Front. Chem. China 2010, 5(1): 88–98
DOI 10.1007/s11458-009-0109-5

RESEARCH ARTICLE

Purification and characterization of two major selenium-containing proteins in selenium-rich silkworm pupas

Qiong LIU (✉)[1], Xueying LIANG[1], Decong HU[2], Ping CHEN[1,3], Jing TIAN[1] and Hong ZHANG[4]

Selenium (Se) is an essential trace element *in vivo* involved in the defense against oxidative stress. Se deficiency is associated with many human diseases. The bioactivity of Se is dose- and species-dependent. Silkworm pupa has been reported to accumulate Se mainly in proteins. Thus the characterization of major Se-containing proteins is very important in the application of Se-rich silkworm pupas in food and drugs. In this study, crude proteins were extracted from Se-rich silkworm pupas, followed by DEAE-Sepharose and Sephedex G-75 chromatography. Se content was measured after each step to determine the highest Se-containing fraction for the next step of separation. The proteins obtained were analyzed using SDS-PAGE, followed by in-gel digestion with trypsin, and were characterized by MALDI-TOF MS and ESI-MS/MS. These data showed two proteins mainly accumulated Se in the silkworm pupas. Those two proteins were proven by mass spectrometry to be arylphorin and sex-specific storage-protein 2 precursor (SP-2), respectively. Both of them belong to the storage proteins of amino acids during metamorphosis and the non-feeding pupal stage. The results suggest that Se could be enriched by storage proteins and be supplied to silkworm pupas in accompany with amino acids for the synthesis of new Se-containing proteins and peptides.

Received September 16, 2009; accepted October 29, 2009

1. College of Life Sciences, Shenzhen University, Shenzhen 518060, China
2. Technology Center of Hubei Entry-Exit Inspection and Quarantine Bureau, Wuhan 430022, China
3. Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun 130022, China
4. College of Physics Science and Technology, Shenzhen University, Shenzhen 518060, China
E-mail: liuqiong@szu.edu.cn

Keywords selenium, silkworm pupa, mass spectrometry, arylphorin, sex-specific storage-protein 2 precursor (SP-2)

1 Introduction

Selenium (Se) is an essential trace element in animals [1]. Se deficiency is associated with many human diseases, such as cancer and cardiovascular disease, as well as virus mutation and infection [2–4]. Experimental animal studies have shown that supranutritional exposure of Se can suppress tumorgerosis [5]. Clinical intervention trials have found Se treatment to be associated with reductions in cancer risks [5–7]. Mechanisms of anti-carcinogenicity of Se are owed to the functions of selenoenzymes and Se-metabolites that can defend against carcinogenic free radicals [5]. In addition, it is also possible that one or more Se-compounds/metabolites can function directly in anti-carcinogenic ways [8,9]. There is evidence of anti-carcinogenic activities for several intermediary metabolites of Se, for example, selenodiglutathione, hydrogen selenide, and the methylated metabolites of selenide [5]. Selenomethionine (Se-Met) and selenocysteine (Se-Cys) from food and feedstuff can be metabolized first to hydrogen selenide, and then to methylselenol to exert anti-carcinogenic activities [10].

However, the anti-cancer activity of Se is dose-dependent and species-specific [6,11,12]. There is a narrow range between the physiological selenium intake levels and toxic levels [13]. Over-administration of Se not only causes toxicosis, but also raises possible adverse cardio-metabolic effects [14]. Thus, it is very important to balance the risks and benefits in relation to Se intake [15]. Potential solutions for increasing Se intake include agronomic biofortification and genetic biofortification. For individuals, it includes increased intake of naturally Se-rich foods, functional foods or supplements [15].

Recently, the completion of silkworm genome sequencing [16,17] and proteomics research [18,19] provided a solid foundation for exploring Se-containing proteins for pharmaceutical purposes. Silkworm pupas have been used as feed, food and Chinese traditional medicine since ancient times [10]. Pharmacological studies show that silkworm pupas are alimental for increasing immunity, protecting the liver and preventing cancer [20]. Ziyang County in Shanxi Province is one of the two Se-rich areas in China. Ziyang silkworm pupas had 215 times higher Se content than that of normal silkworm pupas [10]. Se-rich amino acids and peptides extracted from the Se-rich pupas significantly induced the apoptosis of human hepatoma cells SMMC-7721. In order to identify the effective component responsible for the carcinostatic activity, two major Se-containing proteins were isolated from the

Se-rich silkworm pupas in this study and characterized to be arylphorin and sex-specific storage-protein 2 precursor (SP-2) by matrix assisted laser desorption ionization- time of flight-mass spectrometry (MALDI-TOF MS) and electrospray ionization – quadrupole / time of flight- tandem mass spectrometry (ESI-Q-TOF MS/MS).

2 Materials and methods

2.1 Materials

Se-rich silkworm pupas were produced in the Ziyang County of Shanxi Province, one of the two largest Se-rich areas in China. All chemicals used were of analytical grade. Milli-Q water (Bedford, MA) was used in all experiments. Sequencing-grade trypsin was purchased from Roche (UK). Acetonitrile (ACN; Optima grade) was from Fisher Scientific International Inc. (USA), trifluoroacetic acid (TFA) from J. T. Baker (USA), and 2,5-dihydroxy benzoic acid (DHB) from Sigma (USA). Standards of selenium dioxide (A.P) and sodium selenite (A.P.) were also purchased from Sigma.

2.2 Isolation of the major Se-containing proteins

A portion of 800 g Se-rich silkworm pupas was homogenized in 1000 mL of 30 mmol·L^{-1} Tris-HCl buffer solution (pH 7.5) containing 2 mmol·L^{-1} dithiothreitol (DTT), 1 mmol·L^{-1} EDTA, 1 mmol·L^{-1} $MgCl_2$, 0.1% Tween 20, and 1 mmol·L^{-1} PMSF. The homogenate was centrifuged at 800 g for 15 min at 4°C. The supernatant was lyophilized immediately. Proteins in the supernatant were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Major proteins in the Se-rich silkworm pupas were isolated following the methods described previously [21,22]. The lyophilized proteins were extracted with pre-cooled ethyl acetate at −20°C for 10 h. The mixture was centrifuged and the pellet was dried and extracted again for 12 h with 60 mmol·L^{-1} Tris-HCl buffer (pH 7.5). The extraction was precipitated with acetone pre-cooled at 4°C. After centrifugation, the supernatant was dialyzed against Milli-Q water and lyophilized. A portion of obtained crude proteins was suspended in buffer A (30 mmol·L^{-1} Tris-HCl, pH 7.5) and centrifuged at 4°C under 15000 *g* for 20 min. The supernatant was collected, measuring total protein content, and diluted to the concentration of 2 mg·mL^{-1}. An aliquot of 5–8 mL of the protein solution was applied onto a 2.6 cm×20 cm DEAE-Sepharose Fast Flow column (Pharmacia Biotech, Sweden), eluted with a linear gradient of 0.005–0.6 mol·L^{-1} sodium chloride at a flow rate of 45 mL·h^{-1} and 3 mL per fraction. The protein content of the eluted fraction was monitored at 280 nm and the concentration of Se in each peak was measured. The peak with the largest Se content was dissolved in 5 mL buffer A, loaded onto a Sephedex G-75 column (1.6 cm×80 cm, Pharmacia), and eluted with buffer A at a flow rate of 30 mL·h^{-1} and collected at 3.0 mL per fraction. Protein amount and Se content were also measured. Each protein peak was lyophilized.

