Food Chemistry 424 (2023) 136459

Contents lists available at ScienceDirect

Food Chemistry

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

Nutrition, allergenicity and physicochemical qualities of food-grade protein extracts from *Nannochloropsis oculata*

Sara Hamzelou [a,*], Damien Belobrajdic [a], Angéla Juhász [b], Henri Brook [a], Utpal Bose [b,c], Michelle L Colgrave [b,c], James A Broadbent [c]

[a] CSIRO Health & Biosecurity, Adelaide, SA 5000, Australia
[b] Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, School of Science, Edith Cowan University, Joondalup, WA 6027, Australia
[c] CSIRO Agriculture and Food, St Lucia, QLD 4067, Australia

ARTICLE INFO

Keywords:
Microalgae
Nannochloropsis oculata
Proteomics
SWATH-MS
Allergens
Cross-reactivity
Diet
Protein

ABSTRACT

Microalgae offer an opportunity to act as a sustainable source of dietary protein. This study aimed to evaluate the impact of different protein extraction methods on the nutritional and physicochemical properties of *Nannochloropsis oculata.* Food-grade protein extracts were obtained by hypotonic osmotic shock using milli-Q water. Food grade (FG) and non-food grade (NFG) extraction buffers were compared along with three cell disruption methods including bead beating, probe sonication and a combination of both methods for protein extraction. Mass spectrometry was used for protein and putative allergen identification in FG extracts.

Bead beating led to a slightly higher number of identifiable proteins in FG extracts compared to control condition. Putative allergenic proteins were identified in FG extracts of *N. oculata* using different *in-silico* methods. These findings support the need to further evaluate the potential allergenic proteins in microalgae including *N. oculata* such as immunoglobulin E (IgE) binding tests.

1. Introduction

Microalgae have immense potential as an important source of sustainably produced, high-quality proteins for the human diet. The quantity and quality of microalgal proteins compare favourably with conventional protein sources (Batista, Gouveia, Bandarra, Franco, & Raymundo, 2013). Marine-water unicellular microalgae, like *Nannochloropsis* have been commercially produced, primarily for their lipid content, yet the high levels of protein and essential amino acid (EAA) enables its use in human food applications (Greenly & Tester, 2015; Kent, Welladsen, Mangott, & Li, 2015; Safi, Cabas Rodriguez, et al., 2017; Safi, Olivieri, et al., 2017; Safi et al., 2014; Spolaore, Joannis-Cassan, Duran, & Isambert, 2006). For *Nannochloropsis* species such as *N. oculata*, the robust cell wall is a key challenge in production of protein extracts. While studies to date have primarily focused on optimising the protein extraction process for bulk protein yield and estimation of total nitrogen content using Dumas (combustion) or Kjeldahl methods (Moore, DeVries, Lipp, Griffiths, & Abernethy, 2010), little is known about the quality of the protein extracts for human consumption.

A wide range of physical, mechanical, chemical, and biological cell disruption methods have been used to extract protein from *Nannochloropsis sp*. (Greenly and Tester, 2015; Safi, Cabas Rodriguez, et al., 2017; Safi, Olivieri, et al., 2017; Safi et al., 2014). The combination of multiple disruption methods has also been reported for the release of water-soluble proteins from *Nannochloropsis gaditana* (Safi, Cabas Rodriguez, et al., 2017). Mechanical procedures such as bead milling and high-pressure homogenization (HPH) have been shown to be more effective in extracting protein from *Nannochloropsis* compared to ultrasonication, enzymatic treatment, pH shift methods and pulsed electric field (PEF) (Safi, Cabas Rodriguez, et al., 2017; Safi et al., 2015; Safi, Olivieri, et al., 2017). Although limited, a few studies have also investigated the quality of protein extracted, including amino acid composition and protein bioavailability (digestibility) (El-Hawy et al., 2022; Wild, Steingaß, & Rodehutscord, 2018), In the study by Wild et al the use

Abbreviations: B, Bead beating; BP, Bead beating-Probe sonication; C, Control; FDR, False Discovery Rate; FG, Food-grade; GO, Gene Ontology; IgE, Immunoglobulin E; LHC, Light Harvesting Complex; NFG, Non-food-grade; OPLS-DA, Orthogonal Projections to Latent Structures Discriminant Analysis; P, Probe sonication; PCA, Principal Component Analysis; SWATH, Sequential window acquisition of all theoretical mass spectra; VIP, Variable importance in projection.

* Corresponding author.
E-mail address: sara.hamzelou@csiro.au (S. Hamzelou).

https://doi.org/10.1016/j.foodchem.2023.136459
Received 6 February 2023; Received in revised form 11 May 2023; Accepted 22 May 2023
Available online 24 May 2023
0308-8146/Crown Copyright © 2023 Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

of a vibrating disc mill enhanced the protein digestibility of *Nannochloropsis* by 25% (Wild et al., 2018). Yet, the previous studies lack focus on the food-grade proteome profile, which could identify proteins that may impact health, for instance, modulating immunoglobulin E (IgE)-mediated food allergies or beneficial components, e.g., bioactive peptides.

Although common food allergens are known to be present in a range of commonly consumed food products, little is known about food ingredients derived from microalgae. Recently it was reported that there are only seven articles that have identified food allergy/food hypersensitivity following ingestion *Chlorella sp*. or *Arthrospira platensis* (Spirulina) (James, Welham, & Rose, 2023). Clinical symptoms that were reported include urticaria, gastrointestinal symptoms, oedema, and anaphylaxis, (James et al., 2023). The potential allergenic proteins of *Chlorella* sp. and Spirulina were evaluated in a study by Bianco, Ventura, Calvano, Losito, and Cataldi (2022) in which shotgun proteomics and *in-silico* analysis of sequence-based homology with known allergens was used. Multiple putative allergens were identified such as thioredoxins, superoxide dismutase, and C-phycocyanin beta-subunit in Spirulina, and calmodulin and troponin C in *Chlorella sp*. Subsequently, the potential allergenic response needs to be evaluated for other microalgae species that are consumed by people.

A majority of studies reporting proteomic assessment of *Nannochloropsis* have focused on understanding factors that modulate lipid content for biofuel production, not protein levels or other food related applications (Colina, Carbó, Meijón, Cañal, & Valledor, 2020; Hounslow et al., 2021; Roustan, Bakhtiari, Roustan, & Weckwerth, 2017; Wan Razali, Evans, & Pandhal, 2022; Wei et al., 2019). Subsequently, the current study aimed to understand the proteome of *N. oculata* using a food-grade compatible protein extraction method. Various food-grade (FG) as well as non-food-grade (NFG) protein extraction methods were evaluated for their impact on proteome composition. Further processing methods on the ground (milled) microalgae, including bead beating (B), probe-sonication (P), and the combination of both procedures (BP) were evaluated. To identify and quantify all available proteins in microalgae, a NFG extraction method was also used in conjunction with three cell disruption methods, bead beating, probe sonication and BP. The properties of the extracted proteins were evaluated to identify the nutritional composition and putative allergenic proteins of this promising future protein source.

2. Material and methods

2.1. Microalgae collection

Commercial dried and ground *Nannochloropsis oculata* biomass was provided by Algae Pharm Pty Ltd (Queensland, Australia). Briefly, *N. oculata* was grown in open raceway pounds. Microalgae were harvested by membrane filtration (Liqoflux, Netherlands) followed by centrifugation to further remove water. The sample was hot air dried, and the dried sample was ground to a fine powder using a grinding mill (FA Maker, Australia).

2.2. Amino acid analysis

Amino acid content (excluding tryptophan and cysteine) was determined by acid hydrolysis, followed by liquid chromatography-mass spectrometry (LC-MS). In brief, samples containing approximately 4 mg of protein were hydrolysed in 1 mL of 6 M hydrochloric acid containing phenol and sodium thioglycolate, with 2-aminoisobutyric acid as an internal standard. Tubes were flushed with nitrogen and placed in a heat block at 110 °C for 18 h. Upon completion of the hydrolysis, samples were cooled, diluted with 9 mL of Milli-Q water, and centrifuged at 3000×g for 10 min. An aliquot of supernatant was diluted in acetonitrile and filtered using a 0.2 μm syringe filter prior to analysis.

Samples were analysed by liquid chromatography using a Shimadzu Nexera LC-40 X3 coupled to a Shimadzu LC-MS 9030 QToF mass spectrometer (Shimadzu Corporation, Kyoto, Japan). Amino acids were separated on a zwitterionic HILIC column (2.7 μm, 2.1 × 100 mm, Agilent Technologies, Santa Clara, USA) at 30 °C with 20 mM ammonium formate, pH 3.0 (mobile phase A) and 20 mM ammonium formate in 90% acetonitrile (mobile phase B). An initial flow of 100% mobile phase B was applied at 0.5 mL/min, before a linear gradient from 100 to 70% mobile phase B over 11.5 min was used. The TOF-MS scan was collected in positive mode throughout the mass range 50 – 210 *m/z*.

Cysteine content was determined by oxidation with performic acid, followed by acid hydrolysis and detection by LC-MS. Briefly, 0.5 mL of performic acid solution containing phenol was added to samples containing approximately 4 mg of protein. The tubes were sealed and incubated overnight (16 h) at 0 °C. The excess oxidant was decomposed using sodium metabisulfite and the contents of the tubes were evaporated under a flow of nitrogen at 40 °C. One mL of 6 M hydrochloric acid containing phenol was added and samples were hydrolysed in a heat block at 110 °C for 24 h. The samples were cooled, diluted with 9 mL of purified water, and centrifuged at 3,000 × g for 10 min. An aliquot of 200 μL of supernatant was diluted in 800 μL of acetonitrile, spiked with internal standard (2-aminoisbutyric acid) and filtered using a syringe filter (0.2 μm).

