J Am Oil Chem Soc (2007) 84:793–801
DOI 10.1007/s11746-007-1111-z

ORIGINAL PAPER

A Study on Subunit Groups of Soybean Protein Extracts under SDS-PAGE

Shunhu Liu · Ruibao Zhou · Shaojun Tian · Junyi Gai

Received: 3 February 2007 / Revised: 4 July 2007 / Accepted: 11 July 2007 / Published online: 10 August 2007

Abstract Protein extracts of 640 soybean cultivars and landraces, mainly from China and a few from the US, were analyzed for their components and subunits based on distribution patterns of bands with varying molecular weights (MW) under SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis). The number and molecular weight of the bands in SDS-PAGE varied among materials and showed a tendency of continuous distribution. Accordingly, the SDS-PAGE patterns of the soybean protein extracts were divided into two regions: the region of bands with MW < 44 KDa and that with MW $\geq$ 44 KDa. The first region containing mainly 11S proteins was divided into four parts, called subunit groups, i.e. 11S-1 (14.4–22 KDa), 11S-2 (22–26 KDa), 11S-3 (26–34 KDa) and 11S-4 (34–44 KDa). The second region containing mainly 7S protein was divided into six subunit groups, i.e. 7S-1 (44–49 KDa), 7S-2 (49–55 KDa), 7S-3 (55–67 KDa), 7S-4 (67–73 KDa), 7S-5 (73–82 KDa) and 7S-6 (82–91 KDa). The sum of relative contents of 11S-1–11S-4 was obtained as the relative content of 11S protein, those of 7S-1–7S-6 as that of 7S protein, and therefore, the 11S/7S ratio obtained. The proposed criteria were demonstrated to be simple, stable and feasible. Among the 640 tested materials, 39 lacked 11S-1 but none lacked the other 11S subunit groups, while deficiencies existed in all the six subunit groups of 7S, indicating a great potential for the genetic variation of protein components and subunits for breeding for the improvement of protein qualities.

Keywords 7S · 11S · Molecular weight (MW) · Soybean · Protein extract · SDS-PAGE · Subunit group · Cultivar

S. Liu · J. Gai (✉)
National Center for Soybean Improvement, National Key Laboratory for Crop Genetics and Germplasm Enhancement, Soybean Research Institute of Nanjing Agricultural University, Nanjing, Jiangsu 210095, People's Republic of China
e-mail: sri@njau.edu.cn

R. Zhou · S. Tian
Soybean Processing Research Institute, Henan University of Technology, Zhengzhou, Henan 450012, People's Republic of China

Introduction

The protein in soybean seeds accounts for about 40% of the dry seed weight. It is an important source of plant proteins for human and animal nutrition. A number of researchers have focused on soy protein components with respect to their extraction, separation, classification and physicochemical properties. The classic method of separating soy protein into components is the ultracentrifuge method, developed by Wolf and his colleagues [1, 2] who first classified soy proteins into 2S, 7S, 11S, and 15S components and later on estimated that 2S, 7S, 11S and 15S accounted for 22, 37, 31 and 11% of the total proteins of soybeans, respectively.

The other common method of studying protein components is the gel electrophoresis technique. This method has an advantage of further differentiating sub-units of proteins. Catsimpoolas et al. [3] analyzed soy protein and its 11S by Disc-PAGE and found 12 bands in soybean protein extract and six bands in 11S. Using the similar SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) technique, Hill and Breidenbach [4] found six bands in 7S and seven bands in 11S of soybean cultivar Portage, respectively. Kitamura and his colleagues [5, 6] determined four acidic subunits and three or four basic subunits in

soybean cultivar Raiden using Disc-PAGE. The bands of acidic subunits were named firstly by him as A_1, A_2, A_3 and A_4, respectively. The molecular weight (MW) of basic subunits, A_1 (or A_2 and A_3) and A_4 are 22.5, 37 and 45 KDa, respectively. Then, the subunits of 7S and 11S and their MW were reported repeatedly by a number of researchers. Draper and Catsimpoolas [7] found only two acidic subunits (MW 45 and 42 KDa) and one basic subunit (MW 19 KDa) in 11S of soybean cultivar Corosoy. Thanh and Shibasaki [8] obtained α' and α subunits with MW 57 KDa, and β subunit with MW 42 KDa in 7S of soybean cultivar Raiden. Iibuchi and Imahor [9] reported that the MW of α and β subunits in 7S were 68 and 52 KDa, respectively. In 7S of cultivar Raiden and Century, Fontes et al. [10] found α' with MW 72 KDa, α with MW 68 KDa, β with MW 52 KDa, and in 11S, A_3 with MW 42 KDa, A_{1a}, A_{1b}, A_2 and A_4 with MW 37 KDa, and A_5 with MW 10 KDa. There were two bands with MW 15–20 KDa, the same as Mori et al. [11] in 11S basic subunits. But other results [12] were divergent for A_3 and A_4 in Raiden, and there was a B0-Conglycinin subunit in 7S. Sathe et al. [13] found four subunits in 7S of soybean cultivar Kingwa, i.e. α' with MW 80.22 KDa, α with MW 70.63 KDa, β with MW 48.42 KDa, and γ with MW 46.24 KDa. They got six acidic subunits of 11S: A_{1a}, A_{1b}, A_2, and A_4 with MW 33.57 KDa; A_3 with MW 40.74 KDa; A_5 with MW 10 KDa; five basic subunits (B_{1a}, B_{1b}, B_2, B_3 and B_4) with MW 20.65 KDa. In 19 commercial soy protein isolates, Arrese et al. [14] determined α' with MW 79.8 KDa, α with MW 64.5 KDa and β with MW 46.8 KDa, acidic subunit and basic subunit with MW 36.3 and 19.3 KDa, respectively. Thus it can be seen that the previous results about the number and MW of subunits in 11S and 7S varied with each other. In total, five subunits (α', α, β, γ and B0-Conglycinin) in 7S and six acidic subunits (A_{1a}, A_{1b}, A_2, A_3, A_4 and A_5) and five basic subunits (B_{1a}, B_{1b}, B_2, B_3 and B_4) in 11S have been reported, but with inconsistent molecular weights among some works. The results of Sathe, Kitamura, Arrese, Fontes, Thanh and their groups were adopted, respectively, by different researchers, such as Poysa et al. and others [15–18]. MW 71 KDa for α', 67 KDa for α, 50 KDa for β, and 42 KDa for A_3 were cited by Ruiz-Henestrosa et al. [19]; 68 KDa for α, 48 KDa for β were introduced by Natarajan et al. [20]. Tsumura et al. [21] divided the SDS-PAGE pattern into two regions with 42 KDa as the critical value, 7S in the region with MW over 42 KDa, 11S in the region with MW 20–42 KDa. Samoto et al. [22] and others also reported the protein subunits by using the above proposed names, but without explanation of their molecular weights.

