CHARACTERIZATION OF THE POLYMERIC PROTEINS OF SORGHUM

by

BRIAN PAUL IOERGER

B.S., Kansas State University, 1982
B.S., Kansas State University, 1984
M.S., Kansas State University, 1993

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Grain Science and Industry
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Abstract

The role of sorghum protein cross-linking into high M_w polymeric groups in grain hardness was investigated using a number of protein analytical techniques to study the protein composition (reduced and non-reduced) of isolated vitreous and floury endosperm. The relative molecular weight distributions of polymeric proteins within two of three differentially extracted fractions were determined by size exclusion chromatography (SEC). The proteins in vitreous endosperm showed more protein cross-linking and a larger M_w distribution than found in the floury endosperm. An improved method for fractionating sorghum proteins designed to obtain intact disulfide linked protein polymers was developed. Three protein fractions obtained by application of the method represented proportionally different protein polymer contents as evidenced by comparative SEC and provides an improved tool for polymeric protein content comparison and measurement. The improved method was applied to a highly diverse non-tannin wild-type sorghum sample set spanning a range of *in-vitro* protein digestibility (IVPD) values to determine polymers involved with and influencing IVPD. Grain traits other than cross-linked proteins were also investigated for significant relationships to IVPD. Three protein fractions (F1, F2, F3) containing intact protein polymers were obtained for analysis by SEC and RP-HPLC. Proteins represented by four of five individual SEC peaks from F3 were significantly negatively correlated to IVPD, with three of the correlated peaks being polymeric. A 2-dimensional (2-D) technique involving peak collection after size exclusion chromatography followed by reverse phase high performance liquid chromatography (SEC x RP-HPLC) of the collected peaks was applied to protein polymers previously determined to be correlated with IVPD. RP-HPLC chromatogram patterns unique to each collected SEC peak from three selectively extracted protein fractions allowed qualitative and quantitative comparisons of protein polymer components. A pair of early eluting peaks appearing in the γ-kafirin region of 2nd-dimension RP-HPLC chromatograms from a protein fraction with the largest M_w distribution were significantly correlated to IVPD. The correlated peak of interest was collected and characterized using SDS-PAGE and was preliminarily identified as 27kDa γ-kafirin. By combining techniques using differing selectivity's (solvent based, molecular size based, hydrophobicity based), it was possible to disassemble and compare components of protein polymers significantly correlated to IVPD.

CHARACTERIZATION OF THE POLYMERIC PROTEINS OF SORGHUM

by

BRIAN PAUL IOERGER

B.S., Kansas State University, 1982
B.S., Kansas State University, 1984
M.S., Kansas State University, 1993

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Grain Science and Industry
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Approved by:

Co-Major Professor
Hulya Dogan

Co-Major Professor
Scott R. Bean

Copyright

Abstract

The role of sorghum protein cross-linking into high M_w polymeric groups in grain hardness was investigated using a number of protein analytical techniques to study the protein composition (reduced and non-reduced) of isolated vitreous and floury endosperm. The relative molecular weight distributions of polymeric proteins within two of three differentially extracted fractions were determined by size exclusion chromatography (SEC). The proteins in vitreous endosperm showed more protein cross-linking and a larger M_w distribution than found in the floury endosperm. An improved method for fractionating sorghum proteins designed to obtain intact disulfide linked protein polymers was developed. Three protein fractions obtained by application of the method represented proportionally different protein polymer contents as evidenced by comparative SEC and provides an improved tool for polymeric protein content comparison and measurement. The improved method was applied to a highly diverse non-tannin wild-type sorghum sample set spanning a range of *in-vitro* protein digestibility (IVPD) values to determine polymers involved with and influencing IVPD. Grain traits other than cross-linked proteins were also investigated for significant relationships to IVPD. Three protein fractions (F1, F2, F3) containing intact protein polymers were obtained for analysis by SEC and RP-HPLC. Proteins represented by four of five individual SEC peaks from F3 were significantly negatively correlated to IVPD, with three of the correlated peaks being polymeric. A 2-dimensional (2-D) technique involving peak collection after size exclusion chromatography followed by reverse phase high performance liquid chromatography (SEC x RP-HPLC) of the collected peaks was applied to protein polymers previously determined to be correlated with IVPD. RP-HPLC chromatogram patterns unique to each collected SEC peak from three selectively extracted protein fractions allowed qualitative and quantitative comparisons of protein polymer components. A pair of early eluting peaks appearing in the γ-kafirin region of 2nd-dimension RP-HPLC chromatograms from a protein fraction with the largest M_w distribution were significantly correlated to IVPD. The correlated peak of interest was collected and characterized using SDS-PAGE and was preliminarily identified as 27kDa γ-kafirin. By combining techniques using differing selectivity's (solvent based, molecular size based, hydrophobicity based), it was possible to disassemble and compare components of protein polymers significantly correlated to IVPD.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

4-VP	: 4-vinyl pyridine
AHI	: abrasive hardness index
BME	: beta-mercaptoethanol
FZCE	: free zone capillary electrophoresis
IVPD	: in-vitro protein digestibility
LMWN	: low molecular weight nitrogen
RP-HPLC	: reverse phase high performance liquid chromatography
SDS	: sodium dodecyl sulfate
SEC	: size exclusion chromatography
SDS-PAGE	: sodium dodecyl sulfate polyacrylamide gel electrophoresis
SKCS	: single kernel characterization system
TADD	: tangentially abrasive decortication device
Mw	: molecular weight
Mr	: relative molecular mass
kDa	: kilodaltons
HMW	: high molecular weight
LMW	: low molecular weight
SP	: soluble proteins
IP	: insoluble proteins
RP	: residue proteins

Acknowledgements

I would like to thank my wife, Janet, for sharing the walk on what turned out to be a longer journey than we anticipated. Also thanks to my family for their continual support and encouragement. To my major professors Dr. Hulya Dogan and Dr. Scott Bean, thank you for all your guidance, support, and advice over the course of my studies and research, and to Dr. George Lookhart and Dr. Om Prakash for their assistance and service on my graduate committee. I am also grateful to Dr. Chris Culbertson for agreeing to serve as outside chairperson on my committee with such short notice. Finally I would like to thank all of my coworkers at the USDA for their advice, assistance, support, and friendship.

Preface

The chapters in this dissertation were written to be manuscripts submitted for publication, therefore references are included for each chapter and formatted for the potential submission according to journal guidelines.

Chapter 1 - Literature Review

Portions of Chapter 1 are published as:

Bean, S.R., Ioerger, B.P., Smith, B.M., Blackwell, D.L. (2011). Sorghum protein structure and chemistry: Implications for nutrition and functionality. In: Advances in Cereal Science: Implications to Food Processing and Health Promotion. ACS Symposium Series 1089. Eds: J. Awika, V. Piironen, S. Bean. Oxford University Press, Inc. Pp.132-136.

Bean, S. and Ioerger, B.P. (2015). Sorghum and millet proteins. In: Applied Food Protein Chemistry. Ed: Z. Ustunol. John Wiley and Sons, Ltd. Pp.327-333.

General Background

The importance of grain sorghum to the world is apparent considering it is ranked fifth in the world in terms of cereal crop production (FAO, 1995). Although used in developed countries primarily as an animal feed, it has been estimated up to 35% of world production of sorghum is a food staple for millions of people, especially in the semi-arid tropical regions (FAO, 1995; Awika and Rooney, 2004). According to FAO, 1995, countries in Africa and Asia account for 95% of the total food use of sorghum. Africa (36%) and the Americas (39%) lead the world in sorghum production accounting for approximately three quarters of world production between 2000-2013 (http://faostat3.fao.org/browse/Q/QC/E).

The total protein content and amino acid profile of sorghum is quite comparable to that present in maize or wheat (Hoseney, 1994). However, due to the unique biochemical and structural characteristics of sorghum proteins, wet heat treatment (i.e. cooking) results in substantially lower protein digestibility in comparison to other cereal grains (Duodu et al. 2003; El Nour et al. 1998; Emmambux and Taylor, 2009; Hamaker et al. 1987; Nunes et al. 2004; Nunes et al. 2005; MacLean et al. 1983; Oria et al. 1995; Oria et al. 2000). To date, there is no data providing a complete explanation for this occurrence. The following review examines what is known in relation to the proteins of sorghum and focuses on the research performed in the quest to determine all of the contributing factors involved in their unique digestibility characteristics.

Sorghum Grain

Great variation exists in the phenotypic expression of the many cultivars of sorghum grain; however in general, the sorghum kernel may be described as a spherical, naked caryopsis containing a kernel germ and endosperm. Kernel weights varying from between 16 and 30mg

(Laśztity, 1996; Bean et al. 2006) to 3 and 80mg (Waniska, 2000) have been reported, the former range being typical of commercial US varieties. The sorghum caryopsis includes a protective outer covering called the pericarp that exists in a variety of colors, including white, yellow, brown, red, black, and many subtle permutations of these colors. In sorghum, a testa layer resides beneath the pericarp and may or may not be pigmented. The presence of a pigmented testa layer is indicative of so-called tannin sorghums that contain proanthocyanidins (Waniska, 2000; Waniska et al. 1992). As not all sorghums contain a pigmented testa, not all sorghums contain tannins - a common misperception about sorghum (Dykes and Rooney, 2006). Tannins are powerful anti-oxidants and may have unique human health benefits (Awika and Rooney, 2004; Dykes and Rooney, 2006) as well as being beneficial from an agronomic standpoint by aiding in reduction of grain losses from crop pests and microflora, however numerous studies have shown the presence of condensed tannins have a negative impact on the nutritional availability of sorghum protein (Duodu et al. 2003; Taylor et al. 2007). An example of the relative proportions of the major constituents of a typical sorghum kernel were described by Hoseney (1994), and reported to compose 7.9% pericarp, 9.8% germ, and 82.3% endosperm. The endosperm to germ ratio of a typical sorghum kernel is 8.4:1 (FAO 1995).

Sorghum Proteins

Overview

The total protein content of sorghum grain is on par with that present in wheat or maize (Hoseney, 1994). However, variation in cultivation practices such as fertilization, as well as environmental conditions impact protein levels in sorghum. A mean value of 10 percent total protein is typical, with values commonly ranging from 7 to 15 percent (FAO, 1995; Branlard and Bancel, 2006).

The major protein classes in sorghum grain consist of albumins, globulins, kafirins, and glutelins (Hoseney, 1994; Branlard and Bancel, 2006). Nomenclature for determining sorghum protein class has been historically tied to solubility in various extraction solvent systems. Both the Osborne (Osborne, 1907) and the Landry and Moureaux (Landry and Moureaux, 1970) extraction schemes consist of sequential extractions with water (albumin), saline solution (globulin), aqueous alcohol (prolamin), and alkaline or acidic solutions (glutelin). Addition of reducing agents and detergents like sodium dodecyl sulfate (SDS) have allowed further refinement in the isolation of protein fractions exhibiting more specific properties within given protein classes (Jambunathan et al. 1975; Landry, 1997).

The albumin and globulin proteins are water and salt soluble fractions of sorghum protein, and are found in greatest abundance in the germ (Taylor and Schussler, 1986). Albumin and globulins of cereals have been found to contain proteins that are involved in metabolic processes of the maturing plant and plant defense compounds (Wall and Paulis, 1998; Singh et al. 2002). There is considerable variation in reported relative amount of this protein class, with average levels ranging from approximately 10% to 30+% (FAO, 1995; Wall and Paulis,1998; Taylor et al. 1984a). Albumin and globulin synthesis in the grain was initiated at 7 days after anthesis (Subramanian et al. 1990) and characterization of these proteins by SDS-PAGE has revealed numerous proteins in the 14 – 67 kDa range (Taylor and Schössler, 1986). The albumin and globulin fractions contain significantly higher levels of lysine than do the prolamins and high lysine sorghum lines have been found to contain higher levels of albumin and globulin proteins relative to normal sorghum lines (Taylor and Schössler, 1986; Guiragossian et al. 1978).

Kafirin, a prolamin and the major storage protein of sorghum, has been shown to comprise from 48% to slightly over 70% of the total protein in whole grain sorghum flour

(Taylor et al. 1984b; Hamaker et al. 1995). Typically obtained by extraction with aqueous alcohol and aqueous alcohol plus reducing agent, the kafirins are localized in the kernel endosperm in the form of protein bodies (Taylor et al. 1984b; Taylor et al. 1985; Shull et al. 1992; Chandrashekar and Mazhar, 1999, Shewry and Halford, 2002). Of note, over 30% of the total amino acid residues making up kafirin consist of proline and glutamine (Belton et al. 2006). Synthesis of prolamin in developing grain increased from 14 to 28 days after which there was a decline until maturity (Subramanian et al. 1990). The kafirins function as the major storage proteins of sorghum, and are localized almost exclusively in protein body structures within the kernel endosperm (Shewry and Halford, 2002).

The glutelin protein of sorghum functions as a structural element within the matrix of the peripheral and inner endosperm of the sorghum kernel, as well as a possible source of enzymes involved in starch and protein reserve hydrolysis (FAO, 1995; Taylor et al. 1984a). Beckwith (1972), and Nucere and Sumrell (1979) reported glutelin contents made up 40% to 50% of the total protein content of three sorghum varieties. Other researchers (Taylor et al. 1984a; Virupaksha and Sastry, 1986) found lower levels that ranged from 25% to 34% glutelin. According to Taylor et al. (1984a), such differences may reflect incomplete extractions during preceding sequential extraction steps. Increases in glutelin amounts within the developing grain stopped beyond 14 days after anthesis to maturity (Subramanian et al. 1990). SDS-PAGE of sorghum glutelins has shown proteins in the M_w of 20 – 67 kDa (Taylor, 1983).

Sorghum protein classes

The prolamins of cereals are composed of a homogenous mixture of monomeric, oligomeric, and polymeric protein species (Belton et al. 2006). As the major protein component

of grain sorghum, understanding the organizational composition of kafirin helps provide a rationale upon which to consider the potential nutritional and functional contributions of sorghum grain. Current consensus regarding kafirin subclass nomenclature follows from studies of maize prolamins (zeins). Kafirin and zein similarities based on solubility, molecular weight, structure, and amino acid sequence have been confirmed using comparisons of isolated fractions using SDS-PAGE and immunological cross reactivity (Shull et al. 1991; Shull et al. 1992; Mazhar et al. 1993a). The resulting kafirin subclass groups are known as the α-, β-, and γ-kafirins. An additional kafirin subclass category known as δ-kafirin has been proposed using homology of DNA sequences with δ-zein (Belton et al. 2006).

The α-kafirins form the majority of the kafirins and make up approximately 70-80% of total kernel prolamin protein content (Watterson et al. 1993). The α-kafirins predominate in both the vitreous and floury endosperm, and are reported to contain approximately 1 mol% cysteine (Shull et al. 1992). Utilizing differential solubility, SDS-PAGE, and immunological techniques, Shull et al. (1991), described the α-kafirins being comprised of two bands of M_w 23,000 and 25,000. Others (Mazhar et al. 1993a) have reported molecular weights of 28,000 and 22,000 using similar techniques, and subdivided α-kafirin nomenclature into α1- and α2-kafirins respectively. Although SDS-PAGE reveals only two major α-kafirin bands, it has been recently reported that up to 19 α-kafirin genes may be expressed in sorghum (Xu and Messing, 2008).

The γ-kafirins have been characterized based on differential solubility, SDS-PAGE, and immunological techniques and have been reported to comprise 9 to 12% of vitreous endosperm and 19 to 21% of opaque endosperm by (Shull et al. 1991; Watterson et al. 1993). Once reduced, the γ-kafirins are soluble in water and in aqueous organic solvents that span a wide range of polarity such as 10-80% tert-butanol (Shull et al. 1991), despite being the most hydrophobic of

the kafirins based on free energy of hydration (Belton et al. 2006). The odd solubility of the γ-kafirins may reflect the influence of other non-covalent factors such as electrostatic repulsion from high histidine content (Belton et al. 2006). Cross-reactivity with γ-zein for a migration band at Mr 28,000 provided confirmation of the molecular weight of γ-kafirin (Shull et al. 1991). The cross-linking potential of γ-kafirin is indicated by the cysteine content which is relatively high, having been reported as 7 mol% (Duodu et al. 2003). The extensive participation of γ-kafirin in the formation of kafirin oligomers and polymers has been corroborated in studies by El Nour et al. (1998) and Nunes et al. (2005).

The β-kafirins have also been characterized based on their solubility in 10-60% tert-butanol plus reducing agent in addition to cross-reactivity with β-zein antibodies and amino acid sequence, and three different Mr components have been identified as β-kafirins (16kDa, 18kDa, 20kDa) (Shull et al. 1991; Shull et al. 1992). Later studies by Chamba et al. (2005) using molecular cloning techniques identified a single gene encoding for a mature M_w 18,745 β-kafirin species. Containing 5.8 mol% cysteine and ten cysteine residues, this even number suggests the β-kafirin may be involved in intra and inter-chain disulfide bonding (El Nour et al. 1998). These same researchers found evidence for β-kafirin acting as an oligomer chain extender in predominantly higher molecular weight polymers. The vitreous endosperm has been reported to contain 7 to 8% β-kafirin versus 10 to 13% β-kafirin in the opaque endosperm (Watterson et al. 1993).

Possibly the least characterized of the kafirin sub-classes, especially at the protein level, are the δ-kafirins. By employing molecular cloning experiments of cDNA encoding for δ-kafirin, Izquierdo and Godwin (2005) were able to describe a 147 amino acid polypeptide (Mr 16,000) rich in methionine. Two δ-kafirin DNA sequences have been described by Belton et al.

(2006), GENPEPT AAK72689 and AAW32936, that showed extensive homology with M_w 14,000 δ-zein. Total seed storage protein in mature sorghum grain is thought to be made up of less than 1% δ-kafirin (Laidlaw et al. 2010).

Polymeric proteins and cross-linking

Potential kafirin protein interactions occur within the grain during development, storage, cooking, and digestion and a number of investigations have been undertaken to determine the conditions, order, and products of these events. Early studies applying the protein extraction method of Landry and Moureaux (1970) to sorghum grain resulted in two fractions localized to protein bodies that are now commonly referred to as kafirin-1 (in aqueous alcohol), and kafirin-2 (in aqueous alcohol plus reducing agent)(Taylor et al. 1984c; Mazhar and Chandrashekar, 1993b). The latter group tracked changes in these fractions, as well as their associated kafirin subclass composition over the course of endosperm development. They found α- and γ-kafirins increased early in development, with the β-kafirins appearing slightly later. As maturation proceeded, crosslinking between β- and γ-kafirins increased. It was noteworthy that how kafirin components were deposited in terms of rate, type, and content impacted subsequent storage protein degradation during germination. This may reflect differences in protein packaging that influence subsequent susceptibility to digestion processes.

Attempting to localize kafirin components within the endosperm of mature sorghum using immunocytochemical techniques, Shull et al. (1992) revealed that a differential subclass distribution existed in which α-kafirin was the major constituent of both peripheral and central endosperm. The β- and γ-kafirins were present in both endosperm portions, however a higher proportion were present in the central endosperm. This is interesting because the components

with greater crosslinking potential due to higher cysteine content (i.e. β- and γ-kafirins) appeared to predominate in the part of the endosperm exhibiting less rigid and compactly structured character (central "soft" endosperm). Oria et al. (1995) followed α-, β- and γ-kafirin synthesis over the course of kernel development and found that disulfide-bonded complexes formed from the γ- and β-kafirins occurred late in development after most of the kafirins had been synthesized. They attributed the most significant crosslinking initiation effect to maturation and grain dessication rather than to changes in relative α-, β- or γ-kafirin content.

A significant amount of research has focused on kafirin subclass composition involved in crosslinking with less emphasis on analysis of intact protein polymers. Such research is complicated by the fact that most protein analysis methods require analytes in solution. Extractability and large molecular size are in opposition to one another in terms of solubility, and thus, extractability. The challenge is to extract such cross-linked oligomers efficiently and with minimal to no alteration in native structure, so as to accurately reflect *in-vivo* character. An interesting study by El Nour et al. (1998) used an extraction protocol consisting of a sequence of alternating non-reducing t-butanol, reducing t-butanol, followed by non-reducing sonicated extraction in buffer + SDS conditions applied to whole grain sorghum flour with the goal of obtaining intact kafirin polymers for subsequent analysis by SDS-PAGE and SEC. Cross-linked kafirin oligomers containing α1-, α2- and γ-kafirins were obtained using only non-reducing t-butanol. β-kafirin on the other hand was not present in this extract with the exception of very slight trace amounts in the largest polymers (~100kDa). The sonicated extract, in contrast, contained mostly higher molecular weight oligomers that when reduced were shown to consist of α1-, β- and γ-kafirins, with α2-kafirin conspicuously absent. They concluded the degree of polymerization was a result of competitive disulfide linkage of α2- kafirin or β-kafirin to γ- and

α1-kafirin. As a result, β-kafirin was termed a "chain extender", and α2-kafirin a chain "terminator".

A key concept briefly alluded to in several previous reports (El Nour et al. 1998; Lending et al. 1988; Holding, 2014) and worth expanding upon is that of spatial-temporal associations as they apply to interactions between kafirin proteins and crosslinking of developing polymers. The evidence put forth in studies on polymeric sorghum proteins to date would seem to imply that an important factor regarding polymer composition and size may, in fact, relate not to quantity of individual kafirin components, but rather to what components are available at a given time (temporal) and place (spatial-) within the endosperm of given sorghum (El Nour et al. 1998; Lending et al. 1988; Holding, 2014). As a simplified metaphor, a wall made up of a variety of colored bricks will take on an overall character based on the color and size of brick available (the "where") for the bricklayer to use at a given time (the "when"). It is conceivable that different sorghum varieties express and assemble their protein building blocks (kafirins) based on different spatial-temporal patterns, resulting in differences in the protein polymer "walls", and the consequent contribution to differences in digestibility, endosperm character, and agronomic properties.

Sorghum protein body structure

The endosperm is by far the major repository of sorghum kernel protein. Nutrient availability and functional attributes of sorghum grain protein are greatly affected by the form in which this protein exists within the endosperm. Seed storage proteins evolved as a means of storing nitrogen over potentially long periods of time for later use during plant reproduction and

development, and in mature cereal grains may represent from 50% to more than 80% of total protein (Hamaker et al. 1995; Shewry and Halford, 2002).

One mechanism for achieving nitrogen storage in the seeds of plants was the development of specialized organelles known as protein bodies (Shewry et al. 1995). Amino acids needed for growth at germination can be stored for years, protected in the membrane-bound protein bodies (Müntz, 1998). Formed in the endoplasmic reticulum (ER), the protein bodies consist almost entirely of prolamin proteins (Müntz, 1998; Herman and Larkins, 1999; Seckinger and Wolf, 1973; Taylor et al. 1984b). The protein body prolamins develop within the endoplasmic reticulum (ER) in the form of large oligomeric aggregates (Herman and Larkins, 1999). From a species evolutionary standpoint, the accumulation of large aggregates within the ER is very unusual, but may have been possible due to the programmed cell death the endosperm undergoes during late stages of cereal seed development (Dominguez and Cejudo, 2014; Mainieri et al. 2014). It has been proposed that proteins in the interior of the protein bodies are predominantly comprised of α-kafirin along with smaller quantities of β- and γ-kafirin (Shull et al. 1992). The periphery of the protein body on the other hand, is thought to contain an abundance of the β- and γ-kafirins. Containing high levels of cysteine, it is postulated the β- and γ-kafirins could form a cross-linked shell around the α-kafirins in the interior of the protein body (Belton et al. 2006). Protein bodies located in the vitreous endosperm were described by Shull et al. (1992) as 0.3 to 1.5μm spheroids. Likewise, the floury endosperm protein bodies were similar if somewhat smaller in size, but exhibited somewhat irregular shape.

Although a great deal of research has been done to elucidate the biochemical makeup and fine structure of cereal protein bodies, information regarding the sequence of complex molecular interactions, as well as specific details concerning the nature of the actual protein polymers

involved in protein body formation are lacking (Manieri et al. 2014). Investigations utilizing differential solvation have been applied to isolated protein bodies (Taylor et al. 1984) providing data about kafirin localization and amino acid composition, but sample variety and extracted protein details were limited. Other studies comparing the kafirin contents of uncooked and cooked sorghum samples showed how crosslinking increased over the course of development, but did not elaborate with details on how specific protein polymers relate to protein body structure (Oria et al. 1995).

Transmission and scanning electron microscopy studies of protein bodies involving post-sectioning kafirin extractions of mature endosperm (Seckinger and Wolf, 1973), and protein body imaging over the course of seed development allowed visualization of protein body location and ultrastructure (Taylor et al. 1985). Holding and Larkins (2005) in a review on zein protein bodies provided examples of mutations in maize zein synthesis resulting in altered kernel hardness and associated nutritional effects. Although the reviewed studies illustrated that alterations to single zein component synthesis was capable of profoundly altering protein body ultrastructure, information was lacking that directly connects polymeric protein composition to protein body structure.

Immunocytochemical methods applied to maize endosperm (Lending and Larkins, 1989) and to isolated maize protein bodies (Lending et al. 1988) provided data upon which a protein body model for maize was postulated, and by extension for sorghum. Widely cited in zein and kafirin protein studies, this model illustrates a γ- + β-zein cross-linked protein body shell surrounding a central region consisting primarily of α-zein as well as a minor proportion of δ-zein. Although aiding the rationalization of the general property of decreased protein

digestibility, authors of the studies cautioned "the relative amounts and distribution of these proteins varies substantially among different protein bodies" (Lending et al. 1988).

A review of recent technological advancements in the molecular characterization of prolamin deposition in mutants of maize and sorghum highlighted new information on the genetic basis for changes in endosperm texture related to protein body structure and prolamin synthesis (Holding, 2014). Besides genetic mapping of defects associated with prolamin accumulation and packaging into ER localized protein bodies, opaque mutant characterization studies have suggested roles for involvement of indirect protein body related organizational factors such as ER membrane-specific proteins and other products of unfolded protein response (UPR). Luminal binding protein (BiP) is a 75kDa maize protein plant homolog of mammalian binding protein able to recognize and bind incorrectly folded or unfolded polypeptide chains (Flynn et al. 1989). Immunomicroscopy studies published by Zhang and Boston (1992) found highly elevated levels of BiP in the abnormal protein bodies of three maize endosperm mutants. Amounts of BiP were positively correlated with degree of abnormal protein body morphology. The authors concluded the association might reflect a biological function to mediate protein folding and assembly in maize endosperm. In addition to those already discussed, other protein organizational factors in sorghum were characterized by mass spectrometry in a proteomics study by Cremer et al. (2014). Among such factors were thioredoxins, glutaredoxins, protein disulfide isomerases, peptidyl-prolyl cis/trans isomerases, and heat shock proteins. Each of these non-prolamin proteins function to help control protein crosslinking through redox activity or as mediators of protein polymer folding during grain development or environmental stress. These studies suggest protein body assembly in wild type sorghum may be at least partially mediated via transient binding processes of chaperone proteins like BiP and other ER membrane-specific

proteins. If this is indeed the case, differences among sorghum varieties with regard to protein body assembly mediators could have profound impact on mature protein body structure, and resulting implications for variation in sorghum protein digestibility.

An additional endosperm protein component also deriving from the ER, the endosperm matrix protein appears to provide a connecting structure within which the protein bodies and starch granules reside (Chandrashekar and Mazhar, 1999). In addition to protein storage, the matrix probably also functions as an enzyme source for starch and protein hydrolysis (Wu and Wall, 1980; Taylor et al. 1984b). On a quantitative basis, the protein matrix was considered the second most important endosperm protein fraction in a study by Taylor and Schüssler, 1986. The primary composition of the protein matrix appears to be in the form of glutelins based on solubility characteristics and amino acid composition (Taylor et al. 1984b; Taylor and Schüssler, 1986).

Sorghum Protein Digestibility

Factors involved

Factors involved in sorghum protein digestibility have been divided into two general areas (Duodu et al. 2003). Exogenous influences include grain microstructure, polyphenol content, phytic acid content, and starch and non-starch polysaccharides. Endogenous factors involve disulfide and non-disulfide crosslinking, kafirin hydrophobicity, and alterations in kafirin secondary structure. One study on the effect of grain structure and cooking on sorghum and maize protein digestibility (Duodu et al. 2002) found additional factors for consideration including pericarp and germ components, endosperm cell walls and gelatinized starch, but their effects on digestibility were described as minor. In addition, the presence of non-tannin polyphenols showed no involvement as well. The prevailing consensus is that more than one

factor is likely influencing protein digestibility at any given time, but sorghum digestibility appears to be most affected by the endogenous factor of protein crosslinking (Duodu et al. 2003).

Polymeric proteins and digestibility

What actually happens to sorghum protein bodies from a microstructural standpoint when subjected to digestion conditions? This was the subject of a study in which uncooked and cooked sorghum flour was subjected to pepsin digestion conditions with or without prior reducing agent treatment (Rom et al. 1992). Subsequent SEM scans allowed visualization of the effects of the treatments on the protein bodies. Fifteen minutes of exposure to pepsin of uncooked flour with no prior treatment with reducing agent resulted in the appearance of small pits on the surface of the protein bodies. Increased pepsin exposure times up to a final time of 120min showed more extensive pitting and loss of spherical shape. In contrast, exposure to reducing agent alone resulted in no change in protein body appearance even at 120min of exposure. Similar results were seen in another study involving isolated protein bodies exposed to extracted proteinase from germinated sorghum (Taylor and Evans, 1989). Protein body degradation by pitting or the appearance of holes starting from the periphery was observed by TEM. Degradation of the protein body surface by pitting implies certain areas are susceptible to digestion breakdown while others are not. In light of this data, a uniform digestion-resistant shell of γ- + β-kafirin polymers as described in the widely cited protein body model discussed earlier does not seem adequate for an accurate depiction of how these polymers are associated. It would seem some areas of the protein body surface are digestion resistant while others are not, begging the question, what attributes of protein polymer structure and association ("packaging") are responsible for the differential surface digestion results seen in these studies?

Many of the studies already discussed have noted the negative effect wet heat treatment (cooking) has on sorghum protein digestibility. Most of these mention increased covalent kafirin crosslinking in the form of disulfide bonding resulting from cooking treatments as a primary factor in decreased digestibility (Duodu et al. 2003; El Nour et al. 1998; Hamaker et al. 1987). However, other changes to the proteins accompanying wet heat treatment have been noted, and provide additional clues to other protein polymer associations that may be important to the nutritional value of sorghum. For example, the formation of complex prolamin networks composed of reduction resistant protein polymers in cooked sorghum (Mr > 100kDa) were found to relate to accompanying decreases in the digestibility of starch in the same samples (Ezeogu et al. 2005). The authors surmised the observed reduction in starch digestibility was related to decreased accessibility of alpha-amylase.

Nunes et al. (2004) investigated IVPD within a set of sorghum and maize samples using a sequential procedure to see the effect of cooking on remaining undigested proteins at different times over the course of 120min in-vitro digestions. At given times over the course of digestion, the undigested proteins were extracted under reducing conditions and analyzed by SDS-PAGE. They found cooked IVPD did not correlate to uncooked IVPD, and that the less digestible cooked sorghum samples exhibited higher amounts of γ- and α-kafirin monomers remaining after digestion. Perhaps most interestingly, amounts of unidentified non-reducible 45 and 47kDa proteins were significantly negatively correlated with digestibility, indicating that non-disulfide linked proteins may play a role in digestibility.

In another study, Nunes et al. (2005) investigated the effect of cooking on non-reduced t-butanol kafirin extracts. HMW aggregates (80 – 200kDa), 66 and 45kDa oligomers, and α-, β-

and γ-monomers were found in uncooked sorghum samples. Interestingly, in the cooked sorghums the two oligomers (66 and 45kDa) were found to be partially reduction resistant.

Seeking to characterize intact polymeric proteins in heat treated sorghum and maize meals, Emmambux and Taylor (2009) performed non-reduced extractions of the post-heat treated prolamins with t-butanol, followed by SEC to fractionate the extracted proteins. The fractions obtained were collected and analyzed by SDS-PAGE in reduced and non-reduced form. As in other studies, cooking was found to result in formation of disulfide bonded polymers (M_w > 200kDa) for sorghum but not for maize. However, higher Mr bands found by SDS-PAGE in reduced kafirin showed that other non-disulfide mediated protein interactions may be contributing to observed digestibility differences. This study was limited by the non-reduced extraction conditions used which resulted in only 35% of available protein being extracted, thereby providing only a partial picture of what protein interactions take place.

A high molecular weight (HMW) γ-prolamin homologue exhibiting homology to 50kDa γ-zein, γ-canein, and γ-coixin and represented by two peaks from lab-on-a-chip (LOC) analysis (~44 and 46kDa) was characterized in a study utilizing an integrated proteomics approach to identify proteins with potential impact on sorghum quality traits (Cremer et al. 2014). Persisting even in reduced alcohol solubilized samples, the study's authors speculated this previously uncharacterized HMW γ-prolamin homologue may prove to correlate with sorghum grain quality parameters.

Some have speculated the existence of a distinct subgroup of proteins formed from kafirin monomers capable of forming high molecular weight polymers under non-reducing conditions that are unique to more undigestible sorghums (Belton et al. 2006). Further, they suggest the possibility the constituent monomers may differ in some way from kafirin monomers

that do not participate in the formation of such high molecular weight polymers. Addressing ideas like these will require appropriate new extraction protocols that allow more complete extraction of intact native protein polymers, as well as analytical methods that provide ways of more accurately assessing the protein components from which such polymers are constructed.

Transgenic study implications to digestibility

The distribution of individual zeins and kafirins in the formation of endosperm protein bodies and the effect on protein digestibility has been the topic of many studies, including those involving endosperm mutants (Lending and Larkins, 1992). Other researchers utilized multiple *in-vitro* methods for assessing protein and starch digestibility, applying the concepts of principal component analysis for determination of statistically significant contributing factors (Wong et al. 2010). These studies and others have concluded that protein body assembly or protein digestibility cannot be adequately explained simply by the presence of altered amounts of zein or kafirin subclass proteins. Their data argues for the involvement of additional factors that in some manner contribute to protein assembly control or that influence protein packaging. The focus in the aforementioned studies, as well as many others, has been on the proportions of individual prolamins in relation to protein digestibility. While informative, this research does not address the potential for variation (e.g. changes within primary amino acid sequence) within the cross-linked polymers of kafirins as a consequence of allelic variation or environmentally influenced changes in phenotypic protein expression present within different sorghum genotypes. Variation at the gene level in the α-kafirins is certainly indicated, as it is known to be encoded by a multigenic family of 23 genes of which 19 are known to be expressed (Xu and Messing, 2008). Potential for variation was also reported in a study by Laidlaw et al. (2010), in which the β-, γ-,

and δ-kafirin genes from a 35 sample set of sorghums were sequenced to investigate the allelic diversity of storage proteins. Six alleles were identified for β- and γ-kafirin, and three alleles were found for δ-kafirin.

Borrowing from an earlier metaphor, the characteristics of a wall (protein body) made of bricks (kafirins) will be influenced by the composition of those bricks. The wall may be strong or weak, permeable or solid, influenced at least in part by variations within, and spatial arrangements of, constituent bricks. Variability in the composition of available kafirins potentially influences the character (surface hydrophobicity, surface protein domain composition, non-covalent crosslinking, etc.) of the protein body they become a part of and, in turn, demonstrate the importance of their characterization in relation to understanding protein digestibility.

A series of studies performed by a sorghum research group in South Africa (da Silva et al. 2011a; da Silva et al. 2011b; Grootboom et al. 2014) used transgenic technologies to alter the expression of selected kafirin subclasses to investigate the effects on protein digestibility and protein body morphology. Compositing the findings from the studies appears to confirm some previously postulated ideas regarding the role individual kafirins play in the digestibility of cooked and uncooked sorghum proteins, plus reveal new details about the apparent importance of kafirin interactions to protein body formation. Suppressing the formation of γ-kafirin-1 (25kDa) and γ-kafirin-2 (50kDa) improved protein digestibility, and suppressing more than a single kafirin subclass resulted in further digestibility improvement. The authors concluded that optimal improvement in digestibility required suppression of the high cysteine content kafirins to realize maximum effects. Their conclusion agrees with the bulk of previous research (Duodu et al. 2003; El Nour et al. 1998; Hamaker et al. 1987) attributing disulfide mediated crosslinking as

a major factor in sorghum protein digestibility. They also found that co-suppression of several kafirin sub-classes resulted in floury grain endosperm, and that it required the suppression of just three kafirins (γ-kafirin-1, γ-kafirin-2, α-kafirin A1(25kDa)) to disrupt normal protein body formation. Grootboom et al. (2014) provided data showing that the suppression of these three genes alters the cross-linking of kafirin protein leading to an irregularly invaginated protein body phenotype. The relationship between kafirin crosslinking and protein body structure is a critical point from this work, and warrants expanded attention, particularly in light of the implied connection to protein digestibility revealed in these studies.

Scope of the current study (4 parts)

Because sorghum grain hardness is an important quality trait in terms of processing (milling) and mold and insect resistance, knowledge regarding protein cross-linking of sorghum proteins into larger polymeric groups would be beneficial. To achieve this, a differential extraction method using non-reduced and reducing techniques was applied to physically separated vitreous and floury endosperm to find differences representative of these two endosperm types. Results provided information indicating the vitreous endosperm contained higher levels of polymeric proteins, greater cross-linking, and a larger M_W distribution. However, because one portion of the extracted protein fractions was required to be reduced, cross-linking information was lost, and data critical to accurately assessing polymeric protein character was unavailable.

To address this problem, a better way of extracting native intact sorghum protein polymers was needed. The challenge was to solubilize large proteins while minimizing impact

on polymer structure in order to be able to determine the representative polymeric protein content of wild-type protein varieties. A multi-step sequential procedure involving no reducing agents was developed, and optimized for protein extraction and sample stability. Chromatographic separation conditions were optimized in terms of column and mobile phase composition. SEC analysis indicated extracted fractions represented proportionally different protein polymer contents and importantly, cross-linking information for all extracts was preserved.

The improved extraction method was applied to a diverse sample set with widely varying IVPD to see if protein cross-linking relationships to IVPD could be determined in wild-type (non-mutant) sorghums. The results indicated the final protein fraction obtained from the extraction sequence contained proteins with the largest M_W distribution, and was significantly negatively correlated to IVPD. Additionally, four of the five SEC peaks representing different extracted proteins from the final protein fraction were significantly negatively correlated to IVPD, with three of that four being cross-linked. The results place emphasis for IVPD impact on the polymeric proteins from the final fraction, and imply that how the sorghum protein subclass monomers are packaged within the protein body may play a crucial role in IVPD differences.

Data from the preceding study parts showed significant relationships between IVPD and some but not all of the extracted protein polymers. The question became, what was special or unique about these proteins? An attempt was made to discern differences among extracted protein polymers by applying the 2-D orthogonal analytical technique of SEC x RP-HPLC. These complementary analytical techniques allowed the disassembly of protein polymers into monomer components unique for each cross-linked protein present in the extracts. Statistically significant relationships to IVPD and 2-D peaks in the γ-kafirin region were found, but only in

the largest protein separated from the final extracted fraction. This implies that differences in IVPD may be based on how the protein components are packaged, rather than on simple kafirin subclass content.

References

Awika, J.M., Rooney, L.W. (2004). Sorghum phytochemicals and their potential impact on human health. Phytochemistry, 65:1199-1221.

Bean, S.R., Ioerger, B.P., Smith, B.M., Blackwell, D.L. (2011). Chpt.7: Sorghum protein structure and chemistry: Implications for nutrition and functionality. *In*: Advances in Cereal Science: Implications to Food Processing and Health Promotion. ACS Symposium Series 1089. Eds.: J. Awika, V. Piironen, S. Bean. Oxford Univ. Press.

Bean, S., Ioerger, B.P. (2015). Chpt.16: Sorghum and millet proteins. *In*: Applied Food Protein Chemistry. Ed.: Z. Ustunol. John Wiley & Sons, Ltd., UK.

Bean, S.R., Chung, O.K., Tuinstra, M.R., Pedersen, J.F., Erpelding, J. (2006). Evaluation of the single kernel characterization system (SKCS) for measurement of sorghum grain attributes. Cereal Chem., 83:108-113.

Beckwith, A.C. (1972). Grain sorghum glutellin: isolation and characterization. J. Agric. Food Chem., 20:761-764.

Belton, P.S., Delgadillo, I., Halford, N.G., Shewry, P.R. (2006). Kafirin structure and functionality. J Cereal Sci., 44:272-286.

Branlard, G., Bancel, E. (2006). Plant Proteomics: Methods and Protocols. Eds.: Thiellement, H., Zivy, M., Damerval, C., Méchin, M. Humana Press: Totowa, NJ, pp 15-25.

Chamba, E.B., Halford, N.G., Forsyth, J., Wilkinson, M., Shewry, P.R. (2005). Molecular cloning of β-kafirin, a methionine-rich protein of sorghum grain. J. Cereal Sci., 41:381-383.

