# Effect of Temperature and Genotype on the Crude Glycinin Fraction (11S) of Soybean and Its Analysis by Near-Infrared Reflectance Spectroscopy (Near-IRS)†

D. L. Pazdernik, S. J. Plehn,‡ J. L. Halgerson, and J. H. Orf*

Department of Agronomy and Plant Genetics, University of Minnesota, 1991 Buford Circle, 411 Borlaug Hall, St. Paul, Minnesota 55108-6026

The crude glycinin fraction (11S) of soybean [*Glycine max* (L.) Merr.] protein has desirable quality characteristics. The objectives of this soybean study were the following: (1) to develop a near-infrared reflectance spectroscopy (near-IRS) equation for screening of the 11S fraction; (2) to determine relationships between the 11S fraction, crude protein, and oil concentration; and (3) to study the effect of growing temperature and genotype on this fraction. Three near-IRS 11S equations were developed, with the best equation having a moderate standard error of calibration (7.7 g/kg) and coefficient of determination ($R^2 = 0.84$). A positive correlation was observed between the 11S protein subunit and crude protein concentration ($r = 0.69$, $P < 0.01$), and a negative correlation was found between the 11S protein subunit and oil concentration ($r = -0.67$, $P < 0.01$), based on 319 genotypes. Higher crude protein and 11S protein subunit concentrations were associated with higher growing temperatures, and higher oil concentration was associated with lower growing temperatures, based on 12 genotypes.

**Keywords:** *Soybean; near-infrared reflectance spectroscopy; protein quality; glycinin fraction*

## INTRODUCTION

Soybean [*Glycine max* (L.) Merr.] is known for its high crude protein and oil concentrations, with typical U.S. averages of approximately 410 and 210 g/kg for crude protein and oil concentrations, respectively, on a dry weight basis (Leffel and Rhodes, 1993). However, Breene et al. (1988) reported a gradual decline in the crude protein concentration of soybeans grown in northern states during the past 10−15 years. Additionally, more emphasis has been placed on soybean as a potential food product, which suggests that the nutritional quality of soybean protein will become increasingly important, but information on breeding for improved soybean protein quality is lacking.

Soybean protein (Osborne and Campbell, 1898) can be classified into four general types of proteins on the basis of size and solubility: albumin, prolamin, glutenin, and globulin proteins. The globulin proteins are much more complex in their molecular structure than the other protein types and comprise 65−90% of total soybean protein (Catsimpoolas et al., 1967; Murphy, 1985; Wilson, 1987). The globulin proteins can be further separated into five subunits on the basis of their size via ultracentrifugation (2S, 7S, 9S, 11S, and 15S) (Wolf, 1978), with the 7S and 11S subunits comprising over 75% of the total globulin fraction (Moreira et al., 1979; Wilson, 1987). Glycinin, the purified form of the 11S protein subunit (Morr, 1988; Wolf, 1978), comprises over 40% of the total globulin fraction and is the largest single fraction of total seed protein (25−35%) (Murphy and Resurrection, 1984; Wilson, 1987). In addition, soybean protein is deficient in methionine and cysteine, and glycinin contains a higher concentration of these sulfur-containing amino acids than other protein subunits and is important to the quality of tofu. Saio et al. (1969) reported that the ratio of glycinin to $\beta$-conglycinin is the major factor in the palatability of tofu.

The breeding of cultivars with an enriched 11S protein subunit concentration could be used to improve the nutritional quality of soybean protein, but a rapid 11S protein subunit screening method is needed, and information concerning the relationship between total crude protein and oil concentrations with 11S protein subunit concentration is lacking in soybean. The fastest method for total seed crude protein and oil analysis in soybean is near-infrared reflectance spectroscopy (near-IRS) (Norris, 1989; Panford, 1984), which possibly could be used for rapid 11S protein subunit screening. The objectives of this study were the following: (1) to develop an near-IRS equation for rapid screening of the 11S protein subunit in soybean; (2) to determine the relationship between 11S protein subunit, crude protein, and oil concentration; and (3) to study the effect of growing temperature and genotype on 11S protein subunit concentration in soybean.

