins and non-kafirins (Hamaker, Mohamed, Habben, Huang and Larkins, 1995).

2.3.1 Non-kafirin proteins

The albumin and globulin proteins are concentrated in the germ of the kernel, and are higher in nutritional value compared to the kafirin proteins, due to higher concentrations of lysine (Taylor and Schüssler, 1986). The glutelin proteins (also referred to as endosperm matrix protein) are the second major proteins of the endosperm (Serna-Saldivar and Rooney, 1995). They are higher in the basic amino acids, glutamic acid and lysine, compared to the kafirins. The glutelin proteins are thought to also function as a source of enzymes involved in the hydrolysis of starch and protein reserves, as opposed to being storage type proteins like the kafirins (Taylor, Novellie and Liebenberg, 1984). In maize (and possibly sorghum), the glutelin proteins making up the matrix protein are thought to be derived from cytoplasmic albumin and globulin proteins which form disulphide-linkages during the latter stages of seed development (Wall, Cooker and Bietz, 1988).

2.3.2 Kafirin proteins

The kafirin proteins are the most abundant proteins in whole grain sorghum, making up 50 to 60% of the total protein (Taylor and Schüssler, 1986; Shull, Watterson and Kirleis, 1991). However, values as high as 70% of whole grain protein and about

80% of decorticated grain protein have been reported (Hamaker et al., 1995). Kafirins are rich in glutamine, proline, alanine and leucine, but are essentially free of lysine (reviewed by Lásztity, 1984, reviewed by Shewry and Halford, 2002). It is due to the high content of kafirin prolamin storage proteins, which are deficient in essential amino acids, that sorghum (and other cereals) is rendered inadequate with regard to protein nutritional quality.

Kafirin proteins are known to be more hydrophobic than other prolamin proteins (such as zein) due to a higher proportion of non-polar amino acid residues (Duodu et al, 2003). Proline and glutamine constitute more than 30% of the kafirin proteins. Thus, the optimum solubility condition for kafirin is 60% aqueous tertiary butanol, due to the higher hydrophobicity of this solvent compared to other aqueous alcohol solvents, such as ethanol and propanol which are commonly used (Belton et al., 2006). Small quantities of cystine (sulphur-containing amino acid) contribute to intra- and inter-molecular disulphide bonding of kafirins. Disulphide bonding of kafirin proteins results in the formation of large polymeric kafirin proteins, which are insoluble in the aqueous alcohols and have low digestibility (Hamaker et al., 1995; El Nour, Peruffo and Curioni, 1998). Consequently a reducing agent such as 2-ME is required to break the disulphide bonds and improve extraction of the kafirin proteins.

2.3.2.1 Kafirin sub-classes

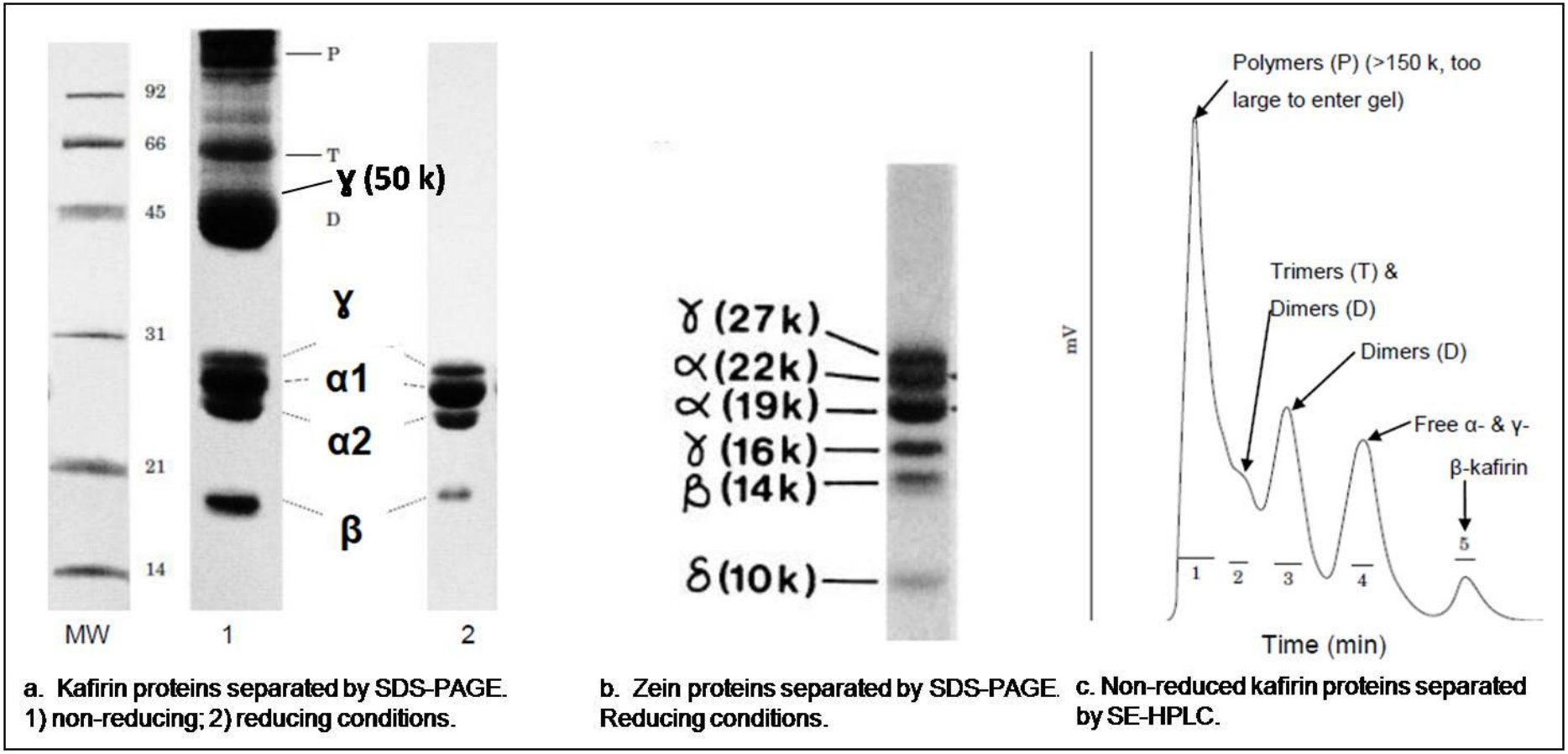
Kafirins are heterogeneous proteins separable into different sub-classes (α-, β- ɣ- and δ-kafirins) of different proportions (Belton et al., 2006). The sub-classes are characterised based on their solubility, amino acid composition, electrophoretic, chromatographic, and immunological property differences and DNA sequencing (Shull et al., 1991; Belton et al., 2006). Table 2.1 summarises the major properties of the different kafirin sub-classes with regard to percentage of total kafirin fractions in both corneous (vitreous) and opaque endosperm, number of resolving bands observed by SDS-PAGE, apparent molecular mass, amino acid composition and polymerisation (crosslinking) behaviour. Kafirin proteins are homologous to the zeins and many comparative studies regarding these two prolamin proteins have been done (Esen, 1987; Shull et al., 1991; Mazhar, Chandrashekar, and Shetty,

