Is Gigartina a potential source of food protein and functional peptide-based ingredients? Evaluating an industrial, pilot-scale extract by proteomics and bioinformatics

Gregersen Echers, Simon; Abdul-Khalek, Naim; Mikkelsen, Rasmus Kranold; Holdt, Susan Løvstad; Jacobsen, Charlotte; Hansen, Egon Bech; Olsen, Tobias Hegelund; Sejberg, Jimmy J.P.; Overgaard, Michael Toft

Published in:
Future Foods

DOI (link to publication from Publisher):
10.1016/j.fufo.2022.100189

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Future Foods 6 (2022) 100189

Contents lists available at ScienceDirect

Future Foods

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

Is Gigartina a potential source of food protein and functional peptide-based ingredients? Evaluating an industrial, pilot-scale extract by proteomics and bioinformatics

Simon Gregersen Echers[a,*], Naim Abdul-Khalek[a], Rasmus Kranold Mikkelsen[b], Susan Løvstad Holdt[b], Charlotte Jacobsen[b], Egon Bech Hansen[c], Tobias Hegelund Olsen[d], Jimmy J.P. Sejberg[e], Michael Toft Overgaard[a]

[a] Department of Chemistry and Bioscience, Aalborg University, Fredrik Bajers Vej 7H, DK-9220 Aalborg, Denmark
[b] Research Group for Bioactives - Analysis and Application, National Food Institute, Technical University of Denmark, Kemitorvet, 201, Kgs Lyngby 2800, Denmark
[c] Research Group for Gut, Microbes and Health, National Food Institute, Technical University of Denmark, Kemitorvet, 202, 1250, Kgs Lyngby 2800, Denmark
[d] Department of Health Technology, Technical University of Denmark, Ørsteds Plads, 345C, Kgs. Lyngby 2800, Denmark
[e] CP Kelco ApS, Ved Banen 16, Lille Skensved, 4623, Denmark

ARTICLE INFO

Keywords:
Rhodophyta
Seaweed protein
Valorization
Proteomics
Bioinformatics
Peptide emulsifiers

ABSTRACT

Seaweeds attract substantial interest as novel sources of sustainable food protein, as they are established sources of industrial hydrocolloids with reasonable protein content. In this study, we investigate the protein composition and nutritional quality of a seaweed protein extract (SPE) from *Gigartina sp*. The SPE displayed low (<2%), but pH-dependent, aqueous solubility likely due to the harsh conditions employed during extraction. Solubility improved using alkaline buffering and detergent addition to facilitate proteomic characterization by quantitative LC-MS/MS. Proteomics analysis revealed that SPE was dominated by proteins related to light harvest and particularly phycobiliproteins (44%), where phycoerythrin was most abundant (28%). Based on subcellular localization, the extraction method was evaluated as good for release of cellular protein. SPE was rich in essential amino acids (36–41%) and particularly branched chain amino acids (22–24%), and thereby a potential source of nutritional food protein. Using bioinformatic prediction and structural modelling, we found abundant SPE proteins contained novel peptides with the amphiphilic properties required to stabilize an oil/water interface, and thereby high probability of being potent emulsifiers. Based on this study, *Gigartina sp.* could serve a good candidate for extraction of sustainable, nutritious food protein, with the possibility of further processing into hydrolysates with strong emulsifying properties for use as natural food ingredients.

Introduction

Global food systems are responsible for more than one third of the total anthropogenic greenhouse gas emission, where animal-based food systems account for more than half (Xu et al., 2021). Meanwhile, the global population continues to increase, thereby also increasing food and protein demand significantly in the future. It is estimated that within the period from 2010 to 2050, global food demand may increase up to 62% (van Dijk et al., 2021). To accommodate this demand, and meanwhile reducing environmental footprint of food systems, development and application of alternative and sustainable protein sources is pivotal. While plant-based protein (green biomass) have received tremendous attention (Aschemann-Witzel et al., 2020), aquatic/marine protein (blue biomass) is making headway in academia and industry as a means to solve global challenges related to sustainable food supply (Lange and Meyer, 2019). Although not by definition a plant (Sudhakar et al., 2018), seaweeds (macroalgae) are regarded as a high potential aquatic source for sustainable and animal-free protein production, thereby also adhering to e.g. vegan dietary preferences (Raja et al., 2022). In addition, seaweeds are often highlighted for their nutraceutical (Tanna and Mishra, 2018) and health beneficial properties (Holdt and Kraan, 2011).

The primary industrial application of seaweeds is for production of hydrocolloids such as alginate, agar and carrageenan, which are used as gelling agents in the food industry (Brown et al., 2014). With approximate 25 million MT cultivated in 2019, where upwards of 70% are

* Corresponding author.
E-mail addresses: sgr@bio.aau.dk (S. Gregersen Echers), nakg@bio.aau.dk (N. Abdul-Khalek), rkrmik@food.dtu.dk (R.K. Mikkelsen), suho@food.dtu.dk (S.L. Holdt), chja@food.dtu.dk (C. Jacobsen), egbh@food.dtu.dk (E.B. Hansen), tobiasheol@gmail.com (T.H. Olsen), Sejberg@cpkelco.com (J.J.P. Sejberg), mto@bio.aau.dk (M.T. Overgaard).

https://doi.org/10.1016/j.fufo.2022.100189
Received 10 June 2022; Received in revised form 14 September 2022; Accepted 20 September 2022

mineral- and protein-rich side- and/or waste-streams (Cai et al., 2021), the global seaweed market is in rapid development and estimated to exceed US$ 20 billion by 2024 (Custódio et al., 2017). Within the hydrocolloid industry, red seaweeds (Rhodophyta) belonging to the *Eucheuma, Chondrus, Kappaphycus, Hypnea*, and *Gigartina* genera are commonly used as sources for carrageenan extraction (Colusse et al., 2022; Singh et al., 2022). In general, red seaweeds are considered of high potential for food application due to their high protein content, reaching upwards of 47 % dry matter basis (Biancarosa et al., 2016) and therefore also highlighted as a promising source of alternative proteins in the quest to increase global food sustainability (Rawiwan et al., 2022). As red seaweed furthermore constitute more than half of the global seaweed market (Cai et al., 2021), side streams from hydrocolloid production represent a high potential protein source which still remains underutilized (Bleakley and Hayes, 2017; Naseri et al., 2019). *Gigartina* is a genus belonging to the *Gigartinaceae* family under the *Gigartinales* order and comprises 43 species (Cotas et al., 2022). *Gigartina* sp are considered to be edible seaweed and have been highlighted for their nutraceutical potential (Cotas et al., 2022) due to high levels of minerals, dietary fibers, and protein (Gómez-Ordóñez et al., 2010). In addition to the many direct applications of *Gigartina* carrageenan, the genus has been reported to show antioxidant, skin protective, wound-healing, and anti-proliferative activity (Jiménez-Escrig et al., 2012; Chrapusta et al., 2017; Cotas et al., 2020). Nevertheless, *Gigartina* protein remains largely unutilized.

Mass spectrometry (MS)-based proteomics analysis has developed into a cornerstone of biotechnological and biomedical research (Armengaud, 2016; Noor et al., 2021), but is still in its infancy for food protein analysis and characterization. In the food context, MS-based proteomics has primarily been employed to investigate food safety, food quality, allergenicity, as well as organism and cellular responses to exogenous stimuli in relation to e.g. crop breeding (Agrawal et al., 2013; Andjelković and Josić, 2018; Mora et al., 2018; Lexhaller, Colgrave and Scherf, 2019; Marzano et al., 2020). Such studies have also been performed for red seaweeds (Beaulieu, 2019). However, the ability of MS-based proteomics to supply qualitative and quantitative insight on the single protein level also opens new possibilities for its application. For instance, MS-based proteomics and downstream bioinformatic data analysis has been employed as a tool to evaluate extraction efficiency from red seaweeds (Gregersen et al., 2022) and how integrated extraction methods such as application of cell wall degrading enzymes can improve recovery of intracellular protein from *Eucheuma denticulatum* (Gregersen et al., 2021).

Food protein is traditionally evaluated based on a range of bulk characteristics related to functional, nutritional, sensory, and safety aspects. Among these, a suitable amino acid (AA) profile is crucial for a protein being compatible with human nutritional requirements (Mæhre et al., 2014). In this context, the content of particularly essential amino acids (EAAs) and the ratio to non-essential amino acids (NEAAs), ΣEAA/ΣNEAA, are central evaluation criteria. For a representative subsample of red seaweeds, this ratio (0.63) is slightly higher than for green (0.61) and brown (0.59) seaweeds. Compared to other alternative protein sources such as soy (0.91), pea (1.02), potato (1.09), and corn (0.78), this makes seaweed protein overall of a somewhat lower nutritional quality (Donadelli et al., 2019). Nevertheless, some red seaweed species such as *E. denticulatum* have been shown to have comparable nutritional potential with ΣEAA/ΣNEAA = 0.86 (Naseri et al., 2019). In this respect it is also important to keep in mind that the AA profile for a whole organism may not at all reflect the profile in a protein extract or isolate. Traditionally, the AA composition is evaluated by means of conventional AA analysis, where protein is fully hydrolyzed in hot hydrochloric acid and subsequently individual amino acids are quantified relative to a standard curve (Rutherfurd and Gilani, 2009). However, conventional AA analysis comes with limitations. For instance, due to deamidation it is not possible to distinguish Asn and Gln from Asp and Glu, respectively, Trp is fully degraded, and substantial losses of hydroxyl- (Thr and Ser) and sulfur- (Cys and Met) containing AAs is observed (Rutherfurd and Gilani, 2009). To alleviate this, the use of peptide- and protein-level MS data has been suggested as a promising alternative to estimate AA composition, and showed good overall correlation with conventional AA analysis (Jafarpour et al., 2020; Gregersen et al., 2021). Moreover, this approach is able to quantify Trp and free Cys as well as distinguish Asn/Gln from Asp/Glu.

A good nutritional quality is important for alternative proteins, but having functional properties increases applicability as food ingredients and thereby also product value (Granato et al., 2010). For many alternative proteins, extraction methodology has the unwanted effect of being denaturing, thereby reducing or completely eliminating functionality and aqueous solubility (Sari et al., 2015). This is particularly relevant when designing a scalable and economically feasible process in the transfer from small-scale lab experiments to full-scale industrial processes. While significant effort is being put into development of novel, green extraction methods that are both environmentally friendly and has potential to produce a non-denatured protein (Picot-Allain et al., 2021), enzymatic hydrolysis offers an alternative solution to the problem. Hydrolysis has been reported to improve solubility, recovery, and functionality for a wide range of alternative proteins such as chickpea (Mokni Ghribi et al., 2015), soy (Sun, 2011), potato (Gregersen Echers et al., 2022), pea (García Arteaga et al., 2020), just to name a few. While this approach holds great potential, the process is usually not well understood or characterized on the molecular level. Rather, hydrolysis is performed by trial-and-error, where specific protease(s) and process conditions are screened and subsequently evaluated based on functionality of the bulk hydrolysate. As the liberated peptides are responsible for the observed functionalities, molecular insight and peptide-level understanding may add value by the possibility of process design based on desired functionality, working more from the bottom-up rather than the top-down.

Recent development in bioinformatic prediction of peptide functionality, has facilitated *de novo* computational screening of proteins for embedded peptides with specific functionalities such as emulsifying (García-Moreno, et al., 2020; García-Moreno, et al., 2020) and antioxidant (Olsen et al., 2020) activity. Moreover, several web-tools have been developed to predict peptide functionality and bioactivity within multiple classes such as Peptide Ranker (Mooney et al., 2012), MultiPep (Grønning et al., 2021), and BIOPEP-UWM (Minkiewicz et al., 2019). While the specific predictors are highly flexible and able to use both protein and peptide sequence input to predict functional potential, the multi-label classifiers come with some limitations. Peptide Ranker and MultiPep are restricted to only peptide input. While BIOPEP-UWM has many flexible features, it is restricted to database searching against known functional and bioactive peptides. These restriction limit their applicability in the discovery phase of novel peptides using protein-level input data, whereas the specific *de novo* predicters have been applied to identify several novel peptides with emulsifying (Yesiltas et al., 2021) and antioxidant (Yesiltas et al., 2022) activity from various alternative protein sources, including seaweeds.

