CHAPTER IV

RESULT AND DISCUSSION

The proteins in *Sesbania grandiflora* were analyzed using all methods as described in the previous chapter. The results will be shown and discussed in each part of this chapter, respectively.

4.1 Determination of protein yield

The protein yields at different stages of the protein extraction and precipitation process are shown in Table 4.1. From this table, the amount of protein is higher when precipitated proteins with 20-60% ammonium sulfate saturation compared to 0-20% and 60-90% saturation are lower in protein. In section 2.2.3, protein yield in dried mass flowers (ZMB) is 14.5 % which higher than over all yield of crude protein from fresh flowers (0.056 %). When using the higher amount of ammonium sulfate, the amount of protein is higher.

Table 4.1 The protein yield of *Sesbania grandiflora* flowers from 2 kg fresh flowers

Crude protein precipitation	Amount of protein (mg)	Protein yield (%)
0-20%	244.4	0.012
20-60%	523.2	0.026
60-90%	361.2	0.018

4.2 Biological activity screening of crude protein

Crude proteins (crude 20%, 60%, 90% ammonium sulfate) were tested to check the biological activity. Two activities were performed. First, Hemagglutination activity of crude protein with red blood cell (RBC) of rabbit and goat was determined. The result (Table 4.2) show that crude protein can agglutinate rabbit RBC but not agglutinate goat erythrocytes. The hemagglutinating unit (HU), the reciprocal of the highest dilution exhibiting hemagglutination, and of specific activity, the number of hemagglutinating units per milligram of protein (HU/mg), was illustrated in Table 4.3. From the result, the order of specific activity from highest to lowest is 60%, 90%, 20% crude protein, respectively. There is a little difference of specific activity between 60% and 90% crude protein (Table 4.3).

Table 4.2 Hemagglutination Activity of crude protein

Crude protein	Hemagglutination Activity	
	Rabbit RBC	Goat RBC
20%	+	-
60%	+	-
90%	+	-

Table 4.3 Hemagglutination activity of crude protein with rabbit erythrocytes

Crude protein	Concentration (mg/ml)	Protein (mg)	HU	Specific activity (HU/mg)
20%	0.260	244.4	15.4	0.06
60%	0.960	523.2	266.7	0.51
90%	1.505	361.2	170.1	0.47

Table 4.4 α-Glucosidase inhitory activity of crude protein

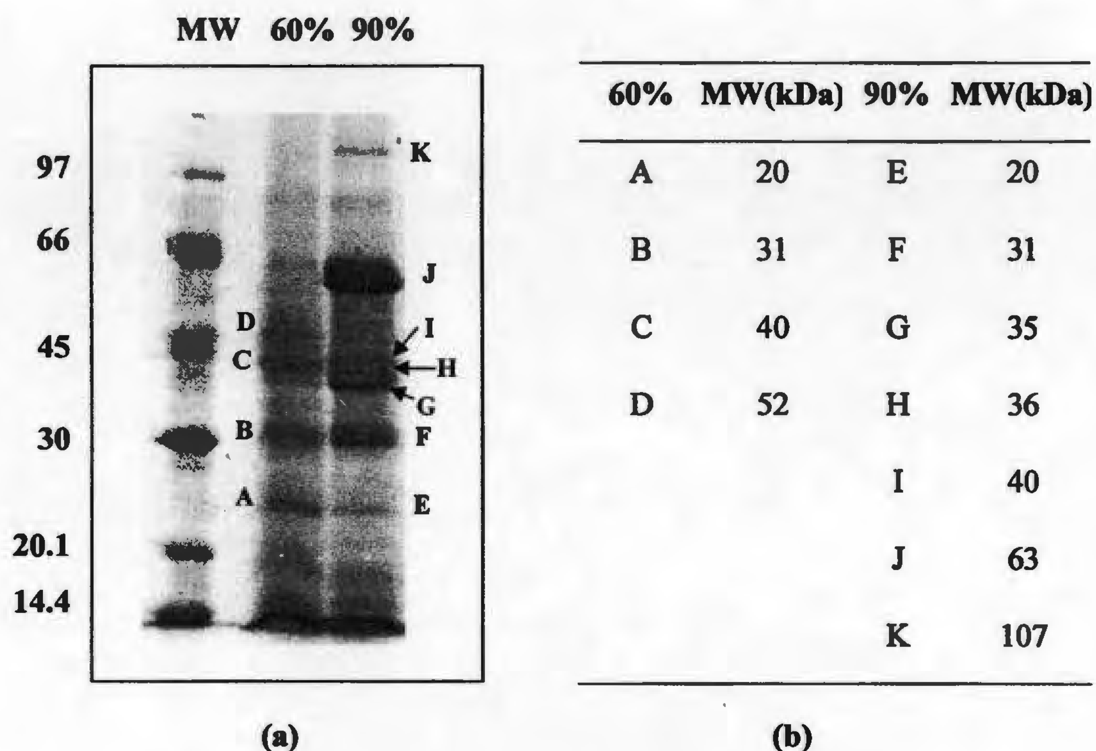
Crude protein	Concentration (mg/ml)	Protein (mg)	% inhibition
20%	0.260	244.4	-
60%	0.960	523.2	49.55
90%	1.505	361.2	82.07