1993; El Nour et al., 1998; Nunes, Correia, Barros, and Delgadillo, 2005). Due to this homology a nomenclature for kafirin has been described, but is constantly being reviewed and modified as new and improved technologies are developed (reviewed by Bean and Lookhart, 2000). In this current research project, the different kafirin sub-classes targeted for suppressed synthesis are defined as, α-kafirin A1 (25 kDa) and α-kafirin B1 and B2 (which correspond to 19 kDa and 22 kDa α-kafirin classes, respectively, (Belton et al., 2006)), ɣ-kafirin 1 (27 kDa) and 2 (50 kDa), and δ-kafirin (15 kDa) (Dr R. Jung, Pioneer Hi-Bred, personal communication).

The crosslinking behaviour of the different monomeric kafirins to form dimers, oligomers and polymers is directly related to the number of cysteine residues per monomer. Figure 2.1 shows the separation of kafirin and zein proteins by SDS-PAGE (sodium dodecyl sulphate-polyacrylamide gel electrophoresis), and the separation of non-reduced kafirin proteins by SE-HPLC (size exclusion high performance liquid chromatography). According to El Nour et al. (1998) α-kafirins are present in whole grain flour as both monomeric and polymeric proteins, under non-reducing conditions (Figure 2.1a). The α-kafirin 1 (26.6 kDa, 2 cysteine residues) can link by disulphide bonds to form different sized oligomers, while α-kafirin 2 (24 kDa, one cysteine residue) forms dimers or small oligomers. Thus, α-kafirin 1 is considered to be a ‘chain extender’ and α-kafirin 2 a ‘chain terminator’. The β-kafirins contain 10 cysteine residues, which enables them to form either inter- or intra-molecular disulphide bonds, forming oligomers and polymers (Belton et al., 2006). According to El Nour et al. (1998), β-kafirin can act as a bridge between oligomers of α-kafirin 1 and ɣ-kafirins resulting in very large polymers, extractable only under reducing conditions. High molecular weight polymeric proteins (>150 kDa) are not observed by SDS-PAGE, as they are too large to enter the stacking gel (El Nour et al., 1998). However, other techniques such as size exclusion chromatography can be used to separate the large polymeric proteins (Figure 2.1c) as was demonstrated by El Nour et al. (1998) and later by Emmambux and Taylor (2009). The ɣ-kafirins contain substantially more cysteine residues (15 residues) and are present in their native state as polymers stabilised by disulphide bonds (Belton et al., 2006) and as disulphide crosslinked polymers with α-kafirin 1 (El Nour et al., 1998).

Table 2.1 Summary of the major properties of the four types of kafirin proteins and their crosslinking behaviour within the sorghum grain (modified from Belton et al. (2006)).

Kafirin type	Total kafirin fraction (%) (in vitreous (corneous) or opaque(floury) endosperm)	Number of resolving bands	Apparent molecular mass ranges reported (kDa)	Amino acid composition	Polymerisation behaviour
α-	80-84% (vitreous) 66-71% (opaque)	Two major bands. Can be resolved into more components which vary in number and properties between genotypes, and separation technique used.	19 to 27 (19, 22, 23, 24, 25)	240-250 amino acid residues. Rich in non-polar amino acids, no lys, 1 typ, 10 blocks of repeated amino acids	Monomers, oligomers and polymers α- 1, “Chain extender” α- 2, “Chain terminator”
β-	7-8% (vitreous) 10-13% (opaque)	Three major bands. Or a single band.	15, 17 and 18 or 18 only	172 amino acid residues. Rich in Met, and Cys, 2 Trp	Monomers with intra-chain disulphide bonds. Oligomers and polymers, stabilised by intra- and inter-chain disulphide bonds. “Chain extender”
ɣ-	9-12% (vitreous) 19-21% (opaque)	A number of resolving bands have been identified.	16 to 28 and 50 (minor band)	193 amino acid residues. Rich in Pro, Cys, His. No Lys, Asn, Asp, Trp. Four repeats (consensus PPPVHL)	Oligomers and polymers, stabilised by inter-chain disulphide bonds. “Chain extender’
δ-	Not known	Two sequences have been reported.	12.961 or 10 and 18	114 amino acid residues. Rich in Met, No Lys, 1 Trp	Not known



a. Kafirin proteins separated by SDS-PAGE. 1) non-reducing; 2) reducing conditions.

b. Zein proteins separated by SDS-PAGE. Reducing conditions.

c. Non-reduced kafirin proteins separated by SE-HPLC.

Figure 2.1 Separation of kafirin and zein proteins using SDS-PAGE, and kafirin proteins using SE-HPLC.
(a) Kafirin proteins separated by SDS-PAGE; MW M_r standard proteins; Lane 1, unreduced 60% (v/v) tertiary butanol extract; Lane 2, reduced 60% (v/v) tertiary butanol extract. P, T and D indicate protein polymers, trimers and dimers, respectively (El Nour et al., 1998). (b) Zein proteins separated by SDS-PAGE (Shewry and Halford, 2002). γ, α- 1, α- 2, and β indicate the zein/kafirin proteins. (c) Non-reduced kafirin proteins separated by SE-HPLC unreduced 60% (v/v) tertiary butanol sorghum flour extract. Numbers indicate collected peaks, and where identified by SDS-PAGE of the eluted peaks (El Nour et al., 1998).

2.4 Synthesis of prolamin storage protein bodies

Cereal seed storage proteins are synthesised in the rough endoplasmic reticulum (rough ER) and are deposited in discrete protein bodies within the grain endosperm (Shewry and Halford, 2002). The synthesis of maize (Gibbon and Larkins, 2005) and sorghum (Taylor, Schüssler and Liebenberg, 1985) protein bodies has been studied extensively, and found to be very similar. Hence the descriptive model for the development of zein protein bodies and the distribution of zein proteins in maize (Figure 2.2), proposed by Lending and Larkins (1989), is commonly used for sorghum. Figure 2.3 shows transmission electron micrographs of protein bodies in developing sorghum endosperm.

