β-Conglycinin and Glycinin in High-Protein Soybean Seeds

Robert W. Yaklich

Soybean and Alfalfa Research Laboratory, Building 006, Room 201, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, Maryland 20705

The agronomic performance and storage proteins of high seed protein lines of soybeans [*Glycine max* L. (Merr.)] were investigated to determine if the two major storage proteins, β-conglycinin and glycinin, contribute to the increased protein content of high seed protein lines. Subunits of these two major storage proteins were estimated by scanning SDS-PAGE gels by scanning densitometry. The relative rankings of the lines with respect to seed size and protein content were not different between years in one environment over 5 years, but oil and total protein and oil contents and the ratio of protein to oil differed. The α′, α, and β subunits of β-conglycinin were significantly higher in the high-protein lines except CX797-115, CX804-108, CX804-3, D81-8498, and NC-2-62. The acidic A_3 polypeptide of glycinin was significantly higher in high-protein lines except 76-48773, CX804-108, CX804-3, D81-8498, and NC-2-62, whereas the acidic polypeptides $A_{1,2,4}$ of glycinin were significantly higher in all of the high-protein lines. The basic polypeptides of glycinin were significantly higher in all high seed protein lines except D81-8259. In conclusion, high-protein lines appear to contain more β-conglycinin and glycinin than normal-protein soybean lines, and the amounts of subunits and polypeptides differ among lines.

Keywords: *Glycine max L. (Merr.); β-conglycinin; glycinin; high-protein seeds; electrophoresis*

INTRODUCTION

Soybeans seeds contain more protein than any other cultivated commercial crop. Approximately 41% of the dry weight of the soybean seed is storage protein, and 21% is oil. The primary products resulting from processing soybean seed are crude soybean oil and protein meal. The oil is refined and used mainly for human consumption, and the protein meal is a source of high-quality protein for animal husbandry.

Historically, soybeans have been bred to increase seed yield and oil content. More recently, with an increase in the worldwide standard of living and a concomitant increase in the consumption of meat, the demand for protein in animal husbandry has increased. Smith (*1*) concluded that 65% of the value of soybean seeds is now attributable to the protein fraction.

Soybean accessions exist that have seed protein content >50%. Breeders have used these accessions and released high seed protein lines in which the protein component exceeds 50% of the seed dry weight (*2*, *3*). However, two major obstacles have hindered the development of high-protein soybeans for commercial use. Soybean protein is negatively correlated to seed oil and to yield (*4*). These correlations are deterrents to the commercial use of high seed protein lines because soybeans are bought and sold on a weight basis and the high seed protein lines yield less and contain less oil.

The two main seed storage proteins in soybean are glycinin and β-conglycinin (*5*, *6*). The glycinins are 11S hexamers with molecular masses reported to be between 320×10^3 and 375×10^3 (*7*, *8*). Each hexamer is composed of two trimers that are rotated 60° with respect to each other to form a triagonal antiprism (*9*, *10*). The trimers consist of three monomers. The monomers consist of subunits that are composed of a specific acidic (A) polypeptide chain (M_r 40000) linked by disulfide bonding to a specific basic (B) polypeptide chain (M_r 20000) and can be one of any of five subunits ($A_{1a}B_{1b}$, A_2B_{1a}, $A_{1b}B_{1a}$, $A_5A_4B_3$, and A_3B_4) (*5*, *6*). The β-conglycinins are 7S trimers with molecular masses of $(150-175) \times 10^3$ (*11*, *12*). They are formed by various combinations of three nonidentical but homologous polypeptide subunits: the α′, α, and β subunits (*13*) with M_r = 76000, 72000, and 53000, respectively (*14*). The two major storage proteins do not contain many sulfur amino acids, although glycinin contains more than β-conglycinin.

Wilcox and Cavins (*15*) found that the protein component of the soybean seed could be increased without significantly affecting yield. This realization has caused an increased interest in high seed protein lines because the negative relationship between yield and protein apparently can be moderated. Relatively little is known about which storage proteins are increased in high seed protein lines. We recently found (*16*) that the quantity of a major soybean seed allergen (P^{34}) was not increased in high seed protein lines relative to normal-protein lines. However, P^{34} accounts for only 1−2% of the seed storage protein. Alternatively, the two major storage proteins (glycinin and β-conglycinin) constitute ∼70% of the storage proteins in normal soybean lines. Therefore, a study was undertaken to determine if an increase in these storage proteins could be detected by using scanning densitometry and immunoblotting to estimate the subunit quantity of β-conglycinin and acidic and basic polypeptide content of glycinin in normal and high seed protein lines.