Samples were analysed by LC-MS using a Shimadzu Nexera LC-40 X3 coupled to a Shimadzu LCMS 9030 Q-ToF mass spectrometer (Shimadzu Corporation, Kyoto, Japan). A zwitterionic HILIC column (2.7 μm, 2.1 × 100 mm, Agilent Technologies, Santa Clara, USA) was used at 30 °C with 20 mM ammonium formate, pH 3.0 (mobile phase A) and 20 mM ammonium formate in 90% acetonitrile (mobile phase B). A linear gradient from 100 to 80% mobile phase B over 9 min was applied. The TOF-MS scan was collected in positive mode throughout the mass range 80 – 200 *m/z* to detect the internal standard, while cysteic acid was quantified by monitoring the transition 168.0 → 81.0 *m/z* in negative mode.

Tryptophan content was determined by base hydrolysis and LC-MS detection. Samples containing approximately 4 mg of protein were hydrolysed in 3 mL of 4 M lithium hydroxide containing ascorbic acid. Sample tubes were sealed and placed in a heat block at 110 °C for 16 h. After hydrolysis, the samples were cooled and neutralised with 6 M hydrochloric acid. An aliquot was filtered through a syringe filter (0.2 μm), diluted with purified water, and spiked with internal standard (Tryptophan-d5, Toronto Research Chemicals, Toronto, Canada) for quantification.

Samples were analysed by LC-MS using a Shimadzu Nexera LC-40 X3 coupled to a Shimadzu LC-MS 9030 QToF mass spectrometer (Shimadzu Corporation, Kyoto, Japan). Separation was performed using an EC-C18 column (2.7 μm, 3.0 × 150 mmm, Agilent Technologies, Santa Clara, USA) at 30 °C. A flow of 20 mM ammonium formate, pH 3.0 (mobile phase A) and 20 mM ammonium formate in 90% acetonitrile (mobile phase B) at 0.5 mL/min was applied to the column. After 2 min at 5% mobile phase B, tryptophan was eluted by applying a linear gradient from 5 to 45% mobile phase B over 4 min. The TOF-MS scan was collected in positive mode over the mass range 200 – 210 *m/z*.

2.3. Protein bound metals

The FG protein solutions were diluted twenty-fold with Milli-Q water prior to analysis by Inductively Coupled Plasma-Optical Emission Spectroscopy using Agilent 5100 (ICP-OES, Agilent Technologies, USA). The samples were aspirated into a nebulizer and ionised by an argon plasma. The non-detected ions were then analysed by Inductively Coupled Plasma-Mass Spectrometer using an Agilent 8900 (Triple Quad ICP-MS, Agilent Technologies, USA) under three different tuning conditions including no gas, hydrogen (H_2) reaction and helium (He) collision with flow rate of 1.07 L/min for H_2 and He. The detailed operating parameters can be found in Supplementary Table S1.

2.4. Protein extraction

A range of FG and NFG protein extraction methods were evaluated (Fig. 1). Two cell disruption methods including bead beating (B), probe sonication (P) and a combination of B and P (BP) as well as control (no cell disruption methods) were applied during both the FG and NFG extractions.

2.4.1. Food-grade (FG) protein extraction

Seven mL of Milli-Q water was added to 140 mg of dried *N. oculata* samples (n = 4 per treatment) to obtain 2% microalgae slurries. The slurries were incubated at 4 °C overnight (16 h) with shaking to induce osmotic shock. Microalgae slurries were centrifuged at 5250×*g* for 20 min at 4 °C. The supernatants were collected and kept at 4 °C (wash). Equivalent volume of water (7 mL) was added to the microalgae pellet and vortex mixed until homogeneous. Subsamples were taken that underwent no further treatment (Control, C) or further mechanical disruption that involved bead beating, probe sonication or BP.

For bead beating, 0.5 mm zirconium oxide beads (ZROB05, Next Advance, USA) were added to the algae slurry and cell disruption was conducted using FastPrep bead beater (FastPrep®24, MP Biomedicals, USA). The algae cells were homogenized in 10 cycles of 60 s at maximum speed (6.5 m/s). Dry ice was used during sonication to prevent heating of samples during homogenization. For probe sonication, the microalgae slurry was sonicated in 10 cycles of 20 s on ice using a laboratory ultrasonicate- model VCX 134ATA (Sonics & Materials, Newtown, CT, USA) equipped with a 1/8″ (3 mm) tip probe at 50% of amplitude for a maximum power of 300 W and 20 kHz frequency. For BP, samples were first subjected to probe sonication followed by beat beating using the same parameters as described individually for probe sonication and bead beating. The resulting protein solution from each treatment were pooled with the previous protein solution (wash) in the Falcon tubes, freeze-dried, 1 mL Milli-Q water was added and stored at −20 °C until protein digestion.

2.4.2. Non-food-grade (NFG) protein extraction

Protein extraction from algae samples was performed as described previously (Hamzelou et al., 2021). In brief, 20 mg dried microalgae was suspended in 1.5 mL of ice-cold 10% (v/v) trichloroacetic acid in acetone, 2% (v/v) β-mercaptoethanol, followed by 30 min vortex at 4 °C and incubation at −20 °C for 2 h. Cell disruption methods as described for FG extraction (including the control with no cell disruption) were applied for NFG extraction. Samples were centrifuged at 16,000×g for 30 min and washed three times with ice-cold acetone 100%. The resulting protein pellets were dried and resuspended in 3% (w/v) SDS in 50 mM Tris–HCl (pH 8.8). The protein pellet was precipitated using chloroform–methanol followed by a methanol wash step. The protein pellet was resuspended in 8 M urea in 100 mM Tris–HCl, pH 8.8.

2.5. Protein assay, and protein digestion

Protein concentration was measured using a bicinchoninic acid (BCA) assay kit (Thermo Scientific, San Jose, CA, USA). Filter-aided sample preparation (FASP) was used for protein digestion. FG and NFG protein extracts (100 μg) were transferred to 30 kDa MWCO filters (Merck Millipore, Bayswater, Vic, Australia). The protein was washed twice with 200 μL of urea buffer (8 M urea in 100 mM Tris-HCl-pH 8.8) followed by centrifugation at 14,000×g for 15 min. Protein samples were reduced with 100 μL of 10 mM dithiothreitol at room temperature for 40 min before centrifugation at 14,000×g for 10 min. Protein samples were alkylated with 100 μL of 25 mM iodoacetamide in urea buffer in the dark for 20 min followed by centrifugation at 14,000×g for 10 min. The buffer was exchanged with 100 mM ammonium bicarbonate (pH 8.0) by a series of two wash and centrifugation steps. The filters containing proteins were transferred to fresh collection tubes and digested at 37 °C overnight using trypsin (Sigma-Aldrich, NSW, Australia) at an enzyme to protein ratio of 1:50. Peptides were collected in the collection tubes by centrifugation at 14,000×g for 15 min. The filters were washed using 100 mM ammonium bicarbonate, and the resulting peptides in the filtrates were vacuum dried by rotary evaporation and resuspended in 1% formic acid. Peptide concentration was quantified by micro-BCA assay (Pierce Biotechnology, Rockford, IL, USA).

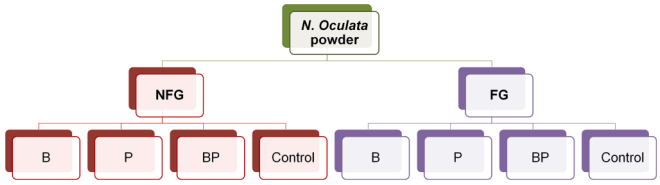

Fig. 1. Schematic overview of protein extraction methods.

2.6. SWATH-Ms

Five micrograms of peptides were analysed by microflow liquid chromatography-tandem mass spectrometry using an Ekspert nanoLC415 (Eksigent, Dublin, CA, USA) coupled to a TripleTOF 6600 MS (SCIEX, Redwood City, CA, USA). The peptides were desalted on-line on a polar C18 ProteCol trap column (Trajan; 3 μm Particle Size × 300 Å Pore Size, 10 mm × 300 μm ID) for 3 min at a flow rate of 10 μL/min of water-formic acid (999:1, v/v), switched on-line and separated on a ChromXP C18 (3 μm, 120 Å, 150 mm × 0.3 mm) column at a flow rate of 5 μL/min. The applied gradient for eluting the peptides consisted of mobile phase A (aqueous 0.1% formic acid / 5% DMSO) and mobile phase B containing formic acid-DMSO-acetonitrile (1:50:949 v/v/v). An 87 min linear solvent gradient was used, starting with 3% to 25% solvent B for 68 min followed by 25–35% B for 5 min, 35–80% for 2 min and then 3 min hold at 80% B. The gradient returned to 3% B for 1 min and 8 min of re-equilibration. The eluent from LC was coupled to the OptiFlow source of the TripleTOF 6600. The ion spray voltage - floating was set to 4500 V. The curtain gas was set to 138 kPa (30 psi), and the ion source gas 1 (nebulizer gas) and 2 (heater gas) were set to 138 and 138 kPa (30 and 30 psi) respectively. The heated interface was set to 100 °C. The TOF-MS scan was collected throughout the mass range of *m/z* 350–1250 with 150 ms accumulation time. The MS2 SWATH (Sequential Window Acquisition of All Theoretical Mass Spectra) spectra were acquired with a mass range of *m/z* 100–1500 with 30 ms accumulation time per window using optimal window boundaries as determined by the SWATH variable window calculator 1.0 (SCIEX). The window centre was used for calculating rolling collision energy (CE) using equations within AnalystTF software (SCIEX, USA) and a CE spread of 5 eV was applied to account for variances in *m/z* across the SWATH windows. A procedural blank between each sample was performed to minimise sample carry-over. An analysis of a pooled biological quality control (PBQC) sample (composed of a 1 ug sub-sample of each experimental sample) was performed prior to and periodically during the batch. Data were collected in a randomised order to reduce batch or injection order bias.

