MDPI

Review

Key Proteomics Tools for Fundamental and Applied Microalgal Research

Maxence Plouviez [1,2,*] and **Eric Dubreucq [3]**

1   School of Agriculture and Environment, Massey University, Palmerston North 4410, New Zealand
2   The Cawthron Institute, Nelson 7010, New Zealand
3   Agropolymer Engineering and Emerging Technologies, L'Institut Agro Montpellier, 34060 Montpellier, France; eric.dubreucq@supagro.fr
*   Correspondence: maxence.plouviez@cawthron.org.nz

Abstract: Microscopic, photosynthetic prokaryotes and eukaryotes, collectively referred to as microalgae, are widely studied to improve our understanding of key metabolic pathways (e.g., photosynthesis) and for the development of biotechnological applications. Omics technologies, which are now common tools in biological research, have been shown to be critical in microalgal research. In the past decade, significant technological advancements have allowed omics technologies to become more affordable and efficient, with huge datasets being generated. In particular, where studies focused on a single or few proteins decades ago, it is now possible to study the whole proteome of a microalgae. The development of mass spectrometry-based methods has provided this leap forward with the high-throughput identification and quantification of proteins. This review specifically provides an overview of the use of proteomics in fundamental (e.g., photosynthesis) and applied (e.g., lipid production for biofuel) microalgal research, and presents future research directions in this field.

Keywords: microalgae; biochemical pathways; biotechnology; proteomics; high-throughput; protein modelling

1. Introduction

Photosynthetic microorganisms can be found in all aquatic ecosystems, in which, as primary producers, they have important ecological functions, either positive (e.g., biogeo-chemical cycle [1–3]), or negative (e.g., when they bloom at the wrong time and place [4,5]). Although they encompass both prokaryotes (cyanobacteria) and eukaryotes ("true" microalgae), they will be collectively referred to as microalgae in the present review [6]. Microalgae such as the model *Chlamydomonas reinhardtii* are widely studied to improve our understanding of key plant functions such as chloroplast-based photosynthesis and nutrient assimilation, and the structure and function of the eukaryotic flagella [7–10]. In addition, due to their physiology and biochemistry, significant research is focusing on the development of microalgae-based environmental biotechnologies for food and high-value molecules production, biofuels generation, wastewater treatment and nutrient recovery [11–14].

In recent years, the so-called "omics" technologies have been shown to be valuable tools in microalgal research [15–18]. Omics, which includes genomics, transcriptomics, proteomics and metabolomics, broadly refers to the comprehensive analyses of classes of biological molecules, i.e., DNA, RNA, proteins and metabolites, and their interactions (interactomics, Figure 1). Omics allow us to determine and study the whole makeup of a cell/biological system at a given time and, therefore, to correlate molecular signatures with phenotypes [17,19–23]. With a wide range of applications, these technologies have considerably accelerated the rate of discoveries and provided a leap forward in fundamental and applied research. For example, the sequencing and study of the *C. reinhardtii* genome in 2007 unraveled the evolution of the eukaryotic flagellum and plastid [7]. In the past decade,

technological developments in hardware and software have allowed omics technologies to become more affordable and more accessible [22,24,25]. Consequently, it is now possible to study the whole genome, transcriptome, proteome and metabolome of an organism in a matter of weeks (assuming the proper steps are considered and followed). As clear evidence, whereas until 2008 the genome of only three microalgae had been sequenced [16], many omics datasets have now been deposited in public repositories.
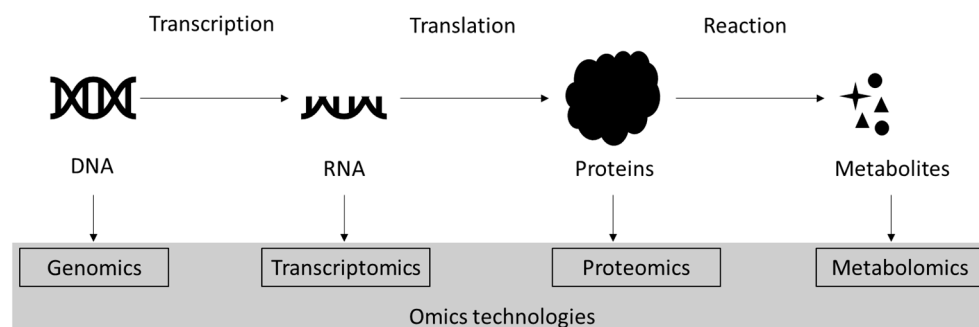

Figure 1. The four main omics. While not represented on the figure, the study of the interactions between and among proteins and other molecules within a cell is referred as interactomics.

While genomics and transcriptomics deliver significant information about genes and their expressions, these approaches do not provide an indication of protein levels, protein turnover and post-translational modifications. Consequently, proteomics, i.e., the study of proteins (from their structure to their interactions with other molecules in a cell), corroborates and builds on genomics and transcriptomics and it is, therefore, a comprehensive approach to characterize a biological system. As proteins are effectors of biological functions, the levels (and forms) of proteins in a cell indeed represent comprehensive information about cellular function, defining the phenotype of a cell in response to genetic or environmental changes [13,19,25].

This review aims to give an overview of proteomics technologies and their applications for microalgal research and focuses on the recent improvements in the use of proteomics in microalgal research. It also presents the future of proteomics in the post-omics era, i.e., the development and use of artificial intelligence (AI)-based technologies. Recent advancements in machine learning, particularly deep learning, have enabled researchers to accurately predict protein structures and functions from their amino acid sequences, improving the quality and reliability of analytical workflows in mass spectrometry-based proteomics [26]. This approach is becoming central to biomarker discovery from proteomics data, and is beginning to outperform existing assays [27]. Furthermore, AI algorithms have been instrumental in analyzing large-scale proteomic datasets [27–30], ultimately enabling the identification of novel protein interactions and pathways that are critical for microalgal adaptation to environmental stresses [16,31]. It is noteworthy that the potential of microalgal biotechnology and that of proteomics technologies, along with their respective advantages and disadvantages, have been reviewed elsewhere [11,16,19,21,32,33]. Therefore, these topics are not the primary focus of this review.

2. Proteomics

According to the central dogma of molecular biology, the genetic information encoded in the nucleotide sequence of DNA is transcribed into messenger RNA (mRNA). This mRNA is then translated into a polypeptide chain of amino acids. Subsequently, the polypeptide undergoes folding, often facilitated by chaperone proteins, as well as post-translational modifications. These processes, along with potential interactions with other proteins or molecules, lead to the formation of a functional protein [19–21,25,33]. The proteome is defined as the whole set of proteins of a biological system, and proteomics refers to methodologies to identify, characterize, structurally study and/or quantify proteins

(further details in Box 1, and see examples of methods used in microalgal research in the next section).

Methodological advancements, including the development of mass spectrometry techniques for precise mass and chemical structure analysis, have revolutionized protein research. In conjunction with growing peptide and protein sequence databases as well as advanced bioinformatics tools for handling and analyzing large datasets, they have enabled rapid and accurate large-scale analyses of proteins [19,21,25,31], a significant shift from the study of single or a few proteins that characterized the field decades ago.

Proteins are complex molecules that serve structural roles, as well as drive and regulate metabolic reactions within cells. Studying the proteome is therefore ideal to characterize a biological system at a particular time or phase [19,21,33]. In addition, as proteins perform a function and are encoded in the genome, proteomics is considered one of the most significant methodologies to characterize and understand gene functions [19]. However, proteomics can be challenging, as proteomes are significantly more complex than genomes because of the vast range of proteins found at different abundances and in different forms (i.e., iso/proteoforms) according to post-translational modifications [19,21,23,34]. This diversity also implies a tremendous number of possible interactions between proteins and between proteins and other molecules. These interactions are the subject of the more recent field of interactomics, which attempts to assess as comprehensively as possible this network of interactions [2,20,35,36] (Figure 1).

The variations in protein turnover time and the influence of post-transcriptional and post-translational regulations can also contribute to the complexity of proteomes. This can explain discrepancies between the transcriptome (i.e., the abundance of transcripts) and the proteome (i.e., the abundance of proteins), as reported for the microalgae *Thalassiosira pseudonana* [37] or *Chlamydomonas reinhardtii* [38]. Indeed, moderate correlations of ~0.4 have been widely reported between the abundance of proteins and their corresponding mRNAs in both prokaryotes and eukaryotes [34,39]. Conversely, Plouviez et al. [40] recently showed, by temporal analysis of both datasets, that the transcriptome and the proteome related to phosphorus metabolism are synchronized in *Chlamydomonas*. This provided insight into the turnover and regulation of Vacuolar Transport Chaperone proteins, key proteins involved in polyphosphate synthesis in microalgae [40–42].

The knowledge provided by proteomic research adds greatly to the genetic information accumulated from genomics and transcriptomic studies, as the field of proteomics is broad and encompasses many different interrelated areas of study, including expression proteomics, functional proteomics and structural proteomics (Box 1).

Box 1. Main field of proteomics.

> *Expression proteomics*: Refers to the measurement and comparison of protein abundance in the entire proteome (or subproteome) between samples that differ by some parameters. For example, the proteome of cells cultivated under different conditions can be compared, as was done by Plouviez et al. [40] with *C. reinhardtii* cultures grown under phosphorus-depleted and repleted conditions. The proteome of two different phenotypes can also be compared, as was done by Sithtisarn et al. [43] with a salt-tolerant mutant and "wildtype" *Chlamydomonas* strains. Expression proteomics complements and builds on proteome identification and mining by identifying proteins that are normally expressed in a cell or organelle at a specific time or phase [44], as in the case of the eyespot of *C. reinhardtii* [45].
>
> *Functional proteomics*: A broad term for many specific, directed proteomics approaches to identify the biological functions of specific proteins or classes of proteins. Due to their key role in functional proteomics, significant research is focused on post-translational modifications (PTMs) of proteins, such as oxidation, phosphorylation or acetylation [20,33,46–49]. Indeed, these modifications regulate protein activity, stability, localization and interactions within the cell [20,33,46]. Specific proteomic studies focused on PTMs in *C. reinhardtii* have revealed the complexity of the redox protein regulation network in this microalga and the importance of PTMs in protein regulation and signal transduction during stress responses [46,47,50–53].
>
> *Structural proteomics*: This area of proteomics focuses on characterizing the structure of proteins or protein complexes. The information gathered helps determine how a protein works and how it is regulated, as was done by Gurrieri et al. [54] for *C. reinhardtii* phosphoribulokinase or by Shen et al. [55] for the PSII-LHCII supercomplex involved in light-induced oxidation of water during photosynthesis. Ultimately, structural studies can provide critical information on protein function. Due to ongoing technological advancements, such as mass spectrometry (MS) analyses, protein–protein interaction network analyses, and proteome-wide imaging of protein localization [31,56,57], it is now possible to identify and locate all proteins within a protein complex or organelle [44,57–59]. Additionally, these technologies enable the characterization of protein–protein interactions (how, which, when and where proteins interact). This field, known as spatial proteomics, allows us to study the overall architecture of organelles and elucidates the roles that proteins play in cellular pathways and organellar dynamics [56,60].

3. Overview of Proteomics Techniques Applied in Microalgae Research

Proteomics can be hypothesis-driven or hypothesis-generating depending on if a proteomic experiment targets certain known proteins of interest (direct approach), or instead measures all the proteins present in a sample (defined as the indirect or shotgun approach) [61].