MATERIALS AND METHODS

High seed protein lines were obtained from several soybean breeders. Seed for the northern-type high seed protein lines, CX797-21, CX797-115, CX804-3, and CX804-108, were ob-

10.1021/jf001110s
Published on Web 12/21/2000

Table 1. Means over Years for Seed Size, Protein, and Oil Content of 19 Normal and High Seed Protein Soybean Lines[a]

year	observations	100-seed wt, g	seed protein, g kg^{-1}	seed oil, g kg^{-1}	total protein and oil, g kg^{-1}	protein/oil ratio
1994	19	15.2 a	486.1 a	173.7 b	659.8 b	2.96 b
1995	19	15.3 a	485.7 a	183.4 a	669.1 a	2.77 b
1996	19	14.9 a	481.8 a	174.1 b	655.9 b	2.91 b
1997	19	14.8 a	482.2 a	160.4 c	642.6 c	3.21 a
1998	19	15.9 a	472.1 a	169.3 b	641.4 c	2.90 b

[a] Means followed by the same letter within the same column are not significantly different at the 0.05 probability level.

tained from J. R. Wilcox, USDA-ARS, Purdue University, and 76-48773 was obtained from A. Matson, Soybean Research Foundation. Seed for the southern-type high protein lines, D76-8070, D80-6931, D81-8259, and D81-8498, were obtained from E. E. Hartwig (deceased) USDA-ARS, Jamie Whitten Delta States Research Center, and NC-2-62 was obtained from J. W. Burton, USDA-ARS, North Carolina State University. The high seed protein lines, BARC-6, BARC-7, BARC-8, and BARC-9, were from R. C. Leffel, USDA-ARS, Beltsville, MD. Several normal seed protein lines (Essex, Williams, Miles, and Clark) were included in the study. The non-nodulated line, Clark *rj1*, which is nitrogen deficient during seed fill, was also included.

The soybean lines listed above were grown using normal agronomic procedures (*17*), for 5 years (1994−1998) at Beltsville, MD. Samples of harvested seed were stored at 4 °C until analyzed. Seed size was determined gravimetrically from three 100-seed samples of each line. Protein and oil content were determined for a single sample of each line by infrared reflectance by the National Center for Agricultural Utilization Research, USDA-ARS, Peoria, IL.

Seeds grown in 1996 were used to estimate β-conglycinin and glycinin contents. Ten seeds from each line were ground in a Wiley mill and passed through a 30-mesh screen. Seed powder was weighed (21 mg) in an Eppendorf tube to which was added 840 μL of extraction buffer [0.17 M sodium dodecyl sulfate, 6.0 M urea, 0.57 mM 2-mercaptoethanol, and 0.05 M tris(hydroxymethyl)aminomethane, pH 6.8]. The protein was extracted by treatment for 45 min in an ultrasonic water bath.

Electrophoresis SDS-PAGE was carried out according to the procedures of Laemmli (*18*) in 1.5 mm thick gels with 12.5% (w/v) separating gels and 4% (w/v) stacking gels. The gels contained 15 wells; one of the outside wells was used for molecular weight standards. The ratio of the storage proteins of the high seed protein lines relative to the cultivar Essex was determined in this manner. The two wells next to the molecular weight standard and the last two wells at the opposite end of the gel contained protein extract from Essex (the control soybean line, used for comparison and ratio calculations). Each group of two of the remaining inside wells received extract from a specific soybean line, allowing five lines to be compared to Essex in each gel. A 12 μL sample of the extract was loaded onto each well. The amounts loaded were equivalent with respect to weight of the ground seed. The gels were run in a buffer solution [0.19 M glycine, 0.1% sodium dodecyl sulfate, 0.025 M tris(hydroxymethyl)aminomethane, pH 8.3] for 4−5 h to completion and then were stained with amido black 10B (1.0 g/L). The protein bands on the destained gels were quantitated in a Molecular Dynamics scanning densitometer. Lanes 3 and 14 were used to obtain a value for the protein bands of Essex. Relative mobilities from published photographs of soybean seed SDS-PAGE gels (*19*−*21*) were used to identify the protein bands that were the subunits of β-conglycinin and the acidic and basic polypeptides of glycinin. The amount of protein was determined relative to the amount present in the cultivar Essex. To quantitate the various protein bands from each soybean line, the area of a specific protein band of a soybean line was compared with the area of the protein band for Essex and a value (ratio) relative to Essex was obtained. Four, 10-seed replications for each line were run and quantitated in this manner.

For immunostaining, gels were prepared as above and the fractionated proteins were electroblotted to Immobilon-P transfer membranes (Millipore Corp., Bedford, MA) in a 10 mM Tris−sodium borate 20% (v/v) methanol, pH 8.2, transfer buffer (*22*). Blots were either stained with amido-black to visualize proteins or immunolabeled. For immunolabeling, air-dried blots were wetted in methanol and washed several times in transfer buffer, washed in TBS [Tris-buffered saline, 100 mM tris(hydroxymethyl)aminomethane, 150 mM NaCl, pH 7.5], and washed and blocked in TBS containing 5% w/v powdered milk for 3 h. Antibodies against glycinin and β-conglycinin, which were provided by Dr. Anna L. Tan-Wilson, were prepared as described in Wilson et al. (*23*). For immunolabeling a 1:2500 dilution of the anti-glycinin serum and a 1:12500 dilution of the anti-β-conglycinin serum were added to the blots in TBS containing 5% powdered milk overnight and washed in TBS. Immunoreactive bands were visualized by labeling the blots in TBS containing 5% powdered milk with a 1:2000 dilution of anti-rabbit immunoglobulin G (IgG) alkaline phosphatase and development with chromogenic substances 5-bromo-4-chloro-3- indolyl phosphate and nitroblue tetrazolium.

The variables including seed size, protein, oil, total protein and oil, and the ratio of protein to oil were analyzed as two-factor main effects as a general linear model using PROC MIXED (*24*). The fixed effects were cultivar and year. Slight variance heterogeneity occurred in the analysis of the protein-to-oil ratio data. To correct for this, the data were transformed using the natural logarithm and then analyzed. Means were compared using pairwise contrasts and are reported in their original units. Means were compared using the LSD ($P < 0.05$).