2.7. Microalgae database creation and protein identification

An *N. oculata* database was generated by combining sequences from the Monodopsidaceae family (25,763 protein sequences) with the *N. oculata* CCMP525 reference genome downloaded from UniProt KB and NCBI, respectively, in July 2022. Contigs derived from the reference genomes were used to predict open reading frames (ORFs) with a minimum cut-off of 100 bp. Using Blast2GO in OmicsBox version 2.1.1.4 (OmicsBox, Biobam), protein sequences translated from the ORFs (49,341 translated sequences) were functionally annotated and merged with the Monodopsidaceae proteins.

DIA-NN (version 1.8) was used to generate a peptide spectrum

reference library from PBQC DIA-MS data with "Deep learning-based spectra" option enabled. The generated database for *N. oculata* was used to annotate the library. Trypsin was specified as the digesting enzyme with 0 missed cleavages. Cysteine carbamidomethylation and methionine oxidation were selected as fixed and variable peptide modifications respectively. The precursor false discovery rate threshold was set to 1% and matching between runs was enabled. Protein grouping was performed by protein names and cross-run normalization was set to RT-dependent. The precursor mass range was set to 400–1250 *m/z* and fragment mass range set 200–1400 *m/z*. Precursor charge range was set to 2–3 and peptide length was set to 7–30. All other parameters were kept as default. The MS data for FG and NFG samples were processed using the spectral library generated using DIA-NN.

2.8. Data processing and analysis of abundance data

Data processing and statistical analysis of DIA-NN outputs (protein groups) was performed in Perseus (version 2.0.6.0). Output data were log2 transformed and filtered using valid values identified in three out of four replicates in at least one treatment of FG extracts. The resulting protein list was annotated as reproducibly identified proteins in FG extracts; all FG proteins were identified in the NFG extracts (Supplementary Table S2). The data was normalised by Z-score and missing values in the filtered dataset were imputed by employing the (KNN) method.

Multivariate data analysis was performed using SIMCA software (Sartorius Stedim Biotech; version 17.0) on FG protein quantitation dataset. An unsupervised Principal Component Analysis (PCA) was initially performed to visualise the data and check the groupings and variation in the dataset. A supervised Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) model was then used based on the grouping information acquired from PCA. To determine the contribution of each variable to the classification, the variable importance in projection (VIP) value was calculated for each variable. Those variables with VIP values exceeding 1.5 were considered discriminating as they contributed the most to stratification of the experimental groups.

2.9. Functional protein annotation

Both FG and NFG protein datasets were functionally annotated using Blast2GO in OmicsBox version 2.1.1.4 (OmicsBox, Biobam). Fisher's Exact Test was used to compare the FG protein list against all identified proteins (NFG). GOs with significant enrichment ($P < 0.05$) were retrieved for cell component for further analysis. Using REVIGO Supek, Bošnjak, Škunca, and Šmuc (2011) redundant GO terms were removed first and the SimRel approach was used to calculate their semantic similarities.

2.10. Physicochemical properties and subcellular localization

Several physicochemical properties of proteins such as charge, theoretical isoelectric point (pI), Hydrophobic moment (Hmoment), etc., were predicted from the amino acid sequence of the identified proteins using the Peptides R package (Team, 2010). Physicochemical properties were compared between FG and NFG proteins using two-sample Student's *t*-tests.

Subcellular localization prediction of proteins was performed using DeepLoc-2.0 (https://services.healthtech.dtu.dk/service.php?DeepLoc-2.0).

2.11. Bioinformatic analysis of allergen proteins

Four different methods were used to identify potential allergenic proteins in *N. oculata*. In all methods, identified proteins in FG extracts were used to isolate potential allergen proteins. This work was conducted using CLC Main Workbench version 22.0.2. Methods included: homology searches against the 1- Allergome database (https://www.allergome.org); 2- AllFam database (AllFam, Medical University Vienna, Austria; https://www.meduniwien.ac.at/allfam/); and, 3- the AllerCatPro online tool (Maurer-Stroh et al., 2019). Sequences were also analysed using 4- allergen linear epitopes retrieved from the Immune Epitope Database and Analysis Resource (IEDB, https://www.iedb.org/). In the first 3 approaches sequence homology higher than 70% between algal proteins and known allergens were considered as more than 70% identity in the primary sequence is considered consensus viewpoint for clinical cross-reactivity (Bublin & Breiteneder, 2020). In the first two methods, the microalgae protein lists were searched against Allergome reviewed allergenic proteins which were retrieved from the UniProt KB as well as the AllFam database. Those proteins with sequence homology higher than 70% and an E-value less than e^{-10} were assumed as allergenic proteins. AllFam merges allergen data from the WHO/ IUIS Allergen Nomenclature Database, supplemented by data from AllergenOnline, along with protein family definitions obtained from the Pfam database (Radauer, Bublin, Wagner, Mari, & Breiteneder, 2008).

The AllerCatPro online tool was chosen as the third method to uncover the potential allergenicity of microalgae proteins. AllerCatPro procures data pertaining to allergenic proteins from multiple databases including WHO/International Union of Immunological Societies (IUIS), Comprehensive Protein Allergen Resource (COMPARE), Food Allergy Research and Resource Program (FARRP), UniProtKB and Allergome. This tool uses the similarity between input proteins and the allergen datasets using their amino acid sequences and predicted 3D structures to identify proteins with high allergen potential (Nguyen, Krutz, Limviphuvadh, Lopata, Gerberick, & Maurer-Stroh, 2022). In the fourth method, linear epitopes with confirmed positive assay results and a minimum length of six amino acids were retrieved from the IEDB. The linear epitope search was conducted on the identified proteins via motif search with 100% sequence identity and proteins with matching epitopes were considered potential allergens.

2.12. Statistical analysis

A series of Student's *t*-tests were conducted to determine the differences in the presence of metal ions among various disruption techniques in FG extracts. The difference in abundance of identified proteins and putative allergen proteins between FG than NFG extracts were also evaluated by Student's *t*-test. Significant differences were defined as those with p-values < 0.05.

3. Results and discussion

3.1. Nannochloropsis protein extracts are beneficial for health

The amino acid content of *N. oculata* was 23.6% of the dry weight and 41.5% of its total amino acid content was essential amino acids (EAA) (Supplementary Table S3). The EAA content of *N. oculata* is lower than egg or soy, as reported by Gorissen et al. (2018). Methionine (Met) was the only EAA that was substantially higher in microalgae (0.6 g/100 g dried material) than soy (0.3 g/100 g dried material). However, less distinct variability was observed in isoleucine (Ile) between microalgae and both plant-based and animal-based reference sources.

Previous studies demonstrated that *Nannochloropsis* intake is beneficial for health as it is a good source of essential fatty acids, some minerals, and EAAs (El-Hawy et al., 2022; Wild et al., 2018). However, the reported values for minerals and amino acids vary considerably between studies, likely due to varied growing conditions.

This may explain the difference in amino acid composition of *N. oculata* samples in this study that were grown in an open pond system than the previous reported results where samples were grown in bioreactors under different growth conditions (Brown, 1991; El-Hawy et al., 2022).

The major electrolytes identified in FG protein extracts were Na, K,

Ca, Mg alongside the anion S (Supplementary Table S4), The presence of these was not significantly different for the sample extraction methods tested. Metals, including Fe, were present at levels lower than 100 μg/L, with only Zn, Sr, Mn, and Ni ions present at higher levels in *N. oculata* samples that were processed using bead beating or BP compared to probe sonication or control (p-value < 0.05).

Previous studies have shown that sodium was the highest mineral ion present in *N. oculata* (1.9% or 1862.70 mg/100 g) followed by potassium (0.8% or 798 mg/100 g) (El-Hawy et al., 2022). A consistent result with previous investigation of *N. oculata* whole biomass mineral quantification (El-Hawy et al., 2022) was acquired for the FG protein extracts. The consumption of 3,500 mg (90 mmol) of K per day has been shown to have favourable effects on blood pressure in adults (Efsa Panel on Dietetic Products et al., 2016). The similar trend of metal ion presence in both *N. oculata* FG and whole biomass may be due to the presence of biomolecules other than proteins such as DNA, RNA, and free metal ions bound to water molecules in aqueous cellular environment (Dean, Qin, & Palmer, 2012). Therefore, the FG protein extracts contain nutritional elements such as potassium, magnesium, and calcium.

