Author/s (year)	[redacted]
Title (version Nº)	**Protein stability of *Pyramimonas cordata* Δ5-Elongase**
Owner	**Nuseed Pty Ltd**
Date	**December 15, 2016**
Project	**Omega 3 Canola**
Report Number	**2016-013**
Testing Facility	**CSIRO Agriculture & Food** **Queensland Biosciences Precinct** **St Lucia, Brisbane, Queensland 4067** **Australia**
Test method	**LC-MRM-MS**
GLP	**No GLP**
Confidentiality	**None**

TITLE:
PROTEIN STABILITY *PYRAMIMONAS CORDATA* Δ5-ELONGASE

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

6500 QTRAP	AB SCIEX 6500 QTRAP LC-MS/MS system
ALA	α-Linolenic acid, $18:3^{\Delta 9,12,15}$ (ω3)
CE	Collision energy
Da	Dalton
DHA	Docosahexaenoic acid, $22:6^{\Delta 4,7,10,13,16,19}$ (ω3)
DHA canola	Genetically modified canola, event NS-B50027-4
DPA	Docosapentaenoic acid, $22:5^{\Delta 7,10,13,16,19}$ (ω3)
DTT	Dithiothreitol
EPA	Eicosapentaenoic acid, $20:5^{\Delta 5,8,11,14,17}$ (ω3)
ETA	Eicosatetraenoic acid, $20:4^{\Delta 8,11,14,17}$ (ω3)
FA	Formic acid
FASP	Filter-assisted sample preparation
FDR	False discovery rate
HPLC	High performance liquid chromatography
IAM	Iodoacetamide
kDa	Kilo dalton
LA	Linoleic acid, $18:2^{\Delta 9,12}$ (ω6)
Lackl-Δ12D	*Lachancea kluyveri* Δ12-desaturase
LC-MS	Liquid chromatography-Mass Spectrometry
Micpu-Δ6D	*Micromonas pusilla* Δ6-desaturase
MMT	Million metric tons
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MW	Molecular weight
MWCO	Molecular weight cut-off
m/z	Mass-to-charge ratio
PBS	Phosphate buffer saline
OA	Oleic acid, $18:1^{\Delta 9}$
ω3 LC-PUFA	Omega-3 long-chain (≥C20) polyunsaturated fatty acids
Pavsa-Δ4D	*Pavlova salina* Δ4-desaturase
Pavsa-Δ5D	*Pavlova salina* Δ5-desaturase
Picpa-ω3D	*Pichia pastoris* Δ15-/ω3-desaturase
PMSF	Phenylmethylsulfonyl fluoride
Pyrco-Δ5E	*Pyramimonas cordata* Δ5-elongase
Pyrco-Δ6E	*Pyramimonas cordata* Δ6-elongase
Q1	Quadrupole 1 (referring to the analysis of the precursor ion)
Q3	Quadrupole 3 (referring to the analysis of the fragment ion)

RT	Retention time (min)
SD	Standard deviation
SDA	Stearidonic acid, $18{:}4^{\Delta 6,9,12,15}$ (ω3)
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SGF	Simulated gastric fluid
S/N	Signal-to-Noise
UA buffer	8 M urea, 0.1 M Tris-HCl, pH 8.5

EXECUTIVE SUMMARY

The purpose of this report was to assess the *in vitro* stability of the *Pyramimonas cordata* Δ5-elongase (Pyrco-Δ5E) protein in simulated gastric fluid (SGF) comprising the proteolytic enzyme, pepsin, and in combination with a novel pepsin-trypsin assay employing state-of-the-art mass spectrometric approaches to monitor the precise degradation products. The extent of protein digestion was evaluated by the appearance of peptic products and the disappearance of tryptic peptide products (as a proxy for intact protein). The allergenic potential of a protein is determined by a weight of evidence approach since 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.

The results of the study demonstrated that greater than 75% of the Pyrco-Δ5E protein digested within 5 min and full-length protein was rapidly digested within 60 min of incubation in pepsin producing a suite of pepsin products <3,000 Da that spanned the entire peptide sequence when analysed by LC-MS/MS.

The results of this study show that the integral membrane protein Pyrco-Δ5E was readily digestible in pepsin and/or trypsin. The Pyrco-Δ5E protein was used as the representative of the two microalgae fatty acid elongases engineered into DHA canola, for stability analysis in this report. Rapid digestion of the full-length protein is one of many factors that indicate protein safety.

I. INTRODUCTION

The omega-3 long-chain (≥C20) polyunsaturated fatty acids (ω3 LC-PUFA) eicosapentaenoic acid (EPA, 20:5ω3), docosapentaenoic acid (DPA, 22:5ω3) and docosahexaenoic acid (DHA, 22:6ω3) are widely recognised for their beneficial roles in human health, particularly those related to cardiovascular and inflammatory health. EPA, DPA and DHA are primarily sourced from wild-caught fish oils and algal oils, with algae being the primary producer in the marine food web. These sources are under pressure due to increasing demand for ω3 LC-PUFA by aquaculture, nutraceutical and pharmaceutical applications. Additional sources of these fatty acids can be produced by engineering land-based oilseed crops to convert native fatty acids to marine-type ω3 LC-PUFA which are then accumulated in seed oil. Canola is a commonly grown oilseed with 67 million metric tons (MMT) of rapeseed produced globally in 2015/16[1].

[1] http://www.ers.usda.gov/data-products/oil-crops-yearbook/oil-crops-yearbook/#World Supply and Use of Oilseeds and Oilseed Products

In collaboration with the Commonwealth Scientific and Industrial Research Organization (CSIRO), Nuseed Pty Ltd has developed genetically modified canola event NS-B50027-4 (DHA canola), which accumulates significant amounts of DHA in the seed oil.

In this DHA canola, seven fatty acid desaturases and elongases were introduced to convert oleic acid (OA) to DHA in a single pathway expression vector (Figure 1). 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). The functionalities and activities of these enzymes have been demonstrated in different heterologous expression systems (see Report Nºs 2016-005, 2016-006, 2016-007, 2016-008, 2016-009, 2016-010, 2016-011) and in transgenic Arabidopsis or Camelina seeds (Petrie et al. 2012, 2014). Based on the sequence similarity and functionality, these seven proteins can be classified into three groups, (1) yeast acyl-CoA type fatty acid desaturases including Lackl-Δ12D and Picpa-ω3D (Figure 1, blue) that introduce a double bond at the Δ12 and Δ15 positions, respectively; (2) algae fatty acid elongases including Pyrco-Δ6E and Pyrco-Δ5E (Figure 1, purple) that add a carbon to the carboxyl end of fatty acids; and (3) algae front-end fatty acid desaturases that introduce a double bond between an existing double bond and the carboxyl end of fatty acids (Zhou et al. 2007) including Micpu-Δ6D, Pavsa-Δ5D and Pavsa-Δ4D (Figure 1, green). One representative from each of these three groups was analysed for protein stability.