Table 1 presents the main methodologies used to study individuals, groups or the whole set of proteins in a sample. The choice of methods is highly dependent on the aim of the study, and often multiple methods are used in parallel, such as in the study reported by Stauber et al. [62]], which used 2D gel electrophoresis to separate *C. reinhardtii* light-harvesting proteins and identify these proteins using mass spectrometry. Data outputs can either be qualitative (i.e., is a protein present) or quantitative (i.e., is a protein found in higher abundance in one particular sample). Notably, there has been significant development in, and interest toward, quantitative high-throughput methodologies in recent years [18,20,21]. As a result, conventional methods are nowadays primarily used for specific applications (i.e., protein separation) or for data validation (Figure 2).

Table 1. Overview of the methodologies available to study proteins/proteome with references to their use in *C. reinhardtii*.

Method		Description	Example(s)
Conventional	Chromatography	Chromatography-based techniques are used for protein separation and purification. In brief, proteins can be separated/purified based on charges (IEC: ion-exchange chromatography), size (SEC: size-exclusion chromatography) or (bio)chemical affinity with a matrix (AC: affinity chromatography).	[63–65]
	Immuno assay/blotting	The enzyme-linked immunosorbent assay (ELISA) detects the presence of a protein by measuring the enzymatic activity of an enzyme-labeled antigen or antibody binding to an immobilized target protein (or antibody). ELISA is widely used for diagnostics. Western blotting allows the separation of proteins based on molecular weight through gel electrophoresis and identification of protein from binding of a labeled antibody to its target antigen (i.e., protein) on a membrane.	[66,67]
	Edman sequencing	The Edman method (or Edman degradation) determines the amino acid sequence in peptides/proteins by sequentially identifying and cleaving amino acids from the N-terminal side of a peptide/protein.	[68]
Advanced	Gel-based	The conventional polyacrylamide gel-based method used for protein separation and identification based on proteins' mass (SDS-PAGE: sodium dodecyl sulfate–polyacrylamide gel electrophoresis) has evolved into more advanced 2D methods based on both mass and charge separation (2D-PAGE: two-dimensional polyacrylamide gel electrophoresis) or using labels with a fluorescent dye (2D-DIGE: two-dimensional differential gel electrophoresis). 2D-PAGE and 2D-DIGE can resolve and investigate the abundance of several thousand proteins in a single sample.	[59,62,69]
	Mass spectrometry	MS is one of the most used analytical techniques to identify and, coupled with (ultra)-high performance chromatography, quantitatively measure protein levels. MS of peptides/proteins ionized via matrix-assisted laser desorption ionization (MALDI), surface-enhanced laser desorption/ionization (SELDI) or, more classically, electrospray ionization (ESI), allow the determination, through deconvolution of the mass spectra obtained, of their molecular mass. In the context of proteomics and peptide analysis, multi-stage tandem mass spectrometry (MSn) is also a powerful technique for obtaining detailed information on the structure and sequence of peptides and proteins, particularly with regard to the localization of post-translational modifications (PTMs). Label-free or labeled approaches can be used for quantification (ICAT: isotope-coded affinity tag; iTRAQ: isobaric tagging for relative and absolute quantification; SILAC: stable isotope labeling by/with amino acids in cell culture; QconCAT: quantification concatemer; TMT: Tandem Mass Tag). MS is often combined with separations and fractionation techniques to identify target proteins or subproteomes. Sample fractionation and enrichment are also important when identifying PTMs (phosphorylation, oxidation, nitrosylation, glycosylation, methylation, etc.).	ICAT: [46] iTRAQ and label-free: [70] iTRAQ: [71] QconCAT: [72] SILAC: [50] TMT: [40] PTMs: [46,47,51,53,73]

Table 1. *Cont.*

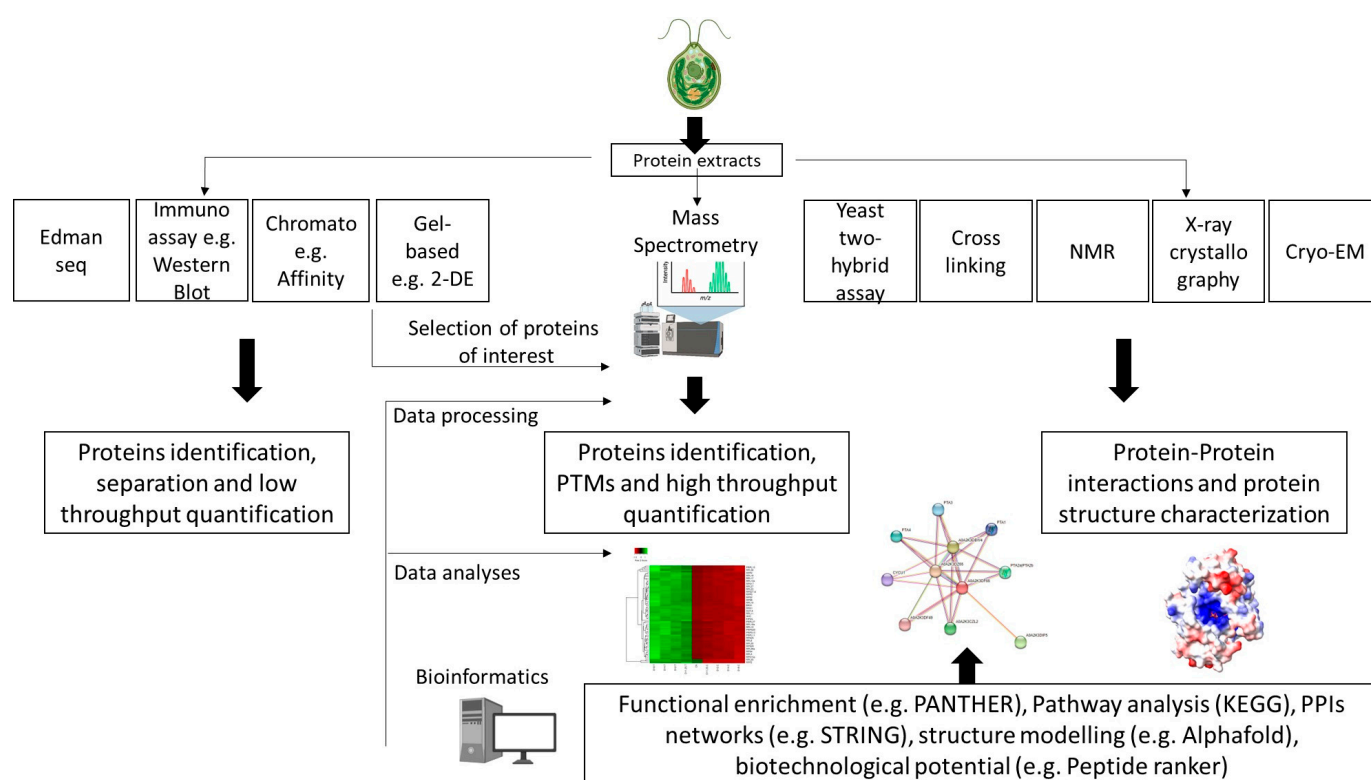|	Method	Description	Example(s)
Advanced	Nuclear Magnetic Resonance	Protein structural determination in solution or solid phase by measuring chemical shifts. NMR structural determination involved several steps, each requiring significant expertise/techniques.	[74]
	X-ray crystallography	Three-dimensional protein structures are determined by exposing highly purified crystallized protein samples to X-rays and measuring diffraction patterns.	[75]
	Cryogenic electron microscopy	Microscopy techniques used to determine 3D structure of proteins or protein complexes via flash freezing and electron bombardment of samples in solution.	[55]
	In silico modelling	Generation and study of protein 3D models using homology modelling or deep-learning-based predictions.	[41]

Figure 2. Proteomics techniques and tools commonly used for microalgae research (seq: sequencing; Immuno: immunodetection; Chromato: chromatography; PTMs: post-translational modifications).

3.1. Quantitative High-Throughput Proteomics: Mass Spectrometry

Mass spectrometry is now commonly used to identify and quantitatively measure protein levels in microalgae samples. MS quantitative proteomics proved useful in fundamental and applied microalgae research (Section 5) to compare the physiological state of different microalgae strains or microalgae grown under different conditions. While top-down (i.e., intact proteins, [76]) and bottom-up (i.e., digested proteins, [77]) analyses are alternative strategies for protein identification and characterization by mass spectrometry, bottom-up approaches have mostly been used for microalgal proteomics.

Label-based and label-free strategies can be used for protein quantification. Label-based techniques (such as ICAT, iTRAQ, SILAC, QconCAT, TMT; Table 1) incorporate stable isotope labels into peptides, creating an anticipated change in mass that can be measured. Conversely, label-free proteomics quantify peptide abundance based on signal

intensity or spectral counting of peptides [78]. Wang et al. [70], previously compared the use of a label-based method (iTRAQ) and label-free approach on *C. reinhardtii*. The authors found a good overlap of proteins identified with both methods, suggesting both provide high-quality quantitative and qualitative data. While quantification accuracy and reproducibility were better when using iTRAQ than the label-free method, more proteins were identified and quantified with the label-free method. Therefore, the authors suggested that extra steps (e.g., pre-fractionation) might have been needed for a higher coverage when using the iTRAQ method, and they recommended considering the experimental design (i.e., sample numbers, sample complexity, and amount) for selecting a suitable proteomic approach. Notably, label-free methods have been the most popular approaches in microalgal proteomics, potentially because label-free methods are cost-effective and do not require expensive labeling reagents. Independently of the method selected, sample preparation is critical for proteomics analyses, as further described in Section 4.

3.2. Protein–Protein Interaction Techniques

Proteins generally drive cellular processes by interacting with other molecules (e.g., metabolites), including other proteins. Protein–protein interactions (PPIs), defined as the physical contacts between two or more proteins, are therefore critical as PPIs can activate or deactivate a specific protein, alter kinetic properties and regulate biological processes. Another key benefit of studying PPIs (or PPI networks) is that it is possible to determine putative functions of uncharacterized proteins. Several methods exist to study PPIs either in vitro or in vivo [79,80]. While X-ray crystallography and NMR are two methods commonly used for in vitro study (see references in Table 1), the yeast two-hybrid assay is widely used in vivo. With this method, the detection of interacting proteins is detected in yeast cells when fused "bait and prey" proteins from a target organism activate reporter genes that enable growth on specific media or a specific reaction leading to color change. This method was successfully applied to characterize the Intraflagellar Transport Complex A Proteins subunit interactions in *C. reinhardtii* [81] or the interaction between the small subunit of Rubisco and the essential pyrenoid component 1, a linker protein required for Rubisco aggregation in algae [82]. The latter improved the understanding of the specific algal CO_2-concentrating mechanism (CCM). Alternatively, Mackinder et al. [83] used affinity purification–mass spectrometry to measure the interaction of proteins involved in the CCM and the pyrenoid in *C. reinhardtii*. Finally, PPIs (and PPI networks) based on known and predicted interactions can also be studied in silico (further discussed in Section 6).

Critically, studying PPIs could prove instrumental for strain improvement. For example, using in silico protein modelling and activity-based chemical crosslinking, Blatti et al. [84] evidenced the critical role of the interaction between a fatty acid acyl carrier protein and thioesterase in governing fatty acid metabolism in *C. reinhardtii*, providing an alternative to manipulating fatty acid biosynthesis for algae biofuel. Considering the tremendous potential of microalgae, further research on PPIs of protein involved in metabolic pathways of use for biotechnology would be valuable.

4. Key Considerations for Successful Proteomics

As seen in Table 1, a wide range of methods are available to study proteins/proteomes. While each of these methods has its own advantages and limitations [19,33], their success is strongly dependent on protein sample preparation [33,85–87]. In addition, especially for high-throughput techniques, huge datasets are being generated and, therefore, require the use of bioinformatics tools and pipelines to acquire and analyze data [24,31,88–90]. The following sections, therefore, focus on the importance of sample preparation and bioinformatics in proteomics.