The heavy metal content of algae is an important component that needs to be determined to ensure that levels, are below a daily maximum safe exposure levels such as 0.3 μg/kg/day for As, 0.26 μg/kg/day for Pb in young children and 0.16 μg/kg/day for older children (Wong, Roberts, & Saab, 2022). For chromium (Cr) the maximum safe daily exposure level is 300 μg/kg/day (Efsa Panel on Dietetic Products & Allergies, 2014). Metal ions including, Cr, As and lead Pb ions were at levels considered safe for human consumption in FG algae extracts (Wong et al., 2022).

3.2. Proteome characterisation of protein extracts

The DIA-NN analysis of *N. oculata* identified 1,373 proteins and protein groups in the NFG protein extracts and 464 in FG extracts in three out of four replicates (Fig. 2a). Using only Milli-Q water as an FG method of protein extraction, we are the first study to show that a total of 338 proteins were identified in FG extracts of *N. oculata* without additional cell disruption methods (C). This number of proteins is similar to a study by Wan Razali et al. (2022) who identified 400 proteins when protein was extracted from *N. oculata* when a 2% sodium dodecyl sulphate (SDS), 40 mM Tris base, and 60 mM dithiothreitol (DTT) extraction buffer was used.

In NFG protein extracts, all cell disruption methods showed a similar yield of total identifiable proteins. However, for FG extraction methods, bead beating resulted in a higher number of proteins compared to the control (p-value < 0.05). Heatmap and hierarchical clustering of the relative abundance of FG proteins (464 proteins) indicated two major groupings of FG and NFG extracts (Fig. 2c). Clustering of those proteins revealed that bead beating, and BP clustered together and separated from the clusters of control and probe sonication in both FG and NFG extracts. Thus, probe sonication has limited additional effect on protein extraction compared to control, whereas bead beating increased the number of different proteins extracted.

Unsupervised PCA assessment of FG and NFG protein extracts confirmed the similarity of proteins extracted using bead beating and PB, which were separated from control and probe sonication. Bead beating and BP samples in NFG solvents and control and probe sonication in both FG and NFG solvents were grouped tightly together (Fig. 3a). Although bead beating and BP in FG extracts were not grouped as closely as control and probe sonication, they still showed an

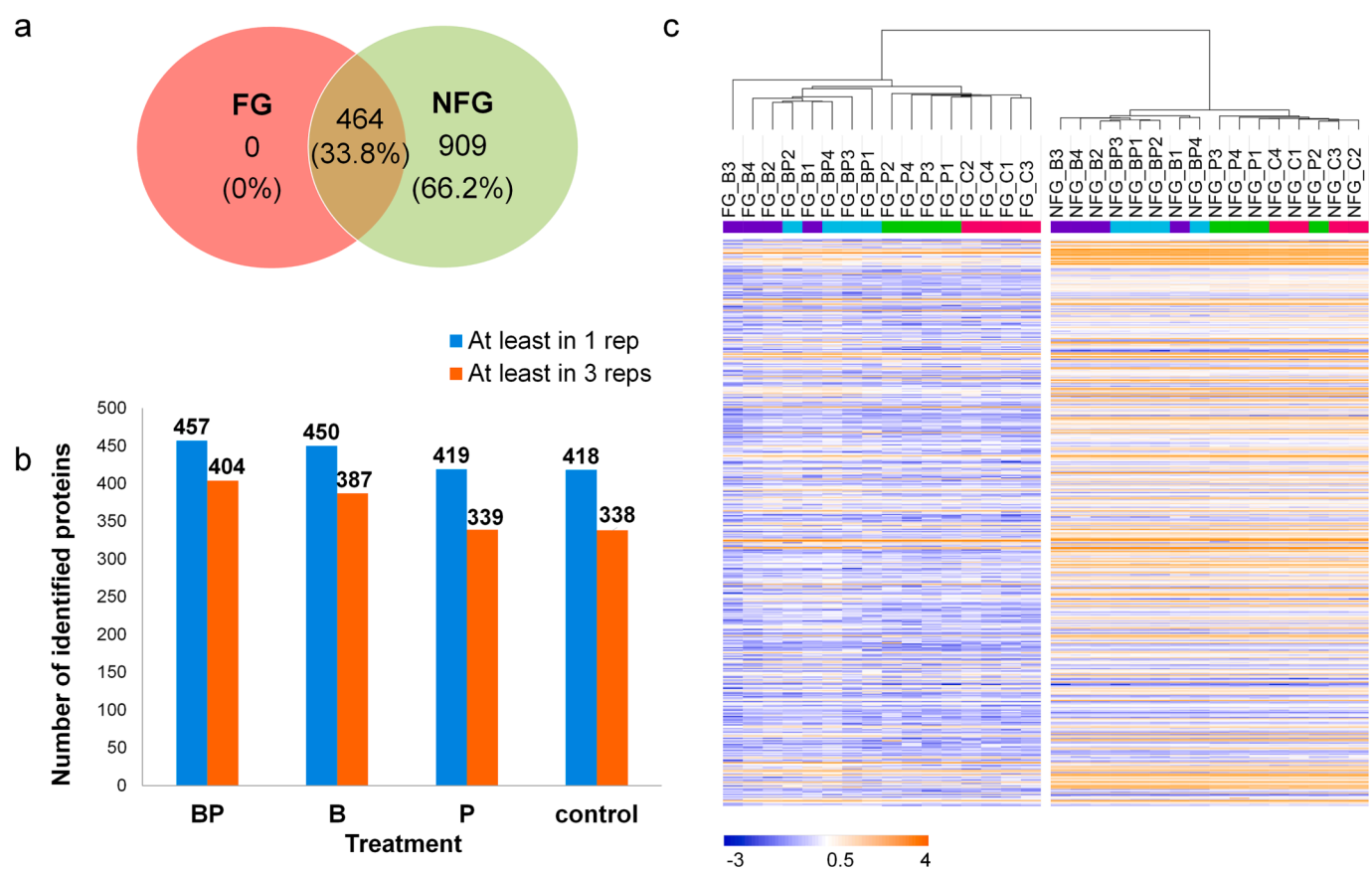

Fig. 2. Profiling of proteins identified in *N. oculata*: (a) Venn diagram of reproducibly identified proteins in food-grade (FG) and non-food-grade (NFG) extracts (b) Bars illustrate the number of identified proteins in different cell disruption methods in FG protein extraction method. (c) Heatmap showing the relative abundance of the protein identified in the FG extract across all FG and NFG samples when bead beating (B), probe sonication (P) and both B and P (BP) were used.

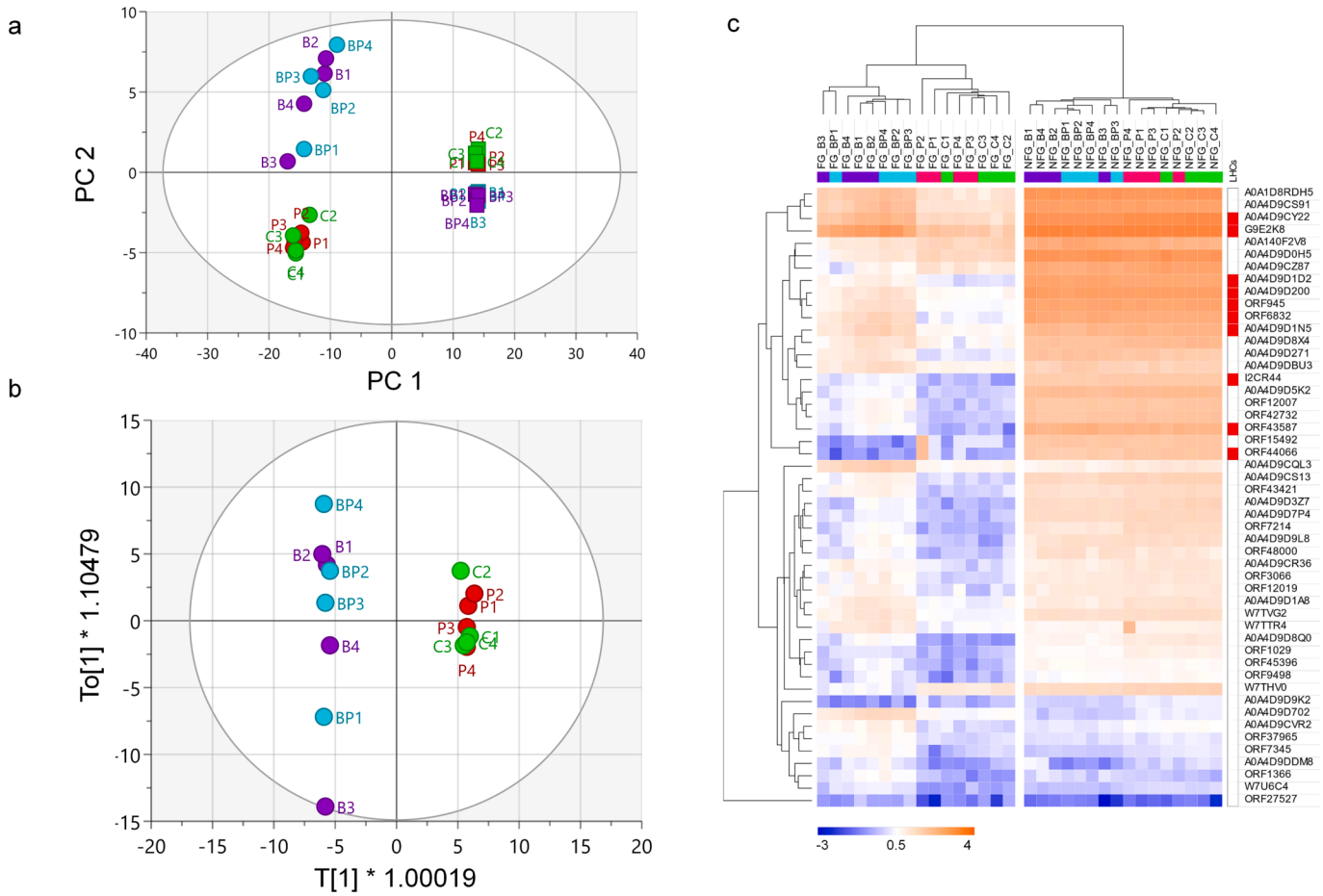

Fig. 3. Overview of the *N. oculata* proteome dataset (a) PCA scores plot indicates the separation of the protein abundance data for different treatments applied to FG and NFG extracts. Each circle and square represent a replicate from FG and NFG treated with different cell disruption methods, respectively. (b) The OPLS-DA score scatter plots of the first group consisting of C and P and second group consisting of B and BP in FG extract proteins. (c) Heatmap showing the relative abundance of the top-50 most discriminating proteins in FG. The red boxes shown in a column on the right-hand side of the heatmap illustrate the presence of light harvesting complex proteins (LHCs) proteins. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