According to Lending and Larkins (1989), protein body formation in maize (and probably sorghum), begins with the initial accumulation of zein proteins within the rough endoplasmic reticulum, which consists of dark-staining deposits of both β- and ɣ-zeins, and contains little or no α-zein (Figure 2.2a). Subsequently, light-staining locules of α-zein begin to accumulate within the β- and ɣ-zeins (Figure 2.2b). As the interior of the protein body fills with α-zein, the locules of α-zein fuse and aggregate to form a central core. Some smaller locules of α-zein remain and are interspersed in the outer region of the protein body (Figure 2.2c). The dark-staining region that contains β- and ɣ-zein forms a continuous layer around the periphery of the protein body. In the final stages of protein body maturation, α-zein fills most of the protein body core and is surrounded by a thin layer of β- and ɣ-zeins (Figure 2.2d). Small, dark-staining patches of ɣ-zein, and to a lesser extent β-zein, may occur within the interior region.

According to Shull et al. (1992) slight differences may exist between the synthesis of sorghum and maize protein bodies. More β- and ɣ-protein crosslinking occurs in older more mature sorghum protein bodies compared to maize protein bodies, which may indicate that differences exist between the development of sorghum and maize protein bodies.

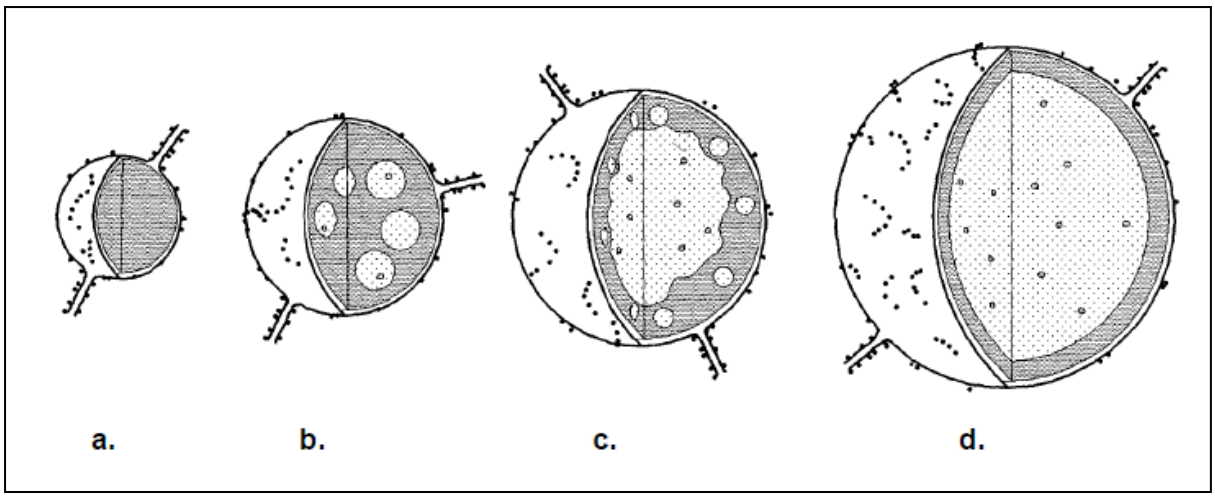


Figure 2.2 A diagrammatic representation of the development of protein bodies in maize endosperm.
The heavily stippled regions correspond to regions that are rich in β- and ɣ-zeins, and the lightly stippled regions correspond to regions rich in α-zein. The protein body is surrounded by rough endoplasmic reticulum (dark dots represent ribosomes). Some β- and ɣ-zeins are found within the regions that consist primarily of α-zein (heavily stippled inclusions). Localization of different zein proteins was done using immunolocalization techniques (taken from Lending and Larkins, 1989). Figure is not to scale.

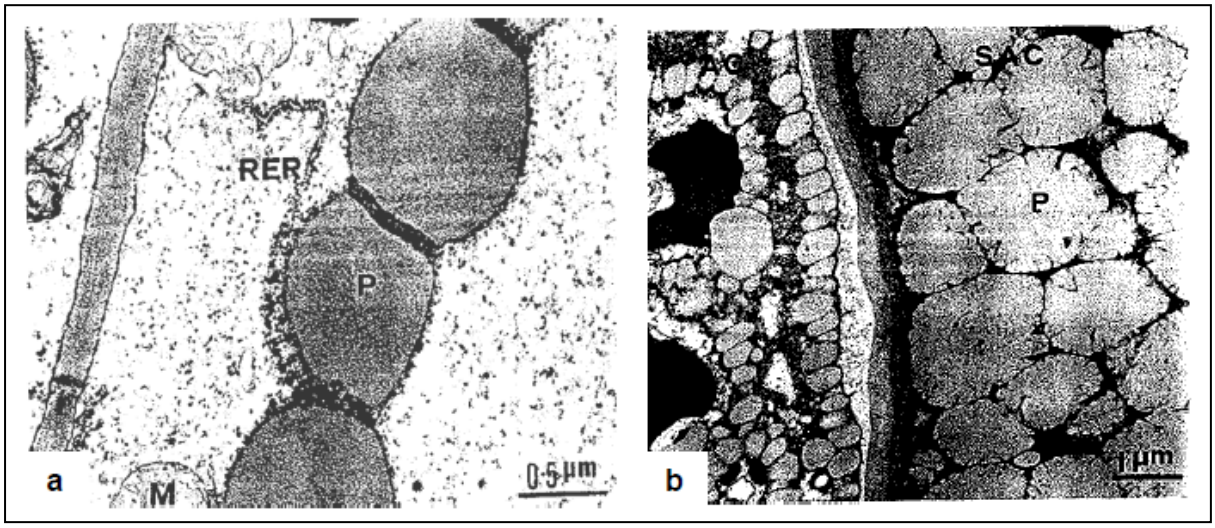


Figure 2.3 Transmission electron micrographs of developing sorghum endosperm.
a) Electron micrograph of sorghum endosperm at soft dough stage of seed development. M – mitochondria; P – protein body; RER – rough endoplasmic reticulum. b) Sorghum aleurone and sub-aleurone cells at the late hard dough stage of seed development; AC – aleurone cell; M- matrix protein; P – protein body; SAC – sub-aleurone cell (taken from Taylor, Schüssler and Liebenberg, 1985).

In addition, in a study conducted by Oria, Hamaker and Shull (1995a), the quantification of different kafirin sub-classes using solvent systems, showed α-kafirin was the first prolamin protein to accumulate in the developing kernel, followed by the accumulation of β- and ɣ-kafirin. It was also stated by these authors that the ratio of crosslinked kafirin fractions to uncrosslinked kafirin fractions increases significantly during grain maturity and desiccation. Kafirin crosslinking is due to the formation of disulphide bonds being formed due to oxidative processes or interchange involving sulfhydryl groups (Oria et al., 1995a). These differences may also be responsible for the lower protein digestibility of sorghum compared to maize (see section 2.6).

