

## Characteristics of Proteins from Normal, High Lysine, and High Tannin Sorghums

Vartan Guiragossian, Bakshy A. K. Chibber, Steve Van Scoyoc, R. Jambunathan, Edwin T. Mertz, and John D. Axtell\*

The purpose of this paper was to study the characteristics of proteins from normal, high lysine, and high tannin *Sorghum bicolor* (L.) Moench. Endosperm preparations were obtained from four inbred lines of sorghum representing a normal, low tannin variety (P-721-N), its mutagenically derived high lysine counterpart (P-721-O), an inbred Ethiopian variety high in lysine (IS-11167), and a high tannin line (IS-4225). Endosperm proteins were separated into five soluble fractions by the Landry-Moureaux method. Whole endosperms and their respective protein fractions were subjected to amino acid analysis. Polyacrylamide gel electrophoresis patterns were determined for the fractionated proteins. The high lysine endosperms had lower levels of kafirins (fractions II and III) than the lysine-deficient, alcohol-soluble protein fraction, when compared with the normal sorghum endosperm preparations. Both high lysine varieties contained elevated levels of albumins and globulins (fraction I) and glutelins (fraction V), which were the highest in lysine content. Differences from normal were observed in the distribution pattern of proteins from a high tannin sorghum. There were no significant differences among the constituent proteins of the identical fractions of these four varieties of sorghum as determined by gel electrophoresis. These results support the general hypothesis that genes affecting protein quality do so by changing the relative quantities of Landry-Moureaux fractions and not by changing the quality of proteins within these fractions.

In recent years a large and concerted effort has been mounted to enhance the nutritional quality of almost all agriculturally significant cereal grains. A considerable portion of this effort is directed toward improving cereal protein quality and is particularly aimed at attaining the most favorable levels of the essential amino acids in cereal grain proteins (Nelson, 1969). *Sorghum bicolor* (L.) Moench ranks fourth in the world production of cereal grains grown for human consumption and is a primary food source for the populations in the semiarid regions of Africa and Asia. The proteins in sorghum, like other cereal grains, may be characterized as albumins, globulins, prolamins, and glutelins (Jambunathan and Mertz, 1973). These four classes of proteins are distinguishable in all cereal grains on the basis of their solubility in water (albumins), salt solutions (globulins), alcohol (prolamins), and alkaline detergents (glutelins). The alcohol-soluble prolamins are extremely low in lysine, which is generally the first limiting amino acid in cereal grains (Mertz et al., 1964). Sorghums

are similar to other cereal grains in many of their nutritional characteristics and yet differ in several important respects, the foremost among which is the presence of tannins. These poorly characterized polyphenolic compounds are present in certain genotypes and adversely affect protein availability and digestion (McGinty, 1969; Jambunathan and Mertz, 1973; Axtell et al., 1975). Jambunathan and Mertz (1973) found that three high tannin varieties of sorghum gave significantly lower growth responses in weanling rats when compared with three low tannin varieties. In this paper we will attempt to examine the proteins of sorghum from a nutritional and biochemical viewpoint and determine the potential for improving sorghum grain protein quality by genetic screening and selection. The objective of this paper was to study the characteristics of proteins from normal, high lysine, and high tannin sorghums.

### MATERIALS AND METHODS

Four inbred lines of sorghum were selected: (1) P-721-N (normal), a low tannin line; (2) P-721-O (opaque), a high lysine variety derived from P-721-N by chemical mutagen treatment (Mohan, 1975); (3) IS-11167, an Ethiopian variety high in lysine (Singh and Axtell, 1973); and (4) IS-4225, a high tannin line. All four varieties were grown at the Agricultural Experiment Station in Puerto Rico during the 1974-1975 winter-spring season.

Department of Agronomy, Purdue University, West Lafayette, Indiana 47907 (V.G., S.V.S., J.D.A.), Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907 (B.A.K.C., E.T.M.), and Central Services Laboratory, ICRISAT, Hyderabad-5000016, AP, India (R.J.).