Chandrashekar, A., Mazhar, H. (1999). The biochemical basis and implications of grain strength in sorghum and maize. J. Cereal Sci. 30:193-207.

Cremer, J.E., Bean, S.R., Tilley, M.M., Ioerger, B.P., Ohm, J.B., Kaufman, R.C., Wilson, J.D., Innes, D.J., Gilding, E.K., Godwin, I.D. (2014). Grain sorghum proteomics: integrated approach toward characterization of endosperm storage proteins in kafirin allelic variants. J. Agric. Food Chem. 62:9819-9831.

da Silva, L.S., Jung, R., Zhao, Z., Glassman, K., Taylor, J., Taylor, J.R.N. (2011a). Effect of suppressing the synthesis of different kafirin sub-classes on grain endosperm texture, protein body structure and protein nutritional quality in improve sorghum lines. J. Cereal Sci. 54:160-167.

da Silva, L.S., Taylor, J., Taylor, J.R.N. (2011b). Transgenic sorghum with altered kafirin synthesis: kafirin solubility, polymerization, and protein digestion. J. Agric. Food Chem. 59:9265-9270.

Dominguez, F., Cejudo, F.J. (2014). Programmed cell death (PCD): an essential process of cereal seed development and germination. Front. Plant Sci. 5:366. doi:103389/fpls.2014.00366

Duodu, K.G., Taylor, J.R.N., Belton, P.S., Hamaker, B.R. (2003). Factors affecting sorghum protein digestibility. J. Cereal Sci 38:117-131.

Dykes, L., Rooney, L.W. (2006). Sorghum and millet phenols and anti-oxidants. J. Cereal Sci 44, 236-251.

El-Nour, I.N.A., Peruffo, A.D.B., Curioni, A. (1998). Characterization of sorghum kafirins in relation to their cross-linking behavior. J. Cereal Sci 28:197-207.

Emmambux, M.N., Taylor, J.R.N. (2009). Properties of heat-treated sorghum and maize meal and their prolamin proteins. J. Agric. Food Chem. 57:1045-1050.

Ezeogu, L.I., Duodu, K.G., Taylor, J.R.N. (2005). Effects of endosperm texture and cooking conditions on the in vitro starch digestibility of sorghum and maize flours. J. Cereal Sci. 42:33-44.

FAO. (1995). Chpt.1 Grains and their structure. Sorghum and millets in human nutrition. (FAO Food and Nutrition Series,No.27); www.fao.org/docrep/T0818E/T0818E00.htm.

FAO. http://faostat3.fao.org/browse/Q/QC/E

Flynn, G.C., Chappell, T.G., Rothman, J.E. (1989). Peptide binding and release by proteins implicated as catalysts of protein assembly. Science 245:385-390.

Grootboom, A.W., Mkhonza, N.L., Mbambo, Z., O'Kennedy, M.M., da Silva, L.S., Taylor, J., Taylor, J.R.N., Chikwamba, R., Mehlo, L. (2014). Co-suppression of synthesis of major α-kafirin sub-class together with γ-kafirin-1 and γ-kafirin-2 required for substantially improved protein digestibility in transgenic sorghum. Plant Cell Rep. 33:521-537.

Guiragossian, V., Chibber, B. A. K., Scoyoc, S. V., Jambunathan, R., Mertz, E. T., Axtell, J. D. (1978). Characteristics of proteins from normal, high lysine, and high tannin sorghums. J. Agric. Food Chem. *26*, 219-223.

Hamaker, B.R., Kirleis, A.W., Butler, L.G., Axtell, J.D., Mertz, E.T. (1987). Improving the in-vitro protein digestibility of sorghum with reducing agents. Proc. Natl. Acad. Sci. USA 84:626-628.

Hamaker, B.R., Kirleis, A.W., Mertz, E.T., Axtell, J.D. (1986). Effect of cooking on the protein profiles and in vitro digestibility of sorghum and maize. J. Agric. Food Chem. 34:647-649.

Hamaker, B.R., Mohamed, A.A., Habben, J.E., Huang, C.P., Larkins, B.A. (1995). Efficient procedure for extracting maize and sorghum kernel proteins reveals higher prolamin contents than the conventional method. Cereal Chem. *72*, 583-588.

Herman, E.M., Larkins, B.A. (1999). Protein storage bodies and vacuoles. The Plant Cell 11:601-613.

Holding, D.R., Larkins, B.A. (2006). The development and importance of zein protein bodies in maize endosperm. Maydica 51:243-254.

Holding, D.R. (2014). Recent advances in the study of prolamin storage protein organization and function. Frontiers in Plant Sci. 5(276):1-9. doi: 10.3389/fpls.2014.00276

Hoseney, R.C. (ed). (1994). Principles of cereal science and technology. 2nd edn, AACC, St. Paul, MN; pp 65-79.

Izquierdo, L., Godwin, I.D. (2005). Molecular characterization of a novel methionine-rich δ-kafirin seed storage protein gene in sorghum (Sorghum bicolor L.). Cereal Chem. 82:706-710.

Jambunathan, R., Mertz, E.T., Axtell, J.D. (1975). Fractionation of soluble proteins of high lysine and normal sorghum grain. Cereal Chem. 52:119-121.

Laidlaw, H.K.C., Mace, E.S., Williams, S.B., Sakrewski, K., Mudge, A.M., Prentis, P.J., Jordan, D.R., Godwin, I.D. (2010). Allelic variation of the β-, γ- and δ-kafirin genes in diverse sorghum genotypes. Theor. Appl. Genet. 121:1227-1237.

Landry, J., Moureaux, T. (1970). Heterogeneity of corn seed glutelin: selective extraction and amino acid composition of the 3 isolated fractions. Bull. Soc. Chim. Biol. 52:1021–1037.

Landry, J. (1997). Comparison of extraction methods for evaluating zein content of maize grain. Cereal Chem. 74:188-189.

Lásztity, R. (1996). The chemistry of cereal proteins. Ed.: R. Lásztity. CRC Press, Boca Raton, FL.

Lending, C.R., Kriz, A.L., Larkins, B.A., Bracker, C.E. (1988). Structure of maize protein bodies and immunocytochemical localization of zeins. Protoplasma 143:51-62.

Lending, C.R., Larkins, B.A. (1989). Changes in the zein composition of protein bodies during maize endosperm development. The Plant Cell 1:1011-1023.

Lending, C.R., Larkins, B.A. (1992). Effect of the *floury-2* locus on protein body formation during maize endosperm development. Protoplasma 171:123-133.

Maclean, W.C., De Romano, G.L., Gastanaduy, A., Graham, G.G. (1983). The effect of decortication and extrusion on the digestibility of sorghum by preschool children. J. Nutrit. 2071-2077.

Manieri, D., Morandini, F., Maitrejean, M., Saccani, A., Pedrazzini, E., Vitale, A. (2014). Protein body formation in the endoplasmic reticulum as an evolution of storage protein

sorting to vacuoles: insights from maize γ-zein. Frontiers in Plant Sci. 5(331):1-11. doi: 10.3389/fpls.2014.00331

Mazhar, H., Chandrashekar, A., Shetty, H.S. (1993a). Isolation and immunochemical characterization of the alcohol-extractable proteins (kafirins) of Sorghum bicolor (L.) Moench. J. Cereal Sci. 17:83-93.

Mazhar, H., Chandrashekar, A. (1993b). Differences in kafirin composition during endosperm development and germination in sorghum cultivars of varying hardness. Cereal Chem. 70(6):667-671.

Müntz, K. (1998). Deposition of storage proteins. Plant Mol. Biol. 38:77-99.

Neucere, N.J., Sumrell, G. (1979). Protein fractions from five varieties of grain sorghum: amino acid composition and solubility properties. J. Agric. Food Chem. 27:809-812.

Nunes, A., Correia, I., Barros, A., Delgadillo, I. (2004). Sequential in vitro pepsin digestion of uncooked and cooked sorghum and maize samples. J. Agric Food Chem. 52:2052-2058.

Nunes, A., Correia, I., Barros, A., Delgadillo, I. (2005). Characterization of kafirin and zein oligomers by preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis. J. Agric. Food Chem. 53:639-643.

Oria, M.P., Hamaker, B.R., Axtell, J.D., Huang, C-P. (2000). A highly digestibile sorghum mutant cultivar exhibits a unique folded structure of endosperm protein bodies. Proc. Natl. Acad. Sci. 97:5065-5070.

Oria, M.P., Hamaker, B.R., Schull, J.M. (1995). In-vitro protein digestibility of developing and mature sorghum grain in relation to α-, β-, and γ-kafirin disulfide crosslinking. J. Cereal Sci. 22:85-93.

Osborne, T. B. (1907). The proteins of the wheat kernel. Carnegie Institute, Washington, D.C.

Rom, D.L., Shull, J.M., Chandrashekar, A., Kirleis, A.W. (1992). Effects of cooking and treatment with sodium bisulfite on in-vitro protein digestibility and microstructure of sorghum flour. Cereal Chem. 69(2):178-181.

Seckinger, H.L., Wolf, M.J. (1973). Sorghum protein ultrastructure as it relates to composition. Cereal Chem. 50:455-465.

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

Shewry, P.R., Halford, N.G. (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. J. Exp. Botany. 53:947-958.

Shull, J.M., Watterson, J.J., Kirleis, A.W. (1991). Proposed nomenclature for the alcohol-soluble proteins (kafirins) of Sorghum bicolor (L. Moench) based on molecular weight, solubility, and structure. J. Agric. Food Chem. 39:83-87.

Shull, J.M., Watterson, J.J., Kirleis, A.W. (1992). Purification and immunocytochemical localization of kafirins in Sorghum bicolor (L. Moench) endosperm. Protoplasma 171:64-74.

Singh, J., Blundell, M., Tanner, G., Skerrit, J.H. (2002). Albumin and globulin proteins of wheat flour: immunological and n-terminal sequence. J. Cereal Sci. 34:85-103.

Subramanian, V., Seetharama, N., Jambunathan, R., Venkateswara, R. P. (1990). Evaluation of protein quality of sorghum [*Sorghum bicolor* (L.) Moench]. J. Agric. Food Chem. 38:1344-1347.

Taylor, J. R.N. (1983). Effect of malting on the protein and free amino nitrogen composition of sorghum. J. Sci. Food Agric. 34:885-892.

Taylor, J., Bean, S.R., Ioerger, B.P., Taylor, J.R.N. (2007). Preferential binding of sorghum tannins with gamma kafirin and the influence of tannin binding on kafirin digestibility and biodegradation. J. Cereal Sci. 46:22-31.

Taylor, J.R.N., Evans, D.J. (1989). Action of sorghum proteinase on the protein bodies of sorghum starchy endosperm. J. Exp. Botany 40(216):763-768.

Taylor, J.R.N., Schüssler, L., van der Walt, W.H. (1984a). Fractionation of proteins from low-tannin sorghum grain. J. Agric. Food Chem. 32:149-154.

Taylor, J.R.N., Novellie L., Liebenberg, N.v.d.W. (1984b). Sorghum protein body composition and ultrastructure. Cereal Chem. 61(1):69-73.

Taylor, J.R.N., Novellie L., Liebenberg, N.v.d. W. (1984c). Location of zein-2 and cross-linked kafirin in maize and sorghum protein bodies. J. Cereal Sci. 2:249-255.

Taylor, J.R.N., Schüssler, L., Liebenberg, N.v.d.W. (1985). Protein body formation in the starchy endosperm of developing *Sorghum bicolor* (L.) Moench seeds. S. Afr. J. Botany 51(1): 35-40.

Taylor, J., Schüssler, L. (1986). The protein compositions of the different anatomical parts of sorghum grain. J. Cereal Sci. 4:361-369.

Virupaksha, T.K., Sastry, L.V.S. (1986). Studies on the protein content and amino acid composition of some varieties of grain sorghum. J. Agric. Food Chem. 16:199-203.

Wall, J.S., Paulis, J.W. (1998). Corn and sorghum grain proteins. In: Advances in Cereal Science & Technology, Pomeranz, J. Ed. American Association of Cereal Chemists: St. Paul, MN, Vol. II, pp 135-219.

Waniska R.D. (2000). Structure, phenolic compounds, and antifungal proteins of sorghum caryopses. *In*: Technical and institutional options for sorghum grain mold management: proceedings of an international consultation. Eds.: A. Chandrashekar, R. Bandyopadhyay R, A.J. Hall, pp72-106. May 18-19; ICRISAT, Patancheru, India. Andhra Pradesh, India.

Waniska, R.D., Hugo, L.F., Rooney, L.W. (1992). Practical methods to determine the presence of tannins in sorghum. J. Appl. Poultry Res. 1:122-128.

Watterson, J.J., Shull, J.M., Kirleis, A.W. (1993). Quantitation of α-, β-, and γ-kafirins in vitreous and opaque endosperm of Sorghum bicolor. Cereal Chem. 70:452-457.

Wong, J.H., Marx, D.B., Wilson, J.D., Buchanan, B.B., Lamaux, P.G. (2010). Principal component analysis and biochemical characterization of protein and starch reveal primary targets for improving sorghum grain. Plant Sci. 179:598-611.

Wu, Y.V. (1978). Protein concentrate from normal and high-lysine sorghums: Preparation, composition, and properties. J. Agric. Food Chem. 26:305-309.

Wu, Y.V., Wall, J.S. (1980). Lysine content of protein increased by germination of normal and high-lysine sorghums. J. Agric. Food Chem. 28:455-458.

Xu, J.-H., Messing, J. (2008). Organization of the prolamins gene family provides insight into the evolution of the maize genome and gene duplication in grass species. PNAS 105:14330-14335.

Zhang, F., Boston, R.S. (1992). Increase in binding protein (BiP) accompany changes in protein body morphology in three high-lysine mutants of maize. Protoplasma 171:142-152.

Chapter 2 - Characterization of Polymeric Proteins from Vitreous and Floury Endosperm

Chapter 2 is published as:

Ioerger, B.P., Bean, S., Tuinstra, M.R., Pedersen, J.F., Erpelding, J.E., Lee, K.M., Herrman, T.J. 2007. Journal of Agricultural and Food Chemistry 55:10232-10239.

Abstract

Differences in protein content and composition between vitreous and floury endosperm were investigated using a number of different techniques. Differences in protein cross-linking between vitreous and floury endosperm were investigated using differential solubility, size exclusion chromatography (SEC), and analysis of sulfhydryl content and composition. Vitreous endosperm was found to have higher levels of total protein and kafirins, but floury endosperm had a higher proportion of γ-kafirins than the vitreous. Floury endosperm was found to have higher levels of SDS soluble proteins than SDS insoluble proteins extracted using sonication than vitreous endosperm. Conversely, vitreous endosperm had a greater proportion of the insoluble proteins. SEC analysis of the polymeric proteins revealed that the insoluble proteins had more polymeric proteins than did the soluble proteins, indicating greater cross-linking and a larger M_w distribution. Vitreous endosperm was also found to have a greater percentage (i.e. a higher ratio of disulfide to total sulfhydryls) of disulfide bonds than floury endosperm. These results show that the proteins in vitreous endosperm have a higher degree of cross-linking and a greater M_w distribution than those found in floury endosperm.

Introduction

Grain hardness is an important grain quality attribute that plays a role in the processing of cereal grains and in the end-use quality of cereal grain based products such as breads and snack foods [1, 2]. Grain hardness also plays a role in plant defense against molds and from insect attack [3]. Therefore, grain hardness is an important economic and end-use quality trait in cereal grains.

Endosperm hardness in maize and sorghum has been positively correlated with both protein content and protein composition [3, 4-9]. The most abundant endosperm proteins of maize and sorghum are the prolamins [10, 11] which have been divided into subclasses based on solubility, structure, and amino acid sequence [12]. The α-subclass comprises the majority of the kafirins and makes up between 60-70% of total protein [10, 11]. The α-prolamins are located primarily in the interior of protein bodies while the β-and γ-prolamins are present on the outer edges of the protein bodies [3, 5, 7].

Past research has associated specific subclasses of kafirin proteins with grain hardness in sorghum and maize. The majority of publications in this area have reported that the vitreous endosperm of these grains is higher in total protein and total prolamin, while the floury endosperm is richer in γ-prolamins compared to vitreous endosperm [3]. Chandrashekar and Mazhar [3] described the relationship between the prolamin subclasses and grain hardness as follows "the γ-prolamins form the cement while the α-prolamins are the bricks." These authors also postulated that both the content and distribution of α- and γ-kafirins (sorghum prolamins) were responsible for modifying endosperm texture with the α-kafirin responsible for protein body size and the γ-kafirin conferred rigidity by cross-linking the outer edges of the protein bodies [7]. Furthermore, these authors reported that for a kernel to be hard the protein bodies needed to be large (high levels of kafirin) with strong cross-linking (high levels of γ-kafirin).

The above studies and most studies in general on sorghum proteins have all focused on the kafirin subclass composition. This is typically done by extracting the proteins under reducing conditions and analyzing the extracts via SDS-PAGE or RP-HPLC. However, recently the polymeric proteins of sorghum were studied using techniques similar to those used in the study of the large gluten protein polymers found in wheat [13] and traditional SDS-PAGE [14]. The polymeric proteins of maize have also been characterized using similar techniques [15].

Considerable work on the differences in protein composition of vitreous and floury endosperm has been done. However, this research has focused mainly on the kafirin subclasses and not the polymeric proteins, even though evidence points in the direction of protein cross-linking as an important part of grain hardness in sorghum [3]. This project addresses the question of the role of cross-linking of sorghum proteins into larger polymeric groups in the role of grain hardness by using a number of protein analytical techniques to study the protein composition (both in terms of reduced proteins and un-reduced polymeric proteins) of isolated vitreous and floury endosperm.

Materials and Methods

Samples

Eight sorghum samples were selected from a large collection of samples held at the USDA-ARS GMPRC lab and were selected to span a range of hardness values. All the sorghum samples were grown in Kansas except for B94C174, a waxy sorghum, which was grown in Nebraska. Two of the samples were the same hybrid, Wheatland x KS115, grown in two different locations in Kansas.

Sample preparation

Sorghum kernels (10g) were decorticated using the tangential abrasive decortication device (TADD) as described previously [16, 17]. Samples were decorticated for various times (1 to 4 min) and kernels were stained according to the method of Scheuring and Rooney [18]. Stained kernels were visually inspected to determine the optimum decortication time for removal of the bran.

Decorticated samples were degermed and cut in half with a scalpel and the floury endosperm removed with a dremel motor tool (Dremel, Racine, WI) and collected. Hard endosperm was then ground in a coffee grinder for 10s followed by grinding in a mortar in pestle and then sieved through a No. 40 mesh screen. Floury endosperm was recovered as a fine powder during the kernel drilling procedure and thus no grinding was necessary. Percent vitreous determined by weight difference after floury endosperm removal.

Protein extraction

Total kafirins were extracted from ground (25mg) whole grain and endosperm fractions as described in Bean et al [19] and analyzed by reversed phase (RP)-HPLC and free zone capillary electrophoresis (FZCE). For analysis of polymeric proteins, a multi-step procedure was used. Soluble proteins (SP) were first extracted from 10mg of ground endosperm using 0.5 mL of a 12.5 mM sodium borate pH 10 buffer with 2% SDS (w/v) for 30 min with continual vortexing (no reducing agent in buffer). After being centrifuged, the supernatant was removed and the insoluble proteins (IP) were extracted from the residue using sonication (30 sec at 10W in the above pH 10 buffer). After centrifugation, the residue proteins (RP) were extracted from the remaining pellet under reducing conditions using the pH 10 sodium borate/SDS buffer with 2% β-mercaptoethanol (v/v) added. Aliquots of each extract (SP, IP, and RP) were removed and

analyzed by size exclusion high performance liquid chromatography (SEC). The percentages of SP, IP, and RP in a given sample were determined by summing the SEC peak areas across each extract (SP area + IP area + RP area = total area) then dividing each individual extracts area by the total area (e.g. SP area/total area). To characterize the composition of each extract (i.e. SP, IP, and RP), individual peaks in the SEC chromatograms were integrated.

To further characterize the proteins in each of these extracts an additional aliquot was taken from each extract and β-mercaptoethanol was added to a final concentration of 2% (v/v). Samples were then allowed to incubate for 30 min at room temperature and then analyzed by RP-HPLC. To evaluate the amount of protein extracted by the multi-step extraction procedure, three of the sorghum hybrids were selected that spanned the range of SCKS hardness values. The above extraction procedure was carried out on a larger scale on whole endosperm (100mg sample to 1 mL solvent) and the amount of protein remaining after all extracts was measured using nitrogen combustion.

Protein analysis

All HPLC separations were carried out using an Agilent 1100 HPLC system. RP-HPLC conditions were as described in Bean et al [19]. Proteins were separated by SE-HPLC using a Biosep-3000 column (Phenomenx, Torrance, CA) with a 50mM sodium phosphate pH 7.0 buffer containing 1% SDS as mobile phase with a flow rate of 1 ml/min as described in Bean et al [20]. HPCE was carried out as described in Bean et al [19] using a Beckman PACE 2100 instrument.

Nitrogen combustion

Protein content was measured using nitrogen combustion via a Leco FP-528 Nitrogen Determinator (St. Joseph, MI) according to AACC method 46-30 (crude protein-combustion method) [21]. Nitrogen was converted to protein using a factor of 6.25.

Grain hardness measures

The single kernel characterization system (SKCS) was used to obtain grain hardness, kernel weight and diameter [17, 22]. The abrasive hardness index (AHI) was measured using the tangential abrasive dehulling device (TADD) [16].

Sulfhydryl measurements

Free sulfhydryl content (SH) was determined using the methods of Thannhauser et al [23] and Chan and Wasserman [24] were used as described in detail in Lee et al [25].

Statistical Analysis

All data were plotted and statistical analysis (averages, standard deviation, correlations) conducted using Microsoft Excel and Microcal Origin.

Results

Physical Grain Attributes

There was a large range of grain hardness among the samples tested as measured by both the SKCS and the TADD (**Table 2.1**). As expected from samples varying in hardness, the other physical properties of the grains also varied among the samples. Kernel weight, for example varied almost twofold (**Table 2.1**). Correlations among the physical properties were as expected with kernel weight highly correlated to kernel diameter ($r = 0.92$, $p<0.05$, data not shown) and

kernel diameter negatively correlated to the % vitreousity ($r = -0.75$, $p<0.05$, data not shown). SKCS hardness values and AHI were not significantly correlated to each other in this study.

Protein and kafirin content and composition

Total protein and kafirin content for vitreous and floury endosperm is shown in **Table 2.2**. As expected from previous studies, vitreous endosperm had greater levels of total protein than the floury endosperm and had much higher amounts of kafirins (both absolute and on a percent basis). The total γ-kafirin peak area and the percent γ-kafirin (γ-kafirin peak area/total peak area) in the vitreous and floury endosperm as determined by RP-HPLC for the individual samples are shown in **Figure 2.1**. The percent of γ-kafirin in both the vitreous and floury endosperm varied among the samples, but in all cases the floury endosperm contained a higher percentage of γ-kafirins than did the vitreous endosperm.

Kafirins extracted from vitreous and floury endosperm were analyzed by RP-HPLC and FZCE to determine the overall kafirin composition in each type of endosperm (selected example shown in **Figure 2.2**). For all samples, kafirin separations were similar to those shown in **Figure 2.2** and visual inspection of the RP-HPLC and FZCE kafirin patterns showed that qualitatively the kafirins were essentially identical between the two endosperm fractions. Quantitatively, however, the kafirins extracted from floury endosperm showed large differences when compared to the kafirins from the vitreous endosperm. The floury endosperm obviously had lower levels of kafirins. The floury endosperm also appeared to have lower levels of the peak in the α region of the FZCE separations just before 18 min (**Figure 2.2**). A similar difference was seen in the α/β range in the RP-HPLC separations at ~43 min in the RP-HPLC separations (**Figure 2.2**). While the absolute γ-kafirin peak content was greater in the vitreous endosperm fractions, the proportion of γ-kafirins to the α/β kafirins in the floury endosperm was greater than in the

vitreous endosperm with γ kafirins in the floury endosperm making up almost twice the percentage of kafirins as in the vitreous endosperm (**Table 2.2**), with only one sample (Mycogen X00ML337) deviating substantially from this trend (**Figure 2.1**).

Sulfhydryl content and composition

Free sulfhydryl (F-SH), total sulfhydryl (T-SH), and disulfide (S-S) contents were measured on vitreous and floury endosperm samples in addition to protein characterization. When comparing averages between the vitreous and floury endosperm across all samples, vitreous endosperm had lower levels of both F-SH and T-SH (on a per protein basis) than did floury endosperm (**Table 2.3**). Disulfide content was also lower in vitreous endosperm than in floury, though the difference was not as great as for the F-SH and T-SH measurements (**Table 2.3**). The ratio of S-S to T-SH was greater in the vitreous endosperm than in the floury endosperm, indicating more disulfide cross-links (i.e. a greater percent of the T-SH were in the form of S-S) in the vitreous endosperm relative to the floury endosperm (**Table 2.3**). When looking at the data from the individual samples, the same trends as for the averages for the endosperm types were seen (**Figure 2.3**). In each sample, free SH and total SH were much higher in the floury endosperm compared to the vitreous endosperm. For the disulfide bonds, differences between vitreous and floury endosperm were readily apparent except for two samples (A8PR1059xLG35 and Mycogen X00ML337) which did not share similar overall grain traits (**Table 2.1**). Likewise the ratio of disulfide to total SH bonds was greater for the vitreous endosperm than floury in all samples, the same as when the averages of vitreous and floury were compared.

Polymeric proteins

To characterize the polymeric proteins in vitreous and floury endosperm, a sequential extraction scheme was used which divided proteins into SDS 'soluble' proteins, insoluble proteins (extracted using ultrasound), and residue proteins (extracted with reducing agent). Data for these solubility classes for each of the sorghum samples is shown in **Figure 2.4**. Floury endosperm either contained the same or higher amounts of SP than did vitreous endosperm among the samples tested (**Figure 2.4A**). Vitreous endosperm contained higher amounts of IP than did floury endosperm and in the majority of samples (**Figure 2.4B**), RP was similar between vitreous and floury endosperm among the samples (**Figure 2.4C**). Comparisons of the averages for these solubility classes across all samples for both vitreous and floury endosperm are shown in **Table 2.4**. On average, the vitreous endosperm had lower levels of the soluble proteins than did floury endosperm, though as mentioned above, this difference was greater in some samples than others. Conversely, the vitreous endosperm showed higher levels of insoluble proteins than did the floury endosperm. Levels of residue protein were not significantly different between the two types of endosperm (**Table 2.4**).

To test the amount of total protein extracted with the above extraction scheme, three samples varying in SKCS hardness values were extracted and the protein in the residue determined by nitrogen combustion. This test revealed that 95-97% of total protein was extracted (data not shown).

To characterize the polymeric protein composition of the SP, IP, and RP extracts, samples were analyzed by SEC and the chromatograms divided into five peak regions (**Figure 2.5**). Based on the elution times of standard M_w marker proteins and by analyzing SEC patterns following reduction, peaks 1 and 2 were judged to contain mostly polymeric proteins while peaks 3-5 contain monomeric proteins, which was in agreement with the results of Bean et al

[20] and El Nour et al [13]. The results of the SEC analysis are shown in **Figure 2.5**. In the vitreous endosperm, the SP extract was mainly peaks 3, 4, and 5 which totaled over 70% of the total SEC area. In the IP extract, however, the percentage of peak 1 was greatly increased and was roughly equal to that of peaks 3 and 4. In the floury SP extract, peaks 3 and 4 dominated, totaling almost 75% of the total SEC peak area. In the floury IP extract, the proportion of peak 1 greatly increased to make up almost 40% of the total SEC area. In both the vitreous and floury RP extracts, peak 4 was the major peak (**Table 2.4**).

To further characterize the proteins present in the SP, IP, and RP extracts from the vitreous and floury endosperm, aliquots of each extract from two sorghum samples were lyophilized and then re-suspended pH 10 SDS buffer with 2% β-ME added and the reduced proteins analyzed by RP-HPLC. The SP, IP, and RP from vitreous endosperm showed little difference in their chromatograms and resembled the chromatograms of a total kafirin extract (data not shown). However, in the floury endosperm fractions, the SP contain large amounts of peaks eluting in the γ kafirin region and the IP chromatograms showed reduced levels of peaks eluting in the γ kafirin region (data not shown).

Discussion

Previous research on prolamins and grain hardness in sorghum and maize has produced a number of ideas on how proteins influence kernel hardness. Most postulate that cross-linking of the γ and possibly β kafirins on the outer edges of the protein bodies plays a role. This crosslinking could be either to matrix protein, which is tightly compressed against protein bodies in the vitreous endosperm or to other prolamins [3, 15, 27, 28].

While several hypotheses about protein cross-linking and grain hardness have been put forward, to date no studies have attempted to look directly at the polymeric proteins in isolated vitreous and floury endosperm. If protein cross-linking between kafirins and matrix protein, or simply within the matrix protein itself, varies between vitreous and floury endosperm, the analysis of unreduced protein extracts from the endosperm fractions should show such differences. To directly gain information on the polymeric protein content and composition of vitreous and floury endosperm, a multi-step extraction procedure was used to divide sorghum proteins into SDS soluble, SDS insoluble, and residue proteins. This is often done in wheat as a method to determine the molecular weight distribution of the polymeric glutenin proteins [29] as the insoluble proteins are hypothesized to have a larger M_w distribution than the soluble proteins (and hence their insolubility); which was recently verified using laser light scattering [30]. Thus, we applied the idea of differential solubility, frequently used to study the molecular weight distribution in wheat proteins, to study the cross-linked proteins of sorghum.

On the basis of the above, we hypothesized that if the protein cross-linking was different between vitreous and floury endosperm, the distribution of SP and IP would also differ between vitreous and floury endosperm. The results of the multi-step extraction procedure showed that the floury endosperm had significantly more SP than did the vitreous endosperm (**Table 2.4**). Conversely, the vitreous endosperm had a higher percentage of the IP fraction than the floury endosperm. Following the wheat protein model, this would suggest that the vitreous endosperm had more of the larger (and therefore insoluble) polymeric proteins than did the floury endosperm. However, other factors such as increased hydrophobicity could also account for lack of solubility of the IP in sorghum. Complete characterization of the SP and IP using multi-angle laser light scattering and mass spectrometry is currently underway.

Early classification schemes for sorghum proteins often utilized an aqueous alcohol solvent to extract a fraction called “kafirin 1”. A second aqueous alcohol extract with a reducing agent was then used to extract a fraction called “kafirin 2”. A similar scheme in the Landry and Moureaux produced extracted fractions labeled “II” and “III”, respectively [31]. The kafirin 1 and fraction II in these earlier papers would be roughly analogous to the SP extract in the current paper. Likewise, kafirins 2 and fraction III would roughly equal to IP in the current study. Similar relationships of kafirin 1 and kafirin 2 to soluble and insoluble extracts were made by El Nour et al [13]. Using these earlier methods, vitreous endosperm was found to have higher levels of kafirin 2 than did floury endosperm during kernel development [6].

To further investigate the M_w distribution of the proteins in the vitreous and floury endosperm, the SP and IP extracts were analyzed by SEC (**Figure 2.5**). Note that in this sense the M_w distribution is referred to not in the ranges of the M_w of the proteins present, but rather in the distribution of the amounts of proteins within the M_w ranges present. This is widely done in wheat proteins where sonication is needed to extract the largest insoluble proteins and which alters the original M_w ranges of the proteins, the largest of which elute in the void volume of the SEC columns in any case, making it difficult to judge the true range of M_w’s present [32]. The use of non-reduced extractions followed by analysis of the proteins by SEC allowed for a direct look at the molecular weight distribution of these extracts, something not done previously (in all previous studies proteins have been reduced prior to analysis). Significant differences in the peak compositions were found between the vitreous and floury endosperm (**Figure 2.6**). In the SP extract, there were no major differences between the vitreous and floury endosperm except that floury endosperm had much higher levels of peak 4. This would seem to make sense if the solubility is related to the M_w distribution, i.e. the extraction process would be expected to

extract similar material in both endosperm fractions. In the vitreous endosperm, IP had more of peak 1 material than peak 2, suggesting a shifted M_w distribution towards larger polymeric proteins. Overall levels of the polymeric peaks (1 and 2) were higher than in the SP, again showing a shift towards more polymeric proteins in IP than SP. The same trend was observed in the floury endosperm, though the lower M_w peaks 3 and 4 were not as prevalent relative to the polymeric proteins (1 and 2).

It was interesting that the floury IP had a higher percentage of its protein in peak 1 than did the vitreous IP, which again indicates differences in cross-linking between the endosperm types. Taken together this seems to indicate that there is a shift towards more polymeric proteins in the IP than SP and that the IP in the floury endosperm has more polymeric proteins relative to monomeric proteins. This could reflect a better extraction of protein in the floury endosperm, however, overall amounts of protein extracted by this procedure were similar between vitreous and floury endosperm and the RP protein content was similar between the vitreous and floury endosperm. It was also interesting to note that the kafirin composition of the SP, IP, and RP from the vitreous endosperm did not vary much while the SP from the floury endosperm was highly enriched with γ kafirins. As noted above RP levels were similar between both vitreous and floury endosperm. These proteins were not extracted with either sonication or in the presence of reducing agents, suggesting that their solubility was not limited due to the M_w, as both these techniques would have reduced their M_w. It is most likely that these are non-prolamin proteins and similar results have been found in wheat [33].

In addition to the above protein studies, the sulfhydryl content and composition of the vitreous and floury endosperm was evaluated. Floury endosperm had higher levels of both F-SH and T-SH on a per protein basis. This could be due to the higher proportion of γ-kafirins in the

floury endosperm, which are known to be high in Cys [34]. However, in the vitreous endosperm the ratio of S-S to T-SH was greater, thus a greater proportion of the T-SH were involved in disulfide bonds, indicating greater cross-linking of proteins in the vitreous endosperm compared to the floury endosperm.

The results of this paper indicate that vitreous endosperm had a greater level of protein cross-linking than did floury endosperm and that this cross-linking produced a larger M_w distribution than found in the floury endosperm. Vitreous endosperm has a much more compact structure with a continuous matrix than floury endosperm and therefore more opportunity for cross-linking between protein components [5]. This was the opposite for the floury endosperm, where protein bodies were not tightly packed and the protein matrix was discontinuous, and thus less opportunity for protein cross-linking [5].

Several relationships were found between the various protein classes measured in this study that may help explain the cross-linking of sorghum proteins. For instance, the amount of γ kafirin in vitreous endosperm was negatively correlated to the amount of SP in vitreous endosperm ($r=-0.85$, $p<0.5$, $n=6$, data not shown). Interestingly, the amount of γ kafirin in vitreous endosperm was positively correlated to the amount of RP in vitreous endosperm ($r=0.92$, $p<0.05$, $n=6$, data not shown). These correlations are based on a limited number of samples and should be regarded as preliminary; however, they do seem to indicate the amount of γ kafirin in the vitreous endosperm is related to the distribution of the solubility classes, and thus the Mw distribution. The RP may be the most difficult to extract due to the most cross-linking of the proteins through the γ kafirins. It remains to be seen if RP proteins have an even larger M_w than the IP, or if they are the most difficult to extract due to some other properties of being highly cross-linked. No significant relationships between kafirin content and composition to SP,

IP, or RP in the floury endosperm were found. Likewise, neither the total kafirins in vitreous endosperm or floury endosperm was correlated to their respective SP, IP, or RP levels. Thus, the γ-kafirins in the vitreous endosperm seem to have the most obvious relationships to indicators of protein cross-linking in the vitreous endosperm. The fact that in the floury endosperm, the SP extract was enriched with γ-kafirins may suggest that whatever cross-links the γ-kafirins in the vitreous endosperm does not cross-link them in the floury endosperm, or perhaps in the floury endosperm they are cross-linked to themselves. El Nour et al [13] also reported that γ-kafirins were important in the cross-linking of kafirins along with β-kafirins when analyzing samples extracted from whole grain, however further study on the composition of sorghum polymeric proteins in sorghum is needed to confirm the roles and mechanisms of γ- and β-kafirins in protein cross-linking. It is also important to note that in studying the polymeric protein composition of sorghum, vitreous and floury endosperm need to be studied separately.

References

1. Bettge, A. D.; Morris, C. F. Relationships among grain hardness, pentosan fractions, and end-use quality of wheat. *Cereal Chem.* **2000**, *77*, 241-247.

2. Cagampang, G. B.; Kirleis, A. W. Relationship of sorghum grain hardness to selected physical and chemical measurements of grain quality. *Cereal Chem*. **1984**, 61, 100-105.

3. Chandrashekar, A.; Mazhar, H. The biochemical basis and implications of grain strength in sorghum and maize. *J. Cereal Sci.* **1999**, *30*,193-207.

4. Wall, J. S.; Bietz, J. A. Differences in corn endosperm proteins in developing seeds of normal and opaque-2 corn. *Cereal Chem*. **1987**, *64*, 275-280.

5. Shull, J.M.; Chandrashekar, A.; Kirleis, W.; Ejeta, G. Development of sorghum endosperm in varieties of varying hardness. *Food Structure* **1990**, *9*, 253-267.

6. Mazhar, H.; Chandrashekar, A. Differences in kafirin composition during endosperm development and germination in sorghum cultivars of varying hardness. *J. Cereal Sci.* **1993**, *70*, 667-671.

7. Mazhar, H.; Chandrashekar, A. Quantification and distribution of kafirins in the kernels of sorghum cultivars varying in endosperm hardness. *J. Cereal Sci*. **1995**, *21*, 155-162.

8. Dombrink-Kurtzman, M.A.; Bietz, J. A. Zein composition in hard and soft endosperm of maize. *Cereal Chem.* **1997**, *70*, 105-108.

9. Pratt, R.C.; Paulis, J. W.; Miller, K.; Nelsen, T.; Bietz, J. A. Association of zein classes with maize kernel hardness. *Cereal Chem*. **1995**, *72*, 162-167.

10. Wallace, J.C.; Lopes, M.A.; Paiva, E.; Larkins, B.A. New methods for extraction and quantification of zeins reveal a high content of □-zein in modified opaque-2 maize. *Plant Physiol*. **1990**, *92*, 191-196.

11. Hamaker, B. R.; Mohamed, A. A.; Habben, J. E.; Huang, C. P.; Larkins, B. A. Efficient procedure for extracting maize and sorghum kernel proteins reveals higher prolamin contents than the conventional method. *Cereal Chem*. **1995**, *72*, 583-588.

12. Shull, J. M.; Watterson, J. J.; Kirleis, A. W. Proposed nomenclature for the alcohol proteins (kafirins) of Sorghum bicolor (L. Moench) based on molecular weight, solubility, and structure. *J. Agric. Food Chem*. **1991**, *39*, 83-87.

13. El Nour, I. N. A.; Peruffo, A. D. B.; Curioni, A. Characterisation of sorghum kafirins in relation to their cross-linking behaviour. *J. Cereal Sci.* **1998**, *28*, 197-207.

14. Nunes, A.; Correia, I.; Barros, A.; Delgadillo, I. Characterization of kafirin and zein oligomers by preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *J. Agric. Food Chem.* **2005**, *53*, 639-643.

15. Wall, J. S.; Cooker, L. A.; Bietz, J. A. Structure and origin of maize endosperm alcohol-insoluble glutelin. *J. Agric. Chem*. **1988**, *36*, 722-728.

16. Oomah, B. D.; Reichert, R. D.; Youngs, C. G. A novel, multi-sample, tangential abrasive dehulling device (TADD). *Cereal Chem*. **1981**, *58*, 392-395.

17. Bean, S. R.; Chung, O. K.; Tuinstra, J. F.; Erpelding, J. Evaluation of the Single Kernel Characterization System (SKCS) for Measurement of Sorghum Grain Attributes. *Cereal Chem*. **2006**, *83*, 108-113.

18. Scheuring, J. F.; Rooney, L. W. A staining procedure to determine the extent of bran removal in pearled sorghum. *Cereal Chem*. **1979**, *56*, 545-547.

19. Bean, S. R.; Lookhart, G. L.; Bietz, J. A. Acetonitrile as a buffer additive for the separation of maize (*Zea mays* L.) and sorghum (*Sorghum bicolor* L.) storage proteins by HPCE. *J. Agric. Food Chem.* **2000**, *48*, 318-327.

20. Bean, S. R.; Ioerger, B. P.; Park, S. H.; Singh, H. Interaction Between Sorghum Protein Extraction and Precipitation Conditions on Yield, Purity, and Composition of Purified Protein Fractions. *Cereal Chem*. **2006**, *83*, 99-107.

21. AACC International. AACC Methods 46-30 and 76-31. *Approved Methods of the American Association of Cereal Chemists*, 10th ed.; AACC International: St. Paul, MN, 2000.

22. Pedersen, J. F.; Martin, C. R.; Felker, F. C.; Steele, J. L. Application of the single kernel wheat characterization technology to sorghum grain. *Cereal Chem*. **1996**, *73*, 421-423.