## MATERIALS AND METHODS

**Sample Selection Using Near-IRS.** Seeds of 319 randomly chosen soybean genotypes, which included plant introductions, cultivars, and breeding lines ranging in maturity from groups OO to III (genotypes adapted to the northern latitudes), were used in this study. Seed samples (20 g) of the 319 genotypes were dried overnight at 60 °C (approximately 8% moisture among all samples) and ground in a Fred Stein mill for 1 min. The ground samples were entered in a NIRS Systems (Silver Spring, MD) Model 6500 near-IRS system. Two Infrasoft International software programs (ISI, Port Matilda, PA), "Center" and "Select", were used to screen samples for spectral outliers and to choose samples that represented the 319 genotypes. The "Select" program eliminates samples with similar spectra, and the "Center" program

* Author to whom correspondence should be addressed [telephone (612) 625-8275; fax (612) 625-1268; e-mail Orfxx001@maroon.tc.umn.edu.

† Journal Series No. 22,312 of the Univ. of Minnesota Agric. Exp. Stn. and the USDA-ARS.

‡ Present address: 527 West 2nd St., Mount Vernon, IN 47620.

S0021-8561(96)00225-7 CCC: $12.00 

**Table 1. Near-IRS Equation Statistics Based on Three Mathematical Treatments of the Spectral Data Used To Predict Soybean Seed 11S Protein Subunit Protein Concentration**

| mathematical treatment | no. of samples | mean (g/kg) | SD[a] (g/kg) | SEC[b] (g/kg) | $R^2$ [c] | SECV[d] (g/kg) | 1 − VR[e] |
|---|---|---|---|---|---|---|---|
| log(1/*R*) | 86 | 204 | 18.0 | 9.7 | 0.71 | 11.2 | 0.62 |
| first derivative | 85 | 205 | 19.1 | 7.7 | 0.84 | 10.7 | 0.69 |
| second derivative | 85 | 205 | 17.2 | 8.3 | 0.77 | 10.9 | 0.64 |

[a] SD, standard deviation of the mean. [b] SEC, standard error of calibration. [c] $R^2$, coefficient of determination. [d] SECV, standard error of cross-validation. [e] 1 − VR, 1 minus the ratio of unexplained variance divided by total variance.

defines spectral boundaries, which eliminates outliers (Shenk and Westerhaus, 1991a). These two programs selected 98 of the 319 genotypes to create a calibration set used to develop the 11S protein subunit equations.

**Analytical Methods for the 11S Protein Subunit.** Duplicate samples (10 g/sample) of the selected 98 genotypes were defatted with 100 mL of petroleum ether in a Soxhlet apparatus (Becker, 1978), and the 11S protein subunit concentration was separated by using a modified method of Thanh et al. (1975, 1976) in which Trizma Buffer (Sigma Chemical Co., St. Louis, MO) was replaced with Bis tris propane. Bis tris propane has a p$K_a$ that more closely accommodates a pH of 6.4, which is essential for the precise separation of the 11S protein subunit from the 7S protein subunit. The 11S protein subunit fraction was then quantified by using a Bio-Rad protein assay (1976). Bovine plasma albumin (BPA) was used as the standard with a conversion factor first determined between the BPA standard and the purified 11S protein subunit (glycinin), which was obtained from the lab of Walter J. Wolf (USDA-ARS, Peoria, IL, 61604). A specified amount of purified glycinin was weighed out multiple times and assayed for protein concentration using BPA as the standard. On the basis of the assay, a conversion factor of 0.894 was calculated that was used to determine the empirical level of the 11S protein subunit on the basis of the BPA standard. The 11S protein subunit concentration of whole seed was calculated as follows: 11S protein subunit concentration whole seed (g/kg soybean meal) = (11S protein subunit concentration defatted meal) × [(100 − oil concentration)/100] × 0.894 (conversion factor).

**Near-IRS Equation Development.** The 11S protein subunit values for the 98 genotypes were entered in an near-IRS calibration file using the ISI program "Calibrate", with the modified partial least-squares option to develop the 11S protein subunit equations (Shenk and Westerhaus, 1991b). The "Calibrate" program was run three times to eliminate outliers, which resulted in either 13 or 14 samples being eliminated depending on the mathematical treatment. Three mathematical treatments [log 1/*R* (*R* = reflectance), first derivative, and second derivative] were compared to determine which treatment provided the best near-IRS equation, with the criteria of minimizing the standard error of cross-validation (SECV) and maximizing 1 minus the ratio of unexplained variance: total variance (1 − VR). The SECV and 1 − VR values provide an independent measure of standard error and explained variance among the three equations.

The 11S protein subunit equation with the first-derivative mathematical treatment was used to predict the 11S protein subunit concentration of all 319 soybean genotypes. Crude protein and oil concentrations also were determined for all genotypes by using near-IRS. The data were analyzed with Proc Reg within SAS, version 6.03 (SAS, 1992), to determine the relationship between 11S protein subunit, crude protein, and oil concentration.

