Pak. J. Bot., 51(4): 1387-1394, 2019. DOI: http://dx.doi.org/10.30848/PJB2019-4(29)

STRUCTURAL CHARACTERIZATION OF 12S SEED CRUCIFERIN FROM *ERUCA SATIVA* IN SOLUTION APPLYING SMALL-ANGLE X-RAY SCATTERING

BINISH KHALIQ[1], SVEN FALKE[2], SADAF IQBAL[3] FRIEDRICH BUCK[4], SEEMA MAHMOOD[1], ZAHID ALI[5], MAQSOOD AHMAD[6], ABDUL WAHID[6], CHRISTIAN BETZEL[2] AND AHMED AKREM[1*]

[1'7]Botany Division, Institute of Pure and Applied Biology, Bahauddin Zakariya University, Multan, Pakistan
[2]Institute of Biochemistry and Molecular Biology, Laboratory for Structural Biology of Infection and Inflammation, University of Hamburg, c/o DESY. Build. 22a, Notkestrasse 85, 22607 Hamburg, Germany
[3]Department of Chemistry, University of Karachi, Karachi, Pakistan
[4]Institute of Clinical Chemistry, University Hospital Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany
[5]Department of Biosciences, COMSATS Institute of Information Technology (CIIT), Islamabad 44000, Pakistan
[6]Department of Environmental Sciences, Bahauddin Zakariya University, Multan, Pakistan
[7]Botany Division, Institute of Molecular Biology and Biotechnology, University of Lahore, Lahore, Pakistan
[*]Corresponding author's email: ahmedakrem@bzu.edu.pk

Abstract

Cruciferin (12S globulin) from seeds of *Eruca sativa* was isolated, purified and thoroughly characterized. The protein indicated 54 % sequence identity with the cruciferins of *Brassica napus* (CRU1) and *Raphanus sativus* (PGCRURSE5) when the obtained amino acid sequence from LC-MS/MS mass spectrometric data was submitted to the UniProtKB. SDS-PAGE exhibited an approx. 50 kDa monomeric cruciferin, which was separated into α-polypeptide with a major band at approx. 30 kDa and a β-polypeptide of approx. 20 kDa under reduced conditions. The secondary structure content of *E. sativa* Cruciferin (*Es*C) was analyzed by Circular Dichroism spectroscopy indicating 7% α-helix, 48% β-sheet, 7% β-turn and 38% disordered conformation. The monodispersity and stability of *Es*C was verified via Dynamic Light Scattering (DLS) and a hydrodynamic radius of *Es*C was calculated to be 5.5 ± 0.3 nm indicating a trimer of the protein in solution. A gyration radius (R_g) of 4.3± 0.30 nm and the globular molecular shape was disclosed by Small-angle X-ray scattering (SAXS) for *Es*C. An incredibly analogous globular shape was obtained when the *ab-initio* dummy model of *Es*C inferring P3 symmetry was carefully compared with the PDB-ID 3KGL; 11S globulin of *Brassica napus*. Moreover, the scattering patterns of both proteins showed a minimized χ^2-value of 2.0 which further confirms the structural similarities. Protruding loops of the *Es*C model were considered as hyper variable region-I (HVR-I) of *Arabidopsis thaliana* Cruciferin C (AtCruC) and variable region II of 3KGL molecular structure and were nominated as the major flexible regions.

Key words: *Eruca sativa*, Cruciferin, Dynamic light scattering, Circular dichroism spectroscopy, Small-angle X-ray scattering (SAXS).

Introduction

Brassicaceae or *Cruciferae* is one of the most widely spreaded and economically important plant family in the plant kingdom (Love *et al.*, 2005). Species from *Brassicaceae* are grown globally for nourishment purposes of animals and humans. *Brassicaceae* seeds are consists of 20 to 35% protein matter and are also rich in oil content. Seed storage proteins (SSPs) are the chief nitrogenous composites in *Brassicaceae* oilseeds. During seed germination, these SSPs are responsible for the release of the N and S residues that are required during seed germination (Müntz, 1998). One of the major types of SSPs is 12S Cruciferin belonging to the cupin super family. In mature *Brassicaceae* seeds, approx. 60% of the total seed storage proteins are present as cruciferin content. In the European rapeseed species, the reported content of cruciferins was 30-50% of the total protein content (Malabat *et al.*, 2003). Cruciferins are known to have molecular masses in the range from 300-350 kDa (Sjodahl *et al.*, 1991).

A variety of chemical and biochemical constituents, glucosinolates, flavonoids, proteins, vitamins A and C, erucic acid and different lipids are present in *Eruca* seeds (Kim & Ishii, 2006; Bell & Wagstaff, 2014). In Pakistan and India, *Eruca* seeds are most commonly utilized as nourishment for animals. *Eruca sativa* is widely used as medicine for the treatment of many diseases; for example, its regular consumption has been linked with the reduction of cardiovascular diseases and lessening of cancer (Higdon *et al.*, 2007; Goloshvili *et al*., 2018; Fuentes *et al.*, 2014). *Eruca sativa* is recognized as diuretic and anti-inflammatory supplement for excellent consequences on the gastrointestinal tract (Yehuda *et al.*, 2012; Björkman *et al.*, 2011 and Traka & Mithen, 2011).

To resolve the low-resolution structure of a biological macromolecule in solution, small-angle X-ray scattering (SAXS) is a powerful method. The technique offers 3D low resolution shape information by using *ab-initio* modeling. Consequently, biological SAXS turn out to be a famous procedure to illustrate overall structure, molecular weight and shape of large macromolecular complexes, proteins and nucleic acids in solution (Hura *et al.*, 2013; Svergun *et al.*, 2013). Further, the oligomeric state of proteins and protein complexes can be investigated (Mertens & Svergun, 2010). In addition, SAXS is highly supportive and complementary to high resolution methods such as X-ray crystallography and Nuclear Magnetic Resonance (NMR). At the moment, SAXS attained a sophisticated state, permitting programmed and quick characterization of protein solutions in terms of conformational analysis, low-resolution modeling,

quaternary structure, and even time-resolved experiments (Graewert & Svergun, 2013; Fang *et al.*, 2015). In this study complementary analytical techniques have been used to elucidate the structural characterization of *E. sativa* Cruciferin (*Es*C).