23. Thannhauser, T.W.; Konishi, Y.; Scheraga, H.A. Analysis for disulfide bonds in peptides and proteins. *Methods Enzymol*. **1987**, *143*, 115-119.

24. Chan, K.Y.; Wasserman, B.P. Direct colorimetric assay of free thiol groups and disulfide bonds in suspensions of solubilized and particulate cereal proteins. *Cereal Chem*. **1993**, *70*, 22-26.

25. Lee, K.-M.; Bean, S. R.; Alavi, S.; Herrman, T. J.; Waniska, R. D. 2006. Physical and biochemical properties of maize hardness and extrudates of selected hybrids. *J. Agric. Food Chem*. **2006**, *54*, 4260-4269.

26. Watterson, J. J.; Shull, J. M.; Kirleis, A. W. Quantitation of □-, □-, and □-kafirins in vitreous and opaque endosperm of Sorghum bicolor. **1993**. *Cereal Chem*., *70*, 452-457.

27. Paiva, E.; Kris, A. L.; Peixoto, M. J. V. V. D.; Wallace, J. C.; Larkins, B. A. Quantitation and distribution of □-zein in the endosperm of maize kernels. *Cereal Chem*. **1991**, *68*, 176-279.

28. Robutti, J. L.; Borras, F. S.; Eyherabide, G. H. Zein compositions of mechanically separated coarse and fine portions of maize kernels.. *Cereal Chem*. **1997,** *74*, 75-78.

29. Southan, M.; MacRitchie. F. Molecular weight distribution of wheat proteins. *Cereal Chem*. **1999**, *76*, 827-836.

30. Bean, S.R.; Lookhart, G. L. Factors influencing the characterization of gluten proteins by SEC-MALLS. *Cereal Chem*. **2001**, *78*, 608-618.

31. Landry, J.; Moureaux, T. Heterogeniety of the glutelins of the grain of corn: Selective extraction and composition in amino acids of the insoluble fractions. *Bull. Soc. Chem. Biol*. **1970**, *52*, 1021-1037.

32. Southan, M.; MacRitchie, F. Molecular weight distribution of wheat proteins. *Cereal Chem*. **1999**, *76*, 827-836.

33. Sapirstein, H. D.; Fu, B. X. Intercultivar variation in the quantity of monomeric proteins, soluble and insoluble glutenin, and residue protein in wheat flour and relationships to breadmaking quality. *Cereal Chem*. **1998**, *75*, 500-507.

34. Belton, P. S.; Delgadillo, I.; Halford, N.G.; Shewry, P. R. Kafirin structure and functionality. *J. Cereal Sci.* **2006**, *44*, 272-286.

Table 2.1 Grain traits and indices for sorghum samples used.

Sample	Key #	AHI[1]	SKCS HI[2,5]	Kernel Weight (mg)[3,5]	Kernel Dia. (mm)[3,5]	% Vitreosity by weight	Crude Protein
ATx623 x RTx430	1	15.9	99.5±3.0	24.4±1.2	2.2±0.1	84.0	9.3
(ATx3042 x Tx435)-F1	2	13.6	80.0±2.4	31.4±1.6	2.4±0.1	81.1	8.1
Pioneer 82G63	3	12.5	71.2±2.2	29.8±1.5	2.3±0.1	79.0	8.3
A8PR1059xLG35	4	15.7	85.2±2.6	26.1±1.3	2.4±0.1	82.9	9.7
[4]Wheatland x KS115	5	15.2	55.4±1.7	47.9±2.4	2.8±0.1	75.6	11.5
[4]Wheatland x KS115	6	14.1	55.3±1.7	41.3±2.1	2.6±0.1	74.2	8.8
Mycogen X00ML337	7	17.3	92.0±2.8	27.0±1.4	2.2±0.1	80.6	11.2
B94C274	8	15.0	75.9±2.3	30.7±1.5	2.5±0.1	80.6	10.9

[1] Abrasive hardness index

[2] Single kernel characterization system (SKCS) hardness index

[3] Determined using the SKCS

[4] Sample nos. 5 and 6 were grown in different locations

[5] Standard error for SKCS measurements calculated from RSD of control sample measured ten times due to limited amounts of experimental samples

Table 2.2 Kafirin composition of endosperm fractions averaged across endosperm fractions obtained from all 8 sample varieties.

	Endosperm fraction	
Measurement	Vitreous	Floury
Total protein[1]	9.8a	7.5b
γ kafirin area[2]	1829a	999b
α/β kafirin area[3]	31612a	8144b
γ kafirin %[4]	5.4a	11.2b
α/β kafirin %[5]	94.1a	88.8b

Values followed by the same letter in the same row are not significantly different ($P<0.05$)

[1] Total protein % of endosperm fractions as determined by nitrogen combustion

[2] RP-HPLC peak area of γ kafirins

[3] RP-HPLC peak area of α and β kafirins

[4] % RP-HPLC area of γ kafirins (γ peak area/ γ peak area + α and β peak area)

[5] % RP-HPLC area of α and β kafirins (α and β peak area/γ peak area + α and β peak area)

Table 2.3 Sulfhydryl content and composition of vitreous and floury endosperm averaged across all 8 sample varieties.

	Endosperm fraction	
Measurement	Vitreous	Floury
Free SH (F-SH)[1]	12.7a	68.6b
Total SH (T-SH)[2]	59.2a	139.2b
Disulfide content (S-S)[3]	23.3a	35.3b
Ratio S-S/T-SH[4]	0.39a	0.25b

Values followed by the same letter in the same row are not significantly different ($P<0.05$)

[1] Free sulfhydryl content, nmol/mg of protein

[2] Total sulfhydryl content, nmol/mg of protein

[3] Disulfide content, nmol/mg of protein

[4] Ratio of disulfide content to total sulfhydryl content

Table 2.4 Amount of soluble, insoluble, and residue protein in vitreous and floury endosperm averaged across all samples.

	SEC Peak area%	
Extraction	Vitreous	Floury
SP[1]	36.7a	47.2b
IP[2]	45.3a	35.9b
RP[3]	18.0a	16.9a

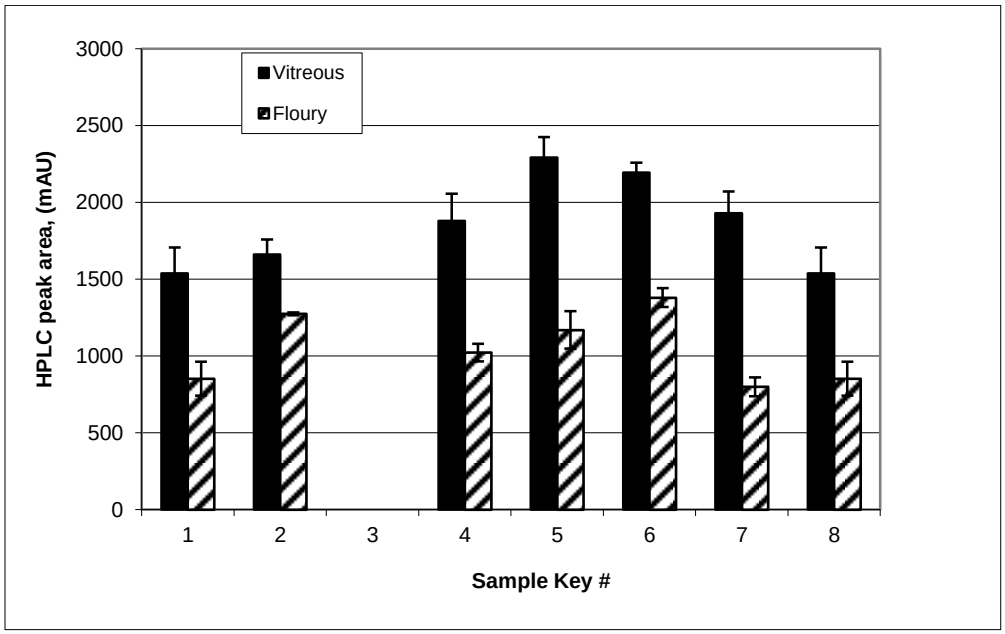
Values followed by the same letter in the same row are not significantly different ($P<0.05$)

[1] Soluble protein

[2] Insoluble protein

[3] Residue protein

Figure 2.1 Gamma kafirin HPLC peak area and % gamma kafirin peak area (of total HPLC area) for individual sorghum samples. Note no data is shown for sample 3 due to lack of available sample for these analyses. Sample numbers correspond to the order the samples are listed in Table 2.1. Error bars represented standard deviation for duplicate extractions (n=2).



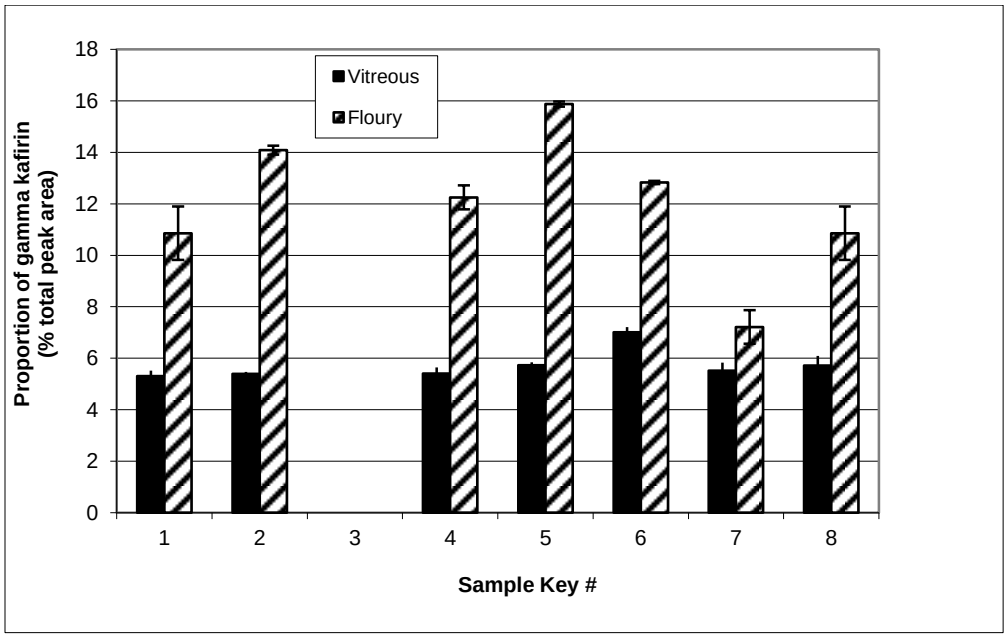
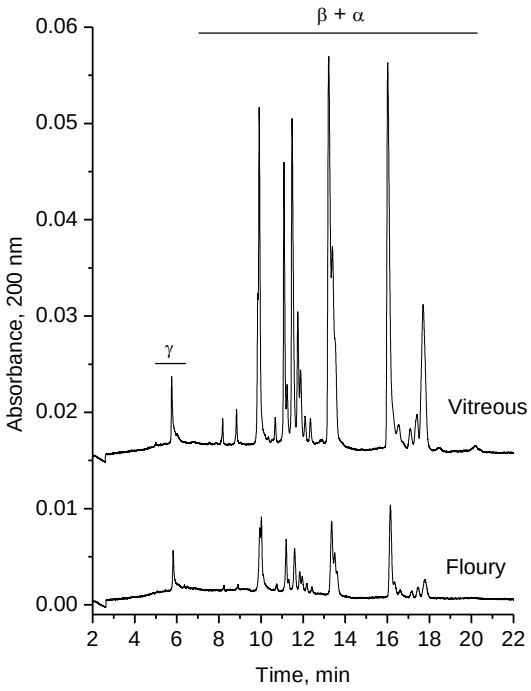


Figure 2.2 A) FZCE and B) RP-HPLC separations of total kafirins extracted from vitreous, and floury endosperm from the sorghum hybrid Mycogen X00ML337.

A



B

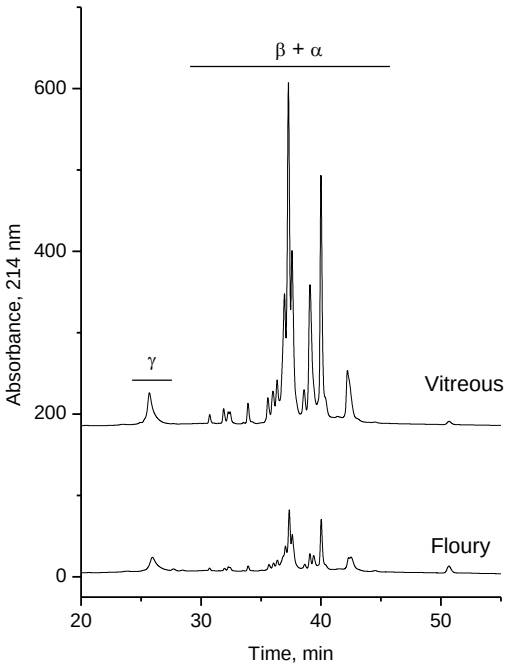


Figure 2.3 A) Free SH bonds for vitreous and floury endosperm for individual samples, nmol/mg protein, B) Total SH bonds Free SH bonds for vitreous and floury endosperm for individual samples, nmol/mg protein, C) Disulfide (S-S) bonds for vitreous and floury endosperm for individual samples, nmol/mg protein, and D) ratio of Disulfide to Total SH for vitreous and floury endosperm for individual samples. Sample numbers correspond to the order in which the samples are listed in table 1. Data not shown for sample 6 due to limited amount of sample available. Error bars represented standard deviation (n=2).

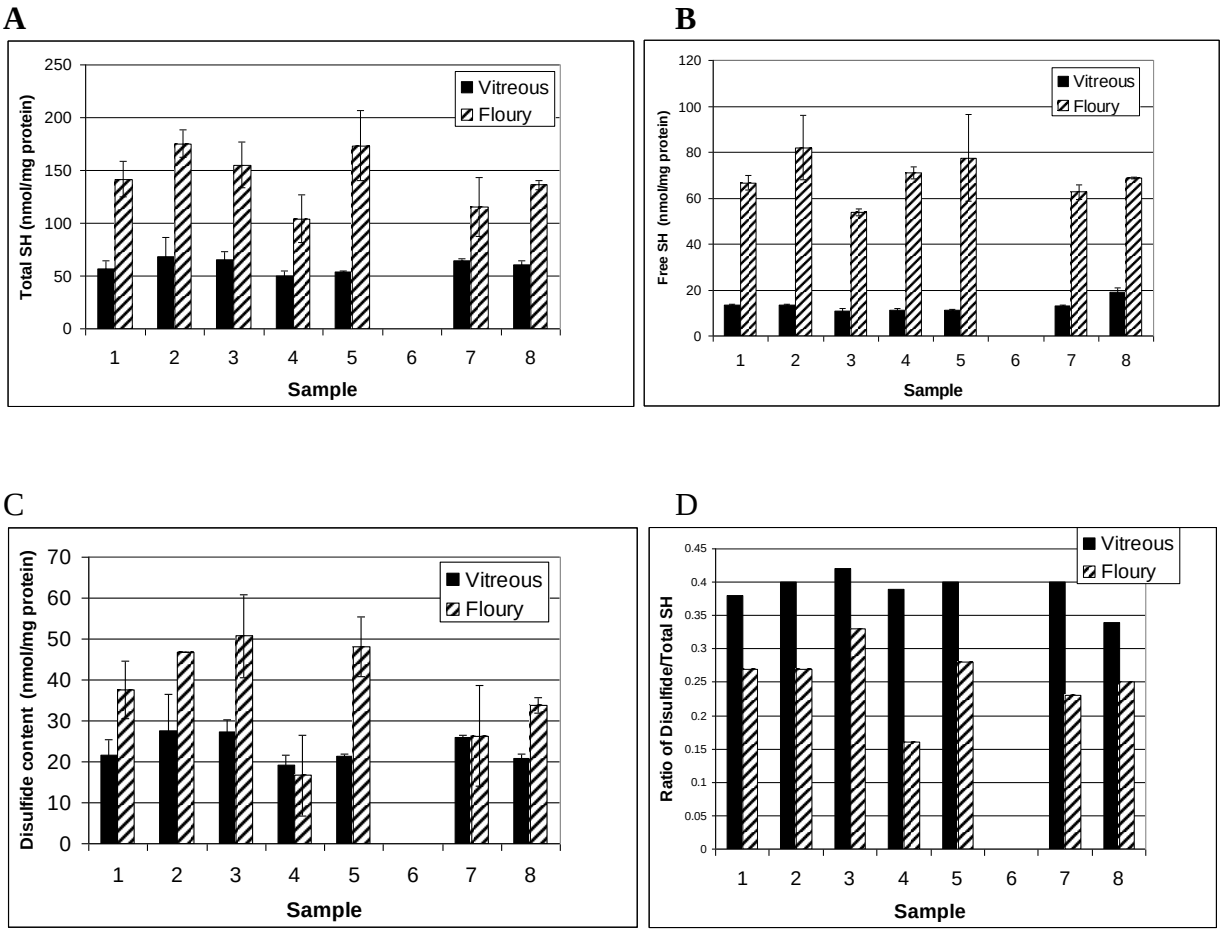
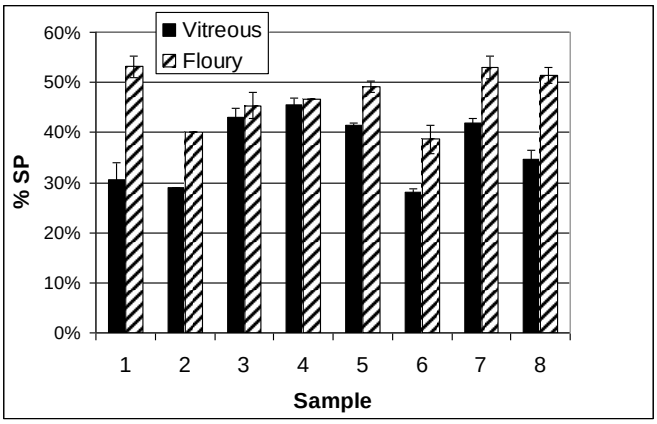
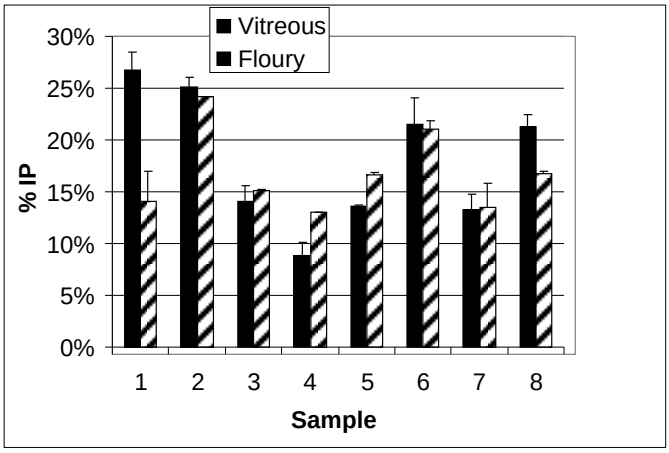


Figure 2.4 Percentage of A) SDS soluble proteins (SP), B) SDS insoluble proteins (IP), and C) residue proteins (RP) in individual sorghum samples. Sample numbers correspond to the order in which the samples are listed in table 1. Error bars represented standard deviation (n=2).

A



B



C

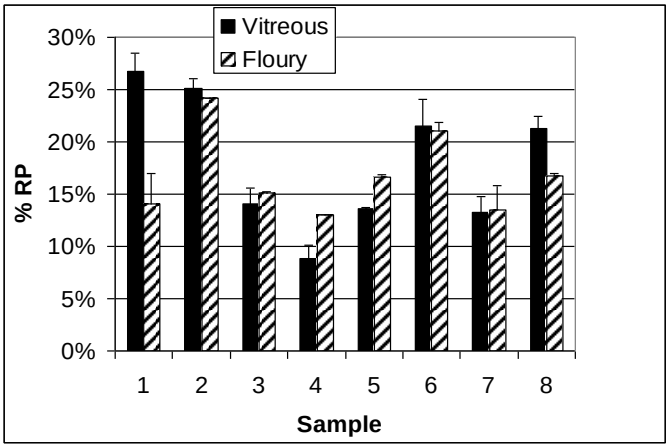


Figure 2.5 SEC separation of A) SDS soluble proteins (SP), and B) SDS insoluble proteins (IP) from vitreous and flour endosperm from the sorghum hybrid Mycogen X00ML337. Arrows indicated the approximate elution position of M_w markers and numbered bars indicated the location of integrated peaks.

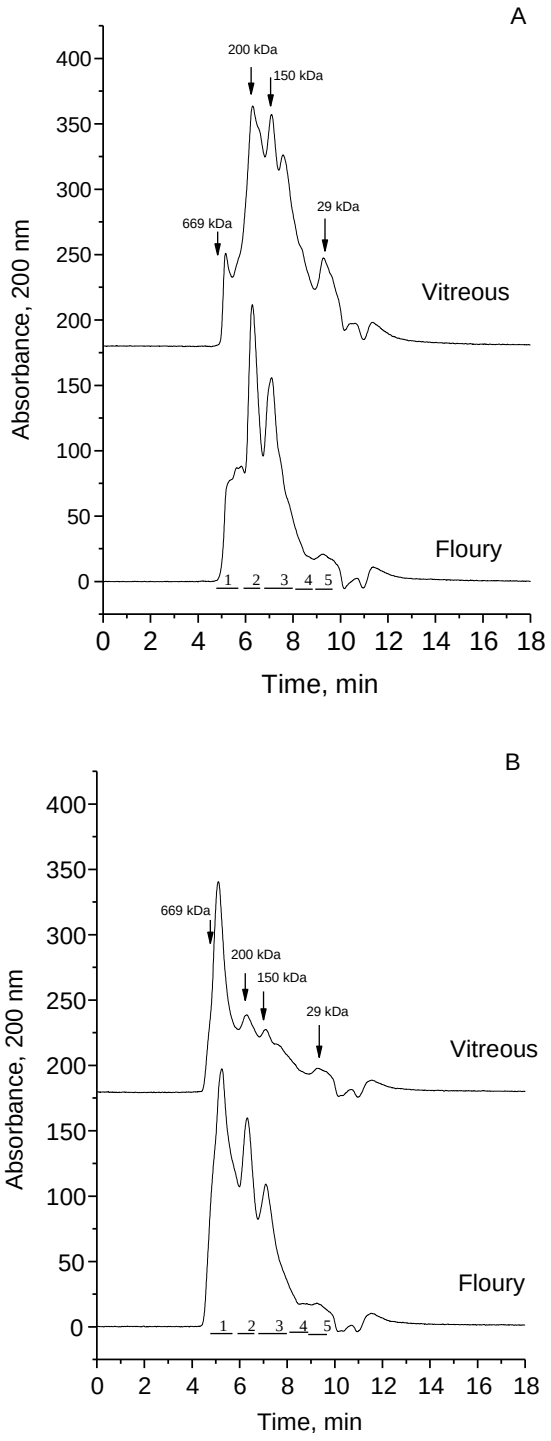
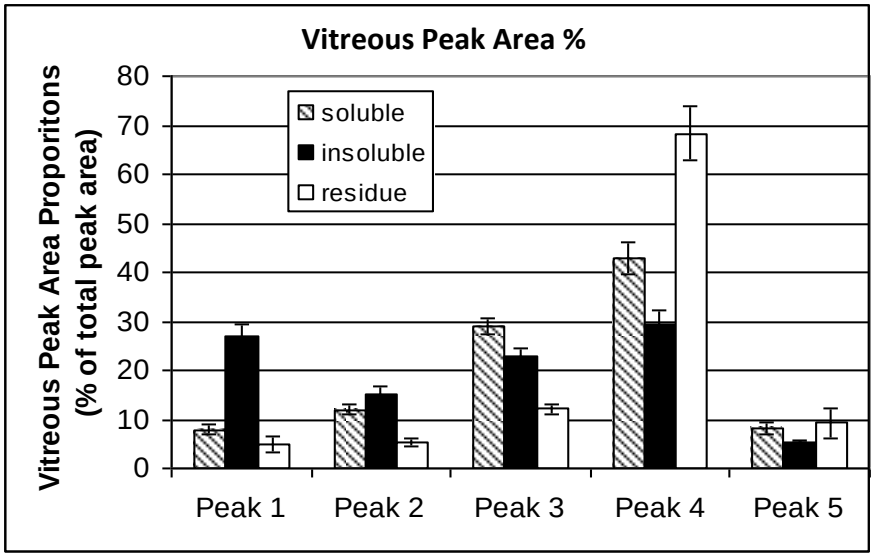
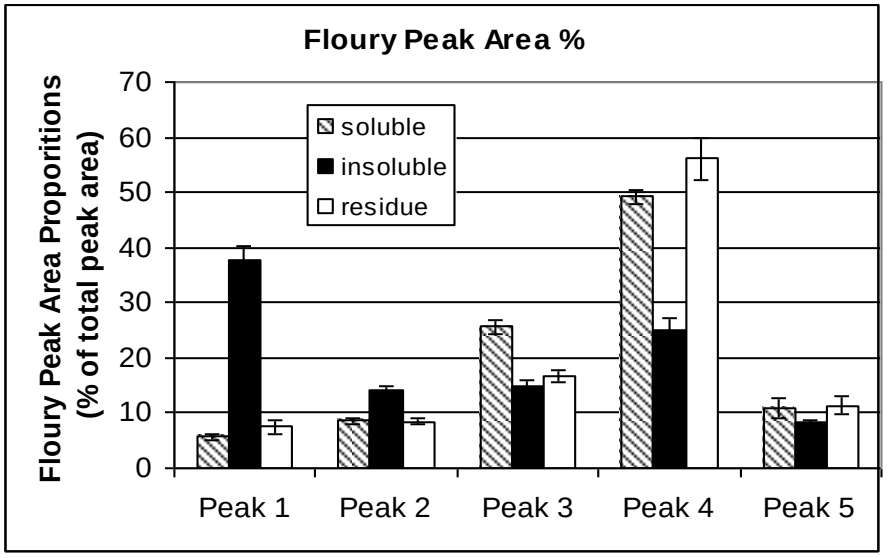


Figure 2.6 SEC peak area % (of total peak area) for A) vitreous endosperm, and B) floury endosperm (averages for 8 samples). Error bars represented standard deviation (n=2).

A



B



Chapter 3 - Improved Method for Extracting Sorghum Protein Polymers for Characterization by Size Exclusion Chromatography

Abstract

A method for fractionating sorghum proteins using extraction solvents and techniques designed to obtain intact disulfide linked protein polymers was developed. Extraction and separation conditions were optimized in terms of completeness of protein extraction, sample stability, and analytical resolution. After pre-extraction of albumins and globulins, a 3-step sequential procedure involving no reducing agents was applied to ground whole sorghum flour. The three fractions obtained represented proportionally different protein polymer contents as evidenced by comparative size exclusion chromatography. Extract stability was maintained for all fractions by application of 80 ˚C heat for 2min. Both silica and polymer based columns were evaluated for separating polymeric sorghum proteins. The best resolution within all fractions was achieved with a silica based column with a nominal M_w range of 5kDa – 700kDa. Complete separation of all extracted proteins typically required 30min or less. Application of this extraction method to sorghum varieties varying in protein digestibility provides a useful tool for polymeric protein content comparison, and should help provide additional insight into how sorghum protein structure relates to digestibility.

Introduction

Although multiple factors may influence sorghum protein digestibility, within tannin-free wild-type sorghums, protein crosslinking is generally considered to have the greatest influence (Duodu et al. 2003; Emmambux and Taylor, 2009; Oria et al. 1995; El Nour et al. 1998; Hamaker et al. 1987; Ezeogu et al. 2008). Thus naturally occurring protein polymers, as well as those formed as a result of processing and feed/food treatment procedures, directly affect the nutrient availability and value of sorghum in feed and foods. A method for characterizing the nature and extent of crosslinking of sorghum varieties intended for feed and food use would be of benefit.

Prospects for gaining greater functionality from processed sorghum products would be improved from a better understanding of sorghum protein structure. The influence of protein polymers in wheat functionality has long been known and correlated to many wheat product properties such as mixing time, extensibility, and loaf volume in bread, as well as quality in durum pasta (Orth and Bushuk 1972; Chakraborty and Khan 1988; Dachkevitch and Autran, 1989; Singh et al. 1990; Gupta et al. 1993; Bean et al. 1998; Sapirstein and Fu, 1998). Because knowledge is lacking about the protein polymers of sorghum and how they relate to protein body structure, details on their effect on potential functionality attributes is unknown.

Size exclusion chromatography (SEC) in conjunction with sequential extractions involving sonication procedures was used by Ioerger et al. (2007) to study differences in protein crosslinking within hard and soft sorghum endosperm. The use of SEC allowed determination of the relative molecular weight distributions of polymeric proteins within two of three differentially extracted fractions. In another study, SEC was used in conjunction with sonicated extraction to predict fermentation quality of sorghum varieties used for ethanol fermentation

(Zhao et al. 2008). High correlation was found between amount of polymeric proteins extracted and ethanol produced by fermentation. Other studies have also used SEC in attempts to characterize extracted kafirins (El Nour et al. 1998; Emmambux and Taylor 2009), however no systematic optimization for increased extraction of the non-reduced polymeric proteins or of the SEC analytical conditions used for analysis was attempted. On the other hand, Bean and Lookhart (2001) did systematically investigate factors influencing the extraction and characterization of wheat gluten proteins by SEC in conjunction with multiangle light scattering. Optimal conditions involved use of SDS solvents for sequential extractions and for SEC analysis.

The objective of the current study was to develop an analytical method for determining the representative polymeric protein content of wild-type sorghum varieties. Optimizing the extraction and SEC separation conditions for characterization of non-reduced polymeric sorghum proteins would provide a valuable method by which digestibility characteristics as well as potential for functionality attributes could be explored.

Materials and Methods

Sample Preparation

Tannin free hybrid sorghum samples were selected for all analyses from samples held at the USDA-ARS CGAHR laboratory. Samples used in this study included NC+371 and F1000. Whole kernel sorghum flour was obtained by grinding samples through a UDY mill (UDY Corp., Ft.Collins, CO) equipped with a 0.5mm screen. Samples were stored desiccated at -20 ˚C, and allowed to equilibrate to room temperature prior to analysis.

Preliminary Extraction Experiments

Various solvents and solvent additive combinations were used for extraction of flour samples (**Table 3.1)** during preliminary experiments. All extraction values, unless otherwise noted, were averages of duplicate analyses. Initial extraction screening involved varying NaOH concentrations, solvent:sample ratios, and high pH buffers (i.e. >pH 10), as listed in **Table 3.1**, and used an extraction procedure consisting of 60min vortex mixing (Vortex Genie2, Fisher, Pittsburgh, PA) in 2mL microtubes. This was followed by centrifugation (4min at 9300g) and pellet lyophilization. Protein removal was quantitated using nitrogen combustion of dried pellets with a LECO FP-528 nitrogen determinator (LECO, Inc., St. Joseph, MI) according to AACC method *46-30.01* to allow determination of total protein extracted (AACC International 2000, crude protein-combustion method).

Further extraction enhancement was attempted using additives commonly used to aid in protein solubilization. These consisted of various salts, detergents, and chaotropes, as well as several organic buffer modifiers (**Table 3.1**). Extraction during this phase of the study consisted of two 5min vortex extractions followed by pooling the resulting supernatants 1:1 post-centrifuge. Use of Tris-borate precluded use of nitrogen combustion for evaluating protein extraction since the extraction buffer itself contains nitrogen. Subsequent comparative quantitative evaluation of extraction efficiency during this phase of method development was accomplished using total SEC peak area from chromatography performed using a Biosep-S3000 column (Phenomenex, Torrance, CA) with a 50mM Na-phos pH 7.0 / 1% SDS mobile phase, and flow rate of 1mL/min.

For extractions using sonication, a Sonic Dismembrator 60 (Fisher, Pittsburgh, PA) equipped with a 0.125in OD probe was used by suspending the probe in the flour/extraction solvent mixture to a depth of approximately 4-7mm from the bottom of the μ-tube. The μ-tube

was suspended in ice water during application of sonication energy to mitigate heating. Preliminary evaluations of sonication time and energy application protocols consisted of: 30s at 10W output; 60s at 10W output; 30s at full power (13-16W); 4 x 15s with 30s pauses at 10W output; and 4 x 15s with 30s pauses at 10W output with a centrifuge and solvent change after the first two 15s sonication intervals (1:1 supernatant pool) . Following centrifugation (4min at 9300g), extracted proteins were analyzed by SEC as described above. Total SEC peak areas as well as polymeric peak area regions (peaks disappearing on reduction) were evaluated.

Extract Stability

To improve protein stability of extracted proteins due to effects of fungal and intrinsic seed proteases (Larroque et al. 2000), several chemical and physical post-extraction treatments were tested. Chemical treatment using phenylmethylsulfonylfluoride (PMSF) was done by adding PMSF to fresh protein extract to achieve a final PMSF concentration of 20mM, followed by 5min vortex mixing. A different chemical treatment using the commercial Protease Inhibitor Cocktail (PIC) Sigma P9599 (Sigma, St. Louis, MO) was also tested. Application rate (per supplied instructions) of PIC was 3uL PIC per mL fresh protein extract and was followed by 5min mixing by vortex. Two heat application protocols were also used, and were accomplished by heating 1mL portions of fresh protein extract while in 2mL polypropylene μ-tubes in a temperature regulated heat block (Reacti-Therm III, Pierce Inc.) for a 2min period at either 80 ˚C or 100 ˚C (Larroque et al. 2000). Post-treatment effects were evaluated by comparing protein peak area regions from SEC chromatograms of separations on a Biosep-S3000 column injected over a period of 96hrs.

Final Extraction Method (3F method)

A water/salt soluble protein fraction (albumins + globulins) and three additional protein fractions (subsequently referred to as F1, F2, and F3) were obtained by application of a sequential extraction scheme to duplicate 100mg milled sample portions in 2.0mL μ-tubes. **Figure 3.1** is a schematic of the entire extraction procedure. Albumins, globulins, and low molecular weight nitrogen (LMWN) were extracted first, using two 5min extractions with 1mL of 1M NaCl (Taylor et al. 1984) each extraction (room temperature vortex at medium speed followed by centrifugation at 9300g for 4min). This fraction was retained or discarded based on the objective of the analysis. This was followed by a 5min wash step with 1mL of deionized H_2O, mixed and centrifuged as before, and discarded. Next, Fraction 1 (F1) was obtained using two sequential 50min extractions with 1mL portions of 60% t-butanol (v/v) / 80mM Tris-borate pH10.0, mixed and centrifuged as before, and the supernatants were combined 1:1 in a clean μ-tube. To obtain Fraction 2 (F2), the same pellet was subjected to two sequential 50min extractions with 1mL of 50mM Tris-borate pH10.0 / 2% SDS (w/v), mixed and centrifuged as before, and the supernatants were combined 1:1 in a clean μ-tube. Sonication was applied to the same pellet from the previous extractions to obtain Fraction 3 (F3) using a single 1mL portion of 50mM Tris-borate pH10.0 / 2% SDS (w/v). A 0.125in OD sonication probe powered by a Sonic Dismembrator 60 unit (Thermo Fisher Scientific, Waltham, MA) was positioned within the tube contents, and centered approximately 4-7mm from the bottom of the μ-tube that was immersed in an ice water bath to reduce solution heating. Three 15s bursts of sonic energy were applied while manually maintaining 10W of output power (sonicator readout), and allowing 30s rests between and after the final burst. After sonication, the mixture was centrifuged as before and the supernatant was transferred to a clean μ-tube.

After each fraction was extracted, the extracted supernatants were subjected to heat deactivation of intrinsic proteases by application of 2min at 80˚ C heat during immersion of the capped µ-tubes in the water-filled wells of a heat block (Pierce, Rockford, IL). After being allowed to cool to room temperature, extracts were filtered through 25mm 0.45µm GHP membrane syringe filters (PALL Life Sciences, Port Washington, NY) into clean µ-tubes or HPLC vials. Fraction extracts intended for size exclusion chromatography (SEC) analysis were analyzed fresh (not lyophilized).

Size Exclusion Chromatography (preliminary experiments)

All high performance liquid chromatography (HPLC) separations during preliminary experiments were carried out using an Agilent 1100 HPLC system (Agilent, Santa Clara, CA) equipped with a variable wavelength UV detector. Separations were monitored at 210nm, flowrate was 1.0mL/min, and all injections were 15µL unless otherwise noted. Column temperatures were evaluated from 30 ˚C to 50 ˚C with negligible differences in separations noted, so 40 ˚C was chosen for the remainder of the study.

Two different types of SEC columns, silica packing based and polymer packing based, were evaluated for separation of non-reduced sorghum protein extracts. Silica packing based columns were of two types. These consisted of a number of the Biosep (Phenomenex, Torrance, CA) series SEC columns (300mm x 7.80mm) differing in nominal separation ranges: S2000 (1 to 300 kD); S3000 (5 to 700 kD); S4000 (15 to 1,500 kD). Biosep columns have a pH stability range of 2.5 to 7.5 (http://phenomenex.com). An additional set of silica based columns were from the Yarra SEC (Phenomenex, Torrance, CA) family of columns (300mm x 7.80mm). Yarra columns have a pH stability range of 2.5 to 7.5. Polymer packing based columns consisted of a number of the Polysep (Phenomenex, Torrance, CA) series SEC columns (300 x 7.80mm)

differing in nominal separation ranges: P2000 (10 to 100 kD); P3000 (75 to 250 kD); P4000 (3 to 400 kD); P5000 (50 kD to 2 M). Polysep columns have a pH stability range of 3.0 to 12.0 (http://phenomenex.com). Several combinations of columns within packing type were also investigated for effect on separations when placed in series during chromatographic analysis.
A variety of mobile phases were studied. Organic solvent based mobile phases investigated included: a) 50% acetonitrile (ACN) / 0.1% trifluoroacetic acid (TFA); b) 20% ACN / 0.1% TFA; c) 50% ACN / 50mM K-phos pH 7.0; d) 50% ACN / 25mM Tris-borate pH 7.0; e) 50% ACN / 20mM NH4OH pH 10.0; f) 50% methanol / 0.1% TFA; g) 50% methanol / 25mM Tris-borate pH 7.0; h) 50% IPA / 0.1% TFA; i) 50% IPA / 25mM Tris-borate pH 7.0; and j) 50% ACN / 20mM K-phos pH 7.0 / 0.5% SDS. Buffer and buffer plus detergent based mobile phases investigated included: a) 50mM Na-phos pH 7.0 / 1% SDS; b) 50mM Na-phos pH 10.0 / 1% SDS; c) 50mM Tris-borate pH 7.0 ; d) 50mM Tris-borate pH 7.0 / 1% SDS; e) 5mM Tris-borate pH 7.0 / 6mM SDS; f) 25mM Tris-borate pH 7.0 / 3mM SDS; and g) 5mM Tris-borate pH 7.0 / 2mM SDS.

Size Exclusion Chromatography (final conditions)

Final SEC protein separation conditions consisted of 50uL injections on silica based Yarra SEC-3000 columns, 300mm x 7.80mm (Phenomenex USA, Torrance, CA). Isocratic runs were accomplished with a mobile phase of 50mM Tris-borate pH7.0 / 1% SDS (w/v) and a flowrate of 0.5mL/min while maintaining a column temperature of 40˚ C. A UV wavelength of 210nm was monitored for analyte detection.

Results and Discussion

Preliminary Extraction Experiments

Initial extraction screening tests using sodium hydroxide (NaOH) solutions alone and in conjunction with the detergent sodium dodecylsulfate (SDS) were conducted (data not shown). Sorghum proteins have low levels of positively charged amino acids (arginine, lysine, histidine), and as a result are more soluble at high pH (de Mesa-Stonestreet et al. 2010). Studies involving sorghum protein extraction for biofilms (Gao et al. 2005), to obtain protein concentrates (Wu 1978), as well as extraction of proteins from whole wheat (Wu and Sexson, 1975) have noted the beneficial effect that addition of NaOH provides in increasing cereal protein extraction efficiency. A 50mM NaOH / 2% (v/v) SDS solution resulted in 90% of available protein being extracted as determined by nitrogen combustion. The measured pH of a 50mM NaOH solution was 12.3.

Subsequent extractions done using 50mM NaOH / 2% (v/v) SDS and analyzed over a period of 37.5hrs indicated a decrease over time in SEC peaks representing the largest polymers, and was presumed due to degradation from the highly alkaline environment (data not shown). Since the goal of characterizing sorghum polymeric proteins would be defeated by solvent degradation of these analytes, less harsh alkaline pH alternatives were sought. Sodium phosphate (Na-phos) pH 10.0 / 2% SDS was chosen for further extraction optimization work because it represented a compromise in pH level, yet retained the relatively high pH and sodium component of the more harsh NaOH solutions. Addition of various extraction solvent additives and combinations to the Na-phos pH 10.0 buffer (**Table 3.1**) were attempted in an effort to enhance protein extraction (Figure 2). No improvement in total protein extracted was realized

over that obtained with Na-phos pH 10.0 / 2% SDS alone, as evidenced by the total peak area from SEC runs.