In this study, we present the first proteomic characterization of a protein-rich extract from *Gigartina*. The seaweed protein extract (SPE) was obtained at pilot-scale using a process that is both scalable and industrially relevant, and is implementable as a direct downstream process following initial carrageenan extraction. Using MS-based proteomics, we present a qualitative and quantitative analysis of the protein composition to obtain insight on protein extractability and the nutritional value of the extract. Moreover, using *in silico* sequence analysis for prediction of potential emulsifier peptides embedded in the most abundant proteins, we relate predictions to the peptide conformation within the native proteins. Through this analysis, we construct a protein- and peptide-level understanding of the proteomic composition which is in important step for *ex vitro* evaluation of a protein source, in relation to its potential as a substrate for production of a functional hydrolysate.

Materials and methods

Materials

The seaweed protein extract (SPE) was supplied by CP Kelco (Lille Skensved, Denmark) in dry form with a protein content of 42.3% (by Kjeldal analysis and an N-to-protein conversion factor of 6.25, CP Kelco supplied information) as a by-product from mild carrageenan extraction in pilot-scale. Extraction of SPE is described below. All used chemicals were of HPLC-grade.

Protein extraction

For the extraction of SPE, CP Kelco uses *Gigartina radula* which is commonly used industrially as a joint description of mixed *Gigartina sp.*, more precisely *Gigartina broad leaf, Gigartina narrow leaf, Gigartina skottsbergii, Gigartina chammisoi,* and *Gigartina pistillata*. Aqueous (aq.) carrageenan extraction was performed at 95 °C and pH 9 for two hours under constant stirring and subsequently centrifuged at 8221 g for 20 min. The sediment (initial byproduct) was subsequently washed with aq. HCl (pH 2) for 30 min to degrade residual carrageenan and centrifuged as above. Lastly, the remaining filter cake was washed with dH_2O until a neutral pH was obtained, centrifuged as above, and lyophilized to yield the final SPE.

Protein solubility

Protein solubility was investigated as pH-dependent aq. solubility and by the addition of buffer and/or detergents. For aq. solubility, ddH_2O was adjusted to six pH values (2, 4, 6, 8, 10, and 12) using dilute hydrochloric acid or sodium hydroxide and 1 mL was added to 22.1 mg SPE corresponding to a total protein concentration of 10 mg/mL (based on supplied protein content) at full solubilization. To investigate the effect of buffer and detergents on protein solubility, 22.1 mg SPE was added 5 mL (final protein concentration 2 mg/mL at full solubilization) of five different solvents: A) 200 mM Ammonium bicarbonate (ABC, Sigma Aldrich, USA) pH; 8.0. B) 2% Sodium dodecyl sulfate (SDS, Applichem, Darmstadt, Germany) in 200 mM ABC pH 8.0. C) 0.2% SDS in 50mM ABC, pH 8.2. D) ddH_2O (no pH adjustment). E) 1% Sodium deoxycholate (SDC, Sigma Aldrich, USA) in 50 mM Triethylammonium bicarbonate (TEAB, Sigma Aldrich, USA), pH 9.5. In all experiments, SPE was mixed with solvent, vortexed for 30 s, sonicated for 30 min, and centrifuged at 3095 RCF and 4 °C for 10 min in a 5810 R centrifuge (Eppendorf, Germany) to precipitate solids prior to aliquoting the supernatant. The total soluble protein (TSP) was determined using both Qubit protein assay (Thermo Scientific, Germany) and A_{280} by Nanodrop (ND, Thermo Scientific) according to the manufacturer guidelines. For Qubit measurements, blank solvents were measured as reference, while for ND, the respective solvent were used to blank the instrument prior to sample analysis. All measurements were performed in triplicates and solubility calculated as % of full solubilization based total protein from Kjeldal analysis (N x 6.25).

SDS-PAGE

Supernatants from all solubilization experiments were analyzed by SDS-PAGE using precast 4–20% gradient gels (Genscript, USA) in a Tris-MOPS buffered system under reducing conditions and according to manufacturer guidelines. As molecular weight marker, 5 µL PIERCE Unstained Protein MW Marker P/N 26610 (ThermoFisher Scientific, USA) was used. Visualization was performed by Coomassie Brilliant Blue G250 (Sigma-Aldrich, Germany) staining and gel imaging conducted with a ChemDoc MP Imaging System (Bio-Rad, USA).

In-solution digest

Proteins were digested in-solution as technical triplicates, according to Zhou et al. (2015) with minor modifications. In brief, aliquots of 50 µg protein (based on nanodrop) in solvent E (1% SDC in 50 mM TEAB) was heated to 99 °C for 10 min and subsequently cooled to room temperature (RT). Next, proteins were reduced using tris(2-carboxyethyl)phosphine (TCEP, Sigma-Aldrich) by adding 1 µg TCEP per 25 µg protein and incubating at 37 °C for 30 min. Then, proteins were alkylated using iodoacetamide (IAA, Fluka Biochemika, Switzerland) by adding 1 µg IAA per 10 µg protein and incubating for 20 min at 37 °C in the dark. Lastly, proteins were digested with sequencing grade modified trypsin (Promega, USA) by adding 1 µg trypsin per 50 µg protein and incubating overnight at 37 °C. Following digestion, samples were acidified by adding formic acid (FA, VWR, Denmark) to a final concentration of 0.5% and incubated at RT for 5 min to precipitate SDC. The samples were centrifuged at 13,000 g for 20 min at 4 °C, whereafter the supernatants were transferred to new tubes, and centrifuged again (same conditions). Final supernatants were purified using in-house prepared C-18 StageTips, dried down, and resuspended in 0.1% FA in 2% acetonitrile (ACN, VWR) for analysis, as previously described (García-Moreno, et al., 2020).

Bottom-up proteomics by LC-MS/MS

Protein digests were analyzed by shotgun, bottom-up proteomics, as previously described (Gregersen *et al.*, 2021). Briefly, LC-MS/MS analysis was performed using a system composed of an EASY nLC 1200 (Thermo Scientific, Germany) coupled in-line to a Q Exactive HF (QE-HF) mass spectrometer (Thermo Scientific) via a Nanospray Flex ion source (Thermo Scientific). Peptides were loaded on a reverse phase Acclaim PEPMAP NANOTRAP column (C18, 100 Å, 100 μm. × 2 cm, (Thermo Scientific)) in solvent A (0.1% FA) followed by separation on a reverse phase ACCLAIM PEPMAP RSLC analytical column (C18, 100 Å, 75 um × 50 cm (Thermo Scientific)). Samples were loaded using 0.1% FA and separated through a gradient from 5-100% of 0.1% FA in 80% ACN over 60 min. The QE-HF was operated in top-20, positive ion, DDA mode, using 28 eV HCD fragmentation and a 1.2 m/z isolation window for MS2. MS1 scans were obtained from 400-1200 m/z with a resolution of 60,000 (at 200 m/z) and MS2 was performed with a resolution of 15,000 (at 200 m/Z). Maximum ion injection time was set to 50 for MS1 and 45 for MS2 scans. ACG target was set at 1e6 and 1e5 for MS1 and MS2, respectively. The underfill ratio was set to 3.5% and a dynamic exclusion of 30 sec was applied. During acquisition, “peptide match” and “exclude isotopes” were enabled.

Protein database construction and LC-MS/MS data analysis

As no specific proteomic, genomic, or transcriptomic databases are available for *Gigartina sp.*, a custom database was created for analysis of LC-MS/MS data. For this, all entries from Uniprot (UniProt Consortium et al., 2021) annotated to the order of *Gigartinales* (16,072 entries) were retrieved. This custom database also includes a references proteome for the related species *Chondrus crispus* (UP000012073, 9598 entries), thereby presumably ensuring sufficient depth and coverage of the *Gigartina sp.* proteome.

LC-MS/MS data were subsequently analyzed in MaxQuant v.1.6.10.43 (Cox and Mann, 2008) against the custom protein database using standard setting for tryptic digest, as previously described (Gregersen et al., 2021).Up to five modifications and two missed cleavages were allowed and a false discovery rate (FDR) of 1% was applied on both peptide (PSM) and protein level. For quantification, both the label-free quantification, MaxLFQ (Cox et al., 2014), and intensity-based absolute quantification, iBAQ (Schwanhüusser et al., 2011), were employed. Quantitative reproducibility for relative iBAQ (riBAQ) quantification was evaluated by the coefficient of variation (CV), determined as the standard deviation of the replicates divided

by the mean for each protein (group). MaxQuant output data were further analyzed in Perseus v.1.6.1.3 (Tyanova and Cox, 2018) to filter contaminants and false positives and investigate reproducibility between technical replicates. Overall quantitative reproducibility by riBAQ was evaluated using the stochastic mean CV for all quantified protein groups and by determining an abundance-weighted mean CV, CV^{w}_{mean}, calculated using riBAQs as weights according to:

$$CV^{w}_{mean} = \sum_{pro=1}^{m} CV^{q} * riBAQ^{q}$$

where CV^{q} is the CV for protein q of m total proteins and $riBAQ^{q}$ is the riBAQ of the same protein. Data was visualized using Perseus and Venny v.2.1 (freely available at https://bioinfogp.cnb.csic.es/tools/venny/) before being exported for further downstream bioinformatic analysis.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Perez-Riverol et al., 2019) with the dataset identifier PXD034435 and DOI 10.6019/PXD034435.

Subcellular localization and signal peptide prediction

The subcellular localization of identified proteins was predicted by deepLoc (Almagro Armenteros et al., 2017) using the open source web-interface (http://www.cbs.dtu.dk/services/DeepLoc/index.php) with the application of BLOSUM62 encoding. The predicted compartment was validated for the most abundant (riBAQ > 1%) by application of LocTree3 (Goldberg et al., 2014), as well as annotated subcellular localization in Uniprot (UniProt Consortium et al., 2021). Potential signal peptides in abundant protein (riBAQ > 1%) was predicted by SignalP (v.6.0) (Armenteros et al., 2019) using the open-source web-interface (http://www.cbs.dtu.dk/services/SignalP-5.0/index.php). Predicted signal peptides were used to further validate subcellular localization prediction (if applicable).

Amino acid composition by peptide- and protein-level MS data

Amino acid composition was estimated using peptide- and protein-level MS data, as previously described (Jafarpour, Gomes, et al., 2020; Gregersen et al., 2021, 2022). For peptide-level MS data, the approach of intensity-weighted, peptide-level AA frequency was applied. Following filtering of reverse and contaminant peptides, peptide MS1 intensities were used as weights to estimate molar AA frequency, $f^{peptide}_{AA}$, as:

$$f^{peptide}_{AA} = \sum_{pep=1}^{n} f^{p}_{AA} * I^{p}_{rel}$$

where f^{p}_{AA} is the relative frequency of a given AA in peptide p and I^{p}_{rel} is the relative MS1 intensity of peptide p (i.e. the intensity of p divided by the sum of intensities for all n peptides.

For protein-level data, AA composition was estimated using the relative molar abundance (iBAQ). For all identified proteins, the relative AA frequency was computed using ProtrWeb (Xiao et al., 2015) (freely available as webtool at http://protr.org/). Using the relative molar frequency of each AA, $f^{protein}_{AA}$, AA composition was calculated as:

$$f^{protein}_{AA} = \sum_{pro=1}^{m} f^{q}_{AA} * riBAQ^{q}$$

where f^{q}_{AA} is the relative frequency of a given AA in protein q and $riBAQ^{q}$is the iBAQ of protein q divided by the sum of iBAQs for all m proteins.