Materials and Methods

Protein purification: The seeds of *Eruca sativa* used in this study were obtained from local market. 50 g seeds were ground to a fine powder and suspended in 100 ml of 10 mM Tris buffer (pH 7.5). The suspension was stirred continuously overnight. The plant debris was pellet down by centrifuging at 12,000 rpm for 25 min and the supernatant was filtered with 8 µm filter paper. To precipitate high molecular weight proteins, 35 % (w/v) ammonium sulfate precipitation was used for the particulate free crude extract. Pellet is formed by allowing centrifugation at 3,000 rpm for the duration of 15 minutes. The obtained pellet was re-dissolved in 30 ml of the 10 mM Tris buffer (pH 7.5). The solution was dialyzed (Millipore, MWCO 3500 Da) against same buffer at 4°C with gentle stirring. Cruciferin was further purified by injecting it onto a pre-equilibrated Superdex G-200 16/60 prep (Product code: 17-5175-01) column. The protein was eluted at a flow rate of 1 ml min^{-1} in 10 mM Tris buffer (pH 7.5) including 150 mM NaCl. The fractions with maximum protein content were combined and analyzed via SDS-PAGE. Final concentration of protein was calculated as 3 mg/ml by Nanodrop 2000c (Thermo Scientific, peqLab, Germany).

LC–MS/MS spectrometry: First, the gel bands were incised and DTT (10 mM, 56°C, 30 min.) was used to reduced the in-gel proteins. The modification in the cysteine residues was performed with 55 mM iodoacetamide at ambient temperature for 20 minutes. The sequencing grade modified trypsin (5 ng trypsin/µL, Promega, Madison, WI, USA) was employed for in-gel protein digestion in 50 mM NH_4HCO_3 at 37°C for 16 hours. The gel slices were extracted many times in 50% acetonitrile and 5% formic acid. Vacuum concentrator was utilized in order to dry the combined extracts and re-dissolved in 20 µl of 0.1% formic acid.

A nano LC system (Dionex Ultimate 3000) connected through the electrospray-ionization (ESI) mass spectrometer (Orbitrap Fusion, Thermo Scientific, Bremen, Germany) was used for recording LC–MS/MS measurements. A trapping column (Acclaim PepMap µ-precolumn, C18) was used on which 5 µl of each sample was loaded per minute using buffer A (0.1% formic acid in H_2O) and 2% buffer B (0.1% formic acid in acetonitrile). Washing was done with 2% buffer B (5 µl/min) and the peptides were obtained (200 nl/min) from the separation column (Acclaim PepMap 100, C18, gradient: 2–30% B). The ion trap of the instrument was utilized for the recording of MS/MS spectra. The processing of the LC–MS/MS raw data was carried out with Proteome Discoverer 2.0 (Thermo Scientific, Bremen, Germany), searched against the *Arabidopsis* protein data base UniProtKB (http://www.uniprot.org/blast/) and the data complemented by manual *de novo* sequencing. The LC–MS/MS exploration was achieved in data dependent acquisition mode (DDA).

Circular dichroism (CD) spectroscopy: Jasco J-815 spectropolarimeter (Jasco, Tokyo, Japan) was employed for CD measurements. The diluted purified *Es*C (1 mg/ml) was prepared in 10 mM Tris-HCl (pH 7.5). The data was measured in the far-UV range from 195 to 260 nm applying a cuvette of 1 mm path length at 20°C. CD spectra were evaluated using Spectra manager™ software (Reed & Kinzel, 1984). Fifteen scans per sample were integrated to obtain one spectrum and the baseline was corrected by subtracting the averaged spectrum of the buffer.

Dynamic light scattering (DLS) analysis: For SAXS measurements, the concentration of the pure *Es*C was adjusted to three different concentrations (2, 5 and 9 mg/ml) using Amicon Ultra-15 (MWCO 10,000 Da, Millipore) concentrating vials. The protein was further analyzed by DLS measurements applying the 300 spectroLight (Xtal Concepts, Germany) to verify the dispersity and size of the protein solutions. Pure transparent cruciferin solution was used to record a series of measurements with a sampling time of 45 s applying a laser with a wavelength of 660 nm. The scattered light was recorded at a fixed angle of 90°. Analysis of the autocorrelation function was done by the CONTIN algorithm (Provencher, 1982). DLS measurements were performed at least twenty times. The hydrodynamic radius of the pure cruciferin was calculated *via* the Stokes-Einstein equation.

Small-Angle X-ray Scattering (SAXS) measurements: SAXS is utilizing the pattern of elastic scattering of X-rays at small angles of typically 0.1 to 10°. For data processing and 3D structural analysis through SAXS, monodisperse and homogenous solutions are required (Jacques *et al.*, 2012). Filtered and monodisperse solutions of *Es*C in 10 mM Tris-HCl (pH 7.5) were applied at the EMBL beamline P12 at storage ring PETRA III (DESY, Hamburg, Germany) (Blanchet *et al.*, 2015) for data collection. A wavelength of λ = 0.124 nm and a sample-detector distance of 3.1 m, scattering data were recorded. For this purpose, a 2D detector (Dectris) was employed with the momentum transfer ranging from 0.03 $nm^{-1} < s < 4.80$ nm^{-1} (where $s = 4\pi \sin\theta/\lambda$, 2θ scattering angle). In order to scrutinize radiation damage 20 consecutive X-ray exposures of 45ms durations were compared. Data were normalized to the intensity of the beam and radially averaged. Further, for the computation of the maximum dimension D_{max}, the radius of gyration R_g along with the particle pair-distance distribution function $p(r)$, the automated data processing pipeline SASFLOW (Petoukhov *et al.*, 2012) was employed. Moreover, verification was performed by using PRIMUS (Konarev *et al.*, 2003). For the computation of *ab-initio* shapes of the protein molecule in low-resolution from the composite

scattering curves, DAMMIF (Franke & Svergun, 2009) was applied. The application exploits a cluster of heavily filled beads signifying the shape of the molecule. Simulated annealing procedures were applied for the calculation of model that corresponds to the experimental scattering intensity data. To verify the obtained structures, ten runs of DAMMIF were executed; resulting in well super imposable models. The obtained modals were evaluated based on their normalized spatial discrepancy (NSD) values. To superpose and compare scattering data with already known homologous high-resolution structures, CRYSOL (Svergun *et al.*, 1995) was employed.

Results

Protein purification: The *Eruca sativa* cruciferin (*Es*C) was purified by means of 35% (w/v) ammonium sulfate precipitation from its crude extract in 100 mM, pH 7.0 sodium phosphate buffer. The precipitates obtained were significantly dialyzed in 10 mM Tris buffer, pH 7.5 to remove traces of salt. The dialyzed fractions were subjected to size-exclusion chromatography to obtain the purified protein as shown in Fig. 1A. Eventually, by performing SDS-PAGE analysis, more than 95% pure was acquired. The reduced and non-reduced SDS-PAGE analysis of the purified *Eruca sativa* cruciferin (*Es*C) strongly indicates the presence of inter-chain disulfide linkages as shown in (Fig. 1B). Further characterizations of *EsC* were done in 10 mM Tris (pH 7.5) buffer.