Table I. Nitrogen Distribution in Sorghum Endosperms<sup>a</sup>

Fraction	Variety			
	P-721-N	P-721-O	IS-11167	IS-4225
Percent protein (g/100 g of endosperm)	12.0	10.6	10.5	9.4
I (albumins and globulins)	9.0	28.6	23.1	6.2
II (kafirin)	25.1	9.9	10.7	10.2
III (crosslinked kafirin)	25.1	15.3	19.0	18.7
IV (glutelin-like)	6.8	4.1	4.8	9.4
V (glutelin)	34.0	42.1	42.4	55.5
Total N extracted, %	98.6	97.9	91.4	80.6

<sup>a</sup> Percent of soluble nitrogen.

One hundred seed weight was recorded by weighing 100 random seeds of each variety. Embryo and endosperm tissue of 100 whole kernels were separated by hand dissection, without removal of the pericarp, to measure their proportions by weight and for chemical analysis. The crude content of whole kernel, embryo, and endosperm samples were determined by micro-Kjeldahl nitrogen times the factor 6.25.

Approximately 2500 embryos were separated from each sorghum variety to provide endosperm for further analysis. The seeds were soaked in water (10-30 min), and the embryos were excised with the aid of a scalpel. Ten-gram quantities of endosperm were collected, air-dried, ground, defatted, air-dried again, and stored at -4 °C. Defatted samples were hydrolyzed and analyzed for amino acid content using an automatic Beckman Model 120C ion-exchange resin amino acid analyzer.

The endosperm proteins were separated into five fractions using the Landry and Moureaux (1970) procedure. The procedure was scaled up to fractionate proteins from 5-g endosperm samples (Misra et al., 1975). The first fraction contains the albumins and globulins, the free amino acids and small peptide fragments, and any other saline-soluble compounds. The second fraction contains the alcohol-soluble proteins and is called kafirin. The third fraction contains the proteins that are alcohol soluble after the disulfide bonds in the protein have been reduced with 2-mercaptoethanol and contains the kafirin-like proteins. The fourth fraction contains the proteins that are alkali soluble after the disulfide bonds are broken and have some of the characteristics of glutelin. The fifth fraction contains the true glutelin, which is a complex, high-molecular weight mixture of proteins that can be solubilized only by treatment with a reducing agent and a detergent, sodium dodecyl sulfate (SDS) at alkaline pH. Individual protein fractions and endosperms were hydrolyzed in acid (6 N HCl, 110 °C, 24 h) and amino acid determinations were made on an automated analyzer (Beckman) following the Spackman et al. (1958) method. Individual protein fractions were electrophoresed on polyacrylamide gels (8.5 and 10% acrylamide) in the presence of 2-mercaptoethanol (2-ME) and sodium dodecyl sulfate (SDS) according to the Weber and Osborne (1969) method as adapted by Misra et al. (1976).

## RESULTS AND DISCUSSION

It is necessary to look at the distribution of proteins in the endosperm to understand the marked changes that occur in the amino acid patterns with the introduction of the high lysine mutant genes. The protein distribution pattern of these sorghum endosperms is shown in Table I. These results show a similarity in nitrogen distribution

Table II. Amino Acid Composition of Defatted Sorghum Kernels, Endosperms, and Embryos