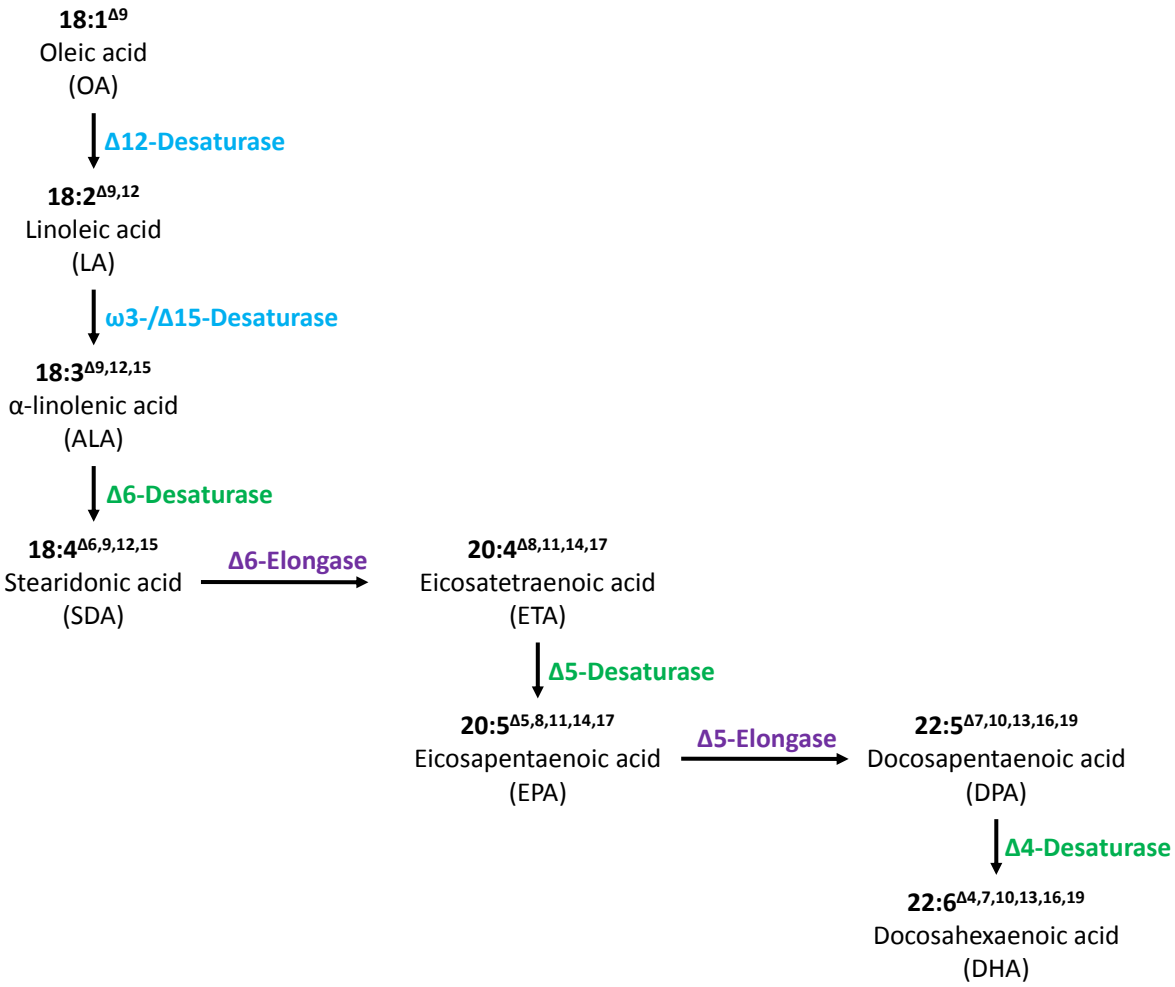


Figure 1. DHA biosynthesis pathway engineered into DHA canola event NS-B50027-4. Seven enzymes introduced in canola to convert oleic acid to final product docosahexaenoic acid were grouped into three classes: two fatty acid desaturases from yeast in blue, two elongases from microalgae in purple, and three front-end desaturases from microalgae in green (see text for detail).

II. PURPOSE

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). The purpose of this study was to assess the *in vitro* digestibility of the fatty acid biosynthesis enzymes introduced into DHA canola by digesting 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.

This particular report focuses on the microalgae fatty acid elongase representative *P. cordata* Δ5-elongase (Pyrco-Δ5E) protein, which was used in the engineering of canola NS-B50027-4 to catalyse the elongation of EPA into DPA ($20{:}5^{\Delta 5,8,11,14,17} \rightarrow 22{:}5^{\Delta 7,10,13,16,19}$)

III. MATERIALS

A. TARGET PROTEIN

The Δ5-elongase gene used in DHA canola was previously cloned from alga *P. cordata* (see Report Nº 2016-010 for details). The Pyrco-Δ5E protein was expressed as a His10::tag fusion in insect cell lines (*Sf9*) infected with *baculovirus* pFastBac vector (Invitrogen, Germany) and then purified. The vector contained coding sequences encoding a His-tag (His_{10}) and a PreScission protease cleavage site (SLEVLFQ$^{\downarrow}$GP) fused to the codon optimized *Pyrco-Δ5E* gene to produce fusion protein His_{10}::Pyrco-Δ5E.

B. OTHER MATERIALS

Sequencing grade porcine trypsin and a highly purified form of pepsin (Catalogue number V195A, ≥2,500 units/mg) were purchased from Promega (Madison, USA). Mouse anti-His antibody (Catalogue number A7058) was purchased from Sigma-Aldrich (Sydney, Australia).

IV. METHODS

A. PROTEIN EXTRACTION

The Pyrco-Δ5E protein was expressed in *Sf9* insect cell line infected with baculovirus as a fusion protein with a 10 histidine residue (His) tag at the N-terminus of the protein (His_{10}::Pyrco-Δ5E). Cells were grown by GeneArt (2L expression in *Sf9* cells infected with 1:100 virus dilution and harvested 48 h post-infection) and the thawed cells were resuspended in lysis buffer (100 mL per 20 g of cell pellet) containing 20 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol, 5 mM DTT, 5 mM EDTA, 1 mM PMSF and 2 protease inhibitor tablets per 100 mL (Roche). The cells were lysed by sonication and the cellular debris removed by centrifugation. The resultant supernatant was centrifuged at 200,000 x *g* for 60 min at 4ºC to isolate the membrane fraction. The pellet was resuspended in 50 mL 20 mM Tris pH 7.5, 150 mM NaCl, 10% glycerol, 5 mM DTT and 10 mM imidazole. To solubilise the His_{10}::Pyrco-Δ5E from the membrane fraction 1% (w/v) FosCholine-16 (Glycon Biochemicals GmbH) was added to the mixture and incubated for 2 h at 4ºC. The mixture was then centrifuged for 60 min at 200,000 x *g* at 4ºC and 10 mL Ni-Sepharose FF (GE Healthcare, Australia) was added to the supernatant and the slurry left to bind overnight at 4ºC on a wheel. After binding, the resin was poured into an empty column and washed with the binding buffer. The protein was eluted with an imidazole gradient. Fractions were analysed by sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting. Fractions containing the His_{10}::Pyrco-Δ5E were pooled and buffer exchanged into MES buffer (20 mM MES pH 6.0, 50 mM NaCl, 10% glycerol, 5 mM DTT and 0.01% FosCholine-16) using a HiPrep 26/10 desalting column (GE Heathcare, Australia). The sample was injected onto a 5mL HiTrap SP column (GE Healthcare, Australia) and eluted with a NaCl gradient. The fractions were analysed by SDS-PAGE and fractions containing the His_{10}::Pyrco-Δ5E were pooled and buffer exchanged using a HiPrep 26/10 column into phosphate buffer saline (PBS) buffer containing 10% glycerol and 0.01% FosCholine-16. The fractions containing the His_{10}::Pyrco-Δ5E were pooled and concentrated to 1.7 mg/mL and flash frozen in liquid nitrogen and stored at -80°C. Concentrated protein was analysed by SDS-PAGE and Western blotting using an anti-His HRP conjugated antibody (A7058, Sigma-Aldrich) (Figure 2). The estimated purity was ~90%.

B. METHOD DEVELOPMENT: LC-MS CHARACTERIZATION OF THE PROTEIN AFTER TRYPSIN DIGESTION

The His_{10}::Pyrco-Δ5E protein after extraction was a solution containing 1.7 mg/mL in PBS, 0.01% FosCholine-16 and 10% glycerol. An aliquot of the protein extract (equivalent to ~5 µg) was subjected to filter-assisted sample preparation (FASP) (Wisniewski et al. 2009). The protein extract was applied to a 10 kDa molecular weight cut-off (MWCO) filter (Millipore,

Australia), diluted to 200 µL with UA buffer (8 M urea, 0.1 M Tris-HCl, pH 8.5) before centrifugation (20,800 x *g*, 15 min). The protein on the filter was washed with two 200 µL volumes of UA buffer with centrifugation (20,800 x *g*, 15 min). To reduce the protein on the filter, dithiothreitol (50 mM, 100 µL) was added and the solution was incubated at room temperature for 50 min with shaking. The filter was washed with two 200 µL volumes of UA buffer with centrifugation (20,800 x *g*, 15 min). To alkylate the cysteine residues, iodoacetamide (50 mM, 100 uL) was applied to the protein on the filter with incubation for 20 min at room temperature in the dark. The filter was washed with two 200 µL volumes of UA buffer with centrifugation (20,800 x *g*, 15 min). The buffer was exchanged using 50 mM ammonium bicarbonate (pH 8.0) by two consecutive wash/centrifugation steps. Sequencing grade porcine trypsin (Promega, Alexandria, Australia) was added (0.5 µg in 200 µL of 50 mM ammonium bicarbonate, 1 mM $CaCl_2$) to the protein on the 10 kDa filters and incubated for 16 h at 37°C in a wet chamber. The filters were transferred to fresh centrifuge tubes and the filtrate (digested peptides) were collected following centrifugation (20,800 x *g*, 10 min). The filters were washed with 200 µL of 100 mM ammonium bicarbonate and the filtrate was combined and lyophilised. The tryptic peptides were resuspended in 50 µL of 1% formic acid (FA) and 25 µL was injected on the LC-MS/MS system.