**Effect of Growing Temperature and Genotype on 11S Protein Subunit Concentration.** Twelve genotypes (Early White Eyebrow cv., M67-68, He Feng 25 cv., Clay cv., Merit cv., Lincoln cv., Anoka cv., M84-756, M55-74, Grande cv., Sturdy cv., and Swift cv.) were grown in one Conviron PGW36 growth chamber (Controlled Environment Ltd., Winnipeg, MB, Canada) at 33/28 °C day/night temperatures and in another Conviron PGW36 growth chamber at 24/19 °C day/night temperatures. Each genotype was grown in two 20-cm pots/growth chamber containing a 1:1 sterilized soil/sand mixture. The two pots/growth chamber were randomly distributed with 10 seeds planted/pot, inoculated with *Bradyrhizobium japonicum* (Kirchner) Buchanan, and thinned to two plants/pot after emergence. Both growth chambers were maintained with a 16-h photoperiod until the onset of flowering when the photoperiod was dropped 0.5 h every 3 days until a 12-h photoperiod was reached. Plants were watered as needed and fertilized with an N-free Hoagland's nutrient solution once/week.

At maturity, seeds from individual plants were harvested, dried overnight at 60 °C, counted, and analyzed for total crude protein and oil concentration and 11S protein subunit concentration by using the previously described near-IRS methods. Samples that contained fewer than 50 seeds were discarded. The near-IRS predicted protein values were verified with a Kjeldahl procedure by Ingman Laboratories, Inc. (Minneapolis, MN), and 11S protein subunit values were verified with the previously described chemical method. The experiment was repeated twice and analyzed as a randomized complete block design with two replications and two fixed factors (temperature and genotype) by using Proc GLM within SAS, version 6.03 (SAS, 1992).

## RESULTS AND DISCUSSION

**Near-IRS Equation Development.** The three mathematical treatments for near-IRS equation development (Table 1) displayed similar results, but did differ with the first derivative outperforming the other two treatments in terms of having lower SEC and SECV (7.7 and 10.7 g/kg, respectively) values and higher $R^2$ and 1 − VR (0.84 and 0.69, respectively) values. These results are similar to those obtained from protein and oil equations previously developed at the University of Minnesota soybean breeding project in which the first derivative was used as the mathematical treatment.

Use of the near-IRS first-derivative equation demonstrates that the 11S protein subunit fraction of soybean seed protein can be reasonably predicted, with the SEC within the expected range of <9 g/kg for crude protein (Marten et al., 1989). But there is some concern that the standard error of cross-validation (SECV) is more than 20% greater than the SEC and that the 1 minus the variance ratio (1 − VR) statistic is low. This high error suggests that either the chemical analysis used to quantify the 11S protein subunit was error prone, the near-IRS was not able to correlate spectra with chemical analysis results, and/or subsampling of soybean seed and ground meal from individual genotypes was variable. The 11S protein subunit chemical analysis (lab mean = 201 g/kg and standard error of lab = 5.86 g/kg) was highly sensitive to specific pH changes, which could have resulted in poor separation of the 11S protein subunit fraction from other protein fractions to cause higher or lower 11S protein subunit values for specific genotypes, causing variable chemical analysis. The standard deviation of 19.1 g/kg for the equation development was small and may explain the low 1 − VR statistic.

**Relationship between 11S Protein Subunit, Crude Protein, and Oil.** The ranges for crude protein (315.4−428.3 g/kg), oil (136.0−214.6 g/kg), and 11S

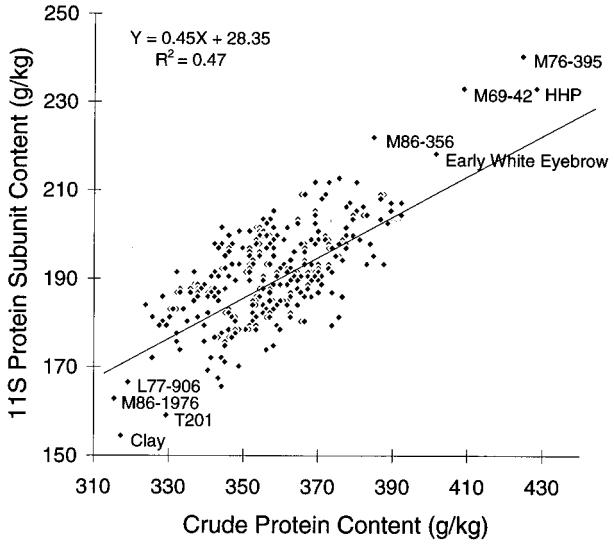


**Figure 1.** Relationship between soybean seed crude protein and 11S protein subunit concentrations (13% moisture basis) for 319 genotypes.