2.3 Gel electrophoresis

Protein purity was detected by SDS-PAGE on a Hoefer™ miniVE vertical electrophoresis system (Pharmacia) [23]. Aliquots of the proteins after each step of separation were mixed with the sample buffer containing DTT, and run on a 12% polyacrylamide gel under 160 V/gel for 3h. Protein markers with molecular weights of 97, 66.2, 43, 31, 20.1 and 14.4 kDa were run simultaneously with the samples. Purity of the isolated proteins after two steps of chromatography was further detected by the analysis of two-dimensional gel electrophoresis (2D-GE) on IPGphor™ isoelectric focusing system (Pharmacia Biotech, Sweden) and Hoefer™ miniVE vertical electrophoresis system [15]. An Immobiline DryStrip of 7 cm length with linear pH gradients 3–10 was rehydrated overnight in 0.5% IPG buffer [24]. Isoelectric focusing (IEF) was performed at 500 V for 30 min, 1000 V for 30 min, and 8000 V for 2 h. The IPG strip was then balanced with equilibrium buffer, positioned on the surface of a 12% SDS-polyacrylamide gel and electrophoresis was run at 160 V/gel for 3 h.

2.4 Selenium measurement

Two types of Se measurements were performed. Se contents in the samples of silkworm pupas, crude proteins, and column-eluted protein peaks were measured quantitatively by the method of hydride generation-atomic fluorescence spectrometer (HG-AFS) [25]. Each sample was added 8 mL mixed acids of ultra-pure nitric acid/perchloric acid (3∶1, v/v), heated until brown smoke disappeared and the solution became clear. 1 mL 5 mol·L^{-1} hydrochloric acid was successively added and heated for another 10 min. The solution was diluted with 5% hydrochloric acid to 50 mL. Se content was measured on a HG-AFS-2201a atomic fluorescence spectrometer (Beijing Vital Instruments, China). Selenium dioxide was used to prepare a series of Se standard solutions of 10, 20, 40, and 60 g·mL^{-1}, respectively.

Cyclic neutron activation analysis using ^{77m}Se was adopted to detect Se in each protein band of SDS-PAGE, because Se level in each protein band was too low to be detected by AFS [26–28]. Cyclic neutron activation analysis was carried out on a miniature neutron source reactor at Shenzhen University, China. Each protein band in the SDS-PAGE was cut from the

gel, and the control was taken from the top blank of the same lane with the same area size. Those gels were separately heat-sealed into a 1.5-mL EP-tube, and further heat-sealed into another polyethylene capsule. The prepared capsule was irradiated in the inner channel of the reactor. The signal of Se was measured by a high pure germanium coaxial detector (18% relative efficiency, ORTEC) coupled with a high throughput digital spectrometer (DSPEC Plus multichannel analyzer, ORTEC), based on a 161.9 keV γ-peak from ^{77m}Se ($t_{1/2}$ = 17.5 s), under a neutron flux of 9.0×10^{11} n·cm^{-2}·s^{-1} and the conditions of $t_i = t_c = 30$ s, $t_d = t_w = 2$ s for 6 cycles, in which t_i is the irradiation time, t_c is the counting time, t_d is the delay time, and t_w is the waiting time. Sodium selenite was used as a standard.

2.5 Peptide mass fingerprinting by MALDI-TOF-MS

The two major protein bands shown on SDS-PAGE were excised individually from gels, washed and dried for 20 min under vacuum. The slices were digested with 5–10 μL 10 mg·L^{-1} sequencing grade trypsin in 25 mmol·L^{-1} ammonium bicarbonate (pH 8.0) for 20 h. The peptides were extracted, and then mixed with the matrix-solution containing DHB in 0.1% TFA and 50% ACN solution [29]. Aliquots were air-dried onto the sample plate and measured *via* Reflex III matrix MALDI-TOF-MS (Bruker-Franzen, Bremen, Germany) in positive ion mode at an accelerating voltage of 20 kV. The spectra were internally calibrated using trypsin autolysis products.

2.6 Peptide sequencing by ESI-MS/MS analysis

The protein that could not be identified by MALDI-TOF-MS was further analyzed by ESI-MS/MS. The protein band from the SDS-PAGE gel was cut into 1-mm^2 pieces and placed in an Eppendorf tube. 100 μL 50% ACN in 100 mmol·L^{-1} NH_4HCO_3 solution was used to wash the gel for 20 min 3 times until the pieces of gel were completely destained. The gel was dried for 20 min in a SpeedVac centrifuge (Savant Instruments, USA). Then, 5–10 μL of 10 mg·L^{-1} trypsin solution in 25 mmol·L^{-1} NH_4HCO_3 buffer (pH 8.0) was added into the sample, left at 4°C for 40 min until the enzyme solution was completely absorbed by the gel. Another portion of 10 μL 25 mmol·L^{-1} NH_4HCO_3 buffer was added, and the sample was incubated at 37°C overnight. The tryptic peptides were extracted by 50–100 μL 5% TFA at 40°C for 1 h, followed by an equal volume of 2.5% TFA-50% ACN solution at 30°C for 1 h, and finally by 50 μL ACN under sonication. The three portions of extraction solution were mixed up and dried with a vacuum desiccator. The dried sample was dissolved in 3–5 μL 0.5% TFA for complete automation CapLC-ESI-MS/MS analysis [30–32]. All MS/MS measurements were performed in positive ion mode on an electrospray Q-TOF2 hybrid quadrupole/TOF mass spectrometer (Micromass, UK) equipped with a desalting pre-column (5 mm × 320 μm i.d.) and a C_{18} capillary column (25 cm × 180 μm i.d.). The mass spectrometer was operated in the positive ion mode. Mass tolerance was set at less than 0.1 Da for the masses of fragment ions. The instrument was operated at a cone voltage of 50 V and nitrogen was employed as the nebulizer and desolvation gas. MS/MS product ion spectra were recorded using a collision energy with argon as the collision gas. An accelerating voltage of 9.1 kV was used for TOF. MCP detection voltage was 2200 V. The applied voltage to the capillary column was 3000 V. Samples were desalted by the pre-column in the first 3 min after injection and washed out by a linear gradient elution. The tryptic peptides were separated *via* C_{18} capillary column and input into a mass spectrometer. Six precursor ions were selected, respectively, from the MS spectrum obtained from a quadrupole mass analyzer for collision-induced dissociation. Six series of product ions were orthogonally pulsed into a TOF mass analyzer to acquire the mass spectrums.

2.7 Data analysis

Protein mass fingerprinting (PMF) obtained by MALDI-TOF-MS was input into a Mascot search engine available at http://www.matrixscience.com to search through the NCBInr, SwissPort and MSDB databases for protein identification. Searching parameters were set up as follows: type of search: PMF; enzyme: trypsin; cleavage by trypsin: cuts C-term side of KR unless next residue is P; variable modifications: carbamidomethyl (C), oxidation (M); mass values: monoisotopic; protein mass: unrestricted; peptide mass tolerance: ±0.3 Da; peptide charge state: 1+; max missed cleavages: 2; number of queries: 68. Protein score is $-10\mathrm{Log}(P)$, where P is the probability that the observed match is a random event. Protein scores greater than 81 are significant ($P < 0.05$).