Secondly, α-glucosidase inhibitory activity of crude protein was identified. The result of this activity illustrate in Table 4.4. From this table show that 90% crude protein has higher α-glucosidase inhitory activity than 60% crude protein activity while 20% crude protein not inhibited α-glucosidase. The 20% crude protein show lowest concentration, hemagglutinating activity and α-glucosidase inhitory, thus only 60% and 90% crude were further tested with antioxidant, antifungal and antimicrobial activity. The result of this examined, both crude proteins not presented these all three activity. After that, 60% and 90% crude protein were chosen to further purified protein which effect hemagglutination or α-glucosidase inhitory.

4.3 Characterization of crude proteins by gel electrophoresis techniques

4.3.1 SDS-PAGE

The 60% and 90% crude protein were separated with SDS-PAGE for molecular weight identification. From Figure 4.1(a), both crude proteins were separated and yield many protein bands which presented similar pattern. When plot graph of log molecular weight of standard protein versus Rf, molecular weight of unknown protein can be estimate. The calibration curve of protein standards and approximates protein mass were demonstrate in Figure 4.2 and Figure 4.1 (b).



60%	MW(kDa)	90%	MW(kDa)
A	20	E	20
B	31	F	31
C	40	G	35
D	52	H	36
		I	40
		J	63
		K	107

Figure 4.1 (a) SDS-PAGE of 60% and 90% crude protein (b) Approximates protein mass from SDS-PAGE and calibration curve

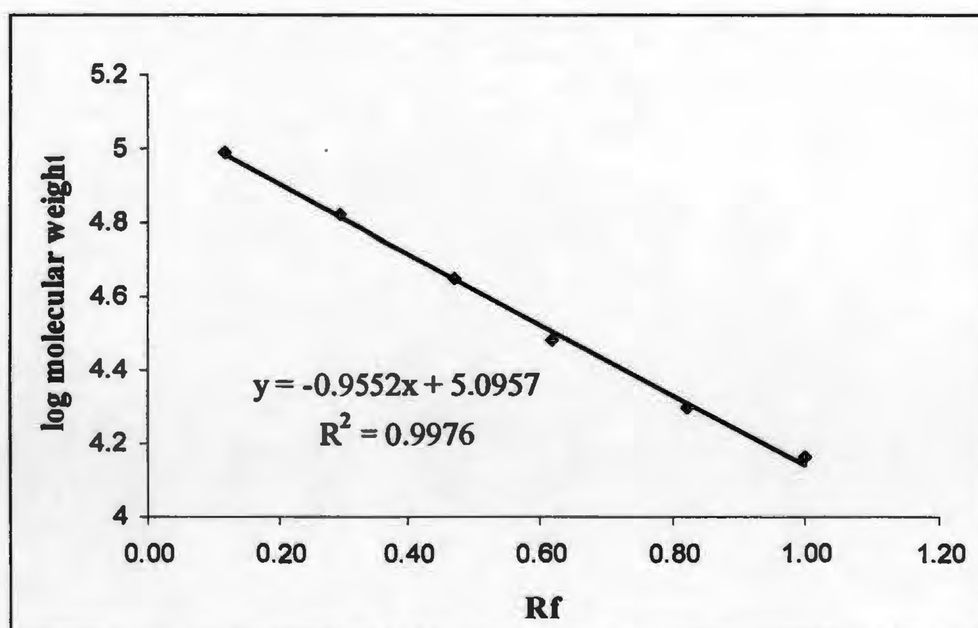


Figure 4.2 Calibration curve of protein standards

4.3.2 2D- gel electrophoresis

Proteins from cell lysates were separated by two-dimensional gel electrophoresis in the first dimension by isoelectric focusing and the second dimension by molecular weight. Separated proteins are visualized by Coomassie straining. The result was shown in Figure 4.3.

