Establishing the red microalga *Porphyridium purpureum* as a novel platform for the production of recombinant proteins

---

Alexander Hammel

Dissertation

zur Erlangung des akademischen Grades
"doctor rerum naturalium"
(Dr. rer. nat.)
in der Wissenschaftsdisziplin "Molekulare Biotechnologie"

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
Institut für Biochemie und Biologie
der Universität Potsdam
und
Max-Planck-Institut für molekulare Pflanzenphysiologie

Potsdam, 04. März 2024

Hauptbetreuer: Prof. Dr. Ralph Bock

Gutachter*innen: Prof. Dr. Ralph Bock
Prof. Dr. Olaf Kruse
Prof. Dr. Jörg Nickelsen

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-63270
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-632709

Eidesstattliche Erklärung

Hiermit versichere ich, die vorliegende Arbeit selbständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben. Ich versichere ebenfalls, dass die Arbeit an keiner anderen Hochschule als der Universität Potsdam eingereicht wurde.

Alexander Hammel

Potsdam, Oktober 2023

Contents

List of figures

List of tables

List of abbreviations

2-ME	2-mercaptoethanol
°C	degree celsius
CAI	dodon adaptation index
ASW	artificial seawater
3xHA	triple hemagglutinin tag
Aa	amino acid
Ab	antibody
Ag	antigen
AGL	antigenic loop
ATP	adenosine tri phosphate
BSA	bovine serum albumin
BUSCO	benchmarking universal single-copy orthologs
CDS	coding sequence
CUB	codon usage bias
cyt	cytosolic
DAAs	direct-acting antiviral medicines
DW	dry weight
EPS	extra(cellular) polysaccharides
ER	endoplasmic reticulum
FACS	fluorescence-activated cell sorting
FW	fresh weight
GFP	green fluorescent protein
HBV	hepatitis b virus
HCV	hepatitis c virus
HVR	hypervariable region
kb	kilo base

kDa	kilo dalton
lsASW	low salt artificial sea water
Man	mannose
Mb	mega base pairs
MW	molecular weight
MWCO	molecular weight cut off
nAbs	neutralizing antibodies
NTCP	sodium taurocholate cotransporting polypeptide
ORF	open reading frame
PBR	photobioreactor
PTM	posttranslational modification
PUFAs	polyunsaturated fatty acids
RBD	receptor-binding domain
RCA	relative codon adaptation
sec	secreted
SP	signal peptide
SRP	signal recognition particles
SVP	subviral particle
TM	tunicamycin
TMD	transmembrane domain
TP	total protein
TSP	total soluble protein
UTR	untranslated region
Xyl	xylose
YFP	yellow fluorescent protein
ε	extinction coefficient

Abstract

Microalgae have been recognized as a promising green production platform for recombinant proteins. The majority of studies on recombinant protein expression have been conducted in the green microalga *C. reinhardtii*. While promising improvement regarding nuclear transgene expression in this alga has been made, it is still inefficient due to epigenetic silencing, often resulting in low yields that are not competitive with other expressor organisms. Other microalgal species might be better suited for high-level protein expression, but are limited in their availability of molecular tools.

The red microalga *Porphyridium purpureum* recently emerged as candidate for the production of recombinant proteins. It is promising in that transformation vectors are episomally maintained as autonomously replicating plasmids in the nucleus at a high copy number, thus leading to high expression values in this red alga.

In this work, we expand the genetic tools for *P. purpureum* and investigate parameters that govern efficient transgene expression. We provide an improved transformation protocol to streamline the generation of transgenic lines in this organism. After being able to efficiently generate transgenic lines, we showed that codon usage is a main determinant of high-level transgene expression, not only at the protein level but also at the level of mRNA accumulation. The optimized expression constructs resulted in YFP accumulation up to an unprecedented 5% of the total soluble protein. Furthermore, we designed new constructs conferring efficient transgene expression into the culture medium, simplifying purification and harvests of recombinant proteins. To further improve transgene expression, we tested endogenous promoters driving the most highly transcribed genes in *P. purpureum* and found minor increase of YFP accumulation.

We employed the previous findings to express complex viral antigens from the hepatitis B virus and the hepatitis C virus in *P. purpureum* to demonstrate its feasibility as producer of biopharmaceuticals. The viral glycoproteins were successfully produced to high levels and could reach their native confirmation, indicating a functional glycosylation machinery and an appropriate folding environment in this red alga. We could successfully upscale the biomass production of transgenic lines and with that provide enough material for immunization trials in mice that were performed in collaboration. These trials showed no toxicity of neither the biomass nor the purified antigens, and, additionally, the algal-produced antigens were able to elicit a strong and specific immune response.

The results presented in this work pave the way for *P. purpureum* as a new promising producer organism for biopharmaceuticals in the microalgal field.

1 Introduction

1.1 The red microalga *Porphyridium purpureum* and its potential as a recombinant protein production platform

1.1.1 Recombinant protein production

Modern biotechnology emerged in the mid-20th century, driven by the discovery of the potential of microorganisms as production platforms in industrial fermentation. A significant milestone was the mass production of penicillin in *Penicillium chrysogenum* starting in 1942 (reviewed in Gaynes, 2017). However, the true potential of biotechnology unfolded after the revelation of DNA as the carrier of genetic information by Watson and Crick in 1953, and the first demonstration of recombinant DNA techniques by Cohen and Boyer in 1973 (Cohen et al., 1973; Crick and Watson, 1953).

Genetic engineering opened the doors to producing recombinant proteins, with the pioneering production of insulin in *Escherichia coli* (Goeddel et al., 1979). This breakthrough allowed the introduction of genetic information into fast-growing microorganisms for mass production of therapeutic proteins, reducing reliance on low-yield animal sources.

Currently, the market for biopharmaceuticals, comprising recombinant proteins and nucleic acid-based and genetically engineered cell-based products, is one of the fastest growing sectors in the health industry (Walsh and Walsh, 2022). Since the first approval of a recombinant protein for human medication in 1982, a total of 541 therapeutic recombinant proteins, including antibodies, vaccines, enzymes, hormones and recombinant growth and clotting factors, have been approved for human use, with 197 approvals between 2018 and 2022 (reviewed in Walsh & Walsh, 2022). Monoclonal antibodies represent a significant portion, accounting for 80.2% of total protein-based biopharmaceutical sales.

Selecting an appropriate heterologous expression system is crucial for successful recombinant protein production. Each system has its advantages and disadvantages, as summarized in Table 1. Mammalian cell lines are the most commonly used expression systems for approved biopharmaceuticals, followed by *E. coli* and yeast. Mammalian cell lines are preferred for proteins requiring post-translational modifications (PTMs), such as disulfide bonds and glycosylation, which is crucial for antibodies due to their glycoprotein nature. Prokaryotic systems lack a glycosylation machinery, and often have problems folding proteins derived from eukaryotes into the correct confirmation, thus resulting in inclusion body formation. Furthermore, there are safety concerns due to endotoxins produced by *E. coli*. On the other hand the bacterial system is relatively cheap, benefits from an advanced collection of genetic tools, shows fast growth kinetics and is the preferred system for easy-to-express proteins or peptides with no PTMs (Rosano and Ceccarelli, 2014)

Table 1: Advantages and disadvantages of different expression organisms. Compiled from Barbosa et al., 2023; Demain and Vaishnav, 2009; Owczarek et al., 2019; Yan et al., 2016.

Expression system	Mammalian cells	Bacteria	Yeast	Insect cells	Plants	Algae
Growth	slow	fast	fast	slow	slow	fast
Protein yield	medium	high	medium	medium	medium	medium
Cost	high	low	medium	medium	low	low
Glycosylation*	correct	none	incorrect	similar	similar	similar
Biosafety	low	low	high	high	high	high
Tools	medium	high	high	low	medium	low
Protein folding	high	low	medium	high	high	high
Resource efficacy	low	medium	medium	low	high	high

* Compared to humans

Yeast systems perform N-glycosylation and often allow for correct protein folding, but usually hyperglycosylate proteins, leading to reduced activity, low receptor binding or unwanted immunogenicity (Demain and Vaishnav, 2009; Li and d'Anjou, 2009). Even though extensive efforts were made to humanize the glycosylation patterns of yeast, insects and plants, the results were not satisfying enough to ensure safe recombinant protein production in these systems, highlighting the complexity of human glycosylation (Khan et al., 2016). Despite disadvantages like high costs, limited scalability, and biosafety risks from potential viral infections, mammalian cell systems remain the preferred platform for biopharmaceuticals used in humans.

All the above-described production platforms require expensive growth media with additional carbon sources, vitamins and growth factors, thus diminishing the resource efficacy of the final product. In light of the current climate crisis, the need for new, green production platforms is higher than ever, and research investments in ecologically sustainable biotechnologies have increased substantially in the last years. Photosynthetic organisms like plants and algae could be the solution to this problem (Barbosa et al., 2023; Xia et al., 2021).

1.1.2 Microalgae as emerging biotechnology platform

Microalgae, which include eukaryotic algae and prokaryotic cyanobacteria, are unicellular, photosynthetic organisms found in diverse terrestrial and aquatic environments. This study

focuses solely on eukaryotic microalgae. These microorganisms trace back to a common ancestor that acquired photosynthetic capabilities through primary symbiosis by incorporating a cyanobacterium as a chloroplast organelle (Chan et al., 2011). They are classified into three main groups: red, green, and glaucophyte algae, with the red group being the phylogenetically most diverse (Guiry, 2012). Through secondary and tertiary endosymbioses, microalgae gave rise to various groups of photosynthetic organisms that play crucial roles in the aquatic food chain and oxygen production (Gould et al.,2008).

In recent years, the potential of microalgae as biotechnological host organisms was recognized and three main application areas were classified. Firstly, the microalgal biomass can be directly used, for example, by extracting the valuable natural products derived from primary or secondary metabolism (value-added products) like polysaccharides, polyunsaturated fatty acids (PUFAs), antioxidants, pigments, and other metabolites with medical properties (Visioli and Artaria, 2017; Alves et al., 2018; Yang et al., 2019; Barta et al., 2021). To truly harness the potential of competitive industrial production of these compounds, methods for bulk production of microalgae in either closed bioreactors or open pond bioreactors are needed. Substantial research efforts are currently focused on the techno-economic properties of these cultivation systems (Bhatt et al., 2022; Vázquez-Romero et al., 2022). Secondly, microalgae are explored for bioremediation and wastewater treatment purposes (Devi et al., 2023). Lastly, they are considered green biofactories for producing high-value recombinant proteins, which has been the focus of this thesis (Banerjee and Ward, 2022).

