Food and Chemical Toxicology 130 (2019) 89–98

Contents lists available at ScienceDirect

Food and Chemical Toxicology

journal homepage: www.elsevier.com/locate/foodchemtox

Proteomics reveals the *in vitro* protein digestibility of seven transmembrane enzymes from the docosahexaenoic acid biosynthesis pathway

Michelle L. Colgrave[a,*], Keren Byrne[a], Joanne Caine[b], Lukasz Kowalczyk[b], Sapna Vibhakaran Pillai[c], Bei Dong[c], George Lovrecz[b], Susan MacIntosh[d], Judith A. Scoble[b], James R. Petrie[c], Surinder Singh[c], Xue-Rong Zhou[c]

[a] CSIRO Agriculture and Food, 306 Carmody Rd, St Lucia, QLD, 4067, Australia
[b] CSIRO Manufacturing, 343 Royal Parade, Parkville, VIC, 3052, Australia
[c] CSIRO Agriculture and Food, GPO Box 1700, Canberra, ACT, 2601, Australia
[d] Nuseed Americas, 11901 S. Austin Avenue, Alsip, IL, 60803, USA

ARTICLE INFO

Keywords:
Protein digestibility
Genetic modification
Omega-3 long-chain fatty acids
Docosahexaenoic acid (DHA)
Canola

ABSTRACT

The measurement of protein digestibility is one of the key steps in determining the safety of a genetically modified crop that has been traditionally accomplished using antibodies. Membrane proteins are often extremely difficult to express with replicated authentic tertiary structure in heterologous systems. As a result raising antibodies for use in safety assessment may not be feasible. In this study, LC-MS based proteomics was used to characterise seven transmembrane enzymes from the docosahexaenoic acid biosynthetic pathway that had been introduced into canola. The application of a two-stage digestion strategy involving simulated gastric fluid followed by trypsin enabled the measurement of protein digestibility *in vitro*. Tryptic peptide markers that spanned the length of each desaturase protein were monitored and showed that these proteins were readily degraded (> 95% within 5 min) and highlighted regions of the elongase enzymes that showed limited resistance to simulated gastric digestion. Traditional gel-based and Western blotting analysis of ω3-desaturase and Δ6-elongase revealed rapid hydrolysis of the intact proteins within seconds and no fragments (> 3 kDa) remained after 60 min, complementing the novel approach described herein. The LC-MS approach was sensitive, selective and did not require the use of purified proteins.

1. Introduction

Crops that express new traits using modern biotechnology were first commercialized with the Flavr Savr™ tomato, which was approved for sale in the USA in 1994. Since that time, most genetically modified crops have been developed with a focus on grower benefits, known as input traits, such as disease/insect protection and herbicide tolerance. Only recently have developers brought output traits to crops that bring benefits directly to the consumer, including improved nutrition and modified oils.

Omega-3 long-chain (≥C20) polyunsaturated fatty acids (ω3 LC-PUFAs) and in particular, eicosapentaenoic acid (EPA, 20:5$^{\Delta5,8,11,14,17}$) and docosahexaenoic acid, (DHA, 22:6$^{\Delta4,7,10,13,16,19}$), have been the focus of interest with medical research highlighting the positive benefits to human health associated with adequate dietary intake of these fatty acids. The positive effects of EPA and DHA have been reported across a range of degenerative and inflammatory disorders such as: heart disease, stroke, rheumatoid arthritis, asthma and some cancers, diabetes mellitus, multiple sclerosis, dementia and clinical depression (Galli and Calder, 2009; Giles et al., 2013).

The main source of dietary EPA and DHA is wild-harvested marine fish whose numbers are generally recognised to be in decline. Oilseed crops, with their large-scale production capacity and relatively low cost, would be an excellent and sustainable source of EPA and DHA oils. The last decade has seen a great deal of research focused on transforming the fatty acid synthesising capability of oilseed crops to synthesise EPA and DHA. The successful accumulation of EPA and DHA into plant hosts has been primarily based upon work performed in the model species *Arabidopsis* and using the aerobic desaturase/elongase pathways (Petrie et al., 2012; Robert et al., 2005; Ruiz-Lopez et al., 2013). This technology was subsequently transferred to *Camelina* as a host species with constructs consisting of five or seven marine microalgal genes for the purpose of engineering either an EPA-only, or an EPA and DHA oil (Ruiz-Lopez et al., 2014). The authors reported an

* Corresponding author.
E-mail address: michelle.colgrave@csiro.au (M.L. Colgrave).

https://doi.org/10.1016/j.fct.2019.05.015
Received 14 November 2018; Received in revised form 8 May 2019; Accepted 10 May 2019
Available online 11 May 2019

EPA content of 24% of total fatty acids in the EPA-only oil, whereas the EPA and DHA oil product had EPA and DHA contents of 11% and 8%, respectively. Similarly, use of a different set of seven yeast and microalgal transgenes, demonstrated fish oil-like levels of DHA production, of up to 12.4% of total fatty acids and an EPA content of 0.8–3.3%, in oil from transgenic *Camelina* (Petrie et al., 2014). Recently we generated transgenic canola with a single vector expressing seven genes for DHA biosynthesis, leading to more than 10% DHA in the seed (Devine et al., 2018). The pathway consisted of the *Lachancea kluyveri* Δ12-desaturase (Lackl-Δ12D (Watanabe et al., 2004)), *Pichia pastoris* Δ15-/ω3-desaturase (Picpa-ω3D (Zhang et al., 2008)), *Micromonas pusilla* Δ6-desaturase (Micpu-Δ6D (Petrie et al., 2010b)), *Pyramimonas cordata* Δ6-elongase (Pyrco-Δ6E (Petrie et al., 2010a)), *Pavlova salina* Δ5-desaturase (Pavsa-Δ5D (Zhou et al., 2007)), *P. cordata* Δ5-elongase (Pyrco-Δ5E (Petrie et al., 2010a)) and *P. salina* Δ4-desaturase (Pavsa-Δ4D (Zhou et al., 2007)). Food, feed and environmental risk assessments of DHA producing canola require evaluation of each of these seven transgenic proteins, including protein stability and plant expression levels.

The allergenic potential of a protein is determined by a weight-of-evidence approach (Joint FAO/WHO Food Standard Programme, 2003; Ladics, 2008) because no single method can predict the allergenicity of a protein. Protein digestibility is one aspect of the overall allergenicity assessment that is conducted for newly introduced proteins into genetically modified crops. Traditionally, protein digestibility has been assessed by subjecting a purified protein of interest to simulated gastric digestion employing pepsin and monitoring the disappearance of the intact protein. This is commonly achieved using gel-based separation (SDS-PAGE) or in combination with antibodies after transfer to nitrocellulose membrane (Western blotting). These techniques have been described as qualitative in nature (Naegeli et al., 2017). It is possible, using SDS-PAGE, to monitor the appearance of large peptide fragments that persist, but often smaller fragments will not be retained on the gel or may be incompatible with subsequent transfer to membranes. Large protein fragments that do transfer may have been hydrolysed at the antigenic site and thus no longer bind monoclonal antibodies. As such, these approaches are largely limited to measuring the loss of intact protein.

Membrane proteins pose further challenges in that they are more difficult to express authentically in recombinant systems especially if they are multi-pass transmembrane proteins. If such proteins are expressed, the yields are often poor and detergents or other membrane mimics must be used to stabilise them in solution. The low expression, difficulties with purification, and the fact that trans-membrane proteins often have only small regions outside the membrane make them poor candidates for the production/generation of antibodies. The target proteins in this study include five desaturase and two elongase enzymes, which are multi-pass transmembrane proteins that exhibit hydrophobic properties and are associated with lipid bilayers (Bushey et al., 2014).

Liquid chromatography-mass spectrometry (LC-MS) offers numerous advantages in protein quantitation (Aebersold et al., 2013). For example, it is possible to conduct LC-MS studies using impure proteins as the detection of the protein is achieved by measuring the derivative peptides and is highly specific and sensitive (Gillette and Carr, 2013). Moreover, peptides representing the entire protein may be identified and monitored enabling a greater understanding of regions of the protein that are more resistant to digestion and hence may warrant further investigation toward their allergenic potential. The LC-MS assay can be multiplexed, such that not only can different regions of a protein be monitored, but multiple proteins can be measured simultaneously. In recent times, digestion proteomics has seen increasing application to not only food safety assessment (De Angelis et al., 2017; Hu and Owens, 2011; Skinner et al., 2016), but to understanding nutrition and elucidating bioactive components of food (Grosvenor et al., 2014).