4.1. Protein Sample Preparation

Sample preparation generally involves multiple steps (e.g., cell disruption, extraction, fractionation, chemical modifications and purification, e.g., [86,87,91]) that are susceptible to losses and contaminations [77,91–93]. Considering the complexity of proteins, their heterogeneity and the cell characteristics, sample preparation can therefore be a time-consuming, limiting step [13,31]. Method optimization is generally required and adds to the labor of sample preparation. Critically, the steps (and options) involved in sample preparation need to be carefully considered to generate quality protein extracts and to prevent failure of downstream analyses. The two sub-sections below focus on protein extraction methods that can be a limiting step in microalgal proteomics and, considering the importance of MS-based techniques in the field, protein sample preparation for high-throughput proteomics.

4.1.1. Protein Extraction

Extraction efficiency varies greatly with the microalgal species considered, due to biochemical and structural differences in membranes [94]. Significant research has been conducted to develop efficient and reproducible sample preparation methods, especially for proteomic analyses of tissues and cells of animals and higher plants [31,86,95,96]. While these protocols have not necessarily been attempted on microalgal biomass, similar cell disruption techniques (i.e., sonication, beading, lysis buffer) and solvents (e.g., trichloroacetic acid, chloroform, methanol) are used for cell lysis and protein solubilization/precipitation, respectively. While proteins can be extracted from wet or dry biomass, lyophilization is generally recommended for sample stability. Anjos et al. [91] recently described a protocol for protein extraction suitable for analyzing the proteome of *Tetraselmis chuii*. However, according to the authors, this protocol effectively generated protein extracts with high yields for larger proteins but was not efficient for proteins with low molecular mass. Importantly, the authors also noted that protein extraction techniques must be specifically selected for each microalgae species. This was also evidenced by Toyoshima et al. [87], who showed that, while sonication in KCl buffer was the most effective method to extract total proteins from *C. reinhardtii*, beads disruption in PTS buffer was better for *Synechocystis*. However, with *Chlamydomonas*, a better detection of individual subunits of the photosynthetic apparatus and other large protein complexes was obtained by the PTS and sonication methods rather than when using the KCl method. Together with the results from Anjos et al. [91], this suggests that, in addition to the species studied, the protocol used for sample preparation must also be selected according to the downstream analyses to be performed. Because of the inherent limitation due to the multiple steps involved for proteomics analyses, as well as the known variability between biological samples, it is highly recommended to conduct preliminary testing and use more than three biological replicates. While there is currently no "standard" protocol for proteomics sample preparation in microalgae, the authors recommend the following literature focusing on microalgal sample preparation: [48,49,85,87,91,93,97].

4.1.2. Sample Preparation for MS-Based Analysis

Bottom-up approaches have mostly been used for microalgal proteomics. For these approaches, extracted proteins are first reduced and alkylated before being digested using proteases to generate peptides that will ultimately be analyzed using MS techniques. While different proteases can be used to broaden proteome coverage, trypsin is the most widely used in proteomics, including in microalgal proteomics [13]. While several proteolytic digestion methods are in development, in-solution digestion is mostly used for microalgae [13]. During in-solution digestion, proteins are digested while remaining in a buffer. Peptide purification/desalting is then applied to remove excess salts and detergents that could interfere with the MS analysis (e.g., [40]). Alternatively to in-solution digestion, in-gel digestion can be performed. In this approach, proteins are separated with gel electrophoresis, and bands or groups of bands of interest are excised and then digested [62].

This allows the analysis of specific groups of proteins, simplify sample complexity and increase the depth of the analysis. However, this process can be time-consuming and prone to error, possibly explaining why in-solution digestion is preferred. Technical progress toward shotgun proteomic analyses, where mixtures of proteins, digested by multiple proteases of different specificities, are simultaneously analyzed by high-resolution MS/MS, has nevertheless modified these practices [98], and the current iterations of this approach have led to the capacity for analyzing and quantitating thousands of proteins per hour from whole-proteome digestions [99,100]. Critically, further investigation of the efficiency and suitability of novel digestion methods and proteases is needed for microalgal proteomics.

4.2. Bioinformatics

Bioinformatics encompasses algorithms and software developed to analyze and interpret biological data. For proteomics, bioinformatics can be used for protein–protein search (i.e., classical BLAST), for in silico analyses (e.g., structure modeling and protein-protein interactions as further discussed in Section 6), and to analyze data from, for example, 2D gel electrophoresis [101]. High-throughput proteomics also critically relies on bioinformatics to acquire and analyze the large datasets generated by mass spectrometry [18,23,61,88,90].

4.2.1. MS Data Processing

While bioinformatics is critical for both top-down and bottom-up approaches, this review focuses on the bottom-up approach that is commonly used in microalgal proteomics (see Section 3.1). Briefly, a standard bottom-up MS proteomics workflow involves several steps [61,77]:

1. Protein samples are prepared for MS, i.e., solubilized proteins are enzymatically digested (usually using trypsin), chemically modified (alkylation) and purified (desalting) to obtain short MS-accessible peptides;
2. Peptides are separated using an LC setup that is coupled to a mass spectrometer (LC-MS);
3. Intact peptide masses and the corresponding masses of fragmented peptide ions are measured by mass spectrometry (i.e., tandem LC-MS/MS or MSn setups). As mentioned in Section 3.1, label-based and label-free strategies can be used for protein quantification, and different techniques are available (Table 1). Notably, different data acquisition modes can also be used (e.g., label-free data-dependent acquisition and data-independent acquisition); refer to Schessner et al. [90] for further details;
4. Based on a reference proteome, the resulting peptide and fragment ion spectra are used to identify the peptides present in the sample;
5. Identified peptide sequences are quantified and assembled to measure protein levels by protein inference. Data from MS analyses are susceptible to systematic, dependent or independent biases (e.g., different handling, equipment calibration) on the measured peptide/protein abundances [102]. Therefore, a key step is to normalize the data to take the bias into account, allowing the data to be comparable and downstream analyses reliable [103–105]. Advanced analysis pipeline frameworks are therefore needed for data normalization but also for protein inference and data analysis [88,89,103,104,106]. The latter has been extensively reviewed by Schessner et al. [90] in their guide to interpreting and generate visual representation of bottom-up proteomics data.

4.2.2. MS Data Interpretation

Interpreting the large amount of data available from high-throughput datasets such as those for quantitative proteomics can be overwhelming. Based on MS data, thousands of peptides and proteins can be identified and quantified. This allows for the determination of absolute levels within a sample and relative levels across different samples [90]. For the latter, statistical analyses (i.e., Student's t-test) and fold change calculations (i.e., relative abundance) are used to identify proteins found in higher or lower abundance between

treatments and/or phenotypes. For instance, a *p*-value < 0.05 and a fold change of −1.20 to 1.20 could be criteria for significance [40]. Several methodologies can then be followed to gather knowledge from the list of identified proteins [90]. As hundreds of proteins can be statistically found at higher/lower abundance, a common approach is to perform functional enrichment analyses [90,107–109]. This bioinformatics method is based on a statistical comparison of a sample dataset against a reference dataset. This allows us to identify classes of proteins (or genes) that are over-represented in a large dataset and may have an association with different phenotypes and/or physiological states. For example, a functional enrichment analysis showed that ribosome synthesis and protein translation are a strong response to the transition from P-depleted to P-repleted conditions in *C. reinhardtii* [40].

During functional enrichment analyses, proteins (or genes) are annotated and classified from manually curated libraries on the basis of their function. Gene Ontology (GO) is the most renowned library that provides terms classified under biological processes, molecular functions and cellular components, providing critical information about the location and function of specific proteins [110]. Other libraries also provide alternative classification, such as the KEGG pathway database, allowing the understanding of metabolic pathways influenced by a specific condition [111]. Functional enrichment analyses can be performed using several online software tools (e.g., MapMan [112]; Panther: [40]; see [90,109] for further examples). Notably, because each software tool uses specific algorithms and statistical tests, performing enrichment analyses with several software applications is recommended for data cross-validation [90].

When groups of proteins have been identified in response to specific conditions, further studies such as protein–protein interactions network analyses can provide a critical understanding of how proteins work together and which proteins play key functions. For example, a PPI network analysis suggested that proteins involved in protein metabolism, energy supply, and photosynthesis act in synergy to reconstitute cellular homeostasis under salt stress in *Dunaliella salina* [113]. PPI networks can be generated from the STRING software v12 (STRING: functional protein association networks (https://string-db.org/; accessed on 1 April 2024) from its wide database of known and predicted protein–protein interactions.

The continuous effort and use of proteomics profit from the increased number of databases, the availability of data and software, and vice versa. Significant protein information is now collated in databases such as UniProt (https://www.uniprot.org/; accessed on the 1 April 2024), a freely accessible database of protein sequences and functional information. Specific to microalgal research, Predalgo is dedicated to the prediction of protein subcellular localization in green algae [83], and the Algal Protein Annotation Suite (Alga-PrAS: Alga-PrAS (riken.jp)) [114] provides protein databases to enable the interpretation of algal proteome features. The increasing availability of software and databases promises a wide range of discoveries for fundamental research but also for applied microalgal research. Based on these, network-based strategies, while primarily discussed in the context of metabolomics, have also been applied to proteomics [115,116]. These strategies encompass association networks based on quantitative information, mass spectra similarity networks to assist metabolite annotation, and biochemical networks for systematic data interpretation.

4.2.3. Functional Characterization

Bioinformatics tools have also proved useful to determine the biotechnological potential of proteins identified during high-throughput proteomics. Peptide Ranker, a server for the prediction of bioactive peptides (PeptideRanker (ucd.ie); accessed on 1 April 2024), was used to predict the bioactivity of 500 peptides in the microalga *Tetradesmus obliquus* [117]. Twenty-five peptides with potential antioxidant and angiotensin-converting–enzyme-inhibitory activities were found, allowing the authors to select four of these peptides for further in vitro testing. In another study, Carrasco-Renado et al. [118] identified 488 proteins with potential industrial applications from the proteome of *Nannochloropsis gaditana*

from a search of the EMBL's European Bioinformatics Institute patented proteins database NRPL2 (Non-redundant Patent Sequences < Patent Data Resources < EMBL-EBI, [32,118]).

5. The Benefit of Proteomics in Microalgal Research

As in other fields of research, proteomics has been instrumental to many discoveries in algal research [15–17]. Most of the proteomics methodologies currently available have been successfully applied to study proteins/proteomes in microalgae, such as *Chlamydomonas* (see references in Table 1). This is not surprising, as *C. reinhardtii* is a model organism not only in phycology but also in plant sciences for key metabolic pathways [8,62,75]. Proteomics approaches are also now widely applied to microalgae species with commercial potential (e.g., *Chlorella vulgaris*, *Nannochloropsis oculata*, *D. salina*), an important step for the development and improvement of microalgal biotechnology. The following sections provide some examples in fundamental and applied microalgae research.

5.1. Fundamental Research

C. reinhardtii is the first microalga to have been sequenced [7]. Consequently, the availability of extensive resources, such as genetic and proteomic databases, mutant collections and the development of protocols for genetic modifications, make it an organism of choice for fundamental research [8]. Significant research has especially focused on *C. reinhardtii* to study oxygenic eukaryotic photosynthesis [9,38,55,72,119]. Thus, many structural and mechanistic understandings about photosynthesis were discovered in *Chlamydomonas*. As the overall architecture of the photosynthetic core complexes are very similar between *C. reinhardtii* and vascular plants [62,120], findings obtained with *C. reinhardtii* have broad implications in plant biology.