C. Method Development: LC-MS Characterization of the Protein after Pepsin Digestion

An aliquot of the His_{10}::Pyrco-Δ5E protein extract (equivalent to ~5 µg) was subjected to FASP digestion. The protein extract was applied to a 10 kDa MWCO filter (Millipore), diluted to 200 µL with UA buffer before centrifugation (20,800 x *g*, 15 min). The protein on the filter was washed with two 200 µL volumes of UA buffer with centrifugation (20,800 x *g*, 15 min). 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 filter was transferred to a fresh centrifuge tube and 15 µg pepsin (150 µL, 0.1 µg/µL in 50 mM ammonium bicarbonate, pH 1.2) was added to obtain an enzyme to protein ratio of 3:1. The filters were incubated at 37°C for 120 min. The filtrate (containing the digested peptides) was collected following centrifugation (20,800 x *g*, 10 min). The filters were washed with 200 µL of 100 mM ammonium bicarbonate and the filtrates were combined and lyophilised and stored at -20°C until analysis. The resultant peptides were reconstituted in 50 µL of 1% FA of which 25 µL was analysed by LC-MS/MS.

D. Method Development: LC-MS Analysis

Proteolytically digested (either pepsin or trypsin) His_{10}::Pyrco-Δ5E protein (25 µL) were analysed as described previously (Colgrave et al, 2014) with chromatographic separation

(0.23%/min linear gradient from 2-40% acetonitrile) using a Nexera UHPLC system (Shimadzu Scientific, Rydalmere, Australia) directly coupled to a TripleTOF 5600 MS (AB SCIEX, Foster City, USA). ProteinPilot™ 4.0 software (AB SCIEX) with the Paragon Algorithm (Shilov et al. 2007) was used for protein identification. Tandem mass spectrometry data was searched against *in silico* tryptic digests of a custom-built database. The database (76,110 sequences) comprised the Noctuidae and Baculovirus proteins of the Uniprot-KB database (version 2015/11) appended with the transgenic proteins and additionally with a database of contaminant proteins (known as the common repository of adventitious proteins). The search parameters were defined as: (1) no modification to cysteine and pepsin as the digestion enzyme; or (2) iodoacetamide modified for cysteine alkylation and trypsin as the digestion enzyme. Additional modifications and cleavages were defined previously (Colgrave et al. 2014). The database search results were manually curated to yield the protein identifications using a 1% global false discovery rate (FDR) determined by the in-built FDR tool within ProteinPilot software (Tang et al. 2008).

E. Identification of Prototypic Peptides for Protein Digestibility

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 unusual or missed cleavages. For the pepsin data, peptide summaries generated by ProteinPilot were used to select peptides that yielded intense peaks after 120 min incubation with pepsin. As pepsin is non-specific, many of these peptide products were overlapping or contained missed cleavages. MRM transitions (Tables 1-2) 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 8 peptic peptides and a single tryptic peptide from His_{10}::Pyrco-Δ5E), wherein the peak area of the three MRM transitions were summed.

F. Digestibility Assay

Two test systems, pepsin digestion (representing simulated gastric fluid, SGF) and a combined pepsin-trypsin digestion, were utilized independently to test the stability of the His_{10}::Pyrco-Δ5E protein. SGF contained the proteolytic enzyme pepsin in a buffer adjusted to an acidic pH 1.2, using a highly purified form of pepsin. The SGF was formulated so that an enzyme:protein ratio of 3:1 would be present in the digestion reactions. The digestion of the Pyrco-Δ5E protein was monitored by LC-MS/MS (as described below).

G. JUSTIFICATION FOR SELECTION OF THE PEPSIN TEST SYSTEM

In vitro digestion models are used widely to assess the nutritional value of ingested proteins based on their amino acid bioavailability. The correlation between protein allergenicity and protein stability in an *in vitro* pepsin digestion assay has been previously established (Astwood et al. 1996). When proteins are found to be highly digestible, the potential for systemic exposure is reduced. The current safety assessment strategy (Codex 2003) is based on a weight-of-evidence approach that recognizes that no single endpoint can predict human allergenicity potential. Based on this strategy, a number of factors are evaluated: the gene source, determining the similarity of amino acid sequence of the newly expressed protein to known allergens, the abundance of the protein in the crop and the digestibility of the protein to *in vitro* digestion. The pepsin digestibility assay protocol used in this study was based on the protocol standardized by the International Life Sciences Institute (ILSI) in a multi-laboratory test. The results demonstrated that the *in vitro* pepsin digestion assay is reproducible when a common protocol is followed (Thomas et al. 2004).

H. JUSTIFICATION FOR DESIGN AND APPLICATION OF THE COMBINED PEPSIN-TRYPSIN TEST SYSTEM

The complete digestion of a protein by a single enzyme is difficult to judge, especially when employing a non-specific enzyme such as pepsin. While 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 2008). This IgE binding places theoretical limits on the peptide size of between 1500 and 3500 Da. Gel analysis with various staining or antibody techniques is typically able to detect peptides down to approximately 3,000 Da. 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. Likewise, cleavage of a protein at a single site in a protein may yield two protein fragments, in which one may contain the epitope (recognised by a monoclonal antibody) while the other does not. In this instance, large protein fragments may evade detection.

By using LC-MS/MS analysis, the peptide products resulting from both pepsin and trypsin digestions could first be determined qualitatively and then subsequently a quantitative LC-MS/MS 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 approach to analyse digestibility in this study mimics the typical mammalian digestive system that exposes food proteins to both pepsin (stomach) and trypsin (intestine) enzymes in transit through the gut.

I. PEPSIN DIGESTION

Aliquots of 6.7 µg of protein (67 µL, n=24 comprising 4 replicate digestions and 6 time points) were applied to a 10 kDa MWCO filter (Millipore) and diluted to 200 µL UA buffer before centrifugation (20,800 x *g*, 15 min). The protein on the filter was washed twice with 200 µL volumes of UA buffer with centrifugation (20,800 x *g*, 15 min). 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 20 µg pepsin (150 µL, 0.133 µg/µL in 50 mM ammonium bicarbonate (pH 1.2) was added to obtain an enzyme to protein ratio of 3:1. The replicate tubes were incubated at 37ºC for 5 time points (5, 10, 15, 30 and 60 mins). 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 served to irreversibly inactivate the enzyme. The 10 kDa filters were immediately centrifuged (20,800 x *g*, 15 min) and the filtrate containing digested peptides were 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 analyzed. The peptic peptides were resuspended in 50 µL of 1% FA and 20 µL was run on the 6500 QTRAP LC-MS system and quantified.

J. TRYPSIN DIGESTION

The 10 kDa filters were transferred to fresh centrifuge tubes and the residual protein 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 x *g*, 15 min). The protein was alkylated with 200 µL of 50 mM IAM, 50 mM ammonium bicarbonate (pH 8.5) in the dark for 20 min prior to centrifugation (20,800 x *g*, 15 min). The 10 kDa filters were transferred to fresh centrifuge tubes and 0.5 µg trypsin (200 µL, 2.5 ng/µL in 50 mM ammonium bicarbonate, pH 8.5, and 1 mM $CaCl_2$) was added to obtain an enzyme to protein ratio of ~1:15. The replicate tubes were incubated at 37ºC for 16 h. The filters were centrifuged (20,800 x *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 analyzed. The tryptic peptides were resuspended in 50 µL of 1% FA and 20 µL aliquots were run on the 6500 QTRAP LC-MS and quantified.

K. LC-MRM-MS Quantification of Digestion Products

Either 20µL of native peptic peptides (Table 1) or reduced and alkylated tryptic peptides (Table 2) were chromatographically separated on a Shimadzu Nexera UHPLC and analyzed on a 6500 QTRAP mass spectrometer (AB SCIEX) as described previously (Colgrave et al. 2014). Quantification was achieved using scheduled MRM scanning experiments using a 60 s detection window for each MRM transition and a 0.5 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 quantification. The graphs showing digestibility of the Pyrco-Δ5E protein were generated in Graphpad Prism v6.

V. Experimental design

A. Specificity of Proteolytic Enzymes Used in this Study

For the digestibility assay, two enzymes trypsin and pepsin were used. Trypsin is a serine protease that is found in the digestive system. Trypsin cleaves polypeptide chains at the carboxyl side of the basic amino acids lysine (K) or arginine (R), but the cleavage is hindered by the presence of proline as the preceding amino acid (P1' position, Figure 2A). Pepsin is a protease produced in the stomach and is efficient at cleaving the peptide bonds adjacent to aromatic and hydrophobic amino acids phenylalanine (F), tyrosine (Y), tryptophan (W) and leucine (L) (Figure 2B). Histidine (H), lysine (K) and arginine (R) at the P3 position act to hinder proteolysis, while proline (P) at P3 or P4 positions promotes proteolysis.

