JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Characterization of *Arabidopsis thaliana* Lines with Altered Seed Storage Protein Profiles Using Synchrotron-Powered FT-IR Spectromicroscopy

Thushan S. Withana-Gamage,[†,§] Dwayne D. Hegedus,[†,§] Xiao Qiu,[§] Peiqiang Yu,[#] Tim May,[⊥] Derek Lydiate,[†] and Janitha P. D. Wanasundara*[,†,§]

[†]Agriculture and Agri-Food Canada, Saskatoon, Saskatchewan S7N 0X2, Canada

[§]Department of Food and Bioproduct Sciences, University of Saskatchewan, Saskatoon, Saskatchewan S7N 5A9, Canada

[#]Department of Animal and Poultry Science, University of Saskatchewan, Saskatoon, Saskatchewan S7N 5A9, Canada

[⊥]Canadian Light Source, University of Saskatchewan, Saskatoon, Saskatchewan S7N 0X4, Canada

Supporting Information

ABSTRACT: *Arabidopsis thaliana* lines expressing only one cruciferin subunit type (double-knockout; *CRUAbc*, *CRUaBc*, or *CRUabC*) or devoid of cruciferin (triple-knockout; *CRU-*) or napin (napin-RNAi) were generated using combined T-DNA insertions or RNA interference approaches. Seeds of double-knockout lines accumulated homohexameric cruciferin and contained similar protein levels as the wild type (WT). Chemical imaging of WT and double-knockout seeds using synchrotron FT-IR spectromicroscopy (amide I band, 1650 cm^{-1}, $\nu C{=}O$) showed that proteins were concentrated in the cell center and protein storage vacuoles. Protein secondary structure features of the homohexameric cruciferin lines showed predominant β-sheet content. The napin-RNAi line had lower α-helix content than the WT. Lines entirely devoid of cruciferin had high α-helix and low β-sheet levels, indicating that structurally different proteins compensate for the loss of cruciferin. Lines producing homohexameric CRUC showed minimal changes in protein secondary structure after pepsin treatment, indicating low enzyme accessibility. The Synchrotron FT-IR technique provides information on protein secondary structure and changes to the structure within the cell.

KEYWORDS: *cruciferin, synchrotron mid-IR, seed storage protein, bioaccessibility, pepsin digestion*

INTRODUCTION

Crucifer oilseeds, such as canola/rapeseed, camelina, and mustard, contain 22−28% protein that is primarily composed of seed storage proteins (SSP).[1] Crucifer SSP are predominantly of the 11S globulin and 2S albumin types, with lesser contributions from oleosins, defensins, and late embryogenesis abundant proteins.[1] Due to their abundance in mature seed, the 11S and 2S proteins determine the properties of the crucifer seed protein fraction. Cruciferin, the most abundant protein, is an 11S hexameric globulin of 300−360 kDa and a member of the cupin superfamily. Cruciferin protomers (subunits) comprise 27 β-strands and 7 helices that are folded into two β-barrel domains and two extended helix domains, each containing two helices.[2] The second most abundant protein is napin, a 12−16 kDa 2S albumin of the prolamin superfamily. The secondary structure of napin is predominantly α-helix, although some β-sheet structure is also present.[3]

Multiple genes encode cruciferins and napins. For example, in the model crucifer *Arabidopsis thaliana* cv. Columbia, three paralogous genes, At5g44120.3, At1g03880.1, and At4g28520.1, encode CRUA, CRUB, and CRUC cruciferin protomers, respectively.[4] Each cruciferin protomer is composed of a heavy acidic (α) and a light basic (β) chain linked by a disulfide bond. In wild type (WT) lines, the resultant hexameric cruciferin is an assortment of these three protomers. Similarly, in *Arabidopsis*, napin is encoded by multiple (five) genes with the mature protein consisting of two disulfide-linked subunits.[3]

Cruciferin secondary, tertiary, and quaternary structure affects biochemical and technofunctional properties such as surface active, thermal, and rheological properties. In addition, enzyme accessibility, which is an important consideration in determining the availability of amino acids and peptides upon digestion, as well as safety (i.e., potential immunogenicity) of the protein is related to the protein structure.[5−7] The native WT cruciferin has a heterogeneous subunit composition; therefore, the ability to extrapolate molecular structure information to functions is limited. Information regarding the contribution of individual cruciferin protomers to the final physicochemical properties of the hexamer is needed to rationalize protein quality improvement through targeted genetic selection. Whereas single SSP protomers expressed in microbial systems have been examined, these lack post-translational modifications that occur during the normal transport and deposition into protein storage vacuole (PSV).[2,8,9] Previously, we predicted that cruciferin CRUA, CRUB, and CRUC protomers can assemble as homomeric

Received: October 11, 2012
Revised: December 19, 2012
Accepted: January 8, 2013
Published: January 8, 2013

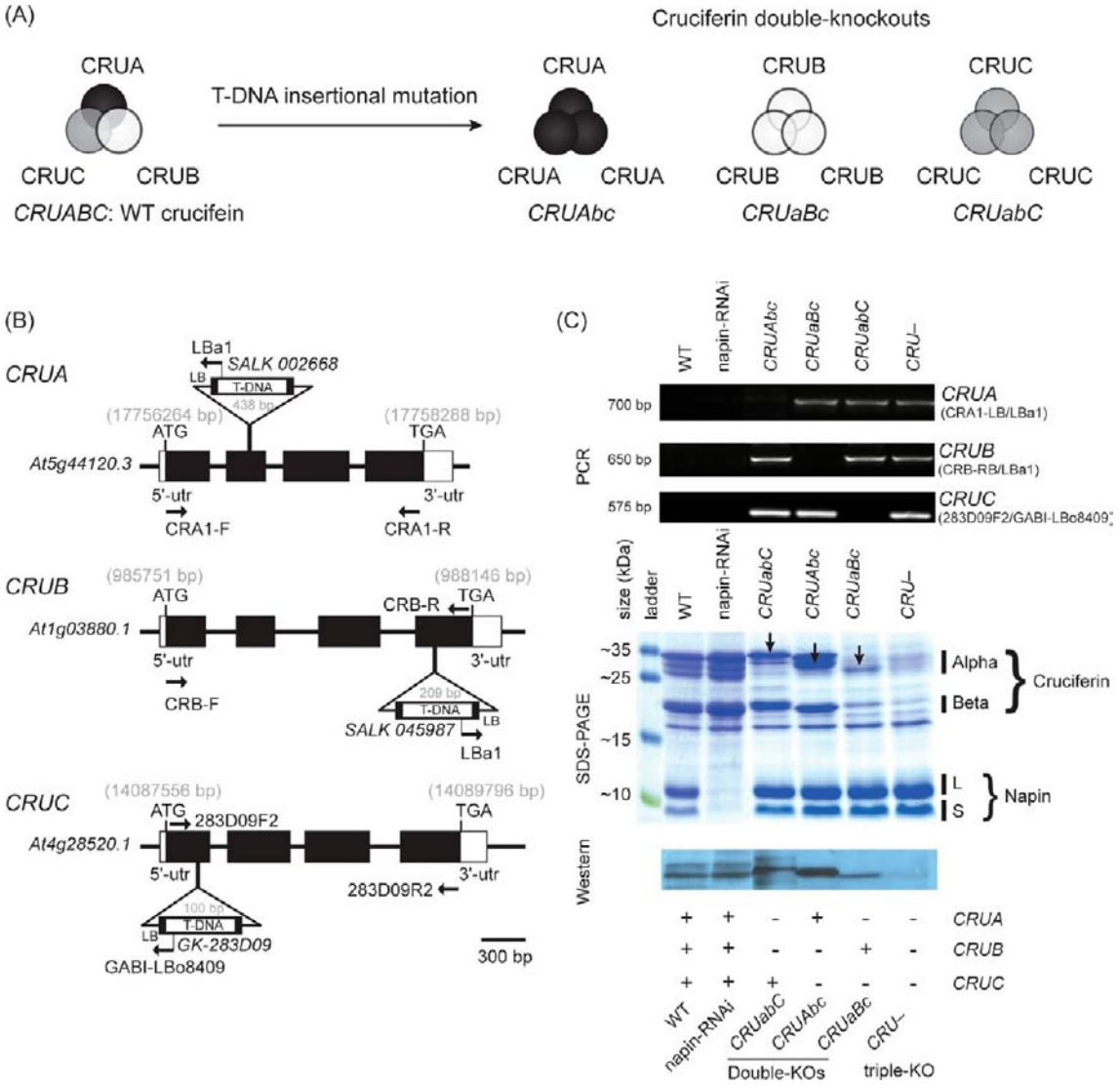

Figure 1. Seed storage protein altered *Arabidopsis*. (A) An illustration of expected cruciferin subunit composition of double-knockout (KO) lines. (B) Schematic diagram of At5g44120.3, At1g03880.1, and At4g28520.1 with T-DNA insertion, and primer location for T-DNA diagnostic PCR. Black boxes represent exons, white boxes represent untranslated region (utr), lines between black boxes represent introns. CRA1-F, CRB-F, and 283D09F2 are forward primers. CRA1-R, CRB-R, and 283D09R2 are reverse primers. LBa1 is a primer specific for the T-DNA left border, and bp is base pairs. Scale bar = 300 bp. (C) T-DNA insertion in the cruciferin gene observed by PCR using gene-specific primers and a T-DNA tag-specific primer suggested by SALK and GABI-Kat and further confirmed by SDS-PAGE of seed extracts and Western immunoblotting with a primary polyclonal antiserum raised against *Arabidopsis* cruciferin α-chain.

proteins and that these would exhibit structural and functional differences.[10] In the present study, *A. thaliana* lines expressing only one type of cruciferin (double-knockout lines) were generated to eliminate the heterogeneity associated with multiple protomer types. In addition, lines devoid of either napin (napin-RNAi) or cruciferin (triple-knockout) were also generated.

