

Profiling and characterization of *Camelina sativa* (L.) Crantz meal proteins

Janitha Wanasundara¹, Tara McIntosh¹, Dwayne Hegedus¹, Suneru Perera², Robert Tyler³, and Cathy Coutu¹

¹Agriculture and Agri-Food Canada

²University of Saskatchewan College of Agriculture and Bioresources

³University of Saskatchewan

January 24, 2022

Abstract

Protein from camelina seed is a valuable co-product that can be derived from the meal remaining after oil extraction. The current study describes the types and physicochemical properties of the major proteins present in camelina meal. Seed coat mucilage, which interferes with protein extraction, was removed from whole seeds by digestion with Viscozyme® and lipids were removed with hexane to obtain demucilaged/defatted meal. Protein comprised 51.3% of meal dry matter and the eight essential amino acids comprised 40.8% of total amino acids. The meal polypeptide profile showed bands originating from cruciferin (~44.1 and 51.7 kDa), napin (~14 kDa) and oil body proteins (OBP; ~15-20 kDa) resembling that of other crucifers. Cruciferins (11 isoforms) were the predominant proteins, while vicilins (6 isoforms) also were identified among the proteins soluble at pH 8.5. Among the proteins soluble at pH 3, napins (5 isoforms) comprised the majority, though late embryogenesis abundant proteins also were found. Camelina cruciferin and napin were confirmed to possess predominantly β -sheet and α -helix secondary structures, respectively. Camelina cruciferin structure was highly sensitive to changes in medium pH and underwent acid-induced denaturation at pH 3, but exhibited high thermal stability (>80°C) at neutral and alkaline pHs. The structure of camelina napins was less sensitive to pH. The major proteins associated with oil bodies were oleosins (6 isoforms). Identification and characterization of the properties of camelina meal proteins will enable strategic paths for co-product valorization.

Introduction

Interest in industrial oilseed crops dedicated to liquid biofuel production has been successful in bringing *Camelina sativa* (L.) Crantz, (camelina), commonly known as false flax or linseed dodder, to the Canadian prairies. Camelina, a member of the *Brassicaceae* family, produces seed containing three valuable biopolymers – oil (36-43%), protein (20-33%) and fibre (10-13%) – with mucilageneous polysaccharides of the seed coat comprising 6-10% of the seed weight and contributing to the soluble fibre fraction (Zubr 2010; Li et al., 2016; Zhu et al., 2016). Camelina oil is generally destined for liquid biofuels; however, the high omega-3 fatty acid content (30-40% α -linolenic acid) makes it suitable for use in human food applications and as a replacement for fish oil in aquaculture feed. Camelina meal, what remains after oil extraction, contains ~40-45% protein (Boyle et al., 2018; Li et al., 2014) and fibre, including polysaccharides. Transcriptome analysis of developing seeds showed that the seed storage proteins (SSPs) of camelina resemble those of other *Brassicaceae* species and primarily consist of cruciferin and napin (Nguyen et al., 2013). According to Russo and Reggiani (2015), a rapid and steady increase in camelina SSP (cruciferin and napin) occurs seven days after pollination and ends at around 42 days after pollination. Nutritional and techno-functional properties of camelina meal protein products relevant in plant-protein-based applications have been reported by few researchers (Boyle

et al., 2018; Kim and Netravali, 2012; Li et al., 2014). Camelina meal contains glucosinolates, namely, glucoarabin [9-(methylsulfinyl)nonylglucosinolate], glucocamelin [(10-(methylsulfinyl)decylglucosinolate)], and 11-(methylsulfinylundecylglucosinolate), at levels of 9-36 $\mu\text{mol/g}$ seed (Matthaus and Zubr, 2000; Schuster and Friedt, 1998). In the context of the bioeconomy, almost all of these components of camelina seed have economic value that can be valorized.

Archeological and historical evidence indicate that camelina was cultivated in Europe since the Bronze Age (Zubr, 1997). Sustainable production of camelina on marginal lands has been proven however, establishment of camelina as a financially competitive, contemporary, oilseed crop necessitates extraction and valorization of both the oil and meal co-products. In the current market, utilization of camelina protein requires nutritional, techno-functional, and safety evaluation. Information on the types of camelina SSPs proteins, and their structural and physico-chemical properties is limited. SSPs of the widely-cultivated crucifer, canola/rapeseed (*Brassica napus*), primarily consist of 11S cruciferin, a cupin family protein, and 2S napin, a prolamin family protein, each with distinct structural and physico-chemical properties that bring diverse end-use opportunities as a plant protein (Perera et al., 2016; Wanasundara, 2011; Wanasundara and McIntosh, 2013). Information on the genetic, chemical and physico-chemical details of camelina seed protein will fill this knowledge gap and optimize their utilization. This study investigated the types of proteins present in camelina meal, and the structural characteristics and physico-chemical properties of the major camelina SSPs.

MATERIALS AND METHODS

Seeds and seed processing

Seeds from a double haploid *C. sativa* cultivar (DH55) were obtained from plants grown under greenhouse conditions (harvesting at 85-100 d after planting) and stored at 4°C. Seed from two biological replicates grown in November-February 2014 and April-July 2015 was used.

Mucilage interferes with aqueous extraction of proteins and was removed (or reduced) before seed disruption. First, whole seeds were placed in deionized water at pH 4.5 (1 g seed:10 mL water) containing Viscozyme® (Sigma-Aldrich Canada Co.) at 0.1 mL enzyme/g seed and incubated at 40°C for 3 h with mixing using an overhead stirrer (850 rpm). The seed slurry was then passed through a #25 sieve (710 μm mesh size) to recover seeds. The seeds were washed 5 times with deionized water (1:10 w/v) to remove residual enzyme and mucilage, and then dried overnight in a forced air oven at 35°C. The seeds were then de-oiled according to AOCS method AM 2-93 (Firestone, 1997) using hexane and the resultant meal was air-dried in a fume hood for 12-16 h. Mucilage-reduced, oil-free meal was stored in airtight containers at ambient temperature.

Microscopy

Light Microscopy : Seeds were examined under a dissecting microscope (Nikon SMZ 1500) attached to a digital camera (Nikon Digital sight DS-5M) at 2 \times zoom range.

Transmission Electron Microscopy (TEM) : See sections (150 nm thick) were prepared using a microtome (Reichert-Jung ultra, Leica Microsystems Inc.), fixed sequentially in glutaraldehyde (3% v/v in 0.1 M sodium cacodylate buffer at pH 7.2) and 1% OSO_4 (in sodium cacodylate buffer), stained with uranyl acetate, embedded in Epon-Araldite resin (Polyscience Inc.) and then mounted on a 200-mesh copper grid. TEM images (5000-50,000 \times magnification) were obtained using a Hitachi HT7700 electron microscope (Hitachi-High Technologies Inc.) set at high contrast field and 80 kV. Oil droplet diameter and the number of droplets per unit area were recorded using ImageJ Fiji software (Life-Line version 2014; <https://imagej.net/software/fiji/>).