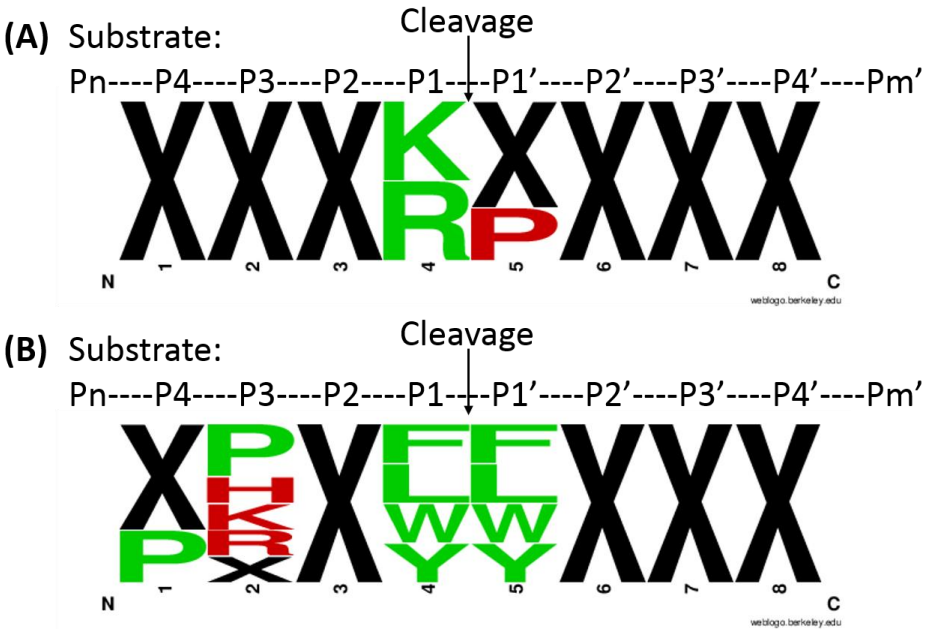


Figure 2. Specificity of proteolytic enzymes used in this study.
(A) Trypsin cleavage site. Proline (P) as the preceding amino acid (P1' position) hinders proteolysis is shown in red. (B) Pepsin cleavage sites at both sides of aromatic and hydrophobic amino acids. Amino acids that act to hinder proteolysis at the P3 position are shown in red, whereas those that promote proteolysis (P at P3 or P4) are shown in green. The images were created using WebLogo (Crooks et al. 2004).

B. THEORETICAL PREDICTION OF DIGESTION CURVES

Upon digestion with pepsin alone, there are a number of scenarios that may occur (Figure 3A). The simplest one is when the protein is rapidly digested to produce fully peptic fragments wherein the response rapidly increases reaching a maximum and creating a plateau (filled circle). The second one involves the slow digestion that does not reach a plateau within the experimental duration (filled triangles). This scenario is difficult to judge for completeness as LC-MS monitors the peptide response (peptide peak intensity or area). The third one involves a rapid, but incomplete digestion that may appear to be complete as judged by the plateau in peptide response (empty circles). Lastly, slow and incomplete digestion may be observed (empty triangles).

By employing trypsin post-pepsin (Figure 3B), 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 appear at the maximum value (in this instance as the MRM peak area). If the protein is not digested, then no decrease in peptide response will be observed (circles, dashed line). If the protein is partially digested, a partial decrease in the peptide response will be observed (squares, dotted line). If the protein is completely digested, the peptide response will drop to zero within the experiment duration (triangles, solid line).

Thus by examining the pepsin proteolytic fragments, the breakdown of a protein could be monitored, but it is noted that determining whether degradation had reached completion is a difficult task. To overcome this deficiency, the tryptic peptide products were used as a proxy for intact protein, wherein in the absence of pepsin, the amount of tryptic peptide was equivalent to 100% of protein being present. In the presence of pepsin (at varying time points during digestion), the level of tryptic peptides would be expected to decrease for peptides that contained a pepsin cleavage site. In this way the complete degradation of the protein could be monitored.

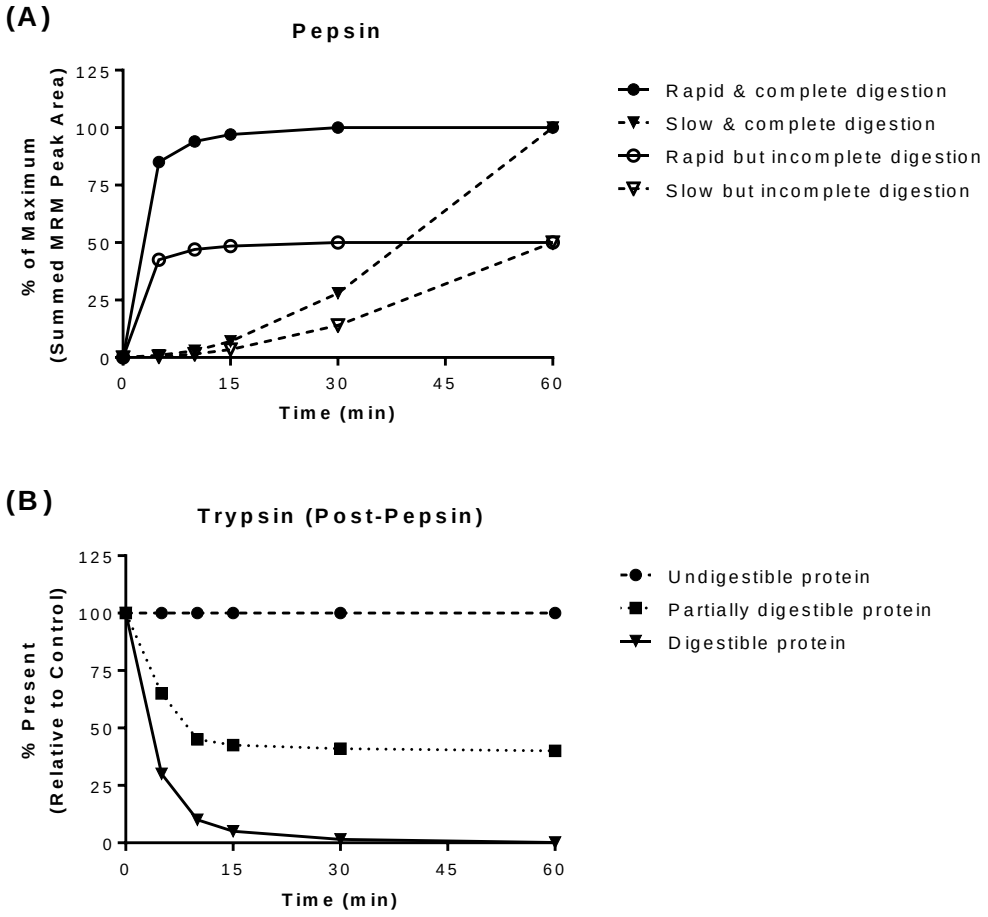


Figure 3. Theoretical digestion curves that could be generated using LC-MS and the proposed digestibility assay.
Theoretical digestion curves for pepsin (A) and trypsin post-pepsin (B).

C. SGF DIGESTION

Simulated gastric fluid (SGF) was represented by the proteolytic enzyme pepsin in a buffer adjusted to an acidic pH 1.2. The digestion was performed for 5, 10, 15, 30 and 60 min, with 0 min (no pepsin added) as the control, and each with five replicates. Due to the practical difficulty that was involved in filtering and washing after pepsin digestion with five replicates, the earliest practical time point was 5 min after the addition of pepsin. The increased abundance of targeted peptic peptides was used as indicator of the protein digestibility.

The SGF digestion was further extended by pepsin digestion at the same time point as above, followed by 16 hrs digestion with trypsin, designated as combined pepsin-trypsin digestion. The relative abundance of tryptic peptides compared to the abundance of the same peptides in no pepsin digestion (0 min) followed by trypsin digestion was used as an indicator of the protein digestibility.

VI. RESULTS

A. PROTEIN EXTRACTION

The total protein extracted was estimated to be 1.7 mg/mL. The total protein from purification was assessed by SDS-PAGE (Figure 4A). The protein was also transferred to PVDF membrane and confirmed with Western blot using an anti His_{10}::tag antibody. The expected molecular weight (MW) of His_{10}::Pyrco-Δ5E is ~33.7 kDa. A specific protein band close to 27 kDa was detected (Figure 4B). Above this band and ranging from 30-120 kDa, there was some smearing apparent in the lane with some faint bands noted within this region. This is duplicated on the Western blot with the faint bands appearing more defined at apparent molecular weights of 30, 40, 50 and 62 kDa. However, the lower than predicted molecular weight on the SDS-PAGE is a common and well documented phenomenon for membrane proteins and is caused by the presence and binding of detergents to the hydrophobic regions (Rath et al. 2009). The presence of the smearing is most likely due to similar detergent effects upon concentration of the protein in detergent micelles and larger than expected molecular weight His_{10}::positive bands could be explained by the formation of multimers of the His_{10}::Pyrco-Δ5E.

The protein was also identified/characterised by LC-MS/MS analysis. The five main bands identified in the gel and Western Blot were excised and subjected to proteolytic digestion with trypsin. All five of the bands were identified with >99% confidence as containing His_{10}::Pyrco-Δ5E (with or without minor contaminating proteins). The higher MW bands may be due to oligomerisation (dimer, trimer, hexamer) or protein-protein interactions.