The values of the dependent variables (seed storage protein bands) were standardized to the values of the soybean line Essex as the relative variable = (variable value/Essex value) × 100. The standardized variables were then analyzed by mixed model analysis of variance techniques (*24*) to determine the fixed effects of soybean lines with four experimental replications (random effect). Assumptions of the analysis of variance were examined and found to be satisfactory. Pairwise means comparisons for significance between Williams and all other lines were carried out.

RESULTS AND DISCUSSION

Soybean breeders have found that genetic traits to increase the seed storage proteins are easily transferred to acceptable agronomic lines suitable for cultivation. However, problems remained with yield depression and reduced oil content in the high-protein lines. Recently, Wilcox and Cavins (*15*) demonstrated that both yield and protein content can be increased in a soybean line, indicating that yield depression can be moderated.

Table 1 illustrates that no significant year-to-year differences in seed size and protein content occurred over the five years of field testing. Likewise, no significant year interaction occurred in the analysis of variance (ANOVA) for either variable (data not presented), indicating that the relative individual protein content and seed size of the 19 lines tested were stable in the one environment from year to year. Piper and Boote (*25*) found that cultivars maintain their ranking across environments with respect to seed oil and protein composition by studying results from check cultivars used in the Soybean Uniform Tests. The results pre-

Table 2. Means over 5 Years for Seed Size, Protein, and Oil Content of 19 Normal and High Seed Protein Soybean Lines[a]

line	observations	100-seed wt, g	seed protein, g kg^{-1}	seed oil, g kg^{-1}	total protein and oil, g kg^{-1}	protein/oil ratio
BARC-6	5	18.2 a	508.0 cde	159.8 f	667.8 b−e	3.21 cd
BARC-7	5	15.3 cd	493.6 de	147.0 g	640.6 ghi	3.39 cd
BARC-8	5	13.0 fg	545.2 a	134.6 h	679.8 abc	4.07 a
BARC-9	5	15.4 c	540.0 ab	141.8 gh	681.8 ab	3.83 ab
CX797-115	5	18.1 a	502.8 cde	184.0 e	686.8 a	2.74 ef
CX797-21	5	18.3 a	502.8 cde	161.8 f	664.6 cde	3.11 de
CX804-108	5	15.3 cd	495.2 de	178.4 def	673.6 a−d	2.79 ef
CX804-3	5	17.1 ab	507.0 cde	160.2 f	667.2 b−e	3.20 cd
D76-8070	5	17.5 ab	485.8 ef	159.0 f	644.8 fgh	3.08 de
D80-6931	5	13.9 def	512.4 cd	145.2 gh	657.6 def	3.57 bc
D81-8259	5	11.8 g	519.0 bc	136.6 gh	655.6 efg	3.87 ab
D81-8498	5	12.5 fg	512.4 cd	145.6 gh	658.0 def	3.54 bc
NC-2-62	5	12.2 g	520.8 bc	146.8 g	667.6 b−e	3.60 bc
76-48773	5	16.0 bc	466.2 f	189.8 de	656.0 efg	2.46 fg
Clark	5	16.0 bc	419.2 g	212.8 b	632.0 hi	1.97 h
Clark *rj1*	5	13.0 fg	342.4 h	250.4 a	592.8 j	1.37 i
Essex	5	13.8 ef	429.4 g	197.8 cd	627.2 i	2.17 gh
Miles	5	14.7 cde	423.4 g	206.4 bc	629.8 hi	2.05 h
Williams	5	17.2 ab	424.6 g	213.2 b	637.8 hi	1.99 h

[a] Means followed by the same letter within the same column are not significantly different at the 0.05 probability level.

sented here, which concur with those findings, indicate that specific end uses that require norms of seed size and protein content are attainable with soybean lines that contain high seed protein content.

A strong negative correlation exists between protein and oil content in soybean seeds (*4*). Normal seed protein lines generally have an average oil content of >200 g kg^{-1}. The average value of the oil content of the lines tested (Table 1) differed between years and was less than the oil content of a normal seed protein line. The ANOVA also showed a highly significant year effect on oil content (data not presented). The oil content of the seed is partially affected by the environment. Howell and Cartter (*26*) found that under controlled conditions, plants at 29 °C produced seeds with oil percentages 2−3% higher than that of plants grown at 22 °C. Howell and Collins (*27*) showed that changes in temperature affected the composition of the oil.

The total protein and oil content and the protein-to-oil ratio also differed between years (Table 1). Average total protein and oil content of normal soybean lines is ∼610 g kg^{-1} (unpublished data, R. W. Yaklich). The values for the high-protein lines presented here are 3−5 percentage points higher than those of normal soybean lines. Normal soybean lines have a protein-to-oil ratio of 1.99 (unpublished data, R. W. Yaklich), whereas the lines tested here had a protein-to-oil ratio approaching 3.0. The data presented in Table 1 indicate that the protein content of a line is more stable than the oil content from year to year.