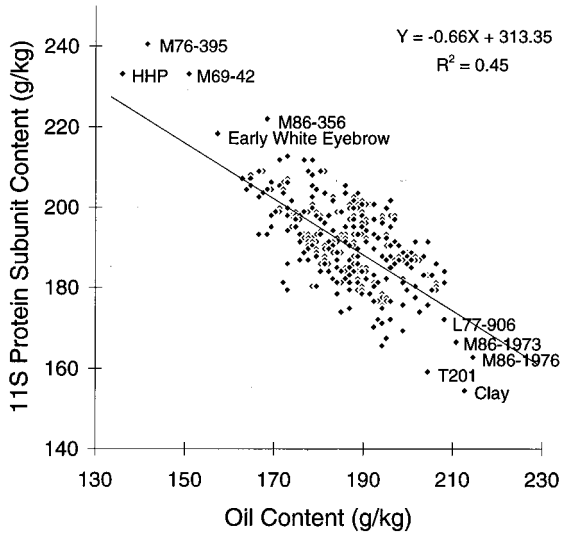


**Figure 2.** Relationship between soybean seed oil and 11S protein subunit concentrations (13% moisture basis) for 319 genotypes.

protein subunit (154.3−240.5 g/kg) concentrations (data not shown) were calculated on a 13% moisture basis among the 319 genotypes. The 11S protein subunit fraction averaged 486.2 g/kg of total seed protein, which is higher than average but within the range of values reported by Murphy and Resurrection (1984). A strong negative correlation existed between protein and oil concentrations ($r = -0.99$, $P < 0.01$), which has been previously well established in the literature (Wilson, 1987; Burton and Brim, 1981). A positive relationship was shown between 11S protein subunit and crude protein concentrations (Figure 1) ($r = 0.69$, $P < 0.01$), with M76-395 (a high-protein line) having the highest 11S protein subunit concentration and HHP (with high 11S protein subunit concentration) having the highest crude protein concentration, suggesting that simultaneous breeding for both traits could be accomplished by combining the highest protein lines with the highest 11S protein subunit lines.

Burton et al. (1982) showed that methionine and total protein concentrations were not correlated in soybean. It is also known that the 11S protein subunit fraction of protein contains a greater percentage of methionine, suggesting that protein and methionine concentrations could be increased simultaneously by selecting for 11S protein subunit concentration (Wilson, 1987).

A negative relationship was shown between 11S protein subunit and oil concentrations (Figure 2) ($r = -0.67$, $P < 0.01$), but the strength of this relationship was weaker than that between protein and oil, suggesting that the 11S protein subunit concentration could be increased while maintaining a constant oil concentration. These results also demonstrated a large range in 11S protein subunit concentration among the 319 genotypes, suggesting that there is ample genetic variation to breed for this trait in soybean, but future work on the heritability of the 11S protein subunit and relationships with seed, agronomic, disease, and insect, and other critical traits needs to be carried out.

**Effect of Growing Temperature and Genotype on 11S Protein Subunit Concentration.** The 12 soybean genotypes used in the temperature study showed a wide range of 11S subunit, protein, and oil concentration values (Table 2), with a positive relationship between 11S protein subunit and crude protein concentrations and a negative relationship between oil and 11S protein subunit concentrations. These results were not surprising since the genotypes were selected to cover the entire range of 11S protein subunit values from the previous experiment. Temperature also had an impact on 10 of the 12 genotypes, with the average 11S protein subunit and crude protein concentrations being significantly higher and the oil concentration

**Table 2. Means for Soybean Seed 11S Protein Subunit, Crude Protein, and Oil Concentrations (13% Moisture) among 12 Soybean Genotypes Grown at 33/28 and 24/19 °C Day/Night Temperatures under Growth Chamber Conditions**