Amino acids <sup>a</sup>	Variety											
	Whole kernel				Endosperm				Embryo			
	P-721-N	P-721-O	IS-11167	IS-4225	P-721-N	P-721-O	IS-11167	IS-4225	P-721-N	P-721-O	IS-11167	IS-4225
Lysine	2.0	2.95	3.2	2.4	1.6	2.6	2.7	2.0	6.4	6.3	7.2	6.8
Histidine	2.2	2.3	2.2	2.1	2.3	2.3	2.4	2.2	3.1	3.1	3.6	3.5
Arginine	3.7	4.5	4.7	3.6	3.6	4.3	4.2	1.0 <sup>b</sup>	10.9	12.1	12.8	10.2
Aspartic acid	6.4	7.5	8.0	6.9	5.7	7.4	9.1	4.1	9.7	7.9	8.5	7.1
Threonine	3.1	3.3	3.4	3.1	2.5	3.3	3.8	3.3	5.1	4.5	4.0	4.9
Serine	4.3	4.2	4.5	4.2	3.2	4.2	5.0	4.5	3.7	4.9	5.1	5.9
Glutamic acid	19.2	20.1	17.3	20.1	21.6	21.5	22.5	24.8	17.1	17.6	17.7	17.9
Proline	8.5	7.6	7.7	7.4	11.3	8.2	9.2	8.3	5.9	5.1	7.7	6.8
Cysteine	1.6	1.5	1.2	1.1 <sup>b</sup>	1.6	1.6	1.7	1.0 <sup>b</sup>	1.5	1.9	1.7	1.4
Glycine	2.9	3.5	3.8	2.9	3.6	3.7	3.8	2.6	6.5	7.1	7.2	7.7
Alanine	9.2	8.4	7.9	8.3	10.5	9.3	10.2	10.1	6.3	5.9	7.4	8.1
Valine	5.1	5.1	5.2	4.7	5.6	5.6	6.1	5.0	6.9	7.4	7.2	7.3
Methionine	1.7	1.6	1.4	1.5	1.9	1.9	1.7	1.2 <sup>b</sup>	1.8	1.7	2.1	1.5
Isoleucine	3.8	3.9	3.9	3.8	4.6	4.1	4.9	4.3	3.3	3.3	3.8	3.9
Leucine	14.2	12.2	11.1	12.4	16.9	13.8	15.1	15.1	6.5	7.0	8.1	9.4
Tyrosine	4.6	4.2	4.0	4.1	5.3	4.3	5.0	4.7	3.4	3.4	3.9	4.2
Phenylalanine	5.5	4.9	5.0	4.8	6.5	5.6	6.3	5.5	4.3	4.4	5.1	5.2
Ammonia	3.7	3.2	2.7	4.8	6.6	3.8	4.7	3.7	2.1	2.6	2.6	1.8
Total recover.	101.5	100.7	97.2	98.2	114.9	107.5	121.4	103.4	104.5	106.1	115.7	113.6

<sup>a</sup> Amino acids expressed as g/100 g of protein (single analysis). <sup>b</sup> Low values not in agreement with analyses of similar low and high tannin sorghums by Jambunathan and Merts (1973). Their values are near those reported for P-721-N in this table.