The average 100-seed weight of the individual lines (Table 2) established that CX797-21 was the largest seeded followed by BARC-6 and CX797-115. The smallest seed size was D81-8259 preceded by NC-2-62. These results indicate that the range in seed size that exists for the high seed protein lines may be useful for cultivar improvement or specialty product applications. The average protein content of most of the high seed protein lines was >500 g kg^{-1} (Table 2), as compared to 420 g kg^{-1} for the normal seed protein cultivars and 340 g kg^{-1} for the non-nodulating line Clark *rj1*. BARC-8 had the highest 5-year average protein content of 545.2 g kg^{-1} or ∼125.0 g kg^{-1} more seed protein (30.0%) than the normal cultivars. Oil content varied from a high of 250.4 g kg^{-1} for Clark *rj1* to a low of 134.6 g kg^{-1} for BARC-8. As mentioned previously, year-to-year variability occurs in oil content that can be attributable to environmental causes. Total protein and oil content was highest in CX797-115 followed by BARC-9 and BARC-8 and lowest in Clark *rj1*, the non-nodulating line. Even though a substantial difference exists between the protein content of high seed protein lines and commercial cultivars, the increase in total protein and oil content present in the high seed protein lines compared to the normal seed protein lines is only ∼40 g kg^{-1}. In normal seed protein lines, the protein-to-oil ratio is 2 to 1. The data in Table 2 for the protein-to-oil ratio in the high protein seed lines indicates that most of the lines produce 3 units of protein to 1 unit of oil. Furthermore, over 5 years, BARC-8 averaged 4 units of protein to 1 unit of oil.

An SDS-PAGE gel of the high-protein lines and Essex (Figure 1) illustrates that the protein bands were similar in the soybean lines. Two bands [M_r of approximately 28400 (upper) and 27300 (lower)] were either present or absent in a particular soybean line. The α′, α, and β subunits of β-conglycinin have M_r values of approximately 76000, 72000, and 53000, respectively. The acidic 3, acidic 1, 2, and 4 (grouped together), and basic polypeptides of glycinin have M_r values of approximately 43000, 40000, and 20000, respectively.

The most logical reason for increased seed protein apparently results from increased quantities of the two main storage proteins, glycinin and β-conglycinin. The results of the scanning densitometry, presented as least-squares means in Table 3, show that of the high-protein lines, D81-8498 contained the greatest amount of α′ and α subunits of β-conglycinin and CX804-3 contained the least. All of the high-protein lines except CX797-115, CX804-108, and CX804-3 contained more of these subunits than Williams. The β subunit of β-conglycinin was highest in D81-8259, and in the high-protein lines CX804-3, D81-8498, and NC-2-62 were not significantly different from Williams.

The values for glycinin (Table 3) indicate that for the acidic 3 (A_3) polypeptide of glycinin, BARC-9 had the highest value, and the high-protein lines 76-48773, CX804-108, CX804-3, D81-8498, and NC-2-62 were not significantly different from Williams. The acidic 1, 2,

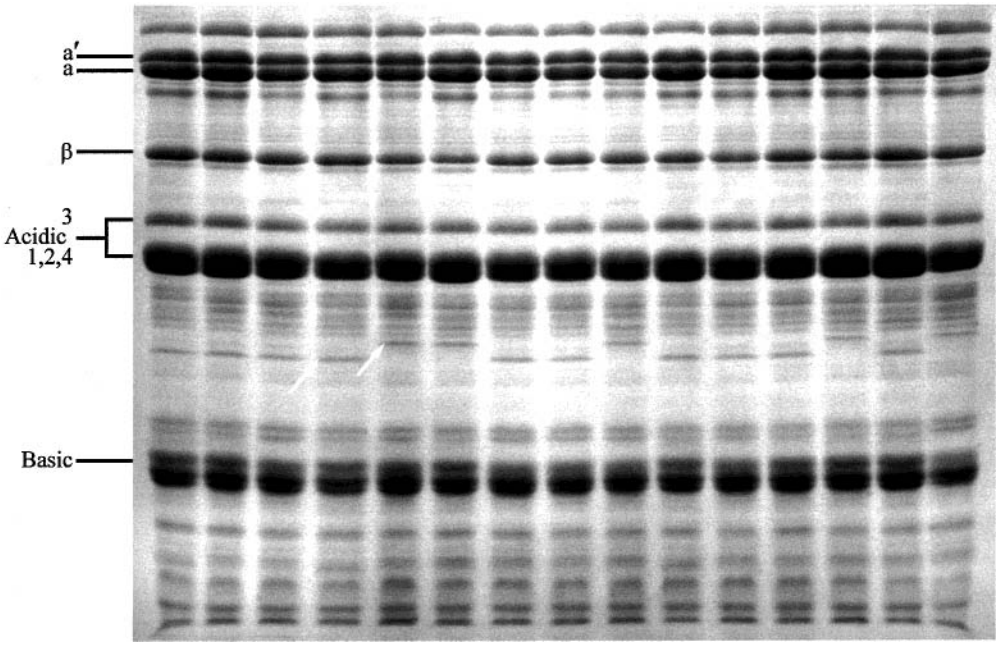

Figure 1. SDS-PAGE gel of the high-protein and Essex soybean lines. The principal protein subunits of β-conglycinin are the α′, α, and β subunits, and the principal polypeptides of glycinin are the acidic A_3 and $A_{1,2,4}$ and basic polypeptides. The lanes contain extract from equivalent weights of seed tissue and represent, from left to right, D80-6931, D76-8070, CX804-108, D81-8259, CX804-3, NC-2-62, CX797-115, CX797-21, 76-48773, D81-8498, BARC-6, BARC-7, BARC-8, BARC-9, and Essex, respectively. The white arrows point to two protein bands of M_r 28400 (upper) and 27300 (lower). The soybean lines contained either one or the other protein band.