Chemical analysis

Moisture, ash and total protein content of whole seed and meal were analysed according to AOAC Official Method 934.01, AOAC Method 942.05 and AOAC Method 990.03 with combustion-based nitrogen content

using 6.25 as a conversion factor (Latimer & Horwitz, 2005), respectively. The level of phytate was determined using the method described by Oomah et al. (2008) with modifications. 2% (v/v) HCl and centrifugation at $1000 \times g$ for 20 min were used to obtain the extract containing phytate. The anion exchange column (AG-1-X8, Bio-k Laboratories Inc.) was equilibrated with 0.08% (v/v) HCl before addition of extract and the eluate (150 μ L) was mixed with Wade reagent (50 μ L; $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ + sulfosalicylic acid) in a microplate well. The extraction of phenolic compounds was carried out using 70% aqueous acetone (1:30, w:v, meal-to-solvent ratio) at room temperature for 1 h. The extract was then centrifuged at $10,000 \times g$ for 20 min and the absorbance of the recovered supernatant was measured at 326 nm (Bio-Rad xMarkTM microplate spectrophotometer, Bio-Rad Laboratories Inc.). Total phenolic content was expressed as sinapic acid equivalent/g of meal as per Oomah et al. (2010). Meal amino acid content (18 amino acids, asparagine and glutamine combined with aspartic and glutamic acid, respectively) was evaluated using AOAC Method 994.12 (Latimer & Horwitz, 2005).

Protein solubility and isolation

Seed storage proteins : Aqueous meal extracts were prepared at pH values between 2 and 12 according to Perera et al. (2016). The soluble protein content of extracts was determined from nitrogen content as above. The polypeptide profile of the soluble protein at each pH was evaluated using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) as described below. Cruciferin and napin were extracted from mucilage-reduced, oil-free meal and purified following the processes described by Perera et al. (2016) for *B. napus*. Purified cruciferin and napin fractions were lyophilized, analyzed for total protein content and stored at -20°C .

Oil body proteins (OBPs) : Oil bodies (OBs) were separated from mucilage-free seeds according to Maurer et al. (2013) with modifications. Seed was soaked overnight in Milli Q water (1g seed:4 mL water) at 4°C , homogenized at 10,200 rpm for 1 min using a Polytron PT3100 homogenizer with PTDA 3020/2T generator (Kinematica Inc.) and filtered through three layers of cheesecloth to obtain a filtrate with OBs. The seed residue was re-extracted with deionized water (1:4, w:v) two more times, the filtrates combined, and sucrose added to achieve 25% (w:v) final concentration. The top cream layer containing OBs was recovered by centrifugation at $5,000 \times g$ for 30 min, mixed with 20% (w:v) sucrose solution at pH 11 and then centrifuged again. The OB layer was recovered and dispersed in 100 mL of deionized water and dialyzed (2 kDa MW cut-off) against deionized water for 24 h at 4°C . The dialysate was centrifuged and the supernatant mixed with 1% SDS (10:1, v:v, dialysate-to-SDS solution). The aqueous medium containing OBPs was recovered by centrifugation ($7500 \times g$ for 30 min) and then dialyzed (2 kDa MW cut-off) for 24 h against deionized water at room temperature. Lyophilized OBPs were reconstituted in deionized water and separated on a HiPrep 26/10 desalting column with deionized water as the mobile phase with detection at 280 nm (ÄKTA Explorer system, GE Healthcare Life Sciences).

Electrophoresis and LC-MS/MS analysis

One dimensional electrophoresis (1DE) : Meal and OB proteins were separated using SDS-PAGE in one dimension (1D) under both reducing and non-reducing conditions as described by Wanasundara et al. (2012). Purified cruciferin (1 mg/mL) and napin (4 mg/mL) were dissolved in 0.1 M phosphate buffer (pH 8.0) containing 0.1 M NaCl and also separated by 1DE. Homogeneous 20% polyacrylamide gels with 1.7-40 kDa molecular weight standards (SpectraTM Molecular Low-Range protein ladder, Thermo Fisher Scientific Inc.) were used for napin and cruciferin. 8-25% polyacrylamide gradient gels with 4.6–170 kDa molecular weight standards (PageRuler pre-stained protein ladder, Thermo Fisher Scientific Inc.) were used for all other proteins. Native-PAGE under non-denaturing conditions was performed for cruciferin and napin with reference proteins bovine serum albumin (66 kDa monomer and 132 kDa dimer) and jack bean urease (272 kDa trimer and 545 kDa hexamer (Sigma Aldrich Co.) as described by Perera et al. (2016).

Two-dimensional electrophoresis (2DE) : Purified proteins (cruciferin and napin at 1 mg protein/mL and OBPs at 1 μ g/mL) were mixed with sample buffer (1:2, v:v) containing 6.7 M urea, 2% (w:v) 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonic acid (CHAPS), 0.5% carrier ampholytes, 0.001% bro-

mophenol blue and deionized water. pH 3-10 ampholytes (Bio-Lyte® 3-10, Bio-Rad Laboratories Inc.) were used for cruciferin and OBPs, and pH 9-11 ampholytes (Zoom™, Thermo Fisher Scientific Inc.) for napin. Immobilized pH Gradient (IPG) strips (pH 3-10 IPG for cruciferin and OBPs; pH 9-12 IPG for napin) were re-hydrated in each protein solution overnight at 4°C prior to isoelectric focusing (IEF). IEF was conducted using an IEF Cell (Bio-Rad Laboratories Inc.) for 2 h at 50 V, then 2 h at a voltage gradient from 200-4,000 V, and finally for 9 h at 4,000 V. Prior to running the second dimension, each protein strip was equilibrated for 15 min in buffer 1 [1.8 g urea, 1 mL of 10% SDS, 1.25 mL of 1.5 M Tris-HCl (pH 8.8), 1 mL of 100% glycerol and 0.6 mL of Milli Q water] followed by 15 min in buffer 2 [1.8 g urea, 1 mL of 10% SDS, 1.25 mL of 1.5 M Tris-HCl (pH 8.8), 1 mL of 100 % glycerol, 125 mg iodoacetamide and 0.6 mL of Milli Q water). The second-dimension separation (SDS-PAGE) was carried out using a Bio-Rad Mini-Protean® tetra cell (Bio-Rad Laboratories Inc.) with hand-cast 12% (for cruciferin) or 16 % polyacrylamide (for napin and OB protein) gels. The gels were stained with 0.1% Coomassie Blue R-250 and de-stained in 50% (v:v) methanol in water with 10% (v:v) acetic acid solution for 1-2 h to visualize separated proteins.

Liquid chromatography-mass spectroscopic analysis (LC-MS/MS)

Protein spots separated by 2DE were digested in-gel with trypsin and the resultant peptides analysed by LC-MS/MS as described by Parker et al. (2005) and Senko et al. (2003). The LC-MS/MS data were processed with XCalibur 3.0.63 software (Thermo Fisher Scientific Inc.) and analysed with the Proteome Discoverer 1.4.0.228 software suite (Thermo Fisher Scientific Inc.). Statistical analysis of MS data was conducted with Scaffold Q+S (Proteome Software, Inc.). The abundance of a particular protein was expressed based on the percentage of Normalized Total Spectra (NTS) of a specific peptide found in an individual protein spot.

Protein structure analyses

Fourier Transform-Infrared (FTIR) spectroscopy : FT-IR spectroscopy was conducted using an Agilent Cary 630 ATR-FTIR analyzer (Agilent Technologies Inc.) according to Perera et al. (2016). Secondary structure details of proteins were obtained by Fourier self-deconvolution (FSD) of the amide I region (1610–1700 nm) from the spectra using Agilent Resolution Pro version 5.2.0 with the following parameters: resolution enhancement factor (K) = 2.5, full width at half height = 14 cm⁻¹ and apodization filter = Bessel for cruciferin, whereas K = 2.8 and full width at half height = 18 cm⁻¹ were used for napin.