Table III. Amino Acid Composition of Laundry-Moureaux Fractions<sup>a</sup>

	Lys	His	Arg	Asp	Thr	Ser	Glu	Pro	Gly	Ala	(Cys) <sub>2</sub>	Val	Met	Ile	Leu	Tyr	Phe	NH <sub>2</sub>	Total recov.	
F-I	P-721-N	4.5	2.1	8.5	8.4	3.3	4.0	11.5	3.6	4.9	5.1	1.9	4.6	1.8	2.6	5.0	2.6	3.2	5.0	82.6
	P-721-O	4.7	2.2	8.6	8.8	3.3	4.0	11.6	3.7	5.0	5.3	1.6	4.7	1.2	2.8	5.0	2.8	3.3	4.0	82.6
	IS-11167	3.8	1.7	5.3	6.8	2.0	2.6	8.8	4.7	3.5	4.2	1.0	3.6	0.7	2.5	4.1	2.3	2.7	5.5	65.8
	Corrected <sup>b</sup>	(4.8)	(2.1)	(6.6)	(8.5)	(2.5)	(3.2)	(11.0)	(5.8)	(4.4)	(5.2)	(1.2)	(4.5)	(0.9)	(3.3)	(5.1)	(2.9)	(3.4)	(6.8)	(82.2)
F-II	IS-4225	3.2	2.5	3.5	9.8	1.6	2.6	11.4	6.4	2.3	3.9	0.6	2.3	0.8	1.6	2.2	1.4	1.2	5.4	62.7
	Corrected	(4.2)	(3.3)	(4.6)	(12.9)	(2.1)	(3.4)	(15.0)	(8.4)	(3.0)	(5.1)	(0.8)	(3.0)	(1.1)	(2.1)	(2.9)	(1.9)	(1.6)	(7.1)	(82.7)
	P-721-N	0.1	0.9	1.1	6.2	1.9	3.4	24.8	8.5	1.2	10.0	0.4	4.0	0.6	4.1	17.0	4.2	5.9	4.4	98.7
F-III	P-721-O	0.3	0.8	1.4	6.4	2.2	3.6	23.5	8.5	1.4	9.8	0.5	4.0	0.5	3.9	16.0	4.4	5.4	3.0	95.6
	IS-11167	0.3	0.5	1.0	6.6	2.3	4.2	23.9	7.6	1.5	10.0	0.4	4.1	0.4	3.4	14.4	4.1	4.1	4.4	93.2
	IS-4225	0.3	0.3	1.2	7.0	2.2	4.0	24.8	9.3	1.5	10.0	0.4	4.4	0.5	4.2	15.2	4.5	6.0	3.2	100.3
F-IV	P-721-N	0.2	1.0	1.7	5.2	2.3	4.0	23.3	9.9	1.5	9.6	1.3	3.9	1.3	3.6	15.9	4.6	5.5	4.0	98.8
	P-721-O	0.1	0.7	1.3	4.5	1.7	2.6	22.4	8.5	1.3	8.7	0.7	3.8	1.2	3.6	15.9	4.0	5.0	3.5	89.5
	IS-11167	0.1	0.9	1.1	5.4	2.0	3.9	24.1	9.6	1.1	9.9	0.5	3.9	0.8	4.1	16.9	4.2	6.1	3.0	97.5
	IS-4225	0.1	1.7	1.3	4.5	1.4	2.4	22.8	6.1	1.1	9.3	0.1	3.9	0.5	3.9	16.7	3.7	5.7	4.4	88.5
F-V	P-721-N	1.3	5.8	4.0	3.2	3.9	3.7	16.2	14.0	5.3	4.0	1.0	5.3	1.1	2.6	7.7	2.8	2.3	3.9	87.9
	P-721-O	1.7	5.0	3.8	3.2	3.1	3.3	15.4	11.5	5.2	4.0	0.4	5.0	1.4	2.6	7.2	2.4	2.4	2.5	80.1
	IS-11167	1.4	3.1	3.0	4.0	2.9	4.4	11.4	8.2	4.0	3.3	0.3	3.8	1.0	2.2	5.4	2.2	2.1	1.5	64.2
	Corrected	(1.9)	(4.2)	(4.1)	(5.5)	(3.9)	(5.0)	(15.6)	(11.2)	(5.5)	(4.5)	(0.4)	(5.2)	(1.4)	(3.0)	(7.4)	(3.0)	(2.9)	(2.1)	(87.8)
F-V	IS-4225	0.6	1.1	2.3	3.5	2.3	3.1	11.1	8.4	3.7	3.3	0.4	3.4	1.1	2.6	4.6	2.3	1.9	2.0	57.7
	Corrected	(0.9)	(1.7)	(3.5)	(5.3)	(3.5)	(4.7)	(16.9)	(12.8)	(5.6)	(5.0)	(0.6)	(5.2)	(1.7)	(3.9)	(7.0)	(3.5)	(2.9)	(3.0)	(87.7)
	P-721-N	2.2	2.2	3.5	6.5	3.1	3.9	25.7	9.3	3.2	10.4	0.4	6.1	1.5	4.8	12.9	4.8	5.3	5.0	110.8
	P-721-O	2.8	2.1	4.1	6.8	3.6	4.6	25.7	8.0	3.5	10.4	0.4	6.8	2.4	3.1	12.7	5.2	5.3	5.8	113.3
F-V	IS-11167	2.0	1.5	2.7	5.6	1.9	2.6	17.8	8.5	2.8	8.1	0.3	4.7	1.4	3.6	11.6	3.0	4.6	4.4	86.8
	Corrected	(2.6)	(1.9)	(3.4)	(7.1)	(2.4)	(2.9)	(22.6)	(10.8)	(3.6)	(10.3)	(0.4)	(6.0)	(1.8)	(4.6)	(14.7)	(3.8)	(5.8)	(5.6)	(110.3)
	IS-4225	1.3	1.5	2.6	5.9	2.7	3.8	19.5	6.4	2.0	8.5	0.2	4.4	1.4	3.7	12.8	4.0	4.9	5.9	91.5
	Corrected	(1.6)	(1.8)	(3.1)	(7.1)	(3.3)	(4.6)	(23.6)	(7.7)	(2.4)	(10.3)	(0.2)	(5.3)	(1.7)	(4.5)	(15.5)	(4.8)	(5.9)	(7.1)	(110.5)