Proteomics has been widely used to identify and study the proteins involved in photosynthesis in *C. reinhardtii*. For example, the first crystal structure of Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase, EC 4.1.1.39: the enzyme binding CO_2 during photosynthesis) in green algae was generated by X-ray crystallography in *Chlamydomonas* in 2001 [75]. Later, by combining 2D gel electrophoresis and mass spectrometry, Stauber et al. [62] generated a detailed map of *C. reinhardtii* light-harvesting proteins (Lhcas and Lhcbs proteins) and provided a first hint about structural differences in the light-harvesting complexes between green microalgae and vascular plants. The structures of light-harvesting complexes are now available. Particularly, the study of the *C. reinhardtii* PSII-LHCII supercomplex (the light-harvesting complex involved in light-induced oxidation of water during photosynthesis) from cryo-electron microscopy data unraveled protein–protein interactions and key features that explain how microalgae harvest energy efficiently at low light intensities [55]. In the last decade, significant research efforts have focused on characterizing the chloroplast proteome of *Chlamydomonas* [44,57]. A fluorescent protein-tagging and affinity purification–MS pipeline developed in *Chlamydomonas* allowed the identification and the analysis of the interactions of known and new proteins of the pyrenoid (a chloroplast microcompartment found in many algae) with those involved in the CCM, significantly improving our understanding of the algal CCM [83]. In addition, using MS, Zhan et al. [58] characterized 190 proteins, among which 81 were of unknown function, in the pyrenoid proteome of *Chlamydomonas*. Their functional analysis showed that, in addition to its known function to promote photosynthetic CO_2 fixation by Rubisco, the pyrenoid may be involved in photoacclimation and/or responses to light-induced stress.

Quantitative proteomics has been widely used to improve our understanding of photosynthesis by investigating *C. reinhardtii* wild types or mutants' responses to changing light [38,72,87,121]. It has also been widely applied to study other metabolisms than photosynthesis in *Chlamydomonas*, such as micronutrient deficiency [122], nutrient use [40,123], salt tolerance [50], etc. Similar approaches were used to study key metabolisms in other microalgae, such as the response of *D. salina* and *Chlorella sorokiniana* to toxic metals [124–126] and of *Thalassiosira pseudonana* to P depletion [37]. The latter provided key knowledge on

the strategies developed by diatoms to survive P depletion such as swapping phospholipids for sulfolipids to maintain intracellular P levels.

5.2. Applied Research

Microalgae farming currently supports a multi-billion USD industry, and significant research is focusing on the development of microalgal biotechnologies for wastewater treatment and the production of high-value molecules and biofuels [11–14]. Proteomics, and especially quantitative proteomics, could provide insight for cultivation or strain improvement for the above-mentioned biotechnological applications [13,16,18,127,128].

Quantitative proteomics has thus been widely applied to microalgae with commercial potential such as *Chlorella* [127,129], *Dunaliella* [113,130,131], *Haematococcus* [132], *Nannochloropsis* [32,133] and *Scenedesmus* [134], to name a few (see Table 1 from [16] and extensive supplementary data from [13]). However, even when considering the tremendous potential of microalgae for biotechnology, an extensive literature search showed that most of the proteomics studies concerned the effects of stresses, such as nitrogen depletion, to trigger lipid production for biofuels [16]. Generally, the majority of differentially abundant proteins are found to have functions in metabolic pathways related to fatty acid and lipid metabolism, carbohydrate metabolism, and photosynthesis [16,17,135]. In brief, during nutrient stress, the cells experience reduced photosynthetic efficiency and divert energy and carbon fluxes toward lipids rather than carbohydrate, protein and chlorophyll biosynthesis.

In addition to generating knowledge that improves the understanding of the metabolic pathways involved during specific growth conditions relevant to biotechnological applications, quantitative proteomics also allows for the identification or functional characterization of proteins of interest for biotechnology ([13] and see Section 6 below). The next logical step in microalgal biotechnology would, therefore, be to focus on the in silico and in vitro study of proteins with biotechnological potential. Structural and functional information about 31 algal proteomes can now be found in the Algal Protein Annotation Suite (Alga-PrAS) [114]. In addition, several software platforms are now available to determine the biotechnological potential of proteins or characterize specific features (e.g., bioactivity) of proteins with biotechnological potential (Section 4.2).

6. Proteomics in the Post-Omics Era

Artificial intelligence (AI) technology, especially based on deep learning, is changing our everyday life. It has been rapidly expanding during the last few years and has already had a strong impact on science [24,27,136]. As mentioned above, bioinformatics play a key part in understanding and unraveling the full potential of proteomics data. Combined with AI, bioinformatics is now on the edge of a huge acceleration and facilitation of data analyses, offering new opportunities and incomparable results in applied research [24,27,29]. The development of AI-based software for MS data analysis is one example. Applications in protein structure prediction is another. AI software such as AlphaFold v2.3.2 or RoseTTAFold v1.0 generate protein structure fast and accurately [137], preventing the need for the laborious, time-consuming and expensive experimental techniques for protein structure determination such as X-ray crystallography [136]. Currently, more than 1 million computed protein structures have been added to the Protein Databank (PDB, https://www.rcsb.org/; accessed on 1 April 2024), the main reference deposit of protein structures, and almost all known proteins have had their structure calculated [138]. Combined with molecular dynamics and ligand docking analyses, in silico protein studies have wide applications in many biological fields such as medicine and drug discovery [136–138]. In silico studies are also critical for fundamental research. In the case of microalgae, Cliff et al. [41] used Alphafold 2 and molecular dynamics to generate the first models of VTC4 protein in *C. reinhardtii*, *C. vulgaris*, *Scenedesmus* sp. and *Gonium pectorale*, based on sequences from databases and their own sequencing data (example in Figure 3). Prior to this study, the function of VTC4 as a polyP polymerase was only confirmed in *C. reinhardtii*, without information on its structure. However, as the models showed conservation of the

VTC catalytic core and the SPX regulatory domains among the microalgae species and yeast, the study of Cliff et al. [41] confirmed the probable function of VTC4 as a polyP polymerase in the other microalgae. In addition, by showing, through molecular docking calculations, the affinity between inositol phosphates and the SPX domain, this study also suggested a regulation of VTC4 proteins in microalgae that may be similar to that in yeast. These findings have significant implications in microalgal biology and biotechnology, as polyPs have important metabolic functions in microalgae and significant research is focusing on the ability of microalgae to accumulate polyPs for P recovery from waste [4,40–42,139].
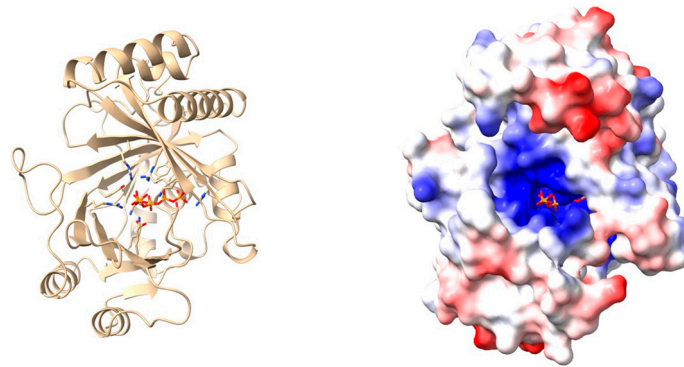

Figure 3. *Gonium pectorale* VTC4 catalytic core domain, shown in complex with a polyP5 molecule (left, cartoon representation, right, surface colors mapped to the electrostatic potential (−10: red to 10: blue KbT/e) at pH 7.0 of the apo form). The models were adapted from Cliff et al. [41] under a Creative Commons license (CC BY-NC-ND 4.0 DEED).

While protein models inform on the putative functions (e.g., substrates and products) and regulatory features of proteins (and, therefore, on the potential differences in the selectivity and specificity between homologous proteins), this should still be validated experimentally, in vitro. Proteins of interest can be isolated from cell extracts using chromatographic methods (Table 1). However, protein isolation can be hindered by the low concentration of certain proteins in cell extracts and/or by the embedding of some proteins in subcellular structures [37,87,91]. Another possibility is to produce proteins of interest using heterologous expression [140]. In recent years, the advances in synthetic biology have facilitated the process of heterologous expression by providing, through automated gene synthesis and molecular biology pipelines, tools for rapid engineering and production of recombinant proteins. Proteins of interest are overexpressed in a host organism, facilitating the extraction and purification of samples suitable for in vitro studies and/or of products for biotechnological applications (e.g., pharmaceuticals). As for proteomics, heterologous expression consists of several steps that can only partly be automated, i.e., host selection, vector design and cloning, transformation of the host, screening of transformants, cultivation of the transformant, heterologous enzyme production, and finally protein recovery and purification. Several hosts can be used for heterologous expression such as *Komagataella phaffii* (formerly *Pichia pastoris*) and *Escherichia coli* [140–142]. The yeast *K. phaffii* is commonly used because of its reliability, the possibility to achieve a high heterologous protein expression level, and the presence of the co-translational and post-translational processing necessary for eukaryotic proteins [140]. Furthermore, the secretory production of recombinant proteins directly into the supernatant of the culture medium is possible with *K. phaffii* [140]. The limited production of endogenous secretory proteins by this yeast is indeed a key attribute for the purification of recombinant protein from supernatants. The microalga *C. reinhardtii* has also been proposed as a promising host to produce recombinant protein [143–145]. The successful expression of several proteins with pharmaceutical relevance has been reported from editing of the nuclear and the chloroplastic genome [145]. While promising for biotechnological applications, considerable developments are, however, still needed before it can be used at large scale [143,144].

7. Future Prospects

The study of model species like *C. reinhardtii* has shed light on a major challenge faced when dealing with un-sequenced microalgal genomes—the lack of comprehensive databases for protein identification. In the past decade, significant progress has been made, as evidenced by the availability of databases with multi-omics data and software to integrate genome sequences and annotation enabling us to broaden ecological and applied research (e.g., PhycoCosm, which integrates genome sequences and annotation for >100 algal genomes, [146]). Aided by AI, it is expected that databases and software will continue developing and improving. As such, the field of microalgal proteomics is on the edge of a huge acceleration of data generation and new opportunities. For instance, while combining omics techniques (multi-omics) may prove critical to building a detailed picture of molecular signatures, their interactions and the phenotypes [16,18,22,24], multi-omics data integration is currently hindered by the heterogeneous nature of data across multi-omics datasets and the natural noise in biological data. AI technology could play a critical role in addressing these challenges [24].

Proteomics has been instrumental improving our understanding of microalgal biology. The continuous development of sample preparation techniques (e.g., extraction, proteases) and high-throughput techniques via AI-based pipelines and software will provide a suitable platform to decipher complex metabolic processes in microalgae. Importantly, many chemicals can be produced by microalgae; therefore, future studies should continue focusing on conditions other than the ones used to trigger lipids production. In addition, the study of PPIs can significantly improve our understanding of protein function, interaction and cellular processes, as shown for the *C. reinhardtii* fatty acid metabolism [84]. Further PPI analyses of key microalgal proteins could, therefore, lead to significant advance for strain (or protein) selection and modifications.

Advances in genetic engineering (e.g., CRISPR/Cas9) have facilitated the creation and analysis of mutants, offering deep insights into vital functions, such as those of global regulators [121,147]. The field of synthetic biology and the production of protein of interest by heterologous expression should therefore continue advancing biological knowledge while also providing the opportunity for new biotechnological applications with the production of proteins of biotechnological interest.