In this study, targeted proteomics was employed to measure the *in vitro* digestibility of seven transmembrane proteins for which no working antibodies were initially available. The proteins were digested with simulated gastric fluid and the peptic products identified and quantified by LC-MS. The amount of intact protein or high molecular weight protein fragments remaining after the peptic digestion was assessed by complete trypsin digestion of the remaining protein and identification and quantitation of these tryptic peptide products. This novel approach allowed the assessment of protein digestibility using peptides spanning the length of each protein and with reference to a control protein that was highly susceptible to pepsin. Subsequently, the method was validated for two of the seven enzymes using traditional SDS-PAGE and Western blotting approaches.

2. Materials and methods

2.1. Reagents

Sequencing grade porcine trypsin and a highly purified form of pepsin (Catalogue number V195A; specific activity > 2500 units/mg) were purchased from Promega (Madison, USA). Mouse anti-His antibody (Catalogue number A7058) was purchased from Sigma-Aldrich (Sydney, Australia). Bovine serum albumin (BSA Fraction V, biotechnology grade, 0332-100 g) was obtained from AMRESCO-VWR International (PA, USA). The BSA concentration (8.264 mg/mL) was determined by high sensitivity AAA at Australian Proteomics Analysis Facility (Sydney, Australia). HRP-tagged mouse anti-His antibody (Catalogue number A7058) was purchased from Sigma-Aldrich (Sydney, Australia) while Goat Anti-Mouse IgG (H + L)-HRP Conjugate (Catalogue number 1706516) was purchased from Bio-Rad (Australia). Polyacrylamide gels and MES running buffer were supplied by Invitrogen (Australia).

2.2. Methods

The methods used for protein extraction, protein characterisation (using trypsin and pepsin) are described in Protocol S1–S2.

2.2.1. Peptide marker selection

For the tryptic data, peptide summaries generated by ProteinPilot were used to select peptides that yielded intense peaks and were fully tryptic, *i.e.* no unexpected or missed cleavages. For the pepsin data, peptide summaries generated by ProteinPilot were used to select peptides that yielded intense peaks and that were consistently observed in the replicate digests and that were present after 120 min incubation with pepsin. As pepsin is relatively non-specific, many of these peptide products were overlapping or contained missed cleavages. Multiple reaction monitoring (MRM) transitions (Tables S2–S17) were determined for each peptide where the precursor ion (Q1) *m/z* and the fragment ion (Q3) *m/z* values were determined from the data collected in the discovery experiments. Three transitions were used per peptide (with up to 10 peptides from each target protein), wherein the peak area of the three MRM transitions were summed.

2.2.2. Combined pepsin-trypsin digestion assay

For each target protein, 25 μg of protein (n = 30 comprising 5 replicate digestions and 6 time points) were applied to a 10 kDa MWCO filter. The exception was Δ5E, because the protein amount was limited so aliquots of 6.7 μg were used. The buffer was exchanged using 50 mM ammonium bicarbonate (pH 8.0) by two consecutive wash/centrifugation steps. The pH was adjusted by further washing with acidified 50 mM ammonium bicarbonate (pH 1.2) by two consecutive wash/centrifugation steps. The 10 kDa filters were transferred to fresh centrifuge tubes and 75 μg pepsin (150 μL, 0.5 μg/mL in 50 mM ammonium bicarbonate (pH 1.2) was added to obtain an enzyme to protein ratio of 3:1 (equivalent to 7.5 U/μg protein). The replicate tubes were incubated at 37 °C for five time points (5, 10, 15, 30 and 60 min).

Pepsin was not applied to the 0 time point, which served as an experimental control for acid hydrolysis. The digestion was stopped by the addition of 200 µL of 50 mM ammonium bicarbonate (pH 8.0) which irreversibly inactivated the enzyme. The 10 kDa filters were immediately centrifuged (20,800 × *g*, 15 min) and the filtrate containing digested peptides collected. The filters were washed twice with 200 µL of 50 mM ammonium bicarbonate (pH 8.0) and the filtrates were combined, lyophilised and stored in a −80 °C freezer until analysed. The peptic peptides were resuspended in 50 µL of 1% formic acid and run on the 6500 QTRAP LC-MS system.

The 10 kDa filters were transferred to fresh centrifuge tubes and the residual protein was reduced with 200 µL of 50 mM DTT, 50 mM ammonium bicarbonate (pH 8.5) on mixer at 600 rpm for 45 min prior to centrifugation (20,800 × *g*, 15 min). The protein was alkylated with 200 µL of 50 mM iodoacetamide, 50 mM ammonium bicarbonate (pH 8.5) in the dark for 20 min prior to centrifugation (20,800 × *g*, 15 min). The 10 kDa filters were transferred to fresh centrifuge tubes and 2 µg trypsin (200 µL, 0.01 µg/mL in 50 mM ammonium bicarbonate, pH 8.5, and 1 mM $CaCl_2$) was added to obtain an enzyme to protein ratio of ~1:12.5. The replicate tubes were incubated at 37 °C for 16 h. The filters were centrifuged (20,800 × *g*, 15 min) and the filtrates containing digested peptides were collected. The filters were washed twice with 200 µL of 50 mM ammonium bicarbonate (pH 8.5) and the filtrates were combined, lyophilised, and stored in a −80 °C freezer until analysis. The tryptic peptides were resuspended in 50 µL of 1% formic acid and run on the 6500 QTRAP LC-MS.

Either 3 µL of the pepsin digested protein (Tables S2–S9) or reduced and alkylated tryptic peptides (Tables S10–S17) were chromatographically separated on a Shimadzu Nexera UHPLC and analysed on a 6500 QTRAP mass spectrometer (AB SCIEX) as described previously (Colgrave et al., 2014). Quantitation was achieved using scheduled MRM scanning experiments using a 60 s detection window for each MRM transition and a 0.3 s cycle time. Peaks were integrated using MultiQuant v3.0 (AB SCIEX) wherein all three transitions were required to co-elute at the same retention time (RT, min) with a signal-to-noise (S/N) > 3 for detection and a S/N > 5 for quantitation. The graphs showing the digestion time course were generated in Graphpad Prism v6.

2.3. Gels and western blotting experiments

Aliquots of each protein containing 20 µg membrane protein in membrane buffer were added to 1.5 mL tubes containing 200 mM ammonium bicarbonate (pH 0.98) to give a final pH of 1.1 and placed at 37 °C. To each tube, 60 µg of pepsin (30 mg/mL solution made up in acidified ultra-pure Milli Q (MQ) water, pH 3.3, incubated on ice for 30 min before use) was added. Aliquots of 30 µL (containing 0.066 µg/µL protein) were removed at ~0.05 and 0.5, 1, 2, 5, 10, 20, 30 and 60 min and immediately added to a tube containing 1 µL of 1M $NaHCO_3$ (pH 11) to raise the pH to above pH 6.8 and irreversibly inactivate the enzyme. Then 5x SDS-SB + βME (200 mM Tris pH 6.8, 15% (v/v) β-mercaptoethanol, 5% (v/v) SDS, 60% (v/v) glycerol and 0.025% (v/v) bromophenol blue) was added to each sample to yield a 1x solution. Pepsin only and protein only controls were set up in parallel and run on the same gel. For a limit of detection (LOD) analysis the protein only control was diluted with 1 x SDS-SB + βME and run on a gel in loadings of 1, 0.8, 0.5, 0.2, 0.1, 0.05, 0.025, 0.01, 0.005, 0.0025, and 0.001 µg per lane.

The samples from the pepsin digest and their controls were centrifuged at 12,000 × *g* for 5 min. Aliquots of 20 µL containing 1 µg membrane protein and 3 µg pepsin were loaded on to a 4–12% Bis-Tris NuPage gel (Invitrogen, Australia) along with one lane of Novex Sharp molecular weight standards (Invitrogen, Australia). The gel was run in NuPage MES running buffer and electrophoresed for 36 min at a constant voltage 200 V. Proteins were fixed by placing the gel in a solution of 50% (v/v) methanol and 7% (v/v) acetic acid for 15 min. The gel was then washed thrice with 100 mL of high purity water and stained overnight with GelCode Blue Stain (24590, Thermo-Fisher, Australia), a colloidal G250 stain. The gel was destained with several changes of high purity water until the background was clear. The gel was imaged with Odyssey Fc Imager (LI-COR Biosciences, USA).