Circular Dichroism (CD) : Far-UV spectra of proteins [1 mg/mL prepared in 10 mM sodium citrate buffer (pH 3), 10 mM sodium phosphate buffer (pH 7) or 10 mM ammonium chloride/ammonia buffer (pH 10)] were acquired at 25°C using a PiStar-180 spectrometer (Applied Photophysics Ltd.) as described by Perera et al. (2016) and Withana-Gamage et al. (2013). Molar ellipticity of proteins was calculated with background correction using the CDNN 2.1 software package (Applied Photophysics Ltd.).

Intrinsic fluorescence : The fluorescence emission spectra of protein slurries (5%, w:v, cruciferin and 10%, w:v, napin prepared at pH 3, 7 and 10 in the buffer systems noted above) were monitored for the tryptophan residue (excitation at 280 nm and emission from 290-450 nm) in the temperature range of 22-93°C using a fluorescence spectrophotometer (Agilent Eclipse, Model G9800A, Agilent Technologies Inc.).

Surface hydrophobicity with anionic fluorescent probing : Binding of 1-anilino-8-naphthalensulfonate (ANS, an anionic fluorescence probe) with hydrophobic regions on the surface of the protein in solution (0.05–0.25 mg/mL at pH 3, 7 and 10) was measured. The fluorescence probe was excited at 390 nm and the emission monitored at 470 nm and 500 nm for cruciferin and napin, respectively, using a Cary eclipse fluorescence spectrophotometer (Agilent Technologies Inc.). Surface Hydrophobicity Index (S₀) was calculated as described in Perera et al. (2016) and Withana-Gamage et al. (2013).

Thermal stability assessment using differential scanning calorimetry (DSC) : Protein slurries (5%, w:v, cruciferin and 10%, w:v, napin) prepared at pH 3, 7 and 10 using the buffer systems noted above were used for DSC analysis. Samples were placed into aluminum pans, hermetically sealed with a Tzero press (TA Instruments) and temperature-induced structural changes recorded from 30degC to 130degC under constant nitrogen purging using a TA Q2000 differential scanning calorimeter (TA Instruments). An empty, hermet-

ically sealed pan was used as a reference and the thermograms analysed using TA universal analysis 2000 software (TA Instruments).

Statistical analysis

Two biological replicates of camelina seed, one grown during the winter and the other during the spring-summer in the greenhouse, were used and chemical analyses were carried out in triplicate (technical replicates). All assessments were performed according to a completely random design (CRD) and data analysis was conducted using the linear model procedure [lm] followed by Tukey's test for mean separation with R statistical software version 3.2.2 (<https://www.r-project.org/>).

RESULTS & DISCUSSION

Removal of seed coat mucilage

Camelina seeds are oblong in shape and comparatively small in size (Figure 1A) with average dimensions of 2.2 mm in length and 1.2 mm in width. Specialized cells in the seed coat of crucifers, such as *Arabidopsis thaliana*, *Capsella bursa-pastoris* and *Lepidium flavum* (Soukoulis et al., 2018), produce extracellular, pectinaceous mucilage that absorbs water and expands and forms a gel around the seed coat. Mucilage aids in fruit and seed dispersal, defends against pathogens, enhances water uptake during germination and acts as a barrier to water and oxygen diffusion into inner tissues to prevent seed germination during undesirable environment conditions (Soukoulis et al., 2018). Interference of seed coat mucilage with protein extraction and recovery has been reported for flax (Wanasundara and Shahidi, 1997) and yellow mustard (Wanasundara et al., 2012), and the same was expected for camelina as the seed contains about 6.7% mucilage by weight.

Li et al. (2016) showed that camelina mucilage is composed of galactose (58.1%), glucose (25%), rhamnose (11.6%) and xylose (5.2%), whereas arabinose, glucose and xylose are more abundant in flax, yellow mustard and chia mucilage, respectively (Soukoulis, et al., 2018). Viscozyme(r), an enzyme complex derived from *Aspergillus* spp. containing arabinase, cellulase, β -glucanase, hemicellulase and xylanase, was more effective in breaking down camelina seed coat mucilage than other carbohydrases (pectinases and cellulases) or mild alkaline (NaHCO_3) solutions (data not shown). The optimum conditions for Viscozyme® pre-treatment for camelina seeds were 3 h of soaking with 0.1 mL enzyme/g seed with vigorous stirring to remove/reduce seed coat mucilage. Microscopic examination of soaked, untreated seed showed water-absorbed mucilage as a halo around the seed (Figure 1B) compared to dry, non-hydrated seed (Figure 1A) or hydrated seed after enzyme treatment (Figure 1C). Although non-enzymatic approaches such as heating of hydrated camelina seed (as slurries) to temperatures up to 85°C to reduce slurry viscosity have been used, the possibility of partial denaturation of proteins at ~50°C during hot-pressing of camelina seed, as has been reported by Boyle et al. (2018), indicates the importance of low temperatures (< 50°C) during seed pre-treatments to minimize heat-induced alterations in proteins.

Seed microstructure and meal chemical composition

TEM microscopy of camelina seed cotyledon cells showed distinct separation of OBs from protein-rich areas or protein storage vacuoles (PSVs) that were, more or less, spherical in shape and evenly distributed within the cytoplasm (Figures 1D and E). The discrete dark areas inside PSVs may be globoids containing phytic acid crystals as reported for other seeds (Weber and Neumann, 1980). Camelina OBs had an average diameter of 0.68 μm with a packing density of ~3 OBs/ μm^2 . The OBs of *Brassica* seeds, including rapeseed, have diameters between ~0.2 and 3.0 μm (Katavic et al., 2006). The comparatively small size of camelina OBs and the high packing density may facilitate larger oil storage capacity within the cotyledon cells. The distinct membrane of camelina OBs (Figure 1E) may prevent coalescence and provide steric hindrance or electrostatic repulsion that may stabilize the OBs similar to other oil-bearing seeds (Tzen, 2012). Mechanical pressing during oil extraction may rupture these membranes and allow oil leakage. Subsequent extraction with non-polar solvents, such as hexane, removes any oil that remains trapped within the cellular matrix.

Cellular components remaining after oil extraction, including organelle membranes, cell wall membranes and proteins, collectively make up the spent meal.

The mucilage-reduced, oil-free meal contained 51.3% protein and 6.8% ash (mineral) (Table 1); these levels were higher than those reported for camelina pressed cake meal [~40-45% protein (Boyle et al. 2018; Li et al., 2014) and 5% ash (Zubr, 1997)]. The levels of phytate and phenolic compounds, which are considered anti-nutrients in many seed -derived meals (Wanasundara, 2011), were at 6.1% and 1.0% (as sinapic acid equivalents), respectively (Table 1).

Characteristics of camelina meal proteins

Amino acid composition

In mucilage-reduced, oil-free camelina, meal protein, essential amino acids (His, Ile, Leu, Lys, Thr, Trp, Val, Met+Cys, Phe+Tyr) represented ~40% of total amino acids (Table 1) and were within the range of values reported by Li et al. (2014) and Zubr (2003). Lysine, which is often the limiting amino acid in meal derived from crucifer seeds, including camelina (Russo and Reggiani, 2015), was at 5%. Camelina meal had a sulphur-containing amino acid (Met + Cys) level of 5% (Table 1), which is comparable to canola meal (Wanasundara, 2011). The level of branched-chain amino acids was ~15%, indicating there is potential for the use of camelina proteins beyond animal feed applications, such as in nutritional supplements.