The concept of using photosynthetic organisms for recombinant protein production, known as molecular farming, emerged around 30 years ago (reviewed, e.g. in Fischer and Buyel, 2020). The first successful production of a functional recombinant antibody in tobacco plants in 1989 marked the inception of molecular farming (Hiatt et al., 1989).

Microalgae share some advantages with plant systems over the commonly used above-described heterologous production systems. Microalgae are photosynthetic organisms and have minimal requirements for their growth medium, being able to grow photoautotrophically without the addition of a reduced carbon source to produce biomass. Plants and microalgae are easily scalable by utilizing existing agricultural processes or open-pond bioreactors, and they pose no risk of disease transmission to humans, eliminating a major biosafety concern. Microalgae can exhibit high protein expression levels, especially when transgenes are expressed from the chloroplast genome. All the above-mentioned advantages render them (theoretically) a very cheap production platform. As eukaryotes, they are capable of performing complex glycosylation and protein folding, which makes them suitable for nuclear expression

of challenging proteins (Bock, 2015; Su et al., 2023). Additionally, their glycosylation pattern is closer to mammalian glycosylation that that of yeasts (Banerjee and Ward, 2022).

Despite intensive ongoing research in molecular farming and the founding of over 50 companies by 2005, only a few products have reached the market due to bottlenecks such as skepticism towards genetically modified organisms (GMOs), especially in plants, and ecological safety concerns related to environmental contamination (Fischer and Buyel, 2020). The lack of protein secretion in plants leads to high costs in the downstream purification process. The yield of purified protein per growth space is relatively low and it can only be overcome by massive upscaling, ultimately leading to space competition with food crop plants. Consequently, the true production costs of recombinant protein production in plants remain higher than those in mammalian cells or bacterial systems (Schillberg et al., 2019).

In contrast, unicellular microalgae address these disadvantages of the plant system. They possess higher theoretical photosynthetic efficacy, faster growth rates, and can be cultivated in dense cultures in enclosed photobioreactors on non-arable land, resulting in higher areal productivity and elimination of competition with crops and the risk of GMO contamination (Bhola et al., 2014; Barbosa et al., 2023). Genetic manipulation of microalgae is faster and often easier compared to plants, further reducing the costs of this production system. Many microalgae are "Generally Recognized As Safe" (GRAS), thus increasing biosafety and reducing downstream processing costs, and enabling direct utilization of algal biomass for oral applications (e.g., oral vaccination Torres-Tiji et al., 2020; Dehghani et al., 2022). Another aspect that reduces downstream processing costs is the possibility to secrete proteins into the medium, making purification of the desired protein much easier.

In summary, microalgae offer the advantages of eukaryotic and photosynthetic production systems and combine them with many of the benefits of prokaryotic microorganisms as biotechnological production platforms.

1.1.3 The search for new microalgal species as green biofactories

The most important prerequisite for the use of microalgae as protein production platforms is the availability of genetic tools. Compared to other organisms like yeast and bacteria, microalgae still lag behind, although a lot of effort was made in the recent years to develop genetic tools for several microalgal species (Fabris et al., 2020). Undoubtedly, the best-studied microalgal model organism to date is the green alga *Chlamydomonas reinhardtii*. The toolbox of this alga includes promoters conferring different expression strengths, synthetic regulatory elements like riboswitches, and well-characterized targeting sequences specific for different compartments. These tools greatly facilitate synthetic biology approaches in this green alga (Crozet et al., 2018; Moulin et al., 2013; Schroda et al., 2000b). A strain capable of high

transgene expression was recently characterized and is commonly used in biotechnological studies (Neupert et al., 2009; Barahimipour et al., 2016; Neupert et al., 2020). Other microalgal species are mostly used for the production of their natural valuable products rather than of recombinant proteins, mostly due to the lack of appropriate molecular tools. Nevertheless, several proof-of-principle studies showed that the expression of antibodies, hormones, cytokines and vaccines in *C. reinhardtii*, *Dunaliella salina* and *Schizochytrium sp*. is technically possible (Bayne et al., 2013; Geng et al., 2003; Mayfield et al., 2003; Ramos-Vega et al., 2021; Rasala et al., 2010; Soria-Guerra et al., 2014a).

Most recently, the SARS-CoV-2 spike protein receptor-binding domain (RBD) was expressed from the nucleus of *C. reinhardtii*, *Phaeodactylum tricornutum* and *Chlorella* sp., demonstrating the capability of expressing complex fragments of viral proteins in their active form in microalgal systems, which can be used as subunit vaccines (Berndt et al., 2021; Malla et al., 2021; Slattery et al., 2022). Even though these studies were an important step in showing the suitability of microalgae to express possible subunit vaccines, the expressed antigens were targeted to the cytosol or to the ER, requiring purification from whole cells. Only one study demonstrated the successful secretion of a functional full-length SARS-CoV-2 spike protein ectodomain into the medium (Kiefer et al., 2022). All studies have in common that the yield of the protein obtained was poor, and by far not sufficient for any biotechnological application. The yield of recombinant RBD in *P. tricornutum* ranged from 28 µg to 34 µg from a 5 L cultivation (Slattery et al., 2022). The yield of RBD in *C. reinhardtii* was 31 µg per gram fresh weight, and even decreased to 1.8 µg per gram fresh weight after purification (Berndt et al., 2021). In *Chlorella*, the yield reached 1.14 µg per gram fresh weight (Malla et al., 2021). The yield for the secreted spike protein in *C. reinhardtii* was 11.2 µg per liter culture volume (Kiefer et al., 2022). The yields of other therapeutic proteins expressed in different microalgae were found to be in the same range of a few µg per gram biomass (Banerjee and Ward, 2022; Rosales-Mendoza et al., 2020)

After decades of research, and despite the recent improvements mentioned above, the main disadvantage of *Chlamydomonas* as a green biofactory is still its poor expression of foreign genes from the nucleus (reviewed in Schroda, 2019). Position effects and gene silencing are likely the main reasons why nuclear transgene expression is challenging in this alga. Protein expression from the chloroplast genome usually leads to higher yields. However, glycosylation and secretion is not achievable in this organelle (Ma et al., 2022). All this demonstrates the need for new microalgal species capable of high-yield nuclear transgene expression. The red microalga *Porphyridium purpureum*, which is already used for the production of natural compounds, offers great potential to become an emerging new platform for recombinant protein production (Li et al., 2019).

1.2 The red microalga *Porphyridium purpureum*

1.2.1 General information

Red algae (Rhodophyta) form a monophyletic group comprised of around 20,000 estimated species, many of which of ecological value, with a wide variety of morphologies (Guiry, 2012). Approximately 95% of red algal species are seaweeds belonging to the class of Bangiophyceae and Florideophyceae, and only 5% are microalgae, belonging to the group of Compsopogonophyceae, Porphyridiophyceae, Rhodellophyceae and Stylonematophyceae (Qiu et al., 2016). The unicellular red microalgae *P. purpureum* belongs to the group of Porphyridiophyceae and was first described in 1841 as *Porphyridium cruentum* Nägeli (Nägeli, 1849). Recent publications commonly refer to this species as *P. purpureum*, but there are still cases where the previous name *P. cruentum* is found in the literature.

P. purpureum is a spherical mesophilic red microalga with an average diameter of 6 to 15 µm. It can be found in a variety of habitats like seawater, soil and brackish water. A single, star shaped chloroplast occupies approx. 65% of the cellular volume (Arteni et al., 2008; Gantt and Conti, 1965). Like all red algae, *P. purpureum* does not possess flagella. Its sexual reproduction cycle is yet to be elucidated, and only the diploid phase is known (Bhattacharya et al., 2013). The cell wall of this red alga is composed of sulfated polysaccharides that are partially (around 30%) released into the medium during growth. The exact structure of the high molecular weight polysaccharides building the cell wall has not yet been resolved, but they are mostly composed of D-xylose, D- and L-lactose, D-glucose, D-glucuronic acid and a variable content of sulfate (1% - 14%) (Bernaerts et al., 2018).

This unicellular red alga naturally accumulates value-added products such as extracellular sulfated polysaccharides (EPS), phycobiliproteins (especially phycoerythrin) and polyunsaturated fatty acids (PUFAs) to high levels, and is thus an attractive species for the industry (Gaignard et al., 2019; Li et al., 2019). In fact, large-scale cultivation of *P. purpureum* was already realized to produce sulphated polysaccharides that were shown to possess antiviral, antitumor and antioxidant activities, and have unique rheological properties, making them excellent thickening agents and biolubricants (Arad and Levy-Ontman, 2010; Arad and Richmond, 2003; Huheihel et al., 2002). The total red algal biomass has also beneficial health properties. After feeding of *P. purpureum* biomass to rats, serum cholesterol levels were reduced, indicating no toxicity and even a positive nutritional value, rendering the alga a promising organism for oral vaccine delivery (Dvir et al., 2000). Phycoerythrin is a protein assembled in the phycobilisome, a water-soluble complex associated to the light-harvesting apparatus of the thylakoid membrane. Phycoerythrin confers the bright red-pinkish color to *P. purpureum*, and is used as a natural food colorant and fluorescent probe (Bermejo et al., 2007; Marcati et al., 2014). These examples show that *P. purpureum* is a commonly used red

microalga species for the production of high-value products, and the knowledge for large-scale production already exists. However, *Porphyridium* is not an established model organism in biology because of the lack of molecular tools.

1.2.2 Molecular tools in *P. purpureum*

As stated above, the availability of molecular tools is one of the most important perquisites for the employment of a species as a successful and cost-effective protein production platform. The next step to further advance the biotechnological utilization of *P. purpureum* is the establishment of genetic modification techniques and the advancement of its molecular toolbox. The first draft of the *P. purpureum* genome was published in 2013 and refined in 2019 (Bhattacharya et al., 2013; Lee et al., 2019). The nuclear genome contains 9898 putative genes and is ~22.1 Mb in size. The diploid genome was sequenced using Oxford Nanopore technology with a genome coverage of 45 times and contained 90.4% of the 429 conserved eukaryotic Benchmarking Universal Single-Copy Orthologs (BUSCO) (Simão et al., 2015). These genes are a standard set of genes found in eukaryotic organisms and can work as a benchmark to assess the quality of the genome assembly. The average GC content was 55.5%, which is significantly lower than that of *C. reinhardtii* (64% GC content; Merchant et al., 2007). Only 2.8% of the genes in *P. purpureum* contain introns and at least 453 genes were acquired by endosymbiotic or horizontal gene transfer indicating a prominent role of this red alga in gene transfer between prokaryotes and photosynthetic eukaryotes (Bhattacharya et al., 2013). The 217 kb plastid genome was first sequenced in 2014 and contains 224 protein-coding genes, which is a larger set than that of other algae (Mohanta et al., 2020; Tajima et al., 2014). The availability of both the nuclear and the plastid genome sequences allows the development of molecular tools for genetic engineering.