Samples for Western blots were prepared by diluting (1:100) the pepsin and protein only incubation time point samples (or controls) in 1 x SDS.SB + βME to give 0.5 ng/µL solution and 10 ng membrane protein (20 µL) loaded per lane. The gels were run as before using SeeBlue Plus2 marker (Invitrogen, Australia) to verify transfer of proteins to the PVDF membrane and estimate the size of the transferred proteins. Electrotransfer onto PVDF membrane (Thermo Fisher, Australia) was performed in Membrane Protein Transfer Buffer, MPTB (48 mM Tris, 39 mM glycine, 0.04% SDS, 20% methanol) for 90 min at 220 mA constant current. After transfer the membranes were blocked for a minimum of 2 h with 1% casein in 1x phosphate buffered saline (PBS, 137 mM NaCl, 2.7 mM KCl, 6.5 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 7.2). The His_{10}:Picpa-ω3D membrane was then probed with 1:1000 dilution of mouse anti-Picpa-ω3D antibody in 1% casein/PBS overnight at 4 °C. Excess antibody was removed and the blot washed (3 × 5 min) with PBS. Secondary antibody, goat-anti-mouse HRP (Bio-Rad, Australia) was incubated with the blot for 1 h then again washed (3 × 5 min) with PBS before adding ECL imaging reagent (Pierce, Life Technologies, Australia). After 5 min incubation with the ECL reagent the blot was imaged with Odyssey Fc Imager (LI-COR Biosciences, USA). For the His_{10}:Pyrco-Δ6E, after blocking the membrane the blot was incubated in anti-His HRP antibody (Sigma-Aldrich, Australia) overnight. Excess antibody was removed, then the membrane was washed, incubated with ECL imaging reagent and imaged as above. Both membranes were analysed to determine the apparent molecular weights (*Mr*) of the digestion products using Image lab software (Bio-Rad, Australia).

3. Results and discussion

3.1. Development and validation of a multiplexed LC-MS assay

3.1.1. Protein digestibility as one measure of food safety

The likelihood of allergic oral sensitization to a protein is first affected by the stability of the protein to gastrointestinal digestion (Astwood et al., 1996; Goodman et al., 2008; Thomas et al., 2004). The

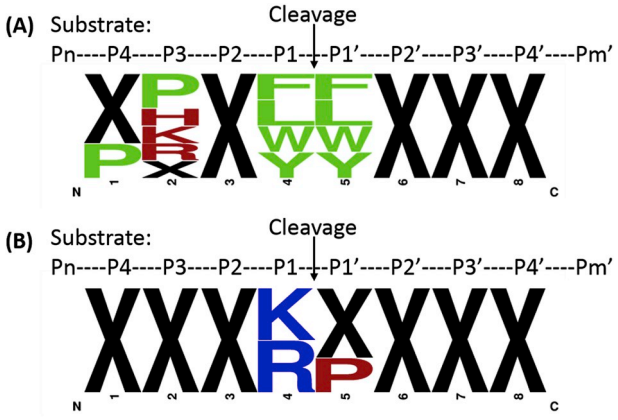

Fig. 1. Specificity of proteolytic enzymes used in this study. (A) Pepsin cleavage sites at both sides of aromatic and hydrophobic amino acids. Amino acids that act to hinder proteolysis at the P3 position are shown in red, whereas those that promote proteolysis (P at P3 or P4) are shown in green. (B) Trypsin cleavage site (blue). Proline (P) as the succeeding amino acid (P1′ position) hinders proteolysis is shown in red. The images were created using WebLogo (Crooks et al., 2004). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

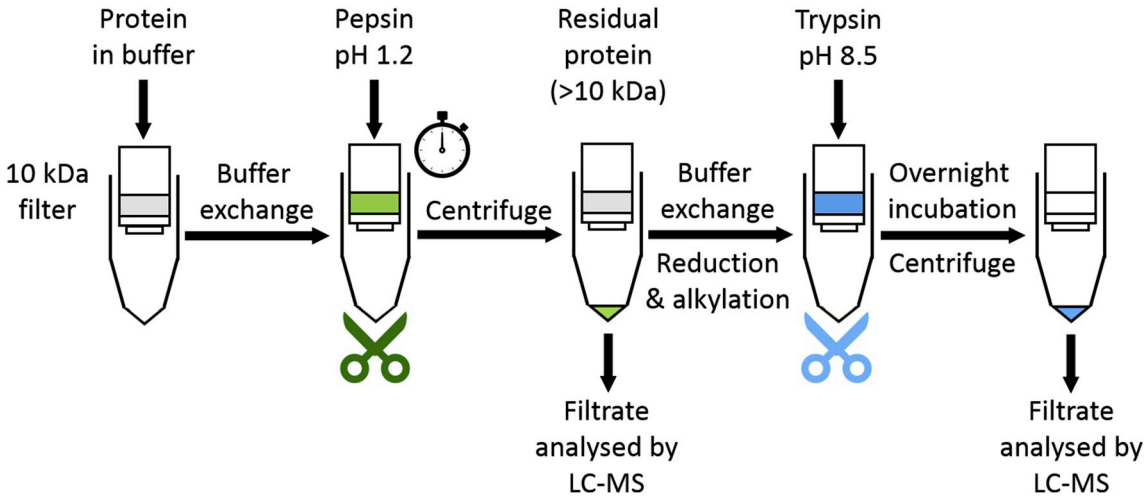

Fig. 2. Workflow used in combined pepsin-trypsin digestion assay. The protein was applied to a 10 kDa molecular weight cut-off filter and the buffer was exchanged to ammonium bicarbonate at pH 1.2. Pepsin was added and the protein was sampled at time points from 5 to 60 min (after quenching by increasing the pH). The filtrate was analysed by LC-MS to reveal the liberation of peptides by pepsin. The residual protein (intact or partially hydrolysed) remained on the filter and after buffer exchange, the protein was reduced, alkylated and then digested by trypsin. The filtrate contained trypsin derived peptides from protein fragments > 10 kDa.

purpose of this study was to assess the *in vitro* digestibility of the fatty acid biosynthesis enzymes, present in DHA canola, by digestion with pepsin. In the absence of functioning antibodies against these integral membrane proteins, as typically used in traditional Western blot analysis, a sensitive LC-MS analytical method was developed.

The complete digestion of a protein by a single enzyme is difficult to judge, especially when employing a non-specific enzyme such as pepsin, adding uncertainty to allergenicity determinations. Although it is possible to judge the disappearance of the intact protein on a gel or by Western blotting techniques, the protein may be hydrolysed once (cleaved at a single site) or multiple times, often resulting in small and overlapping fragments. Allergenic reactions require that a protein or protein fragment simultaneously bind to two IgE molecules in order to induce mast cell degranulation (Goodman and Wise, 2006). In order for antibody binding to occur, the IgE epitopes must be ~15 amino acid in length (Huby et al., 2000). This IgE binding places theoretical limits on