Protein solubility in relation to pH

Crucifer SSPs are known for their distinct solubility characteristics with changing pH (Perera et al., 2016; Wanasundara et al., 2012; Wanasundara and McIntosh, 2013). The protein solubility of camelina meal was similar to that from other eudicot seeds, including crucifers, with high solubility at extremely alkaline (43.6% at pH 10) and acidic pHs (16.2% at pH 1.5) (Figure 2A). The lowest protein solubility (~5% of total meal protein) was at pH 4.5, which is typical for canola and mustards (Wanasundara et al., 2012). In the pH range between 1.5 and 6.5, soluble proteins were of lower molecular weight with the ~15 kDa proteins likely being napin (Figure 2B). Of the proteins soluble above pH 6.5 (Figure 2B), ~40-54 kDa and ~22-33 kDa proteins are most likely cruciferin-derived polypeptides. The more diffuse polypeptides of <31 kDa observed after extraction at pH 12 (Figure 2B) may be indicative of alkali-induced hydrolysis. The present study confirms that varying pHs can be used to selectively recover camelina proteins. Highly alkaline conditions can be used to recover mixtures of SSPs, while lower pH conditions can be used to solubilise lower molecular weight proteins as has been reported for other *Brassica* species (Wanasundara and McIntosh, 2013). Although extreme alkaline pH allows for higher protein yields, products may be subject to alkali-induced denaturation.

Polypeptides associated with camelina meal

The polypeptides associated with mucilage-reduced, oil-free camelina meal were examined by SDS-PAGE (Figure 3A). The 44.1 kDa and 51.7 kDa polypeptides detected under non-reducing conditions are likely monomers of cruciferin similar to those reported for *B. napus* canola (Nietzel et al., 2013; Perera et al., 2016; Wanasundara, 2011). S-S bonds stabilize these polypeptides since bands between 20 and 39 kDa observed under reducing conditions may be similar to the α - and β -chains of cruciferin. The 14 kDa and 10.4 kDa polypeptides observed under non-reducing and reducing conditions, respectively, are characteristic of napins, which also contain a S-S bond (Wanasundara, 2011). Polypeptide profiles of defatted camelina meal reported by Boyle et al. (2018) and Naugen et al. (2013) also indicated bands representing cruciferin and napin. Apart from these, 15-20 kDa polypeptides also were present. Boyle et al. (2018) speculated that the ~15 to 20 kDa polypeptides found in highly alkaline extracts (pH 12) of camelina meal might be glutelins; however, they are most likely oleosins associated with OBPs as reported for canola (Wijesundera et al., 2013). The lightly stained polypeptides between 67 kDa and 69 kDa may be membrane-bound proteins, aggregates of cruciferin free chains or unidentified SSPs.

Polypeptide profiles obtained for the enriched cruciferin (Figure 3B; pH 8.5 extraction) and napin (Figure 3C; pH 3 extraction) fractions showed that the majority of intensely-stained bands were related to cruciferin and napin, respectively, similar to those of other crucifers, such as *B. napus* (Perera et al., 2016), *Sinapis alba*

(Marambe et al., 2015; Wanasundara et al., 2011) and *A. thaliana* (Withana-Gamage et al., 2013) obtained using a similar separation process. Electrophoretic separation of the fractions under native conditions showed that cruciferin was present in both trimeric and hexameric forms (Figure 3D), while napin was monomeric (Figure 3E), keeping in mind that both cruciferin and napin monomers consist of two polypeptides linked via S-S bonds.

Identification of proteins associated with camelina meal

Cruciferins

2DE of the cruciferin-enriched fraction yielded 29 proteins in the neutral pH region in the range of 17-55 kDa (Figures 4A, B). Based on the normalized total spectral (NTS) values (Table 2), these spots consisted primarily of cruciferin isoforms (89.8-100%), with vicillins, napins and non-storage proteins present in minor amounts. Eleven of the twelve cruciferin isoforms encoded by genes in the *C. sativagenome* (Supplementary Table S1) were identified in the isolated protein (Supplementary Table S2). CsCruA-1-G2 was the most prevalent isoform and was present in all of the 29 2DE spots, followed by CsCruC-1-G2 (27 spots), CsCruC-1-G3 (27 spots), CsCruB-1-G3 (25 spots) and CsCruB-1-G1 (22 spots). The only isoform that was not present in the purified cruciferin fraction was CsCruD-1-G3; however, CsCruD-1-G3 was found in the purified oil body fraction (Supplementary Tables S4 & S5). All identified cruciferins were in the molecular weight range of 50-65 kDa, indicative of intact monomers/protomers. Cruciferin protomers consist of α - (~30-40 kDa) and β - (~20 kDa) polypeptide chains attached via a S-S linkage (Wanasundara, 2011). Camelina cruciferins ranged from 453 (CsCruD-1-G1 and CsCruD-1-G2) to 586 (CsCruB-1-G1) amino acid residues with variations depending on the respective gene (Supplementary Figure S1). The cruciferin C isoforms possess the longest α -chain (3-4 kDa longer), the highest isoleucine and tyrosine levels and the lowest alanine, phenylalanine and serine levels compared to the cruciferin A, B and D isoforms. Seeds from camelina lines devoid of cruciferin C isoforms (three cruciferin C genes edited by CRISPR/Cas9 technology) showed differences in amino acid and fatty acid compositions (Lyzenga et al., 2019).

Vicilins

Eight vicilin isoforms are encoded by genes in the camelina genome (Supplementary Table S1), of which six were identified in 13 2DE spots from the cruciferin-enriched fraction (Supplementary Table S2). These six isoforms accounted for 0.2-10.2% of the total protein in the 13 spots. CsVic1B-1-G2 and CsVic2-1-G2 were not found in the cruciferin-enriched, napin-enriched or oil body protein fractions (Supplementary Tables S2-S5). The most prevalent vicilin isoform was CsVic1A-1-G1 (Csa19g031870, 11 spots), followed by CsVic1A-1-G2 (Csa01g025880, 7 spots), CsVic1B-1-G3 (Csa15g039300, 7 spots) and Csa05g038120 (6 spots). Vicilin is a 7S trimeric globulin found mainly in legumes (Shewry et al., 1995). Similar to cruciferin, vicilin is a cupin super-family protein containing a characteristic 'jelly-roll' β -barrel domain (Shewry et al., 1995); however, it forms a trimeric quaternary structure as the protomer is not processed and lacks the internal S-S bond present in cruciferin that is involved in structural changes associated with hexamer assembly. Mature trimeric vicilins have molecular weights of ~150-200 kDa (Shewry, 1998). Camelina vicilin protomers ranged from 462-535 amino acids (Supplementary Figure S2) with molecular weights between 52.0 kDa and 58.5 kDa and migrate with cruciferin subunits in SDS-PAGE gels under non-reducing conditions, but are separable under reducing conditions. The ~69 kDa and ~53 kDa polypeptides observed in camelina meal under reducing conditions (Figure 3A) may be vicilins. Some vicilins, such as that from pea (Casey et al., 1986), are proteolytically-processed yielding smaller polypeptides. If this is also the case for camelina vicilins, they might migrate with the free α - and β -chains of cruciferin (Figure 3B). Delseny and Raynal (1999) and Shewry and Casey (1999) reported two genes, AtPAP85 (Q9LUJ7) and AtVCL22 (Q9SK09), that encode 7S proteins in *A. thaliana* that were highly similar to pea vicilin and related legume 7S proteins.