Identification of *Es*C: Table 1 shows the sequence of *EsC*, obtained from LC-MS/MS mass spectrometry revealing seven tryptic peptides, was searched using BLAST to find homologue proteins in the UniProtKB database. BLAST search discovered the 54 % sequence identity of *EsC* with the cruciferins of *Raphanus sativus* (PGCRURSE5) and *Brassica napus* (CRU1). A multiple sequence alignment was executed between CRU1, PGCRURSE5 and *Es*C to highlight the conserve areas and other structural details as shown in Fig. 2.

Table 1. LC-MS/MS produced amino acid sequences of *E. sativa* Cruciferin (*Es*C).

S.No.	Peptide sequences
1.	VIIEQGGLYLPTFFSSPKISYVVQGMGISGR[31]
2.	LAGNNPQGGSQQQQQQQQNMLSGFDPQVLAQALK[34]
3.	ADVYKPNLGRVTSVNSYTLPILQYIR[26]
4.	GILQGNAMVLPK[12]
5.	QVVNDNGQNVLDQQVQK[18]
6.	TNANAMVSTLAGR[13]
7.	ALPLEVITNAFQISLEEAR[20]

Secondary structure determination of EsC: The existence of α-helical and β-sheeted conformation in the native fold of *Es*C was recorded by CD spectroscopy (Manavalan & Johnson, 1983). The CD spectrum consequent in a computed content of about 7% α-helix, 48% β-sheet, 7% β-turn and 38% disordered conformation as revealed by Fig. 3. The RMS value of 6.9% was calculated for the respective reference fit curve.

Analysis of *Es*C dispersity: For SAXS purposes, the homogeneity and dispersity of purified *Es*C was investigated by applying DLS. The three different concentrations of the purified *Es*C (2, 5 and 9 mg/ml) were centrifuged at 16,000 g for 35 min facilitated by a 0.22 μm syringe filter with the intention of eliminating impurities and bigger aggregates. It was observed from the monodispersity signals of DLS that up to 9 mg/ml, *Es*C is showing a highly monodisperse nature and the globular shape in solution (Fig. 4). A hydrodynamic radius (R_H) of 5.5 ± 0.3 nm acquired from DLS computations is a reflection of a trimeric state of native *Es*C. However, above 9 mg/ml, the *Es*C started to show polydisperse signal.

SAXS investigations of *Es*C: The obtained solution scattering statistics showed the trimeric status of *Es*C molecules in 10 mM, pH 7.5, Tris buffer. As per Guinier approximation, obtained by AUTORG run in PRIMUS, the averaged scattering intensity pattern showed a gyration radius (R_g) of 4.30 ± 0.30 nm (Petoukhov *et al*., 2012) for trimeric molecules of *Es*C. The mean value of collected data indicated D_{max} (maximum dimension) of 15.0 ± 0.4 nm (Fig. 5A and 5B). The scattering amplitude of *Es*C was compared to 11S globulin from *B. napus* (PDB ID: 3KGL) by the program CRYSOL as shown in (Fig. 5A). The scattering amplitude of *Es*C modal was processed by using PRIMUSQT. *B. napus* 11S globulin (PDB ID: 3KGL) was fetched from the PDB and compared with the processed scattering amplitude of *Es*C employing CRYSOL as shown in Fig. 5A. Kratky plot indicated the *Es*C as a globular, rigid and compactly folded protein (Fig. 5B).The scattering pattern of *Es*C (green) and *B. napus* 11S globulin (blue) over a wide range of angles provided minimum χ^2-value of 2.0 (Fig. 5A). The minimized χ^2-value verified high structural similarity between the two models and indicated the existence of a loop revealing flexibility in solution. The forward scattering intensity suggested a similar molecular weight of *Es*C. Likewise, the labor-intensive superimposition further confirms that the two structures are broadly alike, as derived from the volume of the *ab initio* model. An *ab-initio* model of *Es*C displays an extended "loop structure" projecting out of the nearly globular surface in agreement with the p(r) function (Fig. 5C). DAMMIF was further employed for computing ten *ab-initio* models with P3 symmetry because of the symmetry of a putative trimer. These models share a mean NSD value of 0.957 ± 0.271 indicating high similarity of the obtained models; one representative *ab-initio* dummy model is shown in (Fig. 6). A value of 0.76 for the shape factor was computed by dividing R_g with R_H which pointed out a nearly globular particle compatible with the demonstrated *ab-initio* model. Various view of superimpositions between *ab-initio Es*C model and *B. napus* 11S globulins (PDB ID: 3KGL) were demonstrated in Figs. 6, 7 & 8.

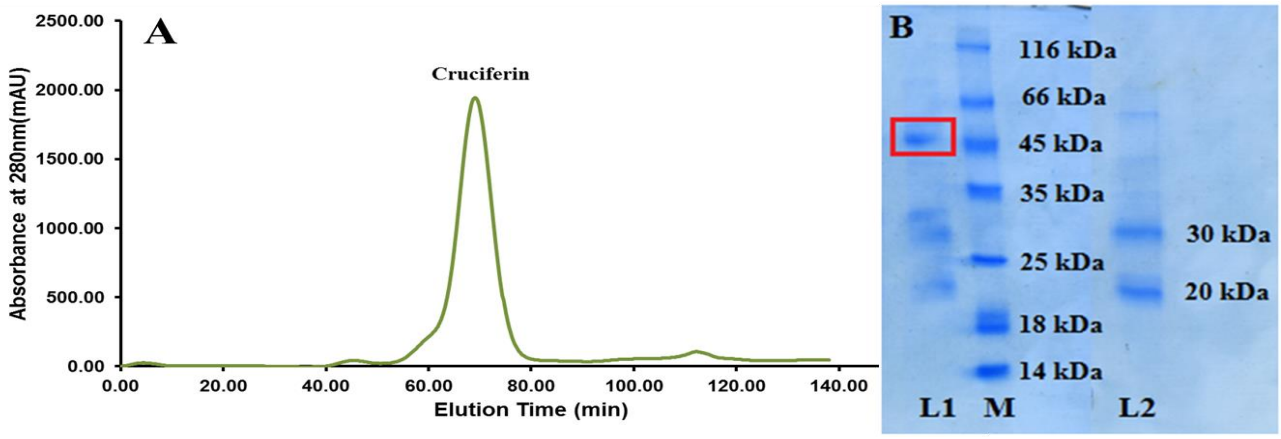

Fig. 1. (A) Size-exclusion chromatogram revealing single band of pure *Es*C and (B) SDS-PAGE analysis of *Es*C displaying pure *Es*C in lane L1 without β-mercaptoethanol (red box), L2 showing splitting into 30 and 20 kDa in presence of β-mercaptoethanol.