Histochemical staining of seed cross sections allows observing details of subcellular components but limits revealing information on the constituent chemical components, their structure details, and levels. Vibrational spectroscopy in the Raman and mid-infrared (IR) regions is very useful in obtaining information on protein secondary structure arrangement by using the unsaturated amide bond in the polypeptide backbone. In general, IR spectroscopy applies for molecules with symmetric vibrations of nonpolar groups and Raman spectroscopy applies for molecules with asymmetric vibration of nonpolar groups. The fluorescence from intrinsic amino acids and the coextracted phenolic compounds interferes with Raman spectroscopy of plant proteins[11] but can be eliminated by Fourier transform near-infrared (FT-NIR) Raman spectroscopy.[12] Infrared spectroscopy in the mid-IR region is nondestructive and provides label-free fingerprint-like spectra originating from the characteristic vibrational frequencies of the chemical bonds of the molecules of the cell and is therefore useful in studying cellular macromolecules.

Synchrotron-powered FT-IR spectromicroscopy was used to examine the SSP from these lines including their secondary structure features, whereas proteins were in their natural forms and setting within the PSV; therefore, modifications or changes that might occur during separation and isolation of the proteins were also eliminated. Furthermore, changes in protein secondary structure features were assessed as an indicator of digestive enzyme accessibility directly within seed tissues.

MATERIALS AND METHODS

Generation of *Arabidopsis* Lines with Altered SSP Profiles. To obtain cruciferin composed of identical subunits (i.e., homomeric cruciferin), two of three cruciferin genes in *Arabidopsis* cv. Columbia were inactivated by genetically combining T-DNA insertions to develop double-knockout lines (Figure 1A). The following T-DNA insertion lines obtained from either the *Arabidopsis* Biological Resources Centre (www.abrc.osu.edu) or GABI-Kat (www.gabi-kat.de) were used: *CRUaBC* (*cruACRUBCRUC*) (SALK 002668, T-DNA inserted in second exon of At5g44120), *CRUAbC* (*CRUAcruBCRUC*) (SALK 045987, T-DNA inserted in fourth exon of At1g03880), and *CRUABc* (*CRUACRUBcruC*) (GK-283D09, T-DNA inserted in first exon of At4g28520) (Figure 1B). The homozygous double-knockout lines referred to as *CRUAbc* (*CRUAcruBcruC*), *CRUaBc* (*CruACRUBcruC*), and *CRUabC* (*cruAcruBCRUC*) were obtained by conventional crossing. The protein from these double-knockout lines is referred to as CRUA, CRUB, and CRUC homohexamers. A homozygous cruciferin null mutant (i.e., triple-knockout, *cruAcruBcruC* or *CRUabc* or *CRU-*) was also generated, and the protein of this line is referred to as CRU-. Table 1 summarizes the relationships of genes and proteins, including nomenclature used throughout this study. To generate a line devoid of napin, an RNAi cassette containing nucleotides 328−513 of At4g27150 was cloned in the sense and antisense orientations between the 3X35S promoter and the OS terminator of pGSA1252, which separated them with a GUS gene fragment to create the RNAi hairpin structure. This was transformed into *A. thaliana* cv. Columbia, with the BAR gene conferring Bialaphos resistance in transformed plants. This napin-depleted line is referred to as napin-RNAi.

Table 1. Relationship of Cruciferin Gene and Proteins and Their Nomenclature in This Study

line	gene	protein[a]
wild type (WT)	*CRUACRUBCRUC* or *CRUABC*	CRUABC
double knockouts	*CRUACruBCruC* or *CRUAbc*	CRUA
	CruACRUBCruC or *CRUaBc*	CRUB
	CruACruBCRUC or *CRUabC*	CRUC
triple knockout	*CruACruBCruC* or *CRUabc* or *CRU-*	CRU-

[a]These acronyms are used as adjective for protomer, trimer, or hexamer of WT and double-knockout proteins (e.g., CRUA protomer, CRUA homotrimer, or CRUA homohexamer).

Seed Production. Seeds from the WT, cruciferin-knockout, and napin-RNAi lines, as well as *Brassica napus* (doubled-haploid line DH12075) were used. All plants were grown in plastic pots containing a COCO/Sunshine soil mixture (1:1) (Michigan Peat, Houston, TX, USA) consisting of coconut fiber/peat moss/vermiculite (1:3:3, w/w/w) with Osmocote PLUS 15−9−12 controlled release fertilizer (Scotts Co. LLC, Marysville, OH, USA), in a greenhouse under the following conditions: 16 h light (~800 Wm^2) at 21 °C and 8 h dark at 16 °C. Plants were bagged at the flowering stage; seeds were harvested at maturity and stored dry at room temperature.

Contents of Lipid and Protein. Protein (based on total N) and lipid contents of the seeds were determined by the combustion method using a Flash EA1112 N-analyzer (Thermo Fisher Scientific) with a conversion factor of 5.64 g protein/g nitrogen[13] and the AOCS method AM 2-93,[14] respectively. All analyses were performed in triplicate.

DNA Extraction and Polymerase Chain Reaction (PCR). Leaves of 4-week-old plants were used for DNA extraction. Briefly, two leaves were ground in liquid N_2 with 400 μL of DNA extraction buffer (250 mM NaCl, 25 mM EDTA, 0.5% w/v SDS, 200 mM Tris-HCl, pH 8.5) and then centrifuged at 13000*g* for 2 min; a 300 μL aliquot of the supernatant was transferred to a new microcentrifuge tube, 300 μL of cold isopropanol added, and the tube gently inverted three times and then incubated for 2 min before centrifugation at 13000*g* for 5 min. The pellet was washed with 70% (v/v) ethanol and resuspended in 50 μL of H_2O. The PCR consisted of 20 ng/μL of DNA, 10× PCR buffer, 25 mM $MgCl_2$, 10 mM dNTPs, 5 μM *CRUA*, *CRUB*, or *CRUC* forward and reverse primers (Supporting Information, Table S1), 5 μM SALK LBa1 or GABI LB08409 primer, and 2 units of Taq DNA polymerase. PCR conditions were as follows: 95 °C for 5 min followed by 29 cycles of 93 °C for 1 min, 50 °C for 1 min, 72 °C for 1.5 min, and a final extension at 72 °C for 10 min.

One-Dimensional Gel Electrophoresis (1D PAGE). Fifteen milligrams of dried seeds was ground in liquid N_2 with 500 μL of extraction buffer consisting of thiourea/urea lysis buffer [7 M urea, 2 M thiourea, 18 mM Tris-HCl, 4% (w/v) CHAPS, 0.2% (v/v) Triton X-100, pH 8.8], 0.34% (w/v) protease inhibitor cocktail (Sigma Steinhein, Germany), 150 mM DNase I, and 10 mM RNase A and incubated at 4 °C for 10 min. Twelve microliters of 100 mM dithiothreitol (DTT) was added and the mixture incubated for another 20 min at 4 °C and then centrifuged at 20000*g* at 4 °C for 20 min. Finally, the supernatant was collected and frozen at −80 °C if PAGE was not carried out on the same day. The supernatant was mixed with 10% (v/v) glycerol and Bromophenol Blue (0.05%, w/v) and electrophoresis carried out in a polyacrylamide gel (12.5% separating and 5% stacking) at 125 V for 2 h. The protein bands was stained with Bio-Safe Coomassie Blue (Bio-Rad Laboratories) or transferred to a nitrocellulose membrane for Western blotting.

Western Blotting. After 1D PAGE, proteins were transferred onto nitrocellulose membrane with 1× Towbin's transfer buffer [25 mM Tris-HCl, pH 7.6, 192 mM glycine, and 20% (v/v) methanol] at 100 V for 1 h. The membranes were incubated with a solution of 5% (w/v) nonfat skim milk powder (Nestle Carnation, Solon, OH, USA) in 1× TBS buffer (20 mM Tris-HCl and 14 mM NaCl, pH 7.5) overnight with gentle shaking at 4 °C. For cruciferin detection, the membranes were probed with a polyclonal antiserum against recombinant *Arabidopsis* cruciferin with affinity for the α-subunit in 1× TBST buffer (0.3% Tween 20 in TBS) at a dilution 1:20000 for 1.5 h. The membrane was washed three times for 10 min with 1× TBST buffer and incubated with a secondary antibody (goat anti-rabbit IgG conjugated to horseradish peroxidase; Bio-Rad Laboratories) at a dilution of 1:20000 in 1× TBST buffer containing 5% nonfat dried skim milk powder for 1 h. The membrane was washed three times using 1× TBST buffer and horseradish peroxidase detected using the ECL Plus Western Blotting system (GE Healthcare).

Sample Preparation and Synchrotron FT-IR Spectromicroscopy. *Arabidopsis* seeds (with or without any treatment) were embedded in optimal cutting temperature (OCT) compound (Sigma Chemical, St. Louis, MO, USA), frozen (−10 °C), and sectioned (~6 μm thickness) using a microtome (Tissue Tech, Edison, NJ, USA). Unstained sections were mounted onto BaF_2 windows (13 mm × 1 mm disk, Spectral Systems, Hopewell Junction, NY, USA) for transmission mode synchrotron FT-IR imaging. Fifteen to twenty seeds were sectioned and mounted on a single BaF_2 disk. Mid-IR spectra of the seed cross sections were collected at Canadian Light Source Inc. (CLSI, University of Saskatchewan, Saskatoon, Canada) on beamline 01B1-1. Microscopic images of seed cross sections were obtained using a charged-couple device (CCD) camera in a Hyperion 3000 confocal microscope using 20× and 36× focusing objectives (Bruker Optik GmbH, Ettlingen, Germany). Spectromicroscopic mapping of seed tissues was performed using a Bruker Optics Vector 70v/S interferometer coupled with a Hyperion 3000 IR confocal microscope equipped with a 36× objective and a 64 × 64 pixel focal plane array (FPA) detector. The mid-IR probe pulses were generated as synchrotron radiation from the bending magnet with an energy range of 0.070−0.744 eV. Raster scans were collected in the mid-IR range of 4000−800 cm^{-1} with a 10 × 10 μm aperture at a resolution of 4 cm^{-1}. The step size of the scanning was 10 μm (1 μm in single cell scanning) in both the *x* and *y* orientations for the maps, and 128 interferograms were co-added to a single spectrum. Sections from 10 seeds were analyzed.