8. Conclusions

In summarizing the recent advances in microalgal proteomics, this review underscores the crucial role of these technologies in understanding both fundamental biological processes and their biotechnological applications. Yet the importance of rigorously maintaining key steps from sample preparation to data analysis for successful proteomics analyses cannot be overstated. Recent years have seen substantial efforts to develop and streamline proteomics frameworks, making this field increasingly accessible to non-experts. This evolution has been complemented by the emergence of companies specializing in affordable data analysis, curation and visualization. With the ongoing integration of AI, these processes are set to become even more efficient and reliable. The insights gained from functional enrichment analyses and protein–protein interactions have been instrumental in addressing complex biological questions and will certainly guide future research directions (e.g., strain improvement). As we transition into the post-omics era, the increasing availability of advanced software and databases promises a broad spectrum of discoveries in both fundamental and applied microalgal research. This journey into the future of microalgal proteomics is not just about enhancing our understanding but also about harnessing the potential of microalgae in innovative and sustainable ways. Notably, microalgae can produce a wide array of chemicals (e.g., nutraceuticals, pharmaceuticals); therefore, the field would benefit from more proteomics studies focusing on conditions triggering the production of these chemicals.

References

1. Falkowski, P.G. The role of phytoplankton photosynthesis in global biogeochemical cycles. *Photosynth. Res.* **1994**, *39*, 235–258. [CrossRef] [PubMed]
2. Bjorbækmo, M.F.M.; Evenstad, A.; Røsæg, L.L.; Krabberød, A.K.; Logares, R. The planktonic protist interactome: Where do we stand after a century of research? *ISME J.* **2020**, *14*, 544–559. [CrossRef] [PubMed]
3. Falkowski, P.G. The power of plankton. *Nature* **2012**, *483*, 17–20. [CrossRef] [PubMed]
4. Plouviez, M.; Oliveira da Rocha, C.S.; Guieysse, B. Intracellular polyphosphate is a P reserve in *Chlamydomonas reinhardtii. Algal Res.* **2022**, *66*, 102779. [CrossRef]
5. Teuma, L.; Sanz-Luque, E.; Guieysse, B.; Plouviez, M. Are Microalgae New Players in Nitrous Oxide Emissions from Eutrophic Aquatic Environments? *Phycology* **2023**, *3*, 356–367. [CrossRef]
6. Thoré, E.S.J.; Muylaert, K.; Bertram, M.G.; Brodin, T. Microalgae. *Curr. Biol.* **2023**, *33*, R91–R95. [CrossRef] [PubMed]
7. Merchant, S.S.; Prochnik, S.E.; Vallon, O.; Harris, E.H.; Karpowicz, S.J.; Witman, G.B.; Terry, A.; Salamov, A.; Fritz-Laylin, L.K.; Marechal-Drouard, L.; et al. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **2007**, *318*, 245–250. [CrossRef] [PubMed]
8. Salome, P.A.; Merchant, S.S. A Series of Fortunate Events: Introducing *Chlamydomonas* as a Reference Organism. *Plant Cell* **2019**, *31*, 1682–1707. [CrossRef] [PubMed]
9. Dupuis, S.; Merchant, S.S. *Chlamydomonas reinhardtii*: A model for photosynthesis and so much more. *Nat. Methods* **2023**, *20*, 1441–1442. [CrossRef]
10. Findinier, J.; Grossman, A.R. *Chlamydomonas*: Fast tracking from genomics. *J. Phycol.* **2023**, *59*, 644–652. [CrossRef]
11. Plouviez, M.; Chisti, Y.; Guieysse, B. Practical guide to algal biomass production: What can we learn from past successes and failures? In *3rd Generation Biofuels*; Woodhead Publishing: Cambridge, UK, 2022; pp. 979–1000. [CrossRef]
12. Chen, H.; Li, T.; Wang, Q. Ten years of algal biofuel and bioproducts: Gains and pains. *Planta* **2019**, *249*, 195–219. [CrossRef] [PubMed]
13. Hamzelou, S.; Belobrajdic, D.; Broadbent, J.A.; Juhász, A.; Lee Chang, K.; Jameson, I.; Ralph, P.; Colgrave, M.L. Utilizing proteomics to identify and optimize microalgae strains for high-quality dietary protein: A review. *Crit. Rev. Biotechnol.* **2023**, 1–16. [CrossRef] [PubMed]
14. Arend, M.; Zimmer, D.; Xu, R.; Sommer, F.; Muhlhaus, T.; Nikoloski, Z. Proteomics and constraint-based modelling reveal enzyme kinetic properties of *Chlamydomonas reinhardtii* on a genome scale. *Nat. Commun.* **2023**, *14*, 4781. [CrossRef]
15. Anand, V.; Singh, P.K.; Banerjee, C.; Shukla, P. Proteomic approaches in microalgae: Perspectives and applications. *3 Biotech* **2017**, *7*, 197. [CrossRef]
16. Kumar, G.; Shekh, A.; Jakhu, S.; Sharma, Y.; Kapoor, R.; Sharma, T.R. Bioengineering of Microalgae: Recent Advances, Perspectives, and Regulatory Challenges for Industrial Application. *Front. Bioeng. Biotechnol.* **2020**, *8*, 914. [CrossRef] [PubMed]
17. Kumar, N.; Kar, S.; Srivastava, A.; Banerjee, C.; Shukla, P. Algal genomics tools: Technological updates and progress. In *Microbial Bioprocesses*; Academic Press: Cambridge, MA, USA, 2023; pp. 67–81. [CrossRef]
18. Chakdar, H.; Hasan, M.; Pabbi, S.; Nevalainen, H.; Shukla, P. High-throughput proteomics and metabolomic studies guide re-engineering of metabolic pathways in eukaryotic microalgae: A review. *Bioresour. Technol.* **2021**, *321*, 124495. [CrossRef]
19. Aslam, B.; Basit, M.; Nisar, M.A.; Khurshid, M.; Rasool, M.H. Proteomics: Technologies and Their Applications. *J. Chromatogr. Sci.* **2017**, *55*, 182–196. [CrossRef] [PubMed]
20. Bludau, I.; Aebersold, R. Proteomic and interactomic insights into the molecular basis of cell functional diversity. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 327–340. [CrossRef]
21. Cui, M.; Cheng, C.; Zhang, L. High-throughput proteomics: A methodological mini-review. *Lab. Investig.* **2022**, *102*, 1170–1181. [CrossRef]
22. Dai, X.; Shen, L. Advances and Trends in Omics Technology Development. *Front. Med.* **2022**, *9*, 911861. [CrossRef]
23. Manzoni, C.; Kia, D.A.; Vandrovcova, J.; Hardy, J.; Wood, N.W.; Lewis, P.A.; Ferrari, R. Genome, transcriptome and proteome: The rise of omics data and their integration in biomedical sciences. *Brief. Bioinform.* **2018**, *19*, 286–302. [CrossRef] [PubMed]
24. Flores, J.E.; Claborne, D.M.; Weller, Z.D.; Webb-Robertson, B.M.; Waters, K.M.; Bramer, L.M. Missing data in multi-omics integration: Recent advances through artificial intelligence. *Front. Artif. Intell.* **2023**, *6*, 1098308. [CrossRef] [PubMed]
25. Veenstra, T.D. Omics in Systems Biology: Current Progress and Future Outlook. *Proteomics* **2021**, *21*, 2000235. [CrossRef] [PubMed]

26. Zeng, W.F.; Zhou, X.X.; Willems, S.; Ammar, C.; Wahle, M.; Bludau, I.; Voytik, E.; Strauss, M.T.; Mann, M. AlphaPeptDeep: A modular deep learning framework to predict peptide properties for proteomics. *Nat. Commun.* **2022**, *13*, 7238. [CrossRef] [PubMed]

27. Mann, M.; Kumar, C.; Zeng, W.F.; Strauss, M.T. Artificial intelligence for proteomics and biomarker discovery. *Cell Syst.* **2021**, *12*, 759–770. [CrossRef] [PubMed]

28. Stransky, S.; Sun, Y.; Shi, X.; Sidoli, S. Ten questions to AI regarding the present and future of proteomics. *Front. Mol. Biosci.* **2023**, *10*, 1–8. [CrossRef]

29. Swan, A.L.; Mobasheri, A.; Allaway, D.; Liddell, S.; Bacardit, J. Application of machine learning to proteomics data: Classification and biomarker identification in postgenomics biology. *Omics* **2013**, *17*, 595–610. [CrossRef]

30. Klukowski, P.; Riek, R.; Guntert, P. Rapid protein assignments and structures from raw NMR spectra with the deep learning technique ARTINA. *Nat. Commun.* **2022**, *13*, 6151. [CrossRef]

31. Yan, S.; Bhawal, R.; Yin, Z.; Thannhauser, T.W.; Zhang, S. Recent advances in proteomics and metabolomics in plants. *Mol. Hortic.* **2022**, *2*, 17. [CrossRef]

32. Carrasco-Reinado, R.; Bermudez-Sauco, M.; Escobar-Nino, A.; Cantoral, J.M.; Fernandez-Acero, F.J. Development of the "Applied Proteomics" Concept for Biotechnology Applications in Microalgae: Example of the Proteome Data in *Nannochloropsis gaditana*. *Mar. Drugs* **2021**, *20*, 38. [CrossRef]

33. Chandramouli, K.; Qian, P.Y. Proteomics: Challenges, techniques and possibilities to overcome biological sample complexity. *Hum. Genom. Proteom.* **2009**, *2009*, 239204. [CrossRef] [PubMed]

34. Vogel, C.; Marcotte, E.M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **2012**, *13*, 227–232. [CrossRef] [PubMed]

35. Davey, N.E.; Simonetti, L.; Ivarsson, Y. The next wave of interactomics: Mapping the SLiM-based interactions of the intrinsically disordered proteome. *Curr. Opin. Struct. Biol.* **2023**, *80*, 102593. [CrossRef] [PubMed]

36. Luck, K.; Kim, D.K.; Lambourne, L.; Spirohn, K.; Begg, B.E.; Bian, W.; Brignall, R.; Cafarelli, T.; Campos-Laborie, F.J.; Charloteaux, B.; et al. A reference map of the human binary protein interactome. *Nature* **2020**, *580*, 402–408. [CrossRef]

37. Dyhrman, S.T.; Jenkins, B.D.; Rynearson, T.A.; Saito, M.A.; Mercier, M.L.; Alexander, H.; Whitney, L.P.; Drzewianowski, A.; Bulygin, V.V.; Bertrand, E.M.; et al. The transcriptome and proteome of the diatom *Thalassiosira pseudonana* reveal a diverse phosphorus stress response. *PLoS ONE* **2012**, *7*, e33768. [CrossRef]

38. Mettler, T.; Muhlhaus, T.; Hemme, D.; Schottler, M.A.; Rupprecht, J.; Idoine, A.; Veyel, D.; Pal, S.K.; Yaneva-Roder, L.; Winck, F.V.; et al. Systems Analysis of the Response of Photosynthesis, Metabolism, and Growth to an Increase in Irradiance in the Photosynthetic Model Organism *Chlamydomonas reinhardtii*. *Plant Cell* **2014**, *26*, 2310–2350. [CrossRef] [PubMed]

39. de Sousa Abreu, R.; Penalva, L.O.; Marcotte, E.M.; Vogel, C. Global signatures of protein and mRNA expression levels. *Mol. Biosyst.* **2009**, *5*, 1512–1526. [CrossRef]