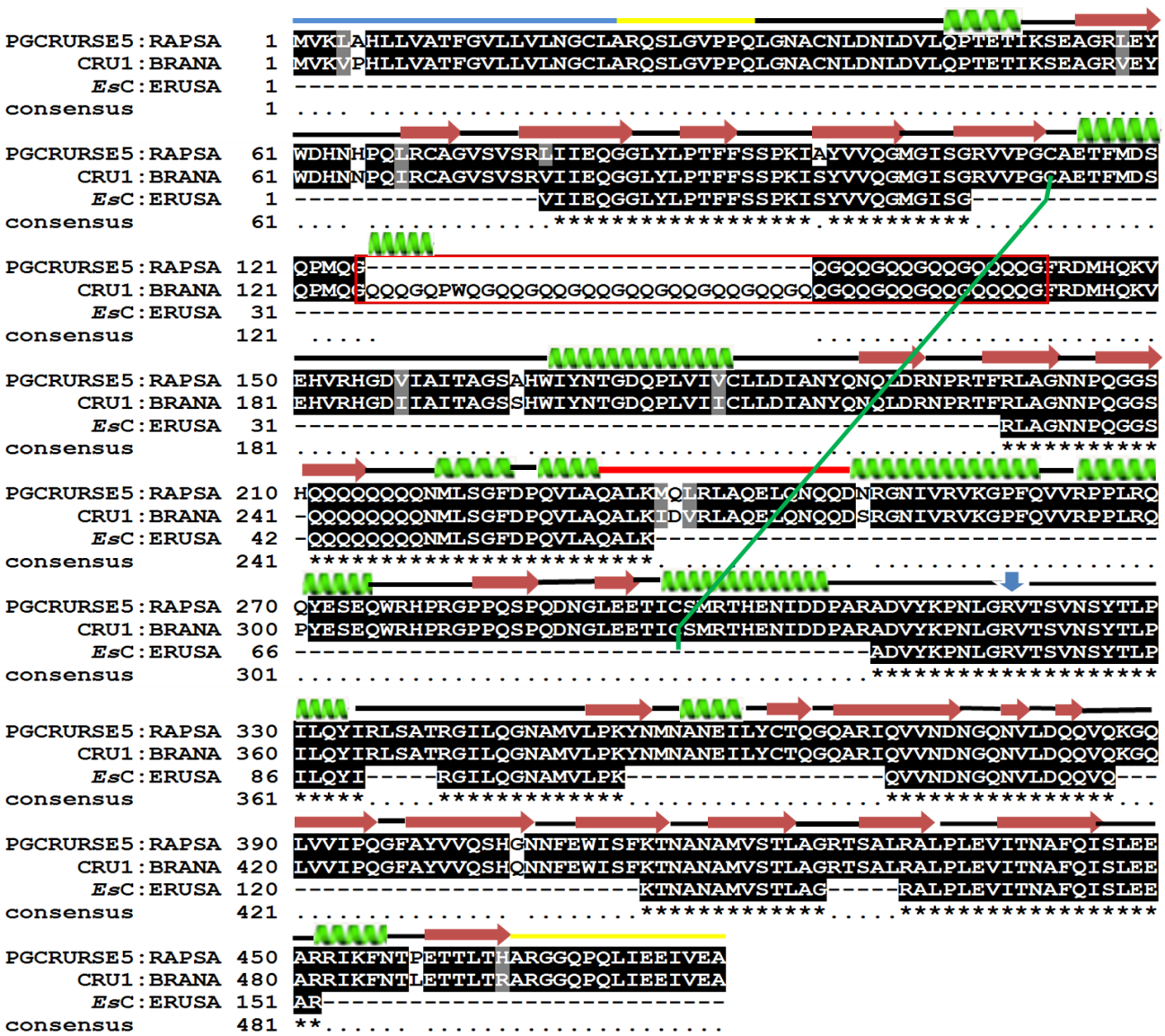

Fig. 2. Sequence alignment of *Es*C to other plant cruciferins; PGCRURSE5 from *Raphanus sativus* (UniProtKB - Q02498 CRU1_RAPSA) and CRU1 from *Brassica napus* (UniProtKB - P33525; CRU3_BRANA). Secondary structure of cruciferins (α-helix and β-sheet) is indicated at the top by using Phyre[2]. Signal peptide, N/C terminal disordered and extended loop regions are indicated with blue, yellow and red bars respectively. Disordered region is shown as red box; while the internal proteolytic possessing site for the separation of acidic and basic chains is indicated by a blue arrow head and disulfide bridge is marked with green rods. Identically conserved residues are represented by Asterisks (*), while semi-conserved substitutions are represented by single dot (.).

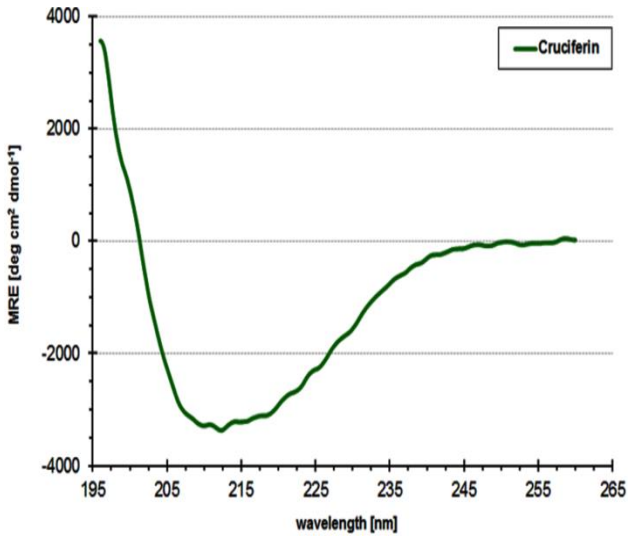

Fig. 3. The predominant β-sheeted fraction of secondary structure, followed by unstructured parts and loop regions, recorded by averaged far-UV CD spectrum of *Es*C.