The amide I band at 1650 cm^{-1} of the synchrotron FT-IR spectra was integrated (baseline 1720−1588 cm^{-1}) to construct protein chemical images of seed tissues using Opus v6.5 software (Bruker Optik GmbH). For lipid distribution, the carbonyl ester peak at 1747 cm^{-1} (integral baseline 1786−1720 cm^{-1}) was used. Fourier self-

deconvolution of the amide I peak was carried out using the same software. Band assignments and quantification of secondary structure elements, α-helix (1659.7 ± 0.7 cm^{-1}), β-sheet (1616.8 ± 0.8, 1627.8 ± 1.3, 1638.4 ± 1.7, and 1694.1 ± 0.2 cm^{-1}), β-turn (1671.0 ± 0.3 and 1682.3 ± 0.3 cm^{-1}), and random structure (1649.1 ± 1.5 cm^{-1}) were made as described previously.[15,16]

Enzyme Accessibility of Tissue Proteins. Seeds were treated with porcine pepsin (2410 units/mg solid, Sigma Chemical Co.) to assess enzyme accessibility as follows. Seeds were soaked in water for 1−2 h, washed three times to remove most of the seed mucilage, and then frozen at −70 °C. Frozen seeds were slightly cracked using a motar and pestle to expose the inner tissues. Simulated gastric fluid (SGF) was prepared by dissolving pepsin in 35 mM NaCl in 0.084 N HCl (pH 2.0) to maintain an enzyme to seed protein ratio of 1:3 (w/w). The cracked seeds were incubated with SGF at 37 °C for 2 h, at which time pepsin digestion was terminated by adjusting the pH to 7.4 with 1 M NaOH. The treated seeds were washed three times with ddH_2O and freeze-dried before being used for synchrotron FT-IR analysis as described above.

Statistical Analysis. To quantify secondary structure components and compare integral peak area of the amide I band, 10 spectral data sets were subjected to one-way analysis of variance (ANOVA) using the General Linear Model (GLM) procedure of SAS version 9.1.[17] If the main effect was significant ($P < 0.05$), and mean separation was done by calculating Fisher's protected least significance difference (LSD). Principal component analysis (PCA) on secondary structure data ($n = 10$) obtained from Fourier self-deconvolution (FSD) was performed using SAS procedure PROC PRINCOMP. A detailed description of the application of PCA method in the FSD amide I components is found in the literature.[5,6] The first three principal components (PC1, PC2, and PC3) were extracted on the basis of the criterion that they accounted for at least 10% of the total variance in the data set (SAS manual).[17]

RESULTS AND DISCUSSION

Confirmation of Lines with Altered SSP Profiles. The proteins (cruciferin and napin) from *A. thaliana* are closely related to those from more economically important oilseed species, such as *B. napus*.[10,18−21] The *A. thaliana* cv. Columbia genome contains four genes that encode cruciferin protomers: At5g44120.3 (gene *CRUA*: syn *CRU4 or CRA1*), At1g03880.1 (gene *CRUB*: syn *CRU3 or CRB*), At1g03890.1 (gene *CRU2*), and At4g28520.1 (gene *CRUC:* syn *CRU1*, *CRC*, or *CRU3*) (http://www.arabidopsis.org/ and http://www.uniprot.org, accessed in July 2012). Most transcripts are derived from At4g28520 and At5g44120 and less so from At1g03880.[22] At1g03890 is linked in tandem to At1g03880 that encodes CRUB, but the former is very poorly transcribed (http://signal.salk.edu/) and contributes little to the SSP profile. In accordance, MS analysis revealed that the relative contributions of cruciferin isoforms to total seed protein are CRUC > CRUA > CRUB with the cruciferin produced by At1g03890 present at much lower levels.[4] In *B. napus*, five cruciferin protomers have been reported: CRUA (gene *CRUA* syn *CRU2/3*), CRU4 (gene *CRU4*), CRU3 (gene *CRU1*), CRU1 (gene *BnC1*), and CRU2 (gene *BnC2*) (http://www.uniprot.org, accessed in July 2012). The *Arabidopsis* CRUA, CRUB, and CRUC subunits show high degrees of identity (50−84%) and similarity (67−91%) with *B. napus* subunits CRUA (CRU2/3), CRU4, and CRU1 (Supporting Information, Table S2), indicating that it is reasonable to extrapolate information pertaining to the properties of *Arabidopsis* cruciferin protomers to those from *B. napus*. The primary sequences of CRUA, CRUB, and CRUC protomers show distinct differences in their hypervariable regions (HVR) but have a highly conserved core structure.[10] For example, the HVR-1 region of CRUC protomer contains 61 residues with at least 8 Gln and Gly repeats, whereas those of CRUA and CRUB protomers contain only 9 and 7 residues with 1 and 2 repeats, respectively.[10] Comparative in silico modeling indicated that identical cruciferin subunits could form stable hexameric structures.[10]

Genetic combination of *CRU* T-DNA insertion mutants yielded double-knockout lines containing a single functional *CRU* gene and a triple-knockout line in which all of the *CRU* genes were inactivated (Figure 1B,C). SDS-PAGE analysis showed bands corresponding to the α- and β-chains of the CRUA, CRUB, and CRUC protomers in the seed extracts (Figure 1C). The WT profile had three protein bands between 27 and 34 kDa corresponding to the α-chains; the 34 kDa band was more intense and possibly two proteins of similar molecular weight, and two bands at approximately 20 kDa, corresponding to the region where the β-chains migrate. The double-knockout line expressing only *CRUC* produced a predominant 34 kDa and a minor 32 kDa α-chain. The line expressing only the *CRUA* gene produced two α-chains of equal abundance, one at approximately 33.5 kDa and the other at 32 kDa. A single β-chain of 27 kDa was predominant for *CRUAbc* and *CRUabC*. The 33.5 kDa band of CRUA protomer and the 34 kDa band of CRUC protomer likely combine to form the band of greater intensity seen in the WT profile. The line expressing only the *CRUB* gene produced a single β-chain of approximately 27 kDa and a single α-chain of 31 kDa. The multiple β-chains seen in the *CRUAbc* and *CRUabC* genotypes likely derived from alternate proteolytic processing of the protein or post-transcriptional modifications, because previous MS analysis did not detect proteins derived from rare mRNA splice variants that appear in transcript analysis data.[4] The triple-knockout line had very weak and diffused protein bands of approximately 30 kDa; however, Western blot analysis with an antibody having affinity for cruciferin α-chain did not detect cruciferin (Figure 1C). All of the lines including the triple-knockout line showed a protein band at 21 kDa, which is different from cruciferin β-chain most likely originating from oleosins. Napin expression was not affected in the cruciferin altered lines, and the protein was separated into large (L) and small (S) chains upon S−S bond reduction (Figure 1C).

Five genes encode napin in *A. thaliana* cv. Columbia: a single gene at At5g54740 and four genes linked in tandem at At4g27140, At4g27150, At4g27160, and At4g27170. It is impractical to combine T-DNA knockouts in the four linked genes through crossing; therefore, RNAi was used to reduce napin gene expression. SDS-PAGE analysis of a line with an especially strong phenotype revealed that little or no napin was produced, whereas the profile of other seed proteins was unaffected (Figure 1C). This phenotype in the line has remained stable through at least six generations. Furthermore, in either the cruciferin and napin-deficient lines no detectable differences in seed size, seed germination, or seedling vigor were found compared to the WT (data not shown).

IR Evidence for Biopolymer Localization in Parenchyma Cells. High-resolution synchrotron FT-IR techniques have been used for subcellular chemical mapping of single cells, such as hybridoma B cells in mice,[23] metal-cyanobacteria,[24] and oral and lung epithelial cells.[25,26] Scanning of seed tissue samples (~6 μm thickness) in the mid-IR spectral range from 4000 (λ = 2.5 μm) to 800 cm^{-1} (λ = 12.5 μm) under transmittance mode allowed information on cellular microstructure to be obtained that was converted into details of the structural properties of the macromolecules within. This was