| genotype | maturity group | 11S protein subunit (g/kg) | | crude protein (g/kg) | | oil (g/kg) | |
|---|---|---|---|---|---|---|---|
| | | 33/28 °C | 24/19 °C | 33/28 °C | 24/19 °C | 33/28 °C | 24/19 °C |
| E. W. Eye[a] | O | 213.7 | 215.5 | 397.8 | 399.6 | 160.1 | 157.3 |
| M67-68 | I | 230.3 | 212.8 | 420.9 | 395.9 | 146.2 | 161.9 |
| He Feng 25 | II | 198.9 | 199.8 | 366.3 | 382.0 | 180.4 | 169.3 |
| Clay | O | 215.5 | 195.2 | 390.4 | 367.2 | 166.5 | 181.3 |
| Merit | O | 212.8 | 194.3 | 395.0 | 360.8 | 162.8 | 185.0 |
| Lincoln | III | 220.2 | 188.7 | 415.3 | 369.1 | 148.9 | 178.5 |
| Anoka | I | 212.8 | 185.1 | 407.9 | 365.4 | 154.5 | 182.2 |
| M84-756 | I | 222.9 | 175.8 | 407.9 | 341.3 | 154.5 | 198.0 |
| M55-74 | I | 190.6 | 174.8 | 359.8 | 329.3 | 185.0 | 204.4 |
| Grande | O | 194.3 | 173.9 | 374.6 | 346.0 | 174.8 | 192.4 |
| Sturdy | I | 195.2 | 172.1 | 381.1 | 346.0 | 172.1 | 195.2 |
| Swift | O | 187.8 | 158.2 | 377.4 | 331.2 | 173.0 | 203.5 |
| LSD (0.05)[b] | | 26.8 | 24.9 | 32.7 | 30.5 | 21.4 | 19.8 |

[a] E. W. Eye, Early White Eyebrow. [b] LSD (0.05), least significant difference used to compare genotype means within columns ($P = 0.05$).

**Table 3. Means for Soybean Seed 11S Protein Subunit, Crude Protein, and Oil Concentrations (13% Moisture) Grown at 33/28 and 24/19 °C Day/Night Temperatures under Growth Chamber Conditions Averaged across 12 Soybean Genotypes**

| day/night temp (°C) | 11S protein subunit (g/kg) | crude protein (g/kg) | oil (g/kg) |
|---|---|---|---|
| 33/28 | 207.7 | 391.2 | 164.9 |
| 24/19 | 187.2 | 361.2 | 184.1 |
| LSD (0.05)[a] | 8.9 | 10.9 | 6.4 |

[a] LSD (0.05), least significant difference used to compare genotype means within columns ($P = 0.05$).

lower at 33/28 °C (Table 3). These findings agree with those of Wolf and Cavins (1982), who reported a positive correlation between protein concentration and temperature. Additionally, there was no significant temperature × genotype interaction, demonstrating that genotypes performed similarly in both temperature regimes under growth chamber conditions, but future studies need to confirm whether a genotype × temperature interaction exists under field conditions using randomly selected genotypes.

Hurburgh and Brumm (1988) reported that soybeans grown in warmer climates generally have higher protein concentrations than those grown in cooler climates, such as Minnesota's. Because of the lower crude protein concentration, Minnesota soybeans have lower nutritional quality and are less desirable for export. The results of this study showed that temperature affected many genotypes, but did not affect the 11S protein subunit concentration of two high 11S protein subunit genotypes (Early White Eyebrow cv. and He Feng 25 cv.) and the crude protein concentration of one high-protein genotype (Early White Eyebrow cv.), suggesting that it may be possible to develop cultivars for Minnesota in which the 11S protein subunit and crude protein concentrations are not as temperature sensitive. Future studies should investigate many sources of germplasm for temperature sensitivity to 11S protein subunit and crude protein concentrations, which could be used to develop high 11S protein subunit and protein lines for cooler climates.

ACKNOWLEDGMENT

We thank Dr. Carroll Vance, USDA, St. Paul, MN, for the use of his lab facilities.

LITERATURE CITED

Becker, K. W. Solvent extraction of soybeans. *J. Am. Oil Chem. Soc.* **1978**, *55*, 754−761.

*Bio-Rad Laboratories Technical Bulletin 1051*; Bio-Rad Laboratories: Richmond, CA, 1976.

Breene, W. M.; Lin, S.; Hardman, L.; Orf, J. Protein and oil concentration of soybeans from different geographic locations. *J. Am. Oil Chem. Soc.* **1988**, *65*, 1927−1931.

Burton, J. W.; Brim, C. A. Recurrent selection in soybeans. III. Selection for increased percent oil in seeds. *Crop Sci.* **1981**, *21*, 31−34.

Burton, J. W.; Purcell, A. E.; Walter, W. M., Jr. Methionine concentration in soybean protein from populations selected for increased percent protein. *Crop Sci.* **1982**, *22*, 430−432.