Discussion

A 50 kDa *E. sativa* cruciferin (*Es*C) and its corresponding isoforms were purified, analyzed and characterized by chromatography and small-angle X-ray scattering techniques. SDS-PAGE under reduced condition revealed the presence of two molecular weights i.e., 30 kDa (α-polypeptide) and 20 kDa (β-polypeptide) as shown in (Fig. 1B, lane 2). Above report is compatible with the reported molecular weights of cruciferin proteins (Khaliq *et al.*, 2017). Likewise, two major bands corresponding to basic (MW: 22,000 to 24,000 Da) and acidic (MW: 30,000 to 37,000 Da) polypeptides linked by a disulfide bond under reduced conditions have already been reported (Dalgalarrondo *et al.*, 1986; Robin *et al.*, 1991). In the same way, the pro-11S globulin of *Brassica napus* protein oligomerizes to a molecular mass of 300-400 kDa and composed of six monomeric subunits with molecular weights in the range of 50-60 kDa. The acidic and basic chains of the 50 kDa monomer are linked to each other by a disulfide linkage and are considered to be synthesized by cleavage of single polypeptide precursor during its processing in the storage vacuoles (Mori *et al.*, 1979; Staswick *et al.*, 1984).

Our CD measurements and SAXS results revealed the ascendancy of anti-parallel β-sheeted secondary structural elements over α-helical content of *Es*C (Fig. 3) and is also in accordance with that of *B. napus* 11S seed globulin structure with PDB-ID: 3KGL (Tandang-Silvas *et al.*, 2010) and CD spectra of *Bn*C (Khaliq *et al.*, 2017). Withana-Gamage *et al.*, (2011) determined the secondary structure of *Arabidopsis* (wild type) cruciferin and it was observed that it consisted of 9.4% α-helix and 44.1% β-sheet content. Schwenke et al. (1983) reported that rapeseed 12S globulin also had a low α-helix content (11%) and relatively high content of β-conformation (31%).

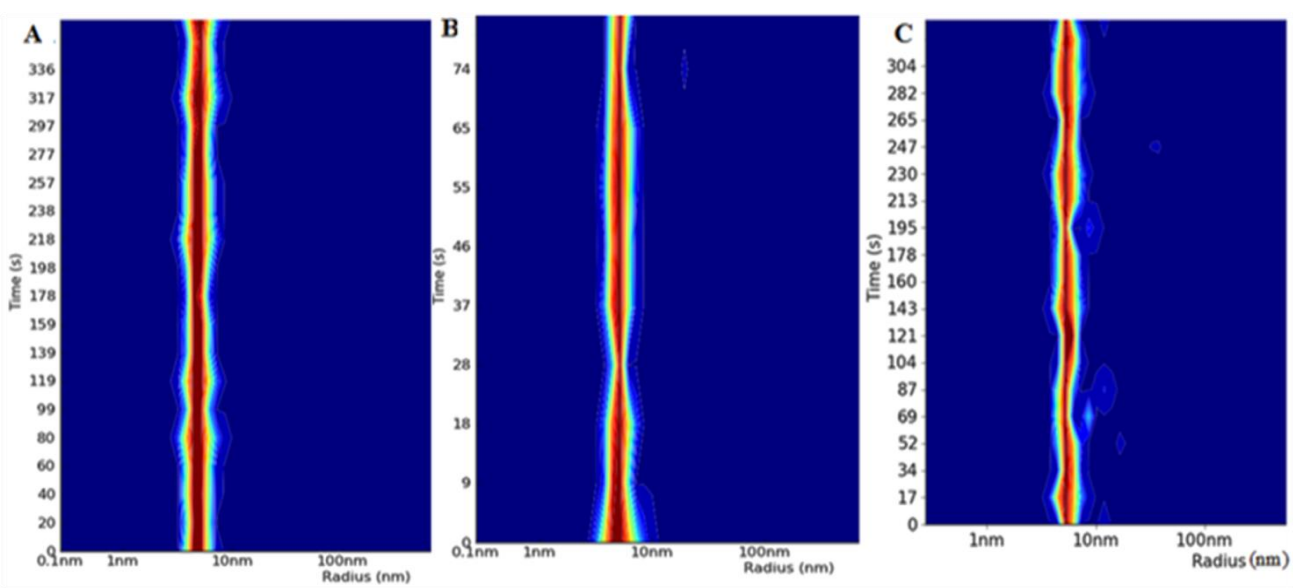

Fig. 4. Dynamic light scattering (DLS) measurements in three different concentrations (A) 2, (B) 5 and (C) 9 mg/ml of *Es*C. However, a closer look is showing a reflection of *Es*C aggregation that started from 9 mg/ml (C) concentration of *Es*C.

The monodispersity of *Es*C was optimized by DLS measurements in Tris buffer (10 mM, pH 7.5). Apart from qualitative optimization of *Es*C, the DLS data was also used for the quantitative optimization of protein for further SAXS analysis. Three different concentrations of the purified *Es*C (2, 5 and 9 mg/ml) were exposed to DLS and it was seen that up till 9 mg/ml protein concentration, a highly monodisperse nature (Fig. 4A and 4B) of *Es*C was maintained, nevertheless, minor aggregates started to appear from 9 mg/ml (Fig. 4C) concentration. Moreover, DLS signal also indicated the trimeric state of *Es*C with an approximate hydrodynamic radius of 5.5 nm indicating the trimeric status of the molecules which was further confirmed by SAXS and subsequent *ab initio* modeling (Figs. 5 and 6). The trimeric state of *Es*C in solution was analyzed by SAXS spectroscopy. The values of R_g and D_{max} obtained from the collected data suggested a

folded but globular structure. Results of Kratky plot curve showed that *Es*C is rigid and compactly folded (Fig. 5B). The trimeric state conservation and a slightly oblate globular shape, deduced from the p(r) function (Fig. 5C) has also been reported for *B. napus* (Perera *et al.*, 2016). However, it is important to mention that majority of 11S globulins are hexameric in nature and are assembled as two trimers and each trimer is further composed of subunits which are isoforms of each other thus adding heterogeneity to the hexameric structure as have already been reported for soybean globulin A3B4 (Adachi *et al.*, 2003), peanut allergen Ara h3 (Jin *et al.*, 2009). Similarly, overall molecular shape of recombinant 3KGL is hexameric with two assembled trimers interacting in face-to-face sandwich conformation (Fig. 7A and B). Withana-Gamage *et al.*, (2011) had extensively studied the *in silico* molecular basis of the hexameric conformation of *Arbidopsis thaliana* Cruciferin C (AtCruC). They have indicated the presence of two Hyper variable regions (HVR-I and II) which are responsible for the hexamer development in AtCruC. Due to steric hindrances of the two regions, HVR-I move itself towards the periphery of the molecule and thus allowing the interdisulfide-containing face (IE) of the two trimers to come in close proximity and the ultimate development of hydrogen bonding and hydrogen-bonded salt bridges between four highly conserved binding sites of the two trimers of a hexamer. Moreover, it has already been mentioned above that each subunit of a trimer is further composed of two polypeptides as α and β-chains which are joined together through disulfide linkage. Nevertheless, the trimeric status of *Es*C predicted by DLS and SAXS could be justified due to the dissociation of native hexameric structure into trimeric conformation by splitting of the two pieces (trimers) of hexameric sandwich (Fig. 7B). According to Adachi et al. (2003), any change in the ambient pH or ionic strength may lead towards the destabilization of the hexameric shape into two trimers.