<sup>a</sup> Amino acids expressed as g/100 g of protein (single analysis). <sup>b</sup> Total recovery of P-721-N amino acids (last column) divided by total recovery of IS-11167 amino acids (last column) x IS-11167 amino acid level.

**Table IV. Molecular Weight Determination of Sorghum Proteins by SDS-Polyacrylamide Gel Electrophoresis<sup>a</sup>**

Protein fraction	Variety			
	P-721-N	P-721-O	IS-11167	IS-4225
I	70 000			
	<u>57 500</u>	<u>57 900</u>	<u>57 000</u>	
	<u>41 500</u>	<u>41 000</u>	<u>39 800</u>	
	<u>37 200</u>	<u>37 000</u>		37 000
		<u>22 900</u>	<u>23 000</u>	
II		<u>18 800</u>		<u>18 500</u>
		<u>12 500</u>	<u>13 000</u>	<u>12 700</u>
	<u>24 000</u>	<u>24 600</u>	<u>25 000</u>	<u>25 000</u>
	<u>22 300</u>	<u>23 000</u>	<u>23 300</u>	<u>22 800</u>
III	<u>51 800</u>			
	<u>23 000</u>	<u>23 100</u>	<u>24 000</u>	<u>24 000</u>
	<u>17 000</u>	<u>17 300</u>	<u>16 800</u>	<u>16 400</u>
IV				<u>65 500</u>
	<u>41 500</u>	<u>42 000</u>		<u>48 500</u>
	<u>24 500</u>	<u>25 000</u>	<u>24 000</u>	<u>24 300</u>
	<u>12 700</u>	<u>13 200</u>	<u>12 800</u>	<u>13 000</u>
V	<u>80 000</u>	<u>80 000</u>	<u>78 000</u>	
	<u>76 000</u>			
	<u>73 000</u>	<u>74 000</u>	<u>73 000</u>	<u>73 000</u>
	<u>24 500</u>	<u>25 100</u>	<u>26 000</u>	<u>25 000</u>
				<u>22 000</u>

<sup>a</sup> Molecular weights of the major protein bands in each fraction are underlined.

to those of high lysine maize endosperm proteins (Misra et al., 1976). A comparison of the isogenic P-721-N and P-721-O varieties shows the effect of the high lysine mutation on protein distribution. The mutation causes a threefold increase in the levels of albumin and globulin proteins, decreasing the levels of kafirin fractions (II and III) by approximately 50%. There is also a significant increase in the true glutelin proteins (fraction V). A comparison between P-721-O and IS-11167 shows that mutagenic alteration caused a redistribution of endosperm proteins such that their distribution pattern is almost identical with that found in the naturally occurring, high lysine variety. The protein distribution observed in the high tannin variety (IS-4225) is distinct from that of the normal variety (P-721-N), and though these represent genetically diverse material, the observed differences cannot be entirely attributed to the variation in their genetic background. High tannin varieties consistently show a decreased level of fraction I proteins when compared to normal, low tannin varieties (Jambunathan and Mertz, 1973). However, a more significant change is the considerable decrease in the kafirin fractions (II and III) coupled with a similar increase in the glutelin fractions (IV and V). The reason for the lower total percent nitrogen recovery in IS-4225 may be attributed to tannin interference during fractionation. It is of interest to note that most of the highly pigmented tannin material appears in fraction V. Detailed studies on the effect of tannins on protein distribution (Chibber et al., 1977) indicated that tannins were predominantly associated with the kafirin fractions, altering the solubility of those proteins such that the kafirin-tannin complexes behaved as true glutelin proteins.