AA composition was evaluated against the minimum adult dietary requirements according to the World Health Organization/Food and Agricultural Organization of the United Nations/United Nations University (WHO/FAO/UNU, 2007), following conversion to relative molar abundance, as previously described (Gregersen et al., 2021).

Prediction of embedded emulsifier peptides and quantitative visualization

For abundant proteins (riBAQ > 1%), protein sequences were analyzed using EmulsiPred (García-Moreno et al., 2020; García-Moreno et al., 2020) to predict embedded, emulsifying peptides. Peptides were predicted as potential emulsifiers if given an amphiphilic score > 2 when projected in either α-helical or β-strand conformation (facial amphiphiles) or in an undefined (γ) conformation (axial amphiphilicity). Predicted (and Uniprot annotated) signal peptides were used to ensure predicted emulsifier peptides did not have overlap with a potential signal region, which would not be present in the mature protein.

Protein modelling and visualization

The structure of selected proteins was modelled using templated homology modelling using the Swiss-Model Workspace (Waterhouse et al., 2018), as previously described (García-Moreno et al., 2020). Phycoerythrin alpha chain (Uniprot AC# I4DEE4) and Phycocyanin, alpha chain (Uniprot AC# M5DDK1) were modelled using the electron microscopy (EM) structures of Phycoerythrin alpha subunit (STML ID 6kgx.453.A) and C-phycocyanin alpha subunit (STML ID 6kgx.59.A) from the red algae *Porphyridium purpureum* as template, respectively. Ribulose bisphosphate carboxylase (RuBisCO) small subunit (Uniprot AC# A0A7M3VH79) was modelled using the X-ray crystal structure of RuBisCO small subunit from the red algae *Galdieria partita* (STML ID 1iwa.1.B). Photosystem II reaction center protein L (Uniprot AC# M5DD55) was modelled using the crystal structure of oxygen-evolving photosystem II from the red algae *Cyanidium caldarium* (STML ID 4yuu.1.K). Cytochrome b6-f complex subunit 5 (Uniprot AC# M5DDH2) was modelled using the cryo-EM structure of cytochrome b6f from spinach (*Spinacia oleracea*) (STML ID 6rqf.1.G). Models were visualized using The PyMOL Molecular Graphics System v. 1.5.0 (Schrödinger, LLC.) and colored using the Swiss-Model hydrophobicity color scheme (from hydrophilic (blue) to hydrophobic (red)).

Statistical analysis

Statistical analysis was performed for experimental data in GraphPad Prism (v.9.3.0) using multiple comparisons of means by ordinary one-way ANOVA and Tukey's HSD test with a 95% confidence level. For proteomics data, technical replicability was evaluated by computing the Pearson correlation coefficient (PCC) between replicates using both riBAQ and LFQ quantitative data in Perseus. Protein-level reproducibility was evaluated by computing the coefficient of variation (CV) as the standard deviation of the triplicate analysis divided by the mean.

Results and discussion

Initially, the aqueous solubility of the seaweed protein extract (SPE) was investigated in a pH-dependent manner (Fig. 1, Table A.1). In general, SPE showed low aq. solubility (total soluble protein, TSP) regardless of pH and quantification method, although Nanodrop indicated significantly higher solubility (1.5–11.6%) than Qubit (0.69–2.1%). Nanodrop is an unspecific quantitative estimate relying only on sample absorbance, and high A260/A280 ratios (1.61–2.08) from Nanodrop measurements (Table A.1) for aq. samples indicate low protein purity in the supernatants due to presence of e.g. nucleic acids (Schultz et al., 1994), thereby likely making measurements overestimates. Consequently, these measurements can be considered unreliable due to sample complexity, and for that reason Qubit measurements are considered more accurate, being a protein-specific fluorescence-based assay.

Both Nanodrop and Qubit TSP measurements show similar trends in terms of pH-dependent aqueous solubility of the SPE. While the solubility is significantly lower (P<0.05) at pH 2, the solubility increases

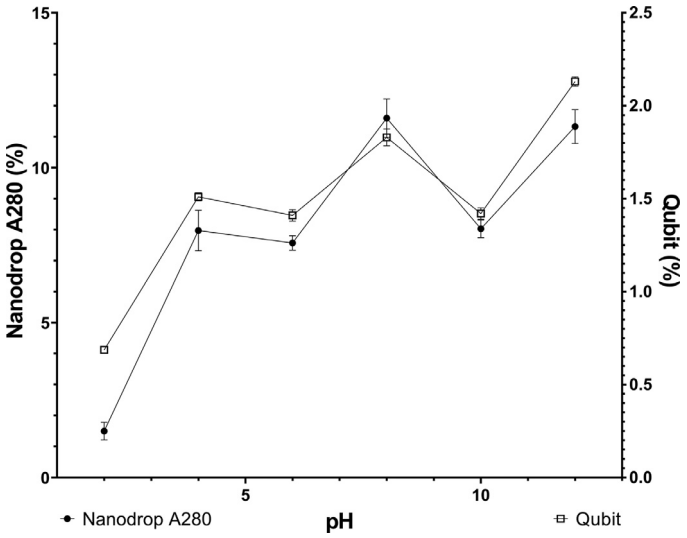

Fig. 1. pH-dependent aqueous protein solubility of SPE. Aqueous solubility of protein in SPE (% relative to supplied protein content by Dumas-N) as a function of pH and quantified using Nanodrop A280 (filled circles, left axis) and Qubit Protein Assay (open squares, right axis). Data points where no error bars are shown is a result of low SD (see Table A.1).

towards pH 8, drops at pH 10, and increases again at pH 12. Based on the upstream processing, these observations are not surprising. Initial carrageenan was performed under slightly alkaline conditions, whereby proteins with high solubility in this range may have been extracted alongside carrageenan, explaining the small drop in solubility at pH 10. Next, the sediment was treated under highly acidic (pH 2) to degrade residual carrageenan, thereby also removing acid-soluble protein from the final SPE, explaining the low solubility at this pH. While the low aqueous solubility of SPE may limit its direct applicability in foods (Amagliani et al., 2017), further processing, such as enzymatic hydrolysis, may significantly improve solubility as well as functionality of the predominantly insoluble protein (Ashaolu, 2020; Gregersen Echers et al., 2022).

To improve protein solubility in order to facilitate deeper proteomic characterization, a range of different solvents were applied to the SPE. These solvents included commonly employed detergents within protein research (SDS and SDC) as well as a slightly alkaline buffer (ABC) at different concentrations. Detergents, buffer components, and their concentration were selected based on compatibility with other employed analytical methods. Moreover, a slightly alkaline pH improved aqueous solubility of SPE in accordance with previous studies on seaweed protein (Veide Vilg and Undeland, 2017; Juul et al., 2022). In contrast to aq. solubilization, Qubit proved inferior due to low compatibility with applied detergents and buffers, where particularly solvent B (2% SDS in 200 mM ABC) and solvent E (1% SDC in 50 mM TEAB) resulted in higher response for solvent references than for the samples (Table A.2). This resulted in a negative protein solubility and is regarded unreliable. Lower detergent and buffer concentrations, i.e. solvent C (0.2% SDS in 50 mM ABC), appeared compatible with Qubit, and significantly ($P<0.05$) improved protein solubility (7.4%) compared to both aqueous solubilization with no pH adjustment (solvent D. 2.2%) and slightly alkaline buffer in solvent A (3.0%). Solvent A improved SPE solubility slightly, but significantly ($P>0.05$) compared to aq. solubility (solvent D) based on Qubit, but both were within range of pH adjusted aq. solubility (Table A.1).

Based on ND analysis of TSP, the addition of detergent, both SDS and SDC, improved protein solubility substantially ($P<0.05$). While SDC facilitated lower solubility (19.5%) of SPE compared to SDS (24.7–27.2 %) (Table A.2), in line with previous studies (Lin et al., 2014), the low content of solubilized protein, evaluated by SDS-PAGE (Fig. A.1, lane 10) was surprising. Nevertheless, the SDC/ABC solvent system was selected for further processing due to compatibility with enzymatic digest and the possibility to selectively precipitate the detergent prior to LC-MS/MS proteomics analysis. This is not possible with e.g. SDS, which is notoriously incompatible with LC (Lin et al., 2008). Moreover, SDC-assisted protein solvation has been reported to improve the number of identified proteins and peptides compared to dilute SDS (Lin et al., 2014); particularly for membrane-associated proteins (Zhou et al., 2006). This could be of particular interest in the characterization of seaweed proteins, where many of the presumably abundant phycobiliproteins (Tandeau De Marsac et al., 2003; Francavilla et al., 2013; Lage-Yusty et al., 2013) are reported to be associated with e.g. thylakoid membranes (Hess et al., 1999).

Regardless of the method used to quantify TSP, the solubility was determined relative to the total protein determined by total-N and a N-to-protein conversion using the universal Jones Factor (Jones, 1931) of 6.25. Although the Jones factor is routinely used in food protein science and industry, species dependent variations have been determined (Mariotti, et al., 2008). For seaweeds in general, this has been speculated to be substantially lower and a global seaweed conversion factor of 5 has been proposed (Angell et al., 2016). Moreover, the conversion factor is based on the whole organism, and an extract may not be representative hereof. It was previously demonstrated how N-to-protein can substantially overestimate protein content (Gregersen et al., 2021), and the obtained solubility of SPE here may therefore be substantially higher than what has been determined using both Qubit and ND.

Proteomics analysis of SPE

Using bottom-up proteomics by LC-MS/MS, we investigated the protein composition of the SPE. Following removal of contaminants and false positive (decoy) hits, we identified 298 protein groups composed of 1,380 potential proteins from the *Gigartinales* protein database, across three technical replicates (Table A.3). The protein groups were identified based on 602 peptide identifications across replicates, following filtering of contaminant and reverse peptides. The vast majority of the peptides (590) were identified with Andromeda scores >40, indicating an overall high level of confidence (Table A.4). The proportion of Andromeda scores >40 were substantially lower on the protein level, where only 68 of the 298 groups had a score >40 (Table A.3). This can primarily be ascribed to low number of peptides within each group (and thus low sequence coverage), as only 84 of the identified protein groups contained >2 peptides (including razor peptides) and only 59 groups contained >2 unique peptides. Of the 298 identified protein groups, 216 (72%) were identified in all three replicates. 38 (13%) in two of three replicates, 28 (9%) in only one replicate, while 16 (5%) were not quantified (Fig. 2A). Overall, there was good reproducibility between replicates with a Pearson correlation coefficient (PCC) >0.98 based on riBAQ quantification (Fig.2B) and PCC >0.97 based on log2 LFQ intensities (Fig 2C). Although the analysis was not performed using a *Gigartina* reference proteome, as such is not currently available, the completeness of the reference proteome for the highly related *Chondrus crispus* (UP000012073), is expected to facilitate sufficient coverage in the applied protein database. The two Rhodophyta species have been co-examined in previous studies showing similar traits (Mathieson and Burns, 1971; Chen et al., 1973; Lipinska et al., 2020) as well as being within close phylogenetic distance, particularly for Gigartina skottsbergii (Hommersand et al., 1994; Chopin et al., 1999).