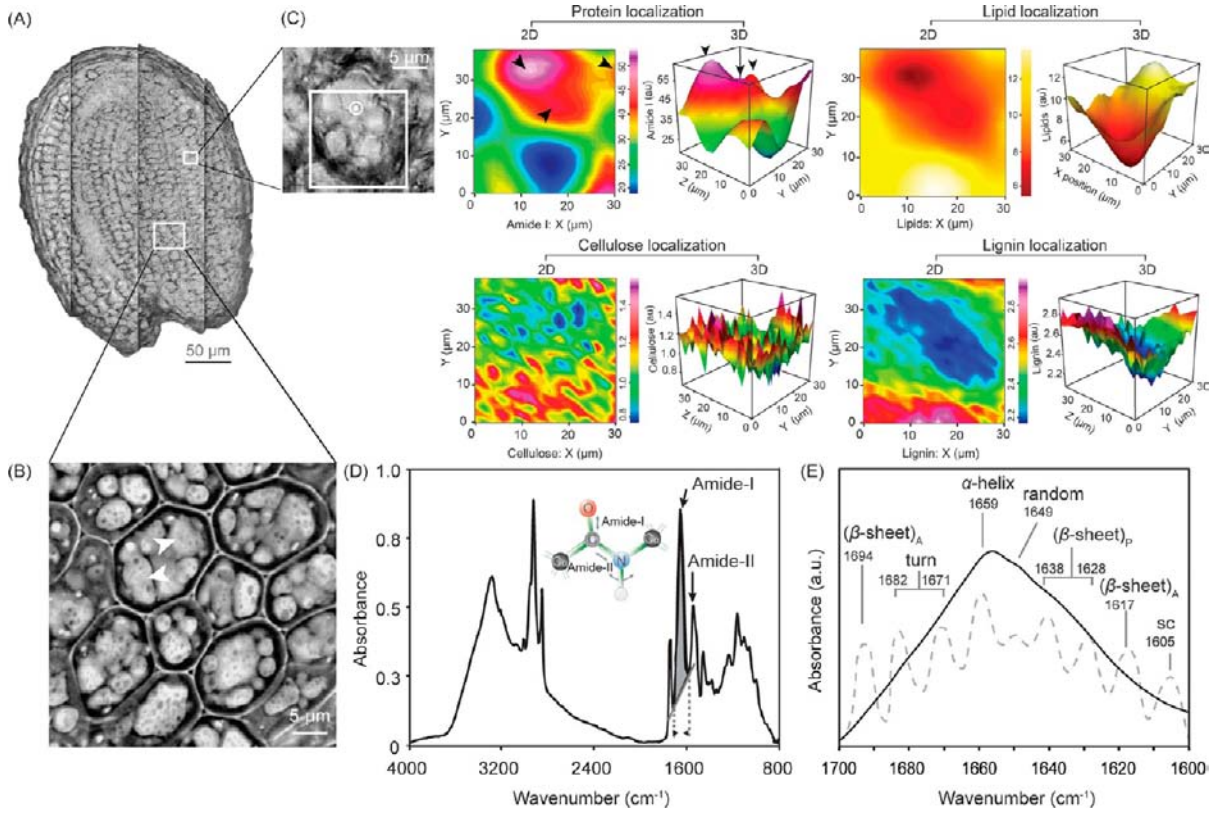

Figure 2. Selection of protein-rich areas in *Arabidopsis* seed tissues and analysis of protein secondary structure components. (A) Visible microscopic image of *Arabidopsis* WT seed. (B) Close-up view of cotyledon parenchyma cells with arrowheads showing the protein storage vacuoles (PSV). (C) 2D and 3D chemical images of protein (amide I), lipid (νC═O of lipid ester), cellulose (1059 cm^{-1}), and lignin (1513 cm^{-1}). Arrowheads show protein-rich area on the protein chemical image. Chemical intensity bars indicate color and scale assigned for IR absorbance; higher value indicates higher biopolymer concentration. (D) FTIR spectra of the white-circled point on the visible image (C) of PSV. The inset shows vibrational modes of amide I and amide II backbone. (E) Fourier self-deconvoluted peaks (dashed gray line) of the amide I band. Labels: $(\beta\text{-sheet})_A$, antiparallel β-sheet; $(\beta\text{-sheet})_p$, parallel β-sheet; SC, side chain.

the first reported use of synchrotron FT-IR analysis of a single cell within an *Arabidopsis* seed. The confocal microscope coupled to the synchrotron IR light source (flux, 1×10^{14} at 10 μm, v/s/0.1%BW at 100 mA) at the CLS is equipped with Schwarzschild objectives with numerical apertures (NA) of 0.4 for 15× and 0.52 for 36× magnifications. The achievable spot size or diffraction-limited spatial resolution for this type of system is 2−10 μm (0.5−1.2λ).[27−29] The single-cell analysis was carried out using an aperture size of 5 × 5 μm with a step size of 1−2 μm in the transmittance mode and scanned with little radiation diffraction; therefore, loss of S/N in the mid-IR range was minimal. Furthermore, the average size of a parenchyma cell in an *Arabidopsis* embryo is approximately 15−20 μm in diameter (Figure 2), whereas the average diameter of PSV is 6.2−8.2 μm,[30] which is much larger than the diffraction-limited spot size used in the mid-IR region.

Protein, oil, cellulose, and lignin distribution was determined in embryonic parenchyma cells of WT *Arabidopsis* (Figure 2). The IR vibrations of N−H in-plane bending (δN−H) and C−N stretching (νC−N) of amide I and II (Figure 2) are highly sensitive to secondary structure folding upon chain packing in the crystalline region.[31] The intensity of the amide II band resulted in these cells being not as strong as the amide I band and overlapped with an absorbance band originating from the νC═C of the lignin aromatic ring (1516 cm^{-1} shoulder at the left side of amide II band, Figure 2D) as reported in other plant tissues.[32] Considering the interference from polysaccharides, oil, and lignin, amide I is the sole absorption band from the protein backbone that can be used to probe protein distribution and secondary structure features. According to the chemical images, more protein accumulated in the cell center (Figure 2C). The lipid distribution map was generated using the band at 1740 cm^{-1} that is due to stretching of carbonyl groups in lipid esters. Lipids accumulated in the cell periphery (Figure 2C). This distribution pattern of protein in the middle and lipid bodies at the edge of the cell has been reported for *Arabidopsis* and other oilseeds.[33] The two main cell wall biopolymers, cellulose and lignin, roughly outlined the cell wall (Figure 2C). The protein-rich areas include protein storage vacuoles and provide less scattering interference from carbohydrates and oil bodies in the synchrotron FT-IR spectra. Visible images of PSVs in the cells showed that they are globular in shape, and the 2D and 3D chemical images showed protein-rich regions that may contain relatively pure protein (Figure 2) and can be used for exploring details of protein structure in its natural setting.

Fourier self-deconvolution (FSD)[34] permits separation of individual component peaks of helical, β-sheet, β-turn, and random structures that overlap within the broad amide I region (Figure 2E). Deconvolution of the amide I region of WT protein resolved nine individual bands;: α-helix at 1659.7 ± 0.7 cm^{-1}; β-sheet at 1616.8 ± 0.8, 1627.8 ± 1.3, 1638.4 ± 1.7, and 1694.1 ± 0.2 cm^{-1}; β-turn at 1671.0 ± 0.3 and 1682.3 ± 0.3 cm^{-1}; and random structure at 1649.1 ± 1.5 cm^{-1}, which were

dx.doi.org/10.1021/jf304328n | J. Agric. Food Chem. 2013, 61, 901−912

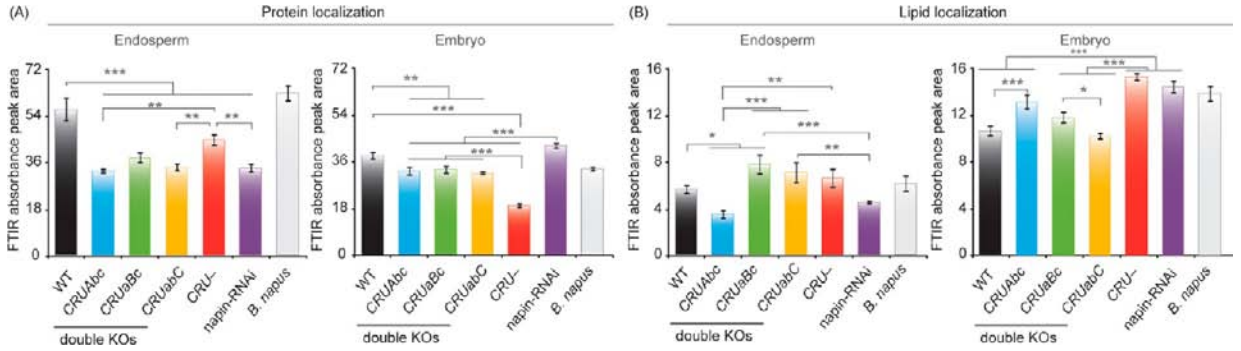

Figure 3. Protein and lipid distribution in *Arabidopsis* and *B. napus* seeds. (A) Protein localization in the endosperm and embryo as calculated by integrating the amide I band at peak 1650 cm^{-1} (integral baseline 1720−1588 cm^{-1}). (B) Lipid localization in the endosperm and embryo according to the area of lipid ester peak at 1747 cm^{-1} (integral baseline 1786−1720 cm^{-1}). (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$ (ANOVA). Error bars, ± SEM ($n = 10$).

Table 2. Secondary Structure Features (Percent) of Embryo Seed Storage Protein of *Arabidopsis* and *B. napus* Seeds Assessed Using Synchrotron FT-IR[a]

line[b]	α-helix	β-sheet	turn	random
Arabidopsis				
WT	15.4 ± 1.1 c	45.3 ± 2.2 a	25.9 ± 1.0 d	13.5 ± 1.0 c
double knockouts				
CRUAbc	15.8 ± 1.1 c	44.9 ± 1.7 a	26.0 ± 1.3 d	13.3 ± 0.6 c
CRUaBc	18.1 ± 0.9 b	40.7 ± 1.3 c	27.7 ± 0.6 ab	13.6 ± 0.3 c
CRUabC	15.6 ± 1.1 c	42.2 ± 1.2 b	26.5 ± 1.1 cd	15.7 ± 0.9 a
triple knockout				
CRU-	24.0 ± 1.5 a	34.4 ± 1.0 d	26.6 ± 1.1 bc	14.9 ± 0.7 b
napin-RNAi	13.4 ± 1.4 d	44.1 ± 2.3 a	28.2 ± 1.4 a	14.3 ± 2.3 bc
B. napus	14.8 ± 1.2	45.0 ± 1.4	27.1 ± 1.6	13.2 ± 0.8
SE	0.41	0.47	0.16	0.17

[a]Means (± SD) of triplicate analyses followed by the same letter within a column do not differ significantly ($P < 0.05$). [b]WT, wild type; double-knockout lines, *CRUAbc*, *CRUaBc*, and *CRUabC*; SE, standard error of the mean. *B. napus* data are presented for comparison.

assigned to secondary structure features according to published references.[15,16]

Similar to the WT line, chemical imaging of a small area (900 μm^2) close to the size of a parenchyma cell in the embryo of double-knockout and napin-RNAi lines showed a high concentration of protein (integral peak area of amide I band at 1650 cm^{-1}) in the mid cell area and a high concentration of lipid (integral peak area of $\nu C{=}O$ lipid ester band at 1740 cm^{-1}) near the cell periphery (Supporting Information, Figure S1). Conversely, the distribution pattern of oil and protein in the cruciferin-depleted triple-knockout line indicated that the PSVs were much smaller and contained very little protein and that the remaining seed protein was more evenly distributed throughout the cytoplasm (Supporting Information, Figure S1). The observations indicated that disruption of any two cruciferin genes or suppression of napin production did not result in an aberrant PSV phenotype. However, complete elimination of cruciferin production caused phenotypic differences in the size and distribution of PSVs. In the cells of *B. napus* seed, both protein and lipids were concentrated in the middle of the cell. This study clearly demonstrates the capability of using synchrotron FT-IR for probing the distribution and abundance of macromolecules in plant cells.