40. Plouviez, M.; Abyadeh, M.; Hasan, M.; Mirzaei, M.; Paulo, J.A.; Guieysse, B. The proteome of *Chlamydomonas reinhardtii* during phosphorus depletion and repletion. *Algal Res.* **2023**, *71*, 103037. [CrossRef]

41. Cliff, A.; Guieysse, B.; Brown, N.; Lockhart, P.; Dubreucq, E.; Plouviez, M. Polyphosphate synthesis is an evolutionarily ancient phosphorus storage strategy in microalgae. *Algal Res* **2023**, *73*, 103161. [CrossRef]

42. Plouviez, M.; Fernandez, E.; Grossman, A.R.; Sanz-Luque, E.; Sells, M.; Wheeler, D.; Guieysse, B. Responses of *Chlamydomonas reinhardtii* during the transition from P-deficient to P-sufficient growth (the P-overplus response): The roles of the vacuolar transport chaperones and polyphosphate synthesis. *J. Phycol.* **2021**, *57*, 988–1003. [CrossRef]

43. Sithtisarn, S.; Yokthongwattana, K.; Mahong, B.; Roytrakul, S.; Paemanee, A.; Phaonakrop, N.; Yokthongwattana, C. Comparative proteomic analysis of *Chlamydomonas reinhardtii* control and a salinity-tolerant strain revealed a differential protein expression pattern. *Planta* **2017**, *246*, 843–856. [CrossRef] [PubMed]

44. Terashima, M.; Specht, M.; Hippler, M. The chloroplast proteome: A survey from the *Chlamydomonas reinhardtii* perspective with a focus on distinctive features. *Curr. Genet.* **2011**, *57*, 151–168. [CrossRef] [PubMed]

45. Schmidt, M.; Gessner, G.; Luff, M.; Heiland, I.; Wagner, V.; Kaminski, M.; Geimer, S.; Eitzinger, N.; Reissenweber, T.; Voytsekh, O.; et al. Proteomic analysis of the eyespot of *Chlamydomonas reinhardtii* provides novel insights into its components and tactic movements. *Plant Cell* **2006**, *18*, 1908–1930. [CrossRef] [PubMed]

46. Perez-Perez, M.E.; Mauries, A.; Maes, A.; Tourasse, N.J.; Hamon, M.; Lemaire, S.D.; Marchand, C.H. The Deep Thioredoxome in *Chlamydomonas reinhardtii*: New Insights into Redox Regulation. *Mol. Plant* **2017**, *10*, 1107–1125. [CrossRef]

47. Morisse, S.; Zaffagnini, M.; Gao, X.H.; Lemaire, S.D.; Marchand, C.H. Insight into protein S-nitrosylation in *Chlamydomonas reinhardtii*. *Antioxid. Redox Signal.* **2014**, *21*, 1271–1284. [CrossRef] [PubMed]

48. Werth, E.G.; McConnell, E.W.; Gilbert, T.S.K.; Couso Lianez, I.; Perez, C.A.; Manley, C.K.; Graves, L.M.; Umen, J.G.; Hicks, L.M. Probing the global kinome and phosphoproteome in *Chlamydomonas reinhardtii* via sequential enrichment and quantitative proteomics. *Plant J.* **2017**, *89*, 416–426. [CrossRef] [PubMed]

49. Wienkoop, S.; Weiss, J.; May, P.; Kempa, S.; Irgang, S.; Recuenco-Munoz, L.; Pietzke, M.; Schwemmer, T.; Rupprecht, J.; Egelhofer, V.; et al. Targeted proteomics for *Chlamydomonas reinhardtii* combined with rapid subcellular protein fractionation, metabolomics and metabolic flux analyses. *Mol. Biosyst.* **2010**, *6*, 1018–1031. [CrossRef] [PubMed]

50. Mastrobuoni, G.; Irgang, S.; Pietzke, M.; Aßmus, H.E.; Wenzel, M.; Schulze, W.X.; Kempa, S. Proteome dynamics and early salt stress response of the photosynthetic organism *Chlamydomonas reinhardtii*. *BMC Genom.* **2012**, *13*, 215. [CrossRef] [PubMed]

51. Mathieu-Rivet, E.; Scholz, M.; Arias, C.; Dardelle, F.; Schulze, S.; Le Mauff, F.; Teo, G.; Hochmal, A.K.; Blanco-Rivero, A.; Loutelier-Bourhis, C.; et al. Exploring the N-glycosylation pathway in *Chlamydomonas reinhardtii* unravels novel complex structures. *Mol. Cell. Proteom.* **2013**, *12*, 3160–3183. [CrossRef]

52. McConnell, E.W.; Werth, E.G.; Hicks, L.M. The phosphorylated redox proteome of *Chlamydomonas reinhardtii*: Revealing novel means for regulation of protein structure and function. *Redox Biol.* **2018**, *17*, 35–46. [CrossRef]

53. Wang, H.; Gau, B.; Slade, W.O.; Juergens, M.; Li, P.; Hicks, L.M. The global phosphoproteome of *Chlamydomonas reinhardtii* reveals complex organellar phosphorylation in the flagella and thylakoid membrane. *Mol. Cell. Proteom.* **2014**, *13*, 2337–2353. [CrossRef] [PubMed]

54. Gurrieri, L.; Del Giudice, A.; Demitri, N.; Falini, G.; Pavel, N.V.; Zaffagnini, M.; Polentarutti, M.; Crozet, P.; Marchand, C.H.; Henri, J.; et al. *Arabidopsis* and *Chlamydomonas* phosphoribulokinase crystal structures complete the redox structural proteome of the Calvin&Benson cycle. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 8048–8053. [CrossRef] [PubMed]

55. Shen, L.; Huang, Z.; Chang, S.; Wang, W.; Wang, J.; Kuang, T.; Han, G.; Shen, J.-R.; Zhang, X. Structure of a C2S2M2N2-type PSII;LHCII supercomplex from the green alga *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 21246–21255. [CrossRef] [PubMed]

56. Lundberg, E.; Borner, G.H.H. Spatial proteomics: A powerful discovery tool for cell biology. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 285–302. [CrossRef] [PubMed]

57. Wang, L.; Patena, W.; Van Baalen, K.A.; Xie, Y.; Singer, E.R.; Gavrilenko, S.; Warren-Williams, M.; Han, L.; Harrigan, H.R.; Hartz, L.D.; et al. A chloroplast protein atlas reveals punctate structures and spatial organization of biosynthetic pathways. *Cell* **2023**, *186*, 3499–3518.e3414. [CrossRef] [PubMed]

58. Zhan, Y.; Marchand, C.H.; Maes, A.; Mauries, A.; Sun, Y.; Dhaliwal, J.S.; Uniacke, J.; Arragain, S.; Jiang, H.; Gold, N.D.; et al. Pyrenoid functions revealed by proteomics in *Chlamydomonas reinhardtii*. *PLoS ONE* **2018**, *13*, e0185039. [CrossRef] [PubMed]

59. Gerin, S.; Mathy, G.; Blomme, A.; Franck, F.; Sluse, F.E. Plasticity of the mitoproteome to nitrogen sources (nitrate and ammonium) in *Chlamydomonas reinhardtii*: The logic of Aox1 gene localization. *Biochim. Biophys. Acta* **2010**, *1797*, 994–1003. [CrossRef] [PubMed]

60. Manjasetty, B.A.; Büssow, K.; Panjikar, S.; Turnbull, A.P. Current methods in structural proteomics and its applications in biological sciences. *3 Biotech* **2011**, *2*, 89–113. [CrossRef]

61. Meyer, J.G. Qualitative and Quantitative Shotgun Proteomics Data Analysis from Data-Dependent Acquisition Mass Spectrometry. *Methods Mol. Biol.* **2021**, *2259*, 297–308. [CrossRef]

62. Stauber, E.J.; Fink, A.; Markert, C.; Kruse, O.; Johanningmeier, U.; Hippler, M. Proteomics of *Chlamydomonas reinhardtii* light-harvesting proteins. *Eukaryot Cell* **2003**, *2*, 978–994. [CrossRef]

63. Quisel, J.D.; Wykoff, D.D.; Grossman, A.R. Biochemical Characterization of the Extracellular Phosphatases Produced by Phosphorus-Deprived *Chlamydomonas reinhardtii*. *Plant Physiol.* **1996**, *111*, 839–848. [CrossRef] [PubMed]

64. Kohinata, T.; Nishino, H.; Fukuzawa, H. Significance of zinc in a regulatory protein, CCM1, which regulates the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant Cell Physiol.* **2008**, *49*, 273–283. [CrossRef]

65. Yang, S.-Y.; Tsuzuki, M.; Muiyachi, S. Carbonic Anhydrase of *Chlamydomonas*: Purification and Studies on its Induction Using Antiserum against *Chlamydomonas* Carbonic Anhydrase. *Plant Cell Physiol.* **1985**, *26*, 25–34. [CrossRef]

66. Kuchitsu, K.; Tsuzuki, M.; Miyachi, S. Characterization of the Pyrenoid Isolated from Unicellular Green Alga *Chlamydomonas reinhardtii*: Particulate Form of RuBisCO Protein. *Protoplasma* **1988**, *144*, 17–24. [CrossRef]

67. Jinag, W.; Cossey, S.; Rosenberg, J.N.; Oyler, G.A.; Olson, B.J.S.C.; Weeks, D.P. A rapid live-cell ELISA for characterizing antibodies against cell surface antigens of *Chlamydomonas reinhardtii* and its use in isolating algae from natural environments with related cell wall components. *BMC Plant Biol.* **2014**, *14*, 244. [CrossRef] [PubMed]

68. Scmitter, J.M.; Jacquot, J.P.; de Lamotte-Guery, F.; Beauvallet, C.; Dutka, S.; Gadal, P.; Decottignies, P. Purification, properties and complete amino acid sequence of the ferredoxin from a green alga, *Chlamydomonas reinhardtii*. *Eur. J. Biochem.* **1988**, *172*, 405–412. [CrossRef] [PubMed]

69. Cid, C.; Garcia-Descalzo, L.; Casado-Lafuente, V.; Amils, R.; Aguilera, A. Proteomic analysis of the response of an acidophilic strain of *Chlamydomonas* sp. (Chlorophyta) to natural metal-rich water. *Proteomics* **2010**, *10*, 2026–2036. [CrossRef]

70. Wang, H.; Alvarez, S.; Hicks, L.M. Comprehensive comparison of iTRAQ and label-free LC-based quantitative proteomics approaches using two *Chlamydomonas reinhardtii* strains of interest for biofuels engineering. *J. Proteome Res.* **2012**, *11*, 487–501. [CrossRef]

71. Helliwell, K.E.; Pandhal, J.; Cooper, M.B.; Longworth, J.; Kudahl, U.J.; Russo, D.A.; Tomsett, E.V.; Bunbury, F.; Salmon, D.L.; Smirnoff, N.; et al. Quantitative proteomics of a B12-dependent alga grown in coculture with bacteria reveals metabolic tradeoffs required for mutualism. *N. Phytol.* **2018**, *217*, 599–612. [CrossRef]

72. Hammel, A.; Zimmer, D.; Sommer, F.; Muhlhaus, T.; Schroda, M. Absolute Quantification of Major Photosynthetic Protein Complexes in *Chlamydomonas reinhardtii* Using Quantification Concatamers (QconCATs). *Front. Plant Sci.* **2018**, *9*, 1265. [CrossRef]

73. Rommelfanger, S.R.; Zhou, M.; Shaghasi, H.; Tzeng, S.C.; Evans, B.S.; Pasa-Tolic, L.; Umen, J.G.; Pesavento, J.J. An Improved Top-Down Mass Spectrometry Characterization of *Chlamydomonas reinhardtii* Histones and Their Post-translational Modifications. *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 1671–1688. [CrossRef] [PubMed]