It seems that the classification criteria of soy protein subunits were mainly from studies on small samples of soybean cultivars, therefore, the number of bands or subunits was limited and easily distinguished. Since there exist a great number of soybean cultivars and landraces (more than 30,000 in the world), especially in China where the cultivated soybean originated and great genetic diversity exists, we need to know the overall variation of the bands, their molecular weight and their frequency distribution in soybean genetic resources of the species (*Glycine max* (L.) Merr.).

Soy protein has been a major protein source used in the food industry. The 11S and 7S are the most important components of soy proteins. Because the two components and their subunits were found to be significantly different in physical–chemical and functional properties in food processing [2, 23], there is a need to select soybeans with different 11S/7S ratios or subunit combinations and to study these protein components further. In our previous studies on protein components in different soybean cultivars, it was found that the number and MW of the bands in SDS-PAGE analysis varied among cultivars, a number of bands appeared which had not been reported before and there is a tendency toward continuous distribution of the bands. Therefore, the present study was aimed at revealing the total distribution of the bands in the population of soybean cultivars and landraces and to determine a MW criterion for distinguishing 11S and 7S of soybean protein extracts as well as their subunits under SDS-PAGE. Based on which, a simple, fast and reliable procedure for classifying soy protein components and their subunits can be tentatively established for identifying differences of soy proteins among cultivars and breeding lines in the improvement of soy protein qualities.

Materials and Methods

Plant Materials

A total of 640 soybean cultivars and landraces were chosen mainly from China to represent six eco-regions [24] and 000 ~ IX maturity groups in addition to a few from the US, and then tested in an incomplete block design (Blocks in Replication Design) with two replications and three rows (2 m × 0.5 m) per plot at the Experimental Station of the National Centre for Soybean Improvement in China in 2002–2003. In 2004, another 18 soybean cultivars selected from the six eco-regions in China were also tested at the station in a randomized block experiment with two replications in order to demonstrate the stability and feasibility of the proposed classification criteria of soy protein components and their subunit groups with molecular weight using SDS-PAGE. Since both experiments showed that the field performance of the cultivars were pretty consistent between the

replications, only the seeds from one replication were used for SDS-PAGE analysis.

Preparation of Soy Protein Extract

The preparation method for soybean protein extracts of Wolf [25] was used extensively for years by a number of researchers, such as Tsumura et al. [21], Puppo et al. [26] and others. The conditions and time in Wolf's method were modified by several workers in order to separate 7S and 11S as well as their subunits (Pesic et al. [16]; Deak et al. [18]; Samoto et al. [22]), but unfortunately some parts of the extract, such as 7S, α', β were enhanced while the other parts might be reduced. It is our understanding that Wolf's method with certain appropriate modifications is suitable for determining the relative contents of 7S and 11S as well as their subunits, rather than for separating them.

In the present study, relatively pure soy protein extract was prepared according to the laboratory method outlined in Fig. 1. It basically resembled a procedure for making soy protein extracts, except that the final precipitate was not dried. Instead it was dissolved.

About 20 g dried seeds of each soybean cultivar was ground with a"1095 Knifetec Sample Mill" (Foss Tecator, Sweden), and passed through a 100 mesh sieve. The defatted powder was obtained by "Soxtec Avanti 2050" (Foss Tecator, Sweden). The remained procedure followed the modified method of Wolf [25] and Puppo et al. [27]. Defatted soybean meal dispersed in distilled water (1:15 W/W). The dispersion was adjusted to pH 7.5 with 2M NaOH, stirred at room temperature for 2 h, and then centrifuged at 9,000*g* for 20 min at 15 °C. The insoluble material was removed. Then, the supernatant was adjusted to pH 4.5 with 2M HCl and the precipitate was collected by centrifugation at 4,500*g* for 20 min at 15 °C. The precipitate obtained was washed with distilled water under centrifugation at 2000 g for 20 min at 15 °C. Then the precipitate was dissolved by stirring in distilled water, adjusted to pH 7.0 with 2M NaOH to yield a soy protein extract with about 10% protein concentration.