Chemical Makeup of Seeds. Crucifer seeds contain primarily lipid, protein, and carbohydrates other than starch; however, only the endosperm and embryo store lipid and protein. Signal intensities for lipid and protein in the endosperm and embryo of the SSP-altered lines were calculated using FT-IR spectral data (Figure 3). According to this estimation, in WT seed, the protein concentration was higher ($P < 0.05$) in the endosperm (55%) than in the embryo (36%), whereas the lipid content was greater in the embryo (10%) than in the endosperm (6%). In the SSP-altered lines, the protein content of the endosperm (Figure 3A) was lower ($P < 0.05$) (32−43%) than the WT (55%). Similarly, the embryo tissues of the double-knockout lines also showed lower ($P < 0.05$) protein content (31−32%) than that of WT seed (37%). As expected, disruption of all three cruciferin genes resulted in markedly reduced (~51% less than WT) SSP accumulation in the embryo, but the accumulated protein content in the endosperm was higher (~40%) than that of double-knockout lines (Figure 3A).

When whole seed protein levels were calculated on the basis of total nitrogen content (Dumas combustion method), the cruciferin altered lines were between 25.7 and 30.6% (Supporting Information, Table S3). The *CRUAbc* and *CRUaBc* lines had significantly higher ($P < 0.05$) protein values than WT, whereas the protein content of the *CRUabC* line was comparable ($P > 0.05$) to that of the WT. The triple-knockout line had a protein content of 25.7%, which was 1.4 percentage points lower ($P < 0.05$) than that of the WT type (27.1%) (Supporting Information, Table S3). However, the protein content calculated using the peak area of the amide I band indicated that the triple-knockout line had at least 50.0% less protein content in the embryo and 20% less protein in the endosperm cells compared to the WT (Figure 3). The

dx.doi.org/10.1021/jf304328n | J. Agric. Food Chem. 2013, 61, 901−912

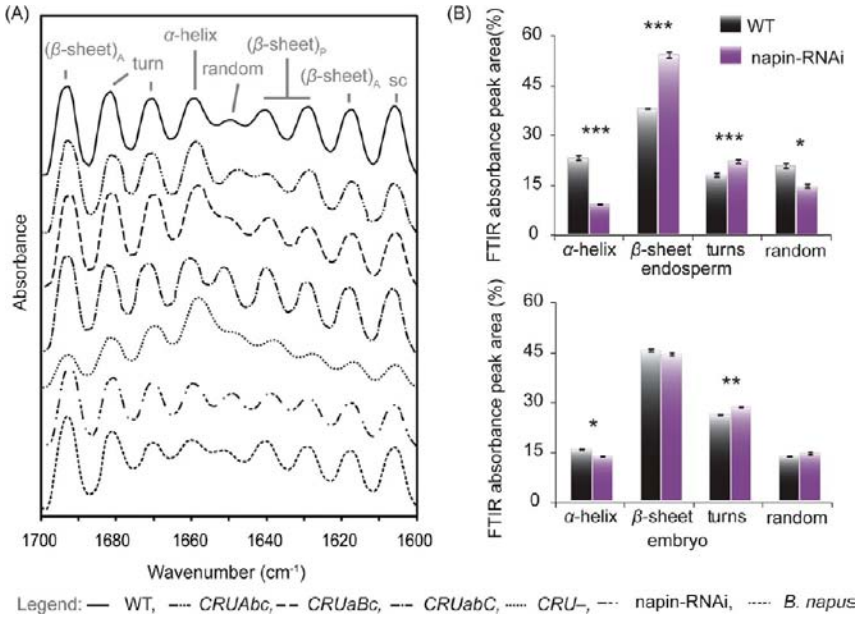

Figure 4. Analysis of deconvoluted synchrotron FT-IR spectra. (A) Identification of secondary structure elements in the spectra of *Arabidopsis* and *B. napus* endosperm tissues. Peak labeling is similar to Figure 2. (B) Comparison of content of secondary structure elements of embryo tissue proteins of WT and napin-RNAi *Arabidopsis* lines. (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$ (ANOVA). Error bars, ± SEM ($n = 10$).

calculation of protein content using the synchrotron FT-IR amide I peak is highly selective for protein molecules because the characteristic in-plane vibrational mode of the peptide chain (80% νC═O, 10% νC—N) generates the particular IR signal.[31,35] The total N-based protein estimation includes protein/peptides, free amino acids, DNA, RNA, and some phospholipids that contain nitrogen, but not necessarily the peptide amide bond. Therefore, the slight change in the total N-based protein value of the cruciferin triple-knockout line compared to the WT indicates the possible accumulation of free amino acids or incorrectly assembled proteins. A similar observation was made in SSP-deficient lines of soybean[36] and common bean,[37] in which rebalancing of proteome via accumulation of free amino acids or sulfur-rich proteins has occurred.

Interestingly, synchrotron FT-IR in situ localization revealed that suppression of napin gene expression (napin-RNAi) did not reduce the total protein level in the embryo, but did so in the endosperm (Figure 3A). Examination of an *Arabidopsis* line expressing an open reading frame encoding a napin-GFP fusion protein driven by a napin promoter revealed that napin accumulation is considerably higher in the endosperm compared to the embryo (data not shown). This preferential accumulation of napin would explain why protein content was reduced only in the endosperm, and not the embryo, in the napin-suppressed line. Seed protein levels based on total N of the napin-RNAi line remained fairly close to the value of the WT (Supporting Information, Table S3).

Protein Secondary Structure Features. The effect of expressed gene composition on the structure of accumulated protein was evaluated by deriving secondary structure features using Fourier self-deconvolution (FSD) and Gaussian curve fitting of amide I bands of the respective synchrotron FT-IR spectra (Table 2 and Figure 4). In the double-knockout lines, genetic alteration of cruciferin expression did not affect napin expression (Figure 1C); therefore, it is expected that the secondary structure feature differences observed for these lines in the synchrotron FT-IR study are due to the changes in cruciferin subunit composition. The proteins of the WT line showed a high β-sheet content (45.3%) compared to α-helix (15.4%), turns (25.9%), or random (13.5%) structures (Table 2), indicating the predominance of cruciferin with its β-barrel structure typical of the cupin superfamily.[2] Comparison of double-knockout lines with the WT line showed that proteins in the *CRUAbc* line had similar levels ($P > 0.05$) of secondary structure components (Table 2); however, proteins in *CRUaBc* and *CRUabC* lines showed some differences, in particular, a low β-sheet content ($P < 0.05$). The level of α-helix and turns ($P < 0.05$) in protein from the *CRUaBc* line was higher than that in the WT. These observed differences in α-helix and β-sheet contents may be related to the level of cruciferin and napin present in the double-knockout lines rather than the structural changes of each protein. However, the homohexameric protein from the *CRUabC* line exhibited significantly ($P < 0.05$) higher value for random structures than protein from WT, *CRUAbc*, or *CRUaBc* lines. This increased random structure in protein in the *CRUabC* line may be related to the disordered regions arising from its especially long HVR-I region.[10] The most notable change was observed for the triple-knockout line, with high ($P < 0.05$) α-helix (24.0%) and low β-sheet (34.4%) contents compared to cruciferin-expressed lines, indicating that the remaining proteins deviate from the native cruciferin structure. Furthermore, the slightly higher contents of turns (26.6%) and random coils (14.9%) of this line indicated the less ordered structure of the contained proteins.

In the napin-RNAi line, embryo proteins showed the lowest α-helix content (13.4%) (Table 2) and indicated lost contribution from napin to the helical components. Besides that, the endosperm of this line had a lower (7.8%) ($P < 0.05$) level of α-helix content than the WT (Figure 4B), indicating that the endosperm and the adjacent cell layer are lower in protein with helical secondary structure than the WT or others. The increased level of β-sheet content (50 vs 34.4% in WT; Figure 4B) of endosperm proteins indicated the reduction of

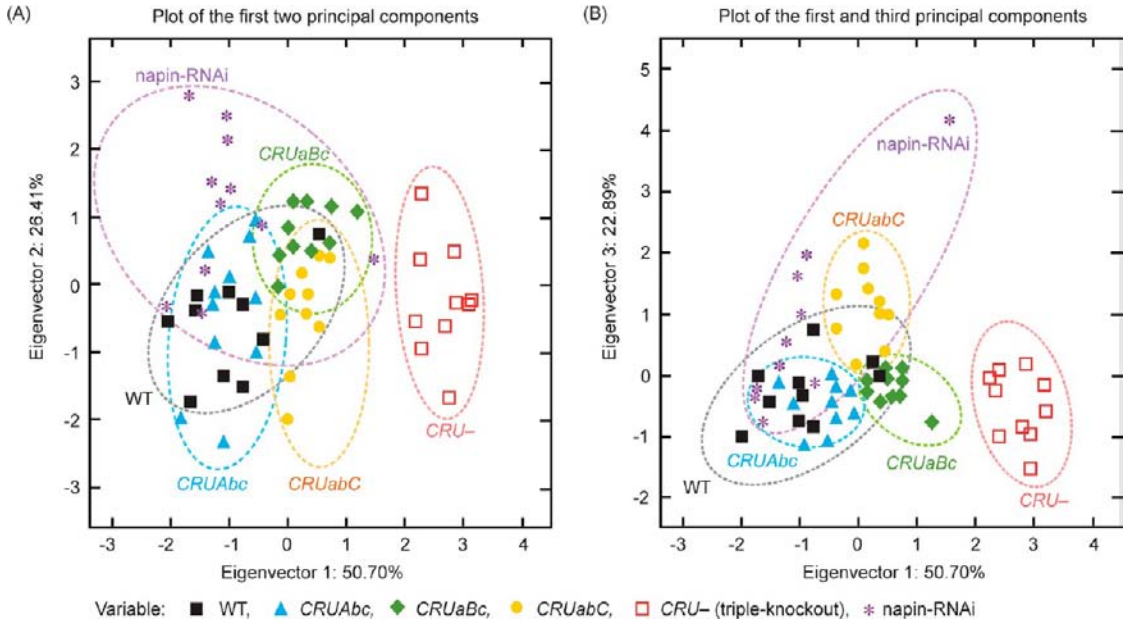

Figure 5. PCA of the deconvoluted amide I band of WT and SSP-altered *Arabidopsis*: biplot of the first two principal components (A) and first and third principal components (B). The first, second, and third principal components explain 50.70, 26.41, and 22.89%, respectively, of total variance in the protein secondary structure elements.

predominantly helical napin in the accumulated protein. Immunolocalization of 2S albumin in developing *Arabidopsis* seed showed that napin is localized in endothelial cells in the inner integument layer in direct contact with the endosperm.[38] Proteins of embryo tissues of the napin-RNAi line had more or less similar secondary structure features as WT.