74. Preimesberger, M.R.; Majumdar, A.; Lecomte, J.T. Dynamics of Lysine as a Heme Axial Ligand: NMR Analysis of the *Chlamydomonas reinhardtii* Hemoglobin THB1. *Biochemistry* **2017**, *56*, 551–569. [CrossRef] [PubMed]

75. Taylor, T.C.; Backlund, A.; Bjorhall, K.; Spreitzer, R.J.; Andersson, I. First crystal structure of Rubisco from a green alga, *Chlamydomonas reinhardtii*. *J. Biol. Chem.* **2001**, *276*, 48159–48164. [CrossRef] [PubMed]

76. Melby, J.A.; Roberts, D.S.; Larson, E.J.; Brown, K.A.; Bayne, E.F.; Jin, S.; Ge, Y. Novel Strategies to Address the Challenges in Top-Down Proteomics. *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 1278–1294. [CrossRef] [PubMed]

77. Dupree, E.J.; Jayathirtha, M.; Yorkey, H.; Mihasan, M.; Petre, B.A.; Darie, C.C. A Critical Review of Bottom-Up Proteomics: The Good, the Bad, and the Future of this Field. *Proteomes* **2020**, *8*, 14. [CrossRef] [PubMed]

78. Anand, S.; Samuel, M.; Ang, C.-S.; Keerthikumar, S.; Mathivanan, S. Label-Based and Label-Free Strategies for Protein Quantitation. In *Methods in Molecular Biology*; Keerthikumar, S., Mathivanan, S., Eds.; Humana Press: New York, NY, USA, 2016; Volume 1549, pp. 31–43.

79. Rao, V.S.; Srinivas, K.; Sujini, G.N.; Kumar, G.N. Protein-protein interaction detection: Methods and analysis. *Int. J. Proteom.* **2014**, *2014*, 147648. [CrossRef] [PubMed]

80. Strotmann, V.I.; Stahl, Y. Visualization of in vivo protein-protein interactions in plants. *J. Exp. Bot.* **2022**, *73*, 3866–3880. [CrossRef] [PubMed]

81. Behal, R.H.; Miller, M.S.; Qin, H.; Lucker, B.F.; Jones, A.; Cole, D.G. Subunit interactions and organization of the *Chlamydomonas reinhardtii* intraflagellar transport complex A proteins. *J. Biol. Chem.* **2012**, *287*, 11689–11703. [CrossRef]

82. Atkinson, N.; Velanis, C.N.; Wunder, T.; Clarke, D.J.; Mueller-Cajar, O.; McCormick, A.J. The pyrenoidal linker protein EPYC1 phase separates with hybrid Arabidopsis-Chlamydomonas Rubisco through interactions with the algal Rubisco small subunit. *J. Exp. Bot.* **2019**, *70*, 5271–5285. [CrossRef]

83. Mackinder, L.C.M.; Chen, C.; Leib, R.D.; Patena, W.; Blum, S.R.; Rodman, M.; Ramundo, S.; Adams, C.M.; Jonikas, M.C. A Spatial Interactome Reveals the Protein Organization of the Algal CO_2-Concentrating Mechanism. *Cell* **2017**, *171*, 133–147.e114. [CrossRef]

84. Blatti, J.L.; Beld, J.; Behnke, C.A.; Mendez, M.; Mayfield, S.P.; Burkart, M.D. Manipulating Fatty Acid Biosynthesis in Microalgae for Biofuel through Protein-Protein Interactions. *PLoS ONE* **2012**, *7*, e42949. [CrossRef] [PubMed]

85. Slocombe, S.P.; Ross, M.; Thomas, N.; McNeill, S.; Stanley, M.S. A rapid and general method for measurement of protein in micro-algal biomass. *Bioresour. Technol.* **2013**, *129*, 51–57. [CrossRef] [PubMed]

86. Wang, W.Q.; Jensen, O.N.; Moller, I.M.; Hebelstrup, K.H.; Rogowska-Wrzesinska, A. Evaluation of sample preparation methods for mass spectrometry-based proteomic analysis of barley leaves. *Plant Methods* **2018**, *14*, 72. [CrossRef] [PubMed]

87. Toyoshima, M.; Sakata, M.; Ohnishi, K.; Tokumaru, Y.; Kato, Y.; Tokutsu, R.; Sakamoto, W.; Minagawa, J.; Matsuda, F.; Shimizu, H. Targeted proteome analysis of microalgae under high-light conditions by optimized protein extraction of photosynthetic organisms. *J. Biosci. Bioeng.* **2019**, *127*, 394–402. [CrossRef] [PubMed]

88. Bichmann, L.; Gupta, S.; Rosenberger, G.; Kuchenbecker, L.; Sachsenberg, T.; Ewels, P.; Alka, O.; Pfeuffer, J.; Kohlbacher, O.; Rost, H. DIAproteomics: A Multifunctional Data Analysis Pipeline for Data-Independent Acquisition Proteomics and Peptidomics. *J. Proteome Res.* **2021**, *20*, 3758–3766. [CrossRef] [PubMed]

89. Pedrioli, P.G.A. Trans-Proteomic Pipeline: A Pipeline for Proteomic Analysis. In *Proteome Bioinformatics, Methods in Molecular Biology*; Hubbard, S.J., Jones, A.R., Eds.; Humana Press: New York, NY, USA, 2010; Volume 604. [CrossRef]

90. Schessner, J.P.; Voytik, E.; Bludau, I. A practical guide to interpreting and generating bottom-up proteomics data visualizations. *Proteomics* **2022**, *22*, 2100103. [CrossRef]

91. Anjos, L.; Estevao, J.; Infante, C.; Mantecon, L.; Power, D.M. Extracting protein from microalgae (*Tetraselmis chuii*) for proteome analysis. *MethodsX* **2022**, *9*, 101637. [CrossRef] [PubMed]

92. Graves, P.R.; Haystead, T.A. Molecular biologist's guide to proteomics. *Microbiol. Mol. Biol. Rev.* **2002**, *66*, 39–63. [CrossRef] [PubMed]

93. Mear, H.; Gillon, P.; Gifuni, I.; Lavenant, L.; Poidevin, A.; Couallier, E. Extraction of soluble proteins by bead milling from *Tetraselmis chui* in two different physiological states. *Algal Res* **2023**, *74*, 103180. [CrossRef]

94. Dixon, C.; Wilken, L.R. Green microalgae biomolecule separations and recovery. *Bioresour. Bioprocess.* **2018**, *5*, 14. [CrossRef]

95. 5 Steps to Fundamental Protein Preparation. Available online: https://www.thermofisher.com/nz/en/home/life-science/protein-biology/protein-purification-isolation/5-steps-fundamental-protein-preparation.html (accessed on 11 December 2023).

96. 5 Steps to Protein Isolation and Purification. Available online: https://www.thermofisher.com/nz/en/home/life-science/protein-biology/protein-purification-isolation/5-steps-protein-purification.html (accessed on 11 December 2023).

97. Knoshaug, E.P.; Gerristen, A.T.; Henard, C.A.; Guarnieri, M.T. Methods for Algal Protein Isolation and Proteome Analysis. In *Metabolic Pathway Engineering*; Himmel, M.E., Bomble, Y.J., Eds.; Humana Press: New York, NY, USA, 2020. [CrossRef]

98. Bandeira, N.; Clauser, K.R.; Pevzner, P.A. Shotgun protein sequencing: Assembly of peptide tandem mass spectra from mixtures of modified proteins. *Mol. Cell. Proteom.* **2007**, *6*, 1123–1134. [CrossRef] [PubMed]

99. Santos, M.D.M.; Lima, D.B.; Fischer, J.S.G.; Clasen, M.A.; Kurt, L.U.; Camillo-Andrade, A.C.; Monteiro, L.C.; de Aquino, P.F.; Neves-Ferreira, A.G.C.; Valente, R.H.; et al. Simple, efficient and thorough shotgun proteomic analysis with PatternLab V. *Nat. Protoc.* **2022**, *17*, 1553–1578. [CrossRef] [PubMed]

100. Meyer, J.G.; Niemi, N.M.; Pagliarini, D.J.; Coon, J.J. Quantitative shotgun proteome analysis by direct infusion. *Nat. Methods* **2020**, *17*, 1222–1228. [CrossRef] [PubMed]

101. Oye, O.K.; Jorgensen, K.M.; Hjelle, S.M.; Sulen, A.; Ulvang, D.M.; Gjertsen, B.T. Gel2DE—A software tool for correlation analysis of 2D gel electrophoresis data. *BMC Bioinf.* **2013**, *14*, 215. [CrossRef] [PubMed]

102. Urban, P.L. Quantitative mass spectrometry: An overview. *Philos. Trans. A Math. Phys. Eng. Sci.* **2016**, *374*, 20150382. [CrossRef] [PubMed]

103. Dressler, F.F.; Bragelmann, J.; Reischl, M.; Perner, S. Normics: Proteomic Normalization by Variance and Data-Inherent Correlation Structure. *Mol. Cell. Proteom.* **2022**, *21*, 100269. [CrossRef] [PubMed]

104. Graw, S.; Tang, J.; Zafar, M.K.; Byrd, A.K.; Bolden, C.; Peterson, E.C.; Byrum, S.D. proteiNorm—A User-Friendly Tool for Normalization and Analysis of TMT and Label-Free Protein Quantification. *ACS Omega* **2020**, *5*, 25625–25633. [CrossRef] [PubMed]

105. Valikangas, T.; Suomi, T.; Elo, L.L. A systematic evaluation of normalization methods in quantitative label-free proteomics. *Brief. Bioinform.* **2018**, *19*, 1–11. [CrossRef] [PubMed]

106. Huang, T.; Wang, J.; Yu, W.; He, Z. Protein inference: A review. *Brief. Bioinform.* **2012**, *13*, 586–614. [CrossRef]

107. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [CrossRef]

108. Koh, C.W.T.; Ooi, J.S.G.; Ong, E.Z.; Chan, K.R. STAGEs: A web-based tool that integrates data visualization and pathway enrichment analysis for gene expression studies. *Sci. Rep.* **2023**, *13*, e7535. [CrossRef] [PubMed]

109. Reimand, J.; Isserlin, R.; Voisin, V.; Kucera, M.; Tannus-Lopes, C.; Rostamianfar, A.; Wadi, L.; Meyer, M.; Wong, J.; Xu, C.; et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat. Protoc.* **2019**, *14*, 482–517. [CrossRef] [PubMed]

110. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **2019**, *47*, D330–D338. [CrossRef] [PubMed]

111. Kanehisa, M.; Furumichi, M.; Sato, Y.; Kawashima, M.; Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **2023**, *51*, 587–592. [CrossRef] [PubMed]

112. May, P.; Wienkoop, S.; Kempa, S.; Usadel, B.; Christian, N.; Rupprecht, J.; Weiss, J.; Recuenco-Munoz, L.; Ebenhoh, O.; Weckwerth, W.; et al. Metabolomics- and proteomics-assisted genome annotation and analysis of the draft metabolic network of *Chlamydomonas reinhardtii*. *Genetics* **2008**, *179*, 157–166. [CrossRef] [PubMed]

113. Wang, Y.; Cong, Y.; Wang, Y.; Guo, Z.; Yue, J.; Xing, Z.; Gao, X.; Chai, X. Identification of Early Salinity Stress-Responsive Proteins in *Dunaliella salina* by isobaric tags for relative and absolute quantitation (iTRAQ)-Based Quantitative Proteomic Analysis. *Int. J. Mol. Sci.* **2019**, *20*, 599. [CrossRef] [PubMed]