Data files obtained from ESI-MS/MS were processed with MaxEnt3 (Micromass, UK) and the amino acid sequences of the peptides were deduced with the peptide sequencing program MasSeq. The set of fragment ions generated by MS/MS acted as a fingerprint for an individual peptide. Search parameters were as follows: precursor ion: singly charged; m/z tolerance: 0.3 Da; intensity threshold: 0.250%. Ion score is $-10\mathrm{Log}(P)$, where P is the probability that the observed match is a random event. Individual ion scores larger than 49 indicate identity or extensive homology ($P < 0.05$). Protein scores are derived from ion scores as a non-probabilistic basis for ranking protein hits. The sequences of three peptides

obtained from MS/MS spectra were input into BLAST program to search through the NCBInr database. Two or more peptides identified in this way are usually sufficient to unambiguously identify a protein [33].

3 Results

3.1 Isolation of the major Se-containing proteins

Total proteins of Se-rich silkworm pupas were obtained after homogenization and centrifugation. The supernatant and pellet were analyzed by SDS-PAGE. Two major proteins were shown on the gel, with molecular weights around 80 kDa (Figure 1(a)). Two major proteins were then isolated through crude protein extraction, ion-exchange chromatography, and gel filtration. Crude proteins were obtained from Se-rich silkworm pupas *via* ethyl acetate precipitation and acetone extraction. Its Se content measured by HG-AFS was 2.26 times higher than that of silkworm pupas. The Se-enriched crude proteins were separated by a DEAE-Sepharose column. Four protein peaks were eluted from this column (Figure 2 (a)), and the largest peak was measured to be the highest Se-containing one. Its Se content was 4.05 times higher than that of silkworm pupas, or 1.79 times higher than crude proteins. This peak was successively separated by a Sephadex G-75 column into another four protein peaks (Figure 2(b)), among which the first one contained the highest Se, measured to be 5.68 times higher than the Se content in silkworm pupas, or 2.51 times higher than crude proteins, or 1.40 times higher than that in the peak 4 from DEAE-Sepharose column.

Proteins obtained from the above steps were further separated by SDS-PAGE and revealed two major proteins indicated as P1 (band 1) and P2 (band 2) in Figure 1(b). Those two bands were cut from the gel in the same area size, individually, for Se detection *via* cyclic neutron activation analysis. The net areas of Se peaks were measured to be 435±33, 506±10, and 514±93 counts/band, respectively, for the control gel, protein band 1, and protein band 2. Although it is difficult to measure the accurate content of Se in each protein band due to the inability of quantitative measurement of protein amount in each band, the result supported the conclusion that Se was contained in two major proteins of Se-rich silkworm pupas.

To investigate whether the two bands are from the subunits of one protein, the obtained proteins were also analyzed by SDS-PAGE using the sample buffer without DTT to break disulfide bridges. Figure 1(c) showed that proteins in reduction (with DTT) and non-reduction (without DTT) solutions had the same bands on SDS-PAGE, implying that those two closely located bands were two proteins with molecular weights around 80 kDa instead of subunits of one protein. To avoid protein overlap at the same position on SDS-PAGE, 2D-GE was also performed for the isolated proteins. As shown in Figure 1(d), only two protein points were observed in the proper range of molecular weights, with isoelectric points (pI values) of 5.1 and 4.9 respectively for P1 (point 1) and P2 (point 2), corresponding to protein band 1 and band 2 in Figure 1(b) & (c).

3.2 Peptide mass fingerprinting by MALDI-TOF MS

To identify the two major Se-containing proteins, two bands of P1 and P2 on the SDS-PAGE gel were analyzed by MALDI-TOF MS. The mass spectra of both proteins were obtained by MALDI-TOF MS. However, the *m/z* data from band 1 was impossible to be identified by Mascot search in the

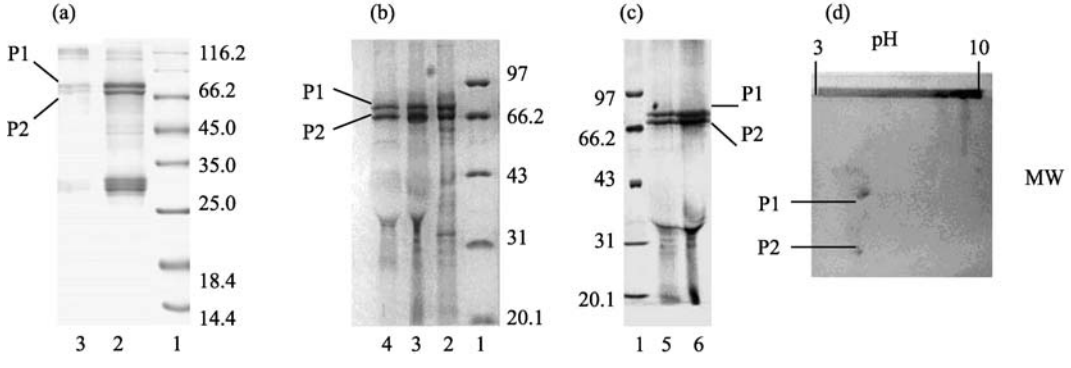

Figure 1 Gel electrophoresis analyses of the proteins isolated from the Se-rich silkworm pupas. (a)~(c) SDS-PAGE analysis of proteins from each separation steps. Lane 1: protein markers with molecular weights shown in the pictures (kDa). (a) Total proteins of the Se-rich silkworm pupas. Lane 2: proteins in the supernatant of homogenate; lane 3: proteins in the precipitate; (b) Stepwise isolated proteins. Lane 2: proteins from crude extraction; lane 3: peak 4 from DEAE Sepharose column; lane 4: peak 1 from Sephadex G-75 column. (c) Proteins without DTT treatment. Lane 5: under non-reductive condition, lane 6: under reductive condition. (d) Two dimensional gel electrophoresis analysis of the isolated protein after crude protein extraction, DEAE Sepharose chromatography, and Sephadex G-75 gel filtration. The Immobiline DryStrip has a 7-cm length and a linear pH gradient of 3–10.