association. Overall, the results indicate that bead beating was responsible for stratifying the experimental groups in FG and NFG; probe sonication did not contribute to the extraction of a higher number of proteins than the control conditions.

3.3. FG proteins discriminate between cell disruption methods

A supervised OPLS-DA model was generated to identify proteins responsible for stratifying FG sample groups. Based on the PCA results, control and probe sonication groups were inseparable. Thus, we classified control and probe sonication as one group and bead beating and BP as another. The OPLS-DA model was performed using bead beating and BP combined in class 1 and control and probe sonication combined in class 2 (Fig. 3b). Subsequently, VIP analysis identified the 50 most discriminating proteins (VIP > 1.5), that differed in abundance for C and P treatments.

Fig. 3c shows the profile of relative protein abundance for the most discriminating FG proteins. The abundance of these proteins is also illustrated for the NFG extracts. Protein abundance was lower in control and probe sonication protein extracts compared to bead beating and BP in FG extracts. A total of 10 proteins (20%) were categorized as light harvesting complex proteins (LHCs). This was confirmed by gene ontology (GO) analysis showing the enrichment of LHC proteins as important discriminators (light harvesting GO category (GO:0009765)) as described in section 3–4. All LHC proteins were more abundant in NFG extracts than FG extracts (Fig. 3c). While no visible differences were identified in the number of proteins using different mechanical procedures in NFG extracts, bead beating enriched LHC proteins in FG extracts.

The instrumentation, time of exposure and intensity of disruption methods employed in this study were different than those used in previous studies on *Nannochloropsis* strains (Safi, Olivieri, et al., 2017; Safi et al., 2014). Using an industrial scale food grinding machine was sufficient to break the cell wall of dried microalgae samples as additional disruption methods did not enhance the amount of protein extracted by the FG or NFG solvents. To minimise the protein degradation and protease activity prior to mass spectrometry analysis of microalgae proteins, all further processing methods were conducted below 4 °C. While time of exposure to disrupt the cells was increased in previous studies regardless of the significant raise of the temperature in microalgae slurries. For example, *Nannochloropsis* samples were bead milled for one hour at a temperature below 35 °C as described by Safi, Olivieri, et al. (2017), while we used FastPrep bead beater and maintained the temperature at 4 °C throughout the entire protein extraction process.

As *Nannochloropsis* protein extracts could potentially be used in human food applications, it is necessary to consider the ease of extraction and the cost of processing such as energy requirements to extract proteins (Soto-Sierra, Stoykova, & Nikolov, 2018). The FG-B method was identified as the most optimal FG method of protein extraction as it resulted in a 16% greater number of proteins identified compared to FG-C, yet there was no significant difference in the protein concentration based on BCA protein assay (Supplementary table S5). As the FG method

involves the addition of water to induce hypotonic osmotic shock this is one of the most cost-effective protein extraction methods possible.

3.4. Functional analysis of algal protein extracts: membrane proteins such as LHCs are still highly abundant in FG extracts

After retrieving gene ontology (GO) categories in FG extracts and NFG extracts, a Fisher's exact test was performed to identify which GO categories in cellular components are enriched in FG solvents compared with NFG solvents (P-value < 0.05). The enrichment analysis also enabled us to determine whether specific GO categories were enriched in the most discriminating proteins. Functional enrichment analysis revealed that the cellular component chloroplast thylakoid (GO:0009534) such as plastids and thylakoid membrane were enriched in FG extracts (Fig. 4a). Membrane protein complex (GO:0016020), integral component of plasma membrane (GO:0005887) and envelope (GO:0031975) were also enriched in the FG extract which are composed of water-soluble proteins.

Membrane proteins were not significantly over-represented in NFG when compared to the FG solvent extraction, which was not expected. This may be explained by the incomplete genome sequence data in *N. oculata* resulting in poor membrane protein peptide spectrum matching. In this regard, some sequences in our database were acquired from other species in the Monodopsidaceae family. With reports indicating that genes encoding for membrane proteins may evolve faster than those encoding water-soluble proteins in *N. oculata*, there may have been a particular loss of PSMs for this class of proteins (Sojo, Dessimoz, Pomiankowski, & Lane, 2016).

The gene ontology results revealed that plastids and thylakoid membrane were enriched in the proteins discriminating FG from NFG extracts. The identification of 10 LHC proteins out of the top-50 discriminating proteins led to the significant over-representation of GO biological process light harvesting (GO:0009765) (Fig. 4b). The molecular functions contributing to transport, including export across the plasma membrane (GO:0140115), and export across cells (GO:0140352) were also significantly enriched in these proteins.

Using our method of FG protein extraction, salts from microalgae are introduced to the aqueous solvent (Milli-Q water) used for protein extraction. Salts can affect the solubility of proteins in aqueous solution as well as their structure, function and dynamics through the Hofmeister effect (Collins & Washabaugh, 1985). So, we assume salt in the microalgae growth media may have aided the solubility of the hydrophobic thylakoid membrane proteins as well as other membrane proteins in water. The use of mechanical processes has been suggested to aid envelope membrane disruption and further release thylakoids and stroma into solution (Rolland et al., 2006). This may elucidate the role of bead beating in thylakoid membrane disruption to release a higher number of LHC proteins when compared to the control condition. Salt treatment followed by sonication was used to extract and solubilise the membrane proteins in a proteomic study of chloroplast envelope membranes in *Arabidopsis thaliana* (Ferro et al., 2003). This may provide an additional avenue to explore for more efficient protein extraction.

3.5. Identification of putative allergens in FG protein extracts

Proteins detected from four different methods were used to identify potential allergen proteins in both FG and NFG extracts as described above. Using immunoinformatic analysis and a combination of four allergen DBs and prediction tools, we identified 31 putative allergens in FG protein extracts. Considering the consumption of FG extracts as food ingredients, the results are only presented for FG extracts. Fig. 5a represents the distribution of evidence across the four analysis methods. Of the total identified putative allergen proteins, 58% and 16% were uniquely identified using AllerCatPro web interface and, the IEDB epitope method respectively. The putative allergen proteins identified using the Allergome and AllFam were also identified by the AllerCatPro method. The information related to sequence identities of the microalgal proteins in relation to the proposed cross-reactive allergens has been provided in Supplementary Table S6.

Fig. 5b shows the abundance of putative allergens identified in FG as well as NFG extracts. A substantial number of putative allergens (70%) are considered as non-food allergen proteins, as their allergenicity is caused by inhalation or skin contact, so their ingestion is not likely to trigger allergenicity. However, contact or inhaled allergens may raise concerns for those who are involved in manufacturing and handling of *N. oculata* products.

Using AllercatPro, we could identify three algal proteins including A0A4D9CQ65, A0A4D9D8Z9 and A0A4D9DDM1 as potential allergens with high homology to known fish allergens i.e., nucleoside diphosphate kinase (G8DZS2 Gad m NDKB) in *Gadus morhua* (Atlantic cod). The

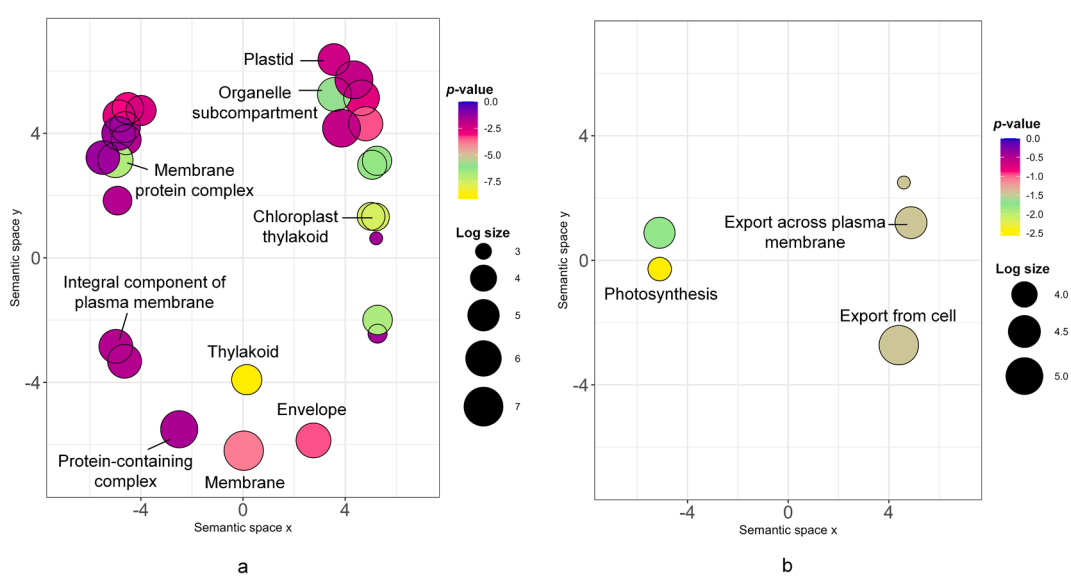

Fig. 4. Gene ontology terms enriched ($P < 0.05$) in (a) food-grade-extract proteins and (b) The top 50 most discriminating proteins using REVIGO. The bubble graph applies multidimensional scaling that relies on the semantic similarity measures between the GO terms. Bubbles indicates the enriched GO terms with different colours representing the Fisher's exact test p-values (log10), and different sizes showing the frequency of the GO terms in the EBI GOA database. The key GO terms as the cluster representatives have been illustrated (dispensability < 0.3) (Supek et al., 2011).