Only a few studies demonstrated the successful transformation of the nuclear genome of *P. purpureum*, using *Agrobacterium*- mediated or biolistic transformation (Li and Bock, 2018; Prasad et al., 2018). Interestingly, when nuclear transformation using the biolistic approach was conducted, the transformation plasmids were stably and episomally maintained at a high copy number (around 20 to 30 copies) in the nucleus of *P. purpureum* and led to high transgenic GFP expression (Li and Bock, 2018). On the other hand, if transformed via *Agrobacterium*-mediated transformation, the authors claimed that the transformation vector stably integrated in the nuclear genome of *P. purpureum*, which was demonstrated by southern blot analysis (Prasad et al., 2018). It is important to note that different strains of *P. purpureum* were used in these two studies, and these strains also differ from the original sequenced strain. As a non-model organism, several different laboratory strains are used and no common consent on which strain to use has yet been agreed on, which would be particularly important for the implementation of reliable and reproducible genetic engineering

protocols. Two studies claimed to have successfully transformed the plastid genome, but showed serious flaws in the methodology and the interpretation of the results, questioning if true transplastomic lines were isolated (Han et al., 2022; Lapidot et al., 2002). The results in the study performed by Lapidot et al. were most likely explained by an episomal transformation as discussed in Li and Bock, 2019. The study performed by Han et al. lacks consistent molecular data on DNA and protein level, and misses crucial controls, which makes the claimed generation of transplastomic lines questionable. The presented results can be explained with an episomal transformation and nuclear expression as well.

The episomal maintenance of the transformation plasmids after biolistic transformation is likely one of the reason why recombinant proteins can accumulate at unprecedented high levels in *P. purpureum*, emphasizing the potential of this microalga as an excellent protein production platform (Li and Bock, 2018). However, fundamental research on red algae is lagging behind that in plants and green algae. Many molecular tools available for *C. reinhardtii* have yet not been tested and implemented in *P. purpureum*.

The goal of this thesis was to expand the molecular toolbox of *P. purpureum* and demonstrate the feasibility of this red alga in becoming a competitive production system for recombinant proteins in algal biotechnology. Different strategies that were previously successful in *C. reinhardtii* for achieving high transgene expression were tested in *P. purpureum* in order to further increase its recombinant protein production capacity. Endogenous expression elements and protein targeting signals (including those for secretion) were tested and established. Finally, the potential of *P. purpureum* as a superb recombinant protein production system was validated by expressing proteins of high biotechnological value, with a focus on candidate vaccine.

1.3 Recombinant protein expression strategies in microalgae

The subsequent section presents construct design strategies and tools for recombinant protein expression in microalgae using mainly *C. reinhardtii* as a model. A comparative analysis will be conducted with the existing knowledge in *P. purpureum*, evaluating the adaptability of these strategies to enhance recombinant protein yields in this red alga.

1.3.1 Selection markers

Stable transformation of an organism requires availability of a functional selection marker. The uptake efficiency of foreign genetic material is usually very low, and cells that acquired exogenous DNA need to be isolated from a large population. In general, there are two different selection approaches that are based either on antibiotic or herbicide resistances, or on metabolic markers that allow cell growth under specific nutritional conditions. A set of

selectable markers is desirable, because it provides the possibility to sequentially transform several constructs into the same line (supertransformation).

Since the first establishment of nuclear transformation in *C. reinhardtii* (Kindle et al., 1989), a wide range of antibiotic and auxotrophic markers have been developed (Doron et al., 2016). Examples for auxotrophic markers used in *C. reinhardtii* are *NIT1* and *ARG7*. Both genes allow prototrophic growth of otherwise auxotrophic strains that are not able to grow under specific nutritional conditions. *NIT1* encodes a nitrate reductase, enabling the transformed cells to use nitrate as nitrogen source, and *ARG7* encodes argininosuccinate lyase, enabling the transformed mutant cells to grow in the absence of arginine in the medium (Debuchy et al., 1989; Kindle et al., 1989). These systems, however, rely on the presence of specific auxotrophic strains that are not available in *P. purpureum*. In order to use auxotrophic markers for selection, *P. purpureum* needs to be screened for mutants susceptible to specific nutritional conditions. So far, the only mutants characterized in this red alga are mutants giving rise to herbicide resistance against diuron and sulfometuron methyl (Lapidot et al., 2002; Sivan and Arad, 1995).

A prerequisite for the use of antibiotic resistance markers is the susceptibility of the alga to the specific antibiotic. Li and Bock (2018) showed that *P. purpureum* is resistant to a wide range of commonly used antibiotics that are effective in *C. reinhardtii*, including hygromycin, kanamycin, paromomycin, spectinomycin and streptomycin. So far, only Zeocin™ and the herbicide norflurazon were successfully used as selection marker for nuclear transformation of *P. purpureum* (Li and Bock, 2018; Prasad et al., 2018). Zeocin™ is a glycopeptide antibiotic intercalating into DNA and inducing double-strand breaks that ultimately lead to cell death. This can be prevented by expressing the *ble* gene that encodes a protein that specifically binds Zeocin™ in a one-to-one stoichiometric ratio, thus inactivating it (Stevens et al., 1996).

Undoubtedly, further selectable markers for *P. purpureum* have to be established in order to extend the molecular toolbox for this organism.

1.3.2 Promoters

A key feature that dictates the expression level of a gene in any organisms are regulatory elements like promoters, terminators and 5’ and 3’ untranslated regions (UTR). For simplification purposes, the terms “promoter” and “terminator” used in this thesis also include 5´UTR and 3´UTR, respectively. Promoters act as binding sites for factors that recruit the RNA polymerase to initiate transcription. Promoters are generally the most decisive elements that dictate transcription rates. Generally, they are classified in constitutive promoters always driving the expression of a gene and inducible promoters that respond to certain biomolecules

or specific growth conditions (e.g. light, temperature) and/or developmental stages (Milito et al., 2023).

Several promoter have been reported to drive high constitutive expression of transgenes in *C. reinhardtii*, in particular the endogenous *PSAD* promoter and the hybrid *HSP70A-RBCS2* promoter (*AR* promoter) (Schroda et al., 2000; Fischer and Rochaix, 2001). These promoters have been employed for recombinant protein production in *C. reinhardtii*, with instances where the *PSAD* promoter exhibited higher activity (Kumar et al., 2013) and cases, where the *AR* promoter displayed superiority (López-Paz et al., 2017). New evidences showed that the expression strength is also dependent on the sequence of the coding region (CDS) of the transgene, as seen by the same promoter leading to different expression levels when used in different sequence contexts (Schroda, 2019). Even though the *PSAD* and the *HSP70A-RBCS2* promoters confer comparatively high expression of transgenes, industrial applications often require higher expression, and the search for stronger promoters lead to the construction of new synthetic promoters (Einhaus et al., 2021). This newly designed synthetic promoter AβSAP(i) was up to four times stronger than the previously used ones and is a new addition to the *C. reinhardtii* toolbox.

In *P. purpureum*, characterized promoters are scarce, with only the endogenous, constitutive *Actin* and *Tubulin* promoters and a heterologous promoter from *H. pulvialis* having been successfully employed. *Actin* and *Tubulin* are so-called “housekeeping” genes that are always active, due to their products being structural proteins of the cytoskeleton (Joshi et al., 2022). These promoters do not necessarily drive high transgene expression, but are a good starting point for constructing transformation plasmids. Another promoter that has been tested was the CaMV 35S, a viral promoter that reliably drives high expression in seed plants, but does not perform similarly well in *C. reinhardtii* (Amack and Antunes, 2020; Díaz-Santos et al., 2013). In *P. purpureum*, the CaMV35S promoter is also ineffective, suggesting missing transcription factors in microalgae that were acquired later in evolution by seed plants. Inducible promoters are desirable for the production of recombinant proteins that have a negative impact on cellular physiology, including toxic proteins, but no example of an inducible promoter is known in this red alga. Examples for inducible promoters in *C. reinhardtii* are the *CAH1* promoter, which is active under low CO_2 concentrations (Kucho et al.,1999), the *NIT1* promoter, that is activated upon nitrate addition (Ohresser et al., 1997), heat shock promoters like *HSP70A*, which are activated upon heat (Schroda et al., 2000) or the *METE* promoter, which is repressed in the presence of vitamin B12 in the medium and activated in its absence (Helliwell et al., 2014).

The characterization of new strong promoters in *P. purpureum* is an important step in the expansion of the molecular toolbox and for establishing this microalga as recombinant protein

production platform. One approach to find strong promoters is to search for highly expressed endogenous genes and use their regulatory elements to drive expression of reporter genes. However, as promoters can function differently in different sequence contexts, validation through testing with heterologous proteins is necessary to ensure their efficacy.

1.3.3 Codon usage and optimization

Protein synthesis relies on ribosomal translation of mRNAs codon by codon. Codons are triplets of nucleotides that dictate incorporation of specific amino acid. The standard genetic code comprises 61 codons coding for 20 amino acids, and three stop codons terminating translation. Some amino acids are encoded by multiple codons, which is commonly referred to as the redundancy of the genetic code. With the exception of methionine and tryptophan, amino acids are encoded by two (Phe, Tyr, His, Gln, Asn, Lys, Asp, Glu, and Cys), three (Ile), four (Val, Pro, Thr, Ala, and Gly) or six triplets (Leu, Ser, and Arg). Codon that specify the same amino acid exchange are called synonymous codons.