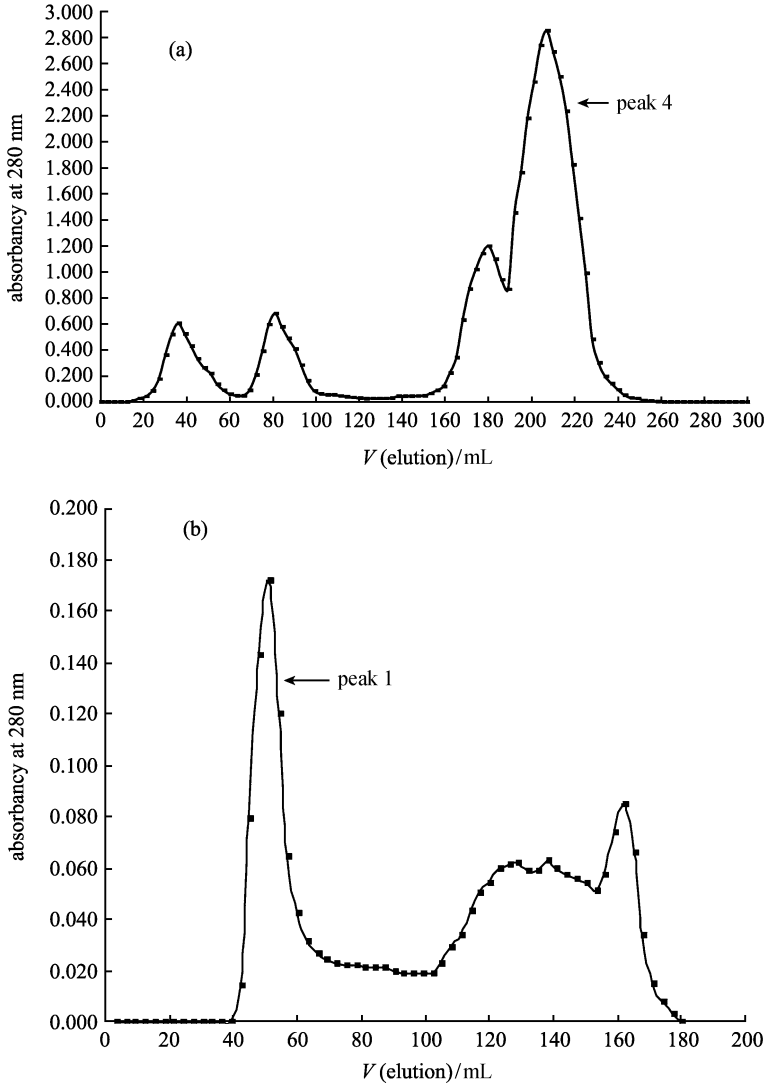

Figure 2 Isolation of Se-containing proteins from the silkworm pupas by DEAE-Sepharose fast flow chromatography (a) and Sephadex G-75 gel filtration (b). The arrows indicate the peaks collected for the next step of separation.

available databases of NCBInr, or MSDB, or SwissPort. On the contrary, protein band 2 was successfully identified by its PMF (Figure 3). The obtained *m/z* data of band 2 were input into the Mascot engine to search through the NCBInr database, and the results contained 6564587 sequences and 2241664500 residues to the date of 20080606. The top Mascot search result matched to a protein, named arylphorin [*Bombyx mori*], with NCBI locus at NP_001037590.1 (top score: 316; expect: $1.6e^{-25}$; queries matched: 46). The rank 2 match was SP-2, with NCBI locus at P20613 (score: 190, expect: $6.6e^{-13}$, queries matched: 37). Figure 4 shows the probability based Mowse score obtained from a Mascot search for band 2. Protein scores greater than 81 are considered significant ($P < 0.05$). Thus arylphorin was the best match to band 2, with a sequence coverage of 68%, as shown in Figure 5 by bold italic letters. Table 1 listed the observed mass values matched and their peptide sequences. Mass values with no match to arylphorin were as follows: 1030.5300, 1038.5200, 1128.4400, 1372.6100, 1394.6400, 1474.5700, 1480.6400, 1573.7900, 1642.7800, 1685.7900, 1768.8700, 1805.9200, 1983.9200, 2026.0400, 2040.0100, 2161.4700, 2183.2200, 2198.0300, 2273.2100, 2653.6800, 2675.4800, 3649.5100. As presented in the NCBI databases, the matched arylphorin has a nominal mass (M_r) 83398 and a calculated pI value of 5.70.

3.3 Peptide sequencing by ESI-Q-TOF MS/MS

Since protein band 1 in the gel of SDS-PAGE could not be identified by MALDI-TOF MS, peptide sequencing was performed by ESI-Q-TOF MS/MS to characterize this protein. Figure 6(a) showed the MS spectrum of proteolytic peptides

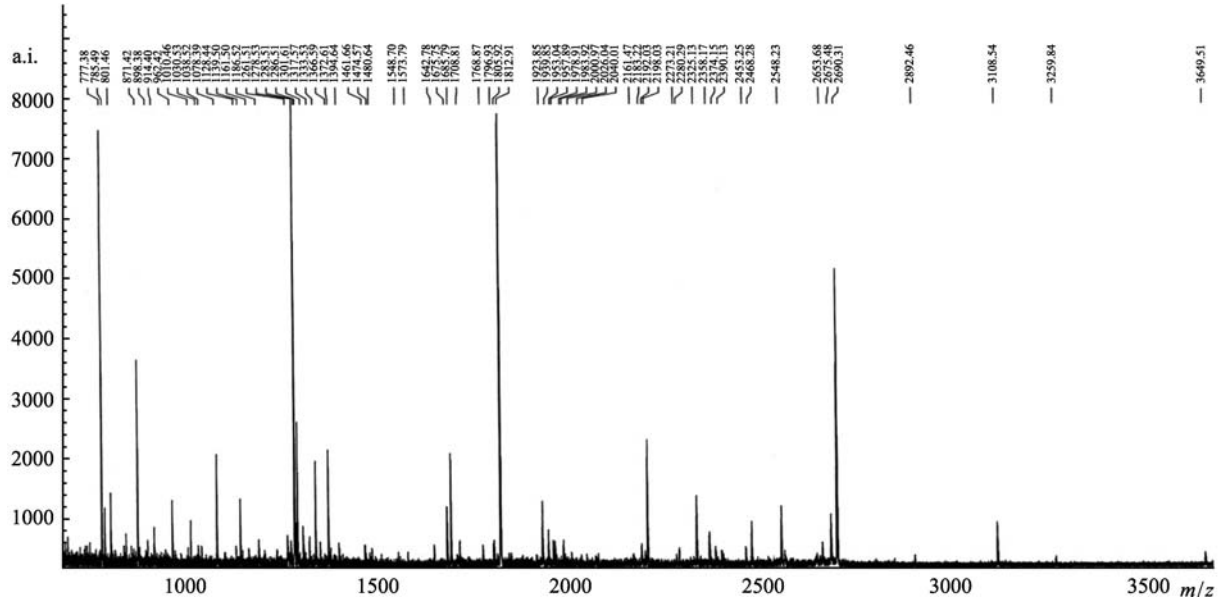

Figure 3 Peptide mass fingerprinting of the protein band 2 (P2) in the SDS-PAGE gel. The proteins were isolated from Se-rich silkworm pupas through three-step separation before they were applied onto SDS-PAGE, characterized by MALDI-TOF MS, and Mascot searched through databases.

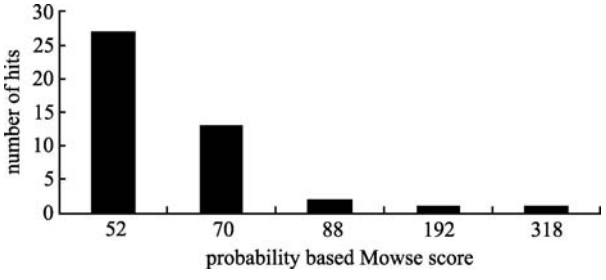

Figure 4 Probability based Mowse score obtained from Mascot search for the protein band 2 (P2) in the SDS-PAGE. Protein score is $-10\mathrm{Log}(P)$, where P is the probability that the observed match is a random event. Protein scores greater than 81 are significant ($P < 0.05$). Search parameters were described in the Materials and Methods.