Because of the difficulty of expressing and purifying membrane proteins in general in prokaryotic and eukaryotic systems, affinity tags like the histidine tag selected here, are commonly used. The insect cell/baculovirus system was selected as it has been widely used for expression of membrane proteins though the yields of the expressed protein is many folds less than that of expressed proteins from systems such as *E. coli*.

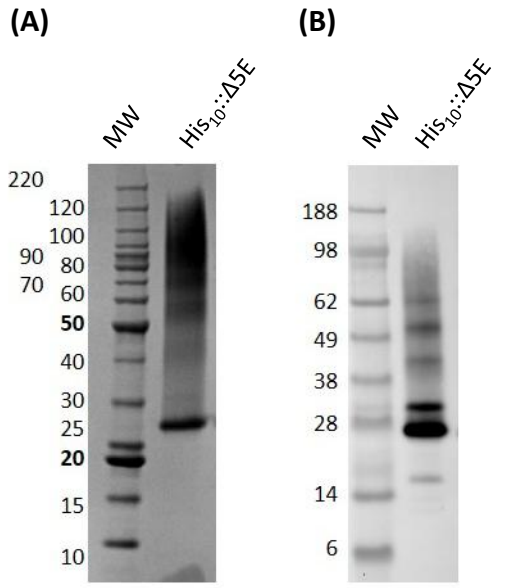


Figure 4. Characterisation of His_{10}::Pyrco-Δ5E protein expressed in baculovirus-infected insect cells.
(A) SDS-PAGE of total proteins from baculovirus infected cells. (B) Western blot analysis of His_{10}::Pyrco-Δ5E developed with anti His_{10}::tag antibody (1:1000 dilution). MW, protein markers with molecular weight indicated aside.

B. CHARACTERISATION OF THE PYRCO-Δ5E PROTEIN USING PEPSIN

As depicted in Figure 2, pepsin is a relatively non-specific enzyme which cleavages at Phe (F), Tyr (Y), Trp (W) and Leu (L) resulting in hundreds of possible peptide fragments wherein missed cleavages are commonly observed. *In silico* analysis of the Pyrco-Δ5E protein with pepsin digestion suggested the theoretical pepsin cleavage map below.

MASIAIPAA**L**AGT**L**G**Y**VT**Y**NVANPDIPASEKVPA**YF**MQVE**YW**GPTIGTIG**YLLF**I**YF**GKRI
MQNRSQP**F**G**L**KNAM**L**V**Y**N**FY**QT**FF**NS**Y**CI**YLF**VTSHRAQG**L**KV**W**GNIPDMTANS**W**GISQVI
WLH**Y**NNK**Y**VE**LL**DT**FF**MVMRKK**F**DQ**L**S**FL**HI**Y**HHT**LL**I**W**S**WF**VVMK**L**EPVGDC**YF**GSSVNT
FVHVIM**Y**S**YY**G**L**AA**L**GVNC**FW**KK**Y**ITQIQM**L**Q**F**CICASHSI**Y**TA**Y**VQNTA**FWL**P**YL**Q**LW**VM
VNM**F**V**LF**AN**FY**RKR**Y**KSKGAKKQ

Figure 5. Theoretical pepsin cleavage map.
The potential pepsin cleavage sites are indicated in bold red font. Pepsin cleaves at both the amino and carboxyl sides of the highlighted residues. The native Pyrco-Δ5E sequence of the fusion protein is shown here.

In this study, the peptide fragments of His_{10}::Pyrco-Δ5E persisting after pepsin digestion for 120 min were characterised by untargeted LC-MS/MS.

**MHHHHHHHHHHSLEVLFQGPMASIAIPAALAGTLGYVTYNVANPDIPASEKVPAYFMQVEY
WGPTIGTIGYLLFIYFGKRIMQNRSQPFGLKNAMLVYNFYQTFFNSY**CIY**LFVTSHRAQGL
KVWGNIPDMTANSWGISQVIWLHYNNKYVEL**LDTFFMVMRKKFDQLSFLHIYHHTLL**IWSW
FVVMKLEPVGDCYFGSSVNTFVHVIMYSYYGLAALGVNCFWKKYITQIQMLQFCICASHSI
YTAYVQNTAFWLPYLQLWVMVNMFVLFANF**YRKRYKSKGAKKQ**

Figure 6. Protein sequence coverage obtained after pepsin digestion.
Green = peptides identified with >95% confidence; yellow = peptides identified with 50-95% confidence; red = peptides identified with <50% confidence; grey = not detected. Wave underlined is the N-terminal His_{10}::tag and protease cleavage site followed by methionine of native Pyrco-Δ5E in the fusion protein.

C. Characterisation of the Pyrco-Δ5E protein using trypsin

Trypsin is a relatively specific enzyme that cleavages at Lys (K) and Arg (R) resulting in 22 possible peptide fragments, of which 8 were in the mass range suited to LC-MS/MS analysis (Report Nº 2016-015).

MASIAIPAALAGTLGYVTYNVANPDIPASE**K**VPAYFMQVEYWGPTIGTIGYLLFIYFG**KR**I
MQN**R**SQPFGL**K**NAMLVYNFYQTFFNSYCIYLFVTSH**R**AQGL**K**VWGNIPDMTANSWGISQVI
WLHYNN**K**YVELLDTFFMVM**RKK**FDQLSFLHIYHHTLLIWSWFVVM**K**LEPVGDCYFGSSVNT
FVHVIMYSYYGLAALGVNCFW**KK**YITQIQMLQFCICASHSIYTAYVQNTAFWLPYLQLWVM
VNMFVLFANFY**RKR**Y**K**S**K**GA**KK**Q

Figure 7. Theoretical trypsin cleavage map.
The potential trypsin cleavage sites are indicated in bold blue font. Trypsin cleaves at the carboxyl side of the highlighted residues. The native Pyrco-Δ5E sequence is shown.

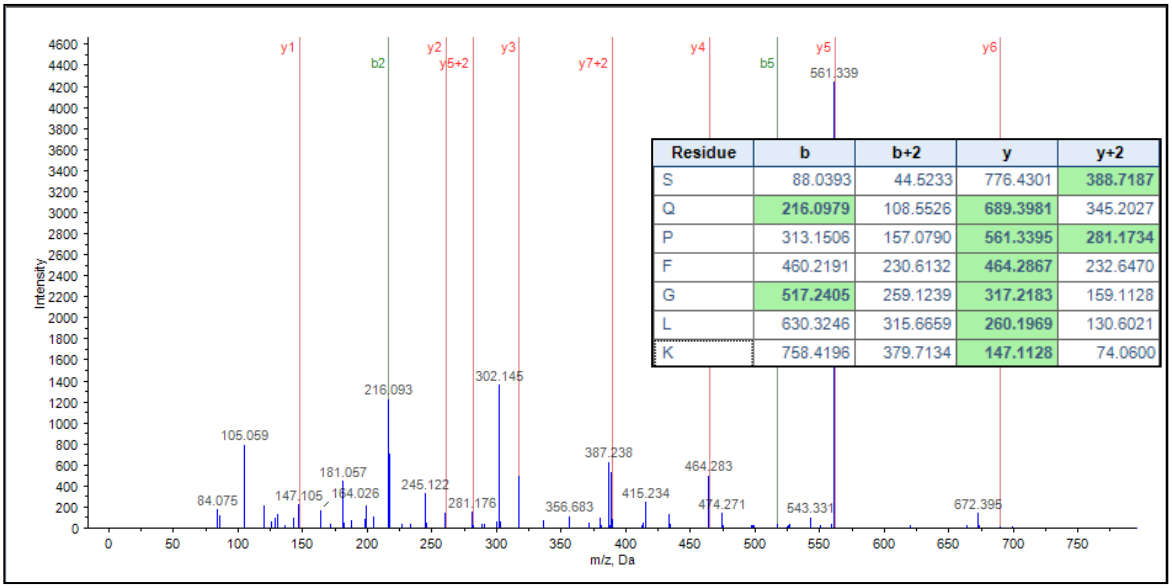
In this study, the peptide fragments present after trypsin digestion for 16 h were characterised by untargeted LC-MS/MS as shown in Figure 8. Owing to the distribution of the tryptic sites

within the Pyrco-Δ5E sequence, there were few tryptic peptides of a size amenable to LC-MS/MS. Furthermore, using a 5 µg protein load, only a single fully tryptic peptide was able to be identified with a confidence of 85%: SQPFLGK.