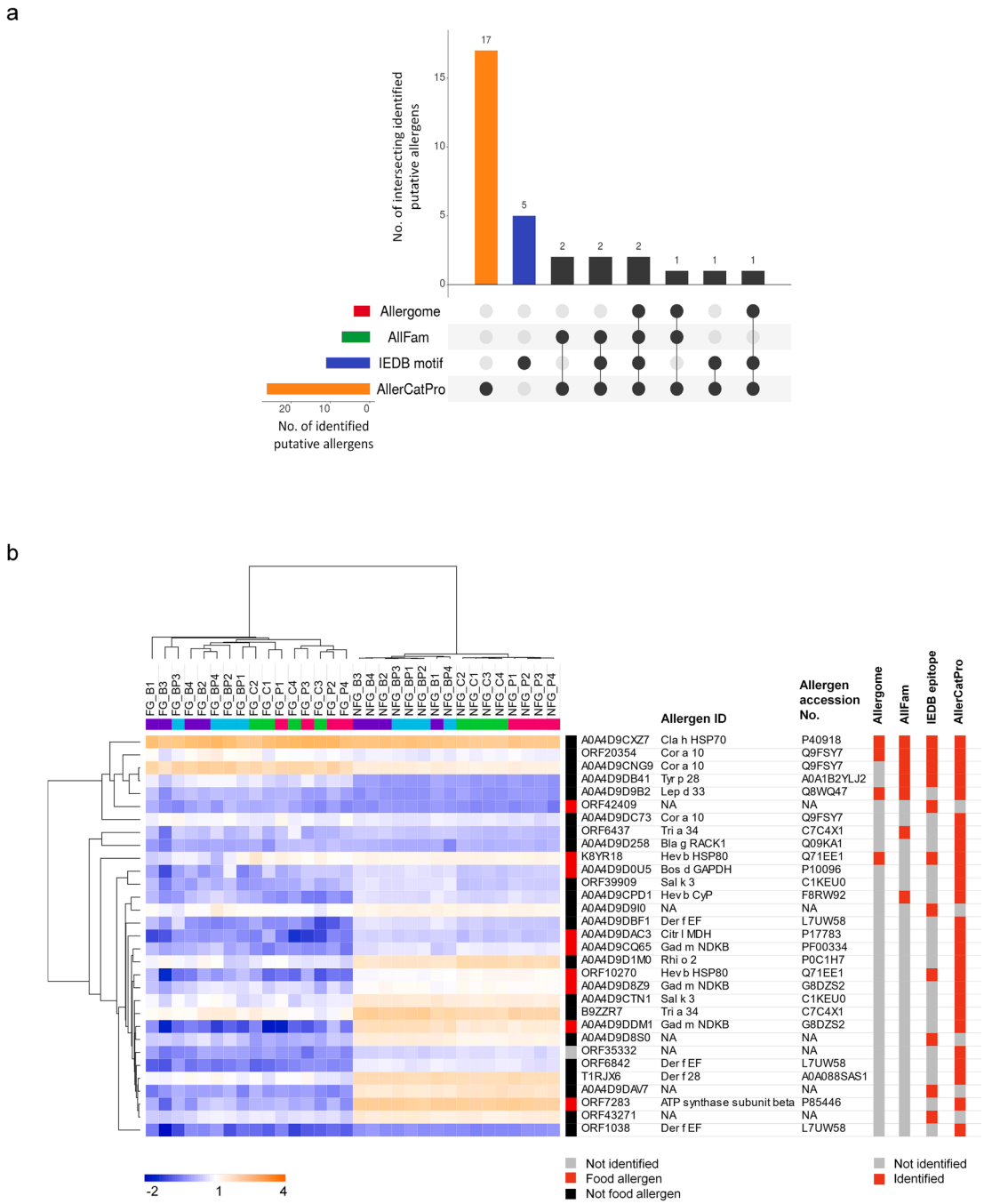

Fig. 5. Allergen proteins identified in food-grade (FG) and Non-food-grade (NFG) extracts using different *in-silico* approaches. (a) UpSet plot summarising the distribution of evidence for putative allergen proteins in FG extracts. The horizontal bar graph illustrates the number of putative allergen proteins identified using the four approaches, including Allergome, AllFam, IEDB epitope, and AllerCatPro. Filled circles and vertical lines correspond to the co-identification of allergen proteins by multiple methods. (b) Heatmap showing relative abundance of putative allergen proteins in FG and NFG for the 47 proteins. Inhalation/contact related reference allergens are indicated by black and proteins that were identified only by IEDB epitope are indicated by grey blocks. Putative allergen identification is indicated by red blocks under the four *in-silico* approaches for each protein. Reference allergens information including UniProt IDs and the name of organisms are shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

alignment of the microalgae proteins with nucleoside diphosphate kinase in both their linear and 3D structures is shown in Supplementary Fig. S7. IgE-reactivity to nucleoside diphosphate kinase B was identified in one individual out of twelve who had previously who had previously experienced an allergic reaction to fish (Tomm et al., 2013). The abundance of all putative fish allergen proteins is significantly lower in FG extracts than NFG ones (p-value < 0.05).

Several inhalation-related and contact allergens were identified with high homology to known plant and insect allergen sources. Using the IEDB method, triosephosphate isomerase (TIM) (A0A4D9DAV7) was identified as a potential allergenic protein (Supplementary Table S6). The linear epitope WSKVVIAYEPVWAIG was studied as part of triosephosphate isomerase from *Blattella germanica* (German cockroach) by Dillon et al. (2015) which is associated with asthmatic reactions.

However, TIM was introduced as a food allergen in plant-derived foods, and fish (Yang et al., 2017) and was introduced as a potential allergenic protein in the *in-silico* study on Spirulina (Bianco et al., 2022). Due to the homology of three algal proteins annotated as heat shock protein 70 (HSP-70) proteins (A0A4D9CNG9, A0A4D9DC73 and ORF20354) with HSP-70 (Q9FSY7- Cor a 10) in hazelnut, these can be categorized as putative inhalation-related allergens. A putative algal allergen was identified with homology with cyclophilin (F8RW92–Hev b CyP) that is recognized as a contact allergen in para rubber tree. Three proteins were also identified in *N. oculata* with homology with an inhaled allergen that is an elongation factor (L7UW58– Der f EF) in house dust mites. The complete list of potential contact/inhalation cross-allergens are shown in Fig. 5b.

Numerous allergen databases are accessible, providing a variety of beneficial information, comprising but not restricted to biochemical functions, allergen origins, sequence homologies, cross-reactive allergens, and protein families (Kleine-Tebbe, Ollert, Radauer, & Jakob, 2017). Our computational analysis of potential allergenic proteins was enhanced by employing multiple databases, resulting in a more comprehensive identification of putative allergens that can be utilised in future research on allergenic proteins in *N. oculata*.

There is a lack of research on identification of microalgae allergen proteins with supporting experimental data on their allergenicity such as serum IgE binding tests. Only a few studies have investigated the allergic reactions after microalgae intake. Besides, these studies are limited to microalgae products currently in the market, including *Chlorella* sp. and Spirulina. Allergic reactions in children (Tiberg, Dreborg, & Björkstén, 1995) and gastrointestinal problems such as nausea and vomiting have been reported for excessive intake of *Chlorella sp.* (Barkia, Saari, & Manning, 2019). This suggests that there may be issues — such as allergic reactions — from the consumption of this algae species. In a study by Pimblett (2020), patients with mild oral reactions to some fresh fruits and raw vegetables developed allergic reactions to *spirulina* intake. Thus, immune cross-reactivity assays along with clinical studies needs to be considered especially with an anticipated increase in consumption of a broader range of algae species.

3.6. Subcellular localization and physicochemical properties of proteins

To uncover the ability of extraction buffers to access the cellular components, subcellular localization of proteins identified in FG and NFG extracts as well as all putative allergens were predicted using DeepLoc-2.0. (Thumuluri, Almagro Armenteros, Johansen, Nielsen, & Winther, 2022). A higher percentage of FG and NFG extracted proteins were identified in plastids and cytoplasm/extracellular space than other subcellular components (Fig. 6a). A small proportion of proteins were predicted to be localized in the cell membrane in both FG and NFG extracts. Fisher's exact test was performed to identify whether FG or putative allergen proteins were significantly over-represented in cell components in comparison to NFG extracts (P-value < 0.05). Our results showed that putative allergen proteins were significantly over-represented in cytoplasm and plastids.