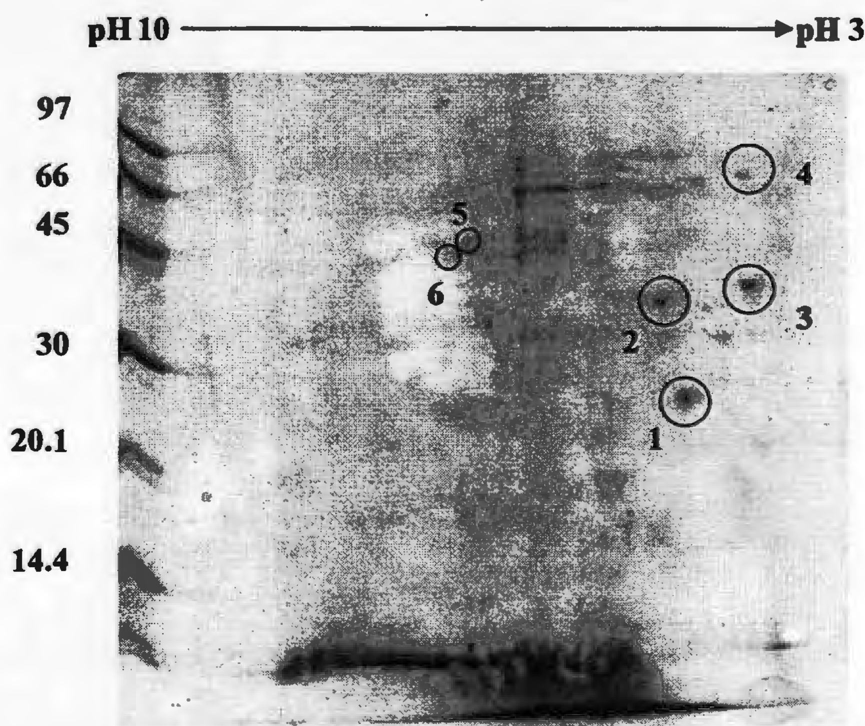


Figure 4.3 2D- gel electrophoresis of *Sesbania grandiflora* flowers crude protein

About six spots of proteins in Coomassie strained gel were presented. The abundance of proteins was summarized in Table 4.5. Protein spots of interest are then excised and subjected to in-gel trypsin digestion as described in section 3.4.1 prior to sequence analysis by MALDI-TOF MS. The mass spectra are shown in Figure 1-6C and the results are shown in Table 1C in Appendices. For proteins identification, the peptide mass fingerprinting data were used for databases searching for family viridiplantae using Mascot program (http://www.matrixscience.com) which set the following criteria: NCBInr database Coverage of the mature protein by the match

must reach a minimum of 15%, at least four independent peptides should match, mass tolerance is 1 Da, and maximum number of missed tryptic cleavages is 1. The searching results of the peptide fingerprint database were shown in Table 4.7. From the search results, there are no signification match between the results and proteins in database. Because of protein from *Sesbania grandiflora* are not in the database.

Table 4.5 Approximation of protein molecular weight and pI from 2D-Gel Electrophoresis

Spot No.	pI	MW(kDa)
1	4.6	30
2	4.8	42
3	4.1	44
4	4.1	67
5	6.9	52
6	7.0	51

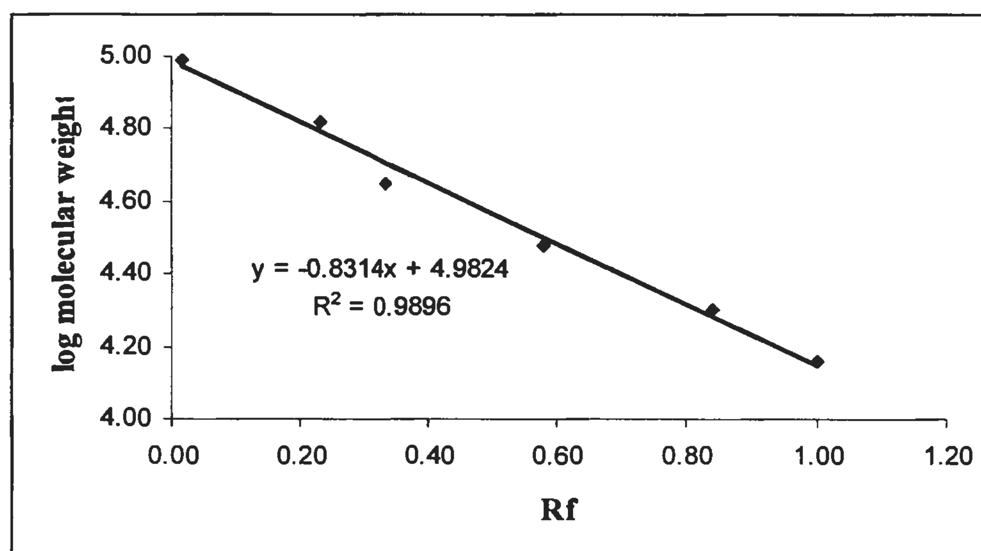


Figure 4.4 Calibration curve of standard protein of 2D-gel electrophoresis

Table 4.6 Peptide Mass Fingerprint Database Search of Selected Spots from 2D Electrophoresis