The codon usage or codon usage bias (CUB) reflects the preference of an organism for specific synonymous codons to encode a specific amino acid. Codon usage varies between different organisms, as well as within the genome of a single organism and impacts protein expression levels, translation efficiency and mRNA stability. Historically, two theories explain CUB. The first relates genomic GC content to codon usage, as the third codon position will predominantly be occupied by either GC or AT. Consequently, the CUB is simply a result of mutational processes favoring either GC or AT (Chen et al., 2004). The second idea attributes CUB variations within genes of an organism to codons that enhance translation via positive correlation with highly abundant isoacceptor tRNAs (Kanaya et al., 2001). The synonymous codons are recognized by the ribosomes, which show a non-uniform decoding rate for the codons. This rate is dependent on the availability of the corresponding tRNA. Highly abundant proteins possess optimal codon adaptation, promoting translation and protein production, while low abundance proteins have suboptimal adaptation (Novoa and Ribas de Pouplana, 2012; Iriarte et al., 2021). The suboptimal synonymous codons slow down the translation process, which may be critical for the formation of the native structure of a protein (Saunders and Deane, 2010).

This suggests that mRNAs exhibiting optimal codon usage are translated more rapidly, resulting in increased protein production, as demonstrated initially in *E. coli* (Sørensen et al., 1989). The CUB also impacts mRNA stability, with lowly translated mRNAs undergoing quicker degradation compared to highly translated ones. Additionally, the mRNA secondary structure is affected by the different synonymous codons, resulting in different mRNA stability

(Bae and Coller, 2022; Hanson and Coller, 2018). Optimal codon usage enhances mRNA stability, leading to elevated protein accumulation.

This mechanism of the codon usage bias can be harnessed for recombinant protein expression, highlighting the significance of codon optimization in construct design. The coding sequence of the gene of interest is *in-silico* optimized based on the host organism's CUB. This has been demonstrated in *C. reinhardtii* where *YFP* gene variants with varying nuclear codon adaptation were constructed, and it was shown that, with increased relative codon adaptation (RCA), the accumulation of mRNA and protein increased proportionally (Barahimipour et al., 2015).

The GC content is the proportion of guanine and cytosine bases among the four possible bases. Variability in GC content occurs across various genomes and organelles throughout all species. The overall genomic GC content, representing the percentage of G and C within an organism's entire genome, being the most common definition of the GC content. Another definition of the GC content acts on the level of coding regions and lastly, one can determine the GC content of the third position of the codons. For the CUB, the GC content of the third position is the most relevant, as it directly reflects the bias of the species for a specific codon. In *Chlamydomonas*, the GC content is approximately 64%, with a strong CUB on guanine or cytosine at the third codon position (Barahimipour et al., 2015; Merchant et al., 2007).

Since the genome of *P. purpureum* was only recently sequenced, and RNAseq data are sparse, the CUB of this organism is not available in common codon usage databases. Determining the CUB is essential for optimizing the coding sequence of the gene of interest. To assess the codon usage and compare its effect with *C. reinhardtii*, the RCA was chosen that outperforms the previously used codon adaptation index (CAI), as it directly into account the effects of mutational bias on codon usage (Fox and Erill, 2010). Whether *P. purpureum* responds similarly to codon usage as *C. reinhardtii* has been the subject of this study.

1.3.4 Introns

Introns are non-coding sequences found within genes and precursor mRNAs. These sequences are removed during mRNA maturation through splicing, leaving only the coding exons. Splicing enables diverse protein production from a single gene via alternative splicing, where various exon combinations are included in or excluded from the mature mRNA, expanding potential protein diversity (Lee and Rio, 2015). Introns also serve as regulatory elements by hosting enhancers or transcription factor binding sites, modulating chromatin structure, enhancing mRNA translation efficiency, or increasing mRNA stability (Rose, 2019; Shaul, 2017).

The positive impact of introns on transgene expression in *C. reinhardtii* was demonstrated by systematic insertion of the first intron of the *RBCS2* gene (rbcs2i1) at different positions within

a reporter gene (Baier et al., 2018). Optimal results were achieved by inserting the intron every ~500 bp from the start codon of the gene. The enhanced expression of the transgene can be attributed to a synergistic effect involving the recruitment of the transcriptional machinery and the adoption of the short average exon lengths naturally occurring in *C.* genome. Around 88% of genes in this green alga contain introns with an average of 8.3 exons per gene and an average exon length of only 240 bp (Merchant et al., 2007).

In contrast, only 2.8% of the genes in *P. purpureum* contain one or maximum two introns. Assessing whether introns can serve as expression-enhancing elements, as observed in C. reinhardtii, presents challenges in *P. purpureum*. Despite possessing a functional spliceosome, *P. purpureum* retains approximately half of its small number of introns (Wong et al., 2021). It remains to be investigated whether this retention also applies to introns inserted into transgenes in *P. purpureum* and whether the addition of introns would enhance transgene expression. Given the small number of introns in the genome of this red alga, it was assumed that additional introns would not confer benefits to transgene expression, and therefore, the effect of introns was not investigated in this study.

1.3.5 Protein targeting signals and secretion

Eukaryotic organisms require precise targeting of proteins to specific subcellular compartments. During the evolution of the eukaryotic cell, the increasing compartmentalization gave rise to metabolic pathways localized in specialized compartments that needed to be provided with their respective proteins to function (Sommer and Schleiff, 2014).

Protein translocation is achieved through short amino acid sequences known as signal peptides (SP). They possess specific biophysical properties that guide proteins to their respective compartments (Blobel and Sabatini, 1971; von Heijne, 1983). This occurs through two primary mechanisms: Firstly, co-translational transport, where proteins are directed across the ER membrane during ribosomal protein synthesis. Signal recognition particles (SRP) recognize the SP of co-translationally imported proteins as they emerge from the ribosome. The SRP guides the complex to the SRP receptor on the ER membrane, allowing protein synthesis directly into the ER (Nyathi et al., 2013). The second mechanism involves post-translational translocation of pre-synthesized proteins using translocon complexes that contain pores or via vesicles. This mode is primarily utilized for tail-anchored proteins located within the ER membrane and for importing proteins into mitochondria, plastids and the nucleus (Johnson et al., 2013).

For biotechnological applications, the translocation of recombinant proteins into the ER is often crucial due to the posttranslational modifications taking place in this compartment like

chaperone mediated folding, glycosylation and disulfide bond formation. Additionally, proteins destined for secretion must traverse the ER and Golgi apparatus before release into the extracellular matrix. Protein secretion offers numerous advantages for biotechnology, including simplified protein purification without the need for cell lysis, leading to reduced downstream processing costs that can amount to up to 80% of the production cost for biopharmaceuticals (Rosales-Mendoza et al., 2020).

Recombinant proteins are targeted for secretion across the outer membrane if they carry signal peptides, short amino acid sequences at the N-terminus of the protein. These signal peptides consist of three key regions: the N-region, a brief segment of 1 to five positively charged amino acids; the H-region, composed of hydrophobic amino acids forming a single alpha helix that is recognized by the SRPs; and the C-region, recognized and cleaved by signal peptidases, usually immediately following ER import (Owji et al., 2018). The mature, translocated protein is further processed in the ER. If the protein bears the C-terminal ER-targeting motif HDEL (histidine, aspartic acid, glutamic acid, and leucine), it remains within the ER. This motif is recognized by receptors in the cis-Golgi apparatus, facilitating the protein's return to the ER. In the absence of this motif, the protein proceeds through the Golgi apparatus, where it is packaged into secretory vesicles, ultimately leading to protein secretion.

Numerous proteins have been effectively secreted into the culture medium of *C. reinhardtii*, including mCherry, venus YFP, luciferase, and human growth factor (Lauersen et al., 2013; Ramos-Martinez et al., 2017; Baier et al., 2018). The effectiveness of secretion varied considerably across the studies. For instance, mCherry was secreted at a concentration of 15 mg/l, while human growth factor was secreted at only 56 µg/l. This was partially due to different signal peptides employed in those studies. It was demonstrated that different signal peptides lead to varying secretion efficacies in mammalian cell lines, but also in *C. reinhardtii* (Kober et al., 2013; Lauersen et al., 2013). A comprehensive investigation assessed the effectiveness of numerous endogenous and exogenous signal peptides for mCherry secretion, including those from *binding immunoglobulin protein 1* (*BiP1*), an ER resident chaperone and *carbonic anhydrase 1* (*CAH1*), a secreted protein. This analysis revealed an up to 10-fold difference in secretion efficiency, depending on the chosen signal peptide (Molino et al., 2018). Another significant advancement in recombinant protein secretion in *C. reinhardtii* was achieved through the C-terminal addition of the SP20 synthetic glycomodule. This module, consisting of a series of 20 serine and proline tandems (SP), led to secretion enhancements of up to 15-fold (Ramos-Martinez et al., 2017).

In *P. purpureum*, only little is known about the targeting and secretion of proteins. In the study of Li and Bock (2018), GFP was targeted into the chloroplast with the endogenous ATPC

transit peptide. To establish this red microalga as a competitive protein production platform, successful secretion has to be demonstrated. A possible hindrance could be the thick shell of polysaccharides, possibly preventing proteins to fully secrete into the medium. The efficiency of recombinant protein secretion in *P. purpureum* was investigated in this study, by employing an endogenous signal peptide to demonstrate secretion of a reporter protein and a therapeutic protein.

1.3.6 Availability of algal strains and episomal plasmid maintenance

A wide strain collection of *C. reinhardtii* is available, including cell wall-less strains, different insertional mutants (CLIP library), and strains engineered for enhanced transgene expression. Among these, the UV-mutagenized cell wall-less strain UVM11 has emerged as a prominent choice for recombinant protein expression, due to its superior transgene expression levels compared to other strains (Neupert et al., 2009; Li et al., 2016). The UVM11 strain was shown to be a sirtuin-type histone deacetylase (*srta*) mutant. The *SrtA* gene encodes a sirtuin-type histone deacetylase that recognizes naked transgenic DNA and then quickly inactivates its expression by the assembly of a repressive chromatin structure composed of deacetylated histones (Neupert et al., 2020). This mutant strain overcomes gene silencing, enabling stable and elevated transgene expression.

In contrast, available strains of *P. purpureum* lack comprehensive descriptions, and the genome was sequenced only for strain CCMP 1328, which differs from the strain employed in this thesis (SAG 1380-1d). There is currently no consensus on the use of a universal strain for specific applications in *P. purpureum*. It is known that about 25% of red algal genera contain more than two episomal plasmid types per species (Goff and Coleman, 1990; Lee et al., 2016). Until now, it is not known how the transformation plasmid replicates in *P. purpureum* and how the copy number is regulated. The attainable levels of recombinant protein accumulation seem to be linked to plasmid copy number (Li and Bock, 2018). Unraveling the plasmid maintenance mechanisms could provide strategies to enhance recombinant protein production by optimizing plasmid copy numbers.