In mature sorghum, the amorphous matrix protein that encases the tightly packed protein bodies and starch granules, is composed of aggregated dense cytoplasm and disintegrated ER and mitochondria (Shull, Chandrashekar, Kirleis and Ejeta, 1990). In normal sorghum, mature protein bodies found in the sub-aleurone layer and vitreous endosperm are round in cross-section, with diameters ranging from 0.5 to 3.5 µm (Adams, Novellie and Liebenberg, 1976; Taylor et al., 1985). Crosslinked polymeric kafirins are commonly found within the protein bodies arranged in concentric rings (Seckinger and Wolf, 1973; Adams et al., 1976; Taylor et al., 1984; Krishnan, White and Pueppke, 1989).

2.4.1 Synthesis of protein bodies in mutant genotypes

In contrast to normal sorghum protein bodies, protein bodies in mutant sorghums (P721Q and derivatives) were found to be remarkably different and irregular shaped (Oria, Hamaker, Axtell and Huang, 2000) (Figure 2.4). These mutant sorghums are also characterised as having high protein digestibility, and are thus termed high-protein digestibility mutants (HPDM) (see section 2.7). The protein bodies of HPDM are characterised by deep invaginations or folds, and in many cases the folds reach the central area of the protein body forming irregular shaped lobes (Figure 2.4b, c) (Oria et al., 2000). The dark inclusions are also seen mainly at the base of the folds (Figure 2.4c), and are composed of crosslinked ɣ- and β-kafirins. Similarly, the protein bodies in developing endosperm of *floury-2* mutant maize (a mutation that results in reduced levels of all zein sub-classes) were also reported to differ in shape

from the discrete, spherical protein bodies formed in normal maize endosperm (Lending and Larkins, 1992). The protein bodies of *floury-2* were reported to be irregular in shape, with irregular clusters (or lobes) of zein proteins within the lumen of the rough ER (Figure 2.5b). This work is in agreement with earlier studies conducted by Wolf and co-workers who isolated small oddly shaped protein bodies in immature *floury-2* maize kernels (Wolf, Khoo and Seckinger, 1967). Of interest, these authors reported that there was an absence of protein bodies in mature *floury-2* maize. It was concluded that the protein bodies disintegrate during grain maturation.

In contrast to both *floury-2* maize mutant and the HPDM, the protein bodies in developing endosperm of *opaque-2* mutant maize (mutation results in considerable reduction in α-zein sub-class, M_r 22,000) are typically smaller than those of the normal genotype, but have a normal spherical shape (Geetha, Lending, Lopes, Wallace and Larkins, 1991) (Figure 2.5c). Decreased α-zein production in high-lysine maize genotypes is commonly associated with smaller and fewer protein bodies in the grain endosperm (Christianson, Khoo, Nielsen and Wall, 1974; Gibbon and Larkins, 2005).

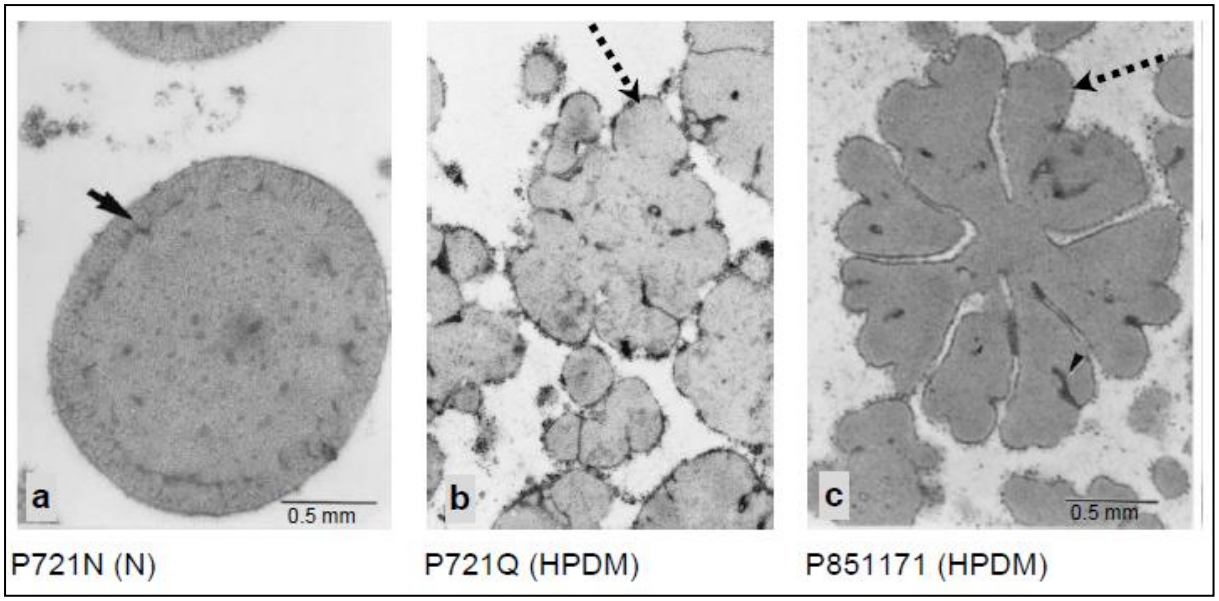


Figure 2.4 Transmission electron micrographs of protein bodies in the peripheral endosperm of normal (N) and high-protein digestibility mutant (HPDM) sorghum cultivars.
Black arrow points at dark-staining peripheral ring; black dashed arrow points at lobes of protein bodies, black arrow head points at dark-staining base folds. (From Oria et al., 2000).

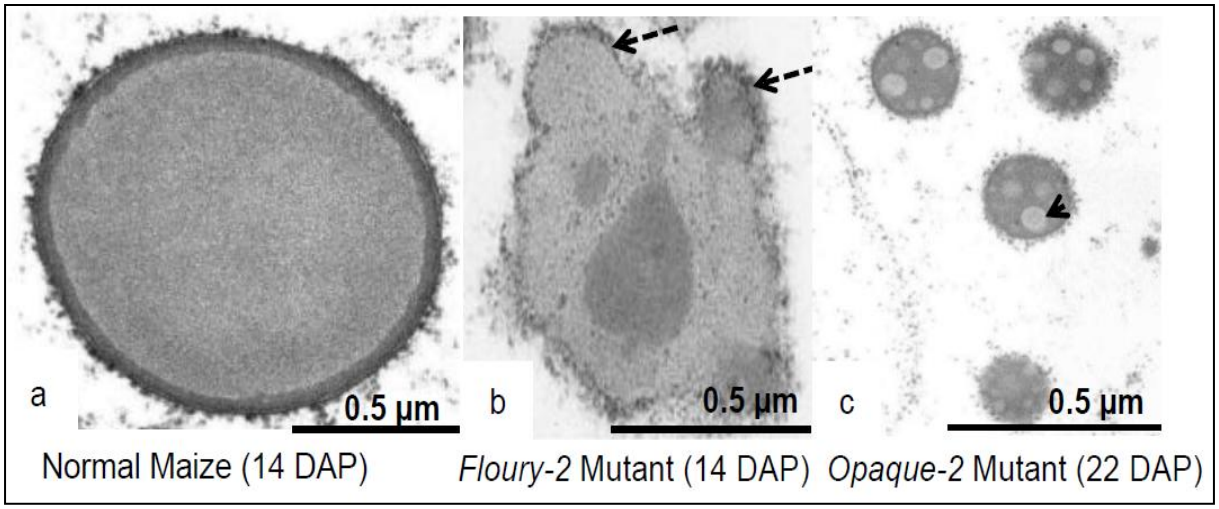


Figure 2.5 Transmission electron micrographs of representative protein bodies from developing endosperm of a) normal maize, b) *floury-2*, c) *Opaque-2* mutant.
Dashed arrows show irregular protein lobs (or clusters); arrow head shows light staining inclusions; DAP - days after pollination. (From: Lending and Larkins (1989); Lending and Larkins (1992) and Geetha et al. (1991), respectively).