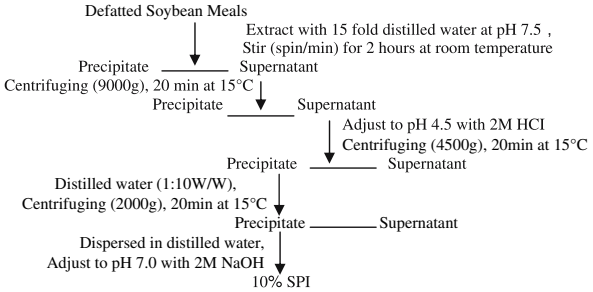

Fig. 1 Flowchart for preparing soybean protein extracts

SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was conducted with 12% polyacrylamide gel according to the modified method of Laemmli [28]. Protein solutions were diluted with an equal volume of a pH 8.0 buffer, and then heated in a boiling water bath for 5 min. A total of 10 μg extracted protein were loaded into each lane. Molecular weight markers used were rabbit phosphorylase *b* (97.4 KDa), bovine serum albumin (66.2 KDa), rabbit actin (43.0 KDa), bovine carbonic anhydrase (31.0 KDa), trypsin inhibitor (20.1 KDa) and chicken egg white lysozyme (14.4 KDa). The staining intensity, relative content and MW of the bands on SDS-PAGE were measured with Bio-Rad Gel Doc™EQ (BIO-RAD Laborations-Segrete, Milan, Italy).

The SDS-PAGE analysis was performed one time for the experiment of 640 cultivars and landraces and four times (or repeats) for the demonstration experiment of 18 cultivars. Analysis of variance was made for the latter experiment according to Gai [29].

Results

SDS-PAGE Patterns of Soy Protein Extracts

As an example, the number of bands of soy proteins from some cultivars are shown in Fig. 2a. The range of clear band number per cultivar was between 10 and 17, with a mean value of 13 within MW 14.4–97.4 KDa (the range of standards of protein MW markers). There were some unclear bands between two clear bands, which might be some low content proteins. The protein extract from each soybean cultivar had a unique number of clear and unclear bands distributed continuously. The diversity of the band numbers depended on the diversity of soybean proteins among cultivars. This was a major reason why different band numbers were obtained from different cultivars by different researchers. But the diversity of band numbers offers an opportunity for breeding for specific soybeans with improved soy protein profiles.

In SDS-PAGE, the MW of the bands was determined by the protein markers. Figure 2b showed that the MW of bands varied among different cultivars too. This was especially true with bands of MW 16.68–72.91 KDa.

Relative Protein Content of Bands

The area and staining intensity of all bands as well as each band varied among cultivars as shown in Fig. 2b. Multiplying anarea by its staining intensity was used to calculate

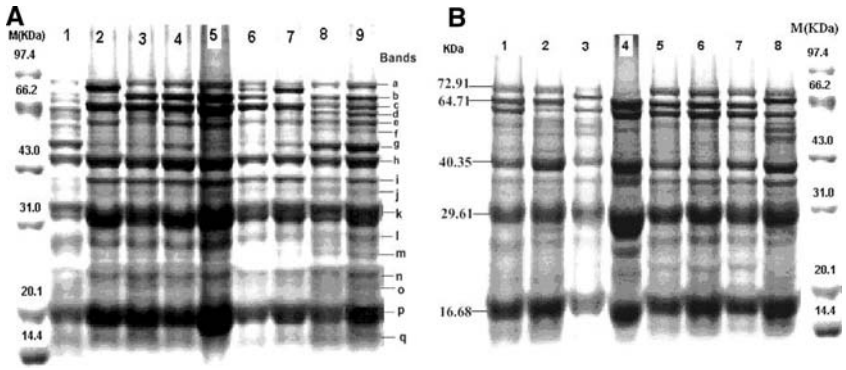

Fig. 2 **a** Diversity of band number (SDS-PAGE) in soybean cultivars 10 μg extracted protein were loaded to each lane. There showed 10–17 bands (a–q) in lane 1–9 where *1* Pella, *2* Meng 9024, *3* Jinda53, *4* Yudou 22, *5* LD 42, *6* Huaidou6hao, *7* He 95–1, *8* Tai 75, *9* Zhongdou 19. **b** Diversity of staining intensity and area of SDS-PAGE bands in soybean cultivars 10 μg extracted protein were loaded to each lane. The MW of the bands were 72.91, 64.71, 40.35, 29.61, and 16.68 KDa, respectively, in cultivar *1* E–M, *2* Huaidou1hao, *3* Jinda 26, *4* Bianjingdadou, *5* T286, *6* Bedford, *7* Corsica, and *8* Xinliuqing

its relative protein content. Here the relative protein content means the percentage of protein content of an electrophoretic band to total protein content of all electrophoretic bands in the same lane (or cultivar). Figure 3 showed the frequency distribution of average relative contents of bands of the 640 cultivars and landraces. Here the average relative content means the relative protein contents of all electrophoretic bands with a same molecular weight averaged over the 640 soybean cultivars. It indicated a great diversity in protein components. The relative content of bands with a MW between 14.4 and 91 KDa was about 2.66–13.00%, with the highest relative content 13.00% corresponding to the band with a MW of 20 KDa, and the lowest relative content 2.66% corresponding to a band with a MW 27 KDa. The difference of relative content of bands provided the background for the improvement of soybean cultivars with specific content of bands for specific functional characteristics of soy protein.