Napins

2DE of the napin-enriched fraction resulted in 18 protein spots in the pH range of 9-12 (Figures 4C, D), similar the pI of ~11 reported for *B. napus* napin (Crouch et al., 1983). Eight genes encoding napin have been identified in camelina (Supplementary Table S1; Nguyen et al., 2013), of which four (CsNap-1-G1,

CsNap-1-G3, CsNap-3-G3 and CsNap-4-G1) contributed between 12.1% and 100% of the combined total napin per spot (Table 3, Supplementary Table S3). CsNap-3-G3 and CsNap-4-G1 were present in all 18 2DE spots, whereas CsNap-1-G1 and CsNap-1-G3 were present in 17 and 12 spots, respectively. The number of amino acid residues in these eight napin isoforms ranged from 161 (CsNap-3-G1) to 166 (CsNap-3-G3; Supplementary Figure S3). Late embryogenesis abundant (LEA) proteins also were present in the napin-enriched fraction (Table 3). Brassica seed LEA proteins are basic with a pI > 7 and have disordered secondary structures (Amara et al., 2014). They are small proteins of ~10 kDa to 30 kDa and accumulate in the embryo during the late stages of seed development (Hong-Bo, et al., 2005). The non-napin polypeptide bands observed at ~15 kDa in the 1DE separation of the napin-enriched fraction (Figure 3C) may be LEA proteins. The LEA proteins contribute to seed longevity and germination by associating with other proteins (e.g., enzymes) and stabilizing their structure during episodes of environmental stress, such as dehydration during seed maturation (Hand et al., 2011). The considerable amount of LEA proteins in camelina seed may be related to its strong drought tolerance making it a suitable crop in arid, low soil moisture areas (Eynck and Falk, 2013).

Oil body proteins

OBPs stabilize the tightly packed oil bodies in cotyledon cells (Figure 1E). Oleosin accounts for 75-80% of the OBPs associated with *B. napus* and *A. thaliana* OBs, while the remaining 20-25% consists of caleosin and steroleosin (Jolivet et al., 2009). Surface-active agents, such as SDS, adsorb to the oil-water interface and results in displacing of OBPs. The polypeptides isolated from camelina OBPs were ~ 15-20 kDa and 26 kDa (Figure 5A) under both non-reducing and reducing conditions and may represent oleosins and caleosins, respectively, while the ~ 37.8 kDa polypeptide may be steroleosins as reported for other crucifers (Jolivet et al., 2009; Katavic et al., 2006; Tzen, 2012). A washing step at pH 11 was included to remove co-extracted SSPs, particularly napin, and very few low molecular weight proteins were visible in the isolated OBPs. Separation of OBPs by 2DE showed protein spots distributed over a broad pH range of 3-12 (Figures 5B-E). Of the twelve oleosin isoforms encoded by genes in the *C. sativa* genome (Supplementary Table S1), seven (CsOle1-1-G1, CsOle1-1-G3, CsOle2-1-G2, CsOle3-1-G1, CsOle3-1-G3, CsOle4-1-G1 and CsOle4-1-G2) and few other proteins related to the oleosin family were detected by LC-MS/MS analysis (Figures 5B-E, Supplementary Tables S4 and S5). CsOle1-1-G3 and CsOle2-1-G2 were present in 18 of 20 2DE spots obtained from the pH 3-10 and pH 9-12 OBP fractions. CsOle3-1-G1 was found in only a single 2DE spot. The five isoforms not observed in the OBP fraction were not present in the cruciferin-enriched or napin-enriched fractions. The largest oleosin isoforms identified were CsOle2-1-G1 and CsOle2-1-G2 which comprise 201 amino acid residues, whereas the smallest isoforms were CsOle3-1-G1, CsOle3-1-G2 and CsOle3-1-G3 which comprise 144 amino acid residues (Supplementary Figure S4). The oleosins have a highly conserved central hydrophobic domain (Supplementary Figure S4) that varies in length depending on the isoform (Tzen, 2012).

The genes that are responsible for caleosin and steroleosin expression in camelina are not well-defined; however, proteomics data revealed that caleosin-like proteins encoded by Csa03g006900, Csa05g023090 and Csa07g038560, were present in the OBP fraction (Supplementary Tables S4 and S5). Hydroxysteroid dehydrogenase also was detected. Katavic et al. (2006) reported that hydroxysteroid dehydrogenase is possibly the steroleosin present in *B. napus*, similar to sesame seed steroleosin. This putative steroleosin in camelina has a molecular weight between ~39 kDa and 42 kDa, which is in agreement with that reported by Tzen, (2012).

Structural characteristics of the major SSPs and their response to pH and temperature

Cruciferin-enriched fraction

Many of the FT-IR spectral features of the camelina cruciferin-enriched fraction (Figure 6) were similar those from *A. thaliana* (Withana-Gamage et al., 2013) and *B. napus* (Perera et al., 2016) with subtle differences observed for the -PO₃ (970 cm⁻¹), C-O-P (1070 cm⁻¹) and -P=O (1170 cm⁻¹) functional groups. The estimated values for secondary structure components based on deconvolution of the amide I band 1600-1690

cm^{-1} (Figure 6) were $43\% \pm 0.9\%$ β -sheets, $10.1\% \pm 0.3\%$ α -helix, $19.8\% \pm 0.8\%$ β -turns and $3.9\% \pm 0.6\%$ random structure. Comparable levels of β -sheet structure were reported for *B. napus* (45.6%; Perera et al., 2016) and *A. thaliana* (~44%; Withana-Gamage et al., 2013). Camelina cruciferin structure exhibited acidic pH-induced, secondary structure changes as observed in the far-UV CD spectra (Figure 7A). The values for molecular ellipticity (θ) changed and α -helix content significantly increased ($11.6\% \pm 0.3\%$) at pH 3, but not at pH 10 (Figure 7A and Table 4), compared to that at pH 7 ($2.9\% \pm 0.2\%$) (Table 4). *Brassica napus* cruciferin also exhibits unfolding at acidic pH (Perera et al., 2016).

An increase in protein surface hydrophobicity values (S_0) at pH 3 (7393.1 ± 32) compared to that at pH 7 (557.8 ± 2.4) (Table 4) indicates exposure of hydrophobic residues that had been buried (Korte and Herrmann, 1994). Together with the changes in intrinsic fluorescence of the Trp residue, the observed decrease in the maximum fluorescence intensity (F_{max}), red shift in the maximum emission wavelength (λ_{max}), and a higher ratio of fluorescence intensity (F) at 350 nm and 330 nm at pH 3 compared to neutral pH ($F_{350}/F_{330} = 1.13$ at pH 3 and 0.71 at pH 7) (Figure 7 C and D) all support acidic pH-induced unfolding of camelina cruciferin tertiary structure. Similar behaviour has been reported in acidic environments for cruciferins from other Crucifer species (Perera et al., 2016; Withana-Gamage, 2013).