114. Kurotani, A.; Yamada, Y.; Sakurai, T. Alga-PrAS (Algal Protein Annotation Suite): A Database of Comprehensive Annotation in Algal Proteomes. *Plant Cell Physiol.* **2017**, *58*, e6. [CrossRef] [PubMed]

115. Goh, W.W.B.; Wong, L. Integrating Networks and Proteomics: Moving Forward. *Trends Biotechnol.* **2016**, *34*, 951–959. [CrossRef] [PubMed]

116. Wu, X.; Hasan, M.A.; Chen, J.Y. Pathway and network analysis in proteomics. *J. Theor. Biol.* **2014**, *362*, 44–52. [CrossRef]

117. Montone, C.M.; Capriotti, A.L.; Cavaliere, C.; La Barbera, G.; Piovesana, S.; Zenezini Chiozzi, R.; Lagana, A. Peptidomic strategy for purification and identification of potential ACE-inhibitory and antioxidant peptides in *Tetradesmus obliquus* microalgae. *Anal. Bioanal. Chem.* **2018**, *410*, 3573–3586. [CrossRef]

118. Carrasco-Reinado, R.; Escobar-Nino, A.; Fajardo, C.; Morano, I.M.; Amil-Ruiz, F.; Martinez-Rodriguez, G.; Fuentes-Almagro, C.; Capilla, V.; Tomas-Cobos, L.; Soriano-Romani, L.; et al. Development of New Antiproliferative Compound against Human Tumor Cells from the Marine Microalgae *Nannochloropsis gaditana* by Applied Proteomics. *Int. J. Mol. Sci.* **2020**, *22*, 96. [CrossRef] [PubMed]

119. Pan, X.; Cao, P.; Su, X.; Liu, Z.; Li, M. Structural analysis and comparison of light-harvesting complexes I and II. *Biochim. Biophys. Acta Bioenerg.* **2020**, *1861*, 148038. [CrossRef] [PubMed]

120. Bujaldon, S.; Kodama, N.; Rappaport, F.; Subramanyam, R.; de Vitry, C.; Takahashi, Y.; Wollman, F.A. Functional Accumulation of Antenna Proteins in Chlorophyll b-Less Mutants of *Chlamydomonas reinhardtii*. *Mol. Plant* **2017**, *10*, 115–130. [CrossRef] [PubMed]

121. Yadav, R.M.; Marriboina, S.; Zamal, M.Y.; Pandey, J.; Subramanyam, R. High light-induced changes in whole-cell proteomic profile and its correlation with the organization of thylakoid super-complex in cyclic electron transport mutants of *Chlamydomonas reinhardtii*. *Front. Plant. Sci.* **2023**, *14*, 1198474. [CrossRef] [PubMed]

122. Hsieh, S.I.; Castruita, M.; Malasarn, D.; Urzica, E.; Erde, J.; Page, M.D.; Yamasaki, H.; Casero, D.; Pellegrini, M.; Merchant, S.S.; et al. The proteome of copper, iron, zinc, and manganese micronutrient deficiency in *Chlamydomonas reinhardtii*. *Mol. Cell. Proteom.* **2013**, *12*, 65–86. [CrossRef] [PubMed]

123. Schmollinger, S.; Muhlhaus, T.; Boyle, N.R.; Blaby, I.K.; Casero, D.; Mettler, T.; Moseley, J.L.; Kropat, J.; Sommer, F.; Strenkert, D.; et al. Nitrogen-Sparing Mechanisms in Chlamydomonas Affect the Transcriptome, the Proteome, and Photosynthetic Metabolism. *Plant Cell* **2014**, *26*, 1410–1435. [CrossRef] [PubMed]

124. Ge, Y.; Ning, Z.; Wang, Y.; Zheng, Y.; Zhang, C.; Figeys, D. Quantitative proteomic analysis of *Dunaliella salina* upon acute arsenate exposure. *Chemosphere* **2016**, *145*, 112–118. [CrossRef] [PubMed]

125. Jiang, X.; Yang, L.; Wang, Y.; Jiang, F.; Lai, J.; Pan, K. Proteomics Provide Insight into the Interaction between Selenite and the Microalgae *Dunaliella salina*. *Processes* **2023**, *11*, 563. [CrossRef]

126. Leon-Vaz, A.; Romero, L.C.; Gotor, C.; Leon, R.; Vigara, J. Effect of cadmium in the microalga *Chlorella sorokiniana*: A proteomic study. *Ecotoxicol. Environ. Saf.* **2021**, *207*, 111301. [CrossRef]

127. Guarnieri, M.T.; Nag, A.; Smolinski, S.L.; Darzins, A.; Seibert, M.; Pienkos, P.T. Examination of triacylglycerol biosynthetic pathways via de novo transcriptomic and proteomic analyses in an unsequenced microalga. *PLoS ONE* **2011**, *6*, e25851. [CrossRef]

128. Patel, A.K.; Huang, E.L.; Low-Decarie, E.; Lefsrud, M.G. Comparative Shotgun Proteomic Analysis of Wastewater-Cultured Microalgae: Nitrogen Sensing and Carbon Fixation for Growth and Nutrient Removal in *Chlamydomonas reinhardtii*. *J. Proteome Res.* **2015**, *14*, 3051–3067. [CrossRef] [PubMed]

129. Rai, V.; Muthuraj, M.; Gandhi, M.N.; Das, D.; Srivastava, S. Real-time iTRAQ-based proteome profiling revealed the central metabolism involved in nitrogen starvation induced lipid accumulation in microalgae. *Sci. Rep.* **2017**, *7*, 45732. [CrossRef] [PubMed]

130. Ben Amor, F.; Elleuch, F.; Ben Hlima, H.; Garnier, M.; Saint-Jean, B.; Barkallah, M.; Pichon, C.; Abdelkafi, S.; Fendri, I. Proteomic Analysis of the Chlorophyta Dunaliella New Strain AL-1 Revealed Global Changes of Metabolism during High Carotenoid Production. *Mar. Drugs* **2017**, *15*, 293. [CrossRef]

131. Wei, S.; Bian, Y.; Zhao, Q.; Chen, S.; Mao, J.; Song, C.; Cheng, K.; Xiao, Z.; Zhang, C.; Ma, W.; et al. Salinity-Induced Palmella Formation Mechanism in Halotolerant Algae *Dunaliella salina* Revealed by Quantitative Proteomics and Phosphoproteomics. *Front. Plant Sci.* **2017**, *8*, 810. [CrossRef]

132. Wang, B.; Pan, X.; Wang, F.; Liu, L.; Jia, J. Photoprotective carbon redistribution in mixotrophic *Haematococcus pluvialis* under high light stress. *Bioresour. Technol.* **2022**, *362*, 127761. [CrossRef]

133. Chen, C.; Harst, A.; You, W.; Xu, J.; Ning, K.; Poetsch, A. Proteomic study uncovers molecular principles of single-cell-level phenotypic heterogeneity in lipid storage of Nannochloropsis oceanica. *Biotechnol. Biofuels* **2019**, *12*, 21. [CrossRef]

134. Roccuzzo, S.; Couto, N.; Karunakaran, E.; Kapoore, R.V.; Butler, T.O.; Mukherjee, J.; Hansson, E.M.; Beckerman, A.P.; Pandhal, J. Metabolic Insights into Infochemicals Induced Colony Formation and Flocculation in Scenedesmus subspicatus Unraveled by Quantitative Proteomics. *Front. Microbiol.* **2020**, *11*, 792. [CrossRef] [PubMed]

135. Salama, E.S.; Govindwar, S.P.; Khandare, R.V.; Roh, H.S.; Jeon, B.H.; Li, X. Can Omics Approaches Improve Microalgal Biofuels under Abiotic Stress? *Trends Plant. Sci.* **2019**, *24*, 611–624. [CrossRef] [PubMed]

136. Service, R.B. 2021 Breakthrough of the year: Protein structures for all. *Science* **2021**, *374*, eabm4805. Available online: https://www.science.org/content/article/breakthrough-2021 (accessed on 1 April 2024). [CrossRef]

137. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373*, 871–876. [CrossRef]

138. Callaway, E. 'The entire protein universe': AI predicts shape of nearly every known protein. *Nature* **2022**, *608*, 15–16. Available online: https://www.nature.com/articles/d41586-022-02083-2 (accessed on 1 April 2024). [CrossRef] [PubMed]

139. Plouviez, M.; Bolot, P.; Shilton, A.; Guieysse, B. Phosphorus uptake and accumulation in *Chlamydomonas reinhardtii*: Influence of biomass concentration, phosphate concentration, phosphorus depletion time, and light supply. *Algal Res.* **2023**, *71*, 103085. [CrossRef]

140. Karbalaei, M.; Rezaee, S.A.; Farsiani, H. Pichia pastoris: A highly successful expression system for optimal synthesis of heterologous proteins. *J. Cell Physiol.* **2020**, *235*, 5867–5881. [CrossRef] [PubMed]

141. Arísticas Ribalta, R.C.; Valdés, L.M.; Lafargue Gámez, M.; Rodríguez Davydenko, S.; Dubreucq, E.; Perrier, V.; Moreau, B.; Vidal, R.F. Constitutive high expression level of a synthetic deleted encoding gene of *Talaromyces minioluteus* endodextranase variant (r–TmDEX49A–ΔSP–ΔN30) in *Komagataella phaffii* (*Pichia pastoris*). *Appl. Sci.* **2022**, *12*, 7562. [CrossRef]

142. Hothorn, M.; Neumann, H.; Lenherr, E.D.; Wehner, M.; Rybin, V.; Hassa, P.O.; Uttenweiler, A.; Reinhardt, M.; Schmidt, A.; Seiler, J.; et al. Catalytic core of a membrane-associated eukaryotic polyphosphate polymerase. *Science* **2009**, *24*, 513–516. [CrossRef]

143. Stoffels, L.; Finlan, A.; Mannall, G.; Purton, S.; Parker, B. Downstream Processing of *Chlamydomonas reinhardtii* TN72 for Recombinant Protein Recovery. *Front. Bioeng. Biotechnol.* **2019**, *7*, 383. [CrossRef]

144. Rosales-Mendoza, S.; Paz-Maldonado, L.M.; Soria-Guerra, R.E. *Chlamydomonas reinhardtii* as a viable platform for the production of recombinant proteins: Current status and perspectives. *Plant Cell Rep.* **2012**, *31*, 479–494. [CrossRef]

145. Masi, A.; Leonelli, F.; Scognamiglio, V.; Gasperuzzo, G.; Antonacci, A.; Terzidis, M.A. *Chlamydomonas reinhardtii*: A Factory of Nutraceutical and Food Supplements for Human Health. *Molecules* **2023**, *28*, 1185. [CrossRef]

146. Grigoriev, I.V.; Hayes, R.D.; Calhoun, S.; Kamel, B.; Wang, A.; Ahrendt, S.; Dusheyko, S.; Nikitin, R.; Mondo, S.J.; Salamov, A.; et al. PhycoCosm, a comparative algal genomics resource. *Nucleic Acids Res.* **2021**, *49*, D1004–D1011. [CrossRef]

147. Nguyen, T.H.T.; Park, S.; Jeong, J.; Shin, Y.S.; Sim, S.J.; Jin, E. Enhancing lipid productivity by modulating lipid catabolism using the CRISPR-Cas9 system in Chlamydomonas. *J. Appl. Phycol.* **2020**, *32*, 2829–2840. [CrossRef]