Table 3. Least-Squares Means for the Area of the α′, α, and β Subunits of β-Conglycinin and the Acidic 3, Acidic 1,2,4, and Basic Subunits of Glycinin Seed Storage Proteins of Soybean Lines[a]

	%				
	β-conglycinin subunits		glycinin polypeptides		
line	α′ and α	β	A_3	$A_{1,2,4}$	basic
BARC-6	100.4*	110.3*	143.6*	125.9*	124.7*
BARC-7	119.8*	108.1*	145.4*	124.9*	118.3*
BARC-8	107.1*	95.5*	134.8*	135.7*	127.1*
BARC-9	108.6*	111.8*	154.9*	156.3*	139.8*
CX797-115	96.4	98.1*	132.0*	109.7*	110.4*
CX797-21	110.3*	111.1*	132.3*	126.5*	128.7*
CX804-108	91.3	101.4*	126.4	116.7*	121.4*
CX804-3	83.3	88.5	119.6	110.9*	115.8*
D76-8070	118.7*	110.3*	136.5*	134.1*	131.4*
D80-6931	113.0*	115.3*	149.5*	141.2*	114.8*
D81-8259	111.0*	122.3*	136.0*	120.3*	89.1
D81-8498	126.6*	88.3	126.0	122.8*	110.3*
NC-2-62	108.2*	82.0	123.3	126.9*	113.0*
76-48773	107.8*	104.4*	125.1	109.4*	111.1*
Clark	73.2	68.2	96.7	67.8	83.5
Clark *rj1*	53.6*	47.1*	70.1*	50.7*	53.8*
Miles	78.7	75.0	97.1	74.8	86.1
Williams	82.4	73.2	105.1	75.0	89.0

[a] The asterisks indicate significance between the value and that of Williams at $P = 0.05$ according to the t test.

and 4 ($A_{1,2,4}$) polypeptides of all the high-protein lines were significantly different from those of Williams. Comparing the high-protein lines, BARC-9 contained the most and 76-48773 contained the least of the $A_{1,2,4}$ polypeptides. All of the high-protein lines contained significantly more of the basic polypeptides of glycinin than Williams with the exception of D81-8259. BARC-9 contained the most basic polypeptides of glycinin.

The above results indicate that glycinin and β-conglycinin are increased in the high seed protein lines. Comparison of the values for the subunits and polypeptides with protein data in Table 2 indicates that D81-8498 and D81-8259 were the highest for the α′ and α and β subunits of β-conglycinin but were the sixth and fourth highest in protein content, respectively. For the polypeptides of glycinin, the acidic and basic polypeptides were highest in BARC-9, which was the second highest in protein content (Table 2). In contrast, BARC-8, which was highest in protein content was 10th, 11th, 7th, 3rd, and 4th in α′ and α, β, A_3, A_{124}, and basic protein contents, respectively. The high-protein seed line with the lowest seed protein content, 76-48773, was 9th, 8th, 12th, 14th, and 12th, in α′ and α, β, A_3, A_{124}, and basic protein contents, respectively. The data indicate that glycinin and β-conglycinin are increased in the high-protein lines; however, estimating the quantity of subunits and polypeptides of these two protein lines did not agree totally with the total protein content of the seed. BARC-9 was second highest in total protein content (Table 2) and contained the highest amounts of subunits of glycinin, indicating that the SDS-PAGE method agreed with the 5-year mean of protein content. The line 76-48773 contained the least total protein of the high seed protein lines, and its glycinin polypeptides were among the lowest of the high seed protein lines. In contrast, BARC-8 had the highest 5-year average protein level, and the subunit and polypeptide values were not indicative of its protein content. These discrepancies could be due to experimental error in the procedure. The acrylamide gels were prepared using standardized procedures and exact times of polymerization and staining in order to reduce the variability inherent in an analysis consisting of multiple gel determinations (*28*). It is known that the quality and quantity of subunits and polypeptides may differ in β-conglycinin and glycinin (*29*). Serretti et al. (*30*) noted that the BARC-8 and BARC-9 lines differed from each other in amino acid content but that these differences were not significant. The BARC lines differed from normal-protein cultivars in that they contained more