Camelina cruciferin showed a significantly ($p < 0.05$) higher maximum denaturation temperature at pH 10 ($T_m = 83.1^\circ\text{C}$) than at pH 7 ($T_m = 80.6^\circ\text{C}$), but similar thermal energies (enthalpy, $-\Delta H$) were required for denaturation (Table 4) indicating high thermal stability of the structure. The absence of a thermal transition peak at pH 3 (Table 4) further supported the structural unfolding indicated in other analyses. The intrinsic fluorescence data (F_{max} , λ_{max} and F_{350}/F_{330}) (Figure 8) obtained for protein heated above 83°C (or the denaturation temperatures) at pH 7 and 10 (Figures 8 B and C) were similar to the values at pH 3 at ambient temperature (22°C). This confirms the existence of unfolded/denatured structure due to heat or acidic pH. The predominant native quaternary structure for camelina cruciferin may be a mixture of trimers and hexamers (as indicated from native PAGE) below 80°C at both pH 7 and 10 (Figures 8B and C). The predominance of unfolded structure was evident as the temperature increased above 83°C , similar to what was observed for *B. napus* cruciferin (Perera et al., 2016). The higher values for denaturation temperature (95.5°C , $\Delta H = \sim 14 \text{ J g}^{-1}$ for cold-pressed meal and 93.3°C , $\Delta H = 9 \text{ J g}^{-1}$ for hot-pressed meal) reported for camelina proteins (Boyle et al., 2018) may be related to other associated proteins/components in the tested material, as indicated from the ~79-82% purity based on total N, or to structural alterations at highly alkaline pH (pH 12, alkali-induced unfolding) during extraction and $(\text{NH}_4)_2\text{SO}_4$ precipitation.

In *B. napus* cruciferin, acid-induced dissociation is considered reversible, whereas heat-induced denaturation causes irreversible changes (Wanasundara, 2011). At pH 3, *B. napus* cruciferin assumes a molten globule structure (Perera et al., 2016; Schwenke and Linow, 1982), a partially-folded conformation of a globular protein with near-native compactness, significant secondary structure, insignificant tertiary structure and a substantial amount of solvent-exposed hydrophobic residues compared to its native state. The behaviour exhibited in the structural features and their changes during pH destabilization indicate that camelina cruciferin also assumes a molten globule state under acidic conditions that is somewhat reversible.

Napin-enriched fraction

The α -helix content of $32.1\% \pm 0.4\%$ (Figure 6) determined for camelina napin was similar to that reported for *B. napus* napin (Pantoja-Uceda, 2003; Wanasundara, 2011). The spectral peaks resembling β -sheets ($1627\text{--}1638 \text{ cm}^{-1}$) and β -turns ($1674\text{--}1684 \text{ cm}^{-1}$) most likely originated from bond vibrations responsible for short, extended chains attached to helical motifs. These peaks resembled those observed for cruciferin (Figure 6); however, they do not originate from true β -structures in napin. This phenomenon is commonly observed in other helical proteins, such as hemoglobin and cytochrome C (Byler and Susi, 1986). Camelina napin showed far-UV CD spectra typical of a helical protein similar to described by Greenfield (2006) with $-\theta$ peaks at $\sim 222 \text{ nm}$ and $\sim 208 \text{ nm}$ and a $+\theta$ peak at $\sim 195 \text{ nm}$ (Figure 7B). The secondary structural features of napin did not show differences in magnitude with changes in pH (Figure 7B and Table 4). Napin from *B. napus* is also structurally stable in response to changes in pH (Perera et al., 2016). The increase in S_0 at pH 3 and the decrease at pH 10 (Table 4), although much less than that observed with cruciferin, suggested

some degree of structural change for napin. Use of intrinsic fluorescence to probe napin tertiary structure details during pH changes was not possible due to the very low number of Trp and Tyr residues. No thermal transition peak above or below 100°C was observed for camelina napin dry powder or slurries at pH 3, 7 and 10. However, Boyle et al. (2018) reported that temperatures of ~104-107°C with lower enthalpy values ($\Delta H = \sim 0.2-3.0 \text{ J g}^{-1}$) are required for denaturation of camelina napin. The high thermal-stability nature of crucifer napin at a temperature of >100°C ($\Delta H = \sim 9-14 \text{ J g}^{-1}$) is reported elsewhere for *B. napus* (Krzyzaniak et al., 1998). The lower sensitivity of napin structure to changes in pH and its high stability in response to heat is a distinct contrast to the properties of cruciferin found in the SSPs of Brassicas, including camelina.

Cruciferin and napin, and to a lesser extent vicilins and OBPs, comprise the bulk of the camelina meal proteins. Camelina protein is rich in S-containing amino acids, contains proteins that are soluble at acidic pH, and is structurally stable at low pH and high temperatures, making camelina seed protein similar to that from other economically-important crucifers. Protein products rich in cruciferin and/or napin could be obtained from camelina using scalable processes described for *Brassica* oilseeds to benefit camelina value chain development. Unique properties of individual proteins (cruciferin and napin) could position camelina alongside commercial ventures being developed for canola/rapeseed protein (<https://www.meritfoods.com/canola.protein/>; <https://napiferyn.com/>; <https://www.dsm.com/corporate/solutions/nutrition-health/canolapro-plant-protein.html>). The ability to change the camelina seed protein profile and consequently influence the quality attributes of SSP fractions (Lyzenga et al, 2019) may cater to specific nutritional or functional applications. Seed coat mucilage is also a valuable hydrocolloid (Soukoulis et al., 2018); however, removal or reduction of camelina seed coat mucilage is critical for high protein recoveries. Therefore, effective technologies for recovering seed coat mucilage are necessary to bring this valuable co-product into an integrated valorization scheme. OBs that are stabilized by OBPs may also have economic value in food, feed and personal care applications (<https://www.botaneco.com/food-and-feed>).

CONCLUSIONS

Microstructural features of camelina seeds confirmed the organization of oil bodies and protein-rich areas (storage vacuoles) in cotyledon cells as being typical of crucifer oilseeds, implying that similar oil extraction and meal co-product processing approaches are applicable. Seed coat mucilage, which is an impediment to recovering proteins from oil-free meal, can be removed/reduced using seed pre-treatments, such as treatment with polysaccharide-degrading enzymes. Several isoforms of both cruciferin and napin comprise the primary SSPs, while minor proteins, such as vicilins, OBPs and LEA proteins, make up the remainder of the meal protein. The OBPs are primarily of oleosins, but other types occur. Camelina cruciferin and napin exhibit structural characteristics and pH and temperature stabilities that correspond to those of proteins from *B. napus*. The composition and properties of camelina seed protein indicate its value as a plant-derived protein with somewhat unique characteristics compared to other plant proteins. This study shows that camelina SSPs have close similarity to *Brassica napus*(canola/rapeseed) proteins with some subtle differences

References

- Amara I, Zaidi I, Masmoudi K, Ludevid M, Pagès M, Goday A, et al. Insights into late embryogenesis abundant (LEA) proteins in plants: From structure to the functions. *Am J Plant Sci.* 2014;5:3440-3455.
- Boyle C, Hansen L, Hinnenkamp C, Ismail BP. Emerging camelina protein: Extraction, modification, and structural/functional characterization. *J Am Oil Chem Soc.* 2018; 95: 1049-1062. doi:10.1002/aocs.12045
- Byler DM, Susi H. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers.* 1986;25:469-487.
- Casey R, Domoney C, Ellis N. Legume storage proteins and their genes. In Mifflin BJ, editor, *Oxford Surveys of Plant Molecular and Cell Biology*. Oxford University Press, Oxford: United Kingdom; 1986. P. 1-95.