Strain choice can also impact growth characteristics. For instance, *P. purpureum* strain CoE1 has been reported to show mixotrophic growth capacity, which remains uncertain for the strain utilized in this study. Exploiting the potential for mixotrophic growth could streamline cultivation practices and accelerate research. Investigating the strain's capability for heterotrophic and mixotrophic growth has been a focal point of this thesis.

1.4 Expression of therapeutic proteins as proof of concept for the utility of *P. purpureum* as a host for recombinant protein production

Therapeutic proteins represent a primary focus for recombinant protein production companies due to their significant cost and profit potential. Typically, these proteins are manufactured in mammalian cells because of the need for post-translational modifications (PTMs). However, this production platform is expensive and vulnerable to contamination risks, raising biosafety concerns. As outlined earlier, microalgae offer an economical and safe alternative. While *Chlamydomonas reinhardtii* is the most common microalga employed for recombinant protein production (Rosales-Mendoza et al., 2020), it suffers from drawbacks such as low transgene expression and frequent gene silencing. Despite years of dedicated research, protein yields remain suboptimal, necessitating consideration of alternative algae for efficient recombinant protein production. *P. purpureum*, due to its remarkable protein accumulation capacity even in non-optimized conditions, emerges as a potential solution.

To assess the potential of using this red microalga for expressing complex therapeutic proteins, two antigens derived from the hepatitis B (HBV) and hepatitis C (HCV) viruses were selected in collaboration with Dr. Norica Nichita from the Institute of Biochemistry of the Romanian Academy. These antigens hold promise as candidates for a subunit vaccine against the respective viruses, and were chosen to demonstrate the capacity of *P. purpureum* to yield complex therapeutic proteins at high levels.

1.4.1 The hepatitis B and hepatitis C virus as global disease agents

Viral diseases pose a significant global public health threat, with hepatitis B virus (HBV) and hepatitis C virus (HCV) being among the most prevalent worldwide. An estimated 296 million individuals suffer from chronic HBV infection and 1.5 million are newly infected each year, resulting in an estimated 820,000 deaths in 2019 (WHO, HBV fact sheet, 2023). Similarly, about 58 million people live with chronic HCV infection, with 1.5 million new infections each year, resulting in an estimated 290,000 deaths in 2019 (WHO, HCV fact sheet, 2023). Together, these diseases contribute to 1.1 million deaths annually, and impose substantial financial burdens on healthcare systems to manage chronic infections.

Transmission of both viruses occurs through exposure to infected blood, unsafe sexual practices, unscreened blood transfusions, injection drug use or vertical transmission from mother to child. Although vaccination can prevent HBV infection, no cure for acute or chronic cases currently exists. Current treatments can suppress the virus, but require lifelong administration to prevent relapse. HCV lacks an effective vaccine, with direct-acting antiviral medicines (DAAs) being the primary treatment option. DAAs can cure up to 95% of HCV

infections, but their accessibility remains limited, particularly in low-income countries (Das and Pandya, 2018).

The hepatitis B virus is a partially double-stranded, enveloped DNA virus with a genome of approximately 3.2 kb and ten known genotypes (Liang, 2009). The genome encodes four highly overlapping ORFs, expressing seven different proteins, which are the HBV DNA polymerase, the hepatitis B core antigen (HBcAg), the hepatitis B envelope antigen (HBeAg), the HBx protein and the three hepatitis B surface antigens (HBsAg). The DNA is surrounded by the nucleocapsid formed by the core antigen, and the envelope is formed by three surface proteins called the large (L), middle (M) and small (S) protein that are encoded by the same ORF. These surface proteins vary in their N-terminus as protein synthesis is initiated at three different translation start sites that are in frame. The so-called Dane particles are the virulent particles causing an infection of liver cells in the human body (Datta et al., 2012). At the same time, small spherical particles, called sub-viral particles (SVPs), only consisting of the three envelope proteins and host lipids, are produced in up to 100,000-fold excess and released into the blood of infected people. The sheer excess of empty SVPs sequesters neutralizing antibodies of the host and thus promotes infection via the Dane particle (Hu and Liu, 2017).

The hepatitis C virus is an enveloped, positive-sense single-stranded RNA virus with a genome of 9.6 kb with a single ORF encoding a single polyprotein. Seven genotypes have been identified that differ at 30-25% of the nucleotide sites (Petruzziello et al., 2016). The polyprotein is cleaved by cellular proteases into 10 smaller proteins: the structural core protein and envelope proteins E1 and E2, and seven nonstructural proteins with various functions (Dubuisson and Fabiani, 2007). The E1 and E2 surface proteins form heterodimers and are highly glycosylated at their ectodomain with six and eleven glycosylation sites, respectively. The heterodimer is essential for cell entry and assembly of viral particles, thus making the two proteins promising candidates for a novel HCV subunit vaccine (Gomez-Escobar et al., 2023).

1.4.2 A chimeric HBV S/preS1 envelope protein as novel vaccine candidate

The majority of current HBV subunit vaccines are based on the small (S) envelope protein of HBV and are produced through yeast-based expression systems. A HBV vaccine was the first human subunit vaccine produced by recombinant DNA technique and approved by the FDA in 1986 (Valenzuela *et al.*, 1982; reviewed in Bucci, 2020). The recombinant S protein undergoes self-assembly into 22 nm spherical sub-viral particles (SVPs) after yeast cell lysis. The availability of genome-free SVPs, which were found to be 1,000 times more immunogenic than S protein dimers, represented a significant advancement in HBV vaccination strategies. This development laid the foundation for effective HBV vaccination by utilizing these highly immunogenic SVPs (Hu and Liu, 2017). The recombinant S vaccine ensures sufficient levels

of protective antibodies in more than 90% of healthy adult recipients, leaving approximately 10% of the recipients unprotected non-responders to the vaccine (Rubin et al., 2014). Another challenge associated with the current HBV subunit vaccine is the ongoing debate regarding the duration of immunization. Over time, the levels of HBs antibodies, which are crucial for protection against HBV infection, tend to decrease. This decline in antibody titers raises concerns about the potential waning of immunity, leaving individuals susceptible to infection decades after initial vaccination (Sahana et al., 2017).

The development of the initial HBV vaccines occurred prior to the comprehensive understanding of the HBV infection cycle. The preS1 domain, a part of the N-terminus of the surface L protein of HBV (Figure 1), plays an essential role in infectivity by facilitating the interaction between the virus and host cells (Herrscher et al., 2020). It was shown that the first 47 amino acids of the preS1 domain are essential for the binding to the cell surface receptor sodium taurocholate cotransporting polypeptide (NTCP), facilitating virus entry (Yan et al., 2012). The variability of the preS1 region is much lower than that of the S region and no escape mutants are known (Gerlich, 2017). The next generation of vaccines tackled the above-mentioned problems by the inclusion of all three (S, M and L) envelope proteins into the SVPs, resulting in enhanced immunogenicity even in individuals who did not respond well to the S protein vaccine (Shouval et al., 2015; Vesikari et al., 2021). However, the production of these vaccines is complex, particularly due to the need to maintain a specific L/M/S ratio, and they have primarily been generated using expensive mammalian systems (Dobrica et al., 2017).

To overcome limitations, a chimeric S/preS1 antigen was designed by our collaboration partners in Norica Nichita's lab. This chimeric antigen incorporates the 16–42 amino acid sequence of the preS1 domain from the L protein, a region crucial for NTCP binding, into the antigenic loop (AGL) of the S protein at amino acid position 127 (Figure 1), resulting in the chimeric S/preS1$^{16\text{-}42}$ protein (Pantazica et al., 2022). An earlier version of this chimeric protein, the S/preS1$^{21\text{-}47}$ protein had been produced in mammalian and plant cells, and showed superior immune response when compared to the S protein, but exhibited poor expression and secretion properties (Dobrica et al., 2017). The new version of the chimeric antigen, incorporating the 16-42 amino acid sequence, has been explored in this work for production in *P. purpureum*, aiming to demonstrate the potential of this red microalga for expressing biotechnologically valuable proteins.

The S protein, including the chimeric S/preS1 protein, is a membrane protein with four transmembrane domains. The AGL is positioned in the ER lumenal region during synthesis and maturation, and it later serves as a major immunogenic site targeted by the host's immune