Spot No.	pI	MW (kDa)	MASCOT results
1	4.6	30	55 = fertilization-related kinase 2*(Solanum chacoense)*
			34 = putative protein*(Arabidopsis thaliana)*
			34 = hypothetical protein*(Arabidopsis thaliana)*
2	4.8	42	37 = GT1a protein*(Arabidopsis thaliana)*
			36 = beta-cyanoalanine synthase like protein *(Solanum tuberosum)*
			36 = putative translation transactivator/inclusion body protein *(Nicotiana tabacum)*
3	4.1	44	38 = OSJNBa0042D13.24 *(Oryza sativa)*
			39 = hypothetical protein *(Cleome spinosa)*
			35 = ser/thr specific protein kinase-like protein *(Arabidopsis thaliana)*
4	4.1	67	43 = ribonuclease II-like protein *(Oryza sativa)*
			38 = maturase *(Chilocarpus suaveolens)*
5	6.9	52	40 = hypothetical protein *(Arabidopsis thaliana)*
			37 = hypothetical protein *(Medicago truncatula)*
6	7.0	51	40 = Citrate synthase 4 *(Oryza sativa)*
			38 = Aminopeptidase *(Oryza sativa)*
			33 = transferase *(Arabidopsis thaliana)*

4.4 Characterization of crude proteins by column chromatography

4.4.1 Characterization of 60% Crude protein

Firstly, three types of ion exchange resin, DEAE-cellulose, Q-sepharose (anion exchange resin) and CM- sepharose (cation exchange resin), and one type of Affinity resin, Affi-gel blue gel, were used to purify 60% crude protein. From the purification result, Q-sepharose, CM-sepharose and Affi-gel blue gel could not separate this crude protein because chromatogram of these (Figure 1-3B in Appendices) shows only unadsorbed fraction. While this protein was loaded on the DEAE-cellulose column, the chromatogram represents a large unadsorbed fraction (D1) and three smaller adsorbed fractions name D2, D3 and D4 (Figure 4.5), which were eluted at a NaCl concentration of 0.1 M (20%B), 0.2 M (40%B) and 0.3 M (60%B), respectively.

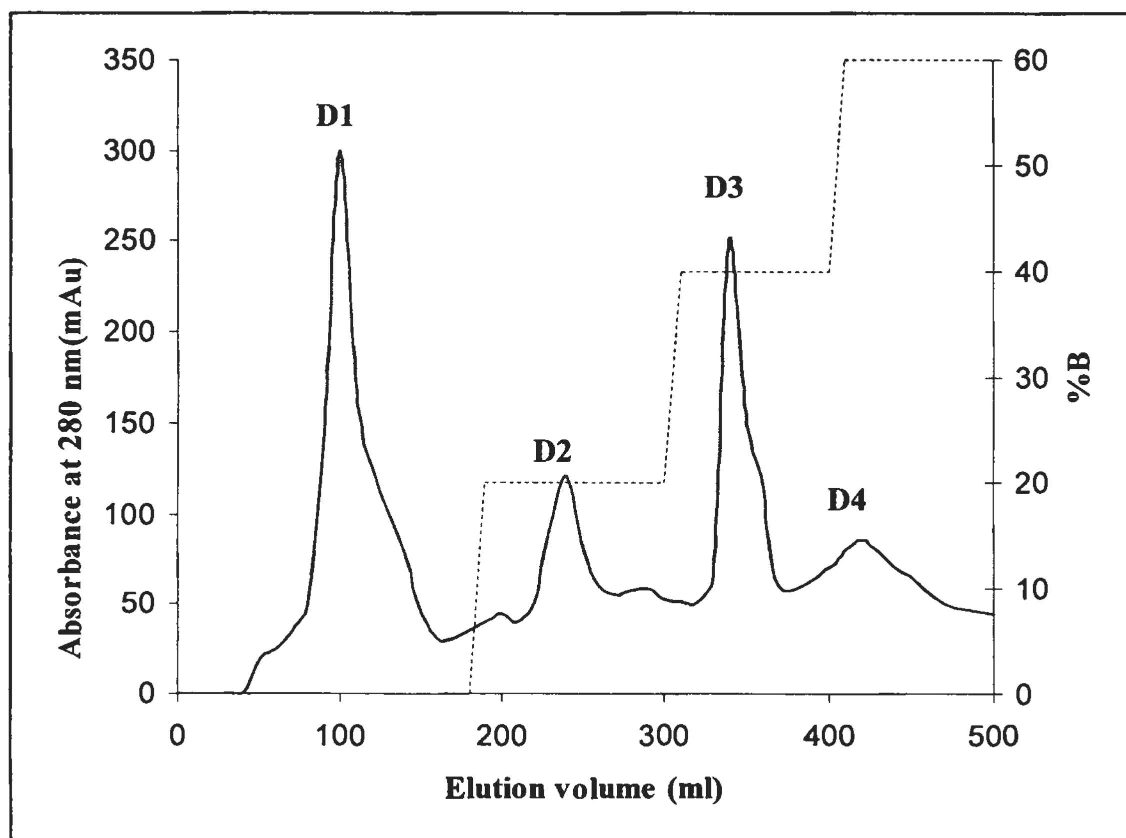


Figure 4.5 Ion exchange chromatography using DEAE-cellulose of 60%crude protein

The protein yields and biological activity of 60% crude protein are presented in Table 4.7. Hemagglutinating activity unit (HU) was found to reside high in the unadsorbed fraction D1 and lower in adsorbed fraction D2 but not found in D3 and D4.