Moreover, the PDB-ID: 3KGL trimer consists of three monomer chains, non-covalently joined with each other. Each monomer expressed a much conserved two jelly-roll β-barrel and two extended α-helix domains (Tandang-Silvas *et al.*, 2010). The number and length of these α-helices are more diverse than the β-stands. Likewise, the structure of amaranth 11S globulin (PDB ID: 3QAC) and recombinant pro-11S globulin from *Cucurbita maxima* (PDB ID: 2E9Q) proteins composite of shorter helices while in others these are noticeably longer. The modeled *Es*C (Fig. 6, grey spheres) demonstrates an extended loop (red ovals) consistent with the p(r) function, corroborated an elongated terminus structure that acquired a little flexibility in solution. These hyper variable regions of *Es*C including loops and HVRs-I and II (Adachi *et al.*, 2003) were not resolved because they are lacking in well defined sheets or helices and are protrude from the compact globulin structure (Tandang-Silvas *et al.*, 2010). In 11S globulin (procruciferin) of *B. napus* (PDB ID: 3KGL), this variable region II (residue 86-138 rich in GQ repeats) is lying on the IE face of the molecule and is directed towards the periphery where it is imparting longer disordered region and is flexible around the vicinity of its IE face (Fig. 8). Based on close analysis of *Es*C SAXS data, we believe that extended loops (Fig. 6; red ovals) of *Es*C *ab initio* model are comparable to HVR-I part of AtCruC as well as variable region II (residue 86-138 rich in GQ repeats) of 3KGL and is highlighted as inset of Fig. 8. Such flexible variable and disordered extended regions are believed to be the real bottleneck towards the development of high diffracting crystals of the native protein parallel to isoform micro heterogeneity.

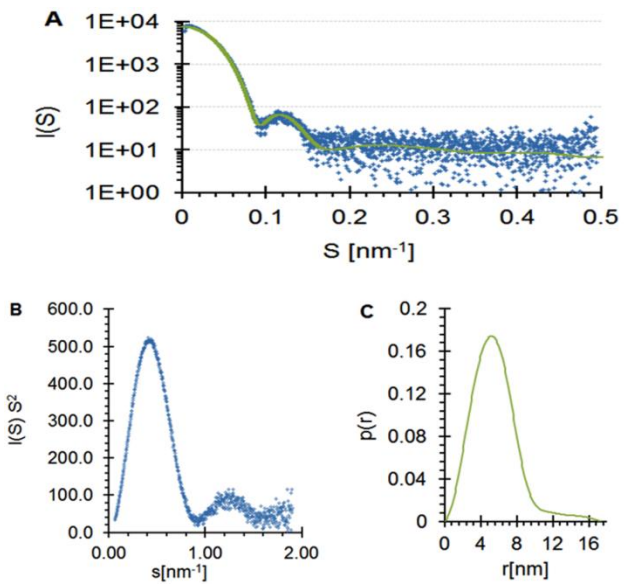

Fig. 5. SAXS data exhibiting important structural information of *Es*C protein. (A) Processed solution scattering pattern of *Es*C (green) in arbitrary intensity units and the fit with *B. napus* 11S globulin (ProCruciferin; PDB ID: 3KGL) (blue) as evaluated by CRYSOL (χ^2 = 2.0). (B) Kratky plot of the scattering data verifying an overall compact and rigid folding. (C) Corresponding pair distance distribution functions p(r).

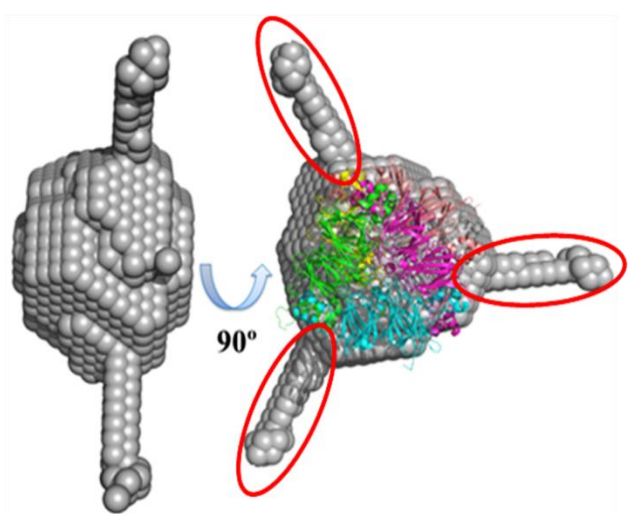

Fig. 6. Superimposition of *Es*C consistent with symmetry of P3 (grey spheres) and *B. napus* pre-cruciferin, the extended loop structures of the *Es*C model are marked by red ovals.