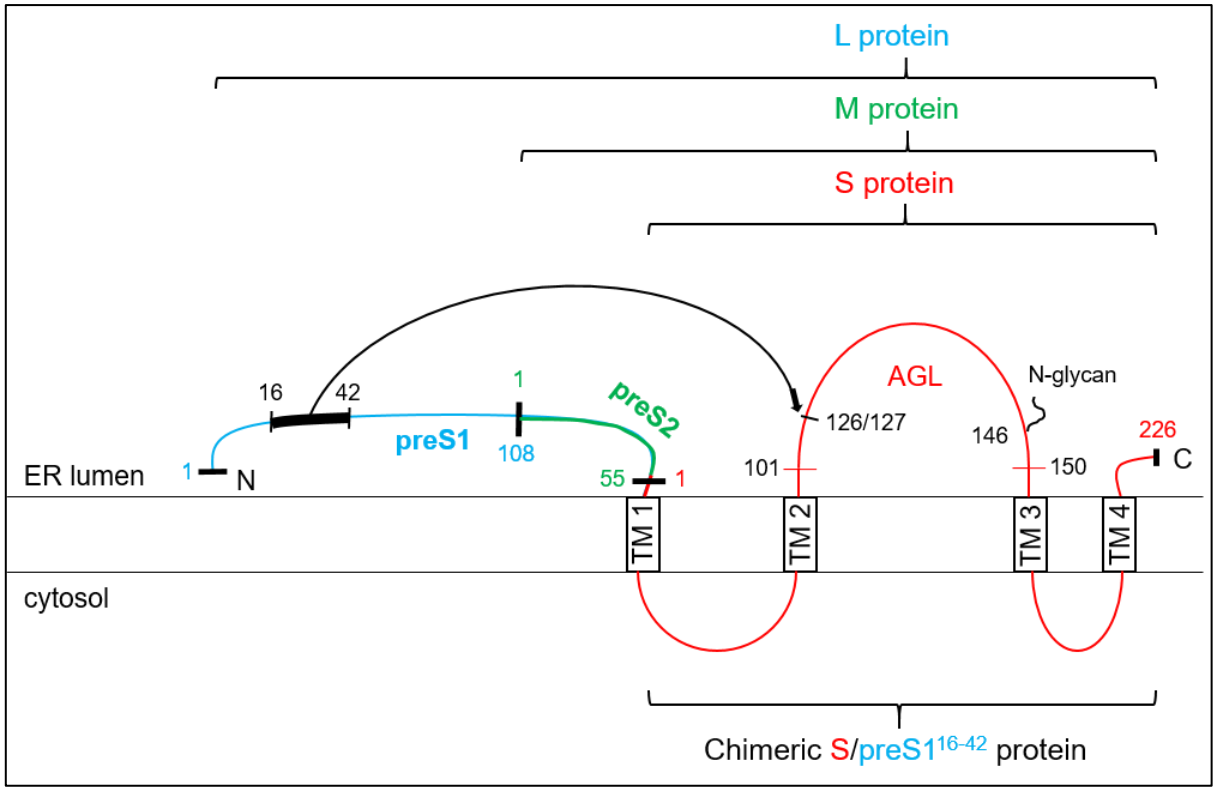


Figure 1: Schematic representation of the HBV envelope antigens and the new chimeric S/preS1 antigen
The HBV envelope antigens are expressed from a single ORF by using different translation start sites resulting in the small (S) protein with 226 amino acids (in red), the medium (M) protein with the additional preS2 domain consisting of 55 amino acids (in green) and the large (L) protein with the additional preS1 domain of 108 amino acids (in blue). Amino acids 16-42 of the preS1 domain were inserted into the antigenic loop (AGL) of the S protein between amino acid 126 and 127, resulting in the chimeric S/preS1$^{16-42}$ protein that was used in this thesis. Numbers indicate amino acid positions of the respective protein. TM: transmembrane domain

response. The N-glycosylation site is located at Asn-146 in the AGL and is only glycosylated in about half of the proteins in a natural infection (Seitz et al., 2020). In yeast-produced S proteins, this glycosylation site remains non-glycosylated. The N-linked glycan of the S domain plays a role during the HBV interaction with the host immune system, but is not essential for success of vaccination (Dobrica et al., 2020). Shortly after co-translational translocation and synthesis, the S proteins form dimers linked by disulfide bridges. The SVPs, consisting of up to 50 S protein dimers and host lipids, are formed at the ER-Golgi intermediate compartment (ERGIC), and the formation is dependent on the right confirmation of the AGL, the second transmembrane domain and the first cytosolic loop (Patient et al., 2007; Suffner et al., 2018). The final secretion of spherical HBV SVPs is not yet fully understood, but seems to follow the constitutive secretory pathway of the host cell for release (Zhao et al., 2021). In yeast cells producing the HBsAg S vaccine, this transport is impaired and the dimers are retained in the ER membrane. Only after cell lysis, the dimers self-assemble into SVPs that are then used for vaccination (Ho et al., 2020).

Successful expression of the chimeric S/preS1 protein in *P. purpureum* would serve as a demonstration of the red alga's capability to effectively express proteins with intricate topologies. Additionally, the investigation aimed to determine the glycosylation state of the protein and its potential to form SVPs, which could further enhance immunogenicity.

1.4.3 The HCV envelope protein E2$^{\Delta HVR1}$ as novel vaccine candidate

Current treatment of HCV infections is based on DAAs that lead to elimination of HCV in approximately 95% of the cases. However, the risk of liver carcinoma post-treatment remains significant. Additionally, resistance to DAAs has been documented, and these medications do not offer protection against reinfection (Gomez-Escobar et al., 2023). Thus, the development of a vaccine would be a big step towards the goal of HCV eradication.

The development of such a vaccine is complicated by the characteristics of HCV as an RNA virus with a high mutation rate. Despite extensive research over the last decades, no HCV vaccine has been approved for human use. Genetic diversity across various HCV genotypes poses an obstacle to developing a universal vaccine against the diverse range of HCV infections. A strong humoral immune response inducing neutralizing antibodies (nAbs) in early infection can lead to clearance and inhibit the development of a chronic infection, and therefore, is the main goal in HCV vaccine development (Walker et al., 2019). Most of the research has been directed towards the structural protein dimer of the HCV envelope, the E1/E2 dimer located at the surface. These proteins are highly glycosylated, disulfide bond-stabilized dimers and anchored to the ER-derived membrane by the TMD anchor of the E2 protein (Dubuisson and Fabiani, 2007).

The E2 protein is involved in cell entry by binding different receptors including CD81. It is the target of most nAbs described in the literature (Pileri et al., 1998; Gomez-Escobar et al., 2023). The E2 protein has a globular ectodomain spanning into the ER lumen with a C-terminal stem region including the anchoring TMD. E2 has 11 potential glycosylation sites that are all occupied by N-glycans during viral maturation in the infected host. Depending on the position, the glycans are involved in modulation of the immune response, receptor binding, cell wall penetration and protein folding (Orlova et al., 2015). Three variable regions are located in the ectodomain that highly differ in amino acid composition between genotypes and help the virus to escape the host’s immune response (Prentoe et al., 2016). The hypervariable region 1 (HVR1) is a sequence located at the N-terminus of E2, which can change rapidly within the same infection in a host due to antibody-driven antigenic drift (Campo et al., 2012). It was demonstrated that HVR1, which itself is an epitope for nAbs, shields other antibody-specific epitopes on E1, E2 or E1/E2, that are more conserved and thus more suited for HCV clearance

(Prentoe and Bukh, 2018). Unlike the surface antigens of HBV, E2 by itself has not shown the capability to form SVPs in any utilized expression system thus far (Dobrica et al., 2021).

In a novel vaccine development approach, our partners in Norica Nichita's lab engineered a truncated form of E2, termed $E2^{\Delta HVR1}$. This version comprises the E2 ectodomain while omitting HVR1, ensuring unobstructed epitope accessibility during immunization. Furthermore, the soluble ectodomain could potentially be secreted into the medium, thus simplifying purification and further reducing production costs.

Successful expression of the $E2^{\Delta HVR1}$ protein in *P. purpureum* would underscore the alga's ability to produce complex, glycosylated therapeutic proteins. Effective secretion would highlight the potential for decreased production costs through algal expression. An interesting question to solve relates to the glycosylation state of this antigen, because very little is known about N-glycosylation in red algae.

1.5 Aim of the project

The objective of this work was to establish *P. purpureum* as a new producer organism for recombinant proteins in the microalgal field. So far, this field has been dominated by *C. reinhardtii*, with some notable exceptions of diatoms and other green algae. *P. purpureum* would be the first red microalga used for recombinant protein production.

To this end, new molecular tools need to be established to allow effective transgene expression in *P. purpureum*. The main determinants of high transgene accumulation were sought to be identified using reporter constructs, focusing on codon usage and promoters. We also tested targeting of foreign proteins to the secretory pathway to allow secretion into the culture medium, which would open up new possibilities for algal biotechnology. Furthermore, we wanted to explore the possibility to produce biotechnologically valuable subunit vaccines as proof of principle that *P. purpureum* is able to express complex proteinaceous biopharmaceuticals. These antigens have been characterized with regard to their glycosylation and yield. The ultimate goal is to test if they are able to elicit a specific and non-toxic immune response in mice.

2 Material

2.1 Common chemicals and consumables

All common chemicals and consumables have been purchased from the following companies:

ABgene (Epsom, United Kingdom), Biozym Scientific (Oldendorf, Germany), Brand (Gießen, Germany), Carl Roth GmbH (Karlsruhe, Germany), Duchefa (Haarlem, Netherlands), Eppendorf (Hamburg, Germany), Fisher Scientific (Schwerte, Germany), Hartenstein (Würzburg, Germany), Invitrogen (Karlsruhe, Germany), Merck KGaA (Darmstadt, Germany), Promega (Mannheim, Germany), Roche (Mannheim, Germany), Sarstedt (Nümbrecht, Germany), Serva (Heidelberg, Germany), Sigma-Aldrich (Steinheim, Germany), Unigloves GmbH (Troisdorf, Germany), VWR International (Darmstadt, Germany), Zymo Research (Freiburg, Germany).

2.2 Special chemicals

[α-32P]dCTP	Hartman Analytic, Braunschweig, Germany
Agarose	Biozym Scientific, Oldendorf, Germany
Ampicillin sodium salt	Carl Roth GmbH, Karlsruhe, Germany
Bovine serum albumin (BSA)	Carl Roth GmbH, Karlsruhe, Germany
Cetyltrimethylammoniumbromide (CTAB)	Carl Roth GmbH, Karlsruhe, Germany
Chelex 100	Bio-Rad, München, USA
Chloramphenicol	Duchefa, Haarlem, Netherlands
Chloroform	Carl Roth GmbH, Karlsruhe, Germany
Complete, protease inhibitor cocktail	Roche, Mannheim, Germany
Coomassie Brilliant Blue G-250	Serva, Heidelberg, Germany
Dithiothreitol (DTT)	Carl Roth GmbH, Karlsruhe, Germany
dNTPs	Promega, Mannheim, Germany
EDTA	Carl Roth GmbH, Karlsruhe, Germany
Ethidium bromide	Carl Roth GmbH, Karlsruhe, Germany
Formaldehyde (37%)	Sigma-Aldrich, Steinheim, Germany
Formamide	Merck KGaA, Darmstadt, Germany
IPTG	Carl Roth GmbH, Karlsruhe, Germany
Isopropanol	Merck KGaA, Darmstadt, Germany
Kanamycin monosulphate	Duchefa, Haarlem, Netherlands
Methanol	Merck KGaA, Darmstadt, Germany
Micro agar	Duchefa, Haarlem, Netherlands

Ni-NTA agarose	Qiagen GmbH, Hilden, Germany
Paromomycin sulphate	Duchefa, Haarlem, Netherlands
Phenol	Carl Roth GmbH, Karlsruhe, Germany
Phenol/Chloroform/Isoamyl alcohol (25:24:1)	Carl Roth GmbH, Karlsruhe, Germany
Rotiphorese® Gel A/B	Carl Roth GmbH, Karlsruhe, Germany
Select-Agar	Invitrogen, Karlsruhe, Germany
Sephadex G-50 Medium	Sigma-Aldrich, Steinheim, Germany
Spectinomycin	Duchefa, Haarlem, Netherlands
ß-Mercaptoethanol	Sigma-Aldrich, Steinheim, Germany
Streptomycin	Duchefa, Haarlem, Netherlands
TEMED	Carl Roth, Karlsruhe, Germany
Tunicamycin	Cayman Chemicals, Ann Arbor, MI, USA
TRIzol®	Thermo Fisher Scientific, Schwerte, Germany
Zeocin™	Invivogen, Toulouse, France