Table 4.7 Concentration and Hemagglutination activity of fractions from DEAE-cellulose ion exchange chromatography

Fraction	Concentration (mg/ml)	Protein (mg)	HU	Specific activity (HU/mg)
D1	0.153	13.5	418	31
D2	0.131	3.93	122	31
D3	0.311	9.33	-	-
D4	0.159	12.72	-	-

After analyzed protein fraction by SDS-PAGE (Figure 4.6), D1 fraction presented a major band protein at 40 kDa in polyacrylamide gel named as SGF60.

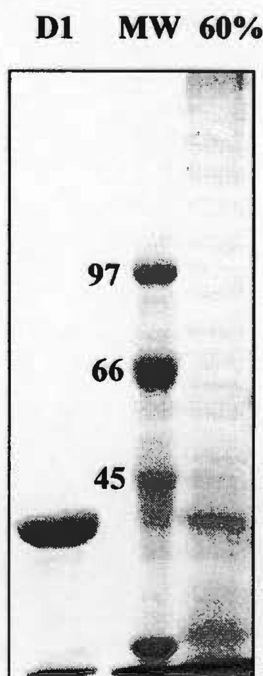


Figure 4.6 SDS-PAGE (12% gel) of fraction D1 (SGF60)

SGF60 was further tested to verify inhibitory activity with α-glucosidase enzyme and the result show that this protein presents the α-glucosidase inhibitory activity (Table 4.8).

Table 4.8 α-Glucosidase inhibitory activity of SGF60

Fraction	Protein yield	AGI(%inhibition)
SGF60	13.5 mg	82

To identify SGF60, The band of this protein was excised and subjected to digestion with trypsin. Amino acid sequences of tryptic fragment were analyzed by ESI-Q-TOF MS. For identification of these proteins, peak lists data (PKL file) containing mass and intensity pairs of precursor peptide has been used. MS/MS Ion search was used to search in databases NCBInr. for viridiplantae using Mascot program (http://www.matrixscience.com). The parameters were set in following criteria: mass tolerance is 200 mmu, and maximum number of missed tryptic cleavages is 1. The results of searching present in Table 4.9.

Table 4.9 MS/MS Ion database searching of SGF60

Protein	MW (kDa)	MASCOT results
SGF60	40	128 = p27SJ *(Hypericum perforatum)*

From Table 4.9, SGF60 matched with p27SJ, a novel protein in St John's Wort (*Hypericum perforatum*) that suppresses expression of HIV-1 genome [68]. This result is significant because the Mowse score of protein matching is higher than the score required to significant (> 36). Even molecular weigh of SGF60 is not identical with mass p27SJ, 26209 Da. Partial amino acid sequence of both should be

relate with SGF60. Sequence coverage percentage of matched peptides of protein SGF60 compare with the protein sequence of p27SJ is 12%. Matched peptides shown in bold line in Figure 4.7.

1 MADINGGGAT LPQALYQTSG VLTAGFAPYI GVGSGNGKAA FLNNDYTKFQ
51 AGVTNK**NVHW AGSDSK**LSAT ELSTYASAKQ PTWGKLIQVP SVGTAVAIPF
101 NKSGTAAVDL SVSELCGVFS GR**ITDWSGIS GSGR**TGAITV VYR**SESSGTT**
151 **ELFTR**FLNAK CAETGTFNIS TTFGTSYTGG LPAGAVSAAG SQGVMTALAG
201 ADGGTTYMSP DFAAPTLAGL DDATKVARVG KDVATNTAGV SPAAANVSAA
251 INAVPVPAST EKP

Figure 4.7 Matched peptide of SGF60 compare with p27SJ

In addition, product ion spectra of tryptic fragments of SGF60 were interpreted by using Peptide Sequencing program on Masslynx software. These sequences (Table1D in Appendix D) were used for protein identification via MS BLAST (http://dove.embl-heidelberg.de/Blast2/msblast.html). Search parameters were set; nrdb 95 database, 30 unique peptides, 100 score table. From the result of database searching, partial amino acid sequences of SGF60 protein matched with p27SJ protein from *Hypericum perforatum* and DING protein from *Solanum tuberosum*. Theses match are statistically significant because HSP (High scoring pair) is higher than the threshold. The results of MS BLAST are show in Table 4.10-4.12. This search result, SGF60 matched with p27SJ protein same as the result from MS/MS ion search which can confirm that SGF60 has amino acid sequences related with p27SJ protein. From MS BLAST result, sequence coverage percentage of matched peptides of protein SGF60 compare with the protein sequence of p27SJ and DING protein is 24.33 and 18%, respectively.