Frequency Distribution of Soy Protein Bands

The bands of soy protein extracts from 640 selected cultivars could be classified according to their molecular weight. For an overall exploration, the bands in the 640 SDS-PAGE analyses were grouped with a 1-KDa interval, the frequency (i.e. the number of cultivars within the same molecular weight interval) was counted for each band MW interval, and the frequency distribution was made as indicated in Fig. 4. The results showed that the frequencies of soy protein bands between MW 14.4–91 KDa were continuously distributed with significant peaks and valleys, in other words, composed of a number of component distributions without distinct separation among bands in the whole population of 640 cultivars. Figure 5 shows the distribution of further grouped bands according to molecular weight which will be explained later. Here it was used to indicate that there also exist differences among band

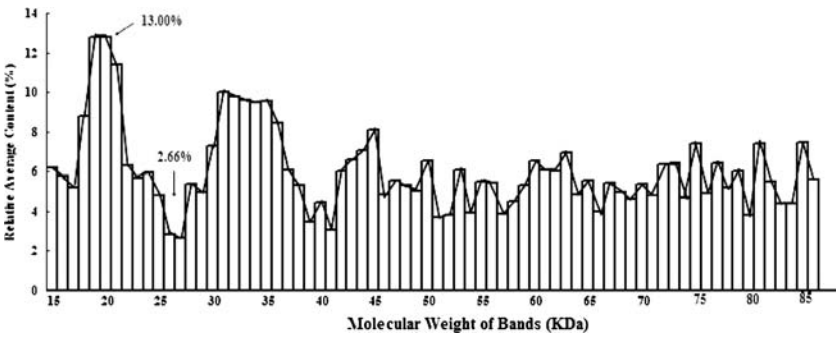

Fig. 3 Distribution of relative average protein content of bands (in both *histogram* and *polygon*) Average content means the relative protein content of all electrophoretic bands with the same molecular weight averaged over the 640 soybean cultivars; relative protein content means the protein content percentage of an electrophoretic band to that of all electrophoretic bands in the same lane. The highest relative content was about 13.00% (corresponding to the band with a MW of 20 KDa) while the lowest relative content was 2.66% (corresponding to a band with a MW of 27 KDa)

frequencies, with bands of MW 18, 25, 30, 40, 46 and 52 KDa more than the others.

Relative Classification of 11S and 7S Components

If 11S and 7S can be distinguished with SDS-PAGE analysis of soy protein extracts, the relative content of 11S and 7S, as well as the subunits can then be determined. The current methods for determining 11S and 7S content and the subunits include procedures such as acidic precipitation, gel filtration, hydroxylapatite chromatography, and ultracentrifuging. Wolf and Briggs [1] classified 11S and 7S relatively according to ultracentrifuge analysis. Based on SDS-PAGE, some works [10, 13, 21] indicated that 11S fell within the band region of MW 14.4–42 KDa and 7S within the band region of 45–91 KDa. However, based on our observation in the 640 cultivars, the frequency distribution was still continuous for bands within MW 42–45 KDa as indicated in Fig. 4, therefore, MW 42 or 45 KDa is not a reasonable critical point for separating the two protein parts. Nevertheless, in the present study on 640 cultivars and landraces, there existed a valley between MW 42 and 45 KDa. Its lowest point corresponding to MW 44 KDa (Fig. 4). So, this point was used to separate the bands into two regions, with the 11S component corresponding to the bands below 44 KDa (14.4–44 KDa) and the 7S component corresponding to bands over 44 KDa (44–91 KDa).

Molecular Weight Region of Subunit Groups

If each band in the region between 14.4 and 91 KDa was regarded as a subunit, every cultivar would have many subunits with varying molecular weight. From a breeding point of view, a simple system of classification criterion considering the whole range of variation of the subunits among various cultivars and landraces is needed, while detailed and complicated subunit classification system is too tedious for practical breeding programs. Here in the total distribution of the 640 SDS-PAGE results, there appeared 25 major peaks (or component distributions), nine of them located in 11S region and 16 of them in 7S in Fig. 4. If each peak is regarded as a band group or a subunit group, there are still too many band groups or subunit groups to count, and also the difference among subunit groups was not distinct enough. The data in Fig. 4 was tried for re-grouping with MW interval of 2, 3 and 4-KDa, respectively, then there appeared differential response and inconsistency in the frequency distribution of 11S and 7S, some reasonable and some not reasonable. Here reasonable means at least reasonable in distinguishing 11S from 7S. The results of re-grouping showed that when 11S and 7S were grouped with MW interval suitable to their own situation, respectively, there appeared four obvious peaks (or component distributions) in 11S with MW interval of 2-KDa and six obvious peaks (or component distributions) in 7S with MW interval of 3-KDa, shown in Fig. 5. Since there showed no obvious peaks and valleys in 11S and 7S under other group intervals, the above best pattern was accepted. Here each peak or component distribution was regarded as a band group or subunit group, therefore, 11S was classified into four subunit groups, i.e. 11S-1 with MW 14.4–22 KDa, 11S-2 with MW 22–26 KDa, 11S-3 with MW 26–34 KDa, and 11S-4 with MW 34–44 KDa; and 7Swas classified six subunit groups, i.e. 7S-1 with MW 44–49 KDa, 7S-2 with MW 49–55 KDa, 7S-3 with MW 55–67 KDa, 7S-4 with MW 67–73 KDa, 7S-5 with MW 73–82 KDa, and 7S-6 with MW 82–91 KDa.