MHHHHHHHHHHSLEVLFQGPMASIAIPAALAGTLGY**VTYNVANPDIPASEK**VPAYFMQVE
YWGPTIGTIGYLLFIYFGKRIMQNR**SQPFGLK**NAMLVYNFYQTFFNSYCIYLFVTSHRAQ
GLK**VWGNIPDMTAN**SWGISQVIWLHYNNKYVELLDTFFMVMRKKFDQLSFLHIYHHTLLI
WSWFVVMKLEPVGDCYFGSSVNTFVHVIMYSYYGLAALGVNCFWKKYITQIQMLQF**CICA**
SHSIYTAYVQNTAFWLPYLQLWVMVNMFVLFANFYRKRYKSKGAKK

Figure 8. Protein sequence coverage obtained after trypsin digestion.
Green = peptides identified with >95% confidence; yellow = peptides identified with 50-95% confidence; red = peptides identified with <50% confidence; grey = not detected. Wave underlined is the N-terminal His_{10}::tag and protease cleavage site followed by methionine of native Pyrco-Δ5E in the fusion protein. It should be noted that the two peptides identified in green (>95% confidence) were not fully tryptic (cleaved at K/R).

The peptide-spectrum match was manually verified by *de novo* peptide sequencing. The presence of six from six possible y-ions and additionally two b-ions confirmed the peptide as confidently identified (Figure 9).



Residue	b	b+2	y	y+2
S	88.0393	44.5233	776.4301	388.7187
Q	216.0979	108.5526	689.3981	345.2027
P	313.1506	157.0790	561.3395	281.1734
F	460.2191	230.6132	464.2867	232.6470
G	517.2405	259.1239	317.2183	159.1128
L	630.3246	315.6659	260.1969	130.6021
K	758.4196	379.7134	147.1128	74.0600

Figure 9. Manual verification of peptide-spectrum match for the single Pyrco-Δ5E tryptic peptide that was identified: SQPFGLK.

D. DEVELOPMENT OF A QUANTITATIVE LC-MRM-MS METHOD TO ASSESS THE PROTEIN DIGESTIBILITY

To assess the digestibility of the His_{10}::Pyrco-Δ5E protein, a targeted LC-MS/MS method based on the use of multiple reaction monitoring (MRM) (Lange et al. 2008) mass spectrometry (MS) was developed. The appearance and the increase of the peptic peptides during the time course of pepsin digestion were used as the evidence of the protein digestibility. Moreover, the rapid decline of the tryptic peptides after the pepsin digestion served as confirmation of the protein digestibility.

In order to select peptides to quantify in this method, the digestion products resulting from both pepsin and trypsin digestion were first characterised as described previously. Pepsin-derived peptides that were identified with 95% confidence and that yielded intense signals in the MS were selected for relative quantification. The eight peptides that were selected from the pepsin digestion of the His_{10}::Pyrco-Δ5E protein and the single tryptic peptide are summarized in Tables 1-2. The selected pepsin-derived peptides spanned the length of the protein.

Table 1. Peptide MRM transitions for Pyrco-Δ5E pepsin products.

Peptide	RT (min)[a]	Q1 m/z [a]	z [a]	Q3 m/z [a]	Fragment	CE [a]
NVANPDIPASEKVPAYF	5.92	916.460	2+	1108.560 1221.650 1335.690	y10+ y11+ b13++	43.9 43.9 43.9
NVANPDIPASEKVPAY	4.73	842.930	2+	961.500 1286.660 724.360	y9+ y12+ b7+	40.3 40.3 40.3
GKRIMQNRSQPFGLKNAML	4.49	548.050	4+	600.320 807.440 746.420	b10++ b14++ y7+	25.4 25.4 25.4
GKRIMQNRSQPFGLKNAM	3.64	519.770	4+	600.320 633.330 586.350	b10++ y6+ b5+	24.0 24.0 24.0
LFVTSHRAQGLKV	3.89	485.950	3+	548.810 598.340 671.880	y10++ y11++ y12++	21.3 21.3 21.3
WGNIPDM	5.63	416.680	2+	471.240 683.310 362.130	b4+ b6+ y3+	19.4 19.4 19.4
LHIYHHTLL	3.99	573.824 382.880 382.880	2+ 3+ 3+	896.490 451.730 508.270	y7+ b7++ b8++	27.1 16.4 16.4
FVHVIM	5.20	373.207	2+	384.200 483.270 596.350	b3+ b4+ b5+	17.3 17.3 17.3

Pyrco-Δ5E sequence:[b]

MASIAIPAALAGTLGYVTY**NVANPDIPASEKVPAYF**MQVEYWGPTIGTIGYLLFIYF**GKRIMQNRSQPFGLKNAM**
LVYNFYQTFFNSYCIY**LFVTSHRAQGLKV**WGNIPDMTANSWGISQVIWLHYNNKYVELLDTFFMVMRKKFDQLSF
LHIYHHTLLIWSWFVVMKLEPVGDCYFGSSVNT**FVHVIM**YSYYGLAALGVNCFWKKYITQIQMLQFCICASHSIY
TAYVQNTAFWLPYLQLWVMVNMFVLFANFYRKRYKSKGAKKQ

a. RT, retention time (min); Q1 m/z, precursor ion mass-to-charge ratio (m/z); z, charge state; Q3 m/z, fragment ion m/z; CE, collision energy in V.

b. Pyrco-Δ5E sequence with mapped peptic peptides (bold, underlined). For pepsin, different cleavage variants were observed owing to the incomplete digestion and these peptides have been differentiated by single or double underline.

The lack of available protein required the digestion protocol to be adjusted to a smaller scale (6.7 µg load). As such, only a single tryptic peptide could be monitored (Figure 9, Table 2).

Table 2. Peptide MRM transitions for Pyrco-Δ5E trypsin product.

Peptide	RT (min)[a]	Q1 m/z [a]	z [a]	Q3 m/z [a]	Fragment	CE [a]
SQPFLGK	3.84	388.719	2+	561.340 317.200 260.200	y5+ y3+ y2+	18.1 26.0 26.0
Pyrco-Δ5E sequence:[b] MASIAIPAALAGTLGYVTYNVANPDIPASEKVPAYFMQVEYWGPTIGTIGYLLFIYFGKRIMQNR**SQPFGLK**NAM LVYNFYQTFFNSYCIYLFVTSHRAQGLKVWGNIPDMTANSWGISQVIWLHYNNKYVELLDTFFMVMRKKFDQLSF LHIYHHTLLIWSWFVVMKLEPVGDCYFGSSVNTFVHVIMYSYYGLAALGVNCFWKKYITQIQMLQFCICASHSIY TAYVQNTAFWLPYLQLWVMVNMFVLFANFYRKRYKSKGAKKQ						

a. RT, retention time (min); Q1 m/z, precursor ion mass-to-charge ratio (m/z); z, charge state; Q3 m/z, fragment ion m/z; CE, collision energy in V.

b. Pyrco-Δ5E sequence with mapped tryptic peptide (bold, underlined). For trypsin, the selected peptide was fully tryptic, *i.e.* contained no missed cleavages.

E. DIGESTIBILITY OF PYRCO-Δ5E PROTEIN

Digestibility of His_{10}::Pyrco-Δ5E in SGF was assessed by LC-MRM-MS method as described above. Characterisation and quantification of the targeted peptic peptides showed the rapid degradation of His_{10}::Pyrco-Δ5E. The pepsin digestion data has been presented in Figure 10 as the mean of four replicate digests relative percentage of the maximum detected MRM peak area (sum of three transitions) per peptide across the time points (0, 5, 10, 15, 30, 60 min).

The rapid degradation of the His_{10}::Pyrco-Δ5E protein demonstrated by the rapid liberation of peptic peptides was further confirmed by the decline of the single tryptic peptide after trypsin digestion (in the combined pepsin-trypsin digestion). Four of the peptides characterised and quantified after pepsin digestion were cleavage variants (Fig. 10A-D). The black arrows in Figure 10 indicate that the peptide in the left panel is cleaved further by pepsin to yield the peptide in the right panel. The N-terminal peptic peptides monitored were produced rapidly (<5 min) and reached an equilibrium over the experimental duration. The peptic peptides monitored may not represent the fully cleaved final product as pepsin is relatively non-specific. All of the displayed pepsin proteolysis products in Figure 10 contained missed cleavages (indicated by red font in peptide sequence) and are therefore susceptible to further degradation. The only peptide that might be considered a final product of pepsin digestion is NVANPDIPASEKVPAY wherein the lysine (K) located in the P3 position is likely to hinder further cleavage before Y (Figure 10B). This peptide reaches an equilibrium (plateau) by 15 min. The appearance of these peptides in the digest is taken as evidence of the degradation and therefore digestibility of the His_{10}::Pyrco-Δ5E protein.