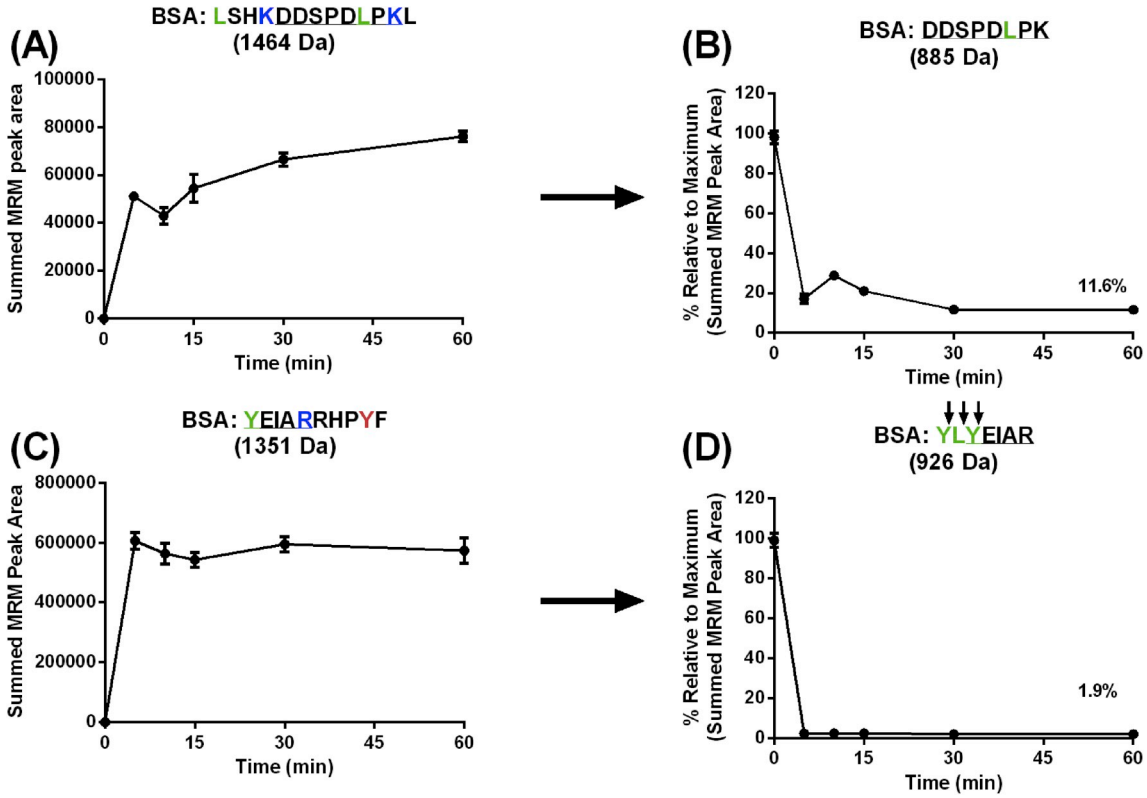

Fig. 3. Pepsin stability assay monitored by LC-MS for the control protein, bovine serum albumin (BSA). Two peptides are shown for pepsin expressed as MRM peak area (A, C) and trypsin post-pepsin expressed as a percentage relative to the maximum MRM peak area (B, D). The selected peptides are derived from the same region of the protein such that the tryptic peptide on the right overlaps with the peptic peptide on the left (underlined). The amino acid sequence (and mass in Da) are shown in the header. The font is colored to show potential cleavage sites: green, missed pepsin cleavage site; red, potentially hindered pepsin cleavage site; blue, trypsin cleavage site. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

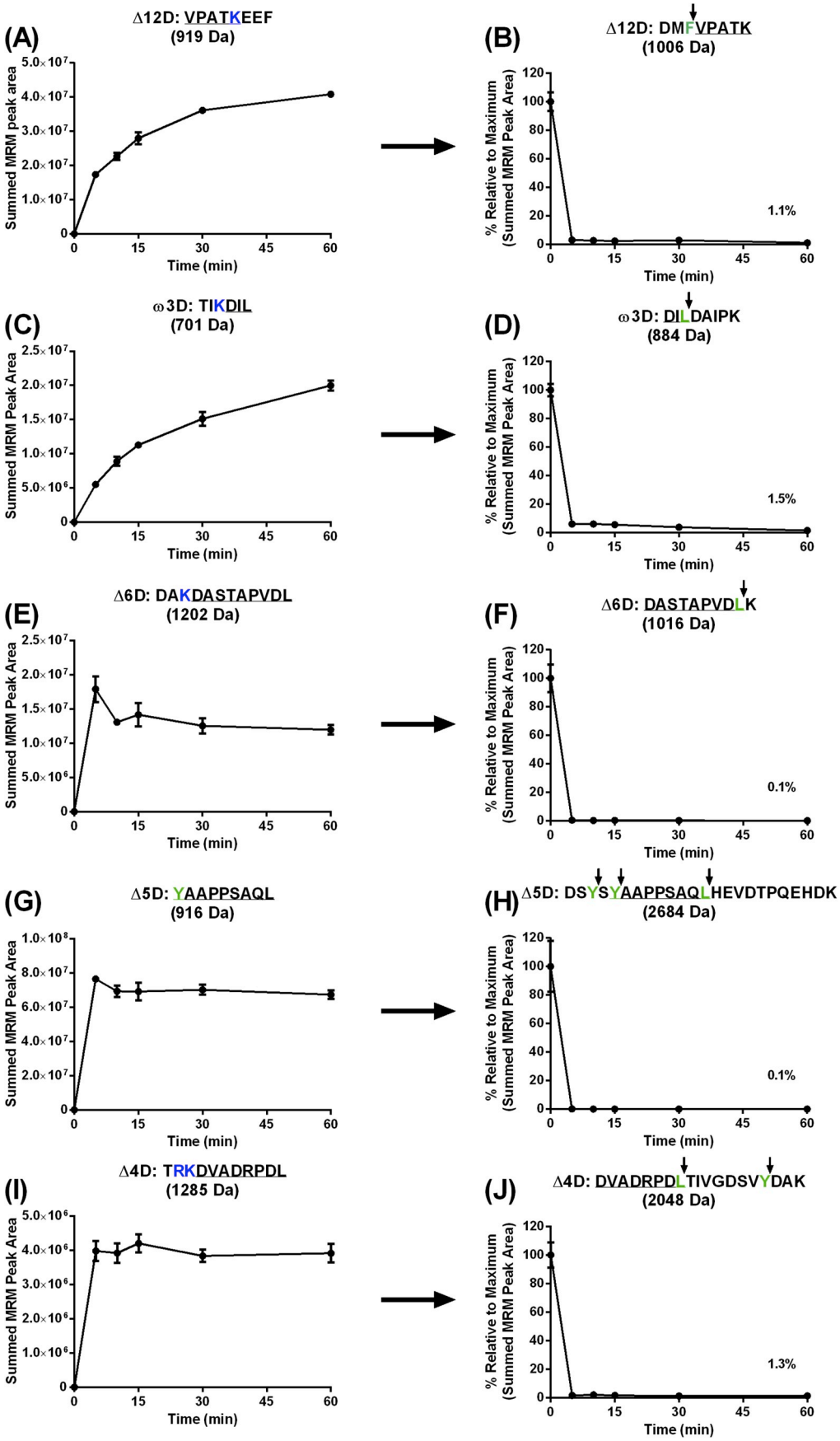

Fig. 4. Pepsin stability assay monitored by LC-MS for the five desaturases. Left panels show peptides derived by pepsin expressed as MRM peak area (A, C, E, G, I) and right panels show peptides derived by trypsin post-pepsin expressed as a percentage relative to the maximum MRM peak area (B, D, F, H, J). The selected peptides are derived from the same region of the protein such that the tryptic peptide on the right overlaps with the peptic peptide on the left (underlined). The amino acid sequence (and mass in Da) are shown in the header. The font is colored to show potential cleavage sites: green, missed pepsin cleavage site; blue, trypsin cleavage site. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

the peptide size of at least 30 amino acids or ~3 kDa. Gel analysis with various staining or antibody techniques is typically able to detect peptides down to approximately 3 kDa, although peptides are detected with lower sensitivity than proteins. When employing gel analysis solely, to judge the completeness of digestion, a high level of purity is required. When employing antibodies, the hydrolysis of a protein by a proteolytic enzyme may result in cleavage of the antigenic site (epitope) thus rendering antibody-based detection methods unsuitable. Conversely, cleavage of a protein at a single site may yield two protein fragments, in which one may contain the epitope (recognised by a monoclonal antibody) while the other does not. In such instances, large protein fragments may evade detection.

3.1.2. A dual-enzyme assay to judge the completeness of digestion

By using LC-MS/MS analysis, the peptide products resulting from both pepsin and trypsin digestions were determined qualitatively, and subsequently a quantitative LC-MS/MS method for the detection of these peptide fragments was developed. LC-MS analysis is capable of simultaneously monitoring peptides spanning the entire protein sequence that are generated by proteolytic digestion. The specificity of the enzymes used is depicted in Fig. 1. Pepsin is efficient at cleaving the peptide bonds adjacent to aromatic and hydrophobic amino acids (F, L, W, Y; Fig. 1A), but may be hindered by basic residues at the P3 position. Trypsin cleaves polypeptide chains at the carboxyl side of the basic amino acids (K, R), but the cleavage is hindered by the presence of proline as the succeeding amino acid (P1' position; Fig. 1B).