Table 4.10 The results of MS BLAST of SGF60 compare with p27SJ protein

Comparing	Score
Query: 1129 LYQTSGVLTAGFA 1141 LYQTSGVLTAGFA Sbjct: 15 LYQTSGVLTAGFA 27	88
Query: 509 SGANAVDLSVSE 520 SG AVDLSVSE Sbjct: 103 SGTAAVDLSVSE 114	60
Query: 663 CGVFSGR 669 CGVFSGR Sbjct: 116 CGVFSGR 122	52
Query: 157 GGGATLPQ 164 GGGATLPQ Sbjct: 6 GGGATLPQ 13	52
Query: 883 TTELFTR 889 TTELFTR Sbjct: 149 TTELFTR 155	51
Query: 443 SGLSGAGR 450 SG+SG GR Sbjct: 127 SGISGSGR 134	43
Query: 674 TLAGLD 679 TLAGLD Sbjct: 216 TLAGLD 221	37
Query: 1252 STTFAASF 1259 STTF S+ Sbjct: 170 STTFGTSY 177	34

Note + are indicated the amino acid which differ from retrieved sequence

Table 4.11 The results of MS BLAST of SGF60 compare with DING protein

Comparing	Score
Query: 663 CGVFSGR 669 CGVFSGR Sbjct: 31 CGVFSGR 37	52
Query: 510 GANAVDLSV 518 G NAVDLSV Sbjct: 19 GTNAVDLSV 27	52
Query: 883 TTELFTR 889 TTELFTR Sbjct: 64 TTELFTR 70	51
Query: 739 TLAGLD 744 TLAGLD Sbjct: 130 TLAGLD 13	37

Table 4.12 Threshold Scores for Statistical Evaluation of MS BLAST Hits of SGF60

Number of reported HSP[a]	Threshold Score	Observe Score	
		p27SJ protein	DING protein
1	75	88[b]	52
2	131		104
3	154		155
4	185		192[c]
5	234		
6	267		
7	278		
Total score		417	192

a: HSP (High scoring pair) is a region of high local sequence similarity between the peptide in the query and the protein in a database that was identified by database searching.

b: HSP is higher than the threshold in first ranked, the match is statistically significant.

c: HSP is higher than the threshold in forth ranked, the match is statistically significant.

4.4.2 Characterization of 90% Crude protein

The purification of 90% precipitating protein involving gel filtration showed the chromatogram in Figure 4.8. The hemagglutination activity was found in fraction G1 and G2 but not in G3, G4 and G5, followed to the result in Table 4.13. Although G1 has specific activity higher than G2, but this fraction has lower concentration. Consequently, G2 was chosen to purify further. SDS-PAGE was used to identification protein pattern of this active fraction which obtained several bands in Figure 4.9.

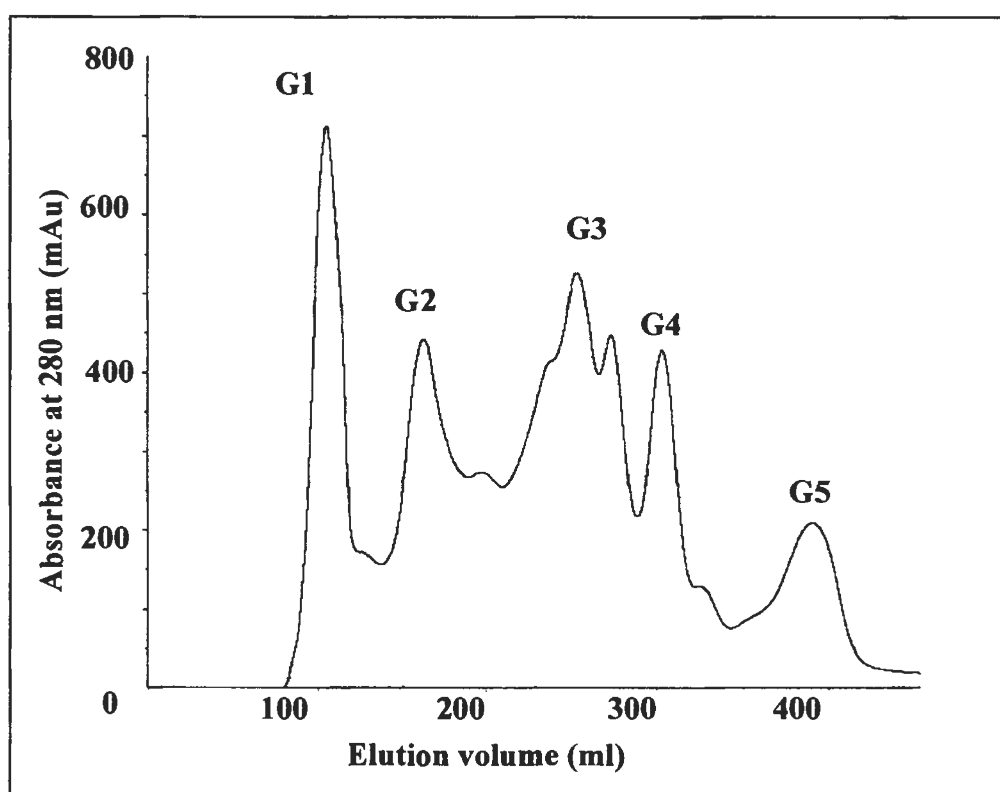


Figure 4.8 Gel filtration chromatography using Superdex-200 of 90% crude protein

Table 4.13 Concentration and Hemagglutination activity of fractions from gel filtration chromatography of 90% crude protein