Determination of the 11S/7S Ratio and Relative Content of their Subunits

Multiplying the area by its staining intensity of a band read by Bio-Rad Gel Doc™EQ was used to calculate the relative

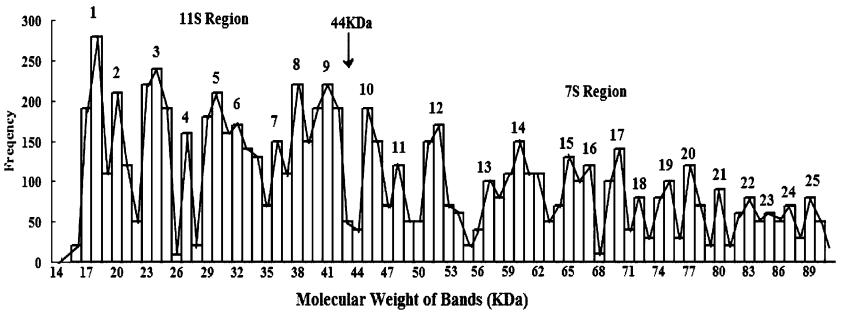

Fig. 4 Frequency distribution of bands with various molecular weights from 14.4 to 91.0 KDa with a1-KDa interval grouping in the 640 soybean cultivars (in both *histogram* and *polygon*). Frequency means the number of cultivars with bands of a same molecular weight. There were 25 component distributions with significant peaks and valleys in the total frequency distribution. 11S region and 7S region were distinguished at MW 44 KDa

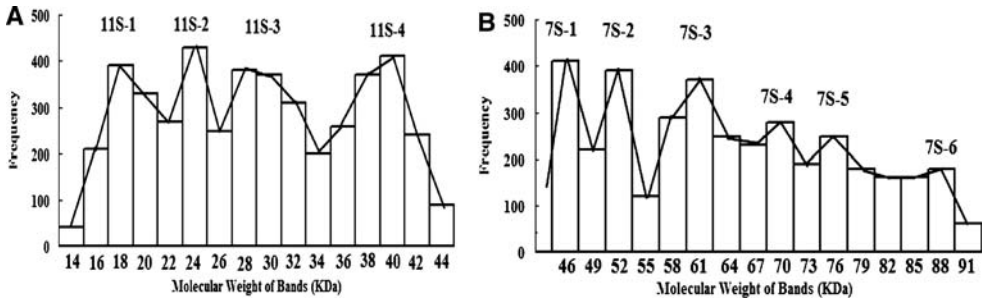

Fig. 5 **a** Frequency distribution of proposed sub-groups in 11S protein (in both *histogram* and *polygon*) The bands were grouped with a 2-KDa interval and resulted in 11S-1 with MW 14.4–22 KDa, 11S-2 with MW 22–26 KDa, 11S-3 with MW 26–34 KDa, and 11S-4 with MW 34–44 KDa. **b** Frequency distribution of proposed sub-groups in 7S protein (in both *histogram* and *polygon*). The bands were grouped with a 3-KDa interval and resulted in 7S-1 with MW 44–49 KDa, 7S-2 with MW 49–55 KDa, 7S-3 with MW 55–67 KDa, 7S-4 with MW 67–73 KDa, 7S-5 with MW 73–82 KDa, and 7S-6 with MW 82–91 KDa

content of a protein band. In a lane (or cultivar), the sum of relative content of all bands in MW 14.4–44 KDa and in MW 44–91KDa was regarded as the relative content of 11S and 7S components of a cultivar, respectively. Then the 11S/7S ratio was obtained from the two relative contents. In turn, the relative content of a subunit group was the sum of all bands in the subunit group.

Comparisons between the Subunit Group Classification and Sathe et al. and Fonte et al.'s Subunits

Table 1 showed the classification criteria of our subunit group classification system along with the subunits proposed by Sathe et al. [13] and Fonte et al. [10]. It is obvious that subunit α' (72 KDa), α (68 KDa), β (52 KDa) of 7S and A_3 (42 KDa), A_{1a}, A_{1b}, A_2, A_4 (37 KDa) of 11S in Fonte et al.'s criterion are located in 7S-4, 7S-2, and 11S-4, respectively; while in Sathe et al.'s criterion, α' (80.22 KDa), α (70.63 KDa), β (48.42 KDa), γ (46.24 KDa) of 7S are located in 7S-5, 7S-4 and 7S-1, respectively, A_{1a}, A_{1b}, A_2, A_4 (33.57 KDa) and A3 (40.74 KDa) of 11S acidic subunits are located in 11S-3 and 11S-4, respectively, and B_{1a}, B_{1b}, B_2, B_4 and B_3 (20.65 KDa) of 11S basic subunits are located in 11S-1. Therefore, Fonte et al. [10] and Sathe et al.'s [13] subunits are only part of our subunit groups, except A_5 with lower molecular weight (10 KDa) is located out of the lower side of 11S-1. It is due to this that only a few genetic materials were used in their studies while 640 cultivars, as a large representative sample, were used in our study. Thus their protein subunit classification system does not cover the whole range of soy protein components, and is inadequate for breeding purposes.