2.5 Organisation and role of prolamin proteins in storage protein bodies

In normal sorghum, the distribution of the different kafirin sub-classes within storage protein bodies is non-homogeneous. Immunocytochemical studies show that the α-kafirins are located mainly in the interior of the protein body, as light staining areas, while the β- and ɣ-kafirins are found inside and at the periphery of the protein bodies, as dark-staining areas (Shull et al., 1992). At the protein body periphery, ɣ-kafirins and to a lesser degree β-kafirins form a disulphide-bound polymeric network which encapsulates the α-kafirins (Figure 2.6a).

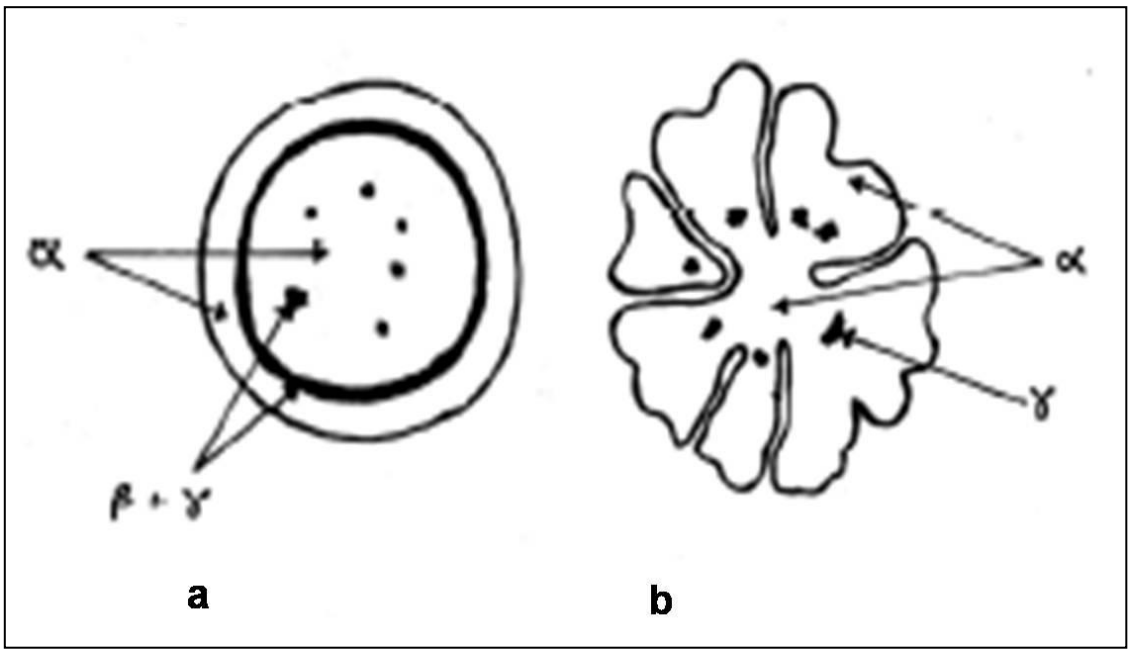


Figure 2.6 A diagrammatic representation of a section through a) a normal sorghum protein body and b) a high-protein digestibility mutant protein body showing differential location of α-, β- and ɣ-kafirins (from Duodu et al., 2003).

In HPDM sorghum protein bodies, the distribution of kafirin proteins is modified compared to that of normal sorghum protein bodies. Immunolocalization studies, conducted by Oria et al. (2000) on protein bodies from HPDM, revealed that α-kafirin is localized homogeneously throughout the light-staining regions of the interior of the protein bodies, β-kafirin is found predominantly as dark staining areas localised at the periphery of the protein body lobes, and ɣ-kafirin is also found as dark staining areas localised at the base of the invaginations and to a lesser extent at the

peripheral regions (Figure 2.6b). Very small amounts of β- and ɣ-kafirins also occur in the light staining area of the protein body.

In a study to characterize the associations and distribution of different zein sub-classes in developing endosperm of the *floury-2* maize mutant, Lending and Larkins (1992) found that the protein bodies were irregular in shape (Figure 2.5b). Also, the zein proteins were also disorganised compared to those observed in normal maize genotypes (Lending, Kriz, Larkins and Bracker, 1988; Lending and Larkins, 1992). In the initial stages of endosperm development, protein body formation followed the same pattern observed in normal maize, namely the initial aggregates within the rough ER consist predominantly of β- and ɣ-zeins, as described by Lending and Larkins (1989), (Figure 2.2). However, as the α-zeins accumulated, the morphology of the protein bodies began to change from that in normal maize. As the protein bodies in *floury-2* endosperm became enlarged, they started to develop irregular lobes. Immunocytochemical staining revealed that the distribution of the different zein sub-classes was disorganised in *floury-2* maize, with patches of β- and ɣ-zein occurring within irregular lobed clusters of α-zeins (Lending and Larkins, 1992). In addition, within the central regions of endosperm cells, clusters of protein deposits were observed to form within the rough ER aggregates and were dispersed throughout the cytoplasm. This was in direct contrast to the formation of protein bodies in normal maize genotypes. Additionally, granular, electron-dense deposits, unique to *floury-2* maize, were also observed between the rough ER and the protein aggregates during all stages of protein body formation.

In the endosperm of *opaque-2* mutant maize, the level of one zein sub-class, M_r 22,000 α-zein, is greatly reduced (Jones, Larkins and Tsai, 1977; Burr and Burr 1982; Kodrzycki, Boston, and Larkins, 1989). The protein bodies in *opaque-2* maize are typically smaller than those of the normal maize, due to the great decrease in the amount of the M_r 22,000 α-zein (Geetha et al., 1991).