In the case of the His_{10}::Pyrco-Δ5E protein, only a single tryptic product could be detected owing to the small-scale (6.7 µg load) digest and also due to the distribution of trypsin sites within the protein sequence resulting in few peptides amenable to LC-MS. The single peptide

monitored, SQPF$^{\downarrow}$GL$^{\downarrow}$K, contained two pepsin cleavage sites (as indicated by the arrows) and it was expected that pepsin would cleave this peptide resulting in a decrease in peptide abundance over the time course of the pepsin digestion. However, after 5 min the peptide peak area was noted to increase 3-fold, remain relatively constant over the next 5 min before proceeding to decline slowly over the next 50 min (Figure 11). A similar phenomenon was observed previously for a peptide derived from Δ4D wherein a 2-fold increase in the peak area of a tryptic peptide (WEGEPISK) was noted after 5 min incubation of the protein with pepsin. Both scenarios are postulated to arise from the peptides monitored residing within the core of the molecule, which in its native conformation is partially protected from trypsin digestion. After a short incubation with pepsin (5 min), the tertiary structure of the protein (Pyrco-Δ5E) is destroyed allowing full access to the tryptic sites and hence liberation of the tryptic peptide (SQPFGLK) at its maximal level (Figure 11). The absence of detectable tryptic peptides derived from the Pyrco-Δ5E protein precluded the determination of the final percentage degradation as determined for Δ4D and ω3D (see Reports 2016-012, 2016-014), however the appearance of peptic products (Figure 10) demonstrated that the Pyrco-Δ5E protein is digested by pepsin over the time course of the experiment with >75% cleavage of the N-terminal region achieved in <5 min (Figure 10A-B).

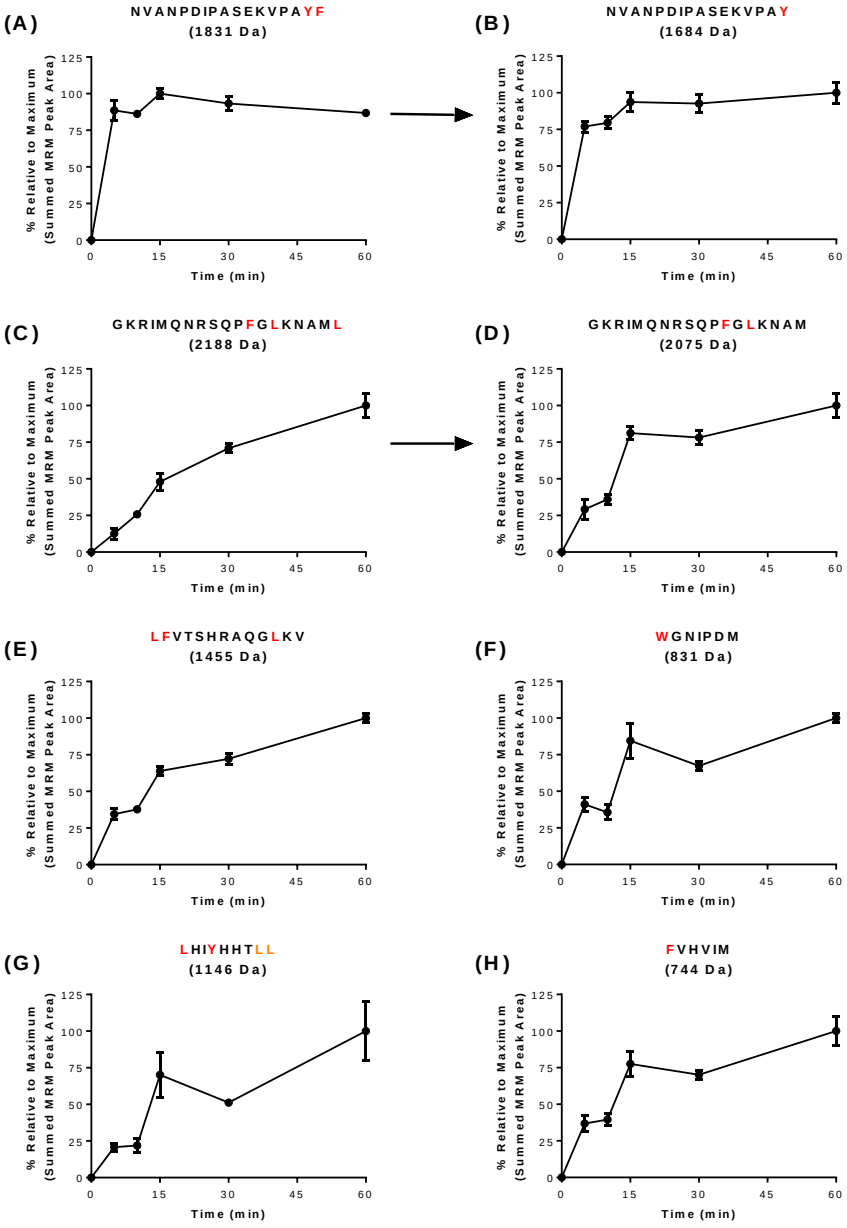


Figure 10. Quantification of the peptide products of His_{10}::Pyrco-Δ5E after pepsin digestion.

LC-MRM-MS analysis of pepsin proteolytic fragments. The response in the LC-MS system (measured as peak area) was converted to a percentage relative to the maximum peak area observed during pepsin digestion. The experimental control was time 0 with no pepsin addition. The peptides are graphed in order from protein N- to C-terminus. The peptide sequence (and calculated molecular weight) are denoted above each graph. Arrows indicate a subsequent cleavage to yield a secondary cleavage variant. The potential sites for secondary pepsin cleavage are indicated in red (expected cleavage) or orange (potentially hindered) font within the sequence. The error bars denote SD.

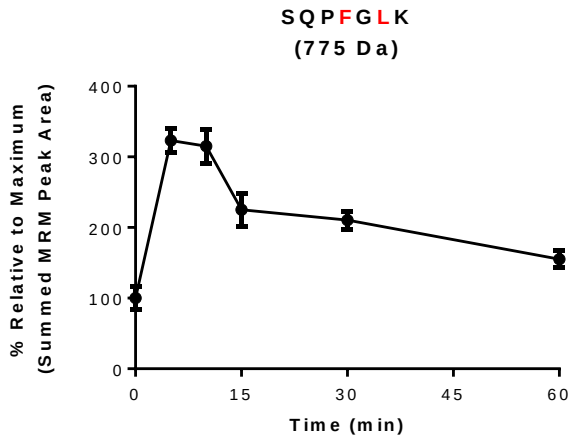


Figure 11. Quantification of the single tryptic peptide product of His_{10}::Pyrco-Δ5E after combined pepsin-trypsin digestion.
LC-MRM-MS analysis of the single trypsin proteolytic fragment that was detected. The response in the LC-MS system (measured as peak area) was converted to a percentage reduction relative to the experimental control (time 0, no pepsin addition). The peptide sequence (and calculated molecular weight) is denoted above the graph. The potential sites for pepsin cleavage of these peptide sequences are indicated in red font.

VII. DISCUSSION

Pyrco-Δ5E is an integral membrane protein. Currently there is no functional antibody for Western blot analysis available to quantify the transgenic protein content in DHA canola, event NS-B50027-4, or detect the stability of Pyrco-Δ5E as native protein. The commercially raised poly and monoclonal antibodies by GenScript (Piscataway, NJ, USA) failed to generate a specific signal towards Pyrco-Δ5E. The antibodies were raised against the synthetic peptides predicted by GenScript as potential epitopes for antigens (Figure 12).

MASIAIPAALAGTLGYVTYNVANPDIPASEKVPAYFMQVEYWGPTIGTIGYLLFIYFGKRIMQNRS
QPFGLKNAMLVYNFYQTFFNSYCIYLFVTSHRAQGLKVWGNIPDMTANSWGISQVIWLHYNNKYVE
LLDTFFMVMRKKFDQLSFLHIYHHTLLIWSWFVVMKLEPVGDCYFGSSVNTFVHVIMYSYYGLAAL
GVNCFWKKYITQIQMLQFCICASHSIYTAYVQNTAFWLPYLQLWVMVNMFVLFANFYRKRYKSKGA
KKQ

Figure 12. Peptides selected for antibody production by GenScript.
The peptides for polyclonal antibodies are highlighted in yellow, while the peptide for both polyclonal and monoclonal antibodies is highlighted in cyan.

Although Pyrco-Δ5E, expressed as the His-tag fusion protein, could be analysed by Western blot using the anti-His-tag antibody, such a Western blot analysis could only monitor the fusion region rather than whole protein, which would be problematic when the His-tag is cleaved off, for example, after SGF digestion. In addition, the anti-His-tag antibody is not suitable for quantification of the native Pyrco-Δ5E (unfused) protein in DHA canola. Thus an alternative approach using LC-MRM-MS analysis was developed here, which can be applied both the quantification of protein expressed in canola and to the stability assay. The results shown here clearly demonstrated that the LC-MS approach is suitable for such an application. This method is as sensitive as traditional Western blot, which can normally detect to a ng to µg protein scale. The LC-MRM-MS approach used was demonstrated to detect Pyrco-Δ5E levels as low as 0.08 femtomoles, which equates to 2.5 pg on a protein scale (Report Nº 2016-015). In addition, Western blot using antibodies might only detect a limited number of epitopes (one or two) from the protein. Here we targeted eight peptides, spanning the intact protein, provides an understanding of the kinetics of digestion and the susceptibility of specific regions of the protein to proteolysis. Due to the technical difficulty that was involved in the filtration and washing steps after pepsin digestion with four replicates, the earliest practical time point was 5 min. Nevertheless, the results have shown the successful application of LC-MRM-MS for protein digestibility analysis.