Stability of Classification of Soy Protein Components and Subunit Groups in a Demonstration Experiment

The SDS-PAGE results from the demonstration test of 18 cultivars with four repeated analyses of SDS-PAGE showed that the protein components and subunit groups of the cultivars were easily distinguishedusing the proposed criteria. The results of analysis of variance (Table 2) showed no significant differences among the four repeated analyses for all soy protein components and subunit groups, and their error mean squares and coefficients of variation were pretty small. The consistency of the results of protein component and subunit group classification indicated good precision and stability of the criteria under SDS-PAGE. Accordingly, the differences of relative protein content among the 18 cultivars were due to genotypes rather than shifts in SDS-PAGE analysis. Therefore, the proposed classification procedure and criteria of protein bands is reasonable, stable and feasible, and can be used in the breeding for protein quality soybeans, which usually needs the technology for evaluating a great number of breeding materials.

Frequency Distribution of Subunit Groups of 11S and 7S of the 640 Cultivars

Based on the above conclusions with regard to the precision and stability of the classification of protein components and subunit groups, the results from the 640 cultivars should be an outline of the soybean cultivar population. The frequency distribution of subunit groups in 640 cultivars is shown in Table 3. In 11S region, there were 39 cultivars lacking 11S-1, but none were lacking with regard to the other subunits; while in the 7S region, there appeared to be deficiencies of all the six subunit groups, especially 7S-4, almost half of the cultivars lacked it, and for 7S-5 and 7S-6, about 1/4 or 1/3 of the cultivars lacked them. It is very interesting that there are a great number of cultivars with subunit groups 11S-2, 7S-3 and 7S-6, respectively, which are absent in both Fonte et al. [10] and Sathe et al.'s [13] subunit system. Figure 6 shows that 11S relative content varied between 38.8–80.1% with an average of 67.4%, 7S relative content

Table 1 Molecular weight criterion for classifying subunit groups of soybean protein extracts and comparisons with Fonte's [10] and Sathe's [13] subunits

Component	Sub-group	MW (KDa)	Fonte et al.'s subunit	Sathe et al.'s subunit
11S	11S-1	14.4–22		B_{1a}, B_{1b}, B_2, B_3, B_4
	11S-2	22–26		
	11S-3	26–34		A_{1a}, A_{1b}, A_2, A_4
	11S-4	34–44	A_3, A_{1a}, A_{1b}, A_2, A_4	A_3
7S	7S-1	44–49		γ β
	7S-2	49–55	B	
	7S-3	55–67		
	7S-4	67–73	α α'	A
	7S-5	73–82		α'
	7S-6	82–91		

The subunit α' (72 KDa), α (68 KDa), β (52 KDa) of 7S and A_3 (42 KDa), A_{1a}, A_{1b}, A_2, A_4 (37 KDa) of 11S in Fonte et al.'s criteria are located in 7S-4, 7S-2, and 11S-4, respectively; while in Sathe et al.'s criteria, α' (80.22 KDa), α (70.63 KDa), β (48.42 KDa), γ (46.24 KDa) of 7S are located in 7S-5, 7S-4 and 7S-1, respectively, A_{1a}, A_{1b}, A_2, A_4 (33.57 KDa) and A3 (40.74 KDa) of 11S acidic subunits are located in 11S-3 and 11S-4, respectively, and B_{1a}, B_{1b}, B_2, B_4 and B_3 (20.65 KDa) of 11S basic subunits are located in 11S-1. Both Fonte et al. and Sathe et al.'s subunit A_5 with MW 10 KDa is not included here

Table 2 Results of *F*-test for protein components and subunit groups in 18 cultivars

Source of variation	*df*	11S	7S	11S/7S	11S-1	11S-2	11S-3	11S-4
Cultivar *F*	17	443.28*	449.66*	506.94*	8654.76*	955.64*	1443.39*	1101.85*
Repeat *F*	3	1.57^{NS}	1.80^{NS}	2.33^{NS}	0.40^{NS}	1.79^{NS}	1.07^{NS}	0.51^{NS}
Error MS	51	0.88	0.91	0.01	1.12	0.44	0.86	0.63
CV (%)		1.29	2.87	0.39	12.29	2.66	3.66	3.33
Source of variation	*df*	7S-1	7S-2	7S-3	7S-4	7S-5	7S-6	$F_{0.05}$
Cultivar *F*	17	227.88*	1937.71*	204.82*	758.17*	1150.09*	4685.95*	1.827
Repeat *F*	3	1.54^{NS}	0.76^{NS}	0.93^{NS}	2.02^{NS}	0.04^{NS}	0.56^{NS}	2.786
Error MS	51	0.04	0.40	0.09	0.31	0.48	0.15	
CV (%)		0.89	5.96	1.70	5.91	6.36	6.65	

The *F* value is calculated from the mean square of cultivar and/or repeat divided by error mean square (MS)

df degree of freedom, $F_{0.05}$ is the significant value at probability 0.05

* Indicates significance at least at the 0.05 level, *NS* non-significant, *CV* coefficient of variation

varied between 15.6–61.1% with an average of 32.2%, and 11S/7S ratio varied between 0.6–4.1 with an average of 2.3.