Although the mechanism of protein body assembly is not fully understood, it has been hypothesized that specific interactions must occur between the various prolamin sub-classes. In maize, it is thought that the sulphur-poor α-zein must pass

into the sulphur-rich β- and ɣ-zeins that initially aggregate within the rough ER (Lending and Larkins, 1992). Research into understanding the mechanisms involved in protein body assembly in cereals is a major focus area (Geli, Torrent, and Ludevid., 1994; Bagga, Adams, Kemp, and Sengupta-Gopalan, 1995; Coleman, Herman, Takasaki and Larkins, 1996; Bagga, Adams, Rodriguez, Kemp and Sengupta-Gopalan, 1997; Kim, Woo, Clore, Burnett, Carneiro and Larkins, 2002; Coleman, Yoho, Escobar and Ogawa, 2004). These authors used *in planta* transgenic tobacco seed or yeast two-hybrid systems, to study the interactions between different zein proteins and their role in the formation of a maize protein body. There is evidence that specific protein–protein interactions play an important role in storage protein body formation. Geli et al. (1994) and Bagga et al. (1995) showed that β-zein and ɣ-zein, when expressed separately, accumulate as protein accretions within tissues of transgenic tobacco seed. Delta-zein was also found to accumulate in transgenic tobacco seed, but the amount of δ-zein was enhanced when it was co-expressed with β-zein (Bagga et al. 1997). However, the δ-zein/β-zein accretions had an irregular structure compared to the spherical protein bodies of normal maize endosperm. This result suggests that formation of a spherical protein body requires the inclusion of both ɣ- and β-zein proteins.

Other studies have shown that α-zein accumulation in transgenic tobacco seed is only possible when either ɣ-zein (Coleman et al., 1996), or β-zein (Coleman et al., 2004) is co-expressed in the transgenic seed. The resulting α-zein/ɣ-zein or α-zein/β-zein accretions formed were similar in appearance to those of maize endosperm protein bodies. These results indicate that either β- or ɣ-zein is able to stabilise the accumulation of α-zein in transgenic tobacco seed endosperm. These also demonstrate the importance of the sulphur-rich prolamin proteins in initiating and maintaining protein body structure.

Considering that β- and ɣ-zein are structurally related and appear to have similar roles in the initiation of protein bodies in developing maize endosperm (Lending and Larkins, 1989; Woo, Hu, Larkins and Jung, 2001; Kim et al., 2002), the findings of Coleman and co-authors (1996 and 2004) are not surprising. The mechanism involved in the initiation and development of zein protein bodies is described as the

aggregation of cysteine-rich β- and ɣ-zein proteins via the formation of inter-molecular disulphide bonds, forming an insoluble protein body core, required for integration of α-zein (Coleman et al., 2004). Both β- and ɣ-zein proteins may possess intrinsic properties that allow them, when assembled as aggregates, to interact with α-zein. Such interactions may involve the hydrophobicity of the α-zein protein or disulphide bonding between α-zein and β- or ɣ-zein (Coleman et al., 2004). A study by Kim et al. (2002) using the yeast two-hybrid system, also confirmed that the various zein proteins either interact with themselves or with other zein proteins during protein body synthesis, but the role of these interactions in initiating or maintaining zein protein body structure was unclear. The model described by Lending and Larkins (1989) shows great similarity to that described above.

2.6 Grain endosperm quality of high-lysine maize and sorghum

The discovery of high-lysine genotypes in maize and sorghum stimulated a great deal of excitement due to the potential of these genotypes for enhancing the nutritional value of these grains (Mertz et al., 1964). However, high-lysine mutant cereals are renowned for having soft, floury endosperm, resulting in lower kernel density and increased susceptibility to insects, pathogens, mechanical damage and reduced yields compared to their normal vitreous counterparts (Mertz et al., 1964; Ortega and Bates, 1983, Lending and Larkins, 1992; Weaver et al., 1998; Tesso et al., 2006; Tesso et al., 2008). These agronomic problems have inhibited the widespread cultivation of high-lysine maize and sorghum.

Grain hardness or endosperm texture, is a key trait in maize and sorghum as it provides resistance to damage during harvesting and marketing, as well as to insect and fungal damage (Maxson, Fryar, Rooney, and Krishnaprasad, 1971). Also grain hardness is the main characteristic that determines milling performance and end-use quality of the decorticated flour (Maxson et al., 1971; Awika, Suhendro and Rooney, 2002). During milling, sorghum and maize with corneous endosperm yield more complete endosperms and fewer broken particles of endosperm than grains with floury endosperm when decorticated (Maxson et al., 1971; Awika et al., 2002). The

poor milling properties of grains with soft, floury endosperm texture is due to the weak protein-protein and protein-starch association in the grain endosperm, and results in shattering of the endosperm during milling (Maxson et al., 1971; reviewed by Chandrashekar and Mazhar, 1999; Awika et al., 2002). In addition, grain endosperm texture also has a profound effect on the food making quality and sensory properties of sorghum and maize foods (Murty and Kumar, 1995; Chandrashekar and Mazhar, 1999).

Endosperm texture is determined in maize and sorghum by the relative amounts of hard (corneous) and soft (floury) endosperm in the grain. There is some correlation between prolamin storage proteins and corneous kernel (Maxson et al., 1971). A 3-point rating system for evaluating sorghum endosperm texture was developed by Taylor and Taylor (2008) (Figure 2.7), and has been adopted by the International Association for Cereal Science and Technology (ICC) as a draft ICC standard method for determining sorghum grain endosperm texture (ICC, 2008). This rating system distinguishes sorghum endosperm texture into corneous, intermediate or floury phenotype.

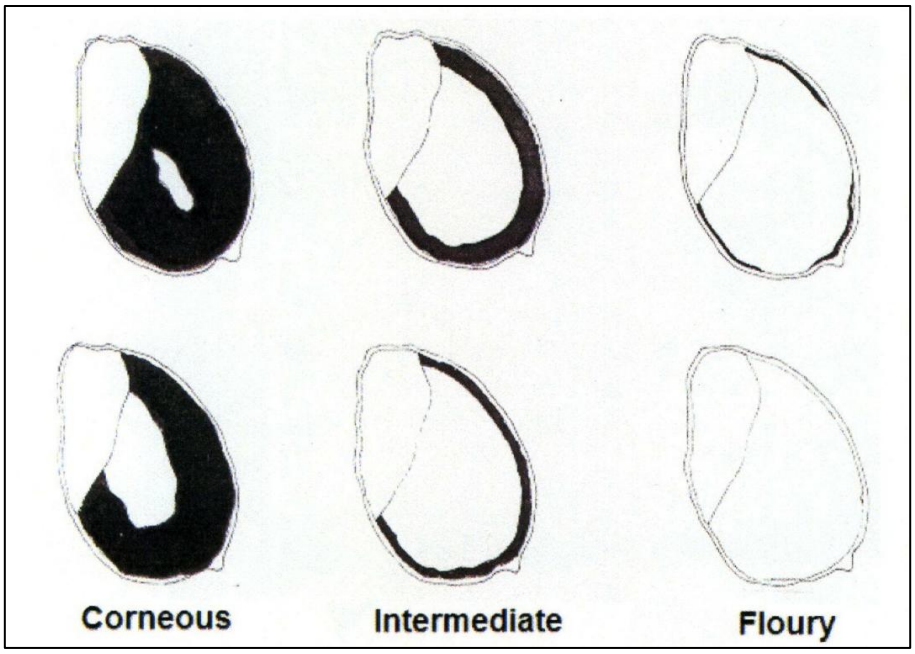


Figure 2.7 Illustration of the 3-point rating system for evaluation sorghum endosperm texture (ICC, 2008).