Fraction	Concentration (mg/ml)	Protein (mg)	HU	Specific activity (HU/mg)
G1	0.260	19.5	61.5	3.1
G2	0.408	36.7	78.4	2.2

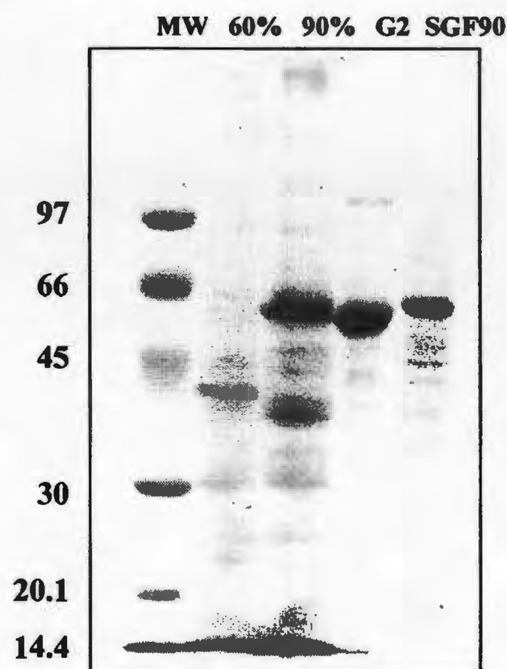


Figure 4.9 SDS-PAGE (15%gel) of fraction G2 and SGF90

When DEAE-cellulose chromatography was used to purify fraction G2, there is sharp peak in unadsorbed fraction (Figure 4.10) ; corresponding to single bane in SDS-PAGE with a molecular mass of 63 kDa and it was called SGF90 (Figure 4.9). And adsorbed fraction is broad peak. These two fractions were then tested with rabbit red blood cell and α-glucosidase enzyme. Only SGF90 protein can agglutinate rabbit erythrocyte and inhibit α-glucosidase. The activity result show in Table 4.14.

Table 4.14 Concentration and Biological activity of fractions SGF90

Fraction	Concentration (mg/ml)	Protein (mg)	HU	Specific activity	AGI (%inhibition)
SGF90	0.048	2.16	8.3	3.8	74