A range of physicochemical properties such as pI (isoelectric point), instability index, hydrophobicity, charge and Hmoment (Hydrophobic moment) were examined for FG and NFG extracts. Proteins identified in FG extract were significantly different in pI, Hmoment and charge (p-value < 0.05) but not in other estimated physicochemical properties from NFG extracts (Fig. 6b). The physicochemical profiling of FG

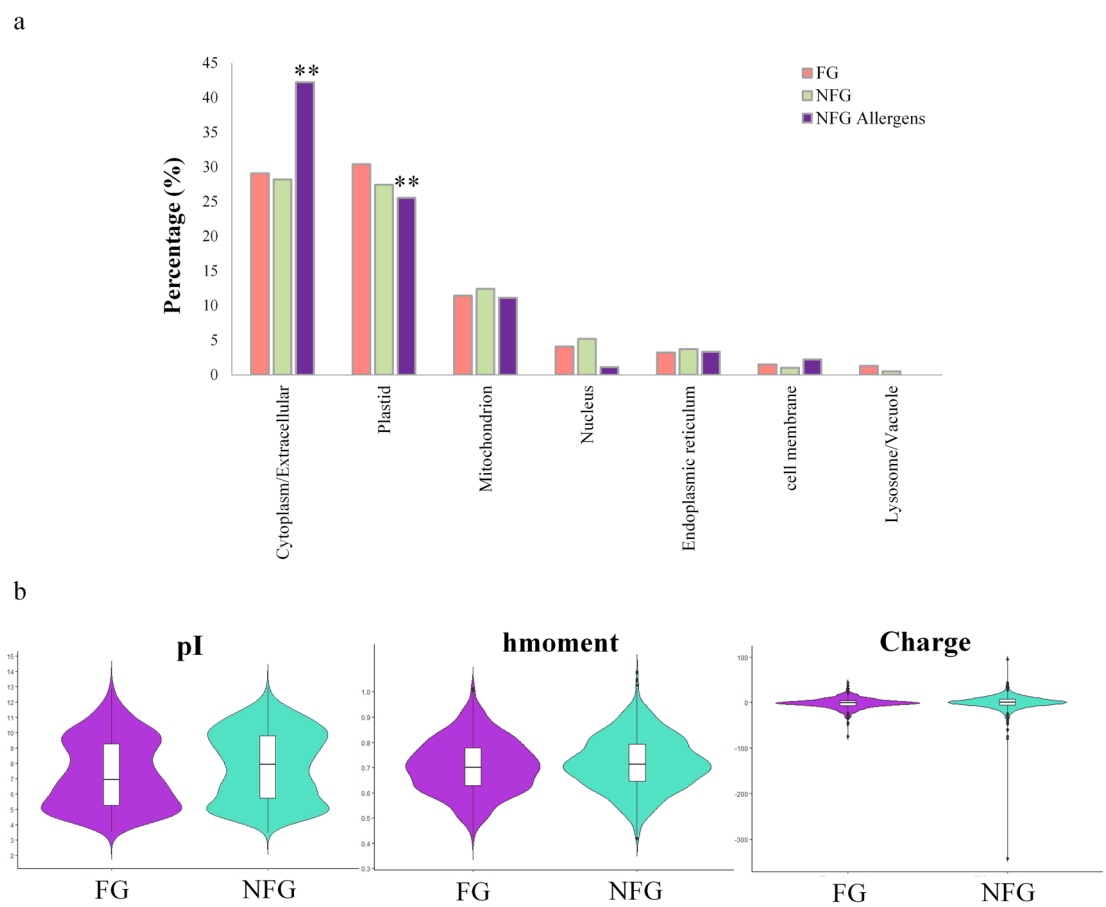

Fig. 6. Subcellular localization and physicochemical properties of proteins. (a) the bars indicate the frequency of proteins localized to different subcellular compartments in FG, NFG and NFG putative allergens. (b) Violin plots comparing the distribution of pI, hmoment, and charge in FG and NFG. ** indicates the p-value < 0.01.

extracts showed that the FG extraction method is not able to extract a broad range of proteins with different pI, charge, and the periodicity in protein hydrophobicity (predicted by Hmoment) in compare with NFG extract buffer.

4. Conclusion

The proteome of *N. oculata* FG and NFG extracts was investigated using data-independent acquisition proteomics and computational tools. Although the proteins identified in NFG were similar for the different cell disruption methods tested, in FG extraction methods, bead beating led to a slightly higher number of identifiable proteins compared to the control extraction method. This resulted in the enrichment of light harvesting complex proteins (LHCs) in the proteins discriminating between FG and NFG extracts. Regarding the safety of *N. oculata* intake, we identified several putative allergenic proteins in FG extracts with high similarity to some known food allergens. This underlines the necessity for establishing the safety of microalgae through further investigation, such as IgE binding assays. It also highlights the necessity for considering allergen labelling for microalgae products. Taken together, our investigation of the *N. oculata* proteome identified our method as robust for extracting food-grade protein from ground commercial microalgae.

CRediT authorship contribution statement

Sara Hamzelou: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Software, Visualization, Writing – original draft. **Damien Belobrajdic:** Supervision, Conceptualization, Writing – review & editing. **Angéla Juhász:** Software, Writing – review & editing. **Henri Brook:** Methodology, Writing – review & editing. **Utpal Bose:** Software, Writing – review & editing. **Michelle L Colgrave:** Supervision, Conceptualization, Writing – review & editing. **James A Broadbent:** Supervision, Conceptualization, Software, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

S.H. would like to acknowledge scholarship support from the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Early Research Career (CERC) postdoctoral fellowship.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2023.136459.

References

Barkia, I., Saari, N., & Manning, S. R. (2019). Microalgae for high-value products towards human health and nutrition. *Marine Drugs, 17*(5). https://doi.org/10.3390/md17050304

Batista, A. P., Gouveia, L., Bandarra, N. M., Franco, J. M., & Raymundo, A. (2013). Comparison of microalgal biomass profiles as novel functional ingredient for food products. *Algal Research, 2*(2), 164–173. https://doi.org/10.1016/j.algal.2013.01.004

Bianco, M., Ventura, G., Calvano, C. D., Losito, I., & Cataldi, T. R. I. (2022). A new paradigm to search for allergenic proteins in novel foods by integrating proteomics analysis and *in silico* sequence homology prediction: Focus on spirulina and chlorella microalgae. *Talanta, 240*, Article 123188. https://doi.org/10.1016/j.talanta.2021.123188

Brown, M. R. (1991). The amino-acid and sugar composition of 16 species of microalgae used in mariculture. *Journal of Experimental Marine Biology and Ecology, 145*(1), 79–99. https://doi.org/10.1016/0022-0981(91)90007-J

Bublin, M., & Breiteneder, H. (2020). Cross-reactivities of non-homologous allergens. *Allergy, 75*(5), 1019–1022. https://doi.org/10.1111/all.14120

Colina, F., Carbó, M., Meijón, M., Cañal, M. J., & Valledor, L. (2020). Low UV-C stress modulates *Chlamydomonas reinhardtii* biomass composition and oxidative stress response through proteomic and metabolomic changes involving novel signalers and effectors. *Biotechnology for Biofuels, 13*(1), 110. https://doi.org/10.1186/s13068-020-01750-8

Collins, K. D., & Washabaugh, M. W. (1985). The Hofmeister effect and the behaviour of water at interfaces. *Quarterly Reviews of Biophysics, 18*(4), 323–422. https://doi.org/10.1017/s0033583500005369

Dean, K. M., Qin, Y., & Palmer, A. E. (2012). Visualizing metal ions in cells: An overview of analytical techniques, approaches, and probes. *Biochimica et Biophysica Acta, 1823*(9), 1406–1415. https://doi.org/10.1016/j.bbamcr.2012.04.001

Dillon, M. B. C., Schulten, V., Oseroff, C., Paul, S., Dullanty, L. M., Frazier, A., … Sette, A. (2015). Different Bla-g T cell antigens dominate responses in asthma versus rhinitis subjects. *Clinical & Experimental Allergy, 45*(12), 1856–1867. https://doi.org/10.1111/cea.12643

Efsa Panel on Dietetic Products, N., & Allergies. (2014). Scientific Opinion on Dietary Reference Values for chromium. *EFSA Journal, 12*(10), 3845. doi: 10.2903/j.efsa.2014.3845.

Efsa Panel on Dietetic Products, N., Allergies, Turck, D., Bresson, J.-L., Burlingame, B., Dean, T., … Naska, A. (2016). Dietary reference values for potassium. *EFSA Journal, 14*(10), e04592. doi: 10.2903/j.efsa.2016.4592.