Eighteen protein groups are considered highly abundant (riBAQ > 1%) and together constitute 79% of the mean, molar protein content in the SPE (Table 1). With the exception of one protein group (M5DD57), the abundant protein groups are found with good quantitative reproductivity between replicates (CV < 23%). Overall, the quantitative reproducibility for the analysis is high, with a stochastic mean CV of 45% and an abundance-weighted mean CV of 15% (Table A.3), illustrating

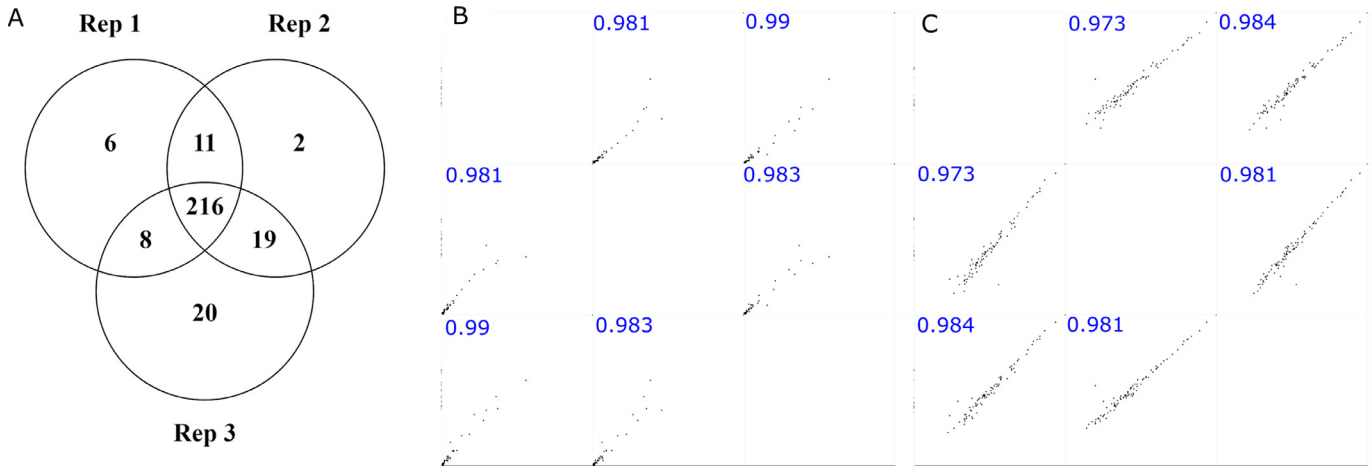

Fig. 2. Reproducibility between technical replicates for LC-MS/MS. Venn diagram showing overlap of identified proteins between sample replicates (A) and scatter plots of quantitative reproducibility using riBAQ (B) and MaxLFQ (C) quantification. For scatter plots, Pearson correlation between individual replicates is indicated in blue.

Table 1
Summary of proteomics data and subcellular localization for the most abundant (>1%) protein (groups) identified in the *Gigartina* SPE. Protein (groups) are annotated by their lead protein ID (Uniprot AC#), name, molecular weight (MW), relative abundance (riBAQ), reproducibility (CV), number of proteins in the group (PIG), identified peptides, experimental sequence coverage (Seq. Cov., %) as well as subcellular localization as predicted by deepLoc and LocTree in addition to Uniprot-annotated compartment.

Lead Protein ID	Protein Name	MW	riBAQ	CV	PIG	Peptides	Seq. Cov.	DeepLoc	LocTree	Uniprot
I4DEE4	Phycoerythrin alpha chain	17.7	18.4%	8%	5	5	34.8	CYT	NUC	CHLM
I4DEE3	Phycoerythrin beta chain	18.5	9.4%	22%	11	8	42.9	CYT	CYT	CHLM
M5DDH2	Cytochrome b6-f complex subunit 5	3.96	7.3%	10%	9	1	16.2	MIT	MITM	THYM
M5DDK1	Phycocyanin, alpha chain	17.6	6.9%	20%	10	7	45.1	CYT	NUC	CHLM
M5DD55	Photosystem II reaction center protein L	4.36	6.8%	8%	9	2	36.8	GLG	ERM	THYM
A0A6H1U7F2	Photosystem II protein Y	3.82	5.4%	5%	6	1	20.6	MIT	EXT	CHLM
A0A7M3VH79	RuBisCO small subunit	16.1	5.2%	14%	4	4	35.5	CYT	CHL	CHL
A0A6H1U732	Allophycocyanin, alpha subunit	17.5	3.8%	9%	6	7	52.8	CYT	NUC	CHLM
A0A6H1U964	Phycocyanin, beta subunit	18.2	3.1%	11%	6	8	58.1	CYT	NUC	CHLM
Q9TN12	RuBisCO large chain	53.2	2.0%	15%	37	10	23.5	PER	CHL	CHL
M5DD57	Photosystem I reaction center subunit XII	3.30	1.8%	52%	2	1	23.3	EXT	EXT	THYM
M5DDA6	Cytochrome c6	11.5	1.7%	4%	3	1	8.3	EXT	CHL	CHLM
M5DD47	Photosystem I iron-sulfur center	8.77	1.6%	8%	7	2	22.2	EXT	EXT	THYM
M5DDG4	Allophycocyanin, alpha chain	17.4	1.5%	7%	4	6	41.6	CYT	NUC	CHLM
Q1AP03	RuBisCO large subunit (Fragment)	1.57	1.0%	9%	1	1	100	PLA	N/A	CHL
M5DDI6	Allophycocyanin, beta chain	17.6	1.0%	23%	10	6	45.3	CYT	NUC	CHLM
A0A7M1VMM8	RuBisCO large subunit	54.2	1.0%	22%	197	9	20.5	CYT	CHL	CHL
A0A342RZ93	Cytochrome c-550	18.1	1.0%	8%	9	3	16	ER	EXT	THYM

Abbreviations for subcellular localization annotation: Cytosol (CYT), mitochondria (MIT), mitochondrial membrane (MITM), golgi apparatus (GLG), peroxisome (PER), extracellular/secreted (EXT), plastid (PLA), endoplasmatic reticulum (ER), endoplasmatic reticulum membrane (ERM), nucleus (NUC), chloroplast (CHL), chloroplast membrane (CHLM), thylakoid membrane (THYM).

that the higher stochastic variability can primarily be ascribed to lower abundance protein groups.

Among the abundant proteins, representing 79% of the identified protein, phycoerythrin (both subunits) constitute 27.9% of the total protein, while phycocyanin and allophycocyanin constitute 10.0% and 6.3%, respectively (Table 1). Ribulose bisphosphate carboxylase (RuBisCO) accounts for 9.2%, while photosystems proteins and cytochromes constitute 15.6% and 9.9%, respectively. That the majority of proteins are related to light harvest is not surprising, as phycobiliproteins are reported to be highly abundant in red seaweeds (Tandeau De Marsac et al., 2003; Francavilla et al., 2013; Lage-Yusty et al., 2013), and RuBisCO is found abundantly in most autotrophic organisms (Andersson and Backlund, 2008). Moreover, our data corporate the general understanding that of the phycobiliproteins, phycoerythrin is the most abundant (Glazer, 1994; Tandeau De Marsac et al., 2003; Pina et al., 2014). In addition, phycoerythrin γ-subunit (R7Q9W8) as well as phycobilisome linker-protein (M5DET7) and anchor-proteins (A0A7M3VH69 and M5DBU8) were identified in lower abundance, strengthening that phycobilisomes were successfully extracted (Table A.3). The high abundance of phycobiliprotein subunits, which all have similar molecular weight (Table 1), correspond with the high intensity band observed in SDS-PAGE analysis (Fig. A.1).

Phycobiliproteins are considered heterodimers of the α- and β-subunits assembled into larger complexes within the phycobilisomes (Ritter et al., 1999), but the equimolar stoichiometry is not reflected in our data. Rather, the α-subunits seem to be substantially enriched (Table 1). This is found for all major phycobiliproteins, where the α:β ratio can be determined for phycoerythrin (1.96:1), phycocyanin (2.23:1), and allophycocyanin (5.20:1). While non-equimolar stoichiometry (3:2) has been suggested for allophycocyanin (Reuter et al., 1990), this does not change that α-subunits appear enriched, suggesting some degree of selectivity in the extraction method or different solubilities of the subunits. This is not observed for RuBisCO, which is known to exist in a heterohexadecamer form but also with equimolar stoichiometry be-

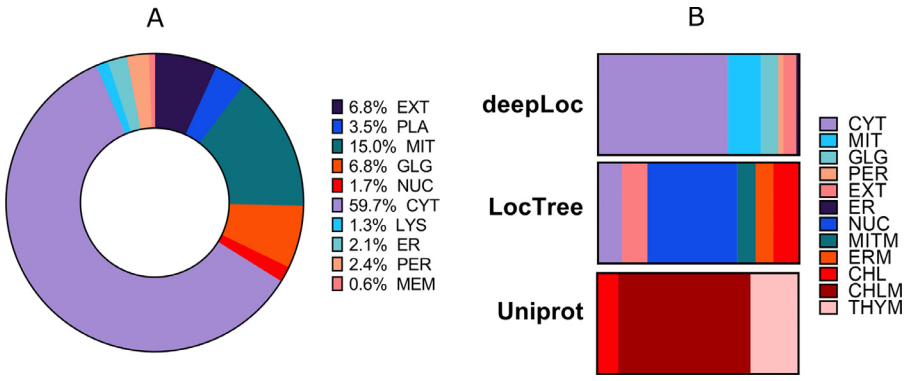

Fig. 3. Subcellular distribution of all identified proteins by deepLoc (A) and abundant (riBAQ > 1%) proteins using different localization prediction algorithms and annotated localization from Uniprot protein homologues (B). Abbreviations for subcellular localization annotation: Cytosol (CYT), mitochondria (MIT), mitochondrial membrane (MITM), golgi apparatus (GLG), peroxisome (PER), extracellular/secreted (EXT), plastid (PLA), endoplasmatic reticulum (ER), endoplasmatic reticulum membrane (ERM), nucleus (NUC), chloroplast (CHL), chloroplast membrane (CHLM), thylakoid membrane (THYM).

tween the large and small subunits (Hauser et al., 2015; Gruber and Feiz, 2018). Here we find the large-to-small ratio for RuBisCO subunits to be 0.91:1 and thus very close to the expected.

Subcellular localization and evaluation of protein extractability

To investigate if the proteomic composition was indeed indicative of intracellular, and thereby efficient, protein extraction, we investigated the subcellular localization of the proteome using both annotated compartments from Uniprot and predictive bioinformatics. From subcellular localization prediction (Table 1), a large discrepancy between prediction algorithms and the annotated compartments in Uniprot is observed for the most abundantly quantified proteins. This may be ascribed to the different architectures and approaches employed in the different models. Nevertheless, as identified proteins are imputed from related seaweeds with proper Uniprot annotation, this is considered a more viable interpretation of the origin of the abundant proteins. As the thylakoid system is not explicitly included in deepLoc, and because the phycobiliproteins have been suggested not to be directly incorporated into the thylakoid membrane, but rather exist in phycobilisomes with larger mobility (Stadnichuk and Tropin, 2017), this may be the reason as to why deepLoc annotates such a large proportion of the protein as cytosolic.

Although prediction therefor may be subject to uncertainties, summarizing across all predictions using deepLoc, the amount of predicted extracellular protein account for only 6.8% of the total protein identified (Fig. 3.A). Moreover, looking only at the most abundant proteins (Fig.3.B), extracellular proteins constitute only a minor part of the extracted proteins according to both deepLoc (5.1%) and LocTree (9.8%) while none are annotated as extracellular or secreted in Uniprot. In a previous study of extracts from *E. denticulatum,* a similar extraction principle was employed, where the soluble fraction following high temperature aq. carrageenan extraction was characterized (Gregersen et al., 2022). In that study, the soluble fraction comprised almost exclusively (97%) extracellular protein. This is in agreement with our findings, where extracellular protein content in the residual solid fraction is depleted and the SPE is constituted almost exclusively by non-extracellular protein regardless of considering subcellular prediction or annotated compartments in Uniprot. Ultimately, this indicates that the SPE is a good source of extractable cellular protein. In particular, phycobiliproteins, RuBisCO, and other light harvest-related proteins.