Table 3 Frequency distribution of subunit groups in 640 cultivars

11S Protein			7S Protein		
Sub-group	MW (KDa)	Frequency	Sub-group	MW (KDa)	Frequency
11S-1	14.4–22	601	7S-1	44–49	506
11S-2	22–26	640	7S-2	49–55	528
11S-3	26–34	640	7S-3	55–67	549
11S-4	34–44	640	7S-4	67–73	333
			7S-5	73–82	460
			7S-6	82–91	412

In conclusion, there exists a great variation in the number of protein bands, their molecular weights and amounts, and therefore, 11S and 7S components and 11S/7S ratios as well as their subunit groups in the population of soybean cultivars, which provides a great potential of the genetic resources for the improvement of protein qualities and functional properties.

Discussion

The protein components of soybeans were classified first by Wolf and Briggs [1] based on an ultracentrifuge technique. It was a relatively rough method. Each component was composed of multiple proteins, and 2S, 7S, 11S and

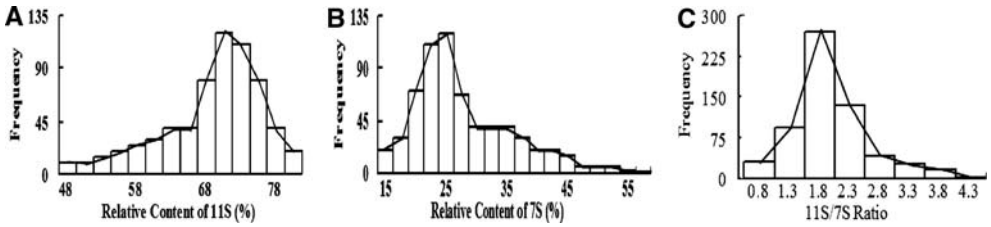

Fig. 6 **a** Frequency distribution of 11S relative content. **b** Frequency distribution of 7S relative content. **c** Frequency distribution of 11S/7S ratio

15S were only the average of each component. Hill and Breidenbach [4] thought that a protein was composed of 2.2S, 7.5S and 11.8S components. Freitas et al. [30] pointed that the range of 11S protein was between 10.8S and 14.0S and the range of 7S was between 6.6S and 8.0S. As was shown on the SDS-PAGE by Mujoo et al. [31], the purified 11S fraction contained about 10% 7S and the purified 7S fraction contained about 10% 11S, therefore, 7S was not separated totally from 11S in his study. In our study, the frequency distribution of SDS-PAGE bands from 640 cultivars also showed that there was no obvious separation between 11S and 7S, but only according to the valley point of the distribution with reference to the molecular weight. Therefore, the separation of protein into components was only a relative criterion, no matter whether the ultracentrifuge method or the SDS-PAGE method was used. However, the ultracentrifuge method is tedious for the separation of 11S from 7S in comparison to SDS-PAGE analysis which is easier and can provide the information on molecular weight and contents of 11S, 7S as well as their subunit groups at the same time, the latter is preferred for handling large samples, such as in the studies on breeding for protein qualities. Anyway, like the sedimentation method, there is still some risk in distinguishing 11S from 7S protein components by using MW 44 KDa as the criterion in SDS-PAGE since the bands in the border area between 11S and 7S are often unstable.

In the present study, our interest was mainly on 11S and 7S from the point of view of functional use, therefore, the soy protein extracts were used to study the protein components and those bands with a MW ofless than 14.4 KDa or more than 97.4 KDa were not detected due to the extraction. Our results were very similar to those obtained from soy protein isolates reported by Arrese et al. [14]. Therefore, our method of classifying subunit groups of soy protein extract should be equivalent to that for soybean protein isolates, and can be used to analyze these isolates.

In our protein component classification system, the neighboring bands were grouped into a subunit group, and four subunit groups were grouped into the 11S fraction and six subunit groups into the 7S fraction. From the results of 640 cultivars and landraces, 11S-1, 11S-3, 7S-1 and 7S-4 accounted for a major part and should be the most important soybean proteins. On the other hand, there appeared a number of cultivars lacked various subunit groups of 7S, which implied that there must be a possibility of using some natural sources for developing 7S subunit group(s) deficient cultivars for functional protein food processing. Of course, not just the existence of protein subunits of 11S and 7S but also the quantity of them should be of concern in soybean breeding since the relative composition determines the functional properties of soy protein [15].

Acknowledgments The project was supported by the Natural Science Foundation of China (9390013, 30490250), the National Key Basic Research Program (2002CB111304, 2004CB7206, 2006CB101708), the National "863" Program (2002AA211052, 2006AA100104), and the Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT). The authors are grateful to Dr. Keshun Liu at USDA-ARS for his valuable comments on and revisions of the manuscript. Specific thanks are also due to the two reviewers for their valuable comments on the manuscript.