Upon digestion with pepsin alone, the target protein may be rapidly digested to produce fully peptic fragments wherein the response rapidly increases reaching a maximum and creating a plateau, indicating that digestion is complete (Fig. S1), but it is possible to reach a plateau with sub-maximal degradation occurring. It is also possible that the digestion rate is slower wherein a visible plateau may not be reached within the experimental duration. By employing trypsin to the protein that remains on the filter (> 10 kDa in size) after pepsin digestion, it is possible to judge the completeness of the digestion by comparison to an experimental control (time 0, no pepsin added) wherein the tryptic peptides liberated, that serve to act as a proxy for intact protein, appear at the maximum value. If the protein is completely digested by pepsin, the tryptic peptide response will drop to zero within the experimental duration.

3.1.3. Implementation of the dual-enzyme assay

Simulated gastric fluid (SGF) was represented by the proteolytic enzyme pepsin in a buffer adjusted to an acidic pH 1.2 (Naegeli et al., 2017). Based on a recent study (Schafer et al., 2016) demonstrating that heterologous proteins were suitable surrogates for proteins in plant matrices in SGF digestion studies, the seven target proteins were first expressed and characterised by SDS-PAGE (Protocol S1). In order to develop the LC-MS assay the peptide fragments present after pepsin (2 h, Fig. S2) and trypsin (16 h, Fig. S3) digestion were identified by untargeted LC-MS/MS with database searching. Peptides that were identified with 95% confidence and that yielded intense signals in the MS were selected for relative quantitation. Where possible for each protein, ten pepsin-derived and ten trypsin-derived peptides were selected that spanned the length of the protein. Notably, for the desaturases the majority of peptides yielding intense peaks were derived from the more hydrophilic loop regions, whereas the paucity of tryptic sites in the elongases limited the analysis to peptides within transmembrane regions.

The digestion was performed on 10 kDa MWCO filters for 5, 10, 15,

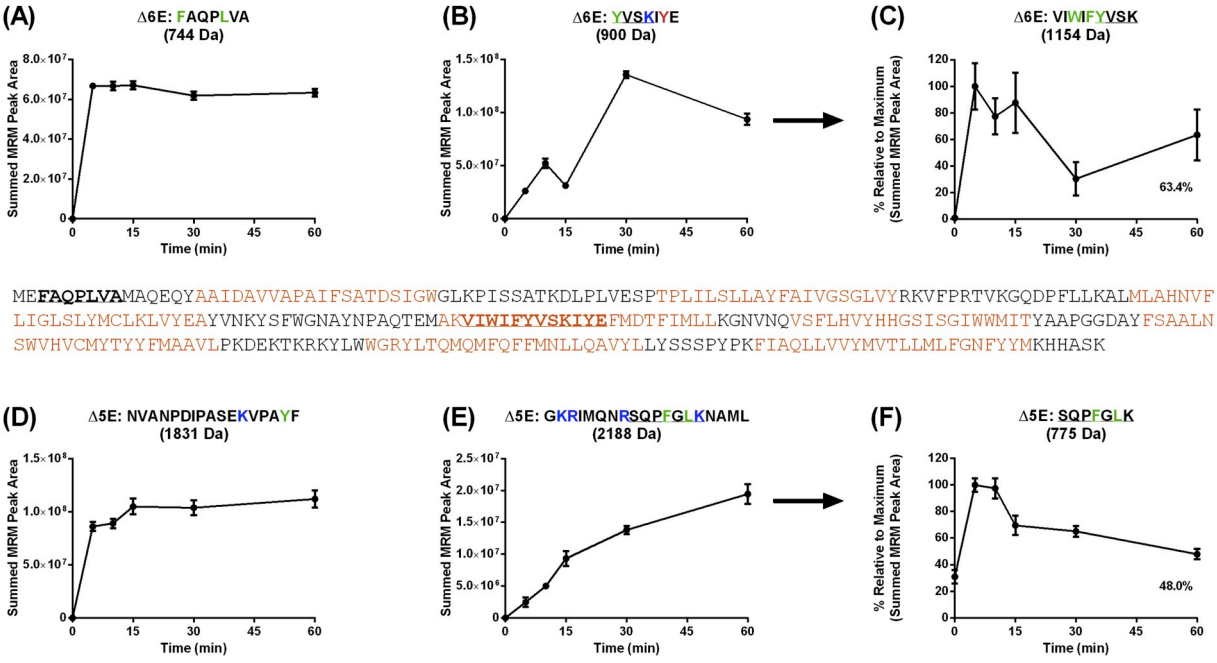

Fig. 5. Pepsin stability assay monitored by LC-MS for the two elongases. Left and middle panels show peptides derived by pepsin expressed as MRM peak area (A, B, D, E) and right panels show peptides derived by trypsin post-pepsin expressed as a percentage relative to the maximum MRM peak area (C, F). Selected peptides (B, C and E, F) are derived from the same region of the protein such that the tryptic peptide on the right overlaps with the peptic peptide on the left (underlined). The amino acid sequence (and mass in Da) are shown in the header. The font is colored to show potential cleavage sites: green, missed pepsin cleavage site; red, potentially hindered pepsin cleavage site; blue, trypsin cleavage site. The protein sequences for Δ6E (top) and Δ5E (bottom) are shown wherein the peptides monitored are depicted by bold, underlined font and the predicted transmembrane domains are shown in orange font. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

30 and 60 min, with 0 min (no pepsin added) as the control, and each with five replicates (Fig. 2). Peptides derived by pepsin digestion (*i.e.* < 10 kDa) were collected by filtration. The undigested or partially digested (> 10 kDa) protein that remained was subjected to buffer exchange followed by 16 h digestion with trypsin. The repeatability of the digestion strategy was assessed using two batches of Δ4-desaturase that were digested a year apart (Fig. S4). The digestion profiles were highly similar, despite protein and enzyme batch variation, demonstrating that this approach is robust.

3.1.4. Validation of assay using a control protein

The pepsin digestion data is presented as the mean of five replicate digests and is expressed as the summed MRM peak area (sum of three transitions). The trypsin digestion data is presented as the percentage relative to the maximum detected MRM peak area per peptide across the time points. As a control bovine serum albumin (BSA, P02769), a readily digestible protein, was first examined. Fig. 3 shows the rapid degradation of the intact protein as evidenced by the increase in peak area for the pepsin digestion products (Fig. 3A and C) and the rapid decrease in the level of the trypsin-derived peptides that served as proxies for the intact protein (Fig. 3B and D). The peptides depicted overlap such that the pepsin-derived fragment in the left panel (A, C) share the same protein region as the trypsin-derived peptides in the right panel (B, D). The first example (Fig. 3A) shows the gradual accumulation of the pepsin fragment, LSHKDDSPDLPKL, over 60 min. After each time point of pepsin digestion the remaining protein (on the 10 kDa filter) was buffer exchanged before being subjected to trypsin digestion. In this example, the tryptic partner DDSPDLPK, which is internal to the pepsin fragment (indicated by underline) was quantified at each time point. With a continual release of the pepsin fragment LSHKDDSPDLPKL over the experimental duration, it is not surprising that the partner tryptic fragment DDSPDLPK was still able to be detected at 12% after 60 min (Fig. 3B). The second example (Fig. 3C) shows the rapid (< 5 min) liberation of the pepsin fragment, YEAIRRHPYF indicating efficient hydrolysis of the peptide at YL$^{\downarrow}$Y. The complete disappearance of the tryptic partner peptide (Y$^{\downarrow}$L$^{\downarrow}$Y$^{\downarrow}$EIAR, Fig. 3D) after 5 min is expected given the three pepsin cleavage sites within its sequence, one of which is cleaved efficiently by pepsin (Fig. 3C). The time course of digestion for the entire suite of peptides monitored are provided in Fig. S5 (pepsin) and Fig. S6 (trypsin) and the percentage of all tryptic peptide markers at each time point is presented in Table S1. These data demonstrate that even in a readily digestible protein such as BSA, there are regions of a protein that show limited resistance to proteolysis. Moreover, the measurement of tryptic peptides that span the length of the protein provides evidence of multi-site hydrolysis that cannot be gleaned from traditional gel-based analyses.

3.1.5. Monitoring the digestion of the transmembrane enzymes

In a similar manner to BSA, Fig. 4 shows related pairs of pepsin fragments (left) and tryptic products (right) for the five desaturase enzymes (with the complete dataset presented in Figs. S7–S20). In each instance the appearance of pepsin products indicates cleavage of the intact protein, while the disappearance of the tryptic peptides evidences rapid hydrolysis (> 94% at 5 min, and > 98% at 60 min) of the intact protein. Because the membrane proteins comprise many hydrophobic amino acids that are potential pepsin cleavage sites, many small overlapping products were generated. For example, the pepsin-derived peptide TIKDIL monitored for Δ12D (Fig. 4C) was derived from a larger fragment, FSVPDF$^{\downarrow}$TIKDIL, which reached a maximum level at 5 min before decreasing in abundance as it suffered a secondary cleavage (F$^{\downarrow}$) to yield FSVPDF and TIKDIL, both of which continued to increase in abundance from 5 to 60 min (Fig. S7, A-C).