Amino acid distribution and nutritional potential

To assess nutritional potential of the SPE, protein- and peptide-level data was used to compute the relative abundance of individual amino acids based on the mean riBAQ and mean relative MS1 peptide intensity, as previously described (Jafarpour et al., 2020; Gregersen et al., 2021, 2022).

Overall, we found good agreement between the use of protein- and peptide-level data for estimating AA composition (Fig. 4). Moreover, we find that, with the exception of His, the SPE contains sufficient EAA levels relative to the minimum requirements by the FAO (WHO/FAO/UNU, 2007). The calculated EAA:NEAA ratio for SPE is 0.69 and 0.56 based of protein-level and peptide-level data, respectively. This is in line with red seaweeds in general (0.63), but still somewhat lower than other alternative protein sources (0.78–1.09), thereby decreasing the nutritional potential (Mæhre et al., 2014; Donadelli et al., 2019). The SPE contains high levels of branched chain AAs (BCAAs), namely Leu, Ile, and Val, corresponding to 24.2% and 21.7% based on protein- and peptide-level data, respectively. High levels of individual EAAs, in particular the BCAAs, is highly desired in fields such as sports nutrition (Campbell et al., 2007). Moreover, BCAAs are linked with decreased food intake (Laeger et al., 2014) and linked to numerous metabolic and physiological functions such as mammary and embryotic health, intestinal development, and immune response regulation (Zhang et al., 2017), and SPE may therefor also be relevant in pregnancy and infant nutrition.

Although the chemical composition of *Gigartina* has been reported not to vary greatly over season (Amimi et al., 2007), the seasonal variation of the proteomic composition has not been investigated. Based on investigations of other red seaweeds such as *Palmaria palmata*, substantial seasonal variation in both protein content and composition are likely (Fleurence, 1999). Quantitative variations of the *Gigartina* proteome will also affect the AA distribution, and should therefore be investigated if moving forward with the seaweed as a source of food protein. As MS-based proteomic determination of AA distribution is still in its infancy and thus related to some uncertainty (Jafarpour et al., 2020; Gregersen et al., 2021, 2022), supporting the analysis with more conventional AA analysis is advisable.

Abundant Gigartina proteins as source of emulsifier peptides

While the *Gigartina* SPE displays some potential as food protein, the high content of phycobiliproteins may also directly imply other potential applications. Phycobiliproteins are membrane-associated (Hess et al., 1999) and furthermore have a very high content of α-helical domains (Apt et al., 1995). Such proteins are known to have highly amphiphilic structural features (Fernández-Vidal et al., 2007), and previous studies have shown that helical and amphiphilic peptides have enormous potential as natural emulsifiers (Malcolm et al., 2007; Dexter and Middelberg, 2008; García-Moreno et al., 2021; Yesiltas et al., 2021). While RuBisCO is primarily localized in the chloroplast stroma (Gruber and Feiz, 2018), it has been shown to possess emulsifying properties in the native form (Di Stefano et al., 2018; Tan et al., 2022). To investigate this potential further, we applied the *in silico* predictive tool, EmulsiPred (García-Moreno et al., 2020; García-Moreno et al., 2020) on the highly abundant SPE proteins.

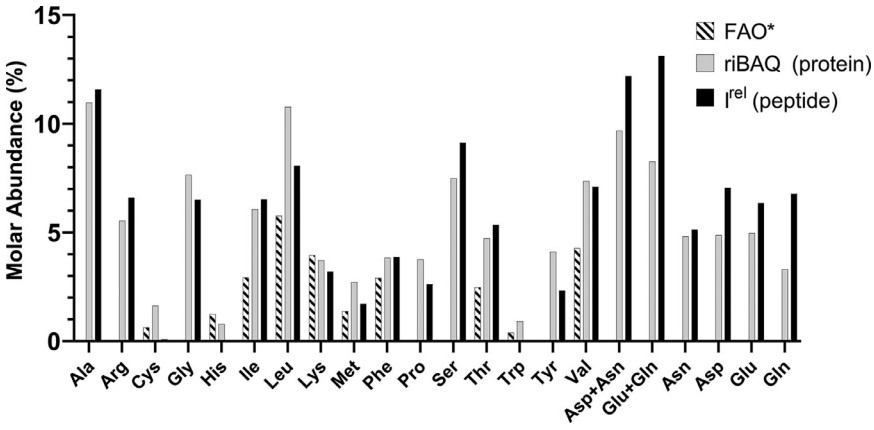

Fig 4. Estimation of amino acid (AA) composition (as relative molar abundance) using peptide- and protein-level MS data for SPE. For reference, minimum requirements according to the FAO* (WHO/FAO/UNU, 2007) are indicated.

Table 2
Summary of bioinformatic prediction of emulsifier peptides in abundant (riBAQ > 1%) SPE proteins. Each protein (group) is annotated with the lead protein ID (Uniprot AC#) and name along with the relative molar abundance in SPE (riBAQ). For each protein (group), the number of potential (score > 2) emulsifier peptides with facial (α and β) or axial (γ) amphiphilicity, as well as total potential emulsifier peptides are indicated. For very high abundance (riBAQ > 5%), the highest scoring peptides within each class (two for α and γ, one for β) are shown by score, type, and sequence.

			Predicted peptides				Selected highest scoring peptide		
Lead Protein ID	Protein Name	riBAQ	α	β	γ	Total	Score	Type	Sequence
I4DEE4	Phycoerythrin alpha chain	18.4%	354	0	214	568	3.90	α	EKVNKCYRDIDHYMRLINYAL
I4DEE3	Phycoerythrin beta chain	9.4%	214	0	460	674			
M5DDH2	Cytochrome b6-f complex subunit 5	7.3%	3	0	183	186	5.16	γ	IVLGLIPVTLAGLLVAAYLQYRRGNQ
M5DDK1	Phycocyanin, alpha chain	6.9%	458	0	66	524	3.88	α	EAAKSLTNNAQRLITGAAQAV
M5DD55	Photosystem II reaction center protein L	6.8%	0	0	293	293	5.84	γ	NPNKEPVELNRTSLFWGLLLIFVLAVLF
A0A6H1U7F2	Photosystem II protein Y	5.4%	57	0	41	98			
A0A7M3VH79	RuBisCO small subunit	5.2%	20	82	588	690	3.95	β	NVYIKIN
A0A6H1U732	Allophycocyanin, alpha subunit	3.8%	162	0	72	234			
A0A6H1U964	Phycocyanin, beta subunit	3.1%	175	0	190	365			
Q9TN12	RuBisCO large chain	2.0%	467	29	1052	1548			
M5DD57	Photosystem I reaction center subunit XII	1.8%	0	0	122	122			
M5DDA6	Cytochrome c6	1.7%	236	0	163	399			
M5DD47	Photosystem I iron-sulfur center	1.6%	0	36	34	70			
M5DDG4	Allophycocyanin, alpha chain	1.5%	266	0	141	407			
Q1AP03	RuBisCO large subunit (Fragment)	1.0%	0	0	0	0			
M5DDI6	Allophycocyanin, beta chain	1.0%	143	0	120	263			
A0A7M1VMM8	RuBisCO large subunit	1.0%	430	30	1113	1573			
A0A342RZ93	Cytochrome c-550	1.0%	139	0	456	595			
	Multiple target proteins		500	25	586	1111			

For the most abundant SPE proteins (Table 1), we predicted a total of 7684 potential emulsifier peptides (score > 2), whereof 1,111 map to multiple target proteins (Table 2). Not surprisingly, the majority of predicted emulsifier peptides are found in α-helical or random (γ) conformation, as light harvest proteins are highly helical by nature. The highest amount of emulsifier peptides (>3000) were predicted from the large subunit of RuBisCO (Q9TN12 and A0A7M1VMM8), with the majority being in γ conformation although a substantial amount were also predicted in helical conformation (ratio γ:α ~2:1). The large number of RuBisCO-derived peptides should also be considered in the light that this protein is also approximately three times the size of phycobiliprotein subunits. Phycobiliproteins were predicted to contain 263–674 potential emulsifier peptides (ratio γ:α ~1:1) per subunit, with the lowest amount found in allophycocyanin, which is also the lowest abundance phycobiliprotein. Photosystem-related proteins and cytochromes are generally enriched in γ-peptides. To investigate high potential emulsifier peptides with a potentially high downstream yield, we focused only on very high abundance (riBAQ > 5%) proteins (Table 2). Looking at the highest scoring α-peptides (Table A.5), specific regions of the α-subunit in phycoerythrin (I4DEE4) and phycocyanin (M5DDK1) stand out (scores 3.6–3.9). For β- peptides, the small RuBisCO subunit (A0A7M3VH79) shows a particularly interesting domain with multiple peptides having scores in the range 3.1–4.0 (Table A.6). In the case of γ-peptides (Table A.7), particular regions from Photosystem II reaction center protein L (M5DD55) and Cytochrome b6-f complex subunit 5 (M5DDH2) are predicted to contain many peptide isoforms of very high emulsification potential (scores 4.5–5.8). Using EmulsiPred, γ-peptides generally have higher prediction scores (García-Moreno et al., 2020; Jafarpour et al., 2020; Yesiltas et al., 2021; Gregersen Echers et al., 2022). This can be attributed to the axial mode of emulsification compared to α- and β-peptides, which exert emulsification through facial amphiphilicity, thereby making the score difficult to compare directly. Compared to previous studies using EmulsiPred prediction (with the same version) and peptide-level *in vitro* functional validation, these scores indicate that the SPE could indeed have very high potential for production of functional, protein-based ingredients (Yesiltas et al., 2021). As all these proteins also have very high abundance (riBAQ > 5.2%), they could serve as sources for release of emulsifier peptides in high total yields.

To better understand and evaluate if predicted peptides are indeed potential emulsifiers, we applied templated homology-modelling of the native protein structure. The conformation of particularly amphiphilic α-peptides within the native protein, has previously been shown to be a good model for oil/water interfacial structure (García-Moreno et al., 2021; Yesiltas et al., 2021). Modelling was performed for the five proteins indicated above to contain particularly high scoring peptides (Fig. 5) yielding high quality models for all protein (Table A.8).