2.3 Consumables

0.5-10 kb RNA Ladder (1 µg/µl)	Invitrogen, Karlsruhe, Germany
Ni-NTA slurry	Qiagen, Hilden, Germany
Amicon® Ultra Centrifugal Filters	Merck KGaA, Darmstadt, Germany
Gene Ruler™, 1 kb DNA Ladder	MBI Fermentas, St. Leon-Rot, Germany
Nitrocellulose membrane Protran 0.45 µm	GE Healthcare, Buckinghamshire, UK
Nylon membrane Hybond-N	GE Healthcare, Buckinghamshire, UK
PageRuler Plus Prestained Protein Ladder	Thermo Fisher Scientific, Schwerte , Germany
Polypropylene Columns	Qiagen GmbH, Hilden, Germany

2.4 Antibodies

3xHA antibody (A01244)	GenScript, Rijswijk, Netherlands
GFP antibody (632381)	Takara, Saint-Germain-en-Laye, France
HBV-antibody (NB100-62652)	Novus biologicals, Centennial, USA
His-tag antibody (MA1-21315)	Invitrogen, Karlsruhe, Germany
Secondary anti mouse HRP (AS111772)	Agrisera Vännäs, Sweden
Secondary anti rabbit HRP (AS09602)	Agrisera Vännäs, Sweden
Secondary anti rabbit HRP (170-6515)	Bio-Rad, München,Germany

2.5 Equipment

Agarose gel electrophoresis system	PeqLab, Erlangen, Germany

Cell counters, Z2 Coulter Particle	Beckman Coulter, Krefeld, Germany
Centrifuge 5415R	Eppendorf, Hamburg, Germany
Centrifuge AllegraTM 25R	Beckman Coulter, Krefeld, Germany
Centrifuge MiniSpin	Eppendorf, Hamburg, Germany
Centrifuge Sorvall® RC-6	Thermo Fisher Scientific, Schwerte, Germany
Concentrator 5301	Eppendorf, Hamburg, Germany
Confocal laser scanning microscope TCS SP5	Leica, Wetzlar, Germany
FPLC, ÄKTAexplorer 100 Air	GE Healthcare, Buckinghamshire, UK
G:Box Chemi XT4 detector	Syngene, Cambridge, UK
Grinding mill, MM301	Retsch, Haan, DE
HiLoad 16/600 Superdex 75 pg column	Merck KGaA, Darmstadt, Germany
Hybaid mini oven	Pegasus Scientific, Rockville, USA
PCR thermocycler Mastercycler EPGradient	Eppendorf, Hamburg, Germany
Percival	Percival Scientific, Perry, USA
PerfectBlueTM „Semi-Dry“-Electroblotter	PeqLab, Erlangen, Germany
Polyacrylamide gel electrophoresis system	Bio-Rad, München, Germany
Power supply PowerPacTM Basic	Bio-Rad, München, Germany
Radioisotope image analyser TyphoonTM Trio+ scanner	GE Healthcare, Buckinghamshire, UK
SevenEasy pH meter	Mettler-Toledo, Zürich, Switzerland
Shaker, Multitron Standard	Infors HT, Bottmingen, Switzerland
Slot blot device SHM-48	Scie-Plas, Cambridge, UK
Sonifier®, W-250 D	G. Heinemann, Lorch, Germany
Spectrophotometer NanoDrop 1000	PeqLab, Erlangen, Germany
Stirrer IKA® RH basic KT/C	IKA® Werke, Staufen, Germany
Storage Phosphor Screens	GE Healthcare, Buckinghamshire, UK
Thermomixer comfort	Eppendorf, Hamburg, Germany
Ultrospec® 3100 pro UV/Visible Spectrophotometer	GE Healthcare, Buckinghamshire, UK
UV-crosslinker BLX-254	Vilber Lourmat, Marne La Vallee, France
UV documentation Quantum CX 5	Vilber Lourmat, Marne La Vallee, France
Vortex-Genie 2	Scientific Industries, New York, USA

2.6 Restriction enzymes

All restriction enzymes used in this study were purchased from New England Biolabs (NEB, Frankfurt a.M., Germany) and Thermo Fisher Scientific Inc. (Schwerte, Germany).

2.7 Enzymes

ABsolute SYBR Green ROX	Thermo Fisher, Schwerte, Germany
DreamTaq DNA Polymerase	Thermo Fisher, Schwerte, Germany
Gibson Assembly® Master Mix	New England Biolabs, Frankfurt, Germany
In-Fusion® Snap Assembly Master Mix	TaKaRa Bio, Saint-Germain-en-Laye, France
Lysozyme	Roth, Karlsruhe, Germany
Phusion™ High-Fidelity DNA polymerase	Thermo Fisher, Schwerte, Germany
RNase A (DNase-free)	AppliChem, Darmstadt, Germany
T4 DNA Ligase	Promega, Mannheim, Germany

2.8 Kits

ECL-Prime™ Western Blotting Detection	GE Healthcare, Buckinghamshire, UK
MegaPrime™ DNA Labeling System	Roth, Karlsruhe, Germany
NucleoBond® Xtra Midi Plus	Macherey-Nagel, Düren, Germany
NucleoSpin® Gel and PCR clean-up kit	Macherey-Nagel, Düren, Germany
NucleoSpin® Plasmid kit	Macherey-Nagel, Düren, Germany
Pierce BCA protein assay kit	Thermo Fisher Scientific, Schwerte, Germany
Roti®-Quant	Roth, Karlsruhe, Germany

2.9 Software

Target P 2.0	DTU Health tech, Lyngby, Denmark
Galaxy platform, version 23.1.rc1	https://usegalaxy.org
CodonWorkbench	Mark Lohse, unpublished
DNAStar Lasergene Suite 17	DNASTAR, Madison, WI USA
ImageJ, 1.52n	Open source
ProtParam webtool	https://web.expasy.org/protparam/
BLASTp	NIH

2.10 Medium and solutions

All media, stock solutions and buffers used are listed.

Table 2: Buffer composition

Medium/Solution	Components	Concentrations
	***E. coli* cultivation**	

Medium/Solution	Components	Concentrations
LB medium	Tryptone	1% (w/v)
	NaCl	1% (w/v)
	Yeast extract	0.5% (w/v)
LB-Agar	LB Medium	1 L
	Agar	15 g
TB medium	HEPES pH 6.7	10 mM
	KCl	250 mM
	$MnCl_2$	50 mM
	$CaCl_2$	15 mM
SOB medium	Tryptone	2% (w/v)
	Yeast extract	0.5% (w/v)
	NaCl	10 mM
	KCl	2.5 mM
	$MgCl_2$	10 mM
	$MgSO_4$	10 mM
***Porphyridium* cultivation**		
Low salt ASW	NaCl	15.2 g/L
	$MgSO_4$	4.9 g/L
	$MgCl_2$	5.6 g/L
	$CaCl_2$	1.5 g/L
	KNO_3	2 g/L
	KH_2PO_4	0.07 g/L
	$NaHCO_3$	0.04 g/L
	Tris HCl pH 7.6	50 mM
	Trace metal solution	1 mL/L
	Chelated iron solution	1 mL/L
Trace metal solution	$ZnCl_2$	40 mg/L
	H_3BO_3	600 mg/L
	$CoCl_2$	15 mg/L
	$CuCl_2$	40 mg/L
	$MnCl_2$	400 mg/L
	$(NH_4)_6 Mo_7O_{24}$	370 mg/L
Chelated iron solution	Na_2EDTA	0.05 M
	$FeCl_3$	2.4 g/L
***Chlamydomonas* cultivation**		
4x Beijerinck´s Solution	NH_4Cl	16 g
	$CaCl_2*2H_2O$	2 g
4x Beijerinck´s Solution (cont.)	$MgSO_4*7H_2O$	4 g
	distilled water	up to 1 L
Trace solutions	Na_2*EDTA	25 mM
	Zn*EDTA	2.5 mM

Medium/Solution	Components	Concentrations
	$(NH_4)_6Mo_7O_{24}$	28.5 µM
	NA_2SeO_3	0.1 mM
	Mn*EDTA	6 mM
	Fe*EDTA	20 mM
	Cu*EDTA	2 mM
Tris-acetate-phosphate medium (TAP)	Tris	20 mM
	Acetic acid	17.5 mM
	Phosphate buffer	1 mM
	4x Bejierinck´s Solution	2.5% v/v
	Trace solutions	0.1% v/v
TAP-Agar	TAP	1 L
	Agar	15 g
Buffers for molecular biology		
Church buffer	BSA	1% (w/v)
	NA_2HPO_4 pH7.2	0.5 M
	SDS	7% (w/v)
	EDTA pH8.0	1 mM
CTAB buffer	CTAB	2% (w/v)
	NaCl	1.4 M
	EDTA pH 8	20 mM
	Tris-HCl pH 8	100 mM
	β-mercaptoethanol	2% (v/v)
10x MOPS buffer	MOPS 3-(N-morpholino) propanesulfonic acid	1 M
	NaAc	3 M
	EDTA	10 mM
10x SSC pH 7.0	NaCl	1.5 M
	Tri-sodium citrate dehydrate	0.15 M
1x TAE	Tris	40 mM
	Acetic acid	0.11 % (v/v)
	EDTA	1 mM
RNA wash buffer I	SSC pH7.0	1x
	SDS	0.2%
RNA wash buffer II	SSC pH7.0	0.5x
	SDS	0.2%
Protein work		
Ni-NTA Lysis buffer pH 8.0	NaH_2PO_4	50 mM
	NaCl	300 mM
	imidazole	10 mM
Washing buffer 1	Ni-NTA buffer	20 ml
	$MgSO_4$	20 mM
	ATP	4 mM