PCA of the values of deconvoluted amide I band also differentiated the aberrant phenotypes (Figure 5). PCA showed clear separation of the *CRU* triple-knockout line from the WT and other mutant lines based on FSD protein secondary structure features (Figure 5). The first two components accounted for 79.13% of the total variance, but separation of the double-knockout lines was not possible (Figure 5A). The first and third principal components accounted for 51.85 and 20.87%, respectively, of the variability and distinguished the double-knockout lines expressing only *CRUA* gene from those expressing the *CRUB* or *CRUC* gene (Figure 5B). The secondary structure differences between the WT and the double-knockout lines or the napin-RNAi line proteins did not separate well in PCA (Figure 5); however, some degree of clustering was observed due to subtle differences in the dominant secondary structure features of the homohexameric cruciferin assemblies.

According to the results of protein distribution and protein secondary structure features of seed storage protein altered seeds, protein accumulation in PSV of double-knockout lines was not much different than in the WT. The differences in secondary structure features indicated the cruciferin core structure is primarily retained in the proteins of these lines. Synchrotron FT-IR revealed subtle differences of accumulated proteins in these lines, showing the usability of this technique in assessing protein modifications without removing these molecules from tissues.

Pepsin-Assisted Degradation of Tissue Proteins. Bioaccessibility of seed tissues and their component molecules is important to release nutrients from seed matrix. The in vitro digestibility of intact starch (e.g., corn, wheat, and potato), lipids, and proteins (almond and peanut) within seeds and cells has been examined using light, scanning electron, and transmission electron microscopy.[39−42] In the case of proteins, release of amino acids and peptides to make them available for absorption in the gastrointestinal tract is a measure of bioavailability or digestibility. Pepsin is the first proteolytic enzyme that starts degrading proteins in a monogastric digestive system; therefore, assessment of pepsin accessibility relates to digestibility and the potential to release small peptides and amino acids of a particular protein-containing substrate. Disruption of enzyme susceptible peptide bonds leads to structural changes of protein and provides information on protein bioaccessibility, which is linked to nutritional quality.[5,6] However, obtaining details about the extent of degradation and quantifying secondary structure are highly unlikely using the above-mentioned microscopic techniques.

The protein chemical maps of untreated and pepsin-treated WT and SSP-altered seed tissues were constructed by integrating the amide I band and compared (Figures 6 and 7A). Proteins in the WT, *CRUAbc*, *CRUaBc*, and *CRU-* lines displayed a significant ($P < 0.05$) reduction in the intensity of integrated amide band, whereas the *CRUabC* seed proteins did not show such a reduction in their pepsin-treated tissues (Figures 6 and 7A). The greatest depletion of SSP was in the napin-RNAi line (Figure 7A). Reduction of protein content in the pepsin-treated tissues indicates the propensity of protein to degradation, because the FT-IR amide I band is related to the intact peptide bond. Although an overall reduction of total protein content was observed due to hydrolysis, details of secondary structure features indicated that the IR absorption peaks of α-helix and β-sheet of FSD spectra responded differently depending on the genetic composition (Figure 7B; Table 3). The β-sheet component of proteins of WT, triple-knockout, and napin-RNAi lines was reduced in pepsin-treated samples, whereas random structure content increased in all samples except proteins of the *CRUabC* line (compare Tables 2 and 3). When cruciferin and napin structures are concerned,

dx.doi.org/10.1021/jf304328n | J. Agric. Food Chem. 2013, 61, 901−912

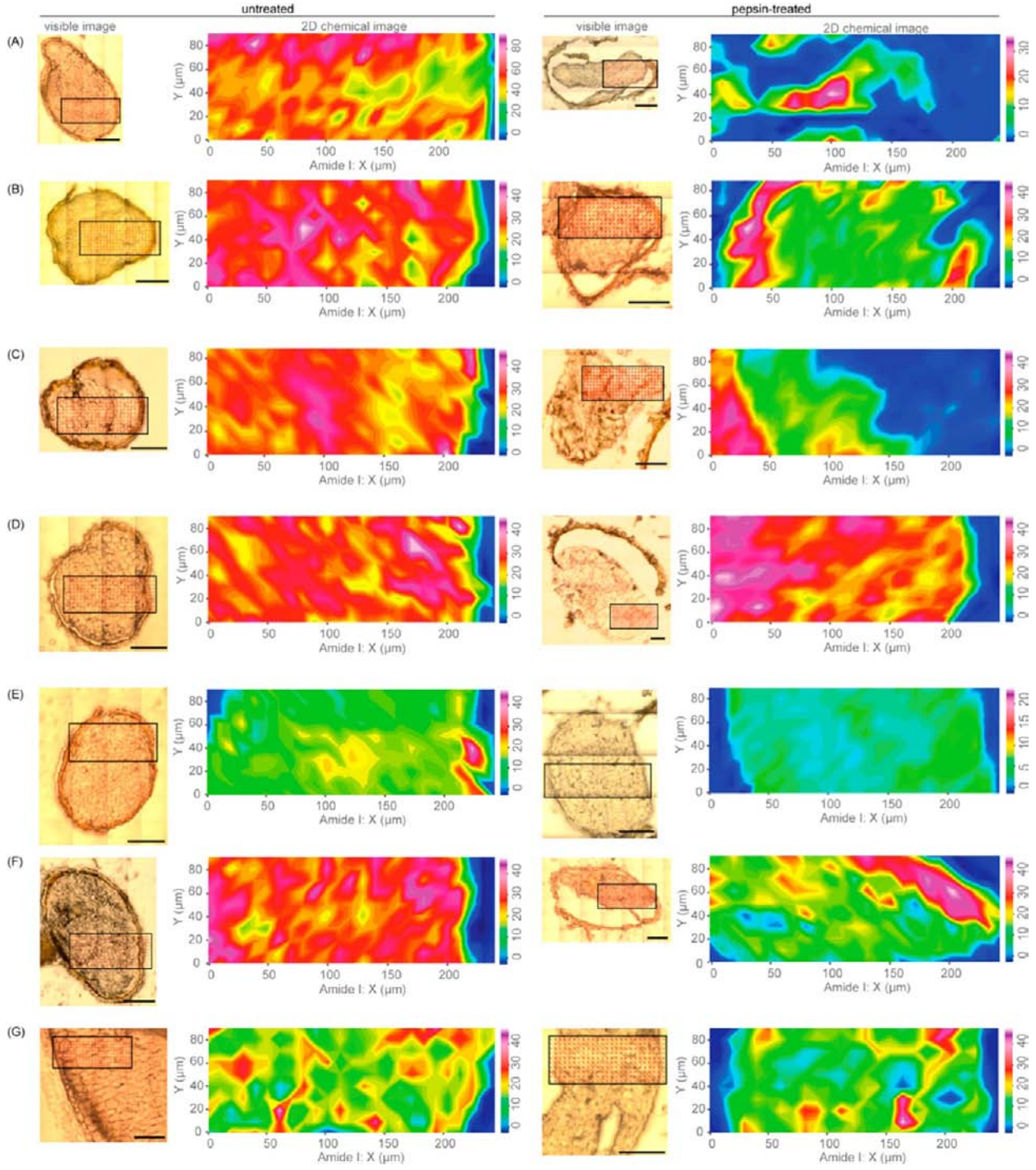

Figure 6. Visible microscopic images (scale bar = 100 μm) and synchrotron FT-IR 2D chemical images of untreated and pepsin-treated seeds of WT (A), *CRUAbc* (B), *CRUaBc* (C), *CRUabC* (D), *CRU-* (triple-knockout) (E), napin-RNAi (F), and *B. napus* (G) seeds. Integration of peak 1650 cm^{-1} (amide I) was used to obtain the chemical images. Chemical intensity bars indicate color and scale assigned for IR absorbance; higher value indicates higher biopolymer concentration.

the α-helix organization is more stable to hydrolysis than β-sheet structure,[43] especially when S−S bonds are involved in stabilizing protein structure such as in napin.[44] The napin-RNAi line does not contain napin and showed the greatest reduction of α-helix and β-sheet contents and the highest increase in random structure content of FSD spectra among the lines showing degradation of cruciferin present in the PSV. The presence of napin in the double-knockout lines and the stability of this protein to pepsin-catalyzed hydrolysis may have contributed to the minor changes observed in α-helix content upon treatment. Therefore, only the changes in β-sheet content may indicate the propensity of respective cruciferin homohexamer to pepsin-catalyzed degradation. Among the cruciferin altered lines, *CRUabC* showed the least changes to the protein secondary structure components due to pepsin treatment.