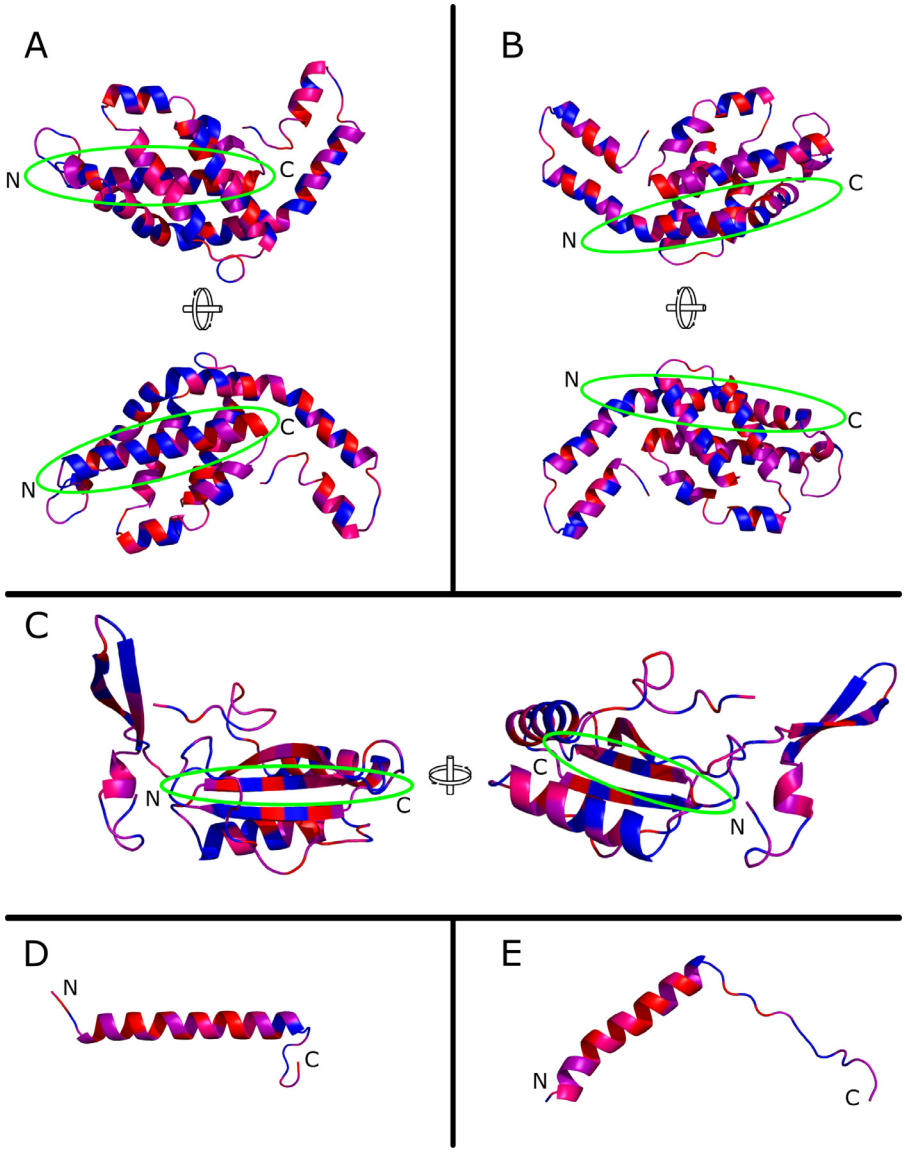

Fig. 5. Structural modelling of proteins embedding particularly high scoring emulsifying peptides. A) Phycoerythrin alpha subunit, B) Phycocyanin alpha subunit, C) RuBisCO small subunit, D) Photosystem II reaction center protein L, and E) Cytochrome b6-f complex subunit 5. For structures A-C, the region containing the highest scoring predicted emulsifier peptides has been highlighted and the N-, and C-termini of the regions are indicated. For structures D and E, predicted emulsifier peptides span the entire (short) length and the N- and C-termini are indicated for the full length proteins.

For phycoerythrin (Fig. 5A), the region containing high scoring peptides is indeed also a highly amphiphilic, surface-exposed helix in the native protein, underlining the emulsifying potential of this protein region. Moreover, the C-terminus of the region is highly hydrophobic, which has been speculated to improve the emulsification potential om amphipathic helixes by serving as an anchor for the oil-phase (Gregersen Echers et al., 2022). In Phycocyanin (Fig. 5B), a highly amphiphilic, surface-exposed helix is also found within the region of interest. Moreover, the region includes a kink within the helix, which is known to play a role in e.g. membrane interactions of antimicrobial peptides (Tuerkova et al., 2020) and have been proposed to also facilitate improved interfacial stabilization in emulsions (García-Moreno et al., 2020). For the small RuBisCo subunit (Fig. 5C), the structure consists of more β-strand, why the large number of β-peptides here from could be expected. High scoring (>3) β-peptides are rather short (7-12 AAs), which may be insufficient to produce a stable emulsion (García-Moreno et al., 2020). Photosystem II reaction center protein L and Cytochrome b6-f complex subunit 5 (Fig. 5D, 5E) are both short, helical proteins involved in larger quaternary complexes. Nevertheless, both were predicted to contain high scoring γ-peptides spanning all of the sequence length. γ-peptides represent axial amphiphilicity, and thus their secondary structure is not crucial for functionality (García-Moreno et al., 2020; García-Moreno et al., 2020), although helical structure in γ-peptides has been seen to facilitate formation of highly stable emulsions (García-Moreno et al., 2021; Yesiltas et al., 2021). Moreover, both proteins are annotated to contain transmembrane helices (UniProt Consortium et al., 2021), indicating that the hydrophobic helix is indeed capable of inserting into a lipid phase. The *in vitro* functionality of the peptides should be further investigated in model emulsion systems to fully determine their potential. Recently, we presented a workflow, where using a basis of quantitative proteomics and bioinformatics, we designed a targeted hydrolysis process where we were able to release emulsifier peptides from potato protein (Gregersen Echers et al., 2022). Such an approach could be valuable in exploring if the *Gigartina* SPE is indeed a good source of emulsifying peptides, releasable through enzymatic hydrolysis. As high scoring peptides were not only restricted to the modelled proteins, other high abundance proteins may also contribute to overall emulsifying properties of a hydrolysate by release of amphiphilic peptides. For instance, RuBisCO large chain had both high scoring α-peptides (3.37) and γ-peptides (4.68) (Tables 2, A.6, A.8), which should also be investigated

further. As the cumulative abundance of RuBisCO large chain isoforms approaches 5% (Table A.3), this subunit may also be a source of potential emulsifier peptides with high yield.

Conclusion

Seaweeds, in particular Rhodophytae, are massively exploited for industrial hydrocolloid production. Side-streams from this production remains underutilized and may be used as a source of food protein and functional ingredients. In pilot scale, a seaweed protein extract (SPE) from *Gigartina radula* was isolated and investigated for its protein composition and quality. The aqueous solubility of SPE was very low (<2%), but did still display pH-dependent solubility, which could be related to the conditions employed during seaweed processing and protein extraction. The addition of detergent improved protein solubility, allowing for visualization by SDS-PAGE and bottom-up shotgun proteomics analysis by LC-MS/MS. Using all known proteins within the *Gigartinales* order as reference database, 298 protein groups were identified and through relative quantification, the most abundant proteins were determined. In general, proteins related to light harvest dominate the SPE and particularly phycobiliproteins (44%) and RuBisCO (9%) were abundant. Within the phycobiliproteins, phycoerythrin constituted the majority of the quantified protein (28%). The expected equimolar stoichiometry of α- and β-subunits of phycobiliproteins was not observed (α-subunit favored), indicating a potential bias in the extraction, whereas the large and small subunits of RuBisCO were identified in equimolar stoichiometry as expected. Investigating the subcellular localization of identified proteins, the extraction method was deemed sufficient for release of intracellular protein. The SPE generally showed sufficient levels (36–41%) of essential amino acids (with the exception of His), indicating that it may be of potential use as a nutritional food protein, although an EAA:NEAA ratio of 0.56–0.69 is somewhat lower than other alternative proteins. High levels of branched chain amino acids (22–24%) indicate potential relevance in e.g. sports and infant nutrition. From the most abundant proteins in the SPE, almost 8000 potential emulsifier peptides were predicted, including many very high potential peptides based on amphiphilic score. Through templated homology modelling of the native protein structure, this potential was verified as the peptides also have similar amphiphilic structure within the native proteins. With this study, we provide fundamental new knowledge within seaweed protein science and the first study of the *Gigartina* proteome. Moreover, we expect that *Gigartina* could be a valuable source of protein-derived natural emulsifiers for foods and this route should be explored further, also investigating the functional properties of the crude extract for comparison.

Funding

This work was supported by Innovation Fund Denmark (Grant number 7045-00021B (PROVIDE)).

Data Availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Perez-Riverol et al., 2019) with the dataset identifier PXD034435 and DOI 10.6019/PXD034435. Data not supplied in PRIDE or appended as supplementary information, may be obtained upon request to the corresponding author.

Ethical Statement

There are no ethical aspects applicable for the study

Declaration of Competing Interest

Jimmy J. P. Sejberg is employed by CP Kelco ApS. All other authors declare no conflicts of interest.

CRediT authorship contribution statement

Simon Gregersen Echers: Conceptualization, Methodology, Investigation, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Naim Abdul-Khalek:** Investigation, Formal analysis, Writing – review & editing. **Rasmus Kranold Mikkelsen:** Formal analysis, Writing – review & editing. **Susan Løvstad Holdt:** Conceptualization, Resources, Writing – review & editing. **Charlotte Jacobsen:** Conceptualization, Resources, Writing – review & editing, Funding acquisition. **Egon Bech Hansen:** Conceptualization, Resources, Writing – review & editing, Funding acquisition. **Tobias Hegelund Olsen:** Methodology, Investigation, Formal analysis, Data curation, Writing – review & editing. **Jimmy J.P. Sejberg:** Methodology, Investigation, Validation, Formal analysis, Resources, Writing – review & editing. **Michael Toft Overgaard:** Conceptualization, Resources, Writing – review & editing, Funding acquisition, Supervision.

Data Availability

The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD034435. Other data may be obtained upon request.

Acknowledgements

The authors would like to thank CP Kelco (Lille Skensved, Denmark) and in particular, Laboratory Technician Helle Bech Olsen, for supplying the seaweed protein extract.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.fufo.2022.100189.

References

Agrawal, G.K., et al., 2013. A decade of plant proteomics and mass spectrometry: translation of technical advancements to food security and safety issues. In: Mass Spectrometry Reviews, 32. John Wiley & Sons, Ltd, pp. 335–365. doi:10.1002/MAS.21365.

Almagro Armenteros, J.J., et al., 2017. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 33 (21), 3387–3395. doi:10.1093/bioinformatics/btx431, Edited by J. Hancock.

Amagliani, L., et al., 2017. The composition, extraction, functionality and applications of rice proteins: a review. Trends Food Sci. Technol. 64, 1–12. doi:10.1016/J.TIFS.2017.01.008, Elsevier.

Amimi, A., et al., 2007. Seasonal variations in thalli and carrageenan composition of gigartina pistillata (Gmelin) stackhouse (Rhodophyta, gigartinales) harvested along the atlantic coast of morocco. Phycol. Res. 55 (2), 143–149. doi:10.1111/J.1440-1835.2007.00457.X, John Wiley & Sons, Ltd.

Andersson, I., Backlund, A., 2008. Structure and function of rubisco. Plant Physiol. Biochem. 46 (3), 275–291. doi:10.1016/J.PLAPHY.2008.01.001, Elsevier Masson.

Andjelković, U., Josić, D, 2018. Mass spectrometry based proteomics as foodomics tool in research and assurance of food quality and safety. Trends Food Sci. Technol. 77, 100–119. doi:10.1016/J.TIFS.2018.04.008, Elsevier.

Angell, A.R., et al., 2016. The protein content of seaweeds: a universal nitrogen-to-protein conversion factor of five. J. Appl. Phycol. 28 (1), 511–524. doi:10.1007/s10811-015-0650-1.

Apt, K.E., Collier, J.L., Grossman, A.R., 1995. Evolution of the phycobiliproteins. J. Mol. Biol. 248 (1), 79–96. doi:10.1006/JMBI.1995.0203, Academic Press.

Armengaud, J., 2016. Next-generation proteomics faces new challenges in environmental biotechnology. Curr. Opin. Biotechnol. 38, 174–182. doi:10.1016/J.COPBIO.2016.02.025, Elsevier Current Trends.

Armenteros, J.J.A., et al., 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks'. Nat. Biotechnol. 37 (4), 420–423. doi:10.1038/s41587-019-0036-z, Nature Publishing Group.

Aschemann-Witzel, J., et al., 2020. Plant-based food and protein trend from a business perspective: markets, consumers, and the challenges and opportunities in the future. Crit. Rev. Food Sci. Nutr. 1–10. doi:10.1080/10408398.2020.1793730, Taylor and Francis Inc..

Ashaolu, T.J., 2020. Applications of soy protein hydrolysates in the emerging functional foods: a review. Int. J. Food Sci. Technol. 55 (2), 421–428. doi:10.1111/IJFS.14380, John Wiley & Sons, Ltd.

Beaulieu, L., 2019. Insights into the regulation of algal proteins and bioactive peptides using proteomic and transcriptomic approaches. Molecules 24 (9), 1708. doi:10.3390/MOLECULES24091708, 2019, Vol. 24, Page 1708. Multidisciplinary Digital Publishing Institute.