El-Hawy, A. S., Abdel-Rahman, H. G., El-Bassiony, M. F., Anwar, A., Hassan, M. A., Elnabtiti, A. A. S., … Kamel, S. (2022). Immunostimulatory effects of *Nannochloropsis oculata* supplementation on Barki rams growth performance, antioxidant assay, and immunological status. *BMC Veterinary Research, 18*(1), 314. https://doi.org/10.1186/s12917-022-03417-y

Ferro, M., Salvi, D., Brugière, S., Miras, S., Kowalski, S., Louwagie, M., … Rolland, N. (2003). Proteomics of the Chloroplast Envelope Membranes from *Arabidopsis thaliana*. *Molecular & Cellular Proteomics, 2*(5), 325–345. https://doi.org/10.1074/mcp.M300030-MCP200

Gorissen, S. H. M., Crombag, J. J. R., Senden, J. M. G., Waterval, W. A. H., Bierau, J., Verdijk, L. B., & van Loon, L. J. C. (2018). Protein content and amino acid composition of commercially available plant-based protein isolates. *Amino Acids, 50*(12), 1685–1695. https://doi.org/10.1007/s00726-018-2640-5

Greenly, J. M., & Tester, J. W. (2015). Ultrasonic cavitation for disruption of microalgae. *Bioresource Technology, 184*, 276–279. https://doi.org/10.1016/j.biortech.2014.11.036

Hamzelou, S., Melino, V. J., Plett, D. C., Kamath, K. S., Nawrocki, A., Larsen, M. R., … Haynes, P. A. (2021). The phosphoproteome of rice leaves responds to water and nitrogen supply. *Molecular Omics, 17*(5), 706–718. https://doi.org/10.1039/D1MO00137J

Hounslow, E., Evans, C. A., Pandhal, J., Sydney, T., Couto, N., Pham, T. K., … Wright, P. C. (2021). Quantitative proteomic comparison of salt stress in *Chlamydomonas reinhardtii* and the snow alga *Chlamydomonas nivalis* reveals mechanisms for salt-triggered fatty acid accumulation via reallocation of carbon resources. *Biotechnology for Biofuels, 14*(1), 121. https://doi.org/10.1186/s13068-021-01970-6

James, C. A., Welham, S., & Rose, P. (2023). Edible algae allergenicity – A short report. *Journal of Applied Phycology, 35*(1), 339–352. https://doi.org/10.1007/s10811-022-02880-2

Kent, M., Welladsen, H. M., Mangott, A., & Li, Y. (2015). Nutritional evaluation of Australian microalgae as potential human health supplements. *PLoS One1, 10*(2), e0118985.

Kleine-Tebbe, J., Ollert, M., Radauer, C., & Jakob, T. (2017). Introduction to molecular allergology: Protein families, databases, and potential benefits. In J. Kleine-Tebbe, & T. Jakob (Eds.), *Molecular allergy diagnostics: innovation for a better patient management* (pp. 3–19). Cham: Springer International Publishing.

Maurer-Stroh, S., Krutz, N. L., Kern, P. S., Gunalan, V., Nguyen, M. N., Limviphuvadh, V., … Gerberick, G. F. (2019). AllerCatPro—prediction of protein allergenicity potential from the protein sequence. *Bioinformatics, 35*(17), 3020–3027. https://doi.org/10.1093/bioinformatics/btz029

Moore, J. C., DeVries, J. W., Lipp, M., Griffiths, J. C., & Abernethy, D. R. (2010). Total protein methods and their potential utility to reduce the risk of food protein adulteration. *Comprehensive Reviews in Food Science and Food Safety, 9*(4), 330–357. https://doi.org/10.1111/j.1541-4337.2010.00114.x

Nguyen, M. N., Krutz, N. L., Limviphuvadh, V., Lopata, A. L., Gerberick, G. F., & Maurer-Stroh, S. (2022). AllerCatPro 2.0: A web server for predicting protein allergenicity potential. *Nucleic Acids Research, 50*(W1), W36–W43. https://doi.org/10.1093/nar/gkac446

Pimblett, P. (2020). Spirulina allergy; a case history of two patients. *World Allergy Organization Journal, 13*(8), Article 100149. https://doi.org/10.1016/j.waojou.2020.100149

Radauer, C., Bublin, M., Wagner, S., Mari, A., & Breiteneder, H. (2008). Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *The Journal of Allergy and Clinical Immunology, 121*(4), 847–852.e847. https://doi.org/10.1016/j.jaci.2008.01.025

Rolland, N., Ferro, M., Ephritikhine, G., Marmagne, A., Ramus, C., Brugière, S., … Garin, J. (2006). A versatile method for deciphering plant membrane proteomes.

Journal of Experimental Botany, 57(7), 1579–1589. https://doi.org/10.1093/jxb/erj162

Roustan, V., Bakhtiari, S., Roustan, P. J., & Weckwerth, W. (2017). Quantitative in vivo phosphoproteomics reveals reversible signaling processes during nitrogen starvation and recovery in the biofuel model organism Chlamydomonas reinhardtii. *Biotechnology for Biofuels, 10*, 280. https://doi.org/10.1186/s13068-017-0949-z

Safi, C., Cabas Rodriguez, L., Mulder, W. J., Engelen-Smit, N., Spekking, W., van den Broek, L. A. M., … Sijtsma, L. (2017). Energy consumption and water-soluble protein release by cell wall disruption of *Nannochloropsis gaditana*. *Bioresource Technology, 239*, 204–210. https://doi.org/10.1016/j.biortech.2017.05.012

Safi, C., Frances, C., Ursu, A. V., Laroche, C., Pouzet, C., Vaca-Garcia, C., & Pontalier, P.-Y. (2015). Understanding the effect of cell disruption methods on the diffusion of *Chlorella vulgaris* proteins and pigments in the aqueous phase. *Algal Research, 8*, 61–68. https://doi.org/10.1016/j.algal.2015.01.002

Safi, C., Olivieri, G., Campos, R. P., Engelen-Smit, N., Mulder, W. J., van den Broek, L. A. M., & Sijtsma, L. (2017). Biorefinery of microalgal soluble proteins by sequential processing and membrane filtration. *Bioresource Technology, 225*, 151–158. https://doi.org/10.1016/j.biortech.2016.11.068

Safi, C., Ursu, A. V., Laroche, C., Zebib, B., Merah, O., Pontalier, P.-Y., & Vaca-Garcia, C. (2014). Aqueous extraction of proteins from microalgae: Effect of different cell disruption methods. *Algal Research, 3*, 61–65. https://doi.org/10.1016/j.algal.2013.12.004

Sojo, V., Dessimoz, C., Pomiankowski, A., & Lane, N. (2016). Membrane proteins are dramatically less conserved than water-soluble proteins across the tree of life. *Molecular Biology and Evolution, 33*(11), 2874–2884. https://doi.org/10.1093/molbev/msw164

Soto-Sierra, L., Stoykova, P., & Nikolov, Z. L. (2018). Extraction and fractionation of microalgae-based protein products. *Algal Research, 36*, 175–192. https://doi.org/10.1016/j.algal.2018.10.023

Spolaore, P., Joannis-Cassan, C., Duran, E., & Isambert, A. (2006). Commercial applications of microalgae. *Journal of Bioscience and Bioengineering, 101*(2), 87–96. https://doi.org/10.1263/jbb.101.87

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One1, 6*(7), e21800.

Team, R. D. C. (2010). R: a language and environment for statistical computing.

Thumuluri, V., Almagro Armenteros, J. J., Johansen, A. R., Nielsen, H., & Winther, O. (2022). DeepLoc 2.0: Multi-label subcellular localization prediction using protein language models. *Nucleic Acids Research, 50*(W1), W228–W234. https://doi.org/10.1093/nar/gkac278

Tiberg, E., Dreborg, S., & Björkstén, B. (1995). Allergy to green algae (Chlorella) among children. *Journal of Allergy and Clinical Immunology, 96*(2), 257–259. https://doi.org/10.1016/S0091-6749(95)70016-1

Tomm, J. M., van Do, T., Jende, C., Simon, J. C., Treudler, R., von Bergen, M., & Averbeck, M. (2013). Identification of new potential allergens from Nile perch (Lates niloticus) and cod (Gadus morhua). *Journal of Investigational Allergology & Clinical Immunology, 23*(3), 159–167.

Wan Razali, W. A., Evans, C. A., & Pandhal, J. (2022). Comparative proteomics reveals evidence of enhanced EPA trafficking in a mutant strain of *Nannochloropsis oculata*. *Frontiers in Bioengineering and Biotechnology, 10*, Article 838445. https://doi.org/10.3389/fbioe.2022.838445

Wei, L., El Hajjami, M., Shen, C., You, W., Lu, Y., Li, J., … Xu, J. (2019). Transcriptomic and proteomic responses to very low CO(2) suggest multiple carbon concentrating mechanisms in Nannochloropsis oceanica. *Biotechnology for Biofuels, 12*, 168. https://doi.org/10.1186/s13068-019-1506-8

Wild, K. J., Steingaß, H., & Rodehutscord, M. (2018). Variability in nutrient composition and *in vitro* crude protein digestibility of 16 microalgae products. *Journal of Animal Physiology and Animal Nutrition, 102*(5), 1306–1319. https://doi.org/10.1111/jpn.12953

Wong, C., Roberts, S. M., & Saab, I. N. (2022). Review of regulatory reference values and background levels for heavy metals in the human diet. *Regulatory Toxicology and Pharmacology, 130*, Article 105122. https://doi.org/10.1016/j.yrtph.2022.105122

Yang, Y., Zhang, Y.-X., Liu, M., Maleki, S. J., Zhang, M.-L., Liu, Q.-M., … Liu, G.-M. (2017). Triosephosphate isomerase and filamin C share common epitopes as novel allergens of Procambarus clarkii. *Journal of Agricultural and Food Chemistry, 65*(4), 950–963. https://doi.org/10.1021/acs.jafc.6b04587