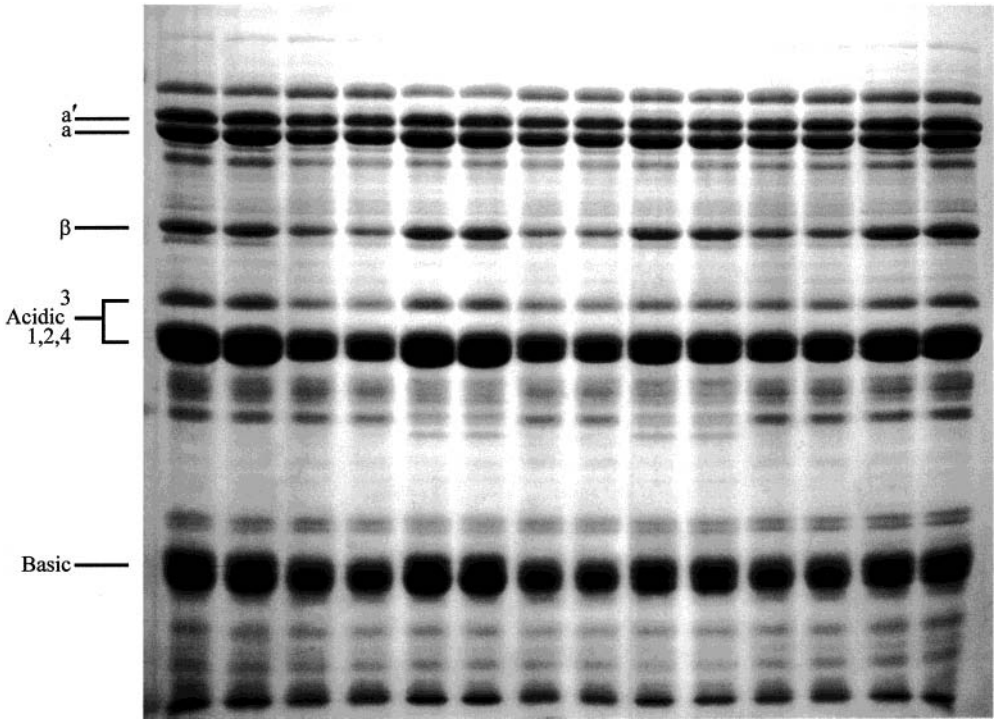

Figure 2. SDS-PAGE gel of soybean lines. The principal protein subunits of β-conglycinin are the α′, α, and β subunits, and the principal polypeptides of glycinin are the acidic A_3, $A_{1,2,4,}$ and basic polypeptides. The lanes contain extract from equivalent weights of seed tissue and represent, from left to right (duplicate lanes), BARC-8, Williams, BARC-9, Williams, CX797-115, Williams, and 76-48773.

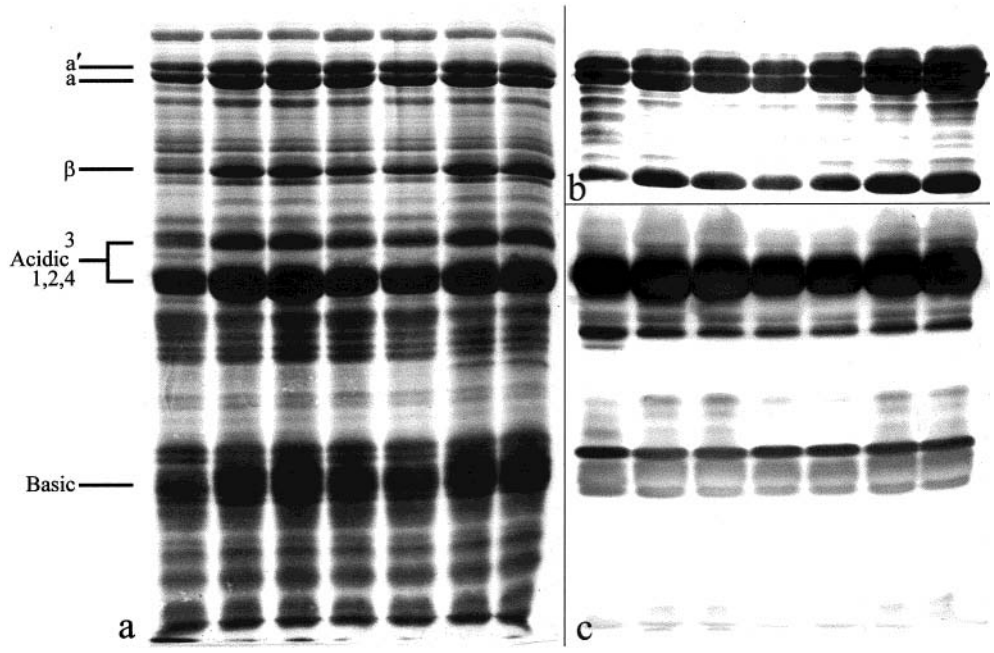

Figure 3. Immunoblot analysis of seed extracts of soybean: (A) amido-black stained blot of total proteins; (B) replicate blot that treated with anti-β-conglycinin seruml (C) replicate blot that treated with anti-glycinin serum. Lane 1 represents Clark *rj1* followed by duplicate lanes of BARC-8, Williams, and BARC-9.

arginine and glutamic acid. Whether these differences in amino acid content account for the levels of β-conglycinin and glycinin measured in BARC-8 is not known

Figure 2 illustrates an SDS-PAGE gel of a few of the high seed protein lines and the cultivar Williams stained with amido-black. Comparison of bands for the subunits of β-conglycinin and acidic and basic polypeptides of glycinin shows that those for BARC-8 and BARC-9 are wider and more intensely stained than those for Williams. Wider and more intensely stained bands of β-conglycinin are also observed for CX797-115 and 76-48773 as compared to Williams, but differences are not as easy to visualize for the acidic and basic polypeptides of glycinin. Western blots of the lines Clark *rj1*, Williams, BARC-8, and BARC-9 stained with amido-black (Figure 3a) also show that the width and intensity of the bands for β-conglycinin and glycinin are more pronounced in BARC-8 and BARC-9; similarly, the

immunoblots from the two proteins for these two high-protein lines (Figure 3b,c) are more intense than those from Clark *rj1* and Williams. For some reason, the antisera for glycinin did not detect the A_3 subunit and showed low affinity for the acidic polypeptides. These indirect methods agree with the results from scanning densitometry and indicate that β-conglycinin and glycinin, the two major seed storage proteins, are increased in high seed protein lines and contribute to the increased protein content of these lines.