1 MKSVLILAGL VAVALSSAVP KPSTIKTKNV DAVFVEKQKK *ILSFFQDVSQ*
51 *LNTDDEYYKI GKDYDIEMNM DNYTNKKAVE EFLKMYRTGF MPKNLEFSVF*
101 *YDKMR*DEAIA LFHLFYYAKD FETFYKTACF AR*VHLNQGQF LYAFYIAVIQ*
151 *RSDCHGFVVP APYEVYPKMF MNMEVLQKIY VTK*MQDGLIN PEAAAK*YGIH*
201 *KENDYFVYK*A NYSNAVLYNN EEQRLTYFTE DIGMNAYYYY FHSHLPFWWT
251 SEKYGALKER *RGEVYFYFYQ QLLARYYFER* LTNGLGK*IPE FSWYSPIKTG*
301 *YYPLMLTKFT PFAQRPDYYN LHTEENYER*V RFLDTYEK*TF VQFLQKDHFE*
351 *AFGQK*IDFHD PK*AINFVGNY WQDNADLYGE EVTKDYQRSY EVFARRVLGA*
401 *APMPFDKYTF MPSAMDFYQT SLR*DPAFYQL YNR*IVEYIVE FKQYLKPYTQ*
451 *DK*LYFDGVKI TDVKVDK*LTT FFENFEFDAS NSVYFSKEEI KNNHVHDVKN*
501 RQPR*LNHSPF NVNIEVDSNV ASDAVVK*IFL APK*YDDNGIP LTLEDNWMKF*
551 *FELDWFTTK*L TAGQNKIIR*N SNEFVIFKED SVPMTEIMKM LDEGKVPFDM*
601 *SEEFCYMPKR LMLPR*GTEGG FPFQLFVFVY PFDNKGK*DLA PFESFVLDNK*
651 *PLGFPLDRPV VDALFKVPNM YFKDIFIYHE GERFPYKFNI PSYDTQSNVV*
701 *PKN*

Figure 5 The amino acid sequence of arylphorin [*Bombyx mori*] (accession number: NP_001037590.1) and its matched peptides from the protein band 2 (P2). Mass data of P2 were obtained by MALDI-TOF MS, and input into Mascot to search for the matched peptides shown in italic bold letters.

digested by trypsin from band 1. Mass values of those precursor ions were also input into the Mascot engine to search through the databases of NCBInr, MSDB, and SwissPort. Similar to the result of MALDI-TOF MS, no protein in the present database matched the band 1. Thus, six precursor ions at *m/z* 443.75, 488.26, 572.29, 697.86, 733.37, and 772.39 were selected, respectively, from the MS spectrum for collision-induced dissociation. Six series of product ions were pulsed into the TOF mass analyzer to acquire their corresponding MS/MS spectra. Data files from those MS/MS spectra were processed using MaxEnt 3 software. Short stretches of sequences can be deduced from the ladder of the y-type fragments. The obtained peptide sequences were searched by BLASTp in the NCBInr database. Three of the six selected peptides matched to one protein, SP-2. Figure 6(b, c,d) demonstrates the MS/MS spectra of three peptides selected from Figure 6(a) at *m/z* 488.26 (b), 572.29 (c), and 697.86 (d), respectively. The quality of the fragmentation data in each MS/MS spectrum was sufficient to determine the amino acid sequence of the peptide as indicated in Figure 6 (b, c,d) and Table 2. The peptide sequences from MaxEnt 3 were searched through the NCBInr database. BLAST results showed that the peptides with *m/z* 488.26 and 572.29 had their top scores matched to SP-2 (NCBI locus: P20613). The peptide with *m/z* 697.86 matched to arylphorin (partial)

Table 1 Mascot search results for protein band 2 characterized by MALDI-TOF MS

start-end	observed	M_r(expt)	M_r(calc)	delta	miss	sequence
40–59	2453.2500	2452.2427	2452.1798	0.0629	1	K.KILSFFQDVSQLNTDDEYYK.I
41–59	2325.1300	2324.1227	2324.0849	0.0379	0	K.ILSFFQDVSQLNTDDEYYK.I
60–77	2192.0300	2191.0227	2190.9925	0.0302	2	K.IGKDYDIEMNMDNYTNKK.A
78–87	1301.6100	1300.6027	1300.6485	−0.0458	1	K.AVEEFLKMYR.T Oxidation (M)
78–93	1978.9100	1977.9027	1977.9692	−0.0665	2	K.AVEEFLKMYRTGFMPK.N 2 Oxidation(M)
94–103	1261.5100	1260.5027	1260.6026	−0.0999	0	K.NLEFSVFYDK.M
94–105	1548.7000	1547.6927	1547.7442	−0.0515	1	K.NLEFSVFYDKMR.D
133–151	2280.2900	2279.2827	2279.2215	0.0612	0	R.VHLNQGQFLYAFYIAVIQR.S
152–183	3811.5700	3810.5627	3810.8170	−0.2542	2	R.SDCHGFVVPAPYEVYPKMFMNMEVLQ KIYVTK.M 3 Oxidation (M)
169–178	1286.5100	1285.5027	1285.5869	−0.0841	0	K.MFMNMEVLQK.I Oxidation (M)
197–209	1675.7500	1674.7427	1674.8042	−0.0615	1	K.YGIHKENDYFVYK.A
261–275	1953.0400	1952.0327	1951.9944	0.0383	1	R.RGEVYFYFYQQLLAR.Y
262–275	1796.9300	1795.9227	1795.8933	0.0294	0	R.GEVYFYFYQQLLAR.Y
276–280	777.3800	776.3727	776.3493	0.0234	0	R.YYFER.L
288–298	1366.5900	1365.5827	1365.6969	−0.1141	0	K.IPEFSWYSPIK.T
299–308	1186.5200	1185.5127	1185.6104	−0.0976	0	K.TGYYPLMLTK.F
309–329	2690.3100	2689.3027	2689.2197	0.0830	0	K.FTPFAQRPDYYNLHTEENYER.V
339–346	1010.4600	1009.4527	1009.5597	−0.1070	0	K.TFVQFLQK.D
347–355	1078.3900	1077.3827	1077.4880	−0.1052	0	K.DHFEAFGQK.I
363–388	3108.5400	3107.5327	3107.4261	0.1067	1	K.AINFVGNYWQDNADLYGEEVTKDYQ R.S
389–395	871.4200	870.4127	870.4236	−0.0108	0	R.SYEVFAR.R
396–407	1317.5700	1316.5627	1316.6911	−0.1284	1	R.RVLGAAPMPFDK.Y Oxidation (M)
397–407	1161.5000	1160.4927	1160.5900	−0.0973	0	R.VLGAAPMPFDK.Y Oxidation (M)
408–423	1957.8900	1956.8827	1956.8750	0.0077	0	K.YTFMPSAMDFYQTSLR.D
434–442	1139.5000	1138.4927	1138.6274	−0.1347	0	R.IVEYIVEFK.Q
443–452	1283.5100	1282.5027	1282.6557	−0.1530	0	K.QYLKPYTQDK.L
468–491	2892.4600	2891.4527	2891.3541	0.0986	1	K.LTTFFENFEFDASNSVYFSKEEIK.N
488–499	1461.6600	1460.6527	1460.7372	−0.0844	1	K.EEIKNNHVHDVK.V
492–499	962.4200	961.4127	961.4730	−0.0603	0	K.NNHVHDVK.V
505–527	2468.2800	2467.2727	2467.2343	0.0384	0	R.LNHSPFNVNIEVDSNVASDAVVK.I
534–549	1923.8500	1922.8427	1922.8720	−0.0293	0	K.YDDNGIPLTLEDNWMK.F
534–549	1939.8500	1938.8427	1938.8669	−0.0242	0	K.YDDNGIPLTLEDNWMK.F Oxidation (M)
550–559	1333.5300	1332.5227	1332.6391	−0.1163	0	K.FFELDWFTTK.L
570–589	2374.1500	2373.1427	2373.1232	0.0195	1	R.NSNEFVIFKEDSVPMTEIMK.M Oxidation (M)
570–589	2390.1300	2389.1227	2389.1181	0.0046	1	R.NSNEFVIFKEDSVPMTEIMK.M 2 Oxidation (M)
579–595	2000.9700	1999.9627	1999.8788	0.0839	1	K.EDSVPMTEIMKMLDEGK.V 3 Oxidation (M)
596–615	2546.2300	2545.2227	2545.1990	0.0237	2	K.VPFDMSEEFCYMPKRLMLPR.G Carbamidomethyl (C)
610–615	785.4900	784.4827	784.4741	0.0086	1	K.RLMLPR.G
610–615	801.4600	800.4527	800.4691	−0.0163	1	K.RLMLPR.G Oxidation (M)
638–666	3259.8400	3258.8327	3258.7329	0.0998	0	K.DLAPFESFVLDNKPLGFPLDRPVVDALF K.V
667–673	898.3800	897.3727	897.4418	−0.0691	0	K.VPNMYFK.D
667–673	914.4000	913.3927	913.4368	−0.0440	0	K.VPNMYFK.D Oxidation (M)
674–683	1278.5300	1277.5227	1277.6040	−0.0813	0	K.DIFIYHEGER.F
674–687	1813.9100	1812.9027	1812.8835	0.0192	1	K.DIFIYHEGERFPYK.F
684–703	2358.1700	2357.1627	2357.1692	−0.0065	2	R.FPYKFNIPSYDTQSNVVPKN.-
688–702	1708.8100	1707.8027	1707.8468	−0.0441	0	K.FNIPSYDTQSNVVPK.N