The thermal stability of the trans-membrane enzyme proteins, such as desaturases and elongases, cannot be easily characterized *in vitro* as it can for non-membrane bound enzymes. Traditionally the isolated protein is heated to various temperatures and its enzymatic activity is assayed to ascertain the level of functionality. A denatured or unfolded protein will lose its

enzymatic activity. Desaturase activity has been assayed in crude extracts when the required substrates are added (Jackson et al. 1998) but with DHA canola it is far more difficult because there are multiple desaturases and elongases expressed in the canola seed and the levels of the transgenic proteins in seed were as low as 20 ng per mg total protein especially for the Pyrco-Δ5E (see Report Nº 2016-015). Solubilisation using detergents to replace the lipid of the membrane and purification can increase the levels of protein but once away from the membranes the desaturase or elongase is not assayable most likely due to the requirement of other proteins co-localized in the membrane as well as cofactors, some yet unknown.

Membrane proteins, and especially trans-membrane proteins, are not very thermal stable and are difficult to refold once they are partially or fully denatured (Bowie 2001). When cells are heated, the lipid membrane becomes more fluid and exposes the hydrophobic regions of the proteins within the membrane. When exposed, the hydrophobic regions tend to lead to protein aggregation and loss of function. This is seen when preparing gel samples of membrane proteins isolated with detergents (Gennis 1989)

SDS-PAGE analysis is not useful for determining thermal denaturation of membrane proteins. It may be used to determine the amount of protein aggregating at a given temperature, with the protein sample being heated, and the aggregated protein removed by centrifugation before the remaining protein in the solution is subjected to SDS-PAGE analysis. However in the absence of detergent, all membrane proteins will be in the precipitate, and in the presence of detergent micelles, unfolded proteins could be artificially solubilised and remain in the soluble fraction, hence providing no useful information about thermal stability or activity.

In the processing of seeds to produce oil, the seed material reaches temperatures ranging from 80ºC to 115ºC. It is improbable that any trans-membrane proteins will be still folded at this temperature and even when cooled down will not refold correctly. Therefore it is expected that Pyrco-Δ5E will not be an active folded protein after processing. Based on these reasons, the thermal stability of Pyrco-Δ5E was not included in this study.

The Pyrco-Δ5E protein belongs to the subfamily of microalgae fatty acid elongases that introduce a carbon to the carboxyl end of fatty acids. The microalgae fatty acid elongases include Δ5-, Δ6- and Δ9-elongases, existing in a wide range of organisms including algae, diatom, fungi, moss and bacteria. Some of these fatty acid elongases are also common in food or in food production (see Report Nº 2016-008, 2016-010). The Pyrco-Δ5E protein was used as the representative of the two microalgae fatty acid elongases (Pyrco-Δ6E and Pyrco-Δ5E) engineered into DHA canola, for stability analysis in this report.

VIII. CONCLUSIONS

The results of this study demonstrated that the His_{10}::Pyrco-Δ5E protein was rapidly digested over the time course of the experiment with >75% cleavage of the N-terminal region achieved in <5 min. Within 60 min of pepsin incubation, a suite of pepsin products <3,000 Da were produced that spanned the entire peptide sequence.

The results of this study show that the integral membrane protein Pyrco-Δ5E was readily digestible in pepsin and/or trypsin. Rapid digestion of the full-length protein is one of many factors that indicate protein safety.

IX. REFERENCES

Astwood, JD, Leach, JN, and Fuchs, RL. 1996. Allergenicity of foods derived from transgenic plants. *Nature Biotechnol* 14:1269-1273.

Bowie, JU. 2001. Stabilizing membrane proteins. *Curr Opin Struct Biol* 11:397-402.

Codex Alimentarius Commission. 2003. Alinorm 03/34. Joint FAO/WHO Food Standard Programme, Codex Alimentarius Commission, Twenty-Fifth Session, Rome Italy 30 June-5 July, 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, pp 47-60.

Colgrave, ML, Goswami, H, Blundell, M, Howitt, CA, Tanner, GJ. 2014. Using mass spectrometry to detect hydrolysed gluten in beer that is responsible for false negatives by ELISA. *J Chromat A* 1370:105-114.

Crooks, GE, Hon, G, Chandonia, JM, Brenner, SE. 2004. WebLogo: A sequence logo generator. *Genome Res* 14:1188-1190.

Gennis, RB. 1989. Biomembranes: Molecular Structure and Function. Springer Advanced Texts in Chemistry. Charles R. Cantor (Editor). Springer Science + Business Media, LLC. Chapter 3.31 pp91-93.

Goodman, R. 2008. Performing IgE serum testing due to bioinformatics matches in the allergenicity assessment of GM crops. *Food Chem Tox*. 46:524-534.

Jackson, FM, Fraser, TC, Smith, MA, Lazarus, C, Stobart, AK, Griffiths, G. 1998. Biosynthesis of C18 polyunsaturated fatty acids in microsomal membrane preparations from the filamentous fungus *Mucor circinelloides*. *Eur J Biochem* 252:513-9.

Lange, V, Picotti, P, Domon, B, Aebersold, R. 2008. Selected Reaction Monitoring for Quantitative Proteomics: a Tutorial. *Mol Syst Biol* 4:222.

Petrie, JR, Liu, Q, Mackenzie, AM, Shrestha, P, Mansour, MP, Robert, SS, Frampton, DF, Blackburn, SI, Nichols, PD, Singh, SP. 2010a. Isolation and characterisation of a high-efficiency desaturase and elongases from microalgae for transgenic LC-PUFA production. *Mar Biotechnol* 12:430-438.

Petrie, JR, Shrestha, P, Belide, S, Kennedy, Y, Lester, G, Liu, Q, Divi, UK, Mulder, RJ, Mansour, MP, Nichols, PD.Singh, SP. 2014. Metabolic engineering *Camelina sativa* with fish oil-like levels of DHA. *PLoS One* 9: e85061.

Petrie, JR, Shrestha, P, Mansour, MP, Nichols, PD, Liu, Q, Singh, SP. 2010b. Metabolic engineering of omega-3 long-chain polyunsaturated fatty acids in plants using an acyl-

CoA Delta 6-desaturase with omega 3-preference from the marine microalga *Micromonas pusilla*. *Metab Eng* 12:233-240.

Petrie, JR, Shrestha, P, Zhou, X-R, Mansour, MP, Liu, Q, Belide, S, Nichols, PD, Singh, SP. 2012. Metabolic engineering plant seeds with fish oil-like levels of DHA. *PLoS One* 7: e49165.

Shilov, IV, Seymour, SL, Patel, AA, Loboda, A, Tang, WH, Keating, SP, Hunter, CL, Nuwaysir, LM, Schaeffer, DA. 2007. The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol Cell Proteomics* 6:1638-1655.

Tang, WH, Shilov, IV, Seymour, SL. 2008. Nonlinear fitting method for determining local false discovery rates from decoy database searches. *J Proteome Res* 7:3661-3667.

Thomas, K, Aalbers, M, Bannon, GA, Bartels, M, Dearman, RJ, Esdaile, DJ, Fu, TJ, Glatt, CM, Hadeld, N, Hatzos, C, Hee, SL, Heylings, JR, Goodman, RE, Henry, B, Herouet, C, Holsapple, M, Ladics, GS, Landry, TD, MacIntosh, SC, Rice, EA, Privalle, LS, Steiner, HY, Teshima, R, van Ree, R, Woolhiser, M, and Zawodny, J. 2004. A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Reg Tox Pharm* 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.

Wisniewski, JR, Zougman, A, Nagaraj, N, Mann, M. 2009. Universal sample preparation method for proteome analysis. *Nature Meth* 6: 359-363.

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, SS, Petrie, JR, Frampton, DMF, Mansour, MP, Blackburn, SI, Nichols, PD, Green, AG, Singh, SP. 2007. Isolation and characterization of genes from the marine microalga *Pavlova salina* encoding three front-end desaturases involved in docosahexaenoic acid biosynthesis. *Phytochem* 68: 785-796.

X. UNPUBLISHED REFERENCES

Report Nº 2016-005. [redacted]

Nuseed Pty Ltd.

Report Nº 2016-006. [redacted]

[redacted] Nuseed Pty Ltd.

Report Nº 2016-007. [redacted] Nuseed Pty Ltd.

Report Nº 2016-008. [redacted] Nuseed Pty Ltd.

Report Nº 2016-009. [redacted] Nuseed Pty Ltd.

Report Nº 2016-010. [redacted] Nuseed Pty Ltd.

Report Nº 2016-011. [redacted] Nuseed Pty Ltd.

Report Nº 2016-015. [redacted] Nuseed Pty Ltd.