In the case of the two elongase enzymes examined (Fig. 5), the majority of the pepsin fragments appeared rapidly, reaching a maximal level within 10–15 min (Δ5E, Fig. S13), whilst some peptide levels continued to increase over the duration of the experiment (increasing to 60 min). Notably, the trypsin digestion did not yield maximum levels at time 0 min, and a three-to four-fold increase was observed at 5 min, indicating that the action of pepsin disrupted the tertiary structure of the proteins, thus exposing the tryptic sites leading to increased liberation of the trypsin products at this time point. Subsequently the tryptic peptides decreased in abundance to 30 min, but at a slower rate than noted for the desaturases. Additionally, we observed a slight increase from 30 to 60 min for three of the tryptic peptides monitored from Δ6-elongase (Fig. S14), which may be due to further disruption of the tertiary structure by the action of pepsin. The elongases (31–33 kDa) are predicted to comprise five to seven transmembrane helices (Fig. 5, Fig. S21), compared with three to five in the desaturases which are also notably larger (48–53 kDa). Rapid liberation of peptides in regions that are predicted to protrude into the cytosol was observed (Fig. 5A; 5D; Fig. S13, A-C, J, Figure S17, A-B), whereas, a lag before pepsin release of peptides predicted to be embedded within the lipid bilayer was observed (Fig. 5B; 5E; Fig. S13E, G-H, Figs. S17G–H). Examining the proteomics data collected at the 60 min time point that was initially used for LC-MS assay development, pepsin digestion yielded

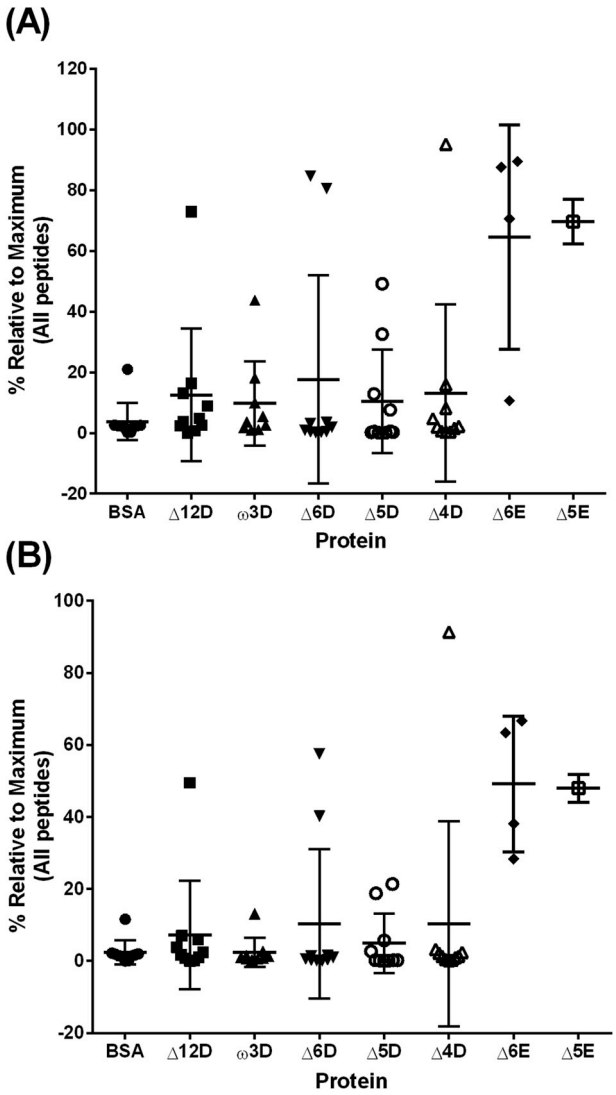

Fig. 6. Stability of target proteins as assessed in a pepsin digestion time course. BSA is used as the control protein for comparison. Scatter plot showing the percentage of tryptic peptides remaining after (A) 15 min; and (B) 60 min (n = 10 peptides for BSA, Δ12D, Δ6D, Δ5D and Δ4D; n = 9 peptides for ω3D, n = 4 for Δ6E, n = 1 for Δ5E). The individual data points represent the mean percentage remaining for each peptide from n = 5 technical replicates.

67% sequence coverage (Fig. S2E), wherein the largest Δ6-elongase fragment detected was 17 amino acids or ~2 kDa and the largest region where no peptides were detected was 20 amino acids in length (~2.4 kDa). Likewise, for the Δ5-elongase 61% sequence coverage was obtained (Fig. S2G) in which all fragments detected/possible were < 3 kDa. These fragments are likely too small to elicit a type I allergenic response (Huby et al., 2000).

The stability of the seven target proteins to SGF conditions was then compared to the readily digestible control protein, BSA (Fig. 6) by plotting the mean percentage remaining at 15 min (Figure 6A) and 60 min (Fig. 6B) for up to ten tryptic peptides monitored for each protein. There was no significant difference in the reported digestibility for the five desaturase enzymes relative to BSA. The elongases showed greater resistance to pepsin digestion with Δ6E showing a significant difference ($p = 0.015$) although only four tryptic peptides could be monitored due to a paucity of trypsin cleavage sites within the amino acid sequence. Qualitatively, Δ5E was similar to Δ6E but because only a single tryptic peptide was available, statistical interpretation could not be conducted. As such, the elongases were determined to be less susceptible to gastrointestinal digestion, however, they were also the proteins that were of lowest abundance (Δ6E and Δ5E were detected as 0.02–0.05% and 0.04–0.13% of seed protein respectively) of the seven target proteins analysed in mature canola seed (Colgrave et al., 2019). The abundance of a protein is also considered a risk factor for allergenicity, with many major food allergens accounting for > 1% of the protein in allergenic foods (Astwood et al., 1996).

3.2. Validation of digestibility using traditional methods

SDS-PAGE analysis was used to determine the stability of the Picpa-ω3D and the Pyrco-Δ6E proteins when treated with pepsin using the same buffer used for the LC-MS analyses. The bands on the gel corresponding to the two full length proteins migrated at a lower molecular weight than predicted by their amino acid sequence, as is common with membrane proteins (Rath et al., 2009), so the gel-derived full-length molecular weight will be referred to as the apparent molecular weight (*Mr*). Fig. 7A shows the gel analysis of the His_{10}:Picpa-ω3D after digestion with pepsin (dark band at *Mr* ~39 kDa) and Fig. 7B shows the corresponding Western blot probed with anti-Picpa-ω3D antibody. The limit of detection (LOD) of the gel-based detection for His_{10}:Picpa-ω3D was determined to be 0.025 μg, which is 2.5% of the total protein loaded (Fig. S22). Visual examination of Fig. 7A shows that 97.5% of the His_{10}:Picpa-ω3D (full length band at *Mr* 40.5) was digested in under 0.05 min (~3 s). Low molecular weight (LMW) fragments of size 10–19 kDa were detected in the first 2 min as evidenced on both the gel (Fig. 7A) and the blot (Fig. 7B). By 5 min these LMW fragments were disappearing to yield major bands between 7.5 and 11 kDa. After 20 min all of the detectable bands were < 7.5 kDa. Fig. 7C shows the