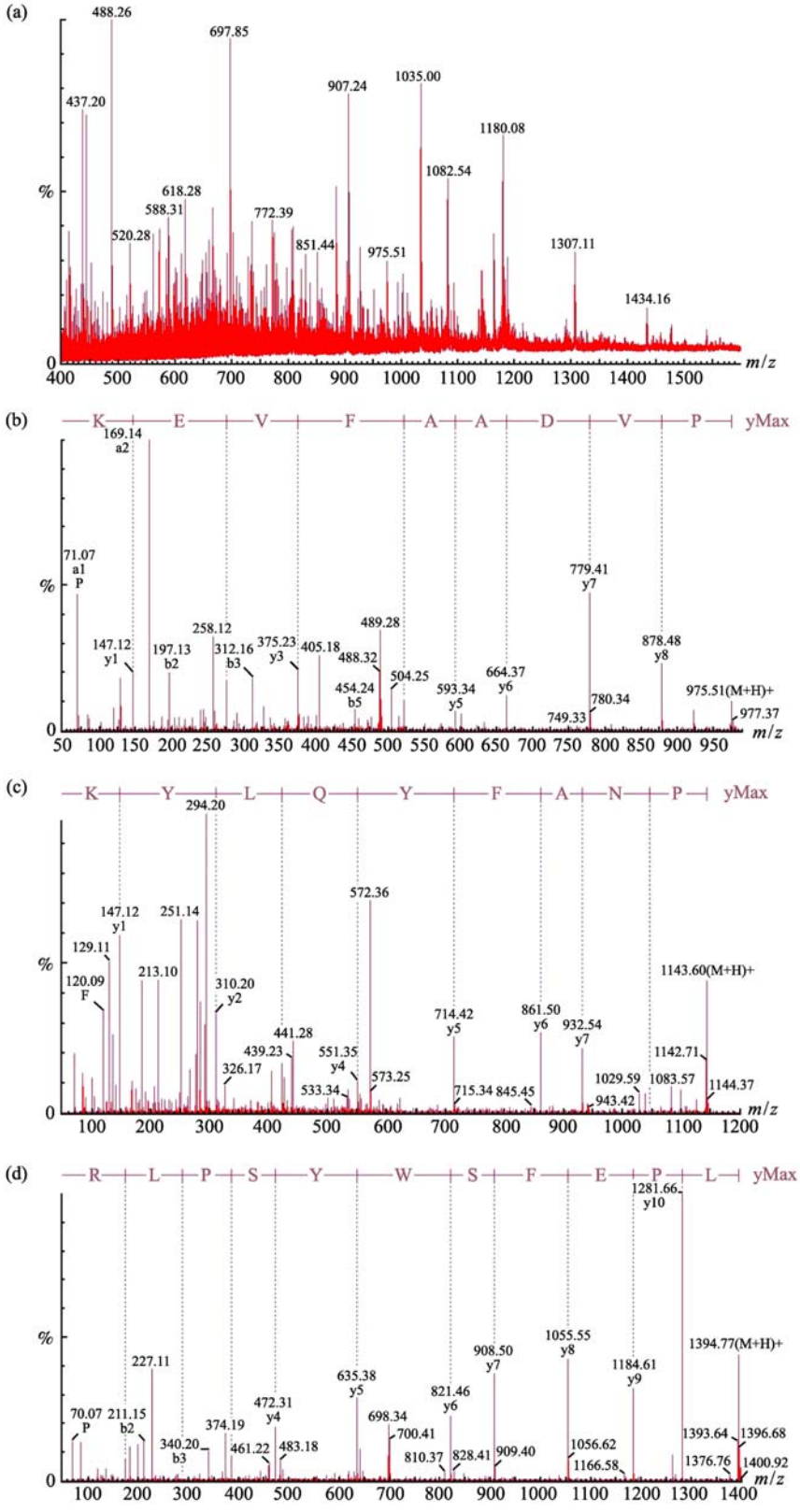

Figure 6 ESI-Q-TOF MS/MS analyses of the protein band 1 (P1) in the SDS-PAGE gel. (a) MS spectrum of precursor ions from the peptides of trypsin-digested P1. (b) (c) (d) MS/MS spectra of the doubly charged precursor ions at *m/z* 488.26 (b), 572.29 (c), and 697.86 (d), respectively. Those product ions were arising from the tryptic fragments in the MS spectrum. Short stretches of sequences were deduced from the ladder of y-type fragments.