A final set of buffer systems based on sodium borate (Na-borate) pH 10.0 and Tris-(hydroxymethyl)-aminomethane borate (Tris-borate) pH 10.0 were investigated for comparison to Na-phos pH 10.0 / 2% SDS. The results of these extractions (**Figure 3.3**) indicated a buffer extraction solution based on Tris-borate pH 10.0 / SDS provided the best total protein as well as polymeric protein extraction. Additional solvents and additives were added to 50mM Tris-borate pH 10.0 buffer (**Table 3.1**) to observe the effects on extractions, but no improvement in either total protein or polymeric protein extracted was realized. **Figure 3.4** shows the effects of varying the levels of SDS in 50mM Tris-borate pH 10.0 and how these compare to the best Na-phos pH 10.0 / SDS and Na-borate pH 10.0 / SDS extraction solutions. The 50mM Tris-borate pH 10.0 in combination with 1% or 2% SDS were the most favorable extraction solvent combinations. A 50mM Tris-borate pH 10.0 / 2% SDS extraction solvent was decided upon for the remainder of the study.

The final portion of preliminary extraction work was to explore application of sonication energy to help improve extraction of non-reduced protein polymers (Singh et al. 1990; Singh and MacRitchie, 2001). Several sonication protocols were evaluated for effectiveness in improving the extraction of non-reduced sorghum protein polymers. Effective sonication was defined as the application method that would extract additional non-reduced large sorghum proteins that simple vortex mixing would not. Each sonication protocol was applied to a given sample three times in succession to see how much additional protein could be extracted. After each application (e.g. one course of 4 x 15s at 10W), the extracted proteins were quantified using SEC peak area. The

4 x 15s at 10W protocol extracted a total of almost 12% more protein over the course of the three applications than did the next most efficient protocol (**Figure 3.5**).

Extract Stability

Because it is frequently necessary to analyze large numbers of samples at a time, it is required that sample extracts do not degrade during lengthy post-extraction periods while waiting in queue for instrumental analysis. To assess extract stability over time, we examined how protein extracts changed over time in the presence of several compounds and physical treatments designed to prevent protein degradation. Naturally occurring serine proteases of fungal origin have been found in sorghum grain (Huang et al. 2000). Such enzymes are active, even under basic extraction conditions, and were shown to effect changes in extracted proteins resulting in disappearance of protein bands during SDS-PAGE analysis (Huang et al. 2000). In that study, phenylmethylsulfonylfluoride (PMSF) effectively prevented degradation of extracted protein. A commercially available protease inhibitor cocktail containing a mixture of protease inhibitors was also tested. Removal of bran by mechanical decortication (TADD, Venables Mfg., Canada) prior to extraction was another physical method examined for reducing enzyme induced changes in extracted proteins. We reasoned removal of the outer seed coat would reduce the presence of fungi and accompanying proteases, as well as possible contributing effects from interfering phenolic compounds in the bran (McGrath et al. 1982). Simple application of heat for inactivation of protease activity has been effectively applied before in wheat flour and meal (Larroque et al. 2000). The effectiveness of the various treatment applications in minimizing extracted protein degradation over time is presented in **Figure 3.6**. The decrease in extracted polymeric protein was quite pronounced in the untreated and cocktail treated samples during the first 24hrs. Decortication or PMSF provided moderately better results, slowing degradation after

24hrs to approximately half the levels seen in the untreated or cocktail treated samples. The two heat treatments provided far better prevention from degradation, with the 80 ˚C treatment slightly preferred. The 80 ˚C treatment reduced degradation to <5% over a twenty-four hour period based on total SEC peak area. To insure heating samples had no effect on chromatographic patterns or peak areas, extracts were divided and one aliquot was subjected to heat treatment prior to analysis while the other was analyzed immediately with no heat treatment. No statistically significant difference in either pattern or peak area was seen (data not shown), indicating the heating process itself as applied to inactivate protein degrading enzymes was not detectably altering the extracted proteins.

Final Extraction Method (3F)

The ultimate goal of the developed extraction procedure was to obtain sorghum proteins in forms reflecting the non-reduced native complement within given samples, thereby allowing meaningful comparisons of samples varying in attributes such as IVPD. As opposed to traditional sorghum protein extraction methods, no reducing agents were used. The more differentiated and discreet the obtained protein fractions can be from among the many proteins making up a total sample complement, the more complete and detailed will be the picture of sample differences. To better accomplish this, an extraction method resulting in multiple fractions based on differential solubility was considered preferable for achieving removal of sequential protein layers, and was incorporated using the results from the preliminary extraction trials already discussed, as well as adaptations from previously published methods (Ioerger et al. 2007; Taylor et al. 1984; Singh et al. 1990).

Taylor et al. (1984) found that a 1.0M NaCl solution, followed by extraction with water, was best for the extraction of sorghum albumin and globulin proteins, and in the current study

this solvent was applied over a series of ten successive 5min extractions to optimize the extraction of these proteins. Based on the total SEC peak area, 99.5% of the albumin, globulin, and LMWN fraction obtained over the course of ten extractions was recovered from the two initial 5min 1mL 1.0M NaCl extractions. This was the procedure implemented in the final method. A 5min 1mL water extraction following the NaCl extractions was also employed to aid removal of residual salt. This fraction was designated albumins/globulins.

To enhance extraction of kafirin proteins in polymeric form, an adaptation of the Landry-Moreaux 60% t-butanol procedure was applied, involving alterations of solvent pH by addition of 0.5% w/v sodium acetate or 60mM Na-phos at pH 3, 7, and 9. **Figure 3.7** shows SEC profiles for these extracts. Although the 60% t-butanol / 60mM Na-phos pH 9 extraction solvent resulted in the most extracted protein (especially in the early eluting polymeric range), buffer insolubility issues precluded its routine use. In fact, difficulty was encountered in maintaining solubility of any of the sodium containing buffers mentioned when combined with 60% t-butanol. A suitable high pH buffer alternative that remained soluble in 60% t-butanol was needed. Of those investigated, Tris-borate pH 10.0 at 80mM was optimal in terms of protein extracted as illustrated in **Figure 3.8**. The optimal application and extraction rate protocol was determined as illustrated in **Figure 3.9**. Little to no additional material was extracted after five sequential extractions. Approximately 86.5% of the protein extracted over five sequential 50min extraction periods was recovered in the first two 50min periods as determined by SEC peak areas. The proteins from the initial two 50min extractions were designated Fraction 1 (F1).

Using the 50mM Tris-borate pH 10.0 / 2% SDS extraction solvent found to recover high levels of polymeric as well as total protein in the preliminary extraction investigation, a series of five repeated vortex extractions on the same sample for periods of 5, 25, or 50min was applied to

sample flours that had already been extracted for albumins/globulins and F1. **Figure 3.10** shows the comparative amounts of total protein extracted over the course of the repeated extractions for different extraction times. The most protein was extracted using the 50min extraction time series. After the first two 50min extractions, approximately 71.5% of the protein obtained over the course of five 50min extractions was recovered when measured as a proportion of the cumulative total SEC peak area. In **Figure 3.11**, the chromatographic traces of five sequential 50min extractions are overlaid to show how protein amounts and relative molecular size distributions changed over the course of the five repeated extractions. The proteins obtained from two 50min vortex applications using 1mL portions of 50mM Tris-borate pH 10.0 / 2% SDS extraction solvent to samples pre-extracted for albumins/globulins and F1 were designated Fraction 2 (F2).

The final extraction was made using sonication to assist in extracting remaining material. Preliminary extraction experiments revealed that longer duration applications of sonic energy (i.e. 30s and 1min @ 10W) did not extract as much protein as multiple shorter duration times. Longer than 1min resulted in the beginning of starch gelatinization, and was presumably a result of localized heating around the probe. This was not an issue for shorter times, and for these reasons shorter sonication times were preferred. A series of 15s sonication extractions were carried out using 50mM Tris-borate pH 10.0 / 2% SDS. Sonication energy bursts were applied for either 1 x 15s, 2 x 15s, 3 x 15s, or 4 x 15s to separate samples already extracted for albumins/globulins, F1, and F2 (detailed description in Material and Methods section). Analysis by SEC of the resulting extracts allowed comparison of total peak areas and profiles. As illustrated in **Figures 3.12** and **3.13**, application of three 15s bursts of sonic energy while maintaining 10W output power (sonicator readout) and allowing 30s rests between and after the

final burst resulted in recovery of the most protein as measured by total SEC peak area. Proteins recovered by application of this protocol were designated Fraction 3 (F3).

An example of the SEC analyses of a sample subjected to the entire extraction sequence is shown in **Figure 3.14**. SEC peaks eluting prior to ~12.5min represent protein polymers as evidenced by a decrease in peak size after treatment with a reducing agent (data not shown). Based on calculations using a SEC standard curve constructed using purified kafirin isolate, application of the entire extraction fractionation method resulted in the extraction of approximately 87.5% of the available protein content. The molecular weight distribution of extracts obtained from application of the full sequence of extraction procedures increases over the course of the sequence (albumin/globulin through F3), indicating relatively more polymeric material is being removed with each successive extraction procedure. It is also worth noting that during the preliminary development of each step, repeated applications of a given solvent/method would result in gradual decreases in the amount of protein being extracted (monomeric as well as polymers). Yet when a different solvent and/or sonication was applied, protein (monomeric and polymeric) was again obtained. It seems reasonable to speculate these results imply there are some as yet uncharacterized differences in how the proteins obtained from each fraction associate or are packaged.

SEC Separation Optimization

Column selection for SEC separation of extracted non-reduced polymeric proteins is a critical factor in being able to accurately assess differences among samples. If the molecular weight cutoff is too low, variations in the largest proteins will be hidden at the exclusion limit of the largest molecules, and this will appear as a large indistinct peak at the beginning of chromatograms. Inversely, too high a cutoff and little or no differentiation based on molecular

size will occur. Long run times and poor resolution among late-eluting, poorly shaped peaks are frequently the result. Other considerations include compatibility with the chosen mobile phase, and minimizing non-specific sample interactions with the column. Therefore, the goal in this portion of the study was to find a size exclusion column with a molecular size separation range that would separate the majority of sorghum proteins extracted in non-reduced form into distinct peaks representing relatively narrow bands of protein molecules of given molecular size.

A variety of mobile phases were studied to observe the effect on protein separations. Conditions used in previous SEC studies on other cereal proteins served as mobile phase composition starting points (Zhao et al. 2008; Ioerger et al. 2007). Mobile phase selections and compositions took into consideration compatibility with a given column packing and particular extraction solvent being used, as well as UV cutoff for use at the desired detector wavelength. The availability of column packings tolerant of very high pH mobile phases (> pH 10), as in the polymer packing based columns, was seen as a possible advantage in allowing the use of a system beneficial to analysis of the sorghum proteins. In contrast, the maximum recommended pH for the silica based columns was pH 7.5.

The resolution of separations achieved using the polymer based columns was never as distinct as that obtained with the silica based columns. Chromatographic separations based on molecular size are really discriminating between protein molecules based on their hydrodynamic radius rather than molecular weight (Potschka, 1988). Therefore, the three-dimensional conformation of a protein in a given mobile phase is a critical factor. How a proteins size and shape (hydrodynamic radius) impacts its mobility as it passes by the pores present in a given column packing will determine the ultimate resolution achieved. It may be that the shape or orientation of the pores present in the polymer packings are inherently less likely to allow

analyte interaction for our extracts than those present in the silica based packings, thus resulting in poorer resolution. Silica based packings on the other hand, possess charged silanol groups that may provide analyte interactions that lend an additional separation mechanism based on the surface ionic character of the protein analyte. This may help explain why there are a greater number of peaks in evidence on separations with the silica based columns in all mobile phases compared to the polymer column. A comparison of the separations achieved with the Yarra SEC-3000 column and the Biosep SB-3000 using 50mM Tris-borate pH 10.0 / 1% SDS as mobile phase is presented in **Figure 3.15**. The Yarra SEC-3000 column provided the best non-reduced sorghum protein extraction separation achieved in this study.

Conclusions

A method for extracting non-reduced sorghum proteins was developed that optimized the recovery of total as well as polymeric protein species. After removal of albumins and globulins, a 3-step sequential procedure (3F) involving no reducing agents was applied to ground whole sorghum flour. The three fractions so obtained represented proportionally different protein polymer contents as evidenced by comparative size exclusion chromatography. The resulting extracts were stabilized against subsequent protease degradation by heating for 2min at 80 ˚C. Evaluation of silica and polymer based SEC columns resulted in selection of a silica column with a 5 kD to 700 kD nominal molecular weight separation range. The optimal mobile phase consisted of 50mM Tris-borate pH 7.0 / 1% SDS. Future work will involve application of the developed method to widely varying sorghum varieties to further characterize and confirm how

the non-reduced polymeric proteins influence sorghum protein digestibility, and should help provide additional insight into sorghum protein structure.

References

Bean S.R., and Lookhart, G.L. 2001. Factors influencing the characterization of gluten proteins by size-exclusion chromatography and multiangle laser light scattering (SEC-MALLS). Cereal Chem. 78(5):608-618.

Bean, S.R., Lyne, R.K., Tilley, K.A., Chung, O.K., and Lookhart, G.L. 1998. A rapid method for quantization of insoluble polymeric proteins in flour. Cereal Chem. 75:374-379.

Chakraborty, K., and Khan, K. 1988. Biochemical and breadmaking properties of wheat protein components. II. Reconstitution baking studies of protein fractions from various isolation procedures. Cereal Chem. 65:340-344.

Dachkevitch, T., and Autran, J.C. 1989. Prediction of baking quality of bread wheats in breeding programs by size-exclusion high-performance liquid chromatography. Cereal Chem. 66:448-456.

De Mesa-Stonestreet, N.J., Alavi, S., and Bean, S.R. 2010. Sorghum proteins: the concentration, isolation, modification, and food applications of kafirins. J. Food Sci. 75(5):R90-R104.

Duodu, K.G., Taylor, J.R.N., Belton, P.S., and Hamaker, B.R. 2003. Factors affecting sorghum protein digestibility. J. Cereal Sci. 38:117-131.

El Nour, I.N.A., Peruffo, A.D.B., and Curioni, A. 1998. Characterization of sorghum kafirins in relation to their cross-linking behavior. J. Cereal Sci. 28:197-207.

Emmambux, M.N. and Taylor, J.R.N. 2009. Properties of heat-treated sorghum and maize meal and their prolamin proteins. J. Agric. Food Chem. 57:1045-1050.

Ezeogu L.I., Duodu, K.G., Emmambux, M.N., and Taylor, J.R.N. 2008. Influence of cooking conditions on the protein matrix of sorghum and maize endosperm flours. Cereal Chem. 85(3):397-402.

Gao, C., Taylor, J., Wellner, N., Byaruhanga, Y.B., Parker, M.L., Mills, E.N., and Belton, P.S. 2005. Effect of preparation conditions on protein secondary structure and biofilm formation of kafirin. J. Agric. Food Chem. 53(2):306-312.

Gupta, R.B., Khan, K., and MacRitchie, F. 1993. Biochemical basis of flour properties in bread wheats. I. Effects of variation in the quantity and size distribution of polymeric protein. J. Cereal Sci. 18:23-41.

Hamaker, B.R., Kirleis, A.W., Butler, L.G., Axtell, J.D., and Mertz, E.T. 1987. Improving the in vitro protein digestibility of sorghum with reducing agents. Proc. Natl. Acad. Sci. USA 84:626-628.

Huang, C.P., Hejlsoe-Kohsel, E., Han, X.Z., and Hamaker, B.R. 2000. Proteolytic activity in sorghum flour and its interference in protein analysis. Cereal Chem. 77(3):343-344.

Ioerger, B., Bean, S.R., Tuinstra, M.R., Pedersen, J.F., Erpelding, J., Lee, K.M., and Herrmann, T.J. 2007. Characterization of polymeric proteins from vitreous and floury sorghum endosperm. J. Agric. Food Chem. 55:10232-10239.

Larroque, O.R., Gianibelli, M.C., Gomez Sanchez, M., and MacRitchie, F. 2000. Procedure for obtaining stable protein extracts of cereal flour and whole meal for size-exclusion HPLC analysis. Cereal Chem. 77(4):448-450.

McGrath, R.M., Kaluza, W.Z., Daiber, K.H., van der Riet, W.B., and Glennie, C.W. 1982. Polyphenols of sorghum grain, their changes during malting, and their inhibitory nature. J. Agric. Food Chem. 30:450-456.

Oria, M.P., Hamaker, B.R., and Shull, J.M. 1995. Resistance of sorghum α-, β-, γ-kafirins to pepsin digestion. J. Agric. Food Chem. 43:2148-2153.

Orth, R.A. and Bushuk, W. 1972. A comparative study of the proteins of wheats of diverse baking qualities. Cereal Chem. 49:268-275.

Potschka, N. 1988. Size-exclusion chromatography of polyelectrolytes: experimental evidence for a general mechanism. J. Chromatogr. 441:239-260.

Sapirstein, H.D., and Fu, B.X. 1998. Intercultivar variation in the quantity of monomeric proteins, soluble and insoluble glutenin, and residue protein in wheat flour and relationships to breadmaking quality. Cereal Chem. 75:500-507.

Singh, N.K., and MacRitchie, F. 2001. Use of sonication to probe wheat gluten structure. Cereal Chem. 78:526-529.

Singh, N.K., Donovan, G.R., Batey, I.L., and MacRitchie, F. 1990. Use of sonication and size-exclusion high-performance liquid chromatography in the study of wheat flour proteins. I. dissolution of total proteins in the absence of reducing agents. Cereal Chem. 67(2):150-161.

Taylor J.R.N., Schüssler, L., and van der Walt W.H. 1984. Fractionation of proteins from low-tannin sorghum grain. J. Agric. Food Chem. 32:149-154.

Wu, Y.V. 1978. Protein concentrate from normal and high-lysine sorghums: preparation, composition, and properties. J. Agric. Food Chem. 26(2):305-309.

Wu, Y.V., and Sexson, K.R. 1975. Preparation of protein concentrate from normal and high-protein wheats. J. Agric. Food Chem. 23(5):903-905.

Zhao, R., Bean, S.R., and Wang, D. 2008. Sorghum protein extraction by sonication and its relationship to ethanol fermentation. Cereal Chem. 85(6):837-842.

Table 3.1 Extraction buffers, solvents and solution additives investigated for sorghum protein extraction.

Solutions / Buffers	Solvent / Additive(s)
10mM, 50mM, 100mM NaOH	n/a
10mM, 50mM, 100mM NaOH	2% SDS
50mM Na-phos pH10.0, pH11.0, pH12.0	n/a
50mM Na-phos pH10.0	4M Guanidine-HCl 60% t-Butanol 1M KI 25%, 50%, 75% Ethylene glycol 25% Ethylene glycol/2% SDS 50% Ethylene glycol/2% SDS 2% SDS
50mM Na-borate pH10.0	2% SDS
2M Sodium isothiocyanate	n/a
50mM Tris-borate pH10.0	0.5%, 1.0%, 2%, 4% SDS 2% Na-octyl sulfate (C8) 2% Na-decyl sulfate (C10) 2% Na-tetradecyl sulfate (C14) C8:C10 combinations 10%, 25%, 50% n-Propanol 1% C10 + 1% C8 / 10%, 25%, 50% n-Propanol 0.25% C10 + 2% C8 / 10%, 25% n-Propanol 1% SDS / 10%, 25%, 50% n-Propanol 2% SDS / 10%, 25%, 50% n-Propanol 4% SDS / 50% n-Propanol 25% BG, PG, DEG or TEG 2M Sodium isothiocyanate 50% acetonitrile / 2% SDS

Figure 3.1 Extraction flowchart for 3F method.

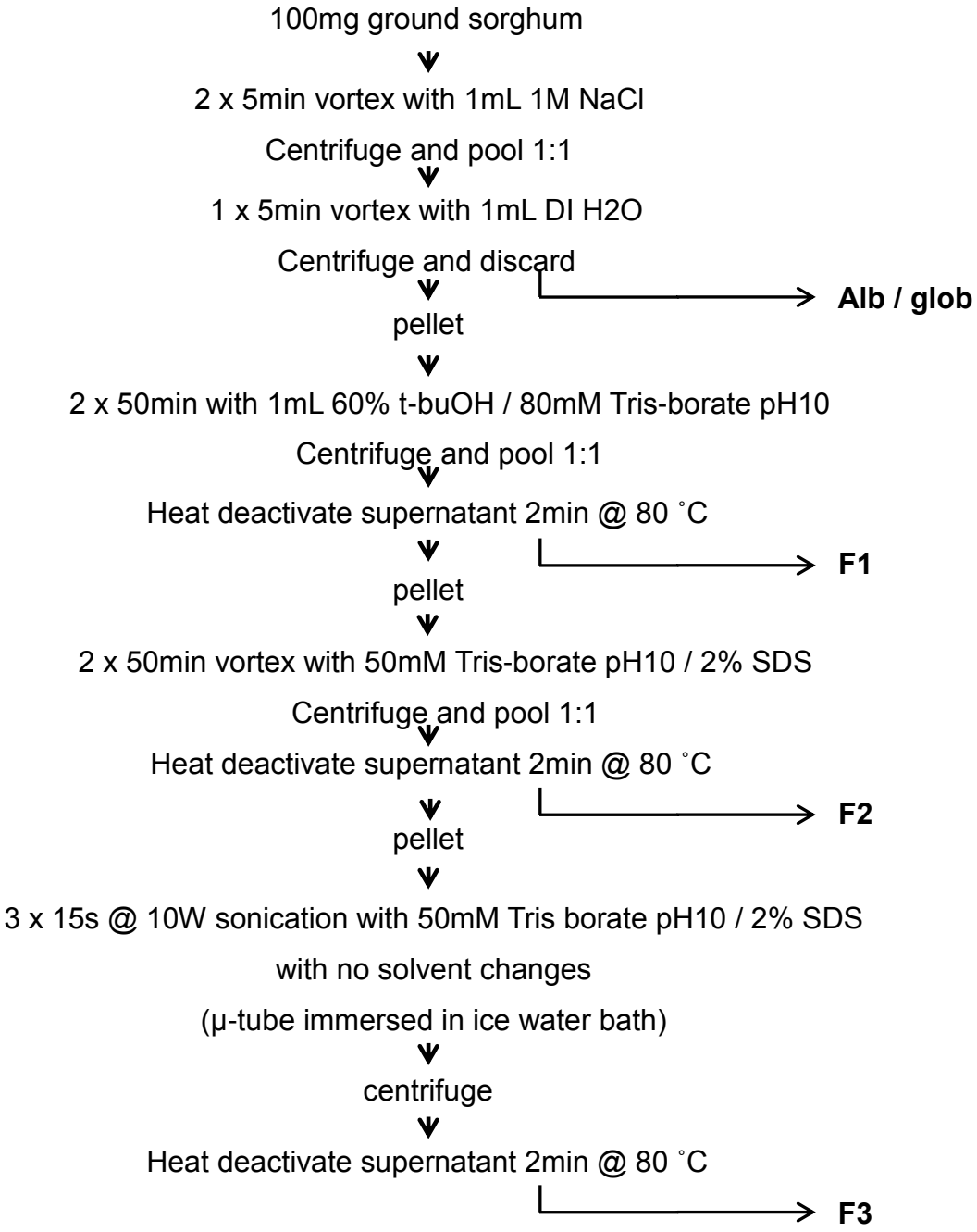


Figure 3.2 Preliminary extraction optimization using 50mM Na-phos pH 10.0 + indicated additive. (error bars = 1SD)

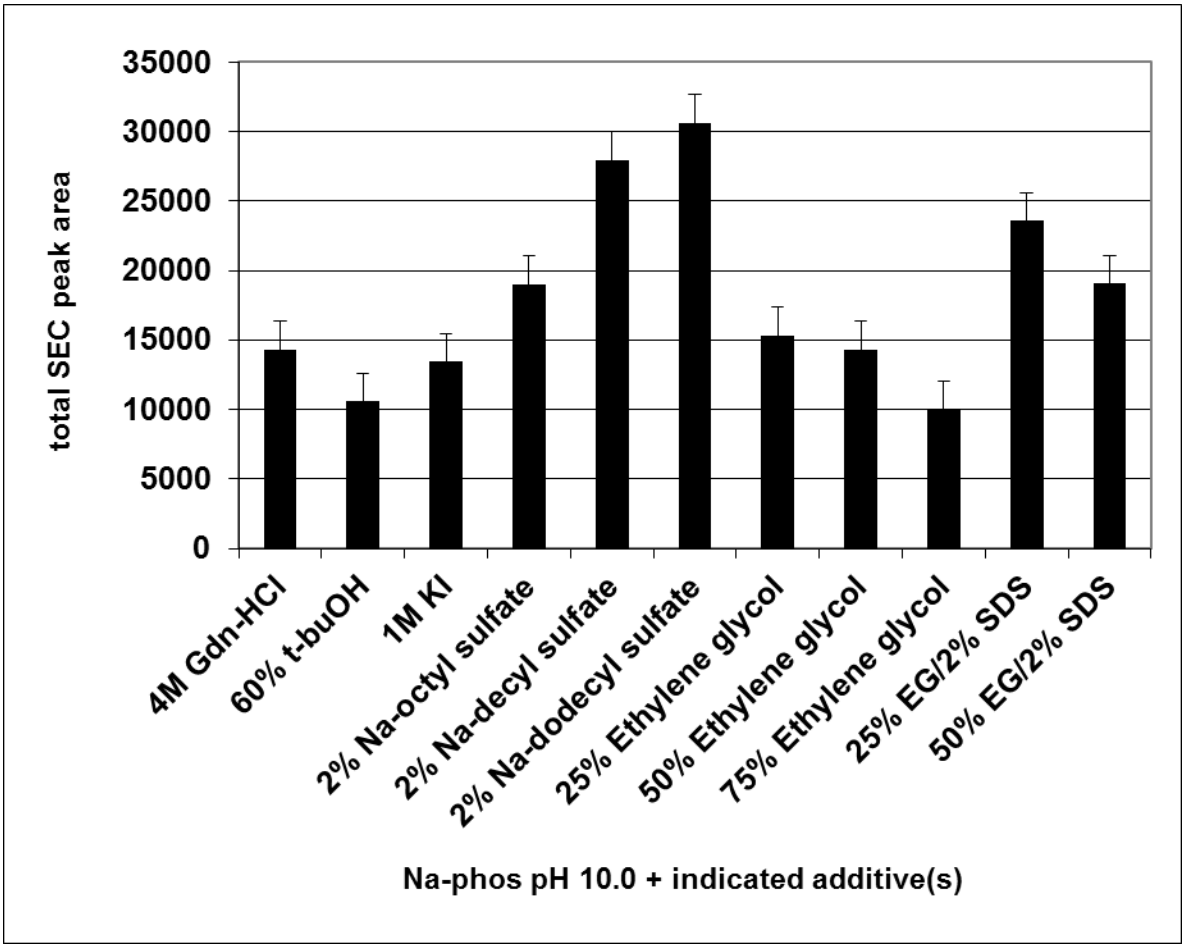


Figure 3.3 Preliminary experiments comparing total protein and polymeric protein peak areas from proteins extracted with Na-phos pH 10.0 / 2% SDS, Tris-borate pH 10.0 / 2% SDS, and Na-borate pH 10.0 / 2% SDS. Polymeric proteins identified using reducing conditions. Extractions and separation conditions as specified in Materials and Methods. (error bars = 1SD)

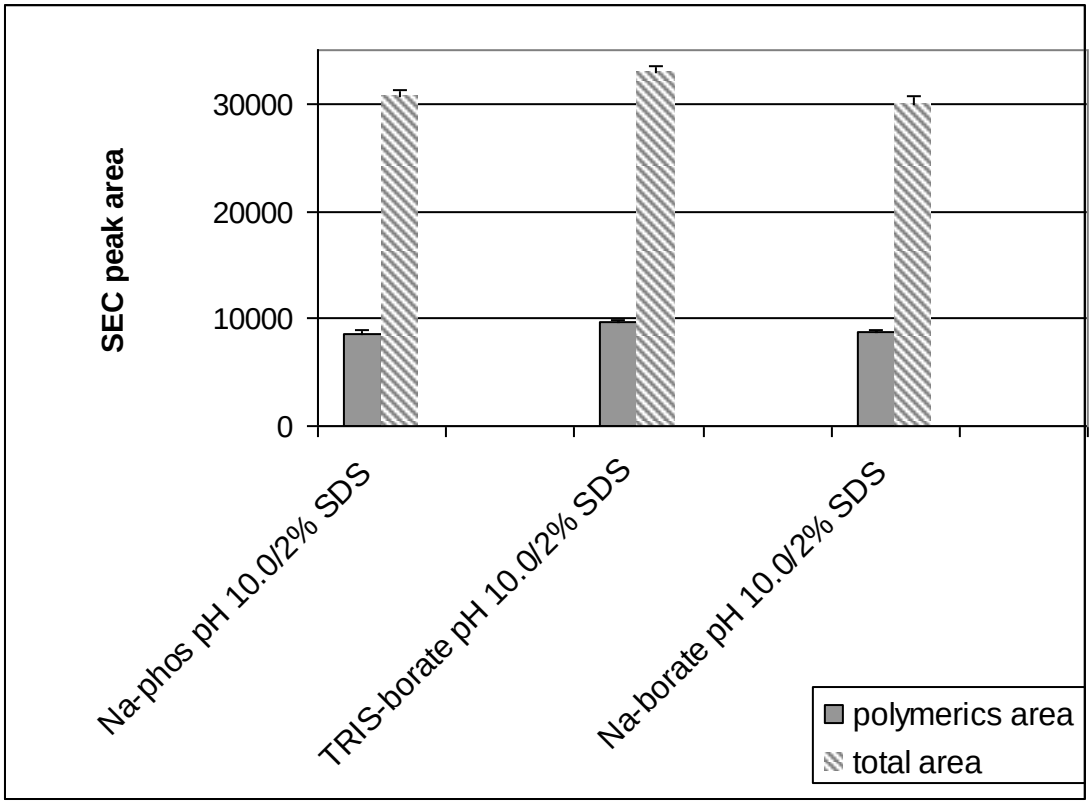


Figure 3.4 Preliminary experiments differentiating polymeric and non-polymeric proteins in extracts from different extraction solvents and additives combinations based on selected SEC peaks (polymeric peaks identified using post-extraction reducing conditions). (error bars = 1SD)

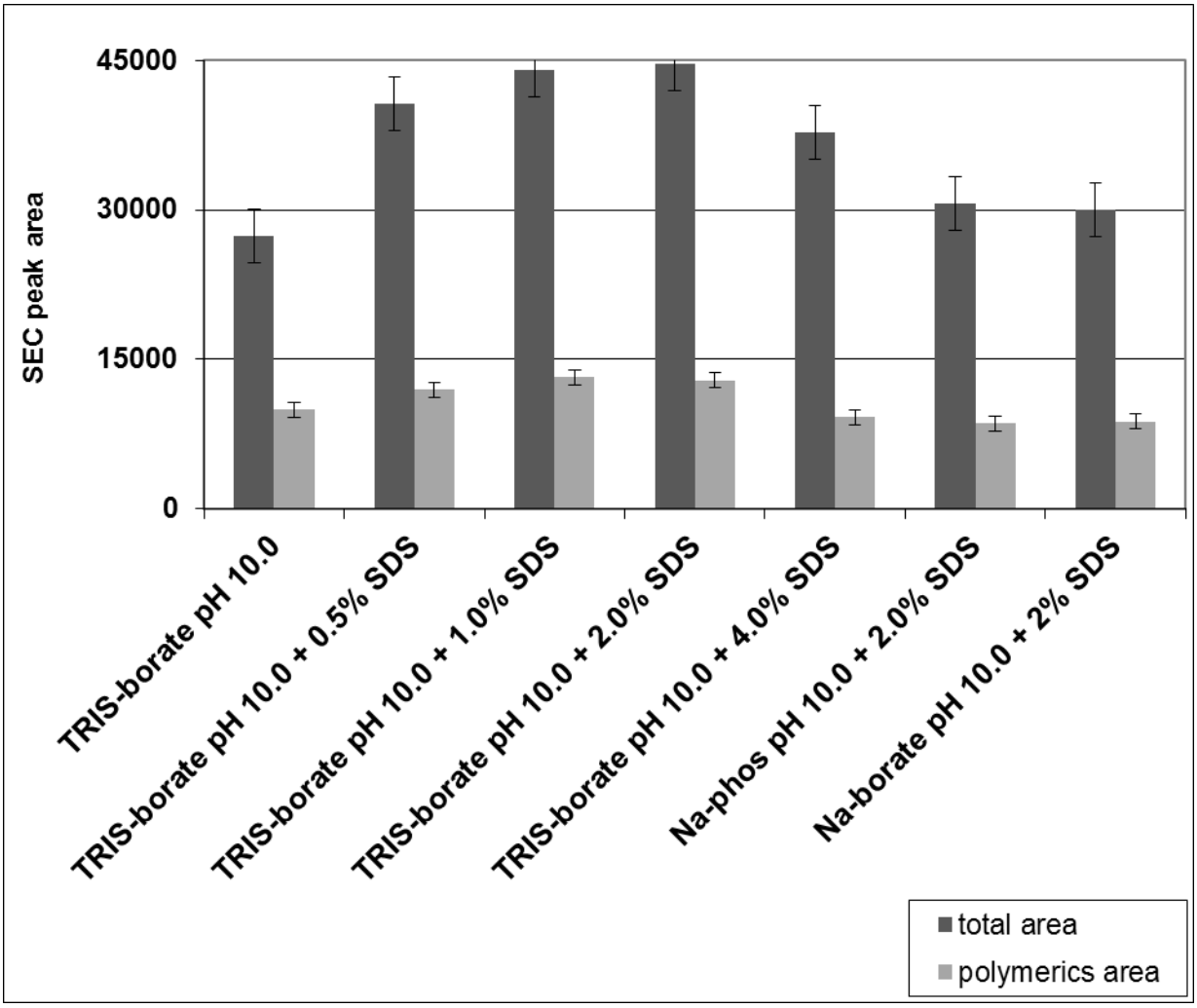


Figure 3.5 Preliminary experiment results from application of different time and sonication energy protocols to the non-reduced extraction of polymeric sorghum proteins in flour suspended in 50mM Tris-borate pH 10.0 / 2% SDS. Peak area data were collected for each of three separate applications of a given sonication time and energy protocol.

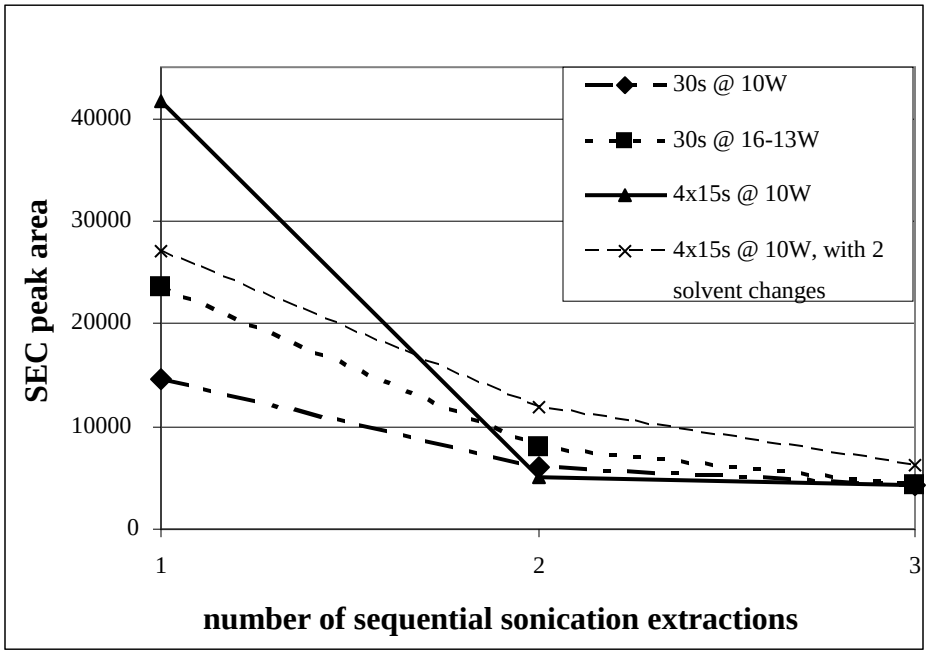


Figure 3.6 Preliminary experiments improving post-extraction protein stability using 2min heat treatments (80 ˚C or 100 ˚C); pre-extraction decortication (decort); 20mM phenylmethylsulfonylfluoride (PMSF); protease inhibitor cocktail, 3uL/50mg (cocktail); or no treatment (No trt). Graph represents analyses of the same extracts after initial treatments over a 96hr time period. Data points are averages of two treatment reps.

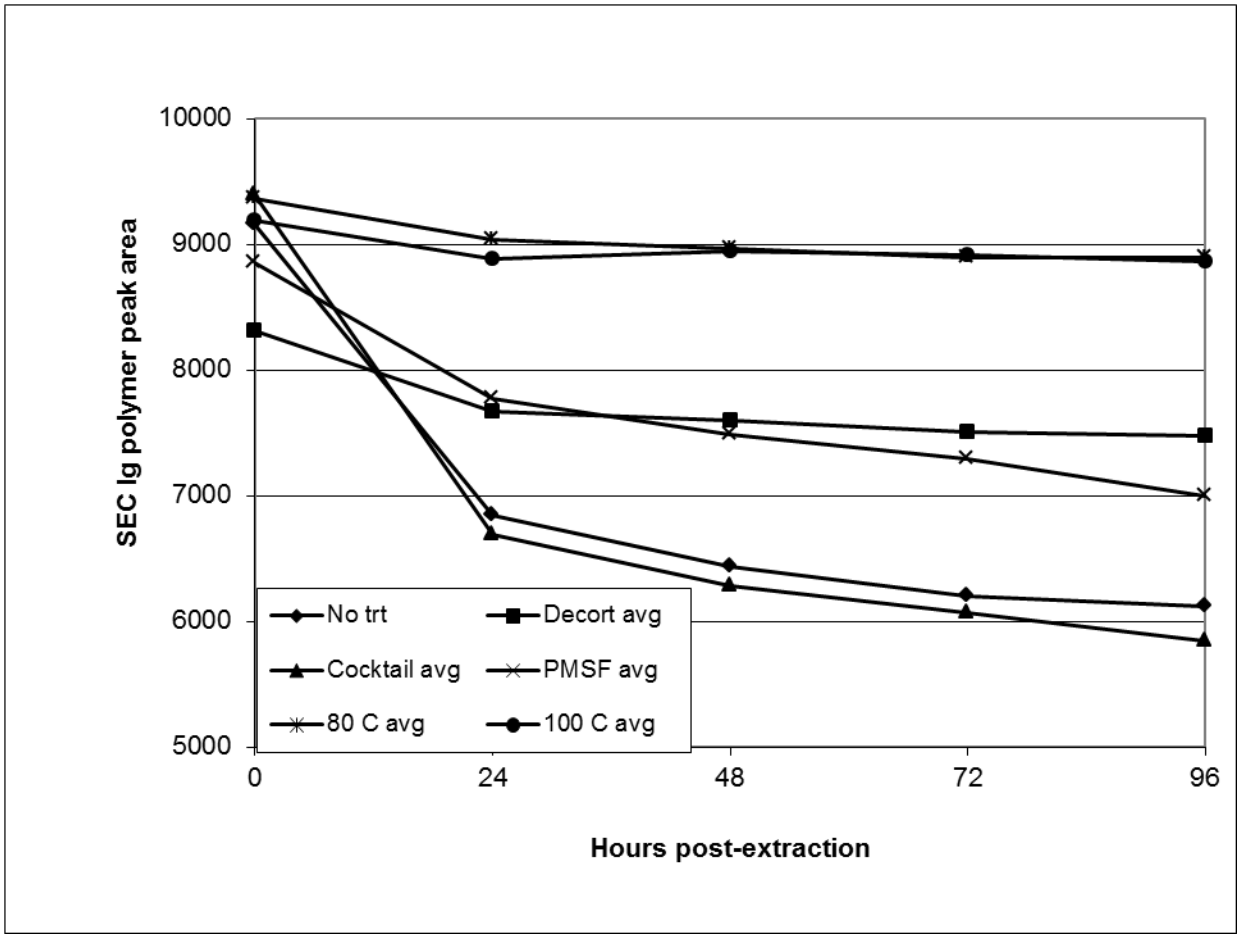


Figure 3.7 Altering the pH of the 60% t-butanol non-reduced kafirins extraction solvent to enhance extraction of protein polymers. These samples were pre-extracted for albumins/globulins.