The data presented in this study indicate that quantitative differences exist in protein bands of two major seed storage proteins, β-conglycinin and glycinin, that were analyzed in the high seed protein lines. These results indicate that there are differences in the genome that are responsible for synthesis of different quantities of subunits and polypeptides. Soybean proteins, and other types of legume seed proteins, are known to be deficient in sulfur amino acids, and β-conglycinin contains less sulfur amino acid than glycinin (*29*). The data indicate that some of the BARC high seed protein lines, which have a higher percentage of glycinin polypeptides than most of the other high seed protein lines, would be beneficial in breeding programs designed to change seed storage protein composition.

The types of amino acids present in the BARC lines have been analyzed previously (*30*). The concentration of cysteine of BARC-8, 13.7 g kg^{-1} of protein, was higher than that of any other entry. The methionine concentration of BARC-7 (10.8 g kg^{-1} of protein) and the lysine concentration of BARC-9 (55.0 g kg^{-1} of protein) were significantly lower than the two normal protein lines used as controls. Subunits of glycinin differ in the amount of sulfur amino acids they contain (*31*). The results of Serretti et al. (*30*) indicate that the high-protein lines may differ in subunit composition and also that the concentration of particular subunits is a heritable trait that can be transferred into high-protein lines. Perhaps conventional breeding could also change the subunit structure of normal-protein lines.

Nielsen and co-workers (*35*) have deduced that five genes regulate the production of the subunits of glycinin, and Harada et al. (*33*) showed that three genes regulate the production of the subunits of β-conglycinin. The acidic and basic polypeptides of each subunit of glycinin are specific to each other because they are synthesized by the same locus. The combinations of these acidic−basic pairings form the subunits of the glycinin molecule in vivo as do the combinations of the α', α, and β subunits of β-conglycinin. Nielsen (*5*) found that the various subunits of glycinin contain different quantities of sulfur amino acids. For this reason Marco et al. (*34*) and Staswick et al. (*31*) studied the A_2B_{1a} subunit. The data presented by Serretti et al. (*30*) also indicate that the BARC lines contain different amounts of the glycinin subunits because of the differences in sulfur amino acids measured between the BARC lines and the check varieties that were used. Further studies to determine the amounts and types of glycinin subunits present in the high-protein lines would be beneficial in soybean breeding because the data presented here show that the amount of glycinin subunits as measured by their polypeptide content is expressed differently and in greater quantity in high-protein lines. Knowledge of subunit concentration and the factors leading to preferential subunit production could significantly affect the quality of protein stored in the soybean seed.

ACKNOWLEDGMENT

I thank Christopher Pooley for preparing the prints of the SDS-PAGE gels.

LITERATURE CITED

(1) Smith, K. Improvement of soybean composition to meet consumer/user demands. In *Report of the 21st Soybean Seed Research Conference*; Wilkinson, D., Ed.; American Seed Trade Association: Washington, DC, 1991; pp 71−78.
(2) Hartwig, E. E. Registration of soybean high-protein germplasm line D76-8070. *Crop Sci.* **1990**, *30*, 764−765.
(3) Leffel, R. C. Registration of high-protein soybean germplasm lines BARC-6, BARC-7, BARC-8, and BARC-9. *Crop Sci.* **1992**, *32*, 502.
(4) Burton, J. W. Quantitative genetics: results relevant to soybean breeding. In *Soybeans: Improvement, Production and Uses*; Wilcox, J. R., Ed.; American Society of Agronomy: Madison, WI, 1987; pp 211−247.
(5) Nielsen, N. C. The structure and complexity of the 11S polypeptides in soybeans. *J. Am. Oil Chem. Soc.* **1985**, *62*, 1680−1686.
(6) Nielsen, N. C. Soybean seed composition. In *Soybean: Genetics, Molecular Biology and Biotechnology*; Verma, D. P. S., Shoemaker, R. C., Eds.; CABI: Wallingford, U.K., 1996.
(7) Badley, R. A.; Atkinson, D.; Hauser, H.; Oldani, D.; Green, J. P.; Stubbs, J. M. The structure, physical and chemical properties of the soybean protein glycinin. *Biochim. Biophys. Acta* **1975**, *412*, 214−228.
(8) Wolf, W. J.; Briggs D. R. Studies on the cold-insoluble fraction of the water-extractable soybean proteins. II. Factors influencing conformation changes in the 11S component. *Arch. Biochem. Biophys.* **1985**, *76*, 377−393.
(9) Reichelt, R.; Schwenke, K. D.; Konig, T.; Pahtz, W.; Wangermann, G. Electron microscopic studies for estimation of the quaternary structure of the 11S globulin (helianthinin) from sunflower seed (*Helianthus annuus* L.). *Biochem. Physiol. Pflanz.* **1980**, *175*, 653−663.
(10) Plietz, P.; Damaschun, G. The structure of the 11S seed globulins from various plant species: comparative investigations by physical methods. *Stud. Biophys.* **1986**, *116*, 153−173.
(11) Thanh, V. H.; Shibasaki, K. β-conglycinin from soybean proteins. Isolation and immunological and physiochemical properties of the monomeric forms. *Biochim. Biophys. Acta* **1977**, *490*, 370−384.
(12) Thanh, V. H.; Shibasaki, K. Major proteins of soybean seeds. Reconstitution of β-conglycinin from its subunits. *J. Agric. Food Chem.* **1978**, *26*, 695−698.
(13) Thanh, V. H.; Shibasaki, K. Major proteins of soybean seeds. Subunit structure of β-conglycinin. *J. Agric. Food Chem.* **1978**, *26*, 692−695.
(14) Sebastiani, F. L.; Farrell, L. B.; Schuler, M. A.; Beachy, R. N. Complete sequence of a cDNA of the a subunit of soybean β-conglycinin. *Plant Mol. Biol.* **1990**, *15*, 197−201.
(15) Wilcox, J. R.; Cavins, J. F. Backcrossing high seed protein to a soybean cultivar. *Crop Sci.* **1995**, *35*, 1036−1041.
(16) Yaklich, R. W.; Helm, R. M.; Cockrell, G.; Herman, E. M. Analysis of the distribution of the major soybean seed allergens in a core collection of glycine max accessions. *Crop Sci.* **1999**, *39*, 1444−1447.
(17) Yaklich, R. W.; Kulik, M. M.; Garrison, C. S. Evaluation of vigor in soybean seeds: influence of date of planting and soil type on emergence, stand, and yield. *Crop Sci.* **1979**, *19*, 242−246.
(18) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680−685.