- Crouch ML, Tenbarger KM, Simon AE, Ferl R. cDNA clones for *Brassica napus* seed storage proteins: evidence from nucleotide sequence analysis that both subunits of napin are cleaved from precursor polypeptide. *J Mol Appl Genet*. 1983;2:273-283.
- Delseny M, Raynal M. Globulin storage proteins in crucifers and non-legume dicotyledonous families. In Shewry PR, Casey R, editors, *Seed Proteins*. Kluwer Academic: Amsterdam, Netherland; 1999. p. 427-451.
- Eynck C, Falk KC. Camelina (*Camelina sativa*). In Sing BP, editor, *Biofuel Crops: Production, Physiology and Genetics*. CAB International; 2013. p.369-391.
- Firestone DE, editor. *Official Methods and Recommended Practices of the AOCS* (5th ed.): AOCS Press: Champaign, IL: 1997.
- Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc*. 2006;1;2876-2890.
- Hand SC, Menze MA, Toner M, Boswell L, Moore D. LEA proteins during water stress: not just for plants anymore. *Annu Rev physiol*. 2011;73:115-134.
- Hong-Bo S, Zong-Suo L, Ming-An S. LEA proteins in higher plants: structure, function, gene expression and regulation. *Colloids Surf B*. 2005;45:131-135.
- Jolivet P, Boulard C, Bellamy A, Larré C, Barre M, Rogniaux H, et al. Protein composition of oil bodies from mature *Brassica napus* seeds. *Proteomics*. 2009;9:3268-3284.
- Katavic V, Agrawal GK, Hajdich M, Harris SL, Thelen JJ. Protein and lipid composition analysis of oil bodies from two *Brassica napus* cultivars. *Proteomics*. 2006;6:4586-4598.
- Kim JT, Netravali AN. Non-food application of camelina meal: Development of sustainable and green biodegradable paper-camelina composite sheets and fibers. *Polym Compos*. 2012;33 :1969-1976.
- Korte T, Herrmann A. pH-dependent binding of the fluorophore bis-ANS to influenza virus reflects the conformational change of hemagglutinin. *Eur Biophys*. 1994;23:105-113.
- Krzyzaniak A, Burova T, Haertlé T, Barciszewski J. The structure and properties of napin-seed storage protein from rape (*Brassica napus* L.). *Food/Nahrung*. 1998;42:201-204.
- Latimer Jr DGW, Horwitz DW, editors. *Official Methods of Analysis of AOAC International* (18th ed.): AOAC Intl: 2005.
- Li N, Guangyan Q, Sun S, Wang D. Characterization of gum isolated from camelina seed. *Ind Crops Prod*. 2016;83:268-274.
- Li N, Qi G, Sun XS, Wang D, Bean S, Blackwell D. Isolation and characterization of protein fractions isolated from camelina meal. *Trans ASABE*. 2014;57:169-178.
- Lyzenga WJ, Harrington M, Bekkaoui D, Wigness M, Hegedus DD, Rozwadowski KL. CRISPR/Cas9 editing of three CRUCIFERIN C homoeologues alters the seed protein profile in *Camelina sativa*. *BMC Plant Biol*. 2019;19:292.
- Marambe PWMLHK, Shand PJ, & Wanasundara JPD. An in-vitro investigation of selected biological activities of hydrolysed flaxseed (*Linum usitatissimum* L.) Proteins. *J Am Oil Chem Soc*. 2008;85:1155-1164.
- Matthaus B, Zubr J. Variability of specific components in *Camelina sativa* oilseed cakes. *Ind Crop Prod*. 2000;12:9-18.
- Maurer S, Waschatko G, Schach D, Zielbauer BI, Dahl J, Weidner T, et al. The role of intact oleosin for stabilization and function of oleosomes. *J Phys Chem*. 2013;117:13872-13883.
- Nguyen HT, Silva JE, Podicheti R, Macrander J, Yang W, Nazarens TJ, et al. Camelina seed transcriptome: a tool for meal and oil improvement and translational research. *Plant Biotechnol J*. 2013;11:759-769.

- Nietzel T, Dudkina NV, Haase C, Denolf P, Semchonok DA, Boekema EJ, et al. The native structure and composition of the cruciferin complex in *Brassica napus*. J Biochem. 2013;288:2238-2245.
- Oomah BD, Blanchard C, Balasubramanian P. Phytic acid, phytase, minerals, and antioxidant activity in canadian dry bean (*Phaseolus vulgaris* L.) cultivars. J Agri Food Chem. 2008;56:11312-11319.
- Oomah BD, Corbe A, Balasubramanian P. Antioxidant and anti-inflammatory activities of bean (*Phaseolus vulgaris* L.) hulls. J Agri Food Chem. 2010;58:8225-8230.
- Pantoja-Uceda D, Bruix M, Gimenez-Gallego G, Rico M, Santoro J. Solution structure of RicC3, a 2S albumin storage protein from *Ricinus communis*. Biochem. 2003;42:13839-13847.
- Parker CE, Warren MR, Loiselle DR, Dicheva NN, Scarlett CO, Borchers CH. Identification of components of protein complexes. Methods Mol Biol. 2005;301:117-151.
- Perera SP, McIntosh TC, Wanasundara JP. Structural properties of cruciferin and napin of *Brassica napus* (Canola) show distinct responses to changes in pH and temperature. Plants. 2016;5:36. doi:10.3390/plants5030036
- Russo R, Reggiani R. Seed Protein in *Camelina sativa* (L.) Crantz var. Calena. Int J Plant Soil Sci. 2015;8:1-6.
- Schuster A, Friedt W. Glucosinolate content and composition as parameters of quality of *Camelina* seed. Ind crop Prod. 1998;7 :297–302.
- Schwenke KD, Linow KJ. A reversible dissociation of the 12 S globulin from rapeseed (*Brassica napus* L.) depending on ionic strength. *Food / Nahrung*. 1982;26:K5-K6.
- Senko MW, Remes PM, Canterbury JD, Mathur R, Song Q, Eliuk SM, et al. Novel parallelized quadrupole/linear ion trap/orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates. Anal Chem. 2003;85:11710-11714.
- Shewry PR, Casey R. The 2S Albumin storage Proteins. In Shewry PR, Casey, R, editors, Seed Proteins. Kluwer Academic: Amsterdam, Netherland; 1999. p. 563–586.
- Shewry PR. Manipulation of seed storage proteins. In Lindsey K, editor, Transgenic Plant Research. Harwood Academic: Amsterdam, Netherland; 1998. p. 135–149.
- Shewry PR, Napier JA, Tatham AS. Seed storage proteins: Structures and biosynthesis. *Plant Cell*. 1995;7 :945–956.
- Soukoulis C, Gaiani C, Hoffmann L. Plant seed mucilage as emerging biopolymer in food industry applications, Curr opin Food Sci. 2018;22:28-42.
- <https://doi.org/10.1016/j.cofs.2018.01.004>
- Tzen JTC. Integral proteins in plant oil bodies. *ISRN Botany*. 2012; <https://doi.org/10.5402/2012/173954>
- Wanasundara JPD. Proteins of Brassicaceae oilseeds and their potential as a plant protein source. Crit Rev Food Sci Nutr, 2011;51:635-677.
- Wanasundara JD, Abeysekara S, McIntosh T, Falk K. Solubility differences of major storage proteins of *Brassicaceae* oilseeds. J Ame Oil Chem Soc. 2012;89:869-881.
- Wanasundara JPD, McIntosh TC. Process of aqueous protein extraction from Brassicaceae oilseeds. US 8557963 (Patent). 2013.
- Wanasundara, PKJPD, Shahidi F. Removal of flaxseed mucilage by chemical and enzymatic treatments. Food Chem. 1997;59:47-55.
- Weber E, Neumann D. Protein bodies, storage organelles in plant seeds. Biochem Physiol Pflanz. 1980;175: 279-306.