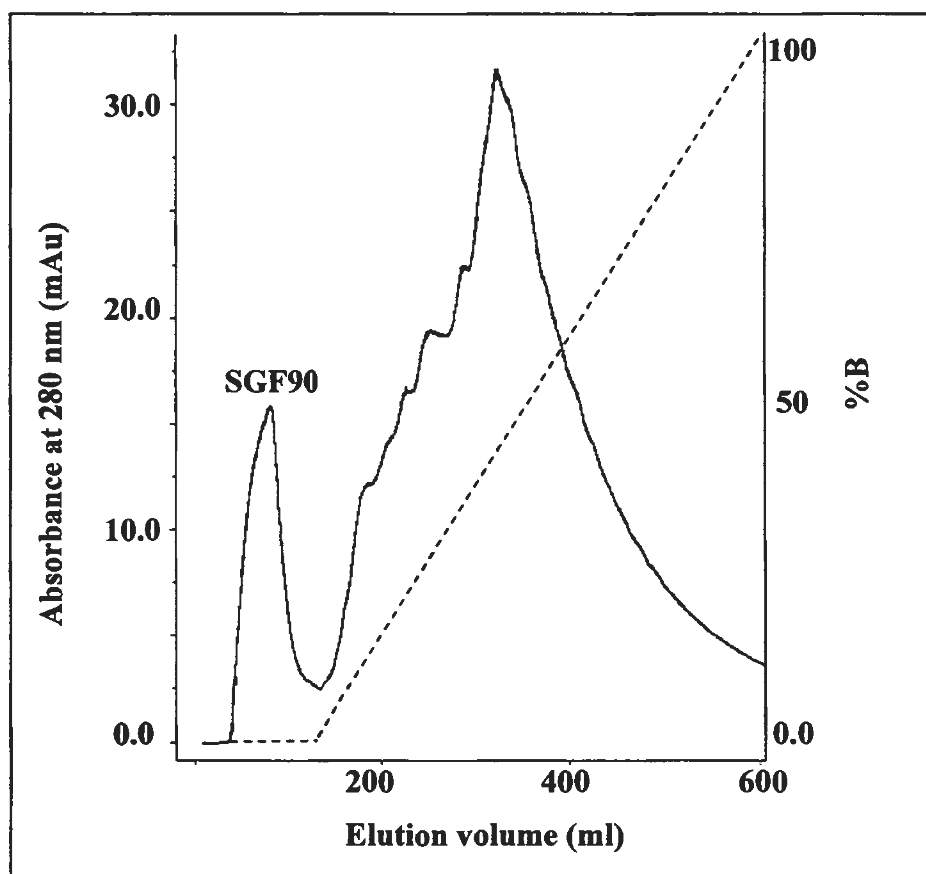


Figure 4.10 Ion exchange chromatography using DEAE-cellulose of G2

To identify the bioactive protein SGF90, single band protein from SDS-PAGE was subjected to tryptic digestion and analyzed peptide mass by using MALDI-TOF MS. The mass spectra of tryptic peptide show in Figure 8C in Appendics. Peptide mass data were used to search in databases NCBInr. for viridiplantae using Mascot program (http://www.matrixscience.com). The parameters were set in following criteria: Coverage of the mature protein by the match must reach a minimum of 15%, at least four independent peptides should match, mass tolerance is 100 ppm, and maximum number of missed tryptic cleavages is 1. The results of searching the peptide fingerprint database were shown in Table 4.15.

Table 4.15 Peptide Mass Fingerprint Database Search of SGF90

Protein	MW app. (Da)	MW cal. (Da)	pI cal	Matching peptide	% Covered, Score	MASCOT results
SGF90	63000	62918	8.74	29(5)	15% , 40	enzyme inhibitor/ pectinesterase *(Arabidopsis thaliana)*

From the result of peptide mass fingerprint, protein scores must be greater than 67 are significant but the score of protein SGF90 is 40 which lower to significant which mean that SGF90 is not pectinesterase but may be has partial amino acid sequences similar to this protein. SGF90 were further identified by tandem mass spectrometry. Tryptic fragments of SGF90 were analyzed using ESI-Q-TOF MS. The conventional mass spectrum is shown in Figure 1D (Appendix D). The thirty product ion spectra were obtained from precursor ions at m/z between 427.47-814.7 kDa and shown in Figure 6-35D in Appendices. These spectra were interpreted to determine the amino acid sequence of peptides by using Peptide sequencing program on Masslynx software. The results of tryptic peptide sequence are shown in Table 2D (Appendix D). These sequences were used for protein identification via MS BLAST (http://dove.embl-heidelberg.de/Blast2/msblast.html). Search parameters were set; nrdb 95 database, 30 unique peptides, 100 score table. From the result of database searching, partial amino acid sequences of SGF90 protein matched with the partial amino acid sequence of beta-glucosidase (At5g36890) protein and beta-glucosidase homolog F8K4.3 protein from *Arabidopsis thaliana* which sequence coverage percentage of matched peptides of protein SGF60 compare with these two beta-glucosidase is 4.9 and 4.7 %, respectively. Comparisons of amino acid sequence between SGF90 and theses beta-glucosidase protein are shown in Table 4.16-4.17.

Table 4.16 The results of MS BLAST of SGF90 compare with beta-glucosidase (At5g36890) protein

Comparing	Score
Query: 1433 GYFAWALLDDFQ 1444 GYFAW LLD+F+ Sbjct: 433 GYFAWSLLDNFE 444	74
Query: 200 YNHPP 204 YNHPP Sbjct: 379 YNHPP 383	43
Query: 135 GLNHY 139 GLNHY Sbjct: 313 GLNHY 317	38
Query: 27 GLAFY 31 G+AFY Sbjct: 113 GIAFY 117	36

Table 4.17 The results of MS BLAST of SGF90 compare with beta-glucosidase homolog F8K4.3 protein

Comparing	Score
Query: 1433 GYFAWALLDDFQ 1444 GYFAW LLD+F+ Sbjct: 470 GYFAWSLLDNFE 481	74
Query: 735 GKYPE 739 GKYPE Sbjct: 314 GKYPE 318	39
Query: 135 GLNHY 139 G+NHY Sbjct: 348 GINHY 352	38
Query: 410 SLADKEA 416 S+ADK A Sbjct: 289 SIADKNA 295	36

Note + are indicated the amino acid which differ from retrieved sequence

Table 4.18 Threshold Scores for Statistical Evaluation of MS BLAST Hits of SGF90

Number of reported HSP[a]	Threshold Score	Observe Score	
		beta-glucosidase (At5g36890)	beta-glucosidase homolog F8K4.3
1	75	74	74
2	131	117	113
3	154	155[b]	151
4	185		187[c]
5	234		
6	267		
7	278		
Total score		191	187

a:HSP(High scoring pair) is a region of high local sequence similarity between the peptide in the query and the protein in a database that was identified by database searching.
b:HSP is higher than the threshold in third ranked, the match is statistically significant.
c:HSP is higher than the threshold in forth ranked, the match is statistically significant.

From Table 4.18, theses match are statistically significant because HSP (High scoring pair) is higher than the threshold. Even though, molecular weigh of SGF90, 63 kDa, is nearly with molecular weigh of beta-glucosidase (At5g36890) protein, 56076 Da, and beta-glucosidase homolog F8K4.3 proteins, 59836 kDa but coverage percentage of these matched protein is lower. So SGF90 should not be these match proteins but partial amino acid sequences of SGF90 similar with these beta-glucosidase.