Table 2 NCBI search results for protein band 1 characterized by ESI-Q-TOF MS/MS

m/*z*	charge	M_r(obs)	MaxEnt 3 sequence	Blast results NCBI locus	name	score	E-value
488.26	2	974.5044	PVDAAFVEK	P20613	Sex-specific storage-protein 2 precursor	24.4	0.30
572.29	2	1142.5643	PNAFYQLYK	P20613	Sex-specific storage-protein 2 precursor	24.4	0.30
697.86	2	1393.7043	LPEFSWYSPLR	BAA07930.1	arylphorin (partial) [*Bombyx mori*]	32.0	5.1
		Combination of three peptides		P20613	Sex-specific storage-protein 2 precursor	35.0	$3e^{-04}$

[*Bombyx mori*], which is also part of the sequence of SP-2. When the selected three peptides were simultaneously input into the BLASTp, NCBI search results revealed a higher score and much lower E value for SP-2, suggesting that protein band 1 is homogenous to SP-2. As presented in the NCBI databases, SP-2 has a nominal mass 83412 and a calculated pI value of 6.04.

4 Discussion

In this paper, two major Se-containing proteins in the Se-rich silkworm pupas were isolated and identified to be arylphorin and SP-2, respectively (Table 3), which were observed as two closely located bands in the gel of SDS-PAGE. Bioinformatics analyses showed that the two proteins contain high proportions of methionines (Met, M) and several cysteines (Cys, C). This could possibly be one reason for them to be the major Se-containing proteins in the Se-rich silkworm pupas since Se can intrude in the sulfur pathways and metabolize along the routes of sulfur metabolism when the living system is exposed to a high Se environment. The nonspecific metabolism rests on the chemical similarity between Se and sulfur. Similar phenomena have already been reported in bacteria [34]. Regarding the Se-rich silkworm pupas, they were fed with Se-rich mulberry leaves grown in Ziyang County [10]. The high Se environment enables Se to intrude in the sulfur pathways of silkworm, through substituting Met for Se-Met and Cys for Se-Cys in proteins. Since both arylphorin and SP-2 are Met-rich proteins, it is reasonable for them to become the major Se-containing proteins in the Se-rich silkworms. This conclusion is also supported by our previous result that Se-Met was the major seleno-amino acid in the proteins of Se-rich silkworm pupas [10].

Se is an essential trace element for mammals. Se-deficiency is associated with many animal diseases, e. g., white muscle disease in sheep, exudative diathesis in poultry, liver necrosis in rats, Keshan disease and Kaschin-Beck disease in human. An adequate amount of Se is mainly present as various kinds of selenoproteins. However, when animals are exposed to a Se-rich environment, Se-containing proteins, rather than selenoproteins, become the major Se-accumulating proteins. In addition, bioinformatics analyses showed that silkworm (*Bombyx mori*) belongs to the selenoproteinless animals [35], although some selenoprotein-like genes were predicted from its genome [36]. The reason why silkworm does not have selenproteins is explained as that the silkworm lacks the components of Se-Cys biosynthesis and insertion machinery [35]. In this study, arylphorin and SP-2 were found to be the major Se-containing proteins, but not selenoproteins, which supported the conclusion that the silkworm belongs to the selenoproteinless animals.

Met is highly susceptible to oxidation to become Met sulfoxide. In cells, the resulting Met sulfoxides are reduced back to Met by stereo-specific reductases MsrA and MsrB. Reversible Met oxidation occurs even in the absence of stress and is elevated during aging and disease [37]. Similar to Met, Se-Met could also be oxidized to Se-Met selenoxide. However, Se-Met selenoxide can be efficiently reduced to Se-Met nonenzymatically by glutathione and other thiol compounds, indicating that Se-Met is more sensitive and flexible than Met to the oxidation state *in vivo*, thus playing an important role in the regulation of a cellular redox state. Since arylphorin and SP-2 contain high-proportions of Met, Se-rich arylphorin and SP-2 should be better than their general forms in resisting peroxidation in cells. In addition, the Se-rich arylphorin and SP-2 in silkworm can also be developed as a good Se resource for animal supplements.

Arylphorin and SP-2 are two storage proteins in silkworm pupas. Storage proteins are well known as amino acid resources during metamorphosis and the non-feeding pupal stage. They are in a stage-specific manner taken up by the fat body cells from the haemolymph due to receptor-mediated endocytosis [38]. Partially catalyzed by a tyrosine kinase, phosphorylation of the receptor is essential for receptor activation and occurs prior to the uptake of storage proteins [38,39]. Thus, Se-rich arylphorin and Se-rich SP-2 in this

Table 3 Two major Se-containing proteins identified in the Se-rich silkworm pupas

protein	name	NCBI locus	mass	aa	Met	Cys	pI (cal)	pI (exp)
band 1	Sex-specific storage-protein 2 precursor	P20613	83412	704	27	5	6.04	5.1
band 2	arylphorin [*Bombyx mori*]	NP_001037590.1	83398	703	23	3	5.70	4.9

study could also be taken up by fat body cells with the aid of their receptors and be utilized as the resources of Se-Met and Se-Cys *in vivo*. According to this deduction, it is possible to propose a new pathway for the transportation of Se-containing proteins into cells, i.e., *via* the storage proteins and their receptors. It also provides an opportunity to investigate the function of new Se-containing proteins or peptides as potential anti-cancer drugs.

5 Conclusions

Two major Se-containing proteins in Se-rich silkworm pupas were isolated by crude protein extraction, ion exchange chromatography, and gel filtration with the aid of detection and Se-measurement. The two obtained Se-rich proteins were further analyzed by gel eletrophoreses and were characterized by MALDI-TOF MS and ESI-MS/MS to be arylphorin and SP-2, respectively. Se-rich arylphorin and SP-2 are possibly used in the silkworms as resources of Se-Met and Se-Cys and as the antioxidants in defending against oxidative stress.

Acknowledgements This work was financially supported by the National Natural Science Foundation of China (Grant Nos. 30570420, 20637010, and 10775099). We thank Mr. Ziqing Mei in Ziyang County, Shanxi Province, for kindly providing us Se-rich silkworm pupas.

Qiong LIU is a professor in the research area of the chemistry, biochemistry and molecular biology of selenium. She obtained her master's degree from the Huazhong University of Science and Technology in China in 1990 and got her Ph.D. degree from the Roskilde University in Denmark in 1998. She was a post-doctoral fellow in the Karolinska Institute in Sweden from 1999 to 2000. Since then she has been an associate professor in the Huazhong University of Science and Technology from 2000 to 2005 and a full professor in the Shenzhen University since 2006 till now.