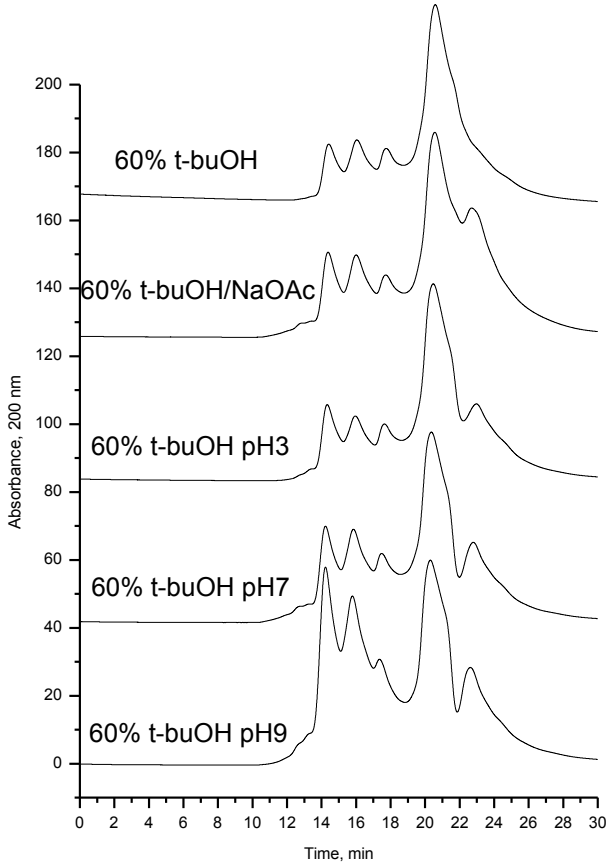


Figure 3.8 Total SEC peak areas of Fraction 1 (F1) extracted with 60% t-butanol + indicated Tris-borate pH10.0 buffer (one extraction per sample per buffer concentration). F1 fractions were obtained following extraction of albumins/globulins. (error bars = 1SD)

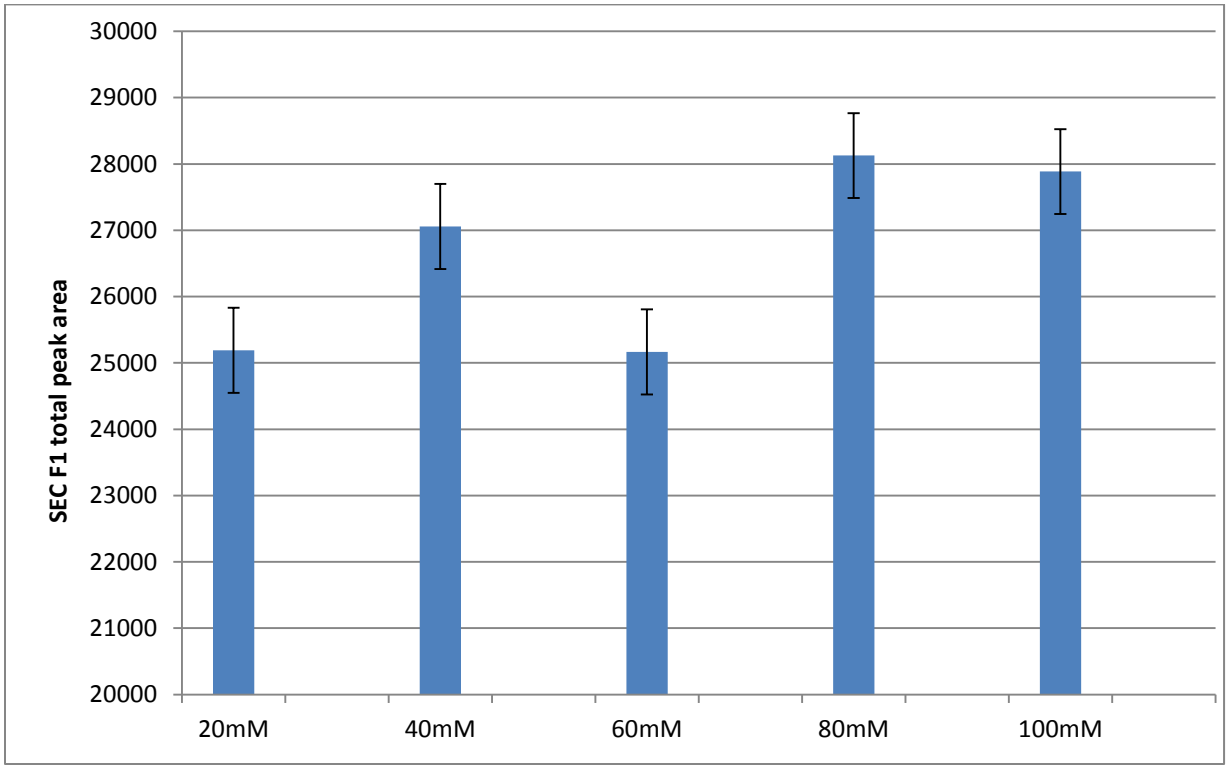


Figure 3.9 Sequential 50min Fraction 1 (F1) extractions with optimal 60% t-butanol + 80mM Tris-borate pH10.0 extraction solvent applied to same sample. F1 fractions were obtained following extraction of albumins/globulins. (error bars = 1SD)

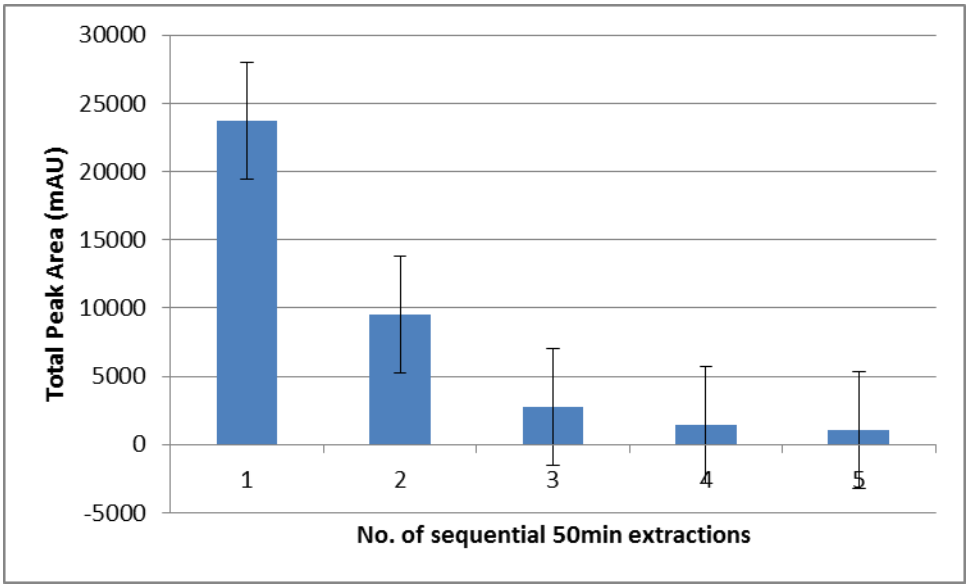


Figure 3.10 Fraction 2 (F2) optimization using repeated vortex extractions with 50mM Tris-borate pH 10 + 2% SDS for indicated times. F2 fractions were obtained following extraction of albumins/globulins and F1. (error bars = 1SD)

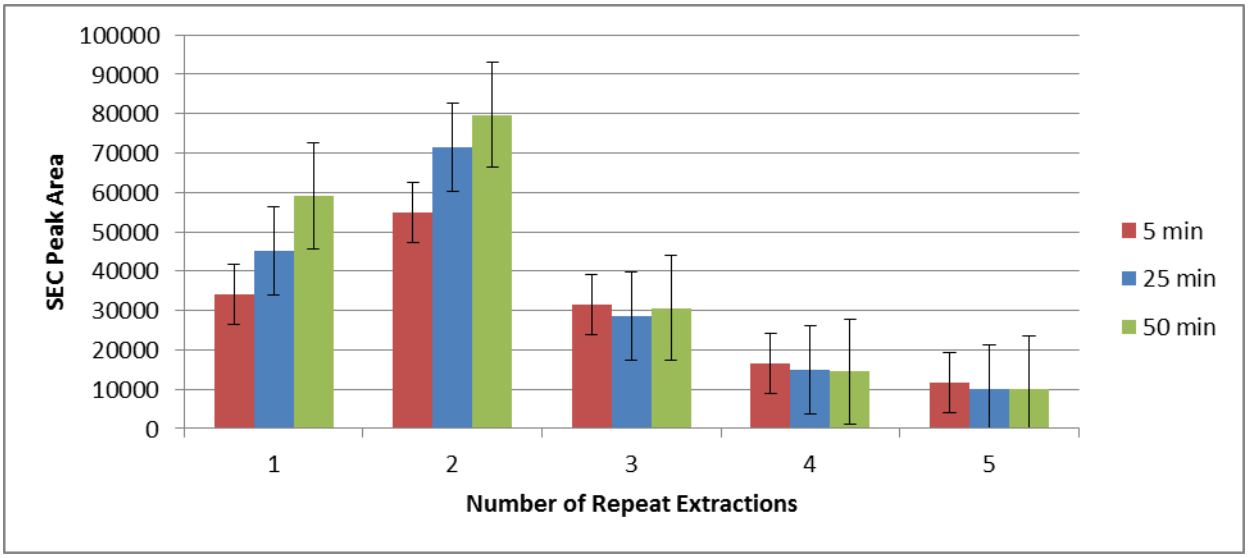


Figure 3.11 SEC chromatograms of repeat 50min sequential extractions of Fraction 2 (F2). These samples were pre-extracted for albumins/globulins and F1. Peak areas eluting prior to ~12.5min represent polymeric proteins based on profiles after reduction.

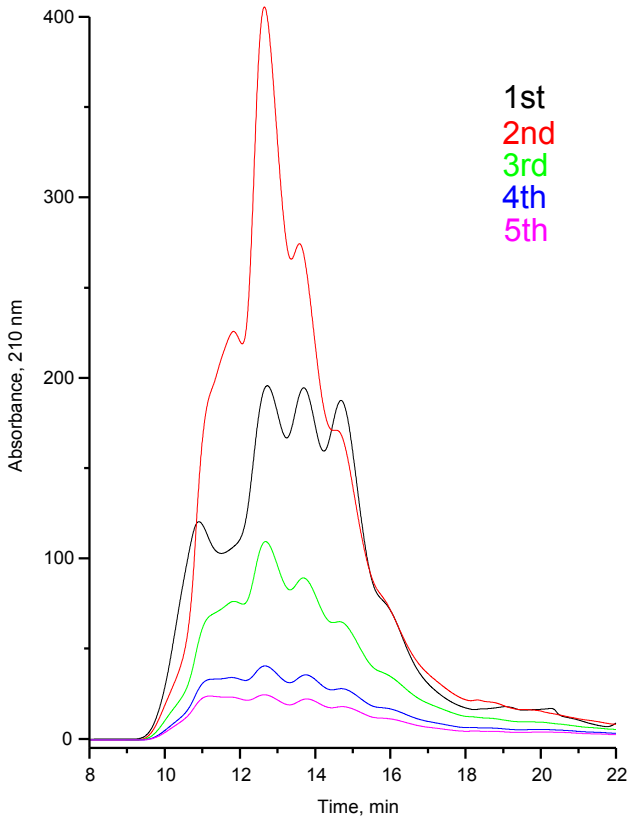


Figure 3.12 Sonication protocol for extraction of Fraction 3 (F3) illustrating the number of timed 10W sonic energy applications to individual samples (30s rests between multiple energy applications) suspended in 50mM Tris-borate pH 10 + 2% SDS. Each bar represents a different sample that was previously extracted to remove albumins/globulins, F1, and F2. (error bars = 1SD)

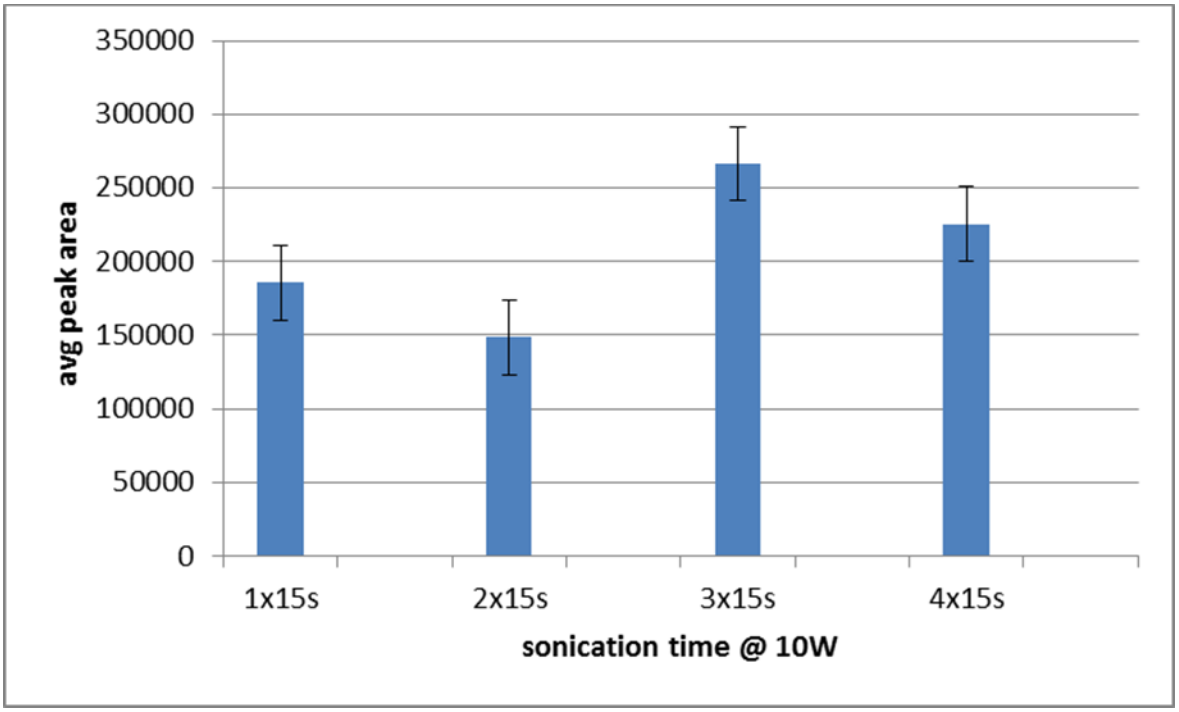


Figure 3.13 SEC chromatograms of Fraction 3 (F3) from timed 10W sonication energy applications to individual samples (30s rests between energy applications) suspended in 50mM Tris-borate pH 10 + 2% SDS. Each trace represents a different sample. These samples were pre-extracted for albumins/globulins, F1, and F2. Peak areas eluting prior to ~12.5min represent polymeric proteins based on profiles after reduction.

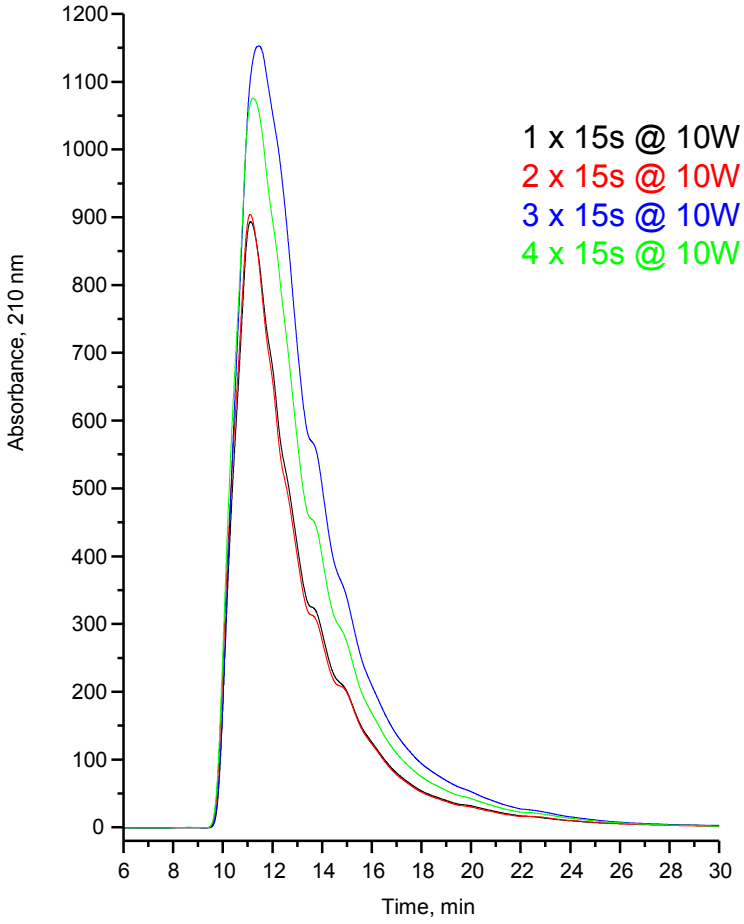


Figure 3.14 SEC chromatograms of a single sample subjected to the entire albumin/globulin through F3 sequential extraction method.

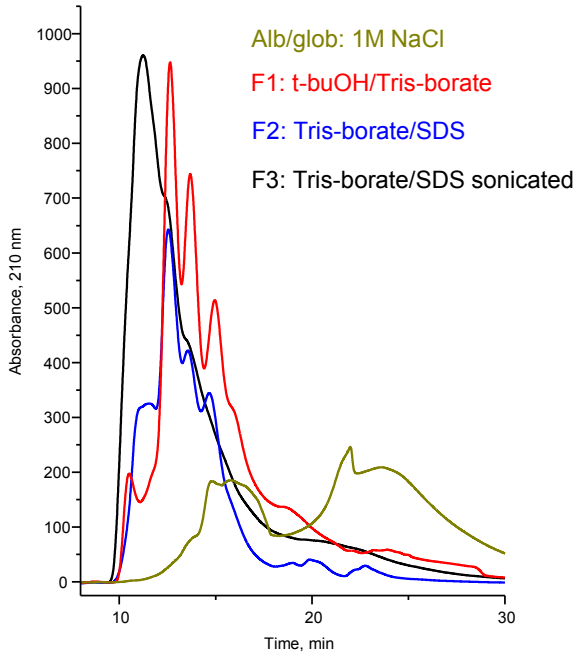
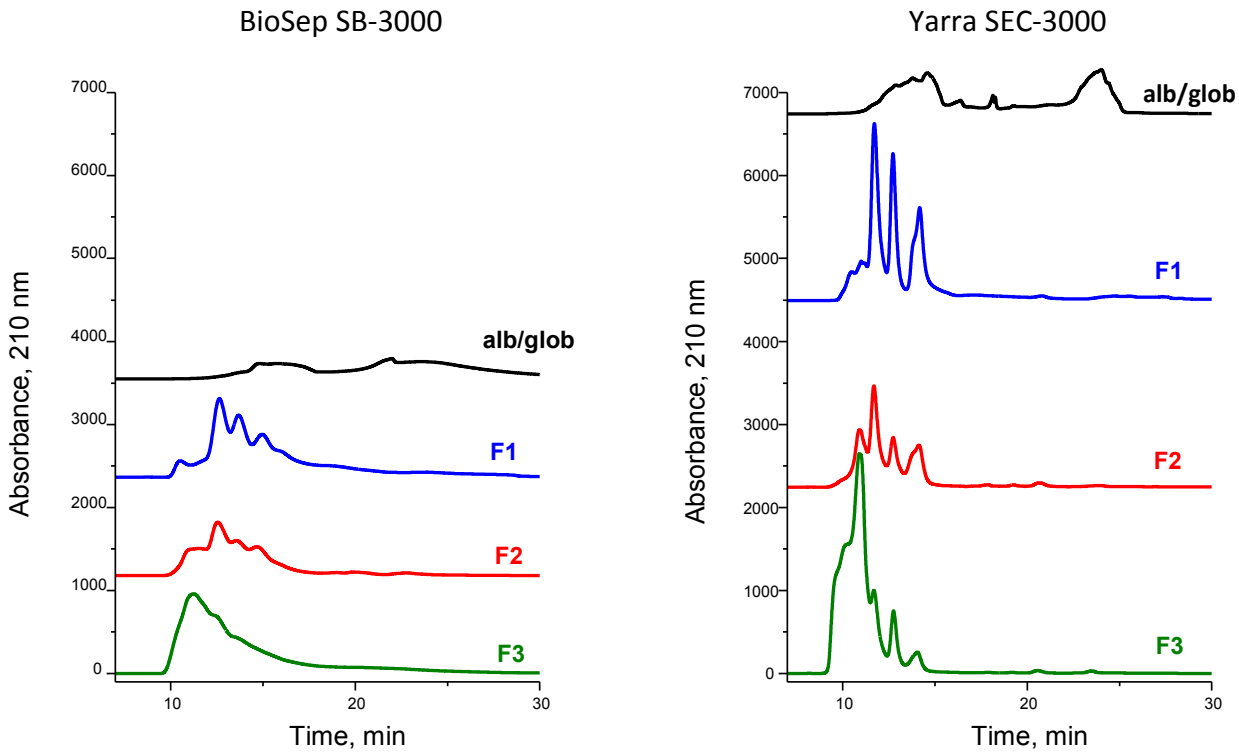


Figure 3.15 Comparison of SEC separations on two different silica based columns (extracts from same sample on both columns).



Chapter 4 - Polymeric Sorghum Proteins and Compositional Relationships to Protein Digestibility

Abstract

The goal of this research was the application of a new extraction method designed to optimize recovery of intact polymeric proteins, thereby allowing determination of the specific polymers involved with and influencing in-vitro pepsin digestibility (IVPD). From a highly diverse 337 member non-tannin association mapping panel, a 27 sample subset was randomly selected to span a range of IVPD values. Subset IVPD values averaged 59.88%, and ranged from 44.32% to 79.27%. Several additional grain traits investigated for potential impact on IVPD included total protein, total phenolic content, phytic acid content, kernel hardness index, kernel weight, and kernel diameter, but none were found to significantly correlate to IVPD. An extraction method designed to obtain three fractions containing intact protein polymers was applied to the subset. Acquired fractions were analyzed by size exclusion chromatography (SEC) in reduced and unreduced form. Molecular weights (M_w) of SEC separated protein peaks were compared to M_w standards, and ranged from <14.3kD to >669kD. Total fraction SEC areas, as well as individual SEC peaks within each fraction representing different proteins were compared to IVPD values. Total fraction areas for fraction 1 and 2 (F1, F2) did not correlate to IVPD. Protein polymers represented by individual SEC fraction peaks from F1 and F2 (with one exception) were also not significantly correlated to IVPD. However, total fraction areas of extracted fraction 3 (F3) were significantly correlated to IVPD ($r = -0.605$ at $p \leq 0.05$). Additionally, four of five observed individual fraction SEC peaks from separations of F3 were significantly correlated to IVPD ($r = -0.465$ to -0.575 at $p \leq 0.05$). This work expands on previous research by focusing on polymeric proteins within sorghum grain found to significantly influence digestibility.

Introduction

The nutritional importance of sorghum to a large proportion of the world's population is well established (Taylor and Belton, 2002; Dendy, 1995). Sorghum is depended upon as an essential protein source by millions of people in the semi-arid regions of the world (FAO 1995). The unique biochemical and structural characteristics of sorghum proteins tend to result in substantially lower digestibility, particularly after wet heat treatment (cooking), compared to proteins found in other cereal grains (Duodu et al. 2003; El Nour et al. 1998; Emmambux and Taylor, 2009; Hamaker et al. 1987; Nunes et al. 2004; Nunes et al. 2005; Maclean et al. 1983; Oria et al. 1995; Oria et al. 2000). Increasing pressure to meet human nutritional needs in developing areas of the world continues to provide impetus for research targeted at improving the nutritional contribution of sorghum for food use. In addition, the importance of grain sorghum protein characteristics in animal nutrition (Selle et al. 2010) and their influence in bio-fuels production (Zhao et al. 2008) provide additional motivation for improvement of our understanding of sorghum protein structure with the goal of optimizing intended use and functionality.

The variability found in the *in-vitro* protein digestibility (IVPD) of wild-type condensed tannin-free sorghum is broad, and reports ranging from 40% to 93% for uncooked sorghum flour are representative (Arbab and El Tinay, 1997; Axtell et al. 1981). The factors responsible for such great variability in wild-type sorghums are not well understood. This is in contrast to the high digestibility sorghum mutants, in which the causes for increased protein digestibility have been thoroughly investigated and determined from the protein body microstructure down to the amino acid and gene sequence level (Oria et al. 2000; Holding, 2014).

Numerous factors have been proposed to explain the lower digestibility of sorghum proteins relative to other cereal grains. Duodu et al. (2003) reported the possible grain components contributing to poor protein digestibility were separated into two categories representing exogenous (non-protein/protein interactions) and endogenous (protein only) factors. Examples of the former include phytates, starch, and polyphenol content. Endogenous factors may consist of variations in protein structure and crosslinking, hydrophobic character, and secondary structure changes.

Duodu et al. (2003) concluded by speculating that protein crosslinking, particularly that involving γ- and β-kafirin proteins located on the periphery of the protein body, exerted the most influence on sorghum protein digestibility. Previous research by El Nour et al. (1998) and Hicks et al. (2001) reinforce this conclusion, and emphasized the importance of kafirin packaging in relation to digestibility. Attempting to expand on these concepts, Ioerger et al. (2007) applied a sequential differential solubility method to separated hard and soft sorghum endosperm in an attempt to determine the role of crosslinking of sorghum proteins into larger polymeric groups. They found positive correlations between the γ-kafirin subclass and the most difficult to extract fraction, and speculated on the relationship to protein polymer M_w profile. Their method, however, required reducing conditions for determination, thereby precluding further polymer characterization.

The goal of this study was the application of an extraction method designed to maintain protein polymer structure (i.e. non-reducing), in order to discover the protein polymer content involved with and influencing IVPD in wild type (non-mutant) sorghum.

Materials and Methods

Sorghum samples

Sorghum samples used in this study were non-tannin varieties from a sorghum association mapping population (diversity panel) grown at the Kansas State University agronomy farm. The full panel consisted of 337 sorghum varieties representing all major cultivated races, and included tropical lines from extensively varied geographic and climactic regions as well as selected U.S. breeding lines (Casa et al. 2008). These included varieties from the Americas, Asia, and all of Africa. After removing tannin positive samples from the full panel, a 27 sample subset was randomly selected to span a range of *in-vitro* pepsin digestibility (IVPD) values. Grain traits of the subset were scrutinized to ensure they represented the larger diversity panel. Several grain traits with potential impact on digestibility were determined including total protein, total phenolic content, phytic acid content, Single Kernel Characterization System (SKCS) hardness index, kernel weight, and kernel diameter.

Grain preparation

Sorghum kernels were milled to flour using a UDY mill (UDY Corp., Fort Collins, CO) equipped with a 0.5mm screen. The milled samples were stored in sealed containers at -20˚ C until needed.

In-vitro pepsin digestibility (IVPD) assay

IVPD assays were performed using a modification of the method of Mertz et al. (1984). For this assay, duplicate 200mg portions of milled sorghum flour were mixed with 35mL of pepsin solution (1.5mg pepsin / mL 0.1M KH_2PO_4 buffer pH2.0). This mixture was incubated for 2h at 37˚ C with continuous shaking. Digestion was stopped by a 2mL addition of 2M NaOH with brief mixing. Decant following centrifugation at 4˚ C for 15min, and wash the remaining

solid material twice with 0.1M KH_2PO_4 buffer pH2.0. The final residue was frozen and lyophilized in the original 50mL tube. Prior to further analyses, dried residues were thoroughly mixed to break up the pellet.

Protein content determination

Total protein content of milled and dried undigested and digested samples was measured by nitrogen combustion using a Leco FP-528 nitrogen determinator (St. Joseph, MI) according to AACC method 46-30 Crude Protein-Combustion Method (AACC International, St. Paul, MN). A factor of 6.25 was used for conversion of percent nitrogen to crude protein percent.

Tannin content

Tannin content of the sorghum diversity panel samples was determined using the vanillin-HCl assay as described in Price et al. (1978). Samples were analyzed in duplicate and reported as positive or negative for tannins.

Total phenolic content

Total phenolics concentration of the sorghum diversity panel samples was determined using a high throughput micro-plate assay as described in Herald et al. (2012) (results reported as mg gallic acid equivalent (GAE) g-1 dried sample). Samples were analyzed in duplicate.

Phytic acid content

Phytic acid was determined colorimetrically using a Megazyme Phytic Acid (Total Phosphorous) Kit according to instructions (Megazyme International, Co. Wicklow, Ireland).

SKCS

Sorghum kernel attributes including kernel hardness index (HI), kernel diameter, and kernel weight were measured using a Perten Single Kernel Characterization System 4100 (SKCS) (Perten Instruments North America, Springfield, IL) as described in Bean et al. (2006).

Extractions (SP, IP, RP method)

A multistep sequential extraction procedure was applied in duplicate to the selected subset samples (n=27), as described in Ioerger et al. (2007). Three fractions designated soluble protein (SP), insoluble protein (IP), and residue protein (RP) are obtained using the procedure. The RP fraction is acquired using a reducing agent.

Extractions (3F method)

A salt soluble protein fraction (F0) and three additional protein fractions (subsequently referred to as F1, F2, and F3) were obtained by application of a sequential extraction scheme to duplicate 100mg milled sample portions in 2.0mL μ-tubes (see dissertation Chapter 3, Figure 3.1). Albumins and globulins (F0) were extracted first, using two 5min extractions with 1mL of 1M NaCl each extraction (room temperature vortex at medium speed followed by centrifugation at 9300g for 4min) (Taylor et al. 1984). The F0 fraction was retained or discarded based on the objective of the analysis. This was followed by a 5min wash step with 1mL of deionized H2O, mixed and centrifuged as before, and discarded. Next, Fraction 1 (F1) was obtained using two sequential 50min extractions with 1mL portions of 60% t-butanol (v/v) / 80mM Tris-borate pH10.0, mixed and centrifuged as before, and the supernatants were combined 1:1 in a clean μ-tube. To obtain Fraction 2 (F2), the same pellet was subjected to two sequential 50min extractions with 1mL of 50mM Tris-borate pH10.0 / 2% SDS (w/v), mixed and centrifuged as before, and the supernatants were combined 1:1 in a clean μ-tube. Sonication was applied to the

same pellet from the previous extractions to obtain Fraction 3 (F3) using a single 1mL portion of 50mM Tris-borate pH10.0 / 2% SDS (w/v). A 3mm x 80mm diameter sonication probe powered by a Sonic Dismembrator 60 unit (Thermo Fisher Scientific, Waltham, MA) was positioned within the tube contents, and centered approximately 4-7mm from the bottom of the µ-tube that was immersed in an ice water bath to reduce solution heating. Three 15s bursts of sonic energy were applied while manually maintaining 10W of output power (sonicator readout), and allowing 30s rests between and after the final burst. After sonication, the mixture was centrifuged as before and the supernatant was transferred to a clean µ-tube.

After each fraction was extracted, the extracted supernatants were subjected to heat deactivation of intrinsic proteases by application of 2min at 80˚ C heat during immersion of the capped µ-tubes in the water-filled wells of a drilled heat block (Pierce, Rockford, IL). After being allowed to cool to room temperature, extracts were filtered through 25mm 0.45µm GHP membrane syringe filters (PALL Life Sciences, Port Washington, NY) into clean µ-tubes or HPLC vials. Fraction extracts destined for size exclusion chromatography (SEC) analysis were analyzed fresh (not lyophilized). Aliquots of fraction extracts destined for RP-HPLC, were lyophilized and then resuspended according to the procedure described in the "RP-HPLC analysis and conditions" section below.

Total kafirin extraction

Total kafirin extracts destined for reverse phase high performance liquid chromatography (RP-HPLC) analysis were obtained by application of a 60% t-butanol (v/v) / 0.5% sodium acetate (v/v) / 2% beta-mercaptoethanol (BME) extraction protocol to duplicate 100mg milled samples as described in Bean et al. (2011).

Size exclusion chromatography (SEC) analysis and conditions

F0, F1 extracts: Non-reduced supernatant was injected as is. Reduced F1 extract was obtained by adding BME at 2% (v/v) of supernatant volume and mixed by vortex for 10min. The reduced supernatant was alkylated by addition of 4-vinylpyridine (4-VP) at the rate of 6.65µL 4-VP / 100µL supernatant and mixed an additional 10min.

F2, F3 extracts: Non-reduced supernatant was injected as is. Reduced F2 and F3 extracts were first diluted 1:1 with a 7M Urea / 2M Thiourea solution and mixed for 10min. The diluted extract was reduced by adding BME at 2% (v/v) of supernatant volume and mixed for an additional 10min. The reduced F2 or F3 extract was alkylated by addition of 4-VP at the rate of 6.65µL 4-VP / 100µL extract and mixed for a final 10min.

SEC analyses were carried out on an Agilent 1100 series HPLC system (Agilent, Santa Clara, CA). Protein separations were performed from 50uL injections on silica based Yarra SEC-3000 columns, 300mm x 7.80mm (Phenomenex USA, Torrance, CA). Isocratic runs were accomplished with a mobile phase consisting of 50mM Tris-borate pH7.0 / 1% SDS (w/v) and a flowrate of 0.5mL/min while maintaining a column temperature of 40˚ C. A UV wavelength of 210nm was monitored for analyte detection.

RP-HPLC analysis and conditions

Lyophilized total kafirin extracts were resuspended prior to RP-HPLC analysis by separate additions of 50mM Tris-borate pH10.0 / 2% SDS and 7M Urea / 2M Thiourea at a 1:1 ratio to achieve the original pre-lyophilized volume. A 10min mixing time was applied after addition of each solvent. To the resolubilized extract, BME was added at 2% (v/v) of extract volume and mixed for an additional 10min. Then extract was alkylated by addition of 4-VP at the rate of 6.65µL 4-VP / 100µL extract and mixed for a final 10min.

Extracted fractions F1, F2 and F3 destined for RP-HPLC were frozen, lyophilized, and subsequently re-suspended prior to RP-HPLC analysis. Dried F1 was resuspended similarly to the total kafirin extracts described above. Dried F2 and F3 extracts were resuspended using the same procedure applied to the total kafirin extracts, with the exception that 50mM Tris-borate pH10.0 with no SDS was used since the lyophilized pellet already contained SDS.

RP-HPLC analyses were carried out on an Agilent 1100 series HPLC system (Agilent, Santa Clara, CA). Chromatography conditions were as described by Bean et al. (2011). Briefly, protein separations were performed from 5uL injections on silica based Poroshell C18, 75mm x 2.1mm (Agilent, Santa Clara, CA). The flow-rate was 0.7mL/min and mobile phase for gradient runs consisted of mobile phase A: deionized water / 0.1% TFA (v/v), and mobile phase B: acetonitrile / 0.07% TFA (v/v). Gradient times were 20 - 40% B from 0 to 5min; 40% - 60% B from 5min to 15min; 60% to 20% B from 15min to 17min; with a 5min post-time at 20% B. The column was maintained at 55 ˚C and UV detection was used at 214nm.

Statistics

All data were plotted, and statistical analyses were conducted using Microsoft Excel (Redmond, WA) and Origin (OriginLab Corp., Northampton, MA). Significant correlations were determined using the Pearson product-moment correlation coefficient and n-2 degrees of freedom where n = sample number. A significance level of $p \leq 0.05$ was used throughout unless noted otherwise in the text.

Results and Discussion

Subset selection (tannin, total protein, and IVPD)

Sorghum samples within the full panel were assayed for condensed tannin content. Only non-tannin samples were considered for in-depth study due to well recognized tannin/protein interactions that would interfere with study results (Taylor et al. 2007, Dykes and Rooney 2006).

Total protein content within the full panel and subset is presented in **Table 4.1**. Full panel sample protein contents averaged 14.42% as is, with a range of 10.50 to 19.94%. The subset sample protein contents averaged 13.85% as is, with a range of 11.41 to 15.85%. Samples selected for the subset were purposely limited to those containing not greater than 16% protein to help reflect more typical protein levels (FAO, 1995; National Research Council, 1996; Deosthale et al. 1970). There was no significant correlation between total protein content and IVPD within the full panel or the selected subset (**Figure 4.1**).

A sample subset from the full non-tannin sorghum panel was selected to span a wide range of protein digestibility. An important criteria for subset selection was for total protein and IVPD values to represent those seen in the larger panel. IVPD values within the full panel and subset are presented in **Table 4.1**. Full panel IVPD values averaged 57.89%, and ranged from 35.48 to 83.83%. Subset IVPD values averaged 59.88%, and ranged from 44.32% to 79.27%. An overlay comparison of total protein and IVPD values for the full panel and subset is presented in **Figure 4.2**, and graphically illustrates how the selected subset is representative of the IVPD values obtained from the larger panel.

Other digestibility factors

Numerous factors within sorghum grain have been hypothesized to have possible roles in protein digestibility (Duodu et al. 2003; Dykes and Rooney, 2006; Selle et al. 2010). Total

protein content and tannin content were discussed above. Four additional physicochemical grain traits with potential for impacting protein digestibility were also assayed within the full panel. These included total phenolic content, kernel hardness index (HI), kernel weight, and kernel diameter. **Table 4.1** provides quantitative comparisons of these group trait values obtained for the full panel and selected subset.

Although no direct links between non-tannin phenolic content (flavonoids and phenolic acids) and decreased protein digestibility in sorghum have been documented, indirect mechanisms for such an effect have been theorized (Damodaran, 1996; Duodu et al. 2003). The possibility exists for oxidation of plant polyphenols to quinones, leading to formation of highly reactive peroxides. Subsequent oxidation of amino acid residues was postulated to result in formation of new protein polymers having a negative effect on protein digestion. To account for this possibility, the non-tannin phenolic content of the full diversity panel was measured. A comparison of the averages and ranges for total phenolic content of the full panel and selected subset are presented in **Table 4.1**. No significant correlation between phenolics content and IVPD was found within the full panel, and this was reflected in the selected subset as well (**Figure 4.3**).

Physical kernel characteristics of the sorghum samples comprising the full diversity panel, including kernel HI, weight, and diameter were determined by SKCS. The initial applicability of the SKCS instrument for use with sorghum grain was demonstrated by Pederson et al. (1996), and optimized by Bean et al. (2006). Statistical comparison of the values obtained for these kernel traits in the current study are presented in **Table 4.1**. Correlations between IVPD and kernel HI ($r = -0.077$ and -0.073), kernel weight ($r = 0.014$ and -0.150), and kernel diameter ($r = 0.028$ and 0.085) were determined for the full panel and selected subset

respectively. There were no significant correlations noted between IVPD and any of the physical kernel traits measured.

Phytic acid levels were determined within the selected subset to see if a significant correlation to digestibility was apparent. In a review discussing exogenous and endogenous factors and their potential effects on sorghum protein digestibility, Duodu et al. (2003) concluded the possibility of phytic acid inhibiting sorghum protein digestibility could not be ruled out. Assay results for the phytic acid content of the subset revealed an average value of 0.86g/100g, and a minimum to maximum range of 0.62g to 1.20g/100g. The IVPD of raw sorghum flour was not significantly correlated to phytic acid content ($r = - 0.247$), thus eliminating phytic acid as a significant contributing factor to sample IVPD variation within this subset.

IVPD correlations to extracts using older method

In a previous study, Ioerger et al. (2007), utilized extraction techniques adapted from earlier studies of large gluten polymers from wheat. In that study, three solubility based fractions were obtained from isolated hard and soft sorghum endosperm. Those three fractions (soluble proteins (SP), insoluble proteins (IP), residue proteins (RP)) proved useful for measuring the protein distribution of each endosperm type, as well as providing a metric for comparisons of how the polymeric protein content varied between endosperm types. The same method was applied to samples in the current study to see how SP, IP, and RP fractions might correlate to IVPD differences within the selected subset. From results based on total fraction SEC peak area, the only fraction to significantly correlate with IVPD was the RP fraction ($r = - 0.664$). This was the most difficult to extract fraction, requiring the most rigorous conditions (followed a sonication step, and required reducing agent).

That the RP fraction correlates with IVPD is significant because it provides a measurable indication of the importance of sorghum protein polymers to observed digestibility differences. Granted, evidence for the relationship between sorghum protein crosslinking and digestibility is not new, and is illustrated in a pair of studies by Hamaker et al. (1986, 1994). Landry-Moureaux fractionation of sorghum and maize samples was carried out to quantitate relative amounts of cross-linked and noncross-linked proteins. Both studies confirmed samples with less cross-linked kafirin exhibited higher IVPD. The goal of the current research, however, was not simply to confirm a known link between degree of protein crosslinking and digestibility. Rather, it was to determine what compositional differences within the protein polymer complement of given non-mutant samples result in significant effects on digestibility. The extracted RP fraction, being reduced, does not retain the structural information required for resolving differences in polymer structure, and consequently, how these differences might relate to variations in protein digestibility. A non-reducing method for obtaining protein polymers influencing sorghum protein digestibility, and in an intact form insofar as possible, was needed.