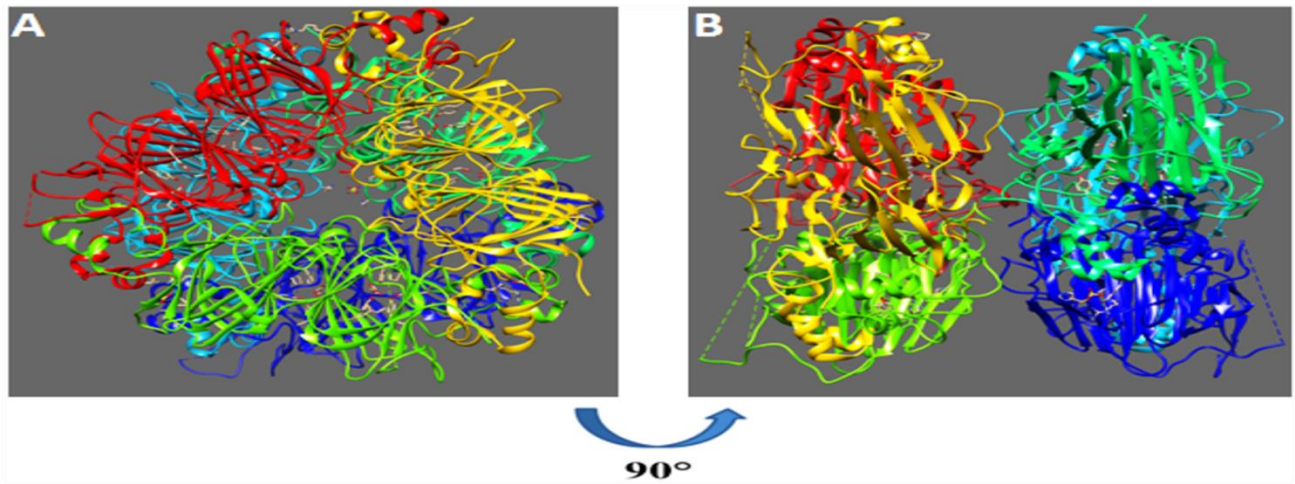

Fig. 7. Overall structure of cloned *Brassica napus* procruciferin (PDB ID: 3KGL). (A) The homohexameric gobular shape molecular structure, (B) the same structure after rotation of 90° clearly indicating the face-to-face joining of two trimers in sandwich conformation and ultimate development of hexameric molecule.

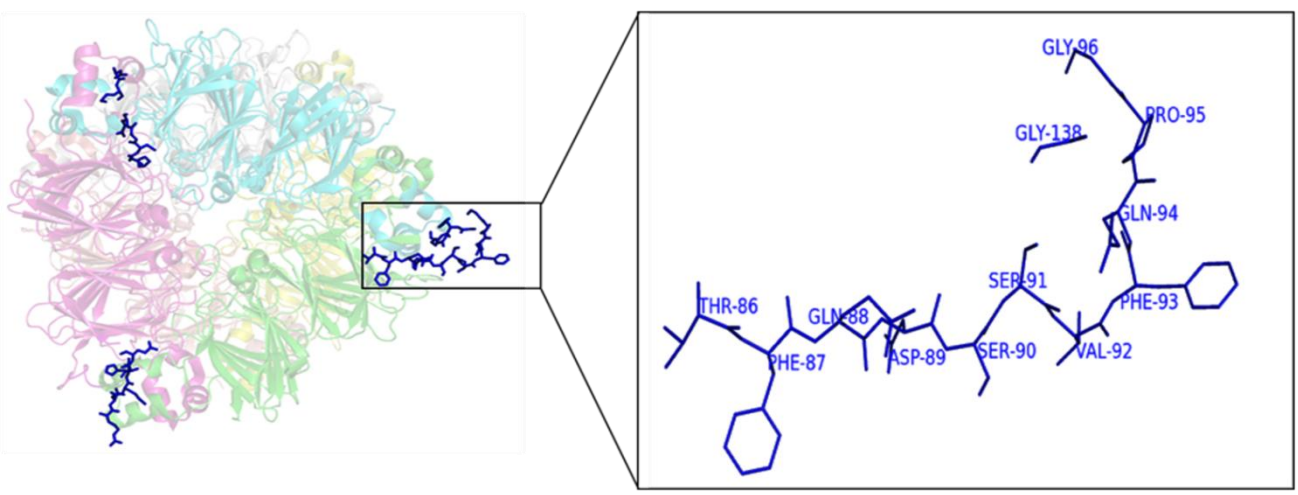

Fig. 8. *B. napus* 11S globulin revealing the disordered variable regions II (residues 86-138) of each monomer; lying on IE face and are protruding out from the compact globular shape.

Conclusions

Our investigation explains the discovery and in-solution characterization of a trimeric/hexameric seed storage *E. sativa* cruciferin (*Es*C). The SDS-PAGE investigation of *E. sativa* cruciferin exhibited protein bands in the range of 50,000 to 52,000 Da molecular weight for monomer. Further cleavage in protein bands occurs into 30,000 Da molecular weight α-polypeptides and 20,000 Da molecular weight β-polypeptides under reduced circumstances. LC-MS/MS mass spectrometric analysis was conducted to recognize peptide fragments. The obtained peptides illustrated an identity of 63% in sequence of *Es*C with that of *R. sativus* and *B. napus*. A hydrodynamic radius of 5.5 ± 0.3 nm further corroborates its trimeric form. The secondary structure content of 7% α-helix, 48% β-sheet, 7% β-turn and 38% disordered regions was computed by circular dichroism (CD) spectroscopy. SAXS data analysis showed that *Es*C has an overall globular shape. Moreover, it is a trimer according to the molecular weight estimation. The oligomeric status of the protein by SAXS results exposed a value of 4.30 ± 0.30 nm as the gyration radius.

Acknowledgements

This research has been financially supported by Higher Education Commission (HEC), Islamabad, Pakistan as a part of awarding Indigenous scholarship under Pin number: 112-22140-2BM1-565.and International Research Support Initiative Program (IRSIP).

References

Adachi, M., J. Kanamori, T. Masuda, K. Yagasaki, K. Kitamura, B. Mikami and S. Utsumi. 2003. Crystal structure of soybean 11S globulin: glycinin A3B4 homohexamer. *Proceed. Nat. Acad. of Sci.*, 100: 7395-7400.

Bell, L. and C. Wagstaff. 2014. Glucosinolates, myrosinase hydrolysis products, and flavonols found in rocket (*Eruca sativa* and *Diplotaxis tenuifolia*). *J. Agri. & Food Chem.*, 62: 4481-4492.

Björkman, M., I. Klingen, A.N. Birch, A.M. Bones, T.J. Bruce, T.J. Johansen, R. Meadow, J. Mølmann, R. Seljåsen and L.E. Smart. 2011. Phytochemicals of Brassicaceae in plant protection and human health–Influences of climate, environment and agronomic practice. *Phytochem.*, 72: 538-556.

Blanchet, C.E., A. Spilotros, F. Schwemmer, M.A. Graewert, A. Kikhney, C.M. Jeffries, D. Franke, D. Mark, R. Zengerle and F. Cipriani. 2015. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J. Appl. Crystallograph.*, 48: 431-443.

Dalgalarrondo, M., J.M. Robin and J.L. Azanza. 1986. Subunit composition of the globulin fraction of rapeseed (*Brassica napus* L.). *Plant Sci.*, 43: 115-124.

Fang, X., J.R. Stagno, Y.R. Bhandari, X. Zuo and Y.X. Wang. 2015. Small-angle X-ray scattering: a bridge between RNA secondary structures and three-dimensional topological structures. *Curr. Opin. in Struct. Biol.*, 30: 147-160.