References

1. Wolf WJ, Briggs DR (1956) Ultracentrifugal investigation of the effect of neutral salts on the extraction of soybean proteins. Arch Biochem Biophys 63:40–49, 88
2. Wolf WJ, Tamura T (1968) Heat denaturation of soybean 11S protein. Cereal Chem 46:331–344
3. Catsimpoolas N, Rogers DA, Circle SJ , Meyer EW (1967) Purification and structural studies of the 11S component of soybean proteins. Cereal Chem 44:631–637
4. Hill JE, Breidenbach RW (1974) Proteins of soybean seeds. I. Isolation and characterization of the major components. Plant Physiol 53:742–746
5. Kitamura K, Shibasaki K (1975) Isolation and Some physicochemical properties of the acidic subunits of soybean 11S globulin. Agric Biol Chem 39:945–951
6. Kitamura K, Takagi T, Shibasaki K (1976) Subunit structure of soybean 11S globulin. Agric Biol Chem 40:1837–1844
7. Draper M, Catsimpoolas N (1977) Isolation of the acidic and basic subunits of glycinin. Phytochemistry 16:25–27
8. Thanh VH, Shibasaki K (1976) Major proteins of soybean seeds. A straightforward fractionation and their characterization. J Agric Food Chem 24:1117–1121
9. Iibuchi C, Imahor K (1978) Heterogeneity and its relation to the subunit structure of the soybean 7S globulin. Agric Biol Chem 42:31–36
10. Fontes EPB, Moreira MA, Davies CS, Nielsen NC (1984) Urea-elicited changes in relative electrophoretic mobility of certain glycinin and β-conglycinin subunits. Plant Physiol 76:840–842

11. Mori T, Utsumi S, Inaba H (1979) Interaction involving disulfide bridges between subunits of soybean seed globulin and between subunits of soybean and sesame seed globulins. Agric Biol Chem 43:2317–2322
12. Zhang GY, Hayashi Y, Matsumoto S, Mvasomotor Y, Mori T (2002) Molecular species of glycinin in some soybean cultivars. Phytochemistry 60:675–681
13. Sathe SK, Liliey GG, Mason AC, Weaver CM (1987) High-resolution sodium dodecyl sulfate polyacrylamide gel electrophoresis of soybean (*Glycine max* L.) seed proteins. Cereal Chem 64:380–384
14. Arrese EL, Sorgentini DA, Wagner JR, Aňón MC (1991) Electrophoretic, solubility, and functional properties of commercial soy protein isolates. J Agric Food Chem 39:1029–1032
15. Poysa V, Woodrow L, Yu K (2006) Effect of soy protein subunit composition on tofu quality. Food Res Int 39(3):309–317
16. Pesic MB, Vucelic-Radovic BV, Barac MB, Stanojevic SP (2005) The influence of genotypic variation in protein composition on emulsifying properties of soy proteins. J Am Oil Chem Soc 82(9):667–672
17. Manjaya JG, Suseelan KN, Gopalakrishna T, (2007) Radiation induced variability of seed storage proteins in soybean [*Glycine max* (L.) Merrill]. Food Chem 100(4):1324–1327
18. Deak NA, Murphy PA, Johnson LA (2007) Characterization of fractionated soy proteins produced by a new simplified procedure. J Am Oil Chem Soc 84(2):137–149
19. Ruiz-Henestrosa VP, Carrera-Sanchez C, Yust MDMY, Pedroche J, Millan F, Patino JMR (2007) Limited enzymatic hydrolysis can improve the interfacial and foaming characteristics of β-conglycinin. J Agric Food Chem 55(4):1536–1545
20. Natarajan SS, Xu C, Bae H, Caperna TJ, Garrett WM (2006) Characterization of storage proteins in wild (*Glycine soja*) and cultivated (*Glycine max*) soybean seeds using proteomic MDM, analysis. J Agric Food Chem 54(8):3114–3120
21. Tsumura K, Saito T, Kugimiya W, Inouye K (2004) Selective proteolysis of the glycinin and β-conglycinin fractions in a soy protein isolate by pepsin and papain with controlled pH and temperature. J Food Sci 69(5):363–367
22. Samoto M, Maebuchi M, Miyazaki C, Kugitani H, Kohno M, Hirotsuka M, Kito M (2007) Abundant proteins associated with lecithin in soy protein isolate. Food Chem 102(1):317–322
23. Mori T, Nakamura T, Utsumi S (1981) Gelation mechanism of soybean 11S globulin: formation of soluble aggregates as transient intermediates. J Food Sci 47:26–30
24. Gai JY, Wang YS (2001) A study on the varietal eco-regions of soybeans in China. Scientia Agric Sinica 34:139–145
25. Wolf WJ (1970) Soybean proteins: their functional, chemical, and physical properties. J Agric Food Chem 18:969–976
26. Puppo C, Chapleau N, Aperoni F, Lamballerie-Anton MD, Michel F, Aňón C, Anton M (2004) Physicochemical modifications of high-pressure-treated soybean protein isolates. J Agric Food Chem 52(6):1564–1571
27. Puppo MC, Speroni F, Chapleau N (2005) Effect of high-pressure treatment on emulsifying properties of soybean proteins. Food Hydrocolloids 19:289–296
28. Laemmli UK (1970) Cleavage of structural proteins during the assembly of head of bacteriophage T4. Nature 227:680–685
29. Gai JY (2000) Methods of experimental statistics. Agric publishing company in China, Beijing, pp 300–335
30. Freitas RL, Ferrira RB, Teixeira AR (2000) Use of a single method in the extraction of the seed storage globulins from several legume species. Application to analyse structural comparisons within the major classes of globulins. Int J Food Sci Nutri 51:341–352
31. Mujoo R, Trinh DT, Ng PKW (2003) Characterization of storage proteins in different soybean varieties and their relationship to tofu yield and texture. Food Chem 82:265–273