According to a review of the major causes of grain hardness in sorghum and maize, it was concluded by Chandrashekar and Mazhar (1999) that grain hardness is multifactorial, which includes: Cell wall structure and the types and concentrations of prolamins present in the endosperm. In hard grains the cell wall polymers are more rigid, with more protein bodies which are evenly distributed in the peripheral endosperm tissue than the protein bodies found in floury grains. In addition, the protein bodies in hard endosperm generally contain more ɣ-prolamins which seem to be crosslinked by disulphide bonds (Chandrashekar and Mazhar, 1999).

2.6.1 Quality Protein Maize (QPM)

Research to improve the endosperm texture of high-lysine maize has resulted in the development of high-lysine phenotypes with corneous endosperm texture for maize, known as Quality Protein Maize (QPM) (Gibbon and Larkins, 2005). QPM was developed by plant breeders in South Africa and the International Maize and Wheat Improvement Centre (CIMMYT) in Mexico. Researchers identified a number of modifier genes that alter the soft, floury texture of the endosperm, giving the grain a normal appearance while maintaining the increased essential amino acid content (due to low α-zein content) of *opaque-2* (Gibbon and Larkins, 2005). The mechanism by which the opaque phenotype of *opaque-2* endosperm is modified in QPM is still under investigation (Wu, Holding and Messing, 2010). Genetic mapping of *opaque-2* modifiers has indicated that there is a link between the locus encoding the 27 kDa ɣ-zein storage protein and a region near the end of the long arm of chromosome 7 (Lopes, Takasaki, Bostwick, Helentjaris, and Larkins, 1995). Other studies on QPM have shown that, compared to *opaque-2*, the 27 kDa ɣ-zein is increased two- to three-fold (Wallace, Lopes, Paiva, and Larkins, 1990; Wu et al., 2010). The increase in ɣ-zein is dependent on the dosage of *opaque-2* modifiers, which appear to act in a semi-dominant manner (Lopes and Larkins, 1991). In addition, numbers of zein protein bodies are also increased, forming a compacted protein mass between the starch granules in the endosperm (Dannenhoffer, Bostwick, Or and Larkins, 1995). According to Wu et al. (2010) ɣ-zeins are essential for ensuring protein body density and starch grain interaction, thus giving QPM a corneous endosperm texture.

A model depicting the corneous or floury starchy endosperm cells for wild type, QPM and *opaque-2* maize, at the mid-maturation stage (18 DAP) is shown in Figure 2.8 (Wu et al., 2010). In the wild-type maize and QPM, compact stable matrices are visible, giving rise to glass-like, vitreous endosperm at maturity. Also of note, the protein bodies in QPM endosperm are generally smaller than that of the wild-type. In the opaque mutant (low-zein line), small, sparsely distributed protein bodies are formed, producing loose, unstable matrices, which shatter during desiccation, resulting in an opaque texture at maturity.

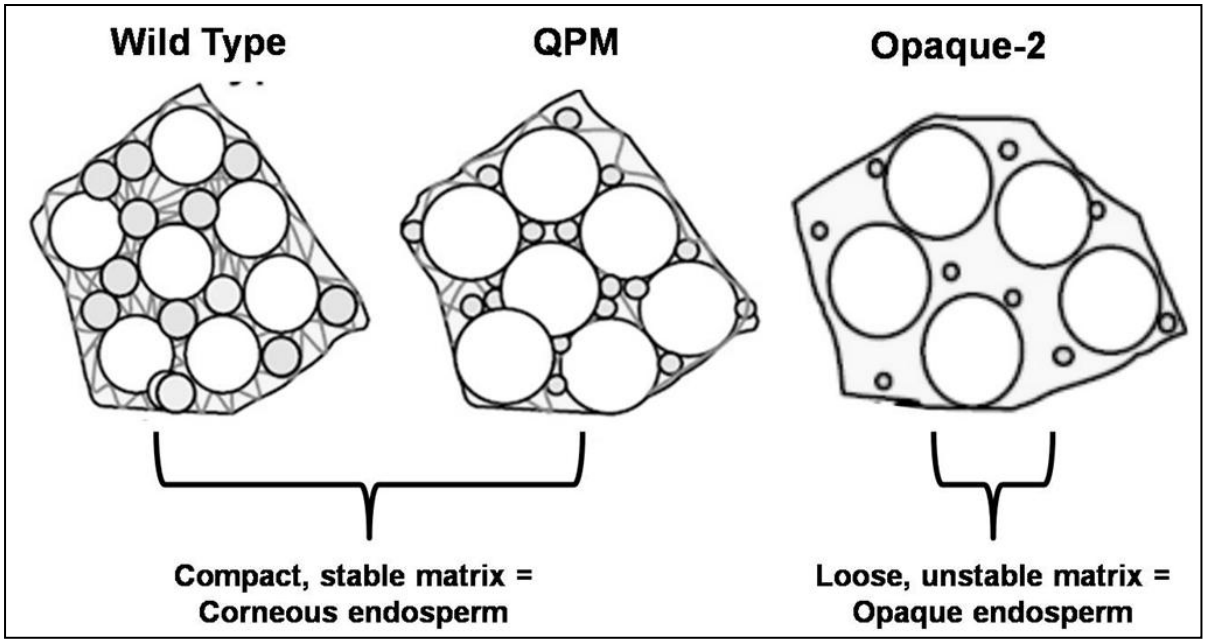


Figure 2.8 Model for corneous endosperm formation in which mid-maturation stage (18 DAP) starchy endosperm cells are depicted for wild type, QPM (Quality Protein Maize) and *opaque-2* mutant.
The protein bodies are represented by small grey spheres, starch grains with white spheres, and proteinaceous matrix with grey lines (from Wu et al., 2010).

In the case of sorghum, Tesso et al. (2006) identified a novel sorghum mutant with both high-protein digestibility and high-lysine (HDP/*hl*) trait and relatively hard endosperm texture. The HPD/hl mutant was an F6 generation of crosses between P721Q mutant and elite hard endosperm sorghum lines. Improved protein digestibility and lysine content were 20% and 44% higher than normal sorghum, respectively. The relatively hard endosperm texture was attributed to areas of densely packed starch granules, but without the typically associated continuous

protein matrix of normal sorghum. In addition, the protein bodies of HDP/*hl* had irregular, invaginated shapes, as previously described for P721Q derived lines (Oria et al., 2000).