Determining the utility of the extraction method

The 3F extraction method described earlier was applied to the selected subset samples. The fractions so obtained were analyzed by SEC in unreduced and post-extraction reduced form. In **Figure 4.4**, examples of SEC chromatograms typical for each fraction are presented. Total peak areas for each fraction (F1, F2, F3) were determined, as well as the areas for individual SEC peaks within each fraction. These values were used to calculate correlations to percent total protein content and to IVPD values.

Significant positive correlations between the total fractions obtained using both the older method (SP, IP, RP) and the newer 3F method would provide good indication they are measuring

similar protein fractions. A correlative comparison of all three total fractions from each method (i.e. SP vs F1, IP vs F2, and RP vs F3) reveals significant relationships to their respective counterparts (**Figure 4.5**), with the correlation between RP and F3 being especially strong (r = 0.865). Also recall that the total extracted RP fraction was significantly negatively correlated with IVPD. If RP and F3 represent similar protein fractions, a similar negative correlation with IVPD could be expected for total F3. This was indeed the case, as total F3 displayed a strong negative correlation to IVPD (r = -0.605). It was concluded the 3F extraction method was useful for the investigation and characterization of extracted non-reduced sorghum polymers, and helpful in the determination of new details in protein polymer relationships to IVPD variation among different sorghum cultivars.

Total protein content correlations to Total fractions and Individual fraction SEC peaks

Significant correlations of total fraction areas and individual peak areas within each fraction to percent total protein content of specific samples were noted. Fraction F2 (r = 0.478) and fraction F3 (r = 0.510) total fractions showed positive correlation to percent total protein content (**Table 4.2**). F0 and F1 on the other hand, were not correlated to percent total protein content. This means within the subset, samples with higher total protein content tended to have greater total fraction F2 and F3 protein levels as well. The SEC elution profiles of fractions F2 and F3 indicate larger molecular weight (M_w) distributions than for fractions F0 or F1, implying fractions F2 and F3 contain relatively more polymeric proteins.

No individual SEC peaks from the F1 fraction were correlated to total protein content. However, two individual peaks from the F2 fraction were significantly positively correlated to total protein content (peaks 3b and 4) (r = 0.457 and r = 0.412 respectively). Similarly, three individual peaks from the F3 fraction were significantly positively correlated with protein

content (peaks 3, 4 and 5) (r = 0.563, 0.553, and 0.445 respectively). Peak locations within respective fractions are shown in **Figure 4.4**. Based on SEC patterns from extracts that were reduced post-extraction, the peaks significantly correlating to total protein content represent predominantly non-crosslinked proteins. This implies as total protein content increases within sorghum samples, the mostly non-crosslinked proteins within F2 and F3 are increasing in a measurably significant way as well. The importance of this relationship is not clear at present, but could indicate the level of protein polymerization within the endosperm slows at some point as protein content increases.

IVPD correlations to Total fractions and Individual fraction SEC peaks

With regard to total fractions, the strong negative correlation of total fraction F3 with IVPD (r = -0.605) has already been discussed. Of the other total fractions within this subset, only F0 was significantly correlated to IVPD (r = -0.543)(see **Table 4.3**). That fraction F0 is negatively correlated to protein digestibility is puzzling, since the water and salt extracts of sorghum are well documented to consist of albumin and globulin proteins (Sastry and Virupaksha, 1967; Taylor et al. 1984), of which both are very digestible.

Determining a significant relationship between individual fraction SEC peaks and IVPD would be invaluable for assisting in the identification of specific protein polymers that impact sorghum protein digestibility. For the locations of individual SEC peaks for each fraction, refer to **Figure 4.4**. **Table 4.3** shows the breakdown of the individual SEC peaks associated with each fraction and displays respective individual peak correlations to IVPD. No significant correlations to IVPD were noted among the five individual peaks from F1. The peak eluting first from fraction F2 (P1) was very weakly correlated to IVPD (r = -0.381). Displaying the shortest SEC retention time, this peak represents the largest polymeric (reducible, see Figure 6) protein(s)

extracted within fraction F2. By far, fraction F3 contained the most peaks (four out of five) displaying significant correlations to IVPD. Significant correlation values of individual peaks from the F3 fraction ranged from r = -0.465 to r = -0.575. Based on SEC analyses of post-extraction reduced F3 extracts, peaks 1, 2 and 3 were polymeric (reducible), and peak 4 was not. The results of the individual fraction peaks provide clear indication that the extracted proteins present within fraction F3 exert the greatest influence on the IVPD of the samples in this subset. Also notable, four of the five peaks significantly correlated to IVPD represent proteins capable of being reduced with a reducing agent (BME), and are therefore polymeric in nature.

Influence of covalent polymerization on IVPD

Characteristics of the polymeric proteins correlated with IVPD were further probed using comparison of data from SEC analysis of extracted fractions after post-extraction breaking of covalent protein linkages with reducing agent. **Figure 4.6** provides examples of SEC chromatograms of the three extracted protein fractions before and after reduction. Proteins capable of being reduced displayed reduced peak areas following exposure to BME. Quantitation of these areas before and after reduction allowed determination of the relationship of covalent bonding within protein polymers to protein content and IVPD.

The covalently bonded polymeric protein content of the subset samples was not statistically related to the total protein content of the samples. No significant correlations of total protein content to the non-reduced or reduced polymeric SEC peak areas of the extracted fractions (F1, F2, F3) were noted.

Integrated areas limited to the peak areas representing polymers were examined before and after reduction to see if relationships to IVPD were apparent. IVPD was significantly correlated to the non-reduced polymeric area of fraction F2 (r = - 0.461). In contrast, the reduced

polymeric area for fraction F2 was not significantly correlated. A possibility for this observation is that the structures of the fraction F2 polymeric proteins are more influenced by non-covalent bonding forces not quantifiably affected by reduction analysis. Recall from previous discussion that only one of the fraction F2 peaks (P1) was related to IVPD, and the correlation was borderline significant at $p \leq 0.05$. However, a more pronounced relationship was seen in fraction F3 where IVPD was significantly correlated to both the non-reduced ($r = -0.560$) as well as the reduced ($r = -0.527$) polymeric peak areas. These results mesh logically with those from the individual peak data previously discussed, where nearly all the significantly correlated fraction F3 protein peaks (4 out of 5) were polymeric in nature, and serves to reinforce the important role protein polymer structure appears to play in sorghum protein digestibility.

In an attempt to determine the effect of "reducibility" of the protein polymers on digestibility, SEC peak area differences between the reduced and non-reduced polymeric peaks within a given sample for each extracted fraction (F1, F2, F3) were calculated. The differences were plotted against IVPD to see if relationships were evident (**Figure 4.7**). The calculated differences between reduced and non-reduced protein polymer contents within extracted fractions F2 ($r = -0.442$) and F3 ($r = -0.540$) were significantly correlated to IVPD. Because the correlations were negative, it suggests those samples with less covalent bonding within their complement of polymeric proteins are more digestible, and is in agreement with results from previous research.

Total protein and IVPD correlations to kafirin content

The total kafirin extraction and RP-HPLC extract preparation methods described in the Material and Methods section were applied to the 27 sample subset. Examples of RP-HPLC chromatograms from that analysis are shown in **Figure 4.8**. Although total protein content of

subset samples was found to significantly correlate to total kafirin content (r = 0.622), there was no significant correlation between total kafirin content and IVPD. Conversely, total protein content of subset samples was not correlated to total γ-kafirin content, and yet a significant correlation was found between total γ-kafirin content and IVPD (r = -0.559). Stated another way, subset samples with higher protein levels tend to have higher kafirin levels as well, but do not necessarily also have increased γ-kafirin amounts. Results from RP-HPLC data also indicated no significant correlation between total kafirin content and total γ-kafirin content (r = 0.062).

Conclusions

A non-tannin sample subset representing a highly diversified sorghum panel with a wide range of IVPD was selected for examination of inherent factors potentially correlating to IVPD. Examined factors included extracted protein fractions, total protein content, total phenolic content, phytate content, kernel hardness index, weight, and diameter. Of these factors, only specific portions of sequentially extracted proteins were found to significantly correlate to IVPD.

Using a sequential extraction method designed to enhance extraction of intact protein polymers, three protein fractions (F1, F2, F3) were obtained for analysis by SEC and RP-HPLC. Fractions obtained later in the extraction sequence contained relatively larger M_w distributions. Total fractions F2 and F3 were significantly positively correlated to total protein content. Several individual SEC peaks from F2 and F3 were found to significantly correlate to total protein content, and represented proteins predominantly non-crosslinked in nature, possibly indicating a pattern in protein polymerization related to protein level.

Correlations to IVPD values showed total fractions F0 and F3 were significantly negatively correlated to IVPD. In addition, proteins represented by four of five individual SEC peaks from F3 were significantly negatively correlated to IVPD, with three of the correlated peaks being polymeric. The findings appear to place emphasis for IVPD impact on polymeric proteins present within fraction F3 in this subset.

Total protein content was correlated with total kafirin content, but not with γ-kafirin content or IVPD. However, total kafirin content was not correlated to γ-kafirin content. There was significant correlation between IVPD and γ-kafirin.

The possible factors governing sorghum protein digestibility in wild-type sorghum populations are undoubtedly complex and likely interdependent. Within the limited variables chosen for consideration in this study, only certain portions of the fractionated protein complement (predominantly polymeric in nature) appear to have a significant effect on protein digestibility. The results indicate a need for more detailed determination of the specific proteins that make up the sorghum polymers significantly affecting IVPD. Utilization of orthogonal analytical methods is currently underway to obtain additional information regarding the composition of these polymers, and should prove relevant to enhancing our understanding of sorghum protein digestibility.

References

Arbab, M.E., and El Tinay, A.H. 1997. Effect of cooking and treatment with sodium bisulphite or ascorbic acid on the in vitro protein digestibility of two sorghum cultivars Food Chemistry, 59:339–343.

Axtell, J.D., Kirleis, A., Hassen, M.M., D'Croz-Mason, N., Mertz, E.T., and Munck, L. 1981. Digestibility of sorghum proteins. Proc. National Acad. Sci. 78:1333–1335.

Bean, S.R., Chung, O.K., Tuinstra, M.R., Pedersen, J.F., and Erpelding, J. 2006. Evaluation of the single kernel characterization system (SKCS) for measurement of sorghum grain attributes. Cereal Chem. 83(1):108-113.

Bean, S.R., Ioerger, B.P., and Blackwell, D.L. 2011. Separation of kafirins on surface porous reversed-phase high-performance liquid chromatography columns. J. Agric. Food Chem. 59:85-91.

Casa, A.M., Pressoir, G., Brown, P.J., Mitchell, S.E., Rooney, W.L., Tuinstra, M.R. Franks, C.D., and Kresovich, S. 2008. Community resources and strategies for association mapping in sorghum. Crop Sci. 48:30-40.

Damadaran, S. 2006. Protein Denaturation. Pages 1-12 in: Handbook of Food Science, Technology, and Engineering, Vol.1. Y.H. Hui, ed. CRC Press: Boca Raton, FL.

Dendy, D.A.V. 1995. Sorghum and Millets: Production and Importance. Pages 11-26 in: Sorghum and Millets: Chemistry and Technology. D.A.V. Dendy, ed. American Association of Cereal Chemists: St. Paul, MN.

Deosthale, Y.G., Mohan, V.S., and Visweswara Rao, K. 1970. Varietal differences in protein, lysine, and leucine content of grain sorghum. J. Agric. Food Chem. 18(4):644-646.

Duodu, K.G., Taylor, J.R.N., Belton, P.S., and Hamaker, B.R. 2003. Factors affecting sorghum protein digestibility. J. Cereal Sci. 38:117-131.

Dykes, L., and Rooney, L.W. 2006. Sorghum and millet phenols and antioxidants. J. Cereal Sci. 44:236-251.

El Nour, I.N.A., Peruffo, A.D.B., and Curioni, A. 1998. Characterisation of sorghum kafirins in relation to their cross-linking behaviour. J. Cereal Sci. 28:197-207.

Emmambux, M.N. and Taylor, J.R.N. 2009. Properties of heat-treated sorghum and maize meal and their prolamin proteins. J. Agric. Food Chem. 57:1045-1050.

FAO. 1995. Production and Utilization. In: Sorghum and Millets in Human Nutrition. FAO Foods and Nutrition Series, No.27.

Hamaker, B.R., Kirleis, A.W., Butler, L.G., Axtell, J.D., and Mertz, E.T. 1987. Improving the in vitro protein digestibility of sorghum with reducing agents. Proc. Natl. Acad. Sci. USA 84:626-628.

Hamaker, B.R., Kirleis, A.W., Mertz, E.T., and Axtell, J.D. 1986. Effect of cooking on the protein profiles and in vitro digestibility of sorghum and maize. J. Agric. Food Chem. 34:647-649.

Hamaker, B.R., Mertz, E.T., and Axtell, J.D. 1994. Effect of extrusion on sorghum kafirin solubility. Cereal Chem. 71(5):515-517.

Herald, T.J., Gadgil, P., and Tilley, M. 2012. High-throughput micro plate assays for screening flavonoid content and DPPH-scavenging activity in sorghum bran and flour. J. Sci. Food Agric. 92:2326-2331.

Hicks, C., Bean, S.R., Lookhart, G.L., Pedersen, J.F., Kofoid, K.D., and Tuinstra, M.R. 2001. Genetic analysis of kafirins and their phenotypic correlations with feed quality traits, in vitro digestibility and seed weight in grain sorghum. Cereal Chemistry 78:412-416.

Holding, D.R. 2014. Recent advances in the study of prolamin storage protein organization and function. Frontiers in Plant Sci. doi: 10.3389/fpls.2014.00276.

Ioerger, B., Bean, S.R., Tuinstra, M.R., Pedersen, J.F., Erpelding, J., Lee, K.M., and Herrman, T.J. 2007. Characterization of polymeric proteins from vitreous and floury sorghum endosperm. J. Agric. Food Chem. 55:10232-10239.

Larroque, O.R., Gianibelli, M.C., Gomez Sanchez, M., and MacRitchie, F. 2000. Procedure for obtaining stable protein extracts of cereal flour and whole meal for size-exclusion HPLC analysis. Cereal Chem. 77(4):448-450.

Maclean, W.C., De Romana, G.L., Gastanaduy, A., and Graham, G.G. 1983. The effect of decortication and extrusion on the digestibility of sorghum by preschool children. J. Nutrit. 2071-2077.

Mertz, E.T., Hassen, M.M., Cairns-Whittern, C., Kirleis, A.W., Tu, L., and Axtell, J.D. 1984. Pepsin digestibility of proteins in sorghum and other major cereals. Proc. Natl. Acad. Sci. 81:1-2.

National Research Council. 1996. Lost Crops of Africa. Volume I: Grains. National Academy Press: Washington, D.C.

Nunes, A., Correia, I., Barros, A., and Delgadillo, I. 2004. Sequential in vitro pepsin digestion of uncooked and cooked sorghum and maize samples. J. Agric. Food Chem. 52:2052-2058.

Nunes, A., Correia, I., Barros, A., and Delgadillo, I. 2005. Characterization of kafirin and zein oligomers by preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis. J. Agric. Food Chem. 53:639-643.

Oria, M.P., Hamaker, B.R., and Shull, J.M. 1995. Resistance of sorghum α-, β-, γ-kafirins to pepsin digestion. J. Agric. Food Chem. 43:2148-2153.

Oria, M.P., Hamaker, B.R., Axtell, J.D., and Huang, C-P. 2000. A highly digestible sorghum mutant cultivar exhibits a unique folded structure of endosperm protein bodies. Proc. Natl. Acad. Sci. 97:5065-5070.

Pedersen, J.F., Martin, C.R., Felker, F.C., and Steele, J.L. 1996. Application of the single kernel wheat characterization technology to sorghum grain. Cereal Chem. 73:421-423.

Price, M.L., Van Scoyoc, S., and Butler, L.G. 1978. A critical evaluation of the vanillin reaction as an assay for tannin in sorghum grain. J. Agric. Food Chem. 26(5):1214-1218.

Sastry, L.V.S. and Virupaksha, T.K. 1967. Disc electrophoresis of sorghum seed proteins in polyacrylamide gels. Anal. Biochem. 19:505-513.

Selle, P.H., Cadogan, X.L., and Bryden, W.L. 2010. Implications of sorghum in broiler chicken nutrition. Animal Feed Sci. Tech. 156:57-74.

Taylor, J., Bean, S.R., Ioerger, B.P., and Taylor, J. 2007. Preferential binding of sorghum tannins with gamma-kafirin and the influence of tannin binding on kafirin digestibility and biodegradation. J. Cereal Sci. 46:22-31.

Taylor, J.R.N. and Belton, P.S. 2002. Sorghum. Pages 26-91 in: Psuedocereals and Less Common Cereals. P.S. Belton and J.R.N. Taylor, eds. Springer-Verlag: Berlin.

Taylor, J.R.N., Schussler, L., and van der Walt, W.H. 1984. Fractionation of proteins from low-tannin sorghum grain. J. Agric. Food Chem. 32:149-164.

Zhao, R., Bean, S.R., Ioerger, B.P., Wang, D., and Boyle, D.L. 2008. Impact of mashing on sorghum proteins and its relationship to ethanol fermentation. J. Agric. Food chem. 56:946-953.

Table 4.1 Diversity panel (n = 337) and Factors subset (n = 27) comparison (non-tannin)

	Total Protein (%)		Raw IVPD (%)		Total Phenolics (mg GAE/g)	
	n = 337	n = 27	n = 337	n = 27	n = 337	n = 27
average	14.42	13.85	57.89	59.88	2.05	1.79
max	19.94	15.85	83.83	79.27	9.70	3.23
min	10.50	11.41	35.48	44.32	0.45	0.89
SD	1.60	1.32	7.03	9.79	1.08	0.58
95% of samples	11.2 - 17.6		43.8 - 71.6		0.97 - 3.13	

	Kernel HI		Kernel Weight (mg)		Kernel Diameter (mm)	
	n = 337	n = 27	n = 337	n = 27	n = 337	n = 27
average	74.97	77.29	28.66	27.30	2.28	2.25
max	114.27	104.11	54.62	36.20	3.00	2.67
min	6.42	60.85	15.85	20.19	1.54	1.71
SD	13.89	10.37	5.57	3.94	0.29	0.25
95% of samples	47.2 - 102.7		17.5 - 39.8		1.69 - 2.87	

Table 4.2 Percent total protein correlations (r) with SEC areas of extracted total fractions and individual peaks within each fraction for a diverse sample set (n=27) of raw flours extracted using a method optimized for polymeric protein extraction (bolded correlations are significant at $p \leq 0.05$).

		Extracted Total Fractions					
Total extracted fraction areas	% total protein vs r =	F0 0.281	F1 0.173	F2 **0.478**	F3 **0.510**		
		Fraction **1** Individual SEC Peaks					
Fraction **1** individual peak areas	% total protein vs r =	P1 0.173	P2 0.077	P3 0.190	P4 0.118	P5 0.212	
		Fraction **2** Individual SEC Peaks					
Fraction **2** individual peak areas	% total protein vs r =	P1 0.122	P2 0.118	P3a 0.234	P3b **0.457**	P4 **0.412**	P5 0.259
		Fraction **3** Individual SEC Peaks					
Fraction **3** individual peak areas	% total protein vs r =	P1 0.344	P2 0.281	P3 **0.563**	P4 **0.553**	P5 **0.445**	

Table 4.3 IVPD correlations (r) with SEC areas of extracted total fractions and individual peaks within each fraction for a diverse sample set (n=27) of raw flours extracted using a method optimized for polymeric protein extraction (bolded correlations are significant at $p \leq 0.05$).

		Extracted Total Fractions					
Total extracted fraction areas	IVPD vs r =	F0 **-0.543**	F1 0.063	F2 -0.263	F3 **-0.605**		
		Fraction **1** Individual SEC Peaks					
Fraction **1** individual peak areas	IVPD vs r =	P1 0.093	P2 0.376	P3 0.198	P4 0.034	P5 0.175	
		Fraction **2** Individual SEC Peaks					
Fraction **2** individual peak areas	IVPD vs r =	P1 **-0.381**	P2 -0.224	P3a -0.341	P3b -0.084	P4 -0.164	P5 -0.045
		Fraction **3** Individual SEC Peaks					
Fraction **3** individual peak areas	IVPD vs r =	P1 **-0.575**	P2 **-0.465**	P3 **-0.521**	P4 **-0.481**	P5 -0.234	

Figure 4.1 Total protein vs IVPD Diversity Panel (r = -0.164) and Panel Subset (r = -0.380) (not significant at $p \leq 0.05$).

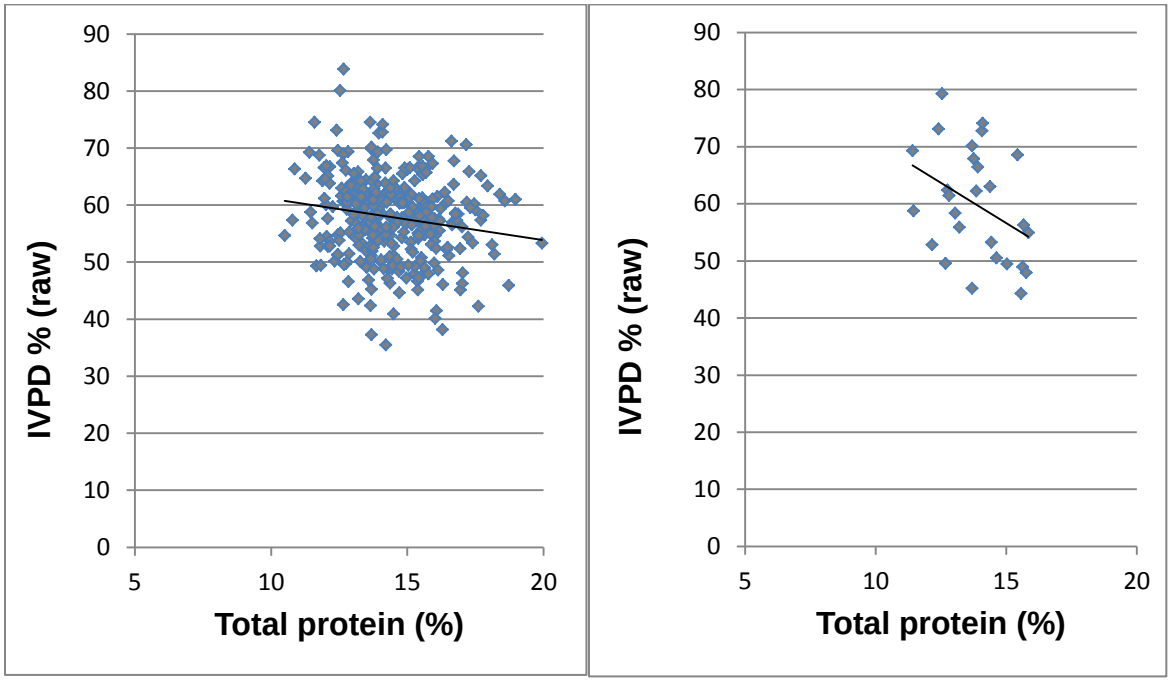


Figure 4.2 Overlay comparison of total protein vs IVPD for the Diversity Panel and Factors Subset.

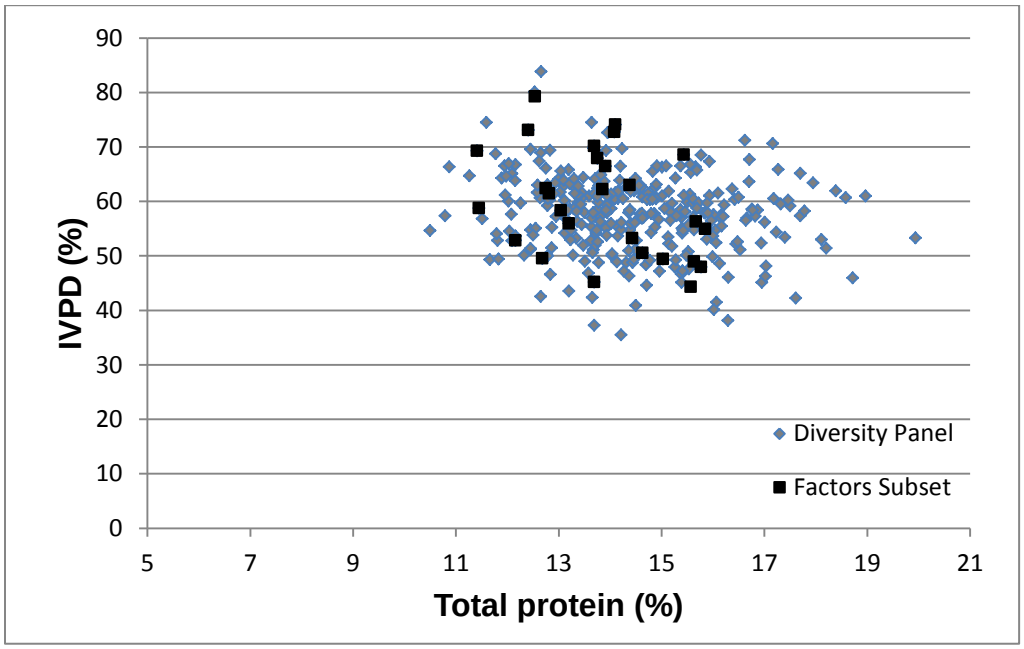


Figure 4.3 Total phenolics vs IVPD for Diversity Panel (r = -0.185) and Factors Subset (r = -0.145) (not significant at $p \leq 0.05$).

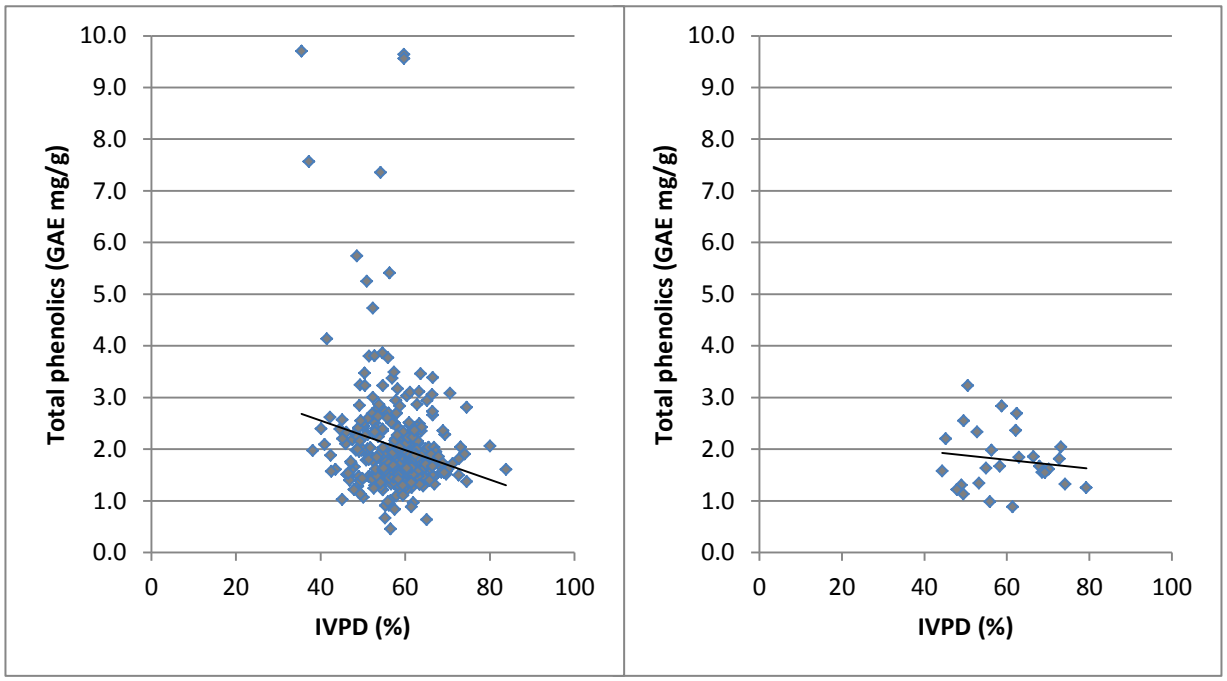


Figure 4.4 Examples of Fractions 1, 2 and 3 analyzed by SEC, as well as locations of individual peaks within each fraction.

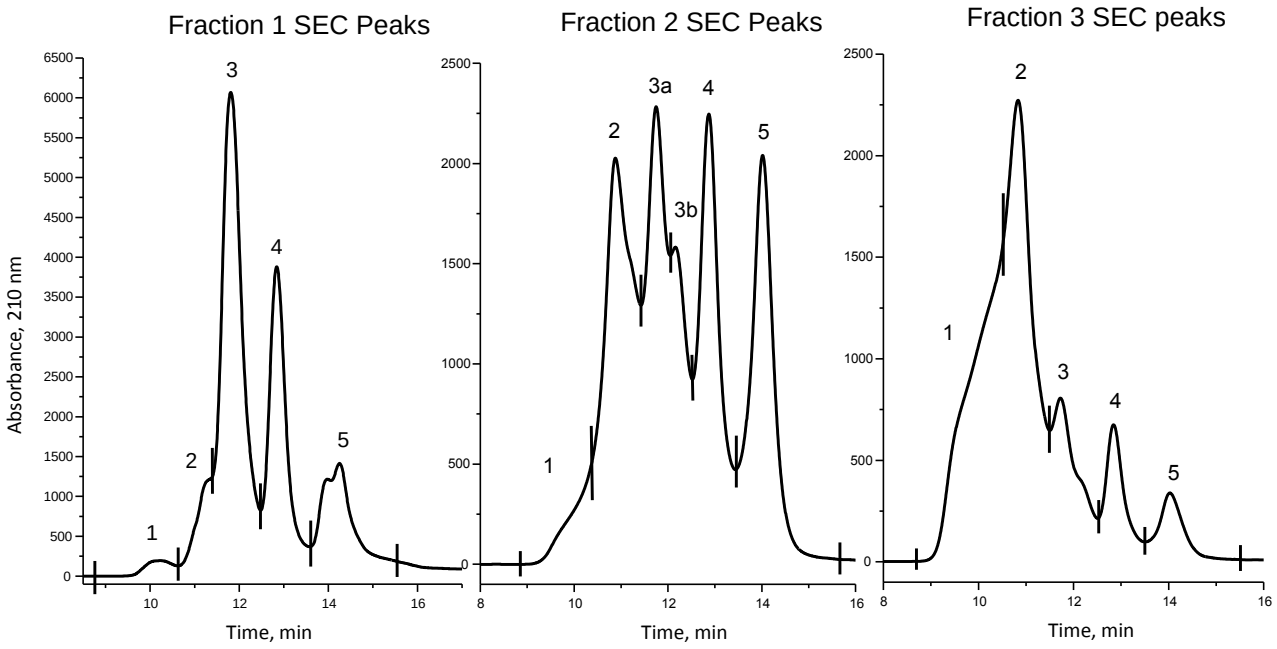
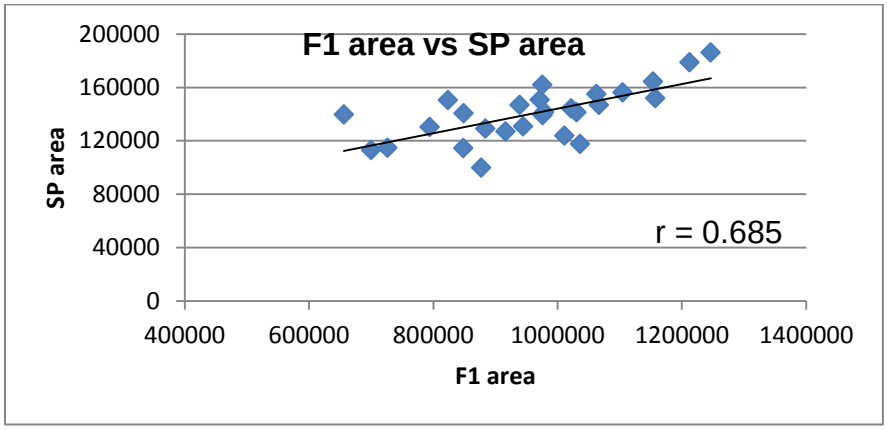
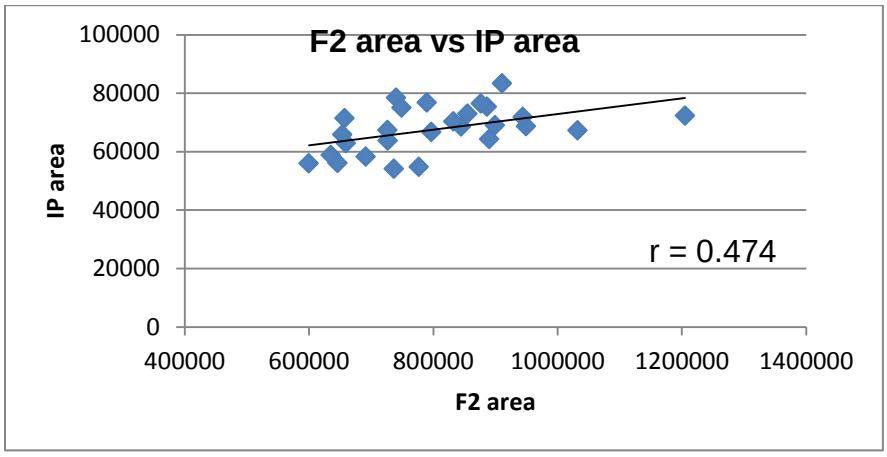


Figure 4.5 Correlations of respective total fractions from the Factors subset (n=27) obtained by application of the SP, IP, RP and the non-reducing 3F methods. (significant at $p \leq 0.05$)





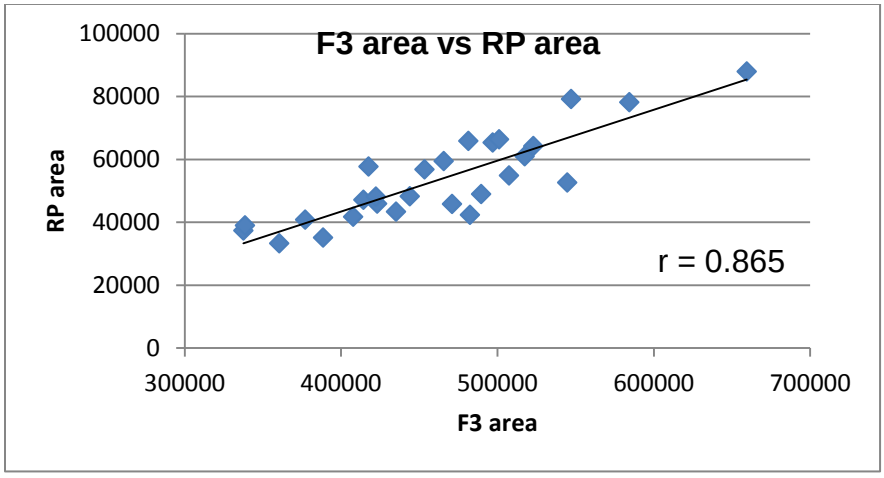


Figure 4.6 Example SEC chromatograms of extracted fractions F1, F2, and F3 non-reduced (solid line) and post-extraction reduced (dotted line).

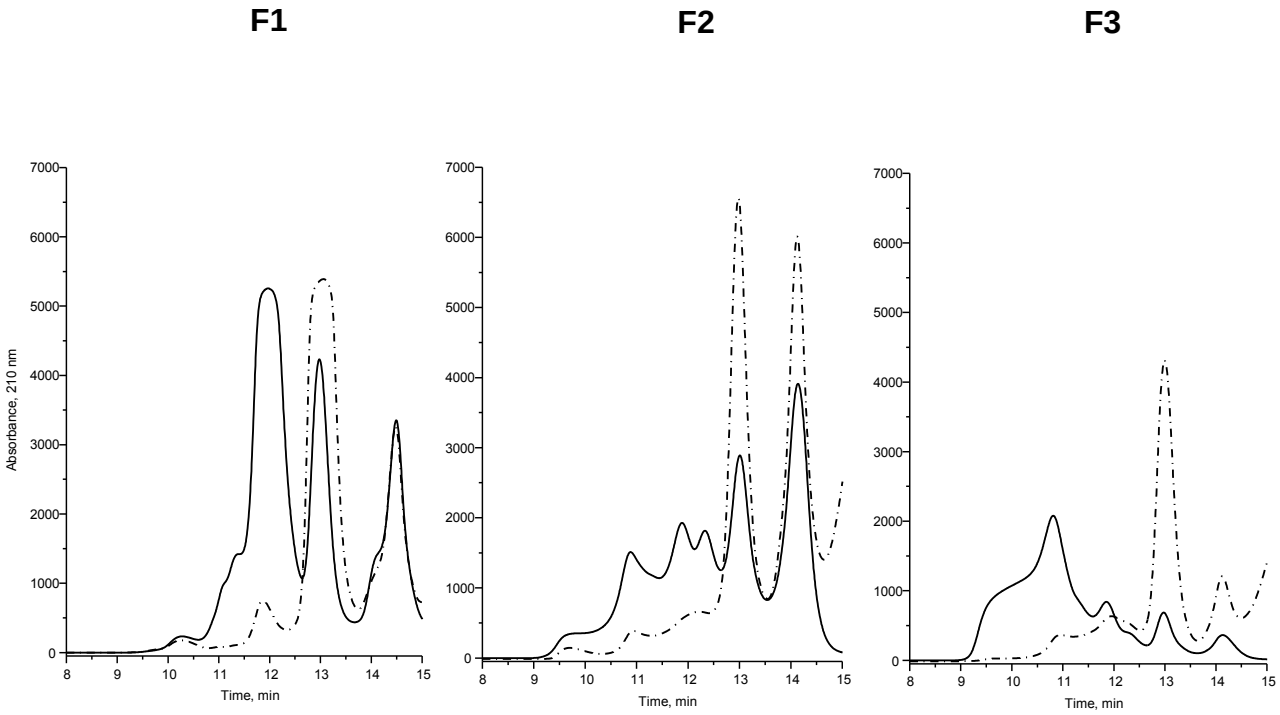
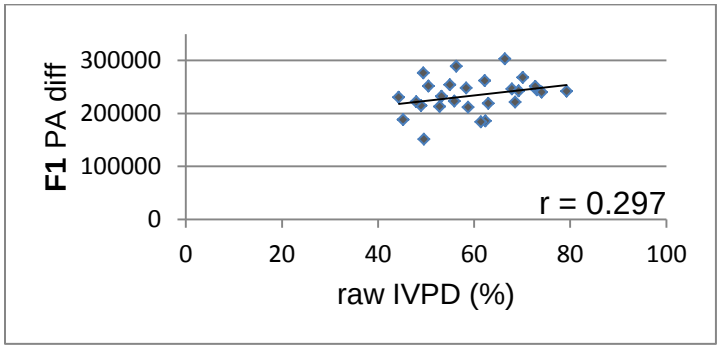
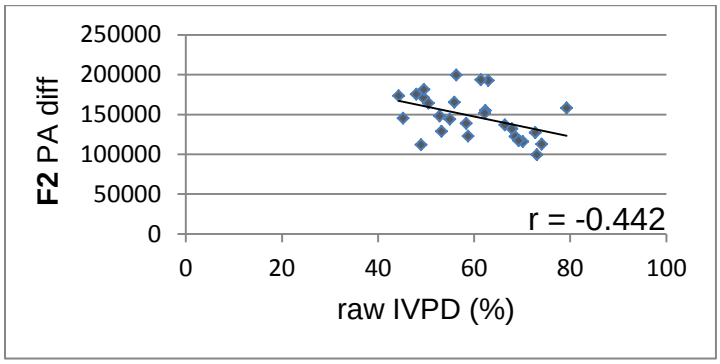


Figure 4.7 IVPD vs SEC reducible peak areas of the polymeric peaks from extracted fractions F1, F2, F3. ($r \geq 0.381$ significant at $p \leq 0.05$)





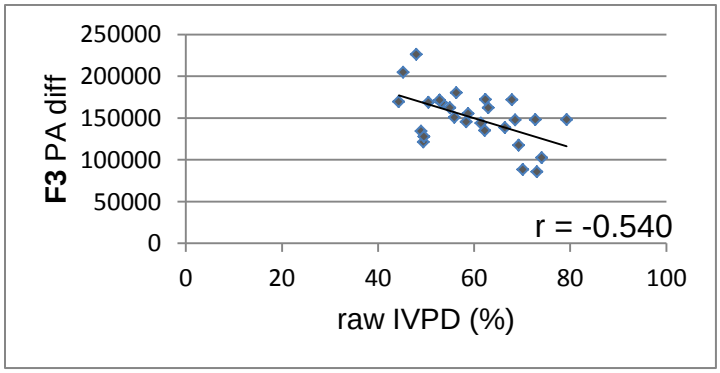
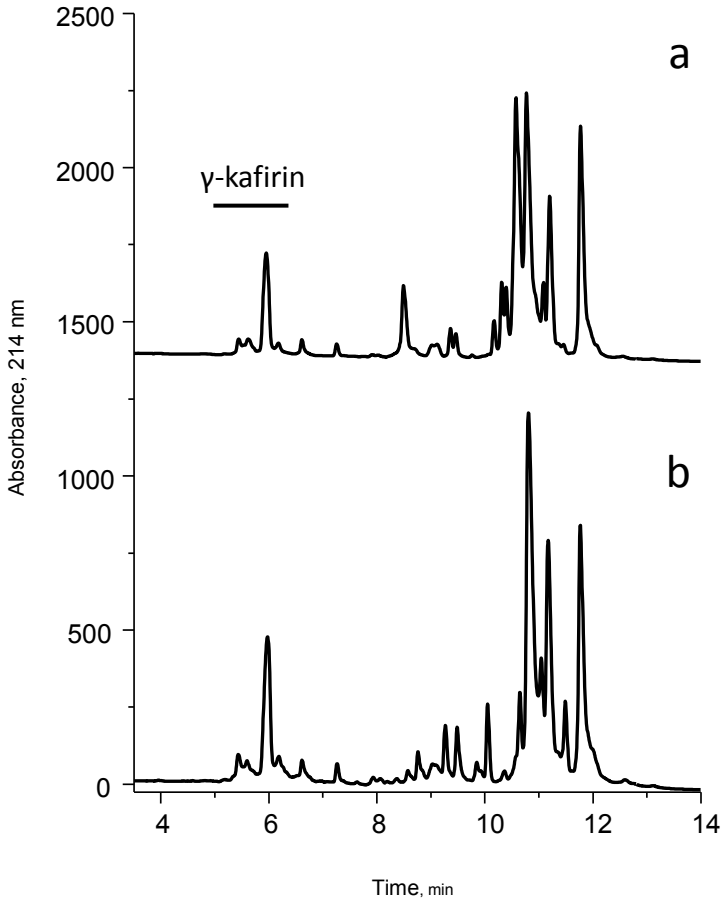


Figure 4.8 Total kafirin by RP-HPLC of high and low IVPD sorghum samples from the 27 sample subset. (IVPD values: a = 79.27%; b = 44.32%)



Chapter 5 - 2-Dimensional Orthogonal Analysis for Characterization of Polymeric Sorghum Proteins Correlated with In-vitro Protein Digestibility

Abstract

The protein compositions of polymeric proteins were characterized within a representative subset from a diverse non-tannin sorghum population exhibiting a wide range of in-vitro protein digestibility (IVPD). Analyses were directed at polymers obtained by application of a fractionation method designed to keep polymer structure intact insofar as possible. A 2-dimensional (2-D) technique involving peak collection after size exclusion chromatography followed by reverse phase high performance liquid chromatography (SEC x RP-HPLC) of the collected peaks was applied to protein polymers previously determined to be correlated with IVPD. RP-HPLC chromatogram patterns unique to each collected SEC peak from three selectively extracted protein fractions allowed qualitative and quantitative comparisons of protein polymer components. A pair of early eluting peaks appearing in the γ-kafirin region of 2nd-dimension RP-HPLC chromatograms from a protein fraction with the largest Mw distribution were significantly correlated to IVPD within the study sample set (summed peaks $r = -0.769$, second peak individually $r = -0.785$, $p \leq 0.05$). Only the γ-kafirin found in this protein fraction correlated to IVPD. The correlated peak of interest was collected and characterized using SDS-PAGE and was preliminarily identified as 27kDa γ-kafirin. By combining techniques using differing selectivities (solvent based, molecular size based, hydrophobicity based), it was possible to disassemble and compare components of protein polymers significantly correlated to IVPD. The complementary (orthogonal) 2-D method used here allowed determination of monomeric compositional differences between sample protein polymers. These differences may provide indications for rationalizing the protein body structure of non-mutant sorghum lines, and the associated impact on protein digestibility.