Franke, D. and D.I. Svergun. 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallograph.*, 42: 342-346.

Fuentes, E., M. Alarcón, M. Fuentes, G. Carrasco and I. Palomo. 2014. A Novel Role of *Eruca sativa* Mill.(rocket) extract: Antiplatelet (NF-κB Inhibition) and antithrombotic activities. *Nutrients.*, 6: 5839-5852.

Graewert, M.A. and D.I. Svergun. 2013. Impact and progress in small and wide angle X-ray scattering (SAXS and WAXS). *Curr. Opin. in Struct. Biol.*, 23: 748-754.

Goloshvili, T., M. Akhalkatsi and G. Badridze. 2018. Characterization of grape seed extracts of native to Georgia varieties of *Vitis vinifera* L. *Pak. J. Bot*., *50*: 245-250.

Higdon, J.V., B. Delage, D.E. Williams and R.H. Dashwood. 2007. Cruciferous vegetables and human cancer risk: epidemiologic evidence and mechanistic basis. *Pharmacol. Res.*, 55: 224-236.

Hura, G.L., H. Budworth, K.N. Dyer, R.P. Rambo, M. Hammel, C.T. McMurray and J.A. Tainer. 2013. Comprehensive macromolecular conformations mapped by quantitative SAXS analyses. *Nature Meth.*, 10: 453-454.

Jacques, D.A., J.M. Guss and J. Trewhella. 2012. Reliable structural interpretation of small-angle scattering data from bio-molecules in solution-the importance of quality control and a standard reporting framework. *BMC Struct. Biol.*, 12: 9.

Jin, T., F. Guo, Y.-w. Chen, A. Howard and Y.-Z. Zhang. 2009. Crystal structure of Ara h 3, a major allergen in peanut. *Mol. Immunol.*, 46: 1796-1804.

Khaliq, B., S. Falke, A. Negm, F. Buck, A. Munawar, M. Saqib, S. Mahmood, M.S. Ahmad, C. Betzel and A. Akrem. 2017. SAXS and other spectroscopic analysis of 12S cruciferin isolated from the seeds of *Brassica nigra*. *J. Mol. Struct.*, 1137: 60-66.

Kim, S.J. and G. Ishii. 2006. Glucosinolate profiles in the seeds, leaves and roots of rocket salad (*Eruca sativa* Mill.) and anti-oxidative activities of intact plant powder and purified 4-methoxyglucobrassicin. *Soil Sci. & Plant Nutr.*, 52: 394-400.

Konarev, P.V., V.V. Volkov, A.V. Sokolova, M.H. Koch and D.I. Svergun. 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallograph.*, 36: 1277-1282.

Love, C.G., A.J. Robinson, G.A. Lim, C.J. Hopkins, J. Batley, G. Barker, G.C. Spangenberg and D. Edwards. 2005. Brassica ASTRA: an integrated database for Brassica genomic research. *Nucl. Acids Res.*, 33: D656-D659.

Malabat, C., H. Atterby, Q.Chaudhry, M. Renard and J.Guéguen. 2003. Genetic variability of rapeseed protein composition. *Proceedings of the 11th International Rapeseed Congress*, Vol. 4., 205–208.

Manavalan, P. and W.C. Johnson Jr. 1983. Sensitivity of circular dichroism to protein tertiary structure class. *Nature.*, 305: 831-832.

Mertens, H.D. and D.I. Svergun. 2010. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.*, 172: 128-141.

Mori, T., S. Utsumi and H. Inaba. 1979. Interaction involving disulfide bridges between subunits of soybean seed globulin and between subunits of soybean and sesame seed globulins. *Agri. & Biol. Chem.*, 43: 2317-2322.

Müntz, K., 1998. Deposition of storage proteins *Protein Trafficking in Plant Cells.* Springer, *pp.* 77-99.

Perera, S.P., T.C. McIntosh and J.P. Wanasundara. 2016. Structural properties of cruciferin and napin of *Brassica napus* (Canola) show distinct responses to changes in pH and temperature. *Plants*, 5: 36.

Petoukhov, M.V., D. Franke, A.V. Shkumatov, G. Tria, A.G. Kikhney, M. Gajda, C. Gorba, H.D. Mertens, P.V. Konarev and D.I. Svergun. 2012. New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallograph.*, 45: 342-350.

Provencher, S.W. 1982. A constrained regularization method for inverting data represented by linear algebraic or integral equations. *Comp. Phy. Comm.*, 27: 213-227.

Reed, J. and V. Kinzel. 1984. Near-and far-ultraviolet circular dichroism of the catalytic subunit of adenosine cyclic 5'-monophosphate dependent protein kinase. *Biochem.*, 23: 1357-1362.

Robin, J., V. Inquello, B. Mimouni and J. Azanza. 1991. Relationship between immunological properties and structural model of 11S rapeseed globulin. *Phytochem.*, 30: 3511-3513.

Sjodahl, S., J. Rodin and L. Rask. 1991. Characterization of the 12S globulin complex of *Brassica napus*. *Europ. J. Biochem.*, 196: 617-621.

Staswick, P., M. Hermodson and N. Nielsen. 1984. Identification of the cystines which link the acidic and basic components of the glycinin subunits. *J. Biol. Chem.*, 259: 13431-13435.

Svergun, D., C. Barberato and M.H. Koch. 1995. CRYSOL–a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallograph.*, 28: 768-773.

Svergun, D.I., M.H. Koch, P.A. Timmins and R.P. May. 2013. *Small angle X-ray and neutron scattering from solutions of biological macromolecules* Oxford University Press.

Tandang-Silvas, M.R.G., T. Fukuda, C. Fukuda, K. Prak, C. Cabanos, A. Kimura, T. Itoh, B. Mikami, S. Utsumi and N. Maruyama. 2010. Conservation and divergence on plant seed 11S globulins based on crystal structures. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics.*, 1804: 1432-1442.

Traka, M.H. and R.F. Mithen. 2011. Plant science and human nutrition: challenges in assessing health-promoting properties of phytochemicals. *The Plant Cell.*, 23: 2483-2497.

Withana-Gamage, T.S., D.D. Hegedus and X. Qiu. 2011. In silico homology modeling to predict functional properties of cruciferin. *J. Agri. Food Chem.*, 59(24): 12925-12938.

Yehuda, H., Y. Soroka, M. Zlotkin-Frušić, A. Gilhar, Y. Milner and S. Tamir. 2012. Isothiocyanates inhibit psoriasis-related proinflammatory factors in human skin. *Inflam. Res.*, 61: 735-742.

(Received for publication 5 April 2018)