Amino acid distribution in sorghum kernels, endosperm, and embryo are shown in Table II. The protein content of these kernels and endosperms, as well as their amino acid composition are within the range of values reported

by Axtell et al. (1975) for various sorghums from the world collection. There was little difference between the naturally occurring, high lysine variety (IS-11167) and the mutationally derived, high lysine variety (P-721-O) in terms of amino acid composition and protein content. There was an increase in lysine and arginine concentrations and a decrease in alanine, proline, and leucine concentrations in the high lysine seeds relative to P-721-N seeds. There was little or no difference among the amino acid compositions of embryos taken from normal, high lysine and high tannin sorghum varieties. The importance of embryo as a source of lysine should be noted. Similar results were obtained from normal and opaque-2 maize embryos (Nelson, 1969).

The amino acid composition of individual protein fractions is listed in Table III. Fractions I and V contained the most lysine rich proteins, an observation which is consistent with the high lysine phenotype of P-721-O and IS-11167, both of which contain much higher levels of these protein fractions. It is apparent that the albumins, globulins, and glutelins offer the best source of protein nutrition in sorghum. While none of the fractions have the minimum required levels of lysine, fraction I has the highest levels observed. In general, the amino acid levels in fractions I, IV, and V of IS-11167 and IS-4225 were lower than in P-721-N and P-721-O because of lower total recovery of amino acids. This has been corrected (See Table III).

Finally, to determine whether the high lysine and high tannin varieties of sorghum were associated with any distinct and unique class of proteins, individual lyophilized endosperm protein fractions were subjected to polyacrylamide gel electrophoresis. The results of these experiments are summarized in Table IV. An examination of the protein bands on densitometric tracings (not shown) revealed no significant differences among the constituent proteins in these endosperms. However, we wish to be cautious in interpreting these results in view of the great heterogeneity of the material examined, as well as the analytical limitations inherent in this technique. Perhaps the only significant difference observed was in the glutelin fraction V, where an additional band was associated with the fraction V proteins from the high tannin endosperm. This observation is consistent with the premise that tannin-kafirin complexes cofractionate with the glutelins. It is interesting to note the similarity between the two kafirin (fraction II) protein bands and those observed by Misra et al. (1976) of zein protein (fraction II) from maize under identical conditions.

#### ACKNOWLEDGMENT

We thank Gina H. Y. Lu and M. M. Hassen for technical assistance.

#### LITERATURE CITED

- Axtell, J. D., Oswalt, D. L., Mertz, E. T., Pickett, R. C., Jambunathan, R., Srinivasan, G., in "High Quality Protein Maize 374-386", Dowden, Hutchinson and Ross, Stroudsburg, Pa., 1975.
- Chibber, B. A. K., Mertz, E. T., Axtell, J. D., submitted to *J. Agric. Food Chem.* (1977).
- Jambunathan, R., Mertz, E. T., *J. Agric. Food Chem.* **21**, 692-696 (1973).
- Mertz, E. T., Bates, L. S., Nelson, O. E., *Science* **145**, 279-280 (1964).
- McGinty, D. D., 6th Biennial Grain Sorghum Research and Utilization Conference, Grain Sorghum Producer's Association, 1969.
- Misra, P. S., Mertz, E. T., Glover, D. V., *Cereal Chem.* **52**, 161-166 (1975).
- Misra, P. S., Mertz, E. T., Glover, D. V., *Cereal Chem.* **53**, 705-711 (1976).