(19) Fontes, E. P. B.; Moreira, M. A.; Davies, C. S.; Nielsen, N. C. Urea-elicited changes in relative electrophoretic mobility of certain glycinin and β-conglycinin subunits. *Plant Physiol.* **1984**, *76*, 840−842.
(20) Ohtake, N.; Suzuki, M.; Takahashi, Y.; Fujiwara, T.; Chino, M.; Ikarashi, T.; Ohyama, T. Differential expression of β-conglycinin genes in nodulated and non-nodulated isolines of soybean. *Physiol. Plant.* **1996**, *96*, 101−110.
(21) Tang, W.; Diaz, P.; Wilson, K. A.; Tan-Wilson, A. L. Spatial distribution of protein accumulation in soybean seed. *Phytochemistry* **1993**, *33*, 1289−1295.
(22) Melroy, D. L.; Herman, E. M. TIP, and integral membrane protein of the protein-storage vacuoles of the soybean cotyledon undergoes developmentally regulated membrane accumulation and removal. *Planta* **1991**, *184*, 113−122.
(23) Wilson, K. A.; Rightmire, B. R.; Chen, J. C.; Tan-Wilson, A. L. Differential proteolysis of glycinin and β-conglycinin polypeptides during soybean germination and seedling growth. *Plant Physiol.* **1986**, *82*, 71−76.
(24) SAS Institute Inc. *SAS/STAT Software: Changes and Enhancements through Release 6.11*; SAS Institute: Cary, NC, 1996; p 1104.
(25) Piper, E. L.; Boote, K. J. Temperature and cultivar effects on soybean seed oil and protein concentrations. *J. Am. Oil Chem. Soc.* **1999**, *76*, 1233−1241.
(26) Howell, R. W.; Cartter, J. L. Physiological factors affecting composition of soybeans: II. Response of oil and other constituents of soybeans to temperature under controlled conditions. *Agron. J.* **1958**, *50*, 664−667.
(27) Howell, R. W.; Collins, F. I. Factors affecting linolenic and linoleic acid content of soybean oil. *Agron. J.* **1957**, *49*, 593−597.
(28) Maytin, E. V.; Levin, E. N.; Anderson, R. R. A laser densitometer for selective spot analysis on dot blots and two-dimensional gel autoradiograms. *Anal. Biochem.* **1991**, *194*, 284−294.
(29) Wilson, R. F. Seed metabolism. In *Soybeans: Improvement, Production and Uses*; Wilcox, J. R., Ed.; American Society of Agronomy: Madison, WI, 1987; pp 643−686.
(30) Serretti, C.; Schapaugh, W. T., Jr.; Leffel, R. C. Amino acid profile of high seed protein soybean. *Crop Sci.* **1994**, *34*, 207−209.
(31) Staswick, P. E.; Hermodson, M. A.; Nielsen, N. C. The amino acid sequence of the A_2B_{1a} subunit of glycinin. *J. Biol. Chem.* **1984**, *259*, 13424−13430.
(32) Nielsen, N. C.; Jung, R.; Nam, Y. W.; Beaman, T. W.; Oliveira, L. O.; Bassuner, R. Synthesis and assembly of 11S globulins. *J. Plant Physiol.* **1995**, *145*, 641−647.
(33) Harada, J. J.; Baker, S. J.; Goldberg, R. B. Soybean β-conglycinin genes are clustered in several DNA regions and are regulated by transcriptional and post-transcriptional processes. *Plant Cell* **1989**, *1*, 415−425.
(34) Marco, Y. A.; Thanh, V. H.; Tumer, N. E.; Scallon, B. J.; Nielsen, N. C. Cloning and structural analysis of DNA encoding an A_2B_{1a} subunit of glycinin. *J. Biol. Chem.* **1984**, *259*, 13436−13441.

Received for review September 8, 2000. Revised manuscript received November 13, 2000. Accepted November 13, 2000. Mention of a trademark or proprietary product does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval to the exclusion of other products that may also be suitable.

JF001110S