Introduction

Grain characteristics contributing to poor sorghum protein digestibility were considered in an extensive review by Duodu et al. (2003). Of the endogenous (protein only) and exogenous (non-protein/protein interactions) factors considered by those authors, they concluded the most significant effect on protein digestibility was exerted by the endogenous factor of protein crosslinking. Emmambux and Taylor (2009), Nunes et al. (2004; 2005), El Nour et al. (1998) and others have pursued this idea with a goal of determining the characteristics of crosslinked protein polymers in sorghum. Much of this work has been successful in terms of a determination of overall individual kafirin composition. However, little work has been done determining the crosslinked composition of individual sorghum protein polymers. Progress has been limited by an inability to extract the largest polymers in an intact form, thereby precluding meaningful sample comparisons. A determination of many of the details of crosslinking, particularly in relation to how differences relate to digestibility has been hindered as a result.

2-dimensional (2-D) analytical methods employing complementary (orthogonal) separation mechanisms provide the advantage of enhanced peak capacity (Giddings, 1987). The resolution of sample components using multidimensional chromatography is greatly improved over single dimension techniques (Giddings, 1987; Nice and Aguilar, 2004). Previous studies on polymeric cereal proteins have relied on SDS-PAGE and SEC x SDS-PAGE for characterization of protein polymer composition (Singh and MacRitchie, 2004; El Nour et al. 1998; Emmambux and Taylor, 2009; Nunes et al. 2005). By taking advantage of the enhanced resolution and quantitative capabilities offered by RP-HPLC, a 2-D method using RP-HPLC as the second dimension should allow a more detailed look at protein polymer composition. 2-D characterization of cereal protein polymers using SEC x RP-HPLC has been used before to

determine the high molecular weight and low molecular weight glutenin subunit composition of gluten protein extracts from wheat (Larroque et al. 1997). The goal of the current study was to investigate the application of a 2-D analytical technique for the determination of the monomeric building blocks of intact sorghum protein polymers previously determined to significantly correlate with IVPD.

Materials and Methods

Sorghum samples

Sorghum samples used in this study were wild-type non-tannin varieties from a sorghum association mapping population (diversity panel) grown at the Kansas State University agronomy farm (see panel description Chapter 3). From a 27 sample subset of the full non-tannin mapping population that had been previously characterized for traits with potential impact on digestibility (see Chapter 3), 8 samples were randomly selected for more in depth study. These 8 samples (hereafter referred to as the 2-D set) spanned a range of IVPD values and were chosen to be representative of the protein contents of the larger subset. Values for protein content and IVPD for the samples in the 2-D set are presented in **Table 5.1**.

Grain Preparation

Sorghum kernels were milled to flour using a UDY mill (UDY Corp., Fort Collins, CO) equipped with a 0.5mm screen. The milled samples were stored in sealed containers at -20˚ C until needed.

Protein content determination

Total protein content of milled and dried undigested and digested samples was measured by nitrogen combustion using a Leco FP-528 nitrogen determinator (St. Joseph, MI) according to AACC method 46-30 Crude Protein-Combustion Method (AACC International, St. Paul, MN). A factor of 6.25 was used for conversion of percent nitrogen to crude protein percent.

Total kafirin extraction

Total kafirin extracts destined for reverse phase high performance liquid chromatography (RP-HPLC) analysis were obtained by application of a 60% t-butanol (v/v) / 0.5% sodium acetate (v/v) / 2% BME extraction protocol to duplicate 100mg milled samples as described in Bean et al. (2011).

Extractions (3F method)

The extraction methodology developed as described in Chapter 2 was used to obtain protein extracts for subsequent SEC analysis, and consisted of the following procedures. A water/salt soluble protein fraction (albumins + globulins) and three additional protein fractions (subsequently referred to as F1, F2, and F3) were obtained by application of a sequential extraction scheme to duplicate 100mg milled sample portions in 2.0mL μ-tubes. Albumins, globulins, and low molecular weight nitrogen (LMWN) were extracted first, using two 5min extractions with 1mL of 1M NaCl each extraction (room temperature vortex at medium speed followed by centrifugation at 9300g for 4min), and this fraction was discarded. This was followed by a 5min wash step with 1mL of deionized H_2O, mixed and centrifuged as before, and discarded. Next, Fraction 1 (F1) was obtained using two sequential 50min extractions with 1mL portions of 60% t-butanol (v/v) / 80mM Tris-borate pH10.0, mixed and centrifuged as before, and the supernatants were combined 1:1 in a clean μ-tube. To obtain Fraction 2 (F2), the same

pellet was subjected to two sequential 50min extractions with 1mL of 50mM Tris-borate pH10.0 / 2% SDS (w/v), mixed and centrifuged as before, and the supernatants were combined 1:1 in a clean μ-tube. Sonication was applied to the same pellet from the previous extractions to obtain Fraction 3 (F3) using a single 1mL portion of 50mM Tris-borate pH10.0 / 2% SDS (w/v). A 0.125in OD sonication probe powered by a Sonic Dismembrator 60 unit (Thermo Fisher Scientific, Waltham, MA) was positioned within the tube contents, and centered approximately 4-7mm from the bottom of the μ-tube that was immersed in an ice water bath to reduce solution heating. Three 15s bursts of sonic energy were applied while manually maintaining 10W of output power (sonicator readout), and allowing 30s rests between and after the final burst. After sonication, the mixture was centrifuged as before and the supernatant was transferred to a clean μ-tube.

After each fraction was extracted, the extracted supernatants were subjected to 2min at 80˚ C heat deactivation of intrinsic proteases during immersion of the capped μ-tubes in the water-filled wells of a heat block (Pierce, Rockford, IL). After being allowed to cool to room temperature, extracts were filtered through 25mm 0.45μm GHP membrane syringe filters (PALL Life Sciences, Port Washington, NY) into clean μ-tubes or HPLC vials. Fraction extracts for size exclusion chromatography (SEC) analysis were analyzed fresh (not lyophilized).

Size exclusion chromatography (SEC) analysis and conditions

SEC protein separation conditions consisted of 50uL injections on silica based Yarra SEC-3000 columns, 300mm x 7.80mm (Phenomenex USA, Torrance, CA). Isocratic runs were accomplished with a mobile phase of 50mM Tris-borate pH7.0 / 1% SDS (w/v) and a flowrate of 0.5mL/min while maintaining a column temperature of 40˚ C. A UV wavelength of 210nm was monitored for analyte detection.

SEC fraction collection

Individual peaks from SEC separations of extracted fractions were collected during runs from duplicate 50μL injections into the same HPLC vial. An Agilent 1100 Series module Model G1364A fraction collector (Agilent, Santa Clara, CA) was programmed to collect separated SEC peaks into empty HPLC vials at timed intervals as described in **Table 4.2**. Aliquots of collected peak volumes were placed in clean labeled μ-tubes and frozen prior to lyophilization.

RP-HPLC analysis and sample prep conditions

Lyophilized aliquots representing collected SEC peaks were resuspended prior to RP-HPLC analysis by separate additions of 50mM Tris-borate pH10.0 / 2% SDS and 7M Urea / 2M Thiourea at a 1:1 ratio to achieve the original pre-lyophilized volume. A 10min mixing time was applied after addition of each solvent. To the resuspended collected peak volume, BME was added at 2% (v/v) of resuspended volume and mixed for an additional 10min. Then the collected peak volume was alkylated by addition of 4-VP at the rate of 6.65μL 4-VP / 100μL resuspended volume and mixed for a final 10min.

RP-HPLC analyses were carried out on an Agilent 1100 series HPLC system (Agilent, Santa Clara, CA). Chromatography conditions were as described by Bean et al. (2011). Briefly, protein separations were performed from 10uL injections on silica based Poroshell C18, 75mm x 2.1mm (Agilent, Santa Clara, CA). The flow-rate was 0.7mL/min and mobile phase for gradient runs consisted of mobile phase A: deionized water / 0.1% TFA (v/v), and mobile phase B: acetonitrile / 0.07% TFA (v/v). Gradient times were 20 - 40% B from 0 to 5min; 40% - 60% B from 5min to 15min; 60% to 20% B from 15min to 17min; with a 5min post-time at 20% B. The column was maintained at 55 ˚C and UV detection was used at 214nm.

SDS-PAGE

Total kafirin and F1, F2, and F3: A precast NuPAGE 12% Bis-Tris polyacrylamide, 1.0mm x 10 well gel (Novex) was pre-run with MOPS SDS Running Buffer (Novex) for 30 min at 60V current. The samples (23 µL each) were mixed with 10 µL SDS sample loading dye that contained no reducing agent in 0.7 mL microtubes. The samples contained urea and were not boiled or heated. After vortexing for 3 seconds, the samples were centrifuged at 3 k g for 5 seconds. Six µL of Precision Plus Protein™ Dual Color standard (Bio-Rad) were loaded into wells 1 and 10 as a molecular weight marker. Eight µL of prepared samples were loaded into corresponding wells. A water blank mixed with a sample loading dye was loaded into well 6 to separate the reduced samples from the non-reduced. The proteins were settled in the wells for 5 min and then run for 10 minutes at 60V. The voltage was raised to 80V and maintained for 3 hours. Upon electrophoresis completion, the proteins were fixed in the gel and stained overnight with Colloidal Blue Stain (Novex). After destaining in diH_2O the gel image was captured on a V700 Photo scanner (Epson).

Collected RP-HPLC peak from SEC F3 peak 1: A 20 µl aliquot of sample was mixed with 5 µl of 4x NuPAGE LDS Sample Buffer loading dye (Novex) containing 1% β-mercaptoethanol. The sample was warmed in an 80 ˚C water bath for 5 minutes, vortexed, centrifuged and then loaded into a high capacity , 4-12%, 12 well, ExpressPlus™ PAGE Gel (GenScript). Precision Plus Protein™ Dual Color Protein Standards (Bio-Rad), Mark 12 Unstained Standards (Novex) and Sharp Pre-stained Protein Standards (Novex) molecular weight markers were loaded into neighboring wells. The gel was run in XT MES buffer (Bio-Rad) at 60V for 1 hour followed by 80V for an additional hour. After electrophoresis, the gel was rinsed in deionized water and silver stained with ProteoSilver ™ Plus Silver Stain (Sigma-Aldrich) per the standard protocol of the manufacture. Briefly, the gel was sensitized, silver stained, and

slightly over developed for maximum band detection. The gel image was captured on aV700 Photo scanner (Epson).

Statistics

All data were plotted, and statistical analyses were conducted using Microsoft Excel (Redmond, WA) and Origin (OriginLab Corp., Northampton, MA). Significant correlations were determined using the Pearson product-moment correlation coefficient and n-2 degrees of freedom where n = sample number. A significance level of $p \leq 0.05$ was used throughout unless noted otherwise in the text.

Results and Discussion

SEC of extracted fractions

Using non-reducing extraction conditions, a sequential protein extraction method (described in Chpt. 3) was applied to sorghum samples comprising the 2-D sample set. The three separate fractions obtained (designated F1, F2, and F3) were each a mixture of polymeric and monomeric proteins. Separation of the proteins contained in each fraction by size exclusion chromatography (SEC) allowed visualization of the respective relative molecular weight distribution within each sample, as well as between samples (**Figure 5.2**). As previously discussed in Chpt. 3, shorter retention time SEC peaks represent larger molecular weight proteins. It is also clear from **Figure 5.2** that as one considers the chromatographic traces going from F1 to F3, the molecular weight distribution of all samples increased from lower to higher relative molecular weight distributions. This indicates a greater proportion of the proteins in F3 are in a more highly polymerized form.

The results of the research reported on in Chpt. 3 allowed determination of statistically significant correlations between some of the SEC separated protein groups and sample *in-vitro* protein digestibility (IVPD). These groups (peaks) were represented by SEC separations of the sequentially extracted proteins from F1, F2, and F3 for the larger sample set (n = 27) considered previously. In that work, the SEC protein peaks most significantly correlated to IVPD were almost exclusively from the F3 fraction, specifically F3 SEC peaks 1 through 4 (P1, P2, P3, P4), and were predominantly polymeric in nature. As a consequence of those results, the distribution and composition of these proteins were of particular interest, and were a focus within the current study.

Figure 5.3 contains examples of SEC separations of reduced and non-reduced low and high IVPD F3 extracts. These and other SEC separations analyzed reduced and non-reduced (data not shown), indicated that at least a portion of the proteins from F3 eluting prior to ~12.2min (represented by P1, P2, and P3) were covalently linked polymers (see **Figure 5.3**). Possibly significant as well is that a portion of these same proteins were not reduced, and may be indicative of non-covalent associations with the potential for impacting digestibility.

SDS-PAGE of total kafirins and total fractions F1, F2, F3

To help visualize the proteins obtained by application of the three step extraction procedure used in the current study, an example of total fractions F1, F2 and F3 as well as a total kafirin extract was analyzed by SDS-PAGE in reduced and non-reduced form (**Figure 5.4**). This also serves as a frame of reference to other studies that have utilized SDS-PAGE for characterization of extracted kafirins. Notable in **Figure 5.4**, is the presence of a significant amount of higher molecular weight material (100+ kDa) in the reduced F3 lane as well as in the non-reduced F3 lane, evidenced by the noticeable darker smear present there. The fact that large

molecular weight material is still present after reduction is in agreement with the results obtained by SEC after reduction that showed the persistence of portions of several peaks representing higher molecular weight distribution proteins (see **Figure 5.3**). Recall also that it was the F3 proteins that exhibited the most significant correlation to IVPD. In fact, large molecular weight material was faintly apparent in all the non-reduced fractions, and provided additional evidence the extraction method was successfully extracting intact polymers.

RP-HPLC of total kafirins and total fractions F1, F2, F3

Total kafirin extractions were performed on all study samples in addition to application of the 3F extraction method previously described. Initial portions of the study involved examination of total kafirin extracts and the three sequentially obtained extracts F1, F2, and F3 directly by RP-HPLC analyses. **Figure 5.5** provides a qualitative comparison between the RP-HPLC chromatograms of low (IVPD = 48.96%) and high (IVPD = 74.11%) protein digestibility samples from the study set. Total RP-HPLC peak areas were determined for the total kafirin extracts and each individual fraction (F1, F2, F3), as well as peak areas corresponding to the γ-kafirin area, to see if relationships to IVPD could be discerned. The γ-kafirin elution time was determined from prior RP-HPLC analysis of isolated γ-kafirin (data not shown). These peak area values were used to calculate correlations to IVPD values.

With regard to RP-HPLC analyses of total kafirin extracts within the study sample set, no correlation was found between total kafirin peak area and IVPD. This finding was consistent with the lack of a significant correlation between IVPD and overall protein content. However, a statistically significant correlation was found between the γ-kafirin peak area within the total kafirin extracts and IVPD ($r = -0.896$). This mirrors results from previous research on sorghum protein digestibility, and appears to confirm the significant role γ-kafirins play in this regard (El

Nour et al. 1998; da Silva et al. 2011; Oria et al. 1995). Aside from affirming the already generally acknowledged capacity for covalent bonding by γ-kafirins, specific details on how this may relate to digestibility differences among sorghum varieties are not as clear. Subsequent portions of the current study sought data useful in answering this question, especially within the context of protein packaging and grain structure.

Within individual fractions (F1, F2, F3) extracted from study samples, two significant relationships to IVPD were noted following RP-HPLC analyses. The total peak area of F2 was significantly correlated ($r = -0.855$), as was the peak area corresponding to the γ-kafirin retention time range within the F3 fraction. The correlation of RP-HPLC total kafirin peak area from F2 was interesting, especially when considered in relation to earlier data from work in our lab (Chpt. 3) in which SEC analysis of a larger sample set that also included the same samples used in the current study specifically did not show a significant correlation between total F2 SEC peak area and IVPD. It is possible the more numerous sample protein complements available in the larger sample set from the earlier work served to cancel out a chance bias within the current study set with regard to F2 proteins. Supporting this idea, the SEC analyses undertaken within the current study of the 8 samples also indicated a significant correlation ($r = -0.733$) between total F2 SEC peak area and IVPD, and thus agrees with the RP-HPLC results. Although significant within the constraints of the smaller data set of the current study, the lack of agreement with data from the larger sample set in the previous work indicates additional research is needed to support further focus on the F2 fraction in relation to IVPD.

The second significant correlation to IVPD noted from RP-HPLC analyses of the individual extracted fractions was from the γ-kafirin peak area within F3 ($r = -0.766$). Considering that peaks within the γ-kafirin retention time range were evident in all three

extracted fractions, it is particularly intriguing only the peaks within the γ-kafirin area obtained from F3 were significantly correlated (**Figure 5.6**). This result begs the question regarding potential difference(s) between the γ-kafirin extracted from the three different fractions, and how that relates to IVPD differences among samples. The possibility for sequence variation at the gene level within the 27kDa γ-kafirin subclass of *Sorghum bi-color* is limited to two protein variants differing by a single amino acid substitution (proline to alanine substitution) as revealed by gene sequencing (Laidlaw et al. 2010). A comparison of the predicted hydrophobicity of γ-kafirin using an example γ-kafirin amino acid sequence from Uniprot ID accession number C5XDL2 reveals a GRAVY score (grand average of hydropathy) of -0.244 and -0.227 with proline or alanine respectively in the substitutable amino acid position (The Uniprot Consortium, 2009; Kyte and Doolittle, 1982). More positive scores represent more hydrophobic character (signal peptides were removed). The proline to alanine substitution does not appear to make a substantial impact to the overall hydrophobic character of γ-kafirin in terms of hydropathicity score (net difference of 0.17). Considered on its own, it is not obvious this difference is enough to influence a significant correlation to IVPD. Additional influences, for example post-translational protein modifications and associated effects on crosslinking and protein body structure must also be considered.

When the individual fraction (F1, F2, F3) γ-kafirin peak areas were plotted against the combined γ-kafirin peak areas (total γ-kafirin) from the samples in the study, it revealed that both F2 and F3 γ-kafirins were significantly positively correlated to total γ-kafirin ($r = 0.779$ and 0.889 respectively). And when the percent γ-kafirin content of individual sample fractions were plotted against IVPD to see the effect of the relative γ-kafirin content on digestibility, no significant relationships were indicated. Together these findings imply there is more to sorghum

protein digestibility variations than simply potential for disulfide cross-linking as reflected by γ-kafirin content. If γ-kafirin content was the overriding factor, it would seem logical for both F2 and F3 γ-kafirins to show significant correlation to IVPD. Recall from previous discussion that only the γ-kafirins from F3 were significantly correlated to IVPD (r = -0.766). The average relative γ-kafirin contents of the F1, F2 and F3 fractions for all samples in the study were 12.8%, 41.8%, and 45.4% respectively.

Considering the varying selectivity for protein fractionation conferred by the respective fraction extraction conditions (solvents, sonication) used in the current study, it may not be unreasonable to wonder if differences in the polymeric associations of some subclass proteins become apparent as intact polymers are isolated using increasingly more rigorous extraction conditions. In this scenario, a picture of unique associations amongst protein polymers based on how they are packaged within protein bodies (perhaps during development, or grain dry-down) for a given sorghum variety's protein body complement might be reflected in the extracted fraction in which they ultimately appear. Variations in IVPD among wild type sorghums might then be indicative of differences in polymeric protein packaging, rather than due solely to differences within specific subclass protein molecule structures per se for a given sample. To help answer this question, additional levels of protein selectivity were applied to help further differentiate the proteins in the next portion of the study.

SEC x RP-HPLC of collected SEC peaks

RP-HPLC for the analysis of cereal proteins is well established (Bietz 1983, 1985; Bean et al. 2011). It typically requires the application of a reducing agent prior to sample introduction for analysis. It also relies on the intrinsic property of molecular hydrophobicity within a given mobile phase system as a means of achieving selective separation. As the second separation

dimension following SEC separation of intact polymeric sorghum proteins, it provides an excellent orthogonal separation solution (Gilar et al. 2005). Due to greatly enhanced peak capacity, the resolution of sample components using multidimensional chromatography is greatly improved over single dimension techniques (Giddings, 1987; Nice and Aguilar, 2004). Three levels of analytical selectivity were applied to differentiate and characterize the kafirin contents of samples within the current study representing differential solubility, molecular size (hydrodynamic radius), and hydrophobicity (**Figure 5.7**).

2-D methods for characterization of cereal protein polymers using SEC as the first dimension have been applied to cereal proteins before (Larroque et al. 1997), and served as a means for determining the high molecular weight and low molecular weight glutenin subunit composition of gluten protein extracts. Examples of the SEC separations of sequentially extracted fractions F1, F2, and F3 are illustrated in **Figure 5.8**. Elution intervals for collection of individual SEC peaks used for subsequent 2-D analysis by RP-HPLC are indicated in **Figure 5.8** as well. These SEC peaks were individually collected, lyophilized, resuspended and reduced for second dimension analysis by RP-HPLC.

Figure 5.9 illustrates the RP-HPLC chromatograms obtained for individual SEC peaks (denoted in the figure as P1 through P5) collected from SEC separations of three selectively extracted fractions from a low and a high IVPD study sample. The peaks in these chromatograms represent the monomeric proteins that make up the proteins in the individual SEC peaks from which they were collected. Qualitatively, differences from comparisons between the low and high IVPD samples seemed fairly subtle. In particular, a close examination of the RP-HPLC peaks within the 5.0min – 6.0min range appearing in the F3 chromatograms did show noticeable differences, especially for the two larger peaks seen there.

Figure 5.10 focuses on only the 2-D results of the SEC x RP-HPLC chromatograms for the collected F3 SEC peaks that were shown to demonstrate significant correlations to IVPD in Chapter 3 (i.e. P1 through P4) for all samples. Quantitatively summing the RP-HPLC peak areas of the two RP-HPLC peaks at 5.0min – 6.0min from F3P1 from all study samples reveals they are significantly correlated to IVPD ($r = -0.769$). When considered separately, only the second peak significantly correlates ($r = -0.785$). Interestingly, the presence of two prominent RP-HPLC peaks in this area seems to be largely limited to the F3 fraction, and to be especially pronounced in the collected P1 and P2 SEC peaks. In contrast, in RP-HPLC runs of SEC peaks collected from F1 and F2, the second peak appears to be largely absent. As discussed earlier, peaks eluting in this area are thought to represent γ-kafirin.

To help characterize the second prominent correlating peak, an additional extraction of KS19 (IVPD = 48.96%) was performed to obtain fresh F3. Repeat SEC separations of fresh KS19 F3 allowed multiple collections of SEC P1 that were subsequently combined, lyophilized and resuspended at a concentrated level (1/10 original collected volume). Second dimension RP-HPLC runs of the concentrated collected F3P1 allowed collection of the peak of interest. In **Figure 5.11**, RP-HPLC chromatogram traces of the concentrated SEC peak F3P1, as well as a trace of the collected RP-HPLC peak of interest (i.e. the correlated second peak referred to previously) are shown. Having successfully isolated the peak of interest with little visible interference, further characterization by SDS-PAGE was attempted. This information could prove useful in rationalizing contribution to IVPD correlation as well as help suggest a possible role in protein body morphology. An indication of relative molecular weight was obtained from analysis by SDS-PAGE (**Figure 5.12**). The band representing the collected peak of interest (circled in the figure) was located within the 25-30kD Mw marker range. This is the molecular

weight range where we would expect γ-kafirin to be and in conjunction with the demonstrated RP-HPLC retention time indicates a good probability for nominal 27kDa γ-kafirin. The other bands visible within the sample lane of the gel were determined to be resuspension solvent components based on examination of solvent blank gels (data not shown).

An additional significant negative correlation (r = -0.833) with IVPD was also noted for the total peak area of a grouping of later eluting proteins (~10.5min – 12.2min) from RP-HPLC runs of collected SEC peak 4 from F3. Peaks in this elution range represent α-kafirins. Improvement in IVPD when the α-kafirins are less abundant has been corroborated by research done by Grootboom et al. (2014) in which the α-kafirin subclass (specifically α-kaf 1 (25kDa)) was genetically suppressed in transgenic sorghum experiments. In that study, interestingly, further improvement in IVPD was realized when the decrease in α-kaf 1 was accompanied by a simultaneous decrease in γ-kaf 1 (25kDa) and γ-kaf 2 (50kDa).

Conclusions

The current study was an extension of previous research (Chpt. 4), and was designed to further characterize the polymeric proteins of sorghum samples varying in IVPD. Application of a non-reducing sequential extraction method designed to obtain intact protein polymers allowed three different protein fractions to be collected from a wild type sorghum sample set exhibiting a broad range of IVPD. These three fractions (F1, F2, F3) were each separated by SEC, and the relative molecular weight distribution as well as quantitation of separated peak areas for each fraction was determined.

Separated SEC peaks from each fraction were collected, lyophilized, and resuspended for subsequent 2-D analysis by RP-HPLC. Much previous research has focused either on relative differences in monomeric kafirin subclass content, or on presence/degree of protein crosslinking as it relates to protein digestibility. Using a 2-dimensional analysis scheme, the current study attempts to go a step further by combining monomeric kafirin content and crosslinking information. The goal was to see if it was possible to determine what the components of protein polymers correlating with digestibility were, and attempt to determine those having the most impact on IVPD. By taking advantage of techniques using multiple selectivities (solvent based, molecular size based, hydrophobicity based), it was possible to disassemble and compare the significant protein polymers and their associated monomeric protein complements.

The monomeric protein complements from within separated and collected SEC protein peaks subsequently analyzed by 2-D RP-HPLC varied within and between samples. Some statistically significant relationships between IVPD and 2-D RP-HPLC peaks were found. One of these peaks ($r = -0.785$) was collected and characterized by SDS-PAGE and tentatively identified as the nominal 27kD γ-kafirin. Although present in most of the collected F3 SEC peaks, only the 27kD RP-HPLC peak found within the first collected SEC peak (P1) from F3 was significantly correlated to IVPD. This peak seemed to be present most predominantly within F3. Another relationship to IVPD was found from 2-D RP-HPLC data of monomeric proteins from within collected SEC peak 4 of the F3 fraction. The combined peak area of a group of late eluting RP-HPLC peaks representing α-kafirins showed a statistically significant negative correlation to IVPD ($r = -0.833$). Together, these data would appear to lend significance to the determination of not only presence and quantity of particular kafirins, but also of what larger polymeric structures those kafirins are a part of.

A detailed understanding of sorghum protein crosslinking, the related involvement in the determination of individual protein polymer structures, and how this effects incorporation into protein body structure has yet to be realized (Manieri et al. 2014). We know about the building blocks (kafirin subclass proteins) and we know the locations of the completed constructs (protein bodies) within the endosperm. Less clear is what happens between those two concepts in terms of factors governing the assembly, crosslinking processes, and polymer packaging that impact protein digestibility variations as displayed among sorghum varieties.

Results from the current study suggest the classic model for sorghum protein body structure is not sufficient to explain, for example, why monomeric proteins were present in all of the unreduced fractions (F1, F2, F3). A continuous γ- plus β-kafirin cross-linked protein "shell" should present a barrier to solvent access and prevent significant protein monomer solubilization similarly to how digestive enzymes have been described to be excluded during digestion (Lending and Larkins, 1989; Lending et al. 1988; Duodu et al. 2003; Belton et al. 2006). Could it be that the surface pitting on protein bodies occurring during protease exposure as described in TEM studies by Rom et al. (1992) and Taylor and Evans (1989) reflects discontinuous or "patch-like" protein polymer associations on the protein body surface? This type of protein body shell might then be susceptible to targeted degradation (digestion or solubilization) at the protein patch junctures. The current study found differences in γ-kafirin correlations to IVPD based on what polymer they were a part of. These results support the idea that to fully understand the relationship between kafirin proteins, sorghum protein body structure, and sorghum protein digestibility, information about protein composition at the protein polymer level is required.

Further work is under way investigating how sorghum protein polymers are influenced by the digestion process in both cooked and uncooked samples. By disassembling extracted

polymers from these treatments using the selective extraction procedures and 2-D analysis method described in this study, the role of cross-linked proteins in protein body structure and, ultimately, protein digestibility may be revealed.

References

Bean, S.R., Ioerger, B.P., Blackwell, D.L. 2011. Separation of kafirins on surface porous reversed-phase high-performance liquid chromatography columns. J. Agric. Food Chem. 59:85-91.

Belton, P.S., Delgadillo, I., Halford, N.G., Shewry, P.R. 2006. Kafirin structure and functionality. J. Cereal Sci. 44:272-286.

Bietz, J.A. 1983. Separation of cereal proteins by reversed-phase high-performance liquid chromatography. J. Chromatography A 255:219-238.

Bietz, J.A. 1985. High performance liquid chromatography: how proteins look in cereals. Cereal Chem. 62(3):201-212.

da Silva, L.S., Jung, R., Zhao, Z., Glassman, K., Taylor, J., Taylor, J.R.N. 2011. Effect of suppressing the synthesis of different kafirin sub-classes on grain endosperm texture, protein body structure and protein nutritional quality in improve sorghum lines. J. Cereal Sci. 54:160-167.

Duodu, K.G., Taylor, J.R.N., Belton, P.S., Hamaker, B.R. 2003. Factors affecting sorghum protein digestibility. J. Cereal Sci 38:117-131.

El-Nour, I.N.A., Peruffo, A.D.B., Curioni, A. 1998. Characterization of sorghum kafirins in relation to their cross-linking behavior. J. Cereal Sci 28:197-207.

Emmambux, M.N., Taylor, J.R.N. 2009. Properties of heat-treated sorghum and maize meal and their prolamin proteins. J. Agric. Food Chem. 57:1045-1050.

Giddings, J.C. 1987. Concepts and comparisons in multidimensional separation. J. High Res. Chrom. & Chrom. Commun. 10:319-323.

Gilar, M., Olivova, P., Daly, A.E., Gebler, J.C. 2005. Orthogonality of separation in two-dimensional liquid chromatography. Anal. Chem. 77:6426-6434.

Grootboom, A.W., Mkhonza, N.L., Mbambo, Z, O'Kennedy, M.M., da Silva, L.S., Taylor, J., Taylor, J.R.N., Chikwamba, R., and Mehlo, L. 2014. Co-suppression of synthesis of major α-kafirin sub-class together with γ-kafirin-1 and γ-kafirin-2 required for substantially improved protein digestibility in transgenic sorghum. Plant Cell Rep. 33:521-537.

Kyte, J., Doolittle, R.F. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.

Laidlaw, H.K.C., Mace, E.S., Williams, S.B., Sakrewski, K., Mudge, A.M., Prentis, P.J., Jordan, D.R., Godwin, I.D. 2010. Allelic variation of the β-, γ- and δ-kafirin genes in diverse sorghum genotypes. Theor. Appl. Genet. 121:1227-1237.

Larroque, O.R., Gianibelli, M.C., Batey, I.L., MacRitchie, F. 1997. Electrophoretic characterization of fractions collected from gluten extracts subjected to size-exclusion high-performance liquid chromatography. Electrophoresis 18:1064-1067.

Lending, C.R., Kriz, A.L., Larkins, B.A. Bracker, C.E. 1988. Structure of maize protein bodies and immunocytochemical localization of zeins. Protoplasma 143:51-62.

Lending, C.R., Larkins, B.A. 1989. Changes in zein composition of protein bodies during maize endosperm development. The Plant Cell 1:1011-1023.

Manieri, D., Morandini, F., Maitrejean, M., Saccani, A., Pedrazzini, E., Vitale, A. 2014. Protein body formation in the endoplasmic reticulum as an evolution of storage protein sorting to vacuoles: insights from maize γ-zein. Frontiers in Plant Sci. 5(331):1-11. doi: 10.3389/fpls.2014.00331

Nice, E.C., Aguilar, M-I. 2004. Chpt.12: Multidimensional HPLC purification of proteins. In: Methods in Molecular Biology, Vol.251, HPLC of Peptides and Proeins: Methods and Protocols. Ed.: M.-I. Aguilar, Humana Press Inc., Totowa, NJ.

Nunes, A., Correia, I., Barros, A., Delgadillo, I. 2004. Sequential in vitro pepsin digestion of uncooked and cooked sorghum and maize samples. J. Agric Food Chem. 52:2052-2058.

Nunes, A., Correia, I., Barros, A., Delgadillo, I. 2005. Characterization of kafirin and zein oligomers by preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis. J. Agric. Food Chem. 53:639-643.

Oria, M.P., Hamaker, B.R., Schull, J.M. 1995. In-vitro protein digestibility of developing and mature sorghum grain in relation to α-, β-, and γ-kafirin disulfide crosslinking. J. Cereal Sci. 22:85-93.

Rom, D.L., Shull, J.M., Chandrashekar, A., Kirleis, A.W. 1992. Effects of cooking and treatment with sodium bisulfite on in-vitro protein digestibility and microstructure of sorghum flour. Cereal Chem. 69(2):178-181.

Singh, H., MacRitchie, F. 2004. Changes in proteins induced by heating gluten dispersions at high temperature. J. Cereal Sci. 39:297-301.

Taylor, J.R.N., Evans, D.J. 1989. Action of sorghum proteinase on the protein bodies of sorghum starchy endosperm. J. Exp. Botany 40(216):763-768.

The Uniprot Consortium. 2009.

Table 5.1 Sorghum samples used in the study, including respective protein contents and *in-vitro* pepsin digestibility (IVPD).

Sorghum	**Flour protein (%)**	**Raw flour IVPD (%)**
KS19	15.63	48.96
SC760	15.85	54.95
SC805	15.67	56.31
Tx2741	13.04	58.33
SC749	13.91	66.44
SC587	11.41	69.27
SRN39	14.08	72.75
SC489	14.10	74.11
min =	11.41	48.96
max =	15.85	74.11
avg. =	14.21	62.64
SD =	1.52	9.24

Table 5.2 Description of peak collection intervals used during size exclusion chromatographic (SEC) separations of three extracted sorghum protein fractions.

Fraction	SEC Peak #	Peak collection interval (min)	Elapsed peak elution time (min)	Collected peak volume (mL)	2 run combined peak volume (mL)
1	1	8.8 - 10.3	1.50	0.75	1.50
	2	10.3 - 11.6	1.30	0.65	1.30
	3	11.6 - 12.85	1.25	0.63	1.25
	4	12.85 - 13.9	1.05	0.53	1.05
	5	13.9 - 15.4	1.50	0.75	1.50
2	1	8.8 - 10.3	1.50	0.75	1.50
	2	10.3 - 11.6	1.30	0.65	1.30
	3a	11.6 - 12.3	0.70	0.35	0.70
	3b	12.3 - 12.85	0.55	0.28	0.55
	4	12.85 - 13.9	1.05	0.53	1.05
	5	13.9 - 15.4	1.50	0.75	1.50
3	1	8.8 - 10.3	1.50	0.75	1.50
	2	10.3 - 11.6	1.30	0.65	1.30
	3	11.6 - 12.85	1.25	0.63	1.25
	4	12.85 - 13.9	1.05	0.53	1.05
	5	13.9 - 15.4	1.50	0.75	1.50

Figure 5.1 Relationship between protein content and IVPD of the sorghum samples used in the study. Correlation = - 0.596 not significant at $p \leq 0.05$.

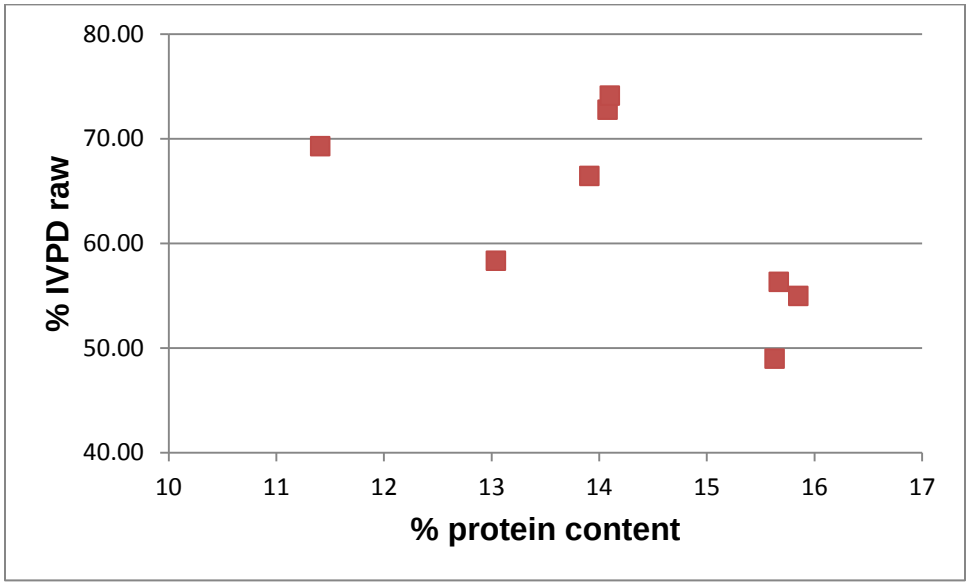


Figure 5.2 SEC chromatograms of F1, F2, and F3 from all study samples displayed with respective IVPD values. Earlier eluting peaks represent larger relative molecular weights.