Biancarosa, I., et al., 2016. Amino acid composition, protein content, and nitrogen-to-protein conversion factors of 21 seaweed species from Norwegian waters. J. Appl. Phycol. 29 (2), 1001–1009. doi:10.1007/S10811-016-0984-3, Springer.

Bleakley, S., Hayes, M., 2017. Algal proteins: extraction, application, and challenges concerning production. Foods 6 (5), 33. doi:10.3390/foods6050033, MDPI AG.

Brown, E.M., et al., 2014. Seaweed and human health. Nutr. Rev. 72 (3), 205–216. doi:10.1111/NURE.12091, Oxford Academic.

Cai, J., Lovatelli, A., Aguilar-Manjarrez, J., Cornish, L., Dabbadie, L., Desrochers, A., Diffey, S., Garrido Gamarro, E., Geehan, J., Hurtado, A., Lucente, D., Mair, G., Miao, W., Potin, P., Przybyla, C., Reantaso, M., Roubach, R., Tauati, M., Yuan, X. 2021. Seaweeds and microalgae: an overview for unlocking their potential in global aquaculture development. FAO Fisheries and Aquaculture Circular No. 1229. Rome, FAO. doi:10.4060/cb5670en.

Campbell, B., et al., 2007. International society of sports nutrition position stand: protein and exercise. J. Int. Soc. Sports Nutr. 4 (1), 1–7. doi:10.1186/1550-2783-4-8/METRICS, BioMed Central.

Chen, L.C.M., et al., 1973. The ratio of kappa- to lambda-carrageenan in nuclear phases of the rhodophycean algae, chondrus crispus and gigartina stellata. J. Mar. Biol. Assoc. UK 53 (1), 11–16. doi:10.1017/S0025315400056599, Cambridge University Press.

Chopin, T., Kerin, B.F., Mazerolle, R., 1999. Phycocolloid chemistry as a taxonomic indicator of phylogeny in the gigartinales, rhodophyceae: a review and current developments using fourier transform infrared diffuse reflectance spectroscopy. Phycol. Res. 47 (3), 167–188. doi:10.1046/J.1440-1835.1999.00170.X, John Wiley & Sons, Ltd.

Chrapusta, E., et al., 2017. Mycosporine-like amino acids: potential health and beauty ingredients. Mar. Drugs 15 (10), 326. doi:10.3390/MD15100326, 2017, Vol. 15, Page 326. Multidisciplinary Digital Publishing Institute.

Colusse, G.A., et al., 2022. Challenges and recent progress in seaweed polysaccharides for industrial purposes. Sustain. Global Res. Seaweeds 2, 411–431. doi:10.1007/978-3-030-92174-3_22, Volume 2. Cham: Springer, Cham.

Cotas, J., et al., 2020. Seaweed phenolics: from extraction to applications. Mar. Drugs 18 (8), 384. doi:10.3390/MD18080384, 2020, Vol. 18, Page 384. Multidisciplinary Digital Publishing Institute.

Cotas, J., et al., 2022. Food applications and health benefits of the genus (Rhodophyta). Sustain. Global Res. Seaweeds 2, 135–144. doi:10.1007/978-3-030-92174-3_6, Volume 2. Cham: Springer, Cham.

Cox, J., et al., 2014. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed maxlfq. Mol. Cell. Proteom. 13 (9), 2513–2526. doi:10.1074/MCP.M113.031591/ATTACHMENT/B414BAF4-20AE-46D2-BB85-A06FBF3A11C3/MMC1.ZIP, American Society for Biochemistry and Molecular Biology Inc..

Cox, J., Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26 (12), 1367–1372. doi:10.1038/nbt.1511.

Custódio, M., et al., 2017. Unravelling the potential of halophytes for marine integrated multi-trophic aquaculture (IMTA)— a perspective on perfor. mance, opportunities and challenges. Aquacult. Environ. Interact. 9, 445–460. doi:10.3354/AEI00244, Inter-Research.

Dexter, A.F., Middelberg, A.P.J., 2008. Peptides as functional surfactants. Ind. Eng. Chem. Res. 6391–6398. doi:10.1021/ie800127f.

Di Stefano, E., et al., 2018. Plant rubisco: an underutilized protein for food applications. J. Am. Oil Chem.' Soc. 95 (8), 1063–1074. doi:10.1002/aocs.12104, Wiley-Blackwell.

Donadelli, R.A., et al., 2019. The amino acid composition and protein quality of various egg, poultry meal by-products, and vegetable proteins used in the production of dog and cat diets. Poultry Sci. 98 (3), 1371–1378. doi:10.3382/PS/PEY462, Elsevier.

Fernández-Vidal, M., et al., 2007. Folding amphipathic helices into membranes: amphiphilicity trumps hydrophobicity. J. Mol. Biol. 370 (3), 459–470. doi:10.1016/J.JMB.2007.05.016, Academic Press.

Fleurence, J., 1999. Seaweed proteins: biochemical, nutritional aspects and potential uses. Trends Food Sci. Technol. 10 (1), 25–28. doi:10.1016/S0924-2244(99)00015-1, Elsevier.

Francavilla, M., et al., 2013. The red seaweed gracilaria gracilis as a multi products source. Mar. Drugs 11 (10), 3754–3776. doi:10.3390/MD11103754, 2013, Vol. 11, Pages 3754-3776. Multidisciplinary Digital Publishing Institute.

García Arteaga, V., et al., 2020. Effect of enzymatic hydrolysis on molecular weight distribution, techno-functional properties and sensory perception of pea protein isolates'. Innovat. Food Sci. Emerg. Technol. 65, 102449. doi:10.1016/J.IFSET.2020.102449, Elsevier.

García-Moreno, P.J., et al., 2021. The structure, viscoelasticity and charge of potato peptides adsorbed at the oil-water interface determine the physicochemical stability of fish oil-in-water emulsions. Food Hydrocolloids 115, 106605. doi:10.1016/j.foodhyd.2021.106605, Elsevier.

García-Moreno, P.J., Gregersen, S., et al., 2020. Identification of emulsifier potato peptides by bioinformatics: application to omega-3 delivery emulsions and release from potato industry side streams. Sci. Rep. 10 (1), 690. doi:10.1038/s41598-019-57229-6, Nature Research.

García-Moreno, P.J., Jacobsen, C., et al., 2020. Emulsifying peptides from potato protein predicted by bioinformatics: stabilization of fish oil-in-water emulsions. Food Hydrocolloids 101, 105529. doi:10.1016/j.foodhyd.2019.105529, Elsevier B.V..

Glazer, A.N., 1994. Phycobiliproteins — a family of valuable, widely used fluorophores. J. Appl. Phycol. 6 (2), 105–112. doi:10.1007/BF02186064, 1994 6:2. Springer.

Goldberg, T., et al., 2014. LocTree3 prediction of localization. Nucl. Acids Res. 42 (W1), W350–W355. doi:10.1093/NAR/GKU396, Oxford Academic.

Gómez-Ordóñez, E., Jiménez-Escrig, A., Rupérez, P., 2010. Dietary fibre and physicochemical properties of several edible seaweeds from the northwestern Spanish coast. Food Res. Int. 43 (9), 2289–2294. doi:10.1016/J.FOODRES.2010.08.005, Elsevier.

Granato, D., et al., 2010. Functional foods and nondairy probiotic food development: trends, concepts, and products. Compr. Rev. Food Sci. Food Saf. 9 (3), 292–302. doi:10.1111/J.1541-4337.2010.00110.X, John Wiley & Sons, Ltd.

Gregersen, S., et al., 2021. Enzymatic extraction improves intracellular protein recovery from the industrial carrageenan seaweed eucheuma denticulatum revealed by quantitative, subcellular protein profiling: a high potential source of functional food ingredients. Food Chem.: X 12, 100137. doi:10.1016/J.FOCHX.2021.100137.

Gregersen, S., et al., 2022. Proteomic characterization of pilot scale hot-water extracts from the industrial carrageenan red seaweed eucheuma denticulatum. Algal Res. 62, 102619. doi:10.1016/J.ALGAL.2021.102619, Elsevier.

Gregersen Echers, S., et al., 2022. Targeted hydrolysis of native potato protein: a novel route for obtaining hydrolysates with improved interfacial properties. bioRxiv doi:10.1101/2022.05.25.493405, Cold Spring Harbor Laboratory.

Grønning, A.G.B., Kacprowski, T., Sché Ele, C., 2021. MultiPep: a hierarchical deep learning approach for multi-label classification of peptide bioactivities. Biol. Methods Prot. 6 (1). doi:10.1093/BIOMETHODS/BPAB021, Oxford Academic.

Gruber, A.V., Feiz, L., 2018. Rubisco assembly in the chloroplast. Front. Mol. Biosci. 5 (MAR), 24. doi:10.3389/FMOLB.2018.00024/BIBTEX, Frontiers Media S.A..

Hauser, T., et al., 2015. Role of auxiliary proteins in rubisco biogenesis and function. Nat. Plants 1 (6), 1–11. doi:10.1038/nplants.2015.65, 2015 1:6. Nature Publishing Group.

Hess, W.R., et al., 1999. Phycoerythrins of the oxyphotobacterium prochlorococcus marinus are associated to the thylakoid membrane and are encoded by a single large gene cluster. Plant Mol. Biol. 40 (3), 507–521. doi:10.1023/A:1006252013008, 1999 40:3. Springer.

Holdt, S.L., Kraan, S., 2011. Bioactive compounds in seaweed: functional food applications and legislation. J. Appl. Phycol. 543–597. doi:10.1007/s10811-010-9632-5, Springer.

Hommersand, M.H., Fredericq, S., Freshwater, D.W., 1994. Phylogenetic systematics and biogeography of the gigartinaceae (Gigartinales, rhodophyta) based on sequence analysis of rbcL. Bot. Mar. 37 (3), 193–204. doi:10.1515/BOTM.1994.37.3.193, De Gruyter.

Jafarpour, A., Gomes, R.M., et al., 2020. Characterization of cod (Gadus morhua) frame composition and its valorization by enzymatic hydrolysis. J. Food Compos. Anal. 89, 103469. doi:10.1016/j.jfca.2020.103469, Academic Press Inc..

Jafarpour, A., Gregersen, S., et al., 2020. Biofunctionality of enzymatically derived peptides from codfish (Gadus morhua) frame: bulk in vitro properties, quantitative proteomics, and bioinformatic prediction. Mar. Drugs 18 (12), 599. doi:10.3390/md18120599, Multidisciplinary Digital Publishing Institute.

Jiménez-Escrig, A., Gómez-Ordóñez, E., Rupérez, P., 2012. Brown and red seaweeds as potential sources of antioxidant nutraceuticals. J. Appl. Phycol. 24 (5), 1123–1132. doi:10.1007/S10811-011-9742-8/FIGURES/5, Springer.

Jones, D.B., 1931. Factors For Converting Percentages of Nitrogen in Foods and Feeds Into Percentages of Protein. US Department of Agriculture.

Juul, L., et al., 2022. Combining pressing and alkaline extraction to increase protein yield from ulva fenestrata biomass. Food Bioprod. Process. 134, 80–85. doi:10.1016/J.FBP.2022.05.006, Elsevier.

Laeger, T., et al., 2014. Leucine acts in the brain to suppress food intake but does not function as a physiological signal of low dietary protein. Am. J. Physiol. - Regul. Integr. Comp. Physiol. 307 (3), 310–320. doi:10.1152/AJPREGU.00116.2014/ASSET/IMAGES/LARGE/ZH60151485000007.JPEG, American Physiological Society.

Lage-Yusty, M.-A., Caramés-Adán, P., López-Hernández, J., 2013. Determination of phycobiliproteins by constant-wavelength synchronous spectrofluorimetry method in red algae. In: CyTA - Journal of Food, 11. Taylor & Francis, pp. 243–247. doi:10.1080/19476337.2012.728629.