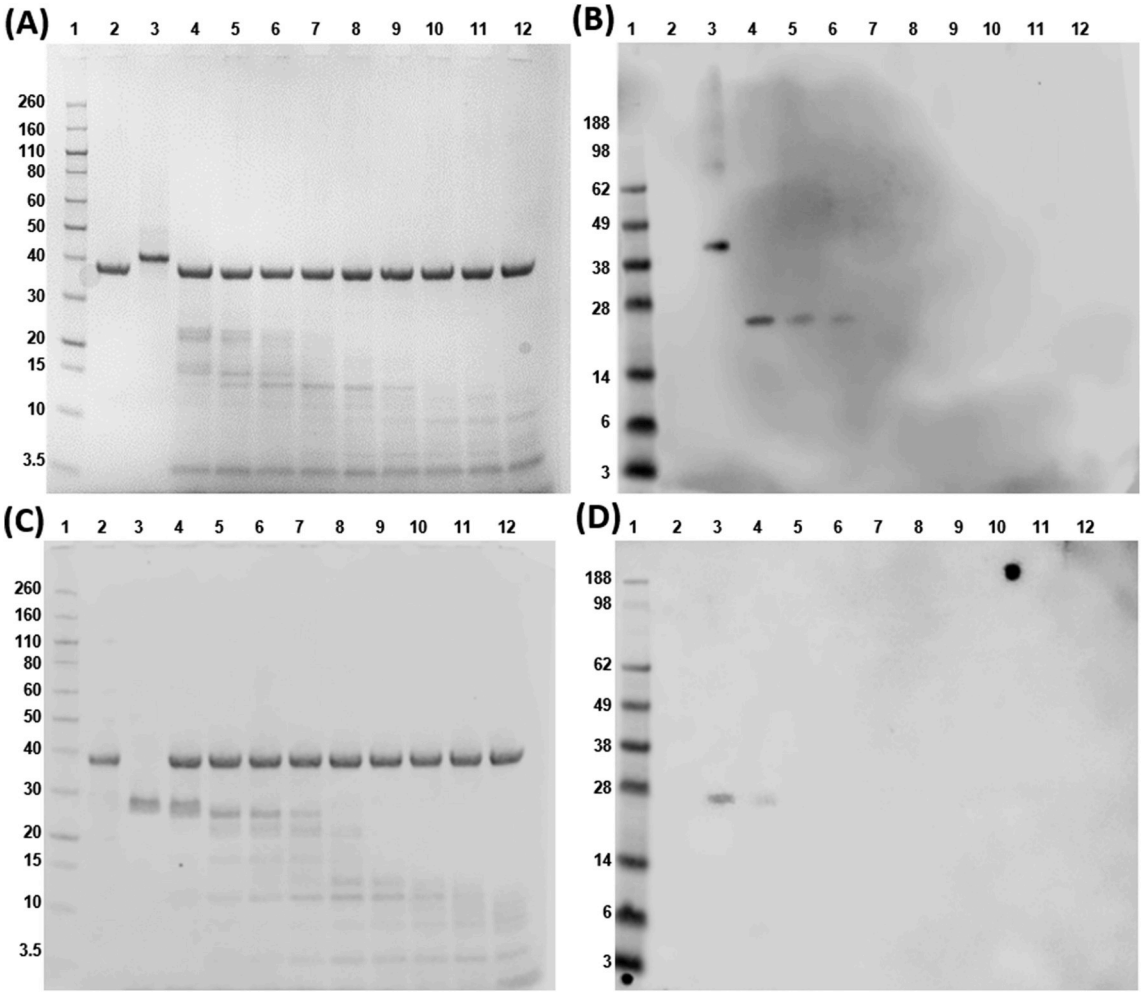

Fig. 7. SDS-PAGE gels (left) and Western blots (right) analysis of His_{10}:Picpa-ω3D (A, B) and His_{10}:Pyrco-Δ6E (C, D). Gels were stained with GelCode Stain and show pepsin (lane 2) and undigested target protein (lane 3). All lanes were loaded with 20 μL of solution comprising 1 μg equivalent protein with 3 μg equivalent pepsin. Aliquots were taken at different time points in the pepsin digestion time course with lanes 4–12 representing < 0.05, 0.5, 1, 2, 5, 10, 20, 30 and 60 min. The corresponding Western blot using the same samples loaded (but with 1/100 dilution to give 10 ng per lane) are shown on the right. The blots were probed with anti-Picpa-ω3D (F7) antibody (B) or anti-His-tag HRP antibody (D).

gel analysis of the His_{10}:Pyrco-Δ6E after pepsin digestion and Fig. 7D shows the corresponding Western blot probed with anti-His HRP antibody. The LOD of the His_{10}:Pyrco-Δ6E band on the gel was determined to be 0.05 μg (5% of protein loaded, Fig. S23). Visual examination of Fig. 7C shows that 95% of full-length His_{10}:Pyrco-Δ6E (*Mr* ~28 kDa) was digested in under 0.5 min. After ~3 s the gel reveals a band that appears to be ~1 kDa smaller than the full-length protein (Fig. 7C) observed as a faint band on the corresponding Western blot (Fig. 7D). The early time points (0.5–2 min) revealed bands that are ≥3 kDa lower than the full-length protein. After 5 min, the bands > 20 kDa are noted to be very faint and LMW protein fragments of size 7–11 kDa are the major components observed in the time points taken from 10 to 60 min.

The gels and blots constitute a traditional method for demonstrating protein digestibility as evidenced by the loss of intact protein and the appearance of protein fragments of size 3 kDa up to the size of the intact protein. These data correlate with and complement the LC-MS analyses wherein the appearance of small (< 3 kDa) peptides derived from pepsin digestion were monitored under the same conditions. Moreover, the LC-MS study utilized peptide markers of each target protein derived from the application of trypsin post-pepsin. This novel method enabled the measurement of the degradation of each region of the protein. Together these complementary studies show that His_{10}:Picpa-ω3D and His_{10}:Pyrco-Δ6E were readily digested under conditions mimicking those expected under physiological conditions. The LC-MS approach provided greater depth in our understanding of the end point(s) of digestion, however, and was capable of detecting the lag in the digestion of His_{10}:Pyrco-Δ6E that was not obvious when using gels/Western blots alone.

4. Conclusions

The results of this study show that these integral membrane proteins were readily digestible by pepsin and/or trypsin. The combined pepsin-trypsin assay showed a rapid decline in the tryptic peptides that were used as a proxy for the presence of intact protein. More specifically, the results of this study demonstrate that no full-length target proteins remained after 5 min incubation in pepsin (the earliest time point sampled by LC-MS) and in fact this was likely to be within seconds as evidenced by the gel-based analyses of the two exemplar proteins ω3-desaturase and Δ6-elongase. LC-MS offers a number of benefits over traditional analysis employing gels and Western blotting (Aebersold et al., 2013) including pre-analytical savings: (1) no need for protein purification; and (2) no requirement for generation of antibodies, which can be a costly and timely process. LC-MS is robust, offers speed, sensitivity and selectivity, but moreover provides the ability to multiplex and hence monitor multiple proteins simultaneously. This approach also enables the measurement of peptides spanning the length of the protein providing greater understanding of the kinetics of digestion and potentially highlighting regions of increased resistance that could warrant further investigation. Rapid digestion of the full-length proteins as demonstrated in this study is one of many factors that indicate a reduced risk of allergenic potential.

Additional Information

Data availability statement

All relevant data are within the paper and its Supporting information files.

Conflicts of interest

The authors report grants from Nuseed Pty Ltd, during the conduct of the study. In addition, the authors have a patent US 7807849 B2 “Synthesis Of Long-chain Polyunsaturated Fatty Acids By Recombinant Cells” issued.

Author contributions

Conceptualization: MC, JP, SS, SM, XZ. Investigation: KB, JC, LK, GL, SVP, BD. Methodology: MC. Formal analysis and data curation: JS, XZ, KB, SM, MC. Writing: KB, MC, XZ.

Acknowledgments

The authors thank Alun Jones and the Molecular and Cellular Proteomics Mass Spectrometry Facility at the University of Queensland for access to some of the mass spectrometers used in this study. This study was in part funded by Nuseed Pty Ltd.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fct.2019.05.015.