The location of cleavage sites in a solvent-accessible area of the molecule may provide reasonably good digestibility.[10] Among the double-knockout lines, the low protein degrad-

dx.doi.org/10.1021/jf304328n | *J. Agric. Food Chem.* 2013, 61, 901−912

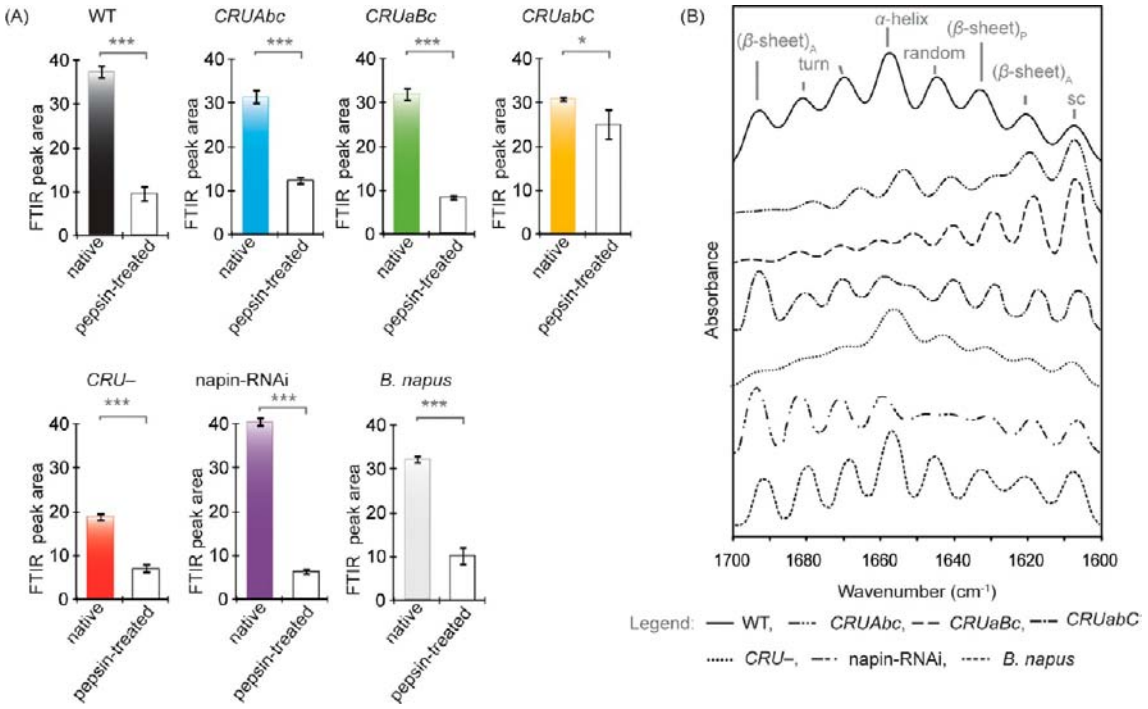

Figure 7. Effect of pepsin treatment on *Arabidopsis* and *B. napus* SSP. (A) Comparison of protein concentration calculated by integrating the amide I band. (∗) $P < 0.05$, (∗∗) $P < 0.01$, (∗∗∗) $P < 0.001$ (ANOVA). Error bars, ± SEM. ($n = 10$). (B) Deconvoluted synchrotron FT-IR spectra showing secondary structure elements of pepsin-treated *Arabidopsis* and *B. napus* seeds. Peak identification and spectra of untreated samples are similar to Figure 2.

Table 3. Protein Structure Features (Percent) of Pepsin-Treated Arabidopsis and *B. napus* Seeds[a]

line[b]	α-helix	β-sheet	turn	random
Arabidopsis				
WT	16.8 ± 3.6 b	34.2 ± 3.3 d	25.4 ± 1.5 b	23.6 ± 1.0 a
double knockouts				
CRUAbc	15.4 ± 1.5 b	47.5 ± 3.0 a	18.3 ± 2.8 d	18.8 ± 2.1 b
CRUaBc	15.7 ± 1.3 b	44.2 ± 3.1 b	21.8 ± 3.4 c	18.8 ± 5.1 b
CRUabC	15.6 ± 0.8 b	45.0 ± 1.8 ab	24.9 ± 2.1 b	14.2 ± 1.4 c
triple knockout				
CRU-	30.2 ± 3.6 a	24.5 ± 3.5 e	26.8 ± 4.1 b	18.5 ± 3.5 b
napin-RNAi	8.1 ± 0.8 c	38.1 ± 1.7 c	30.2 ± 1.6 a	23.6 ± 1.3 a
B. napus	17.4 ± 3.1	42.5 ± 6.7	24.4 ± 4.0	15.7 ± 2.2
SE	0.89	1.29	0.57	0.56

[a]Means (± SD) of triplicate analyses followed by the same letter within a column do not differ significantly ($P < 0.05$). [b]WT, wild type; double-knockout lines *CRUAbc*, *CRUaBc*, and *CRUabC*; SE, standard error of the mean. *B. napus* data are presented for comparison.

ability of the *CRUabC* line may be related to the structural features of dominant homohexameric CRUC. According to homology modeling and proteolytic site prediction, CRUA and CRUB cruciferin homotrimers have pepsin cleavage sites in both HVR-I and HVR-II, with HVR-I directed toward the trimer periphery and HVR-II in the trimer center and on the exposed trimer IE face. Although the HVR-I of the CRUC hexamer structure is long and projects toward the solvent, the absence of pepsin cleavage sites could lead to slow hydrolysis. However, dissociation of the hexamer into trimers may allow solvent-accessible pepsin-labile sites of HVR-II of CRUC.[10] In the tissue intact PSVs, dissociation of cruciferin hexamers into trimers is less likely, although a pH as low as 2 was provided to facilitate pepsin-catalyzed hydrolysis. These factors may have contributed to the low degradation of the CRUC homohexamer containing proteins during pepsin treatment.

Even though the protein secondary structure content of *B. napus* was similar to that of *Arabidopsis* (Table 2), the former was more resistant to pepsin-assisted degradation (Table 3). Less ordered protein structure was expected in the triple-knockout line and possibly contributed to easy degradation of the containing proteins by pepsin.

When all data are considered, pepsin-treated tissues of WT, *CRUAbc*, *CRUaBc*, and triple-knockout (*CRU-*) lines showed that the first three PCs of FSD secondary structure data differentiated the nontreated counterparts (Supporting Information, Figure S2). Secondary structure assessment of pepsin-treated tissues may be able to predict the propensity of the containing proteins to pepsin-assisted degradation.

Summary. The synchrotron FT-IR technique is very useful in mapping biopolymer distribution within a seed cell. This technique also allows investigation of the secondary structure features of protein directly within plant tissues without

extraction. When proteins of different secondary structure features are present, changes in their composition can be probed using this technique as evidenced by cruciferin double-knockouts and napin-RNAi lines. A single cell can be used as a unit of identifying changes in its protein composition using synchrotron FT-IR; this is especially valuable in evaluating protein modification due to changes in genetic composition. Assessment of secondary structure features can be linked to the accessibility of proteins to hydrolytic enzymes such as pepsin, which are important in assessing nutritional quality of proteins.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Figures S1 and S2 and Tables S1, S2, and S3. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Postal address: Agriculture and Agri-Food Canada, Saskatoon, SK S7N 0X2, Canada. Phone: (306) 956-7684. Fax: (306) 956-7247. E-mail: janitha.wanasundara@agr.gc.ca.

Funding

SaskCanola provided a Dr. S. Roger Rimmer Graduate Scholarship to T.S.W.-G., and this work was funded by the AAFC Brassica Protein Production Platforms (RBPI 26) project.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge Tor Pedersen at Canadian Light Source (CLS) Saskatoon, Canada, for his technical support with the Synchrotron mid-IR beamline and Helen Lui and Brad Hope of AAFC (Saskatoon) for assistance in generating plants.

■ REFERENCES

(1) Wanasundara, J. P.D. Proteins of Brassicaceae oilseeds and their potential as a plant protein source. *Crit. Rev. Food Sci. Nutr.* **2011**, *51*, 635−677.

(2) Tandang-Silvas, M. R. G.; Fukuda, T.; Fukuda, C.; Prak, K.; Cabanos, C.; Kimura, A.; Itoh, T.; Mikami, B.; Utsumi, S.; Maruyama, N. Conservation and divergence on plant seed 11S globulins based on crystal structures. *Biochim. Biophys. Acta* **2010**, *1804*, 1432−1442.

(3) Schmidt, I.; Renard, D.; Rondeau, D.; Richomme, P.; Popineau, Y.; Axelos, M. A. V. Detailed physicochemical characterization of the 2S storage protein from rape (*Brasica napus* L.). *J. Agric. Food Chem.* **2004**, *56*, 5995−6001.

(4) Wan, L.; Ross, A. R. S.; Yang, J.; Hegedus, D. D.; Kermode, A. R. Phosphorylation of the 12 S globulin cruciferin in wild-type and abi1-1 mutant *Arabidopsis thaliana* (thale cress) seeds. *Biochem. J.* **2007**, *404*, 247−256.

(5) Yu, P. Applications of hierarchical cluster analysis (CLA) and principle component analysis (PCA) in feed structure and feed molecular chemistry research, using synchrotron-based Fourier transformed infrared (FTIR) microspectroscopy. *J. Agric. Food Chem.* **2005**, *53*, 7115−7127.

(6) Yu, P.; Doiron, K.; Liu, D. Shining lights on the differences in molecular structural chemical makeup and the cause of distinct degradation behavior between malting- and feed-type barley using synchrotron FTIR microspectroscopy: a novel approach. *J. Agric. Food Chem.* **2008**, *56*, 3417−3426.

(7) Todd, A. E.; Orengo, C. A.; Thornton, J. M. Evolution of function in protein superfamilies, from structural perspective. *J. Mol. Biol.* **2007**, *307*, 1113−1143.