References

1. Stadtman, T. C., *Annu. Rev. Biochem*. **1996**, *65*, 83–100
2. Liu, Q.; Jiang, L.; Tian, J.; Ni, J. Z., *Prog. Chem*. **2009**, *2*, 819–830
3. Ge, K.; Yang, G., *Am. J. Clin. Nutr. Supplement* **1993**, *57*, 259–263
4. Gromadzińska, J.; Reszka, E.; Bruzelius, K.; Wasowicz, W.; Akesson, B., *Eur. J. Nutr*. **2008**, *47*, 29–50
5. Combs, Jr.; Gerald, F.; Lu, J. X.. *Selenium as a Cancer Preventive Agent*. in: Hatfield, D. L.; Berry, M. J.; Gladyshev, V. N., *SELENIUM: Its Molecular Biology and Role in Human Health* 2nd Ed.; Springer Science Business Media, LLC: USA, 2006, p. 249–264
6. Sinha, R.; Pinto, J. T.; Facompre, N.; Kilheffer, J.; Baatz, J. E.; El-Bayoumy, K., *Nutr. Cancer* **2008**, *60*, 267–275
7. Bleys, J.; Navas-Acien, A.; Guallar, E., *Arch. Intern. Med*. **2008**, *168*, 404–410
8. Whanger, P. D., *Br. J. Nutr*. **2008**, *91*, 11–28
9. Lu, J.; Jiang, C., *Antioxidants Redox Signaling* **2005**, *7*, 1715–1727
10. Hu, D. C.; Liu, Q.; Cui, H. R.; Wang, H.; Han, D. X.; Xu, H. B., *Life Sci*. **2005**, *77*, 2098–2110
11. Hurst, R.; Elliott, R. M.; Goldson, A. J.; Fairweather-Tait S. J., *Cancer Lett*. **2008**, *269*, 117–126
12. Rayman, M. P., Infante, H. G., Sargent, M., *Br. J. Nutr*. **2008**, *100*, 238–253
13. Institute of Medicine and Panel on Dietary Antioxidants and Related Compounds, *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids: A Report of the Panel on Dietary Antioxidants and Related Compounds, Subcommittees on Upper Reference Levels of Nutrients and Interpretation and Uses of Dietary Reference Intakes, and the Standing Committee on the Scientific Evaluation of Dietary Reference Intakes*, Food and Nutrition Board, Institute of Medicine, National Academy Press: Washington DC, 2000
14. Laclaustra, M.; Navas-Acien, A.; Stranges, S., Ordovas, J. M.; Guallar, E., *Environ Health Perspect* **2009**, *117*, 1409–1413
15. Rayman, M. P., *Br. J. Nutr*. **2008**, *100*, 254–268
16. Xia, Q. Y.; Zhou, Z. Y.; Lu, C.; Cheng, D. J.; Dai, F.; Li, B.; Zhao, P.; Zha, X.; et al., *Science* **2004**, *306*, 1937–1940
17. Xia, Q.; Guo, Y.; Zhang, Z.; Li, D.; Xuan, Z.; Li, Z.; Dai, F.; Li, Y.; et al, *Science* **2009**, *326*, 433–6
18. Yao, H. P.; Xiang, X. W.; Chen, L.; Guo, A. Q.; He, F. Q.; Lan, L. P.; Lu, X. M.; Wu, X. F., *Biosci. Rep*. **2009**, *29*, 363–373
19. Yang, H.; Zhou, Z.; Zhang, H.; Chen, M.; Li, J.; Ma, Y.; Zhong, B., *Amino Acids* **2009**, DOI 10.1007/s00726-009-0342-8
20. Yang, H. X.; Zhu, X. R.; Lu, H. S., *Bull. Sci. Tech*. **2002**, *18*, 318–322
21. Dong, X. Z.; Xu, H. B.; Huang, K. X.; Liu, Q., *Chinese Chem. Lett*. **2003**, *14*, 54–55
22. Dong, X. Z.; Xu, H. B.; Huang, K. X.; Liu, Q.; Zhou, J., *Protein Expres. Purif*. **2002**, *26*, 235–242
23. Ausubel, F. M.; Brent, R.; Kingston, R. E.; Moore, D. D.; Seidman, J. G.; Smith, J. A.; Struhl, K., *Short Protocols in Molecular Biology 3rd Ed.*, John Wiley & Sons Inc.: Chichester, 1995, p 333–351
24. Berkelman, T.; Stenstedt, T., *2-D Electrophoresis Using Immobilized pH Gradients Principles and Methods*; Amersham Pharmacia Biotech Inc: Sweden, 1998, p 1–43
25. Chen, C. Y.; Zhang, P. Q.; Hou, X. L.; Chai, Z. F., *Biochim. Biophys. Acta*. **1999**, *1427*, 205–215
26. Zhang, H.; Chai, Z. F.; Qing, W. Y.; Chen, H. C., *J. Radioanal. Nucl. Chem*. **2009**, *281*, 23–26.
27. Chatt, A.; Desilva, K. N.; Holzbecher, J.; Stuart, D. C.; Tout, R. E.; Ryan, D. E., *Can. J. Chem*. **1981**, *59*, 1660–1664
28. Shi, Y.; Sullivan, E. E.; Holzbecher, J.; Chatt, A., *Biol. Trace Elem. Res*. **1999**, *71–72*, 377–386

29. Ranasinghe, C.; Akhurst, R. J., *J. Invertebr. Pathol*. **2002**, *79*, 51–58
30. Jin, B. F.; He, K.; Wang, H. X.; Wang, J.; Zhou, T.; Lan, Y.; Hu, M. R.; Wei, K. H.; et al, *Oncogene* **2003**, *22*, 4819–4830
31. Wang, H. X.; Jin, B.; Wang, F. J.; He, K.; Yang, S. C.; Sheng, B. F.; Zheng, X. M., *Acta Bioch. Bioph. Sin*. **2002**, *34*, 630–634
32. Wang, J. L.; Zhang, Y. J.; Cai, Y.; Li, L.; Liu, S. Y.; Yang, H. Y.; Qian, X. H., *Acta Bioch. Bioph. Sin*. **2003**, *35*, 459–466
33. Simpson, R. J., *Proteins and Proteomics: A Laboratory Manual, the Use of Mass Spectrometry in Proteomics*, Cold Spring Harbor Laboratory Press: New York, 2003, p 425–596
34. Bock, A.; Rother, M., Selenium Metabolism in Prokaryotes. in: Hatfield, D. L., Berry, M. J., Gladyshev, V. N., *Selenium: Its Molecular Biology and Role in Human Health 2nd Ed.*, Springer Science Business Media, LLC, 2006, p 9–28
35. Lobanov, A. V.; Hatfield, D. L.; Gladyshev, V. N., *Protein Sci*. **2008**, *17*, 176–182.
36. Chen, P.; Duan, J.; Jiang, L.; Liu, Q.; Zhao, P.; Xia, Q. Y.; Xu, H. B., *Chinese Sci. Bull*. **2006**, *51*, 2860–2867
37. Le, D. T.; Liang, X.; Fomenko, D. E.; Raza, A. S.; Chong, C.; Carlson, B. A.; Hatfield, D. L.; Gladyshev, V. N., *Biochemistry* **2008**, *47*, 6685–6694
38. Arif, A.; Scheller, K.; Dutta-Gupta, A., *Insect Biochem. Mol. Biol*. **2003**, *33*, 9219–9228
39. Arif, A.; Manohar, D.; Gullipalli, D.; Dutta-Gupta, A., *Insect Biochem. Mol. Biol*. **2008**, *38*, 307–319