Transparency document

Transparency document related to this article can be found online at https://doi.org/10.1016/j.fct.2019.05.015

References

Aebersold, R., Burlingame, A.L., Bradshaw, R.A., 2013. Western blots versus selected reaction monitoring assays: time to turn the tables? Mol. Cell. Proteom. 12, 2381–2382.
Astwood, J.D., Leach, J.N., Fuchs, R.L., 1996. Stability of food allergens to digestion in vitro. Nat. Biotechnol. 14, 1269–1273.
Bushey, D.F., Bannon, G.A., Delaney, B.F., Graser, G., Hefford, M., Jiang, X., Lee, T.C., Madduri, K.M., Pariza, M., Privalle, L.S., Ranjan, R., Saab-Rincon, G., Schafer, B.W., Thelen, J.J., Zhang, J.X., Harper, M.S., 2014. Characteristics and safety assessment of intractable proteins in genetically modified crops. Regul. Toxicol. Pharmacol. 69, 154–170.
Colgrave, M.L., Goswami, H., Blundell, M., Howitt, C.A., Tanner, G.J., 2014. Using mass spectrometry to detect hydrolysed gluten in beer that is responsible for false negatives by ELISA. Journal of Chromatography A 1370, 105–114.
Colgrave, M.L., Byrne, K., Vibhakaran Pillai, S., Dong, B., Leonforte, A., Caine, J., Kowalczyk, L., Scoble, J.A., Petrie, J.R., Singh, S.P., Zhou, X.R., 2019. Quantitation of seven transmembrane enzymes from the DHA biosynthesis pathway in genetically engineered canola by targeted mass spectrometry. Food Chem. Toxicol. 126, 313–321.
Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res 14, (6), 1188–1190.
De Angelis, E., Pilolli, R., Bavaro, S.L., Monaci, L., 2017. Insight into the gastro-duodenal digestion resistance of soybean proteins and potential implications for residual immunogenicity. Food Funct 8, 1599–1610.
Devine, M., Leonforte, A., Gororo, N., Buzza, G., Tang, S., Gao, W., Petrie, J.R., Singh, S.P., 2018. Elite Event Canola NS-B50027-4. Nuseed Pty Ltd., pp. 1–103.
Galli, C., Calder, P.C., 2009. Effects of fat and fatty acid intake on inflammatory and immune responses: a critical review. Ann. Nutr. Metab. 55, 123–139.
Giles, G.E., Mahoney, C.R., Kanarek, R.B., 2013. Omega-3 fatty acids influence mood in healthy and depressed individuals. Nutr. Rev. 71, 727–741.
Gillette, M.A., Carr, S.A., 2013. Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. Nat. Methods 10, 28–34.
Goodman, R.E., Vieths, S., Sampson, H.A., Hill, D., Ebisawa, M., Taylor, S.L., van Ree, R., 2008. Allergenicity assessment of genetically modified crops–what makes sense? Nat. Biotechnol. 26, 73–81.
Goodman, R.E., Wise, J., 2006. Predicting the allergenicity of novel proteins in genetically modified organisms. In: Maleki, S.J., Burks, A., Helm, R. (Eds.), Food Allergy. ASM Press, Washington D.C., USA, pp. 219–247.
Grosvenor, A.J., Haigh, B.J., Dyer, J.M., 2014. Digestion proteomics: tracking lactoferrin truncation and peptide release during simulated gastric digestion. Food Funct 5, 2699–2705.
Hu, X.T., Owens, M.A., 2011. Multiplexed protein quantification in maize leaves by liquid chromatography coupled with tandem mass spectrometry: an alternative tool to immunoassays for target protein analysis in genetically engineered crops. J. Agric. Food Chem. 59, 3551–3558.
Huby, R.D., Dearman, R.J., Kimber, I., 2000. Why are some proteins allergens? Toxicol. Sci. 55, 235–246.
Joint FAO/WHO Food Standard Programme, 2003. Appendix III, Guideline for the Conduct of Food Safety Assessment of Foods Derived from Recombinant-DNA Plants, and Appendix IV, Annex on the Assessment of Possible Allergenicity. Codex Alimentarious Commission, Rome, Italy, pp. 47–60.

Ladics, G.S., 2008. Current codex guidelines for assessment of potential protein allergenicity. Food Chem. Toxicol. 46 (Suppl. 10), S20–S23.
Naegeli, H., Birch, A.N., Casacuberta, J., De Schrijver, A., Gralak, M.A., Guerche, P., Jones, H., Manachini, B., Messean, A., Nielsen, E.E., Nogue, F., Robaglia, C., Rostoks, N., Sweet, J., Tebbe, C., Visioli, F., Wal, J.M., Eigenmann, P., Epstein, M., Hoffmann-Sommergruber, K., Koning, F., Lovik, M., Mills, C., Moreno, F.J., van Loveren, H., Selb, R., Dumont, A.F., 2017. Guidance on allergenicity assessment of genetically modified plants. Efsa J 15 EFSA Panel Genetically Modified Organisms.
Petrie, J.R., Liu, Q., Mackenzie, A.M., Shrestha, P., Mansour, M.P., Robert, S.S., Frampton, D.F., Blackburn, S.I., Nichols, P.D., Singh, S.P., 2010a. Isolation and characterisation of a high-efficiency desaturase and elongases from microalgae for transgenic LC-PUFA production. Mar. Biotechnol. 12, 430–438.
Petrie, J.R., Shrestha, P., Belide, S., Kennedy, Y., Lester, G., Liu, Q., Divi, U.K., Mulder, R.J., Mansour, M.P., Nichols, P.D., Singh, S.P., 2014. Metabolic engineering *Camelina sativa* with fish oil-like levels of DHA. PLoS One 9, e85061.
Petrie, J.R., Shrestha, P., Mansour, M.P., Nichols, P.D., Liu, Q., Singh, S.P., 2010b. Metabolic engineering of omega-3 long-chain polyunsaturated fatty acids in plants using an acyl-CoA Δ6-desaturase with ω3-preference from the marine microalga *Micromonas pusilla*. Metab. Eng. 12, 233–240.
Petrie, J.R., Shrestha, P., Zhou, X.R., Mansour, M.P., Liu, Q., Belide, S., Nichols, P.D., Singh, S.P., 2012. Metabolic engineering plant seeds with fish oil-like levels of DHA. PLoS One 7, e49165.
Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G., Deber, C.M., 2009. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. Proc. Natl. Acad. Sci. U. S. A. 106, 1760–1765.
Robert, S.S., Singh, S.P., Zhou, X.-R., Petrie, J.R., Blackburn, S.I., Mansour, P.M., Nichols, P.D., Liu, Q., Green, A.G., 2005. Metabolic engineering of Arabidopsis to produce nutritionally important DHA in seed oil. Funct. Plant Biol. 32, 473–479.
Ruiz-Lopez, N., Haslam, R.P., Napier, J.A., Sayanova, O., 2014. Successful high-level accumulation of fish oil omega-3 long-chain polyunsaturated fatty acids in a transgenic oilseed crop. Plant J. : for cell and molecular biology 77, 198–208.
Ruiz-Lopez, N., Haslam, R.P., Usher, S.L., Napier, J.A., Sayanova, O., 2013. Reconstitution of EPA and DHA biosynthesis in arabidopsis: iterative metabolic engineering for the synthesis of n-3 LC-PUFAs in transgenic plants. Metab. Eng. 17, 30–41.
Schafer, B.W., Embrey, S.K., Herman, R.A., 2016. Rapid simulated gastric fluid digestion of in-seed/grain proteins expressed in genetically engineered crops. Regul. Toxicol. Pharmacol. 81, 106–112.
Skinner, W.S., Phinney, B.S., Herren, A., Goodstal, F.J., Dicely, I., Facciotti, D., 2016. Using LC-MS based methods for testing the digestibility of a nonpurified transgenic membrane protein in simulated gastric fluid. J. Agric. Food Chem. 64, 5251–5259.
Thomas, K., Aalbers, M., Bannon, G.A., Bartels, M., Dearman, R.J., Esdaile, D.J., Fu, T.J., Glatt, C.M., Hadfield, N., Hatzos, C., Hefle, S.L., Heylings, J.R., Goodman, R.E., Henry, B., Herouet, C., Holsapple, M., Ladics, G.S., Landry, T.D., MacIntosh, S.C., Rice, E.A., Privalle, L.S., Steiner, H.Y., Teshima, R., Van Ree, R., Woolhiser, M., Zawodny, J., 2004. A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. Regul. Toxicol. Pharmacol. 39, 87–98.
Watanabe, K., Oura, T., Sakai, H., Kajiwara, S., 2004. Yeast Δ12 fatty acid desaturase: gene cloning, expression, and function. Biosci. Biotechnol. Biochem. 68, 721–727.
Zhang, X., Li, M., Wei, D., Xing, L., 2008. Identification and characterization of a novel yeast ω3-fatty acid desaturase acting on long-chain n-6 fatty acid substrates from *Pichia pastoris*. Yeast 25, 21–27.
Zhou, X.R., Robert, S.S., Petrie, J.R., Frampton, D.M., Mansour, M.P., Blackburn, S.I., Nichols, P.D., Green, A.G., Singh, S.P., 2007. Isolation and characterization of genes from the marine microalga *Pavlova salina* encoding three front-end desaturases involved in docosahexaenoic acid biosynthesis. Phytochemistry 68, 785–796.