2.7 Factors affecting *in vitro* protein digestibility of sorghum

The protein digestibility of normal sorghum is particularly low, especially after wet cooking. According to MacLean, de Romańa, Placko and Graham (1981), the apparent protein digestibility of cooked sorghum porridge can be as low as 46%, compared to values for cooked wheat (81%), rice (66%) and even maize (73%), despite the fact the proteins of sorghum and maize are very similar. Factors affecting the digestibility of sorghum proteins have been studied by various authors and can be divided into two broad categories: endogenous and exogenous factors, and have been reviewed extensively by Duodu et al. (2003). The endogenous factors include factors that arise out of changes within the sorghum proteins themselves, namely disulphide and non-disulphide crosslinking, kafirin hydrophobicity and changes in protein secondary structure. The exogenous factors include grain organisational structure, and the interaction of sorghum proteins with non-protein compounds inherent in the whole grain, namely polyphenols, non-starch polysaccharides, starch, phytate and lipids. Reduced protein digestion is due to either the formation of chemically indigestible products or the formation of a physical barrier preventing access of proteases to the proteins during processing. All the above factors have been shown to influence sorghum protein digestibility to different degrees and more than one factor can be involved at any time. The nature or state of the sorghum grain (whole grain or endosperm, protein body preparation, and tannin content) has also been reported to influence the degree of protein digestibility.

Of the factors listed above, it is the formation of disulphide crosslinked kafirins, specifically between the cysteine-rich ɣ- and β-kafirins, during grain development, desiccation and maturity (Hamaker, Kirleis, Mertz and Axtell, 1986; Oria et al., 2000), as well as those that are formed during cooking (Axtell, et al., 1981; Hamaker et al., 1986; Rom, Shull, Chandrashekar and Kirleis, 1992; Oria, Hamaker and Shull, 1995b; Ezeogu, Duodu and Taylor, 2005), that are considered to play major roles in

lowering the protein digestibility of sorghum. Hamaker, Kirleis, Butler, Axtell and Mertz (1987) demonstrated that kafirins are the last proteins to be digested and require the addition of a reducing agent such as 2-ME to break the disulphide crosslinks between the kafirin proteins to improve their uncooked and cooked digestibilities. According to Ezeogu, Duodu, Emmambux and Taylor (2008), cooking in the presence of 2-ME also improves the starch digestibility of wet cooked sorghum and maize. This is due to 2-ME breaking the disulphide bonds in the prolamin-protein matrix, causing an expansion of the corneous endosperm matrix protein mesh and resulting in improved access of α-amylase to the gelatinised starch.

In addition, isolated native (unreduced) α-kafirins are in fact highly digestible and retain their high digestibility after cooking (Hamaker et al., 1986; 1987). These results, and a scanning electron microscopy study, indicate that the breakdown of kafirin protein bodies starts at the protein body periphery and progresses inward (Rom et al., 1992). Hence, the poor digestibility of α-kafirin is caused by its location within the interior of the protein body, as it is surrounded by an enzyme-resistant structure (disulphide crosslinked ɣ- and β-kafirins) that retards digestion of the α-kafirin (Figure 2.4a) (Oria et al., 2000).

Mutant sorghum lines, developed by crossing P721Q with normal hard endosperm lines, have been found to have substantially improved *in vitro* protein digestibility, 10-15% higher in uncooked flour and 25% higher in cooked flour, compared to normal sorghum (Weaver et al., 1998). The improved protein digestibility is reported to be due to change in the shape of the kafirin protein bodies from spherical to invaginated (Figure 2.4b,c) where the ɣ-kafirin is concentrated at the base of folds, exposing the more digestible α-kafirins (Oria et al., 2000). As a result, the α-kafirins are more exposed to digestive enzymes compared to that of normal sorghum protein bodies.

2.8 Africa Biofortified Sorghum (ABS) Project

Due to the fact that sorghum is a major staple food in developing countries and has major limitations regarding protein nutritional quality, it has been identified as one of the crops for biofortification, namely the Africa Biofortified Sorghum (ABS) project

(reviewed by O'Kennedy, Grootboom and Shewry, 2006; Henley et al., 2010). The ABS project is employing recombinant DNA technology with the aim of improving both sorghum lysine content and wet-cooked protein digestibility. Sorghum transformation is brought about using a ''super-binary'' *Agrobacterium*-mediated vector (Zhao, Cai, Tagliani, Miller, Wang, Pang, Rudert, Schroeder, Hondred, Seltzer and Pierce, 2000; Zhao, Glassman, Sewalt, Wang, Miller, Chang, Thompson, Catron, Wu, Bidney, Kedebe and Jung, 2003), or particle bombardment (Grootboom, Mkhonza, O'Kennedy, Chakauka, Kunert and Chikwamba, 2010) of immature zygotic embryos as explant.

Approaches for improving the protein nutritional quality traits of sorghum include:
1) Suppressing the synthesis of specific kafirin sub-classes that are very low in lysine and that are responsible for poor protein digestibility (i.e., suppression of the synthesis of various combinations of α-, ɣ-, and δ-kafirins) (O'Kennedy et al., 2006; Henley et al., 2010). Decreased zein and kafirin synthesis commonly occurs in high-lysine mutants, however, this results in floury endosperm texture. Hence, genetic engineering approaches have been used in maize to specifically down-regulate zein synthesis, with the aim of avoiding the deleterious effects associated with the high-lysine mutations (reviewed by Shewry, 2007).
2) Increasing the lysine content of sorghum by the synthesis of lysine-rich proteins (such as HT12, an analogue of barley hordothionin), expressed in the grains endosperm, and reduced catabolism of free lysine in the grain endosperm by suppressed synthesis of the enzyme lysine ketoglutarate reductase (LKR) (Zhao et al., 2003).

2.9 Conclusions

Sorghum contributes significantly to the protein content of many people's diets, even though it has low protein nutritional quality with regard to low lysine content and poor cooked protein digestibility. The search for high-lysine and high-protein digestibility mutant lines has been somewhat successful but very slow. It would appear that the mutants discovered so far all have floury endosperm texture, requiring considerable effort through breeding programmes to slightly improve the endosperm texture.

Recombinant DNA technology is now being used to co-suppress the synthesis of target kafirin proteins with the aim of developing biofortified sorghum with improved protein nutritional quality. These transgenic sorghums are the first of their kind. Thus, little is known about the effects the different kafirin protein co-suppressions will have on the physical (endosperm texture, protein body structures) and chemical (protein composition, amino acid composition, cooked IVPD and kafirin polymerisation behaviour) characteristics of the transformed grain. Research is thus needed to understand the effect the target kafirin co-suppressions will have on the physical and chemical characteristics of the transformed grains before these biofortified sorghums can be considered for commercial production and processed into foods.