Lange, L., Meyer, A.S., 2019. Potentials and possible safety issues of using biorefinery products in food value chains. Trends Food Sci. Technol. 84, 7–11. doi:10.1016/J.TIFS.2018.08.016, Elsevier.

Lexhaller, B., Colgrave, M.L., Scherf, K.A., 2019. Characterization and relative quantitation of wheat, rye, and barley gluten protein types by liquid chromatography–tandem mass spectrometry. Front. Plant Sci. 10, 1530. doi:10.3389/FPLS.2019.01530/BIBTEX, Frontiers Media S.A..

Lin, Y., et al., 2008. Sodium-deoxycholate-assisted tryptic digestion and identification of proteolytically resistant proteins. Anal. Biochem. 377 (2), 259–266. doi:10.1016/J.AB.2008.03.009, Academic Press.

Lin, Y., et al., 2014. Improvement of a sample preparation method assisted by sodium deoxycholate for mass-spectrometry-based shotgun membrane proteomics†. J. Sep. Sci. 37 (22), 3321–3329. doi:10.1002/JSSC.201400569, John Wiley & Sons, Ltd.

Lipinska, A.P., et al., 2020. To gel or not to gel: differential expression of carrageenan-related genes between the gametophyte and tetasporophyte life cycle stages of the red alga chondrus crispus. Sci. Rep. 10 (1), 1–14. doi:10.1038/s41598-020-67728-6, Nature Publishing Group.

Mæhre, H.K., et al., 2014. Characterization of protein, lipid and mineral contents in common Norwegian seaweeds and evaluation of their potential as food and feed. J. Sci. Food Agric. 94 (15), 3281–3290. doi:10.1002/JSFA.6681, John Wiley & Sons, Ltd.

Malcolm, A.S., Dexter, A.F., Middelberg, A.P.J., 2007. Peptide surfactants (Pepfactants)

for switchable foams and emulsions. Asia-Pacific J. Chem. Eng. 2 (5), 362–367. doi:10.1002/apj.66.
Mariotti, F., Tomé, D., Mirand, P.P., 2008. Converting nitrogen into protein—beyond 6.25 and jones' factors. Crit. Rev. Food Sci. Nutr. 48 (2), 177–184. doi:10.1080/10408390701279749, Taylor & Francis Group.
Marzano, V., et al., 2020. Perusal of food allergens analysis by mass spectrometry-based proteomics. J. Proteom. 215, 103636. doi:10.1016/J.JPROT.2020.103636, Elsevier.
Mathieson, A.C., Burns, R.L., 1971. Ecological studies of economic red algae. I. photosynthesis and respiration of chondrus crispus stackhouse and gigartina stellata (Stackhouse) batters. J. Exp. Mar. Biol. Ecol. 7 (2), 197–206. doi:10.1016/0022-0981(71)90031-1, Elsevier.
Minkiewicz, P., Iwaniak, A., Darewicz, M., 2019. BIOPEP-UWM database of bioactive peptides: current opportunities. Int. J. Mol. Sci. 20 (23). doi:10.3390/ijms20235978, NLM (Medline).
Mokni Ghribi, A., et al., 2015. Effects of enzymatic hydrolysis on conformational and functional properties of chickpea protein isolate. Food Chem. 187, 322–330. doi:10.1016/J.FOODCHEM.2015.04.109, Elsevier.
Mooney, C., et al., 2012. Towards the improved discovery and design of functional peptides: common features of diverse classes permit generalized prediction of bioactivity. PLoS One 7 (10), e45012. doi:10.1371/journal.pone.0045012, Edited by L. Kurgan. Public Library of Science.
Mora, L., Gallego, M., Toldrá, F., 2018. New approaches based on comparative proteomics for the assessment of food quality'. Curr. Opin. Food Sci. 22–27. doi:10.1016/j.cofs.2018.01.005, Elsevier Ltd.
Naseri, A., Holdt, S.L., Jacobsen, C., 2019. Biochemical and nutritional composition of industrial red seaweed used in carrageenan production. J. Aquat. Food Prod. Technol. 28 (9), 967–973. doi:10.1080/10498850.2019.1664693.
Noor, Z., et al., 2021. Mass spectrometry–based protein identification in proteomics—a review. Brief. Bioinform. 22 (2), 1620–1638. doi:10.1093/BIB/BBZ163, Oxford Academic.
Olsen, T.H., et al., 2020. AnOxPePred: using deep learning for the prediction of antioxidative properties of peptides. Sci. Rep. 10 (1), 21471. doi:10.1038/s41598-020-78319-w, Nature Research.
Perez-Riverol, Y., et al., 2019. The pride database and related tools and resources in 2019: improving support for quantification data. Nucleic. Acids. Res. 47 (D1), D442–D450. doi:10.1093/NAR/GKY1106, Oxford Academic.
Picot-Allain, C., et al., 2021. Conventional versus green extraction techniques — a comparative perspective. Curr. Opin. Food Sci. 40, 144–156. doi:10.1016/J.COFS.2021.02.009, Elsevier.
Pina, A.L., et al., 2014. An evaluation of edible red seaweed (Chondrus crispus) components and their modification during the cooking process. LWT - Food Sci. Technol. 56 (1), 175–180. doi:10.1016/J.LWT.2013.08.006, Academic Press.
Raja, K., Kadirvel, V., Subramaniyan, T., 2022. Seaweeds, an aquatic plant-based protein for sustainable nutrition - a review. Fut. Foods 5, 100142. doi:10.1016/J.FUFO.2022.100142, Elsevier.
Rawiwan, P., et al., 2022. Red seaweed: a promising alternative protein source for global food sustainability. Trends Food Sci. Technol. 123, 37–56. doi:10.1016/J.TIFS.2022.03.003, Elsevier.
Reuter, W., Nickel, C., Wehrmeyer, W., 1990. Isolation of allophycocyanin b from rhodella violacea results in a model of the core from hemidiscoidal phycobilisomes of rhodophyceae. FEBS Lett. 273 (1–2), 155–158. doi:10.1016/0014-5793(90)81073-W, John Wiley & Sons, Ltd.
Ritter, S., et al., 1999. Crystal structure of a phycourobilin-containing phycoerythrin at 1.90-Å resolution. J. Struct. Biol. 126 (2), 86–97. doi:10.1006/JSBI.1999.4106, Academic Press.
Rutherfurd, S.M., Gilani, G.S., 2009. Amino acid analysis. Curr. Protocols Protein Sci. 58 (1). doi:10.1002/0471140864.PS1109S58, John Wiley & Sons, Ltd11.9.1-11.9.37.
Sari, Y.W., et al., 2015. Towards plant protein refinery: review on protein extraction using alkali and potential enzymatic assistance. Biotechnol. J. 10 (8), 1138–1157. doi:10.1002/BIOT.201400569, John Wiley & Sons, Ltd.
Schultz, D.J., et al., 1994. RNA isolation from recalcitrant plant tissue. Plant Mol. Biol. Rep. 12 (4), 310–316. doi:10.1007/BF02669273, Springer.
Schwanhüusser, B., et al., 2011. Global quantification of mammalian gene expression control. Nature 473 (7347), 337–342. doi:10.1038/nature10098, Nature Publishing Group.
Singh, Shivdayal, et al., 2022. Antiviral applications of macroalgae. Sustain. Global Res. Seaweeds 2, 455–471. doi:10.1007/978-3-030-92174-3_25, Cham: Springer, Cham.
Stadnichuk, I.N., Tropin, I.V., 2017. Phycobiliproteins: structure, functions and biotechnological applications. Appl. Biochem. Microbiol. 53 (1), 1–10. doi:10.1134/S0003683817010185, Springer.
Sudhakar, K., et al., 2018. An overview of marine macroalgae as bioresource. Renew. Sustain. Energy Rev. 91, 165–179. doi:10.1016/J.RSER.2018.03.100, Pergamon.
Sun, X.D., 2011. Enzymatic hydrolysis of soy proteins and the hydrolysates utilisation. Int. J. Food Sci. Technol. 46 (12), 2447–2459. doi:10.1111/J.1365-2621.2011.02785.X, John Wiley & Sons, Ltd.
Tan, Y., et al., 2022. Comparison of emulsifying properties of plant and animal proteins in oil-in-water emulsions: whey, soy, and rubisco proteins. Food Biophys. 17, 409–421. doi:10.1007/S11483-022-09730-1, Springer.
Tandeau De Marsac, N., et al., 2003. Phycobiliproteins and phycobilisomes: the early observations. Photosynth. Res. 76 (1), 193–205. doi:10.1023/A:1024954911473, Springer.
Tanna, B., Mishra, A., 2018. Metabolites unravel nutraceutical potential of edible seaweeds: an emerging source of functional food. Compr. Rev. Food Sci. Food Saf. 17 (6), 1613–1624. doi:10.1111/1541-4337.12396, John Wiley & Sons, Ltd.
Tuerkova, A., et al., 2020. Effect of helical kink in antimicrobial peptides on membrane pore formation. eLife 9. doi:10.7554/eLife.47946, eLife Sciences Publications Ltd.
Tyanova, S., Cox, J., 2018. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. Humana Press, New York, NY, pp. 133–148. doi:10.1007/978-1-4939-7493-1_7.
UniProt Consortium, T., et al., 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic. Acids. Res. 49 (D1), D480–D489. doi:10.1093/NAR/GKAA1100, Oxford Academic.
van Dijk, M., et al., 2021. A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. Nat. Food 2 (7), 494–501. doi:10.1038/s43016-021-00322-9, 2021 2:7. Nature Publishing Group.
Veide Vilg, J., Undeland, I., 2017. pH-driven solubilization and isoelectric precipitation of proteins from the brown seaweed saccharina latissima—effects of osmotic shock, water volume and temperature. J. Appl. Phycol. 29 (1), 585–593. doi:10.1007/s10811-016-0957-6, Springer Netherlands.
Waterhouse, A., et al., 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic. Acids. Res. 46 (W1), W296–W303. doi:10.1093/nar/gky427, Narnia.
WHO/FAO/UNU, 2007. Protein And Amino Acid Requirements In Human Nutrition. Geneva: World Health Organization Available at:.
Xiao, N., et al., 2015. 'protr/ProtrWeb: r package and web server for generating various numerical representation schemes of protein sequences. Bioinformatics 31 (11), 1857–1859. doi:10.1093/BIOINFORMATICS/BTV042, Oxford Academic.
Xu, X., et al., 2021. Global greenhouse gas emissions from animal-based foods are twice those of plant-based foods. Nat. Food 2 (9), 724–732. doi:10.1038/s43016-021-00358-x, Nature Publishing Group.
Yesiltas, B., et al., 2021. Emulsifier peptides derived from seaweed, methanotrophic bacteria, and potato proteins identified by quantitative proteomics and bioinformatics. Food Chem. 362, 130217. doi:10.1016/J.FOODCHEM.2021.130217, Elsevier.
Yesiltas, B., et al., 2022. Antioxidant peptides derived from potato, seaweed, microbial and spinach proteins: oxidative stability of 5% fish oil-in-water emulsions. Food Chem. 385, 132699. doi:10.1016/J.FOODCHEM.2022.132699, Elsevier.
Zhang, S., et al., 2017. Novel metabolic and physiological functions of branched chain amino acids: a review. J. Animal Sci. Biotechnol. 1–12. doi:10.1186/s40104-016-0139-z, BioMed Central Ltd..
Zhou, J., et al., 2006. Evaluation of the application of sodium deoxycholate to proteomic analysis of rat hippocampal plasma membrane. J. Prot. Res. 5 (10), 2547–2553. doi:10.1021/PR060112A, American Chemical Society.
Zhou, N.A., et al., 2015. Identification of putative genes involved in bisphenol a degradation using differential protein abundance analysis of sphingobium sp. bid32. Environ. Sci. Technol. 49 (20), 12231–12241. doi:10.1021/acs.est.5b02987, American Chemical Society.