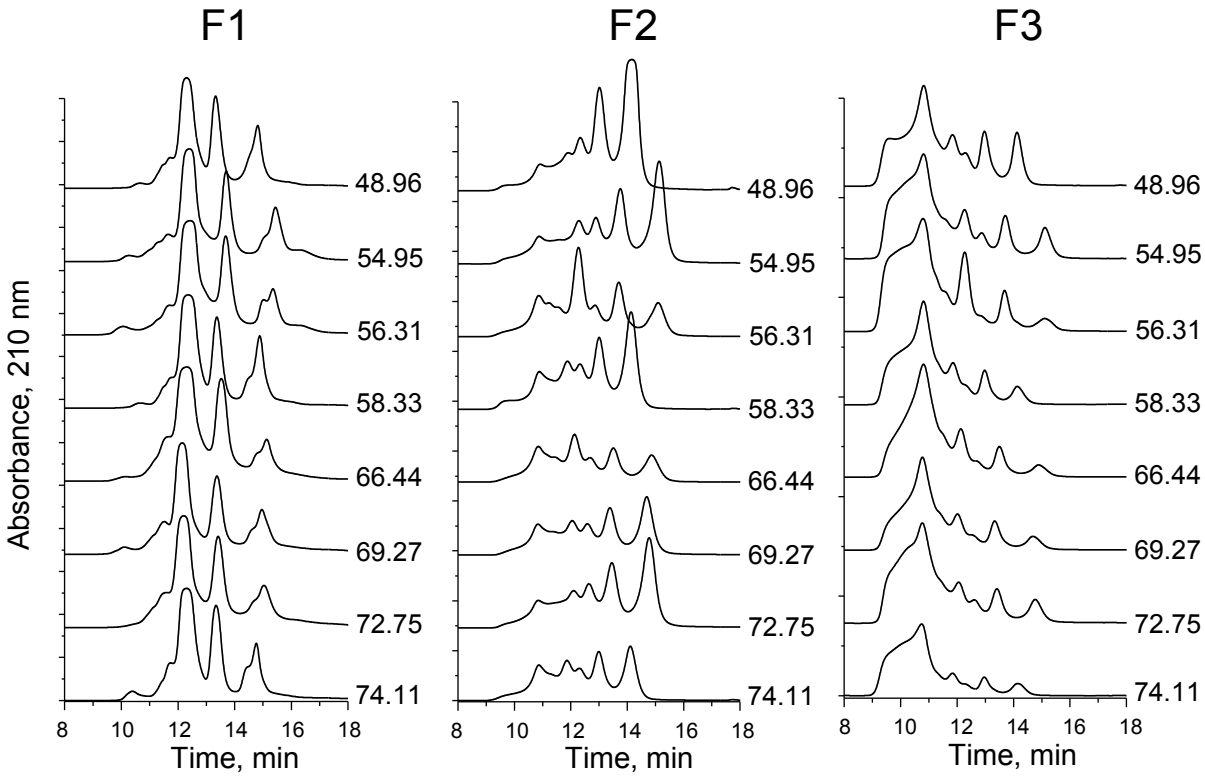


Figure 5.3 Examples of Fraction 3 (F3) SEC chromatograms from low (48.96%) and high (74.11%) IVPD samples. (reduced - - - - -) (non-reduced ———)

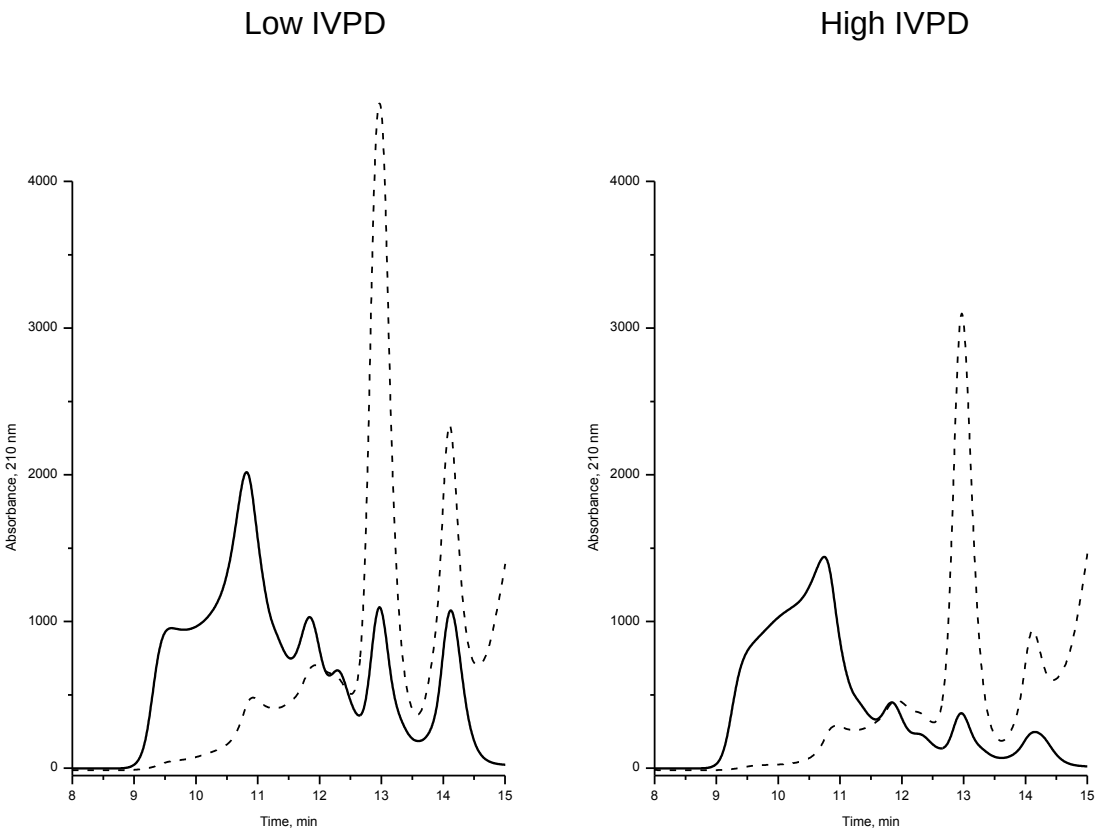


Figure 5.4 Example of SDS-PAGE analysis of a total kafirin extraction and three fractions (F1, F2, F3) obtained as in the current study (extraction conditions as described in Materials and Methods). Extracts were analyzed in reduced and non-reduced form. (sorghum sample: F1000)

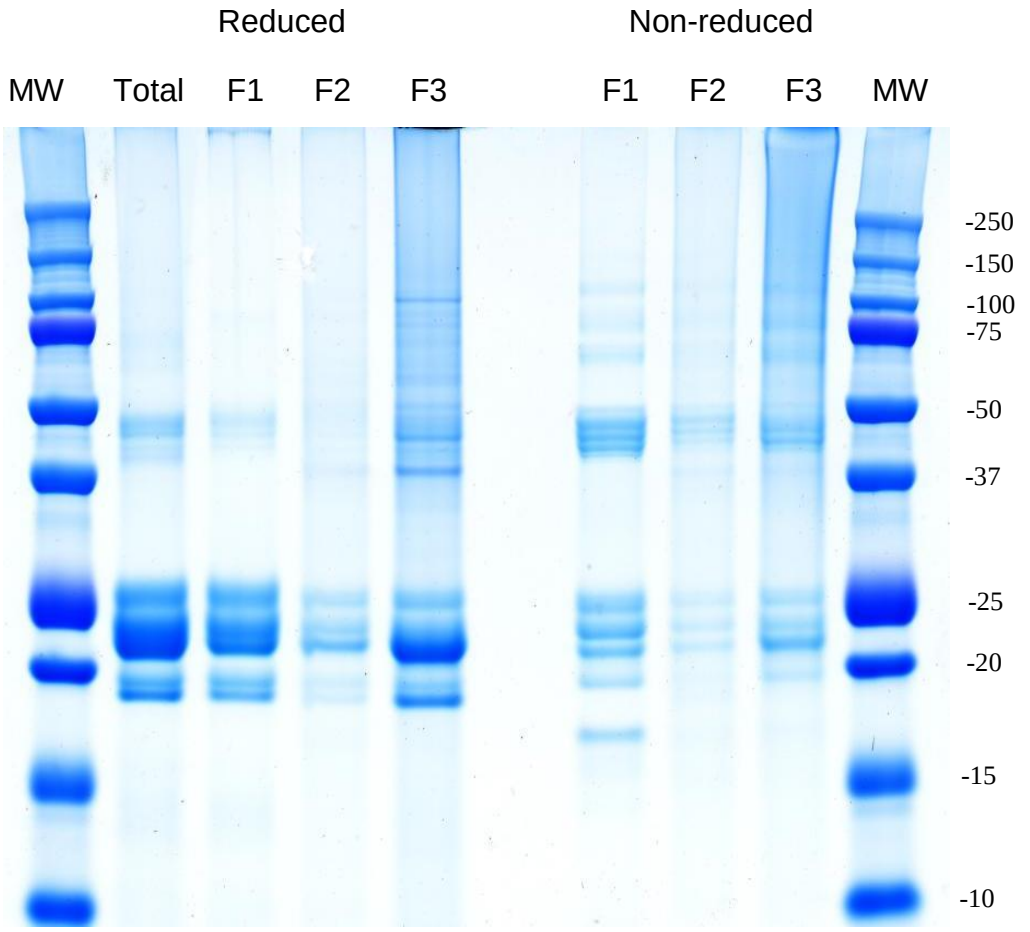


Figure 5.5 RP-HPLC chromatograms of the total kafirin extracts and three sequentially extracted fractions (F1, F2, F3) from low (48.96%) and high (74.11%) IVPD samples. The γ-kafirin peak area is indicated in the graphs.

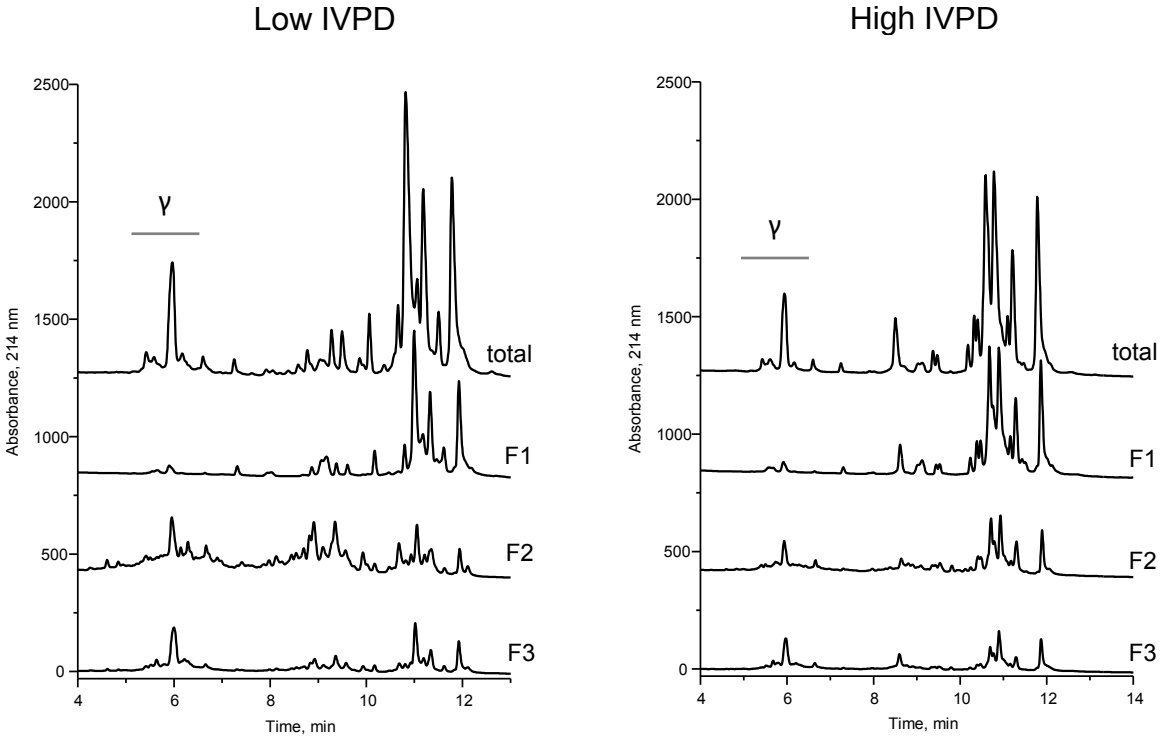


Figure 5.6 Correlation of IVPD with total F1 or total F2 or total F3 γ-kafirin RP-HPLC peak area of study samples. (F1 r = 0.089; F2 r = -0.697; F3 r = -0.766) (only total F3 γ-kafirin peak area significant @ $p \leq 0.05$)

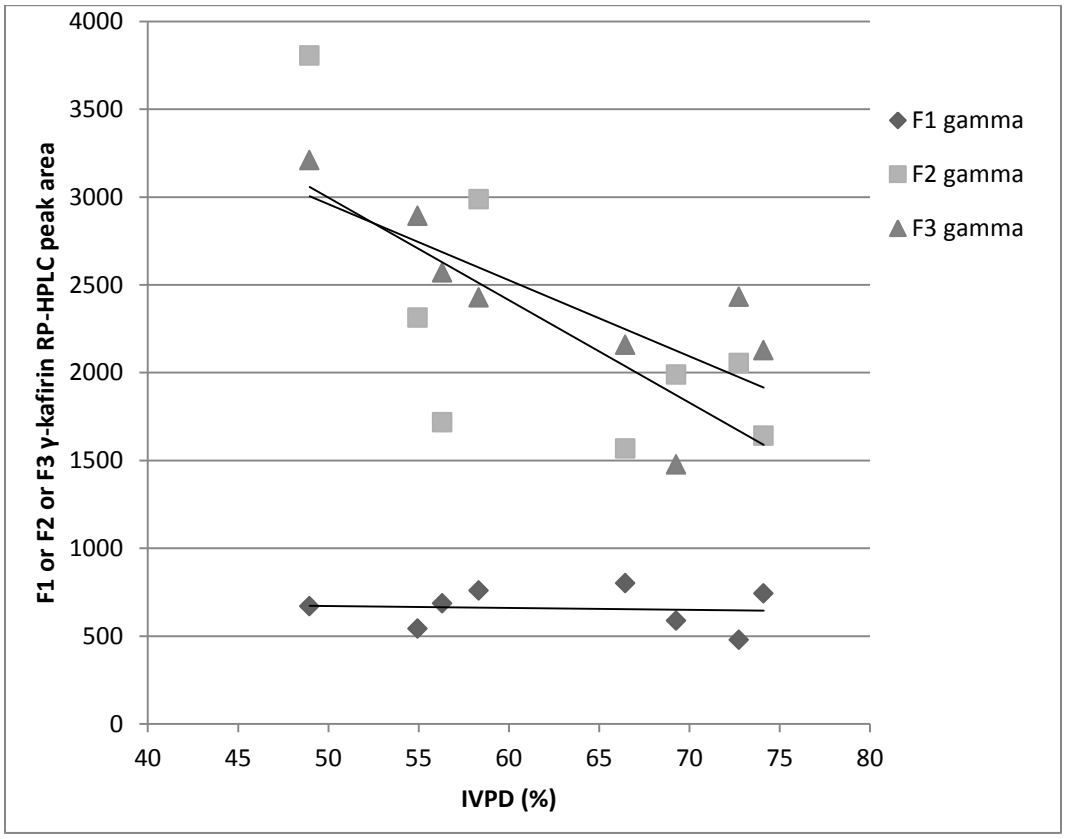


Figure 5.7 Three levels of selectivity used to probe characteristics of sorghum protein composition and structure using selective fractionation, and 2-D SEC x RP-HPLC. (F = protein fraction; P = analytical peaks)

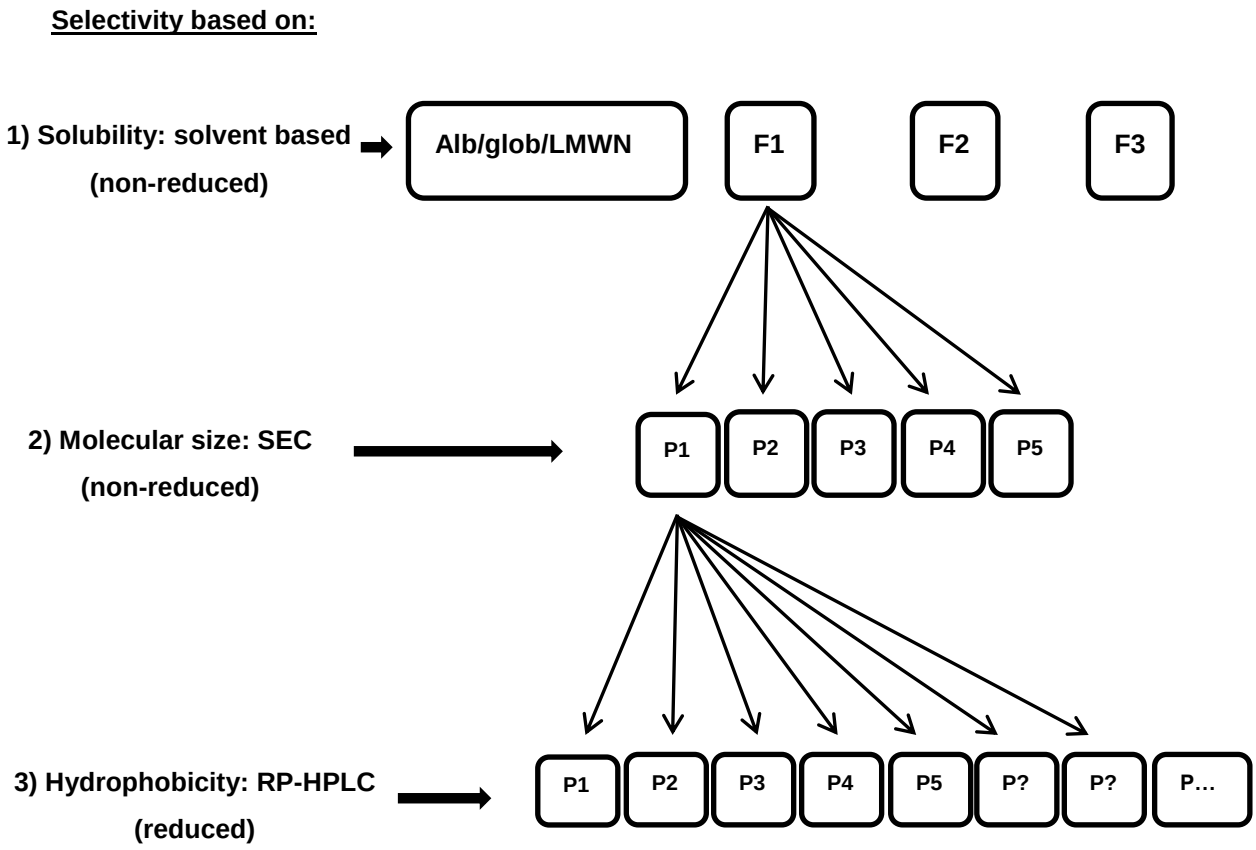


Figure 5.8 Examples of Fractions 1, 2 and 3 analyzed by SEC, as well as peak collection intervals for individual peaks collected from within each fraction for use in subsequent 2-D analysis by RP-HPLC. Relative Mw from analysis of Mw standards as indicated.

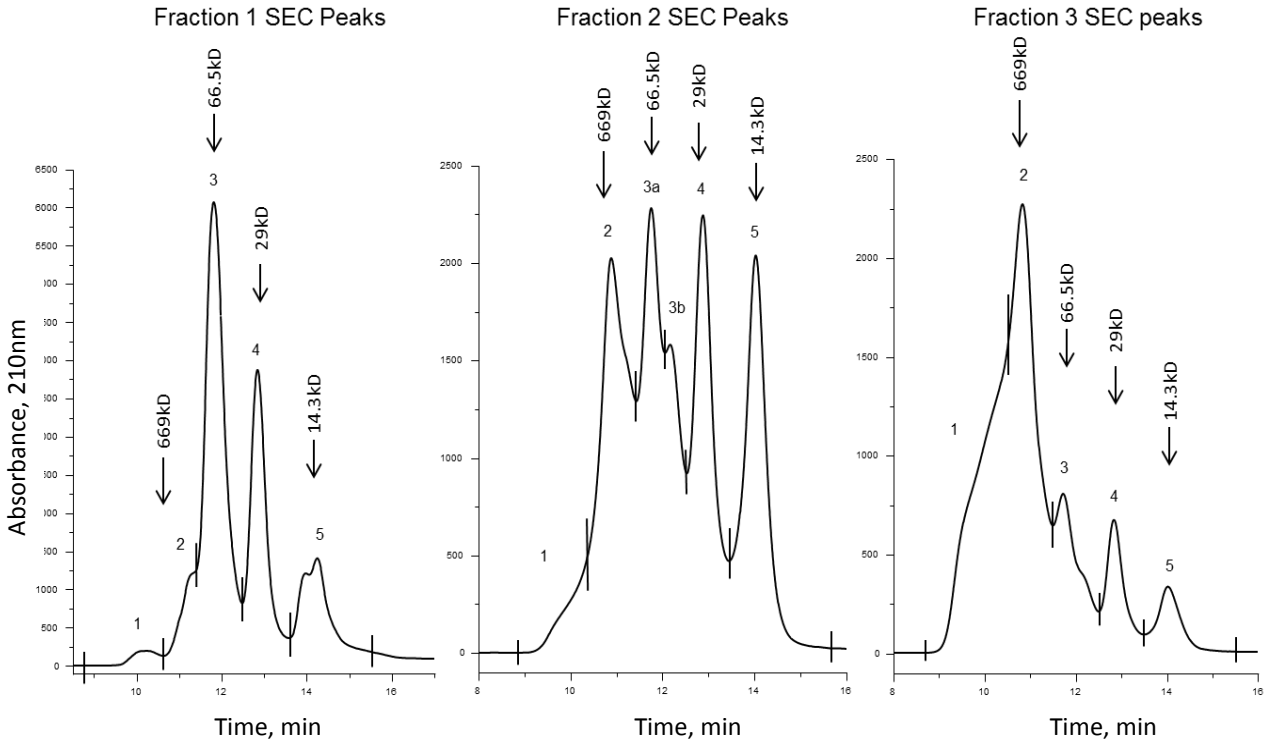
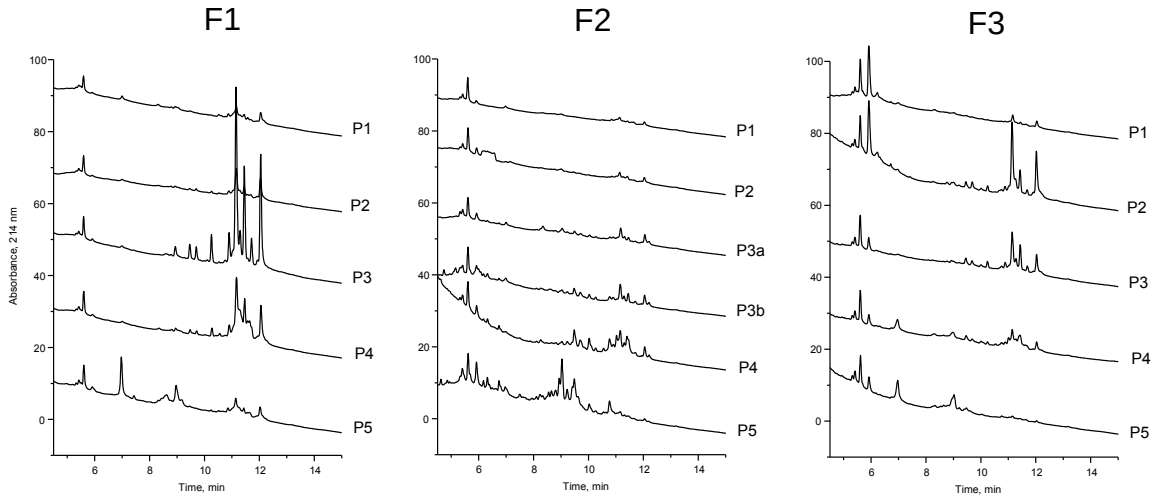


Figure 5.9 RP-HPLC chromatograms of collected SEC peaks from three different protein fractions extracted from sorghum flours exhibiting low (a) KS19 (IVPD = 48.96%) and high (b) SC489 (IVPD = 74.11%) IVPD values. Refer to Figure 5.8 for positions of SEC peaks at time of collection. (P1 = SEC collected peak 1, P2 = SEC collected peak 2, etc.)

a)



b)

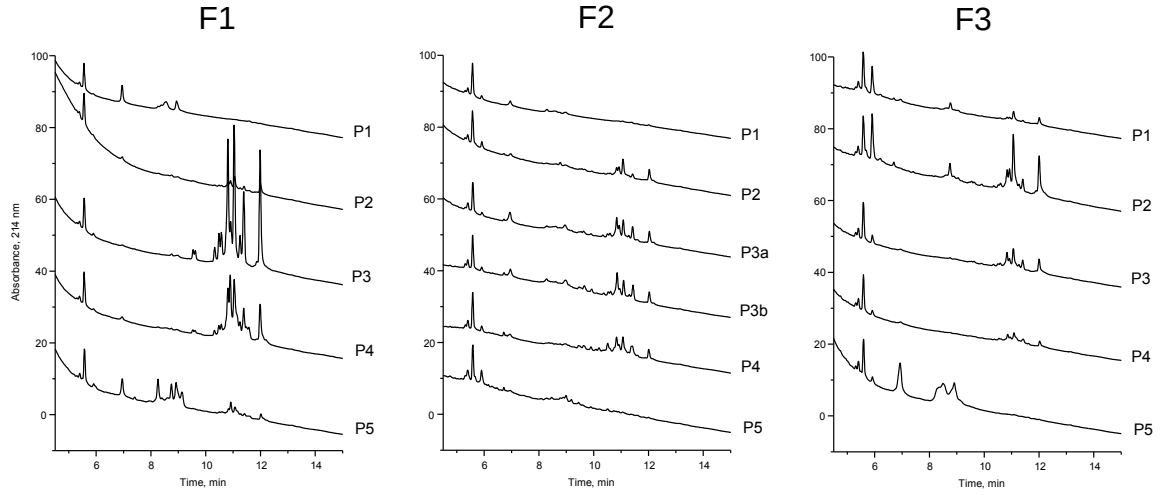


Figure 5.10 Comparison of RP-HPLC Chromatograms of collected SEC peaks 1, 2, 3 and 4 from Fraction 3 with respective IVPD values from all study samples.

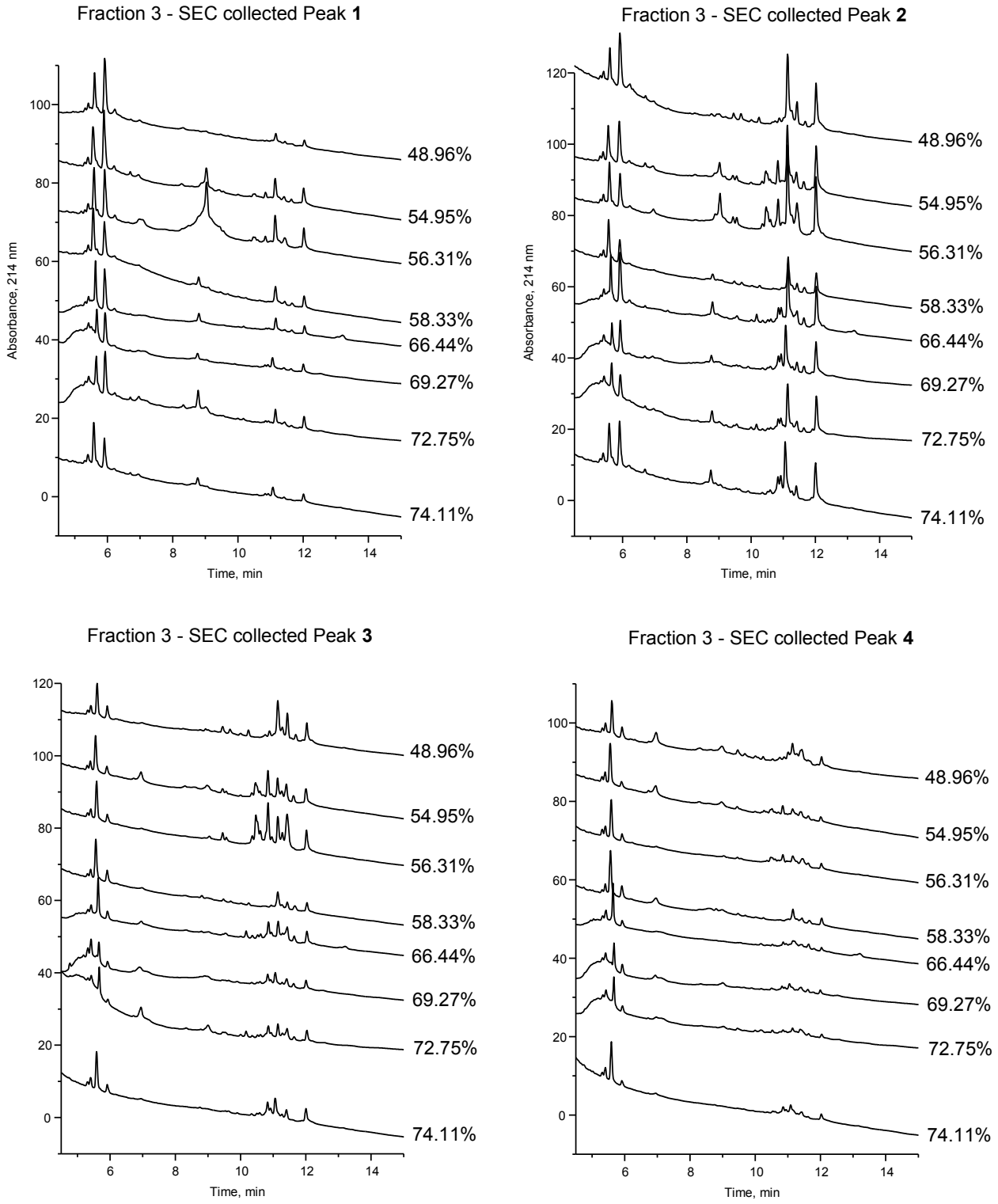


Figure 5.11 A second dimension RP-HPLC chromatogram of a concentrated suspension of the collected SEC peak F3P1 (-----) overlaid on a chromatogram of a peak of interest (———) isolated from it. The isolated peak (boxed) was collected because it showed significant negative correlation ($r = -0.785$, $p \leq 0.05$) to IVPD.

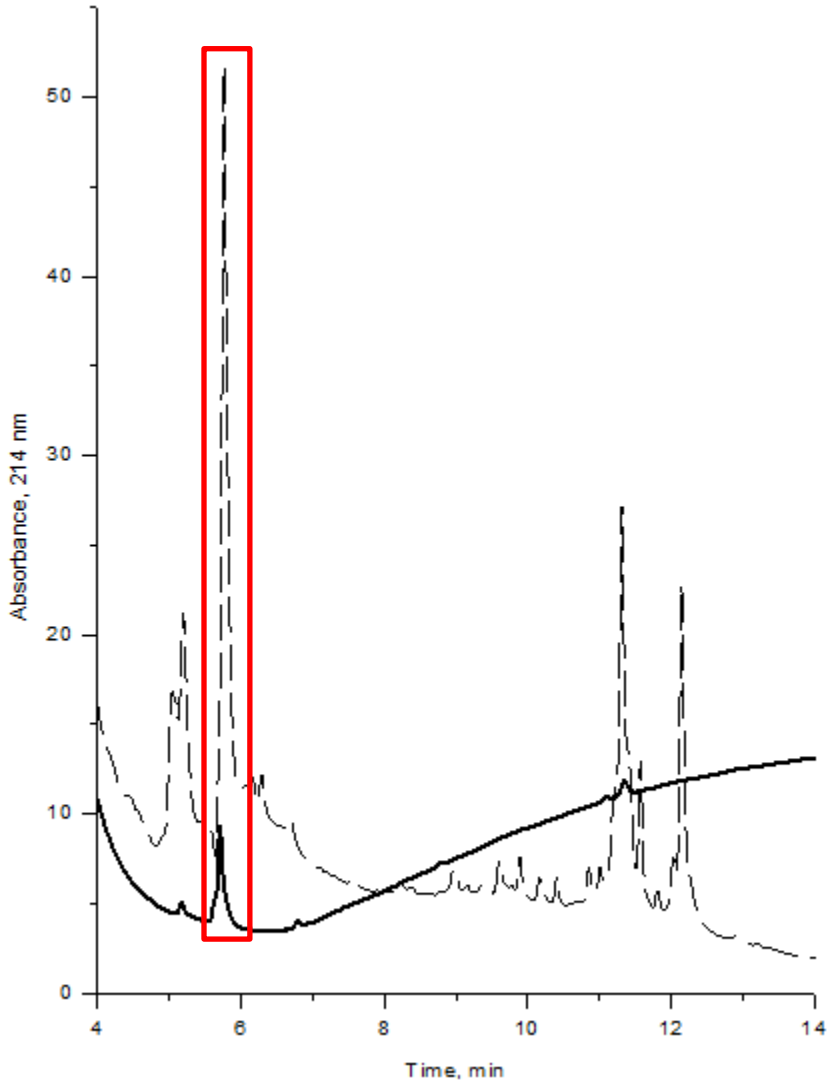
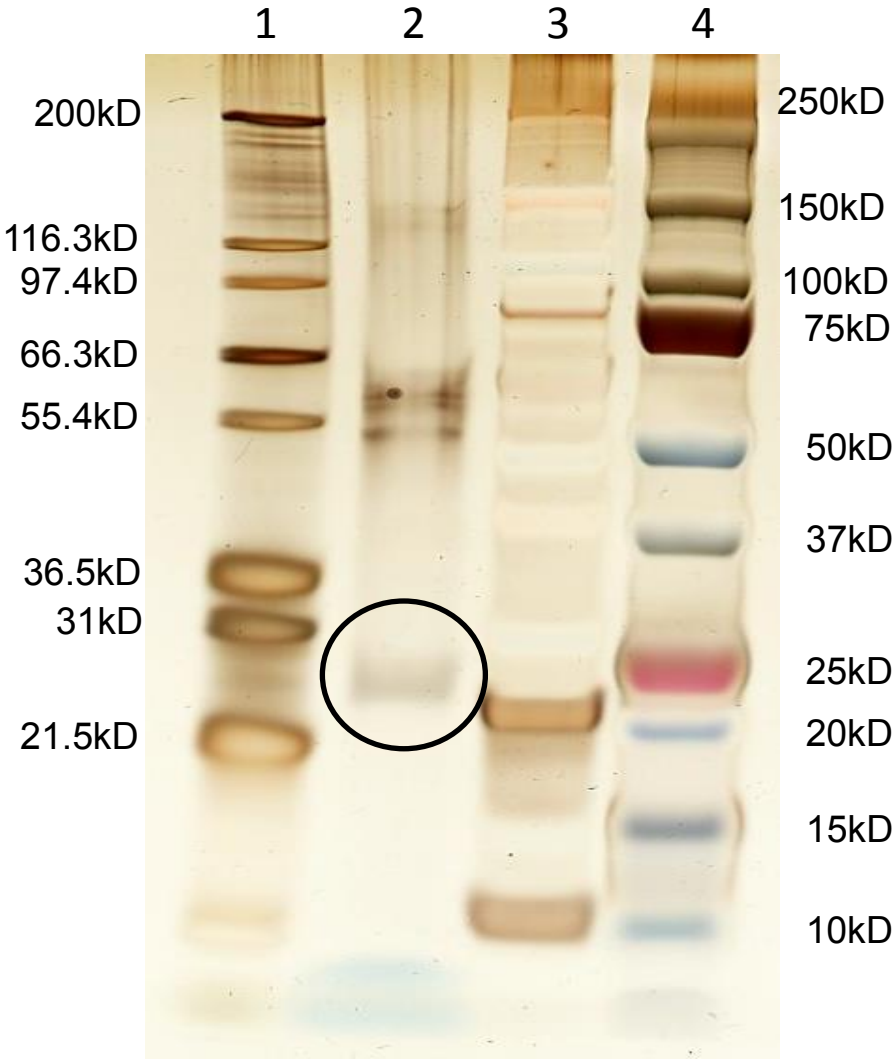


Figure 5.12 SDS-PAGE of collected peak of interest correlating to IVPD (lane 2 circled). Lanes 1, 3 and 4 contain Mw standards as indicated.



Chapter 6 - Summary

Hypothesis 1: The proteins of vitreous and floury sorghum endosperm will differ.

The proteins of the vitreous endosperm have a higher degree of cross-linking and a greater M_w distribution than those found in the floury endosperm.

Hypothesis 2: A method can be developed for determining the representative polymeric protein content of wild-type non-tannin sorghum varieties.

A method was developed that allowed acquisition of three unreduced protein fractions that represented proportionally different protein polymer contents as evidenced by comparative size exclusion chromatography.

Hypothesis 3: Sorghum protein in-vitro digestibility is involved with and influenced by protein polymer content and composition.

This hypothesis was found to be true. Significant negative correlations were found between different extracted protein fractions representing different protein polymer contents. Polymeric proteins from groups with greater M_w distribution appeared to have the greatest impact on in-vitro protein digestibility.

Hypothesis 4: Application of 2-D SEC x RP-HPLC analysis to sorghum protein polymers determined to be correlated with in-vitro protein digestibility will allow determination of monomeric compositional differences between polymers.

Protein polymers were collected and covalently disassembled to monomeric constituents for 2-D analyses thus allowing differences related to *in-vitro* digestibility to be discerned, and provided a new basis for rationalizing sorghum protein body structure.

Appendix A - Supplemental material

Figure A.1 A speculative alternative protein body model to the classic zein-based model. It is based on data from the current study showing kafirin subclass correlation to IVPD is dependent on the selectively fractionated protein polymers they are a part of, rather than simply on total subclass content.

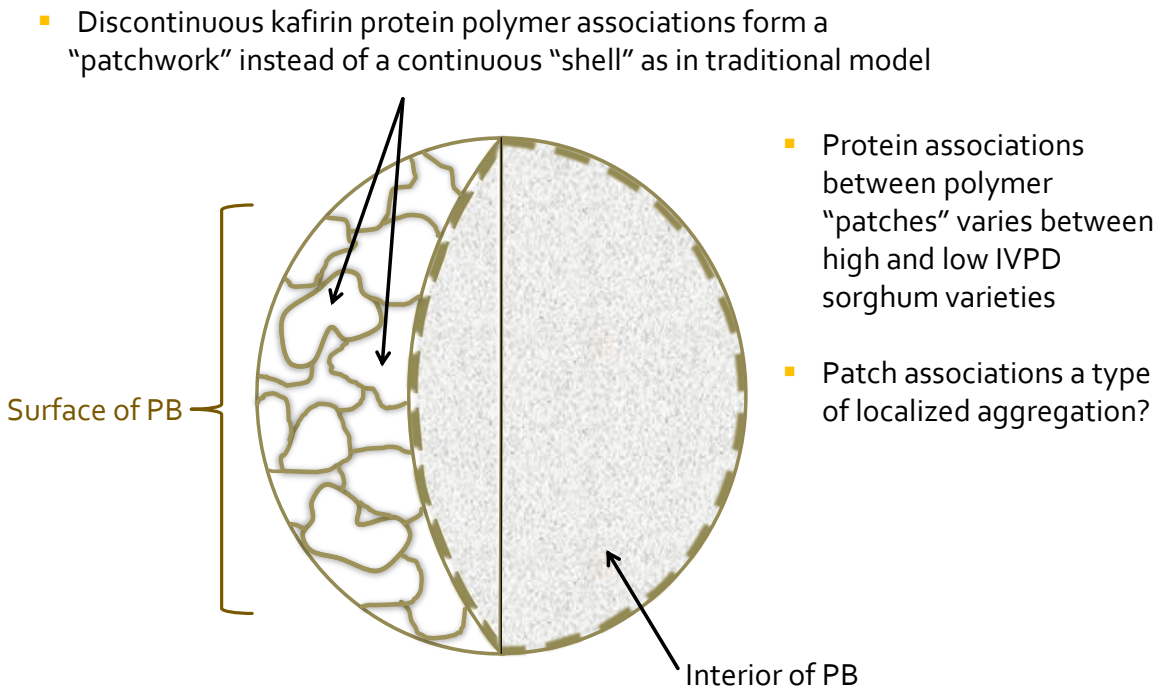


Figure A.2 Future sorghum research suggested by the results of the current study.

- "3F" fractionation applied to isolated floury and vitreous endosperm
- "3F" fractionation applied to isolated protein bodies
- Positively identify individual proteins involved in cross-linking
- Compare polymeric protein in wild-type to HD mutants
- Investigate changes to protein structures after digestion and cooking by application of "3F" fractionation
- Use "3F" fractionation to compare polymeric protein content of diverse maize population to sorghum