Wijesundera C, Boiteau T, Xu X, Shen Z, Watkins P, Logan A. Stabilization of fish oil-in-water emulsions with oleosin extracted from canola meal. *J Food Sci.* 2013;78:C1340-1347. doi: 10.1111/1750-3841.12177

Withana-Gamage TS, Hegedus DD, Qiu X, McIntosh T, Wanasundara JPD. Structural and physicochemical property relationships of cruciferin homohexamers. *J Agri Food Chem.* 2013;61:5848–5859.

Zhu X, Wang D, Sun XS. Physico-chemical properties of camelina protein altered by sodium bisulfite and guanidine-HCl. *Ind Crops Prod.* 2016;83:453-461.

Zubr J. Oil-seed crop: *Camelina sativa* . *Ind Crops Prod.* 1997;6:113-119.

Zubr J. Dietary fatty acids and amino acids of *Camelina sativa* seed. *J Food Qual.* 2003;26: 451-462.

Zubr J. Carbohydrates, vitamins and minerals of *Camelina sativa* seed. *Nut Food Sci.* 2010;40: 523-531.

ACKNOWLEDGMENTS

Acknowledgements :

This study is supported by the project “Developing *Camelina sativa* as a modern crop platform” that received funds from Global Institute of Food Security, University of Saskatchewan (Project # J-001661) and Agriculture and Agri-Food Canada (Project # J-001589).

Author contributions

Suneru P. Perera: Investigation, Data analysis, Writing - original draft & editing

Tara McIntosh: Investigation, Methodology, Data analysis

Cathy Coutu: Methodology, Software, Validation, Data curation

Robert T. Tyler: Supervision, Review & editing

Dwayne D. Hegedus: Funding acquisition, Project administration, Writing - review & editing.

Janitha P. D. Wanasundara: Conceptualization, Funding acquisition, Project administration, Supervision, Writing - review & editing.

Ethics Statement

This manuscript is not published elsewhere.

Conflict of interest

The authors declare that they have no conflict of interest

Figure legend

Figure 1 Microscopic images of camelina seeds showing the effectiveness of Viscozyme® treatment on removing seed coat mucilage and cotyledon ultrastructure. Upper panels show dissecting light microscopic images. A dry seed with average length L and width W indicated (A), a seed showing the hydrated mucilage layer 1 h after hydration (B), and a dried seed after Viscozyme® treatment (C). Lower panels show TEM images of a cross-section of cotyledon cells with PSVs pointed by white arrows and OBs by white arrow heads (D), and OBs at a higher magnification with black arrows showing the oil body protein (OBP) layer (E).

Figure 2 Solubility of camelina meal proteins in relation to pH. Panel A: percentage solubility of the meal protein (%N × 6.25; one-time extraction) at different pH values. Panel B: polypeptide profile of soluble proteins at different pH values. SDS-PAGE separation was under non-reducing conditions using 8-25% gradient gels. MWM=Molecular weight markers.

Figure 3 SDS-PAGE of polypeptides from camelina meal, purified cruciferin and napin. SDS-PAGE profiles of mucilage-reduced, oil-free meal (A), cruciferin-enriched fraction (B), and napin-enriched fraction (C) under

reducing (+ME) and non-reducing (-ME) conditions. Molecular weight markers (MWM) are on the left-hand side of each panel with molecular weights (kDa) of the polypeptide bands shown in right hand margin of each sample lane in Panel A. Panels D and E show the native PAGE profile of purified cruciferin and napin, respectively. An 8-25% gradient gel was used for A, B, D and E, whereas a 20% gel was used for C. Bovine serum albumin (BSA) and urease were used as standards for native PAGE.

Figure 4 Two-dimensional electrophoresis separation profiles and reconstructed images for purified cruciferin (A and B) and napin (C and D) from camelina meal. Numbers in B and D depict the protein spots in A and C, respectively, that were subjected to LC-MS/MS analysis. Cartoons in B and D approximate the size and staining intensity observed in A and C. MWM=Molecular weight markers (kDa).

Figure 5 Polypeptide profiles of OBP from camelina. Panel A: SDS-PAGE profile in one dimension of meal proteins, intermediate products (subnatant) and OBP under non-reducing (-ME) and reducing (+ME) conditions with estimated molecular masses (kDa) of the purified OBPs obtained using 8-25% gradient precast gels. (MWM=Molecular weight markers). Panels B and D: Two-dimensional electrophoresis separation of OBPs on 14% polyacrylamide gels obtained using first dimension IEF pH 3-10 (B) and pH 9-12 (D) followed by SDS-PAGE under non-reducing conditions for second dimension. Panels C and E: Graphical representation of B and D, respectively. Numbers in C and E depict the protein spots observed in A and D, respectively, that were subjected to LC-MS/MS analysis. Protein spot sizes in C and E approximate the size and staining intensity observed in B and D. (MWM=Molecular weight markers).

Figure 6 FTIR spectra of camelina cruciferin and napin together showing differences in Amide I region and the regions of phosphorous containing functional groups. Separate small graphs are for the Fourier Self Deconvoluted (FSD) spectra of amide I region ($1600-1690\text{ cm}^{-1}$) of cruciferin and napin showing peaks assigned for secondary structural components. FSD parameters: resolution enhancement factor (K) = 2.5 (cruciferin) and 2.8 (napin); full width of half height (FWHH) = 14 cm^{-1} (cruciferin) and 18 cm^{-1} (napin); apodization filter = Bassel (cruciferin and napin).

Figure 7 Changes observed in the far-UV CD spectra of purified camelina cruciferin (A) and napin (B) with the pH of the medium and the changes of Fluorescence intensity (C) and the maximum emission wavelength (λ_{max}) (D) of Tryptophan residues of purified cruciferin as at ambient temperature (22°C).

Figure 8 Influence of temperature on tryptophan fluorescence of purified camelina cruciferin at different pH values. Change of fluorescence intensity (top plot in each panel) and maximum emission wavelength (λ_{max}) (bottom plot in each panel) at pH 3 (A), pH 7 (B) and pH 10 (C). Values for F_{350}/F_{330} ratio are provided in the top panel.

Table Legend

Table 1 Chemical composition and amino acid composition of mucilage-reduced-oil-free camelina meal.

Table 2 Abundance of cruciferin, vicilin, napin and other proteins of 2DE separated cruciferin from camelina based on normalized total spectral (NTS) values. (Derived from the supplementary Table 2)

Table 3: Abundance of napin, late embryogenesis abundance (LEA) protein, cruciferin and other proteins of 2DE separated napin from camelina based on normalized total spectral (NTS) values (Derived from the supplementary Table 3).

Table 4 Changes of secondary structure components (%), surface hydrophobicity and thermal denaturation parameters as influenced by the pH of the medium observed for purified camelina cruciferin napin

Hosted file

Table 1_Camelina proteins.docx available at <https://authorea.com/users/456860/articles/553818-profiling-and-characterization-of-camelina-sativa-l-crantz-meal-proteins>

Hosted file

Table 2_Camelina proteins.docx available at <https://authorea.com/users/456860/articles/553818-profiling-and-characterization-of-camelina-sativa-l-crantz-meal-proteins>

Hosted file

Table 3_Camelina proteins.docx available at <https://authorea.com/users/456860/articles/553818-profiling-and-characterization-of-camelina-sativa-l-crantz-meal-proteins>

Hosted file

Table 4_Camelina proteins.docx available at <https://authorea.com/users/456860/articles/553818-profiling-and-characterization-of-camelina-sativa-l-crantz-meal-proteins>