Catsimpoolas, N.; Rogers, D. A.; Circle, S. J.; Meyer, E. W. Purification and structural studies of the 11S component of soybean proteins. *Cereal Chem.* **1967**, *44*, 631−637.

Hurburgh, C. R.; Brumm, T. J. *Improving the Quality of Exported U.S. Soybeans*; American Soybean Association Trade Mission to Japan, 1988.

Leffel, R. C.; Rhodes, W. K. Agronomic performance and economic value of high-seed protein soybean. *J. Prod. Agric.* **1993**, *6*, 365−368.

Marten, G. C., Shenk, J. S., Barton, F. E., II, Eds. *Near infrared reflectance spectroscopy (near-IRS): Analysis of forage quality*; Agricultural Handbook 643; USDA-ARS: Washington, DC, 1989; p 100.

Moreira, M. A.; Hermodson, M. A.; Larkins, B. A.; Nielson, N. C. Partial characterization of the acidic and basic polypeptides of glycinin. *J. Biol. Chem.* **1979**, *254*, 9921−9926.

Morr, C. V. Chemistry of soybean protein. In *Soybean Utilization Alternatives*; McCann, L., Eds.; The Center for Alternative Crops and Products, University of Minnesota: St. Paul, MN, 1988; pp 107-115.

Murphy, P. A. Structural characteristics of soybean glycinin and $\beta$-conglycinin. In *World Soybean Research Conference III Proceedings*; Shibles, R., Ed.; Westview Press, Inc.: Boulder, CO, 1985; pp 143-151.

Murphy, P. A.; Resurrection, A. P. Varietal and environmental differences in soybean glycinin and $\beta$-conglycinin concentration. *J. Agric. Food Chem.* **1984**, *32*, 911−915.

Norris, K. Workshop notes from the 9th Annual National Agricultural near-IRS Research Network, St. Paul, MN, 1989.

Osborne, T. B.; Campbell, G. F. Proteids of the soybean (*Glycine hispida*). *J. Am. Chem. Soc.* **1898**, *20*, 419−428.

Panford, J. A. Application of near-infrared reflectance spectroscopy in North America. In *Near-Infrared Technology in the Agricultural and Food Industries*. Williams, P., Norris, K., Eds.; American Association of Cereal Chemists, Inc.: St. Paul, MN, 1984; pp 201-211.

SAS Institute, Inc. *SAS/STAT Guide for Personal Computers*, 6th ed.; SAS Inst., Inc.: Cary, NC, 1992.

Saio, K.; Kamiya, M.; Watanabe, T. Food processing characteristics of soybean 11S and 7S proteins. *Agric. Biol. Chem.* **1969**, *33*, 1301−1308.

Shenk, J. S.; Westerhaus, M. O. Population definition, sample selection and calibration procedures for near infrared reflectance spectroscopy. *Crop Sci.* **1991a**, *31*, 469−471.

Shenk, J. S.; Westerhaus, M. O. Population structuring of near infrared spectra and modified partial least squares regression. *Crop Sci.* **1991b**, *31*, 1548−1555.

Thanh, V. H.; Shibasaki, K. Major proteins of soybean seeds. A straightforward fractionation and their characterization. *J. Agric. Food Chem.* **1976**, *24*, 1117−1121.

Thanh, V. H.; Okubo, K.; Shibasaki, K. Isolation and characterization of the multiple 7S globulins of soybean proteins. *Plant Physiol.* **1975**, *56*, 19−22.

Wilson, R. F. Primary constituents of soybean seed. In *Soybeans: Improvement, Production and Uses*, 2nd ed.; Wilcox, J. R., Ed.; American Society of Agronomy: Madison, WI, 1987; pp 646-686.

Wolf, W. J. Purification and properties of the proteins. In *Soybeans: Chemistry and Technology. Vol. 1 Proteins*; Smith, A. K., Circle, S. J., Eds.; AVI Publishing: Westport, CT, 1978; pp 93-143.

Wolf, W. J.; Cavins, J. F. Temperature effects on soybeans. *J. Am. Oil Chem. Soc.* **1982**, *59*, 230−232.

Received for review April 1, 1996. Revised manuscript received May 17, 1996. Accepted May 21, 1996.[⊗] This research was supported in part from grants from the Minnesota Soybean Research and Promotion Council, the Agricultural Utilization Research Institute, and the Minnesota Department of Agriculture.

JF960225M

[⊗] Abstract published in *Advance ACS Abstracts*, July 15, 1996.