(8) Tandang, M. G.; Adachi, M.; Inui, N.; Maruyama, N.; Utsumi, S. Effects of protein engineering of canola procruciferin on its physicochemical and functional properties. *J. Agric. Food Chem.* **2004**, *52*, 6810−6817.

(9) Tandang, M. G.; Atsuta, N.; Maruyama, N.; Adachi, M.; Utsumi, S. Evaluation of the solubility and emulsifying property of soybean proglycinin and rapeseed procruciferin in relation to structure modified by protein engineering. *J. Agric. Food Chem.* **2005**, *53*, 8736−8773.

(10) Withana-Gamage, T. S.; Hegedus, D. D.; Qiu, X.; Wanasundara, J. P. D. In silico homology modeling to predict functional properties of cruciferin. *J. Agric. Food Chem.* **2011**, *59*, 12925−12938.

(11) Ma, C.-Y.; Rout, M. K.; Chan, W.-M.; Phillips, D. L. Raman spectroscopic study of oat globulin conformation. *J. Agric. Food Chem.* **2000**, *48*, 1542−1547.

(12) Schrader, B.; Hoffman, A.; Simon, A.; Sawatzki, J. Can a Raman renaissance be expected via the near-infrared Fourier transform technique? *Vib. Spectrosc.* **1991**, *1*, 239−250.

(13) Lonien, J.; Schwender, J. Analysis of metabolic flux phenotypes for two *Arabidopsis* mutants with severe impairment in seed storage lipid synthesis. *Plant Physiol.* **2009**, *151*, 1617−1634.

(14) AOCS Official Method Am 2-93, Determination of oil content in oilseeds. In *Official Methods and Recommended Practices of the AOCS*, 5th ed.; Firestone, D. E., Ed.; AOCS Press: Champaign, IL, 1997.

(15) Byler, D. M.; Susi, H. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers* **1986**, *25*, 469−487.

(16) Dong, A.; Huang, P.; Caughey, W. S. Protein secondary structures in water from second-derivative amide I infrared spectra. *Biochemistry* **1990**, *29*, 3303−3308.

(17) SAS Institute Inc. *SAS User'S Guide: Statistical Analysis System*; SAS Institute: Cary, NC, 2004.

(18) Inquello, V.; Raymond, J.; Azanza, J. L. Disulfide interchange reactions in 11S globulin subunits of cruciferae seeds. Relationships to gene families. *Eur. J. Biochem.* **1993**, *217*, 891−895.

(19) Pang, P. P.; Pruitt, R. E.; Meyerowitz, E. M. Molecular cloning, genomic organization, expression and evolution of 12S seed storage genes of *Arabidopsis thaliana*. *Plant Mol. Biol.* **1988**, *11*, 805−820.

(20) Rödin, J.; Ericson, M. L.; Josefsson, L.-G.; Rask, L. Characterization of a cDNA clone encoding a *Brassica napus* 12S protein (cruciferin) subunit. *J. Biol. Chem.* **1990**, *265*, 2720−2723.

(21) Sjödahl, S.; Rödin, J.; Rask, L. Characterization of the 12S globulin complex of *Brassica napus*. Evolutionary relationship to other 11−12S storage globulins. *Eur. J. Biochem.* **1991**, *196*, 617−621.

(22) White, J. A.; Todd, T.; Newman, T.; Focks, N.; Girke, T.; de Ilarduya, O. M.; Jaworski, J. G.; Ohlrogge, J. B.; Benning, C. A new set of Arabidopsis expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil. *Plant Physiol.* **2000**, *124*, 1582−1594.

(23) Jamin, N.; Dumas, P.; Moncuit, J.; Fridman, W. H.; Teillaud, J. L.; Carr, G. L.; Williams, G. P. Highly resolved chemical imaging of living cells by using synchrotron infrared microspectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4837−4840.

(24) Yee, N.; Benning, L. G.; Phoenix, V. R.; Ferris, F. G. Characterization of metalcyanobacteria sorption reactions: a combined macroscopic and infrared spectroscopic investigation. *Environ. Sci. Technol.* **2004**, *38*, 775−782.

(25) Diem, M.; Romeo, M.; Matthaus, C.; Miljkovic, M.; Miller, L.; Lasch, P. Comparison of Fourier transform infrared (FTIR) spectra of individual cells acquired using synchrotron and conventional sources. *Infrared Phys. Technol.* **2004**, *45*, 331−338.

(26) Holman, H. N.; Martin, M. C.; Blakely, E. A.; Bjornstad, K.; McKinney, W. R. IR spectroscopic characteristics of cell cycle and cell death probed by synchrotron radiation based Fourier transform IR spectromicroscopy. *Biopolymers* **2000**, *57*, 329−335.

(27) Carr, G. L.; Reffner, J. A.; Williams, G. P. Performance of an infrared microspectrometer at the NSLS. *Rev. Sci. Instrum.* **1995**, *66*, 1490−1492.

(28) Carr, G. L. Resolution limits for infrared microspecroscopy explored with synchrotron radiation. *Rev. Sci. Instrum.* **2000**, *72*, 1613−1619.
(29) Holman, H. N.; Bechtel, H. A.; Hao, Z.; Martin, M. C. Synchrotron IR spectromicroscopy: chemistry of living cells. *Anal. Chem.* **2010**, *82*, 8757−8765.
(30) Shimada, T.; Fuji, K.; Tamura, K.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I. Vacuolar sorting receptor for seed storage proteins in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 16095−16100.
(31) Miyazawa, T.; Shimanouchi, T.; Mizushima, S. Characteristic infrared bands of monosubstituted amides. *J. Chem. Phys.* **1956**, *24*, 408−418.
(32) Suarez-Garcıa, F.; Martınez-Alonso, A.; Tascon, J. M. D. A comparative study of the thermal decomposition of apple pulp in the absence and presence of phosphoric acid. *Polym. Degrad. Stab.* **2002**, *75*, 375−383.
(33) Otegui, W. M. S.; Herder, R.; Schulze, J.; Jung, R.; Staehelind, L. A. The proteolytic processing of seed storage proteins in *Arabidopsis* embryo cells starts in the multivesicular bodies. *Plant Cell* **2006**, *18*, 2567−2581.
(34) Kauppinen, J. K.; Moffatt, D. J.; Mantsch, H. H.; Cameron, D. G. Fourier self-deconvolution: a method for resolving intrinsically overlapped bands. *Appl. Spectrosc.* **1981**, *35*, 271−276.
(35) Susi, H. Infrared spectroscopy − conformation. *Methods Enzymol.* **1972**, *26*, 455−472.
(36) Schmidt, M. A.; Herman, E. M. Proteome rebalancing in soybean seeds can be exploited to enhance foreign protein accumulation. *Plant Biotechnol. J.* **2008**, *6*, 832−842.
(37) Marsolais, F.; Pajak, A.; Yin, F.; Taylor, M.; Gabriel, M.; Merino, D. M.; Ma, V.; Kameka, A.; Vijayan, P.; Pham, H.; Huang, S. Z.; Rivoal, J.; Bett, K. E.; Hernández-Sebastià, C.; Liu, Q.; Bertrand, A.; Chapman, R. A. Proteomic analysis of common bean seed with storage protein deficiency reveals up-regulation of sulfur-rich proteins and starch and raffinose metabolic enzymes, and down-regulation of the secretory pathway. *J. Proteomics* **2010**, *73*, 1587−1600.
(38) Ondzighi, C. A.; Christopher, D. A.; Cho, E. J.; Chang, S.-C.; Staehelin, L. A. *Arabidopsis* protein disulfide isomerase-5 inhibits cysteine proteases during trafficking to vacuoles before programmed cell death of the endothelium in developing seeds. *Plant Cell* **2008**, *20*, 2205−2220.
(39) Chambers, S. J.; Wickham, M. S. J.; Bertelli, E.; Regoli, M.; Gunning, A. P.; Nicoletti, C. Rapid in vivo transport of proteins from digested allergen across pre-sensitised gut. *Biochem. Biophys. Res. Commun.* **2004**, *325*, 1258−1263.
(40) Ellis, P. R.; Kendall, C. W. C.; Ren, Y.; Parker, C.; Pacy, J. F.; Waldron, K. W.; Jenkins, D. J. A. Role of cell walls in the bioaccessibility of lipids in almond seeds. *Am. J. Clin. Nutr.* **2004**, *80*, 604−613.
(41) Lee, T.-T.; Huang, Y.-F.; Chiang, C.-C.; Chung, T.-K.; Chiou, P. W-S.; Yu, B. Starch characteristics and their influences on in vitro and pig prececal starch digestion. *J. Agric. Food Chem.* **2011**, *59*, 7353−7359.
(42) Mandalari, G.; Faulks, R. M.; Rich, G. T.; Turco, V. L.; Picout, D. R.; Lo Curto, R. B.; Bisignano, G.; Dugo, P.; Dugo, G.; Waldron, K. W.; Ellis, P. R.; Wickham, M. S. J. Release of protein, lipid, and vitamin E from almond seeds during digestion. *J. Agric. Food Chem.* **2008**, *56*, 3409−3416.
(43) Yu, P. Protein secondary structures (α-helix and β-sheet) at a cellular level and protein fractions in relation to rumen degradation behaviors of protein: a new approach. *Br. J. Nutr.* **2005**, *94*, 655−665.
(44) Koppelman, S. J.; Nieuwenhuizen, W. F.; Gaspari, M.; Knippels, L. M. J.; Penninks, A. H.; Knol, E. F.; Hefle, S. L.; de Jongh, H. H. J. Reversible denaturation of Brazil nut 2S albumin (Ber e 1) and implication of structural destabilization on digestion by pepsin. *J. Agric. Food Chem.* **2005**, *53*, 123−131.

NOTE ADDED AFTER ASAP PUBLICATION

This paper published January 18, 2013 with incorrect citations of Table 1. These citations were changed to Table 2, and the correct verion published January 22, 2013.