Medium/Solution	Components	Concentrations
Washing buffer 2	Ni-NTA buffer	20 ml
	Imidazole	20 mM
Washing buffer 3	Ni-NTA buffer	20 ml
	Imidazole	50 mM
Elution buffer	Ni-NTA buffer	20 ml
	Imidazole	500 mM
PBS	NaCl	137 mM
	KCl	2.7 mM
	Na_2HPO_4	10 mM
	KH_2PO_4	1.8 mM
TE	Tris HCl pH 8	10 mM
	EDTA	1 mM
Total soluble protein lysis buffer (TSP buffer)	Tris HCl pH 7.6	50 mM
	NaCl	100 mM
	KAc	10 mM
	MgAc	5 mM
	EDTA	1 mM
	Protease inhibitor cocktail	1 x
Total protein extraction buffer (TP buffer)	Sucrose	30% (w/v)
	SDS	2% (w/v)
	Tris HCl pH 8.0	100 mM
	β-mercaptoethanol	2% (v/v)
	Protease inhibitor cocktail	1 x
5 x SDS sample buffer	Tris HCl pH 6.8	0.25 M
	Glycerol	50% (v/v)
	SDS	10% (w/v)
	β-mercaptoethanol	10% (w/v)
	Bromophenol blue	0.15% (w/v)
SDS-PAA: separation gel 12%	Rotiphorese Gel A	7.8 mL
	Rotiphorese Gel B	3.1 mL
	1.5 M Tris-Hcl pH 8.8	5 mL
	20% SDS	0.1 mL
	APS	0.15 mL
	TEMED	0.015 mL
	destilled water	to 20 mL
SDS-PAA: stacking gel 3%	Rotiphorese Gel A	0.517 mL
	Rotiphorese Gel B	0.24 mL
	0.5 M Tris-Hcl pH 6.8	1 mL
	20% SDS	0.02 mL
	APS	0.04 mL
	TEMED	0.004 mL
	destilled water	to 4 mL

Medium/Solution	Components	Concentrations
TBS-T	Tris pH 7.6	20 mM
	NaCl	150 mM
	Tween 20	0.1% (v/v)
Protein resuspension buffer	Tis-HCl pH 7.6	50 mM
	NaCl	100 mM
	KAc	10 mM
	MgAc	5 mM
	SDS	0.5% (w/v)
	Protease inhibitor cocktail	1x
Blotting transfer buffer T3	Tris-HCl pH 10.4	300 mM
	Isopropanol	20% (v/v)
Blotting transfer buffer T2	Tris-HCl pH 9.8	25 mM
	Isopropanol	20% (v/v)
Blotting transfer buffer T1	Blotting transfer buffer T2	50 mL
	ε-Aminocaproic acid	225 mg

2.11 Plasmids

Plasmid	Vector backbone	Insert description	Cloning procedure	Primer forward	Primer reverse	Transformed organism
pASH1	pZL22 [1)]	PpYFP, fully codon optimized YFP variant for the *Porphyridium* nuclear genome	pZL22 was digested with XhoI and NheI to excise *GFP*. The insert was amplified using primer pair oASH19 and oASH12 from a plasmid carrying the synthetized codon optimized PpYFP. The fragments were ligated using Gibson assembly.	oASH19	oASH12	*P. purpureum*
pASH2	pZL22 [1)]	cpYFP, YFP codon optimized for the *Chlamydomonas* chloroplast	pZL22 was digested with XhoI and NheI to excise *GFP*. The insert was amplified using primer pair oASH15 and oASH4 from plasmid pRMB8 [2)]. The fragments were ligated using Gibson assembly.	oASH15	oASH4	*P. purpureum*
pASH3	pZL22 [1)]	venus YFP (vYFP)	pZL22 was digested with XhoI and NheI to excise *GFP*. The insert was amplified using primer pair oASH14 and oASH2 from plasmid pRMB11 [2)]. The fragments were ligated using Gibson assembly.	oASH14	oASH2	*P. purpureum*
pASH5	pZL22 [1)]	laYFP, lowly codon adapted YFP for the *Chlamydomonas* nuclear genome	pZL22 was digested with XhoI and NheI to excise *GFP*. The insert was amplified using primer pair oASH16 and oASH6 from plasmid pRMB13 [2)]. The fragments were ligated using Gibson assembly.	oASH16	oASH6	*P. purpureum*
pASH7	pZL22 [1)]	CrYFP, codon optimized for the *Chlamydomonas* nuclear genome	pZL22 was digested with XhoI and NheI to excise *GFP*. The insert was amplified using primer pair oASH14 and oASH2 from plasmid pRMB12 [2)]. The fragments were ligated using Gibson assembly.	oASH14	oASH2	*P. purpureum*
pASH8	pZL22 [1)]	insertion of a SpeI restriction site between the *Actin* promoter and CDS of *cat*	pZL22 was digested with XhoI and NheI to excise *GFP*. The insert was amplified using primer pair oASH13 and oASH8 from plasmid pRMB7 [3)]. The fragments were ligated using Gibson assembly	oASH13	oASH8	*P. purpureum*
pASH15	pASH8	3xHA-tagged HBV S/preS1, codon optimized for the *Porphyridium* nucleus, cytosolic	pASH8 was digested with SpeI and NheI to excise *cat*. The insert was amplified using oASH129 and oASH130 from a synthetic plasmid carrying the *Porphyridium* codon optimized S/preS1 protein. The fragments were ligated using Gibson assembly.	oASH129	oASH130	*P. purpureum*
pASH16	pASH8	3xHA-tagged HBV S/preS1, codon optimized for the *Porphyridium* nucleus, targeted for secretion	pASH8 was digested with SpeI and NheI to excise *cat*. The insert was amplified using oASH131 and oASH130 from a synthetic plasmid carrying the *Porphyridium* codon optimized S/preS1 protein. The fragments were ligated using Gibson assembly.	oASH131	oASH130	*P. purpureum*
pASH17	pASH8	3xHA-tagged HBV S/preS1, codon optimized for the *Porphyridium* nucleus, ER-targeted	pASH8 was digested with SpeI and NheI to excise *cat*. The insert was amplified using oASH131 and oASH132 from a synthetic plasmid carrying the *Porphyridium* codon optimized S/preS1 protein. The fragments were ligated using Gibson assembly.	oASH131	oASH132	*P. purpureum*

pASH18	pASH8	3xHA-tagged HCV E2$^{\Delta HVR1}$ protein, codon optimized for the *Porphyridium* nucleus, cytosolic	pASH8 was digested with SpeI and NheI to excise *cat*. The insert was amplified using oASH136 and oASH130 from a synthetic plasmid carrying the *Porphyridium* codon optimized E2$^{\Delta HVR1}$ protein. The fragments were ligated using Gibson assembly.	oASH136	oASH130	*P. purpureum*
pASH19	pASH8	3xHA-tagged HCV E2$^{\Delta HVR1}$ protein, codon optimized for the *Porphyridium* nucleus, targeted for secretion	pASH8 was digested with SpeI and NheI to excise *cat*. The insert was amplified using oASH131 and oASH130 from a synthetic plasmid carrying the *Porphyridium* codon optimized E2$^{\Delta HVR1}$ protein. The fragments were ligated using Gibson assembly.	oASH131	oASH130	*P. purpureum*
pASH20	pASH8	3xHA-tagged HCV E2$^{\Delta HVR1}$ protein, codon optimized for the *Porphyridium* nucleus, ER-targeted	pASH8 was digested with SpeI and NheI to excise *cat*. The insert was amplified using oASH131 and oASH132 from a synthetic plasmid carrying the *Porphyridium* codon optimized E2$^{\Delta HVR1}$ protein. The fragments were ligated using Gibson assembly.	oASH131	oASH132	*P. purpureum*
pASH23	pRMB19 [3]	3xHA-tagged HBV S/preS1, codon optimized for the *Chlamydomonas* nucleus, cytosolic	pRMB19 was digested with NdeI and EcoRI to excise p24. The insert was amplified using oASH172 and oASH175 from a synthetic plasmid carrying the *Chlamydomonas* codon optimized S/preS1 protein. The fragments were ligated using Gibson assembly.	oASH172	oASH175	*C. reinhardtii*
pASH24	pRMB19 [3]	3xHA-tagged HBV S/preS1, codon optimized for the *Chlamydomonas* nucleus, targeted for secretion	pRMB19 was digested with NdeI and EcoRI to excise p24. The insert was amplified using oASH173 and oASH175 from a synthetic plasmid carrying the *Chlamydomonas* codon optimized S/preS1 protein. The fragments were ligated using Gibson assembly.	oASH173	oASH175	*C. reinhardtii*
pASH25	pRMB19 [3]	3xHA-tagged HBV S/preS1, codon optimized for the *Chlamydomonas* nucleus, ER-targeted	pRMB19 was digested with NdeI and EcoRI to excise p24. The insert was amplified using oASH173 and oASH176 from a synthetic plasmid carrying the *Chlamydomonas* codon optimized S/preS1 protein. The fragments were ligated using Gibson assembly.	oASH173	oASH176	*C. reinhardtii*
pASH26	pRMB19 [3]	3xHA-tagged HCV E2$^{\Delta HVR1}$ protein, codon optimized for the *Chlamydomonas* nucleus, cytosolic	pRMB19 was digested with NdeI and EcoRI to excise p24. The insert was amplified using oASH174 and oASH175 from a synthetic plasmid carrying the *Chlamydomonas* codon optimized E2$^{\Delta HVR1}$ protein. The fragments were ligated using Gibson assembly.	oASH174	oASH175	*C. reinhardtii*
pASH27	pRMB19 [3]	3xHA-tagged HCV E2$^{\Delta HVR1}$ protein, codon optimized for the *Chlamydomonas* nucleus, targeted for secretion	pRMB19 was digested with NdeI and EcoRI to excise p24. The insert was amplified using oASH173 and oASH175 from a synthetic plasmid carrying the *Chlamydomonas* codon optimized E2$^{\Delta HVR1}$ protein. The fragments were ligated using Gibson assembly.	oASH173	oASH175	*C. reinhardtii*
pASH28	pRMB19 [3]	3xHA-tagged HCV E2$^{\Delta HVR1}$ protein, codon optimized for the *Chlamydomonas* nucleus, ER-targeted	pRMB19 was digested with NdeI and EcoRI to excise p24. The insert was amplified using oASH173 and oASH176 from a synthetic plasmid carrying the *Chlamydomonas* codon optimized E2$^{\Delta HVR1}$ protein. The fragments were ligated using Gibson assembly.	oASH173	oASH176	*C. reinhardtii*
pASH41	pRMB25 [2]	Addition of an N-terminal 3xHA-tag to a YFP-6xHis	pRMB25 was linearized with NcoI. The 3xHA-tag was amplified using oASH170 and oASH171 from plasmid pDK231 and inserted via Gibson assembly.	oASH170	oASH171	*E. coli*

pASH42	pASH1	PpYFP under the control of the endogenous carbonic anhydrase promoter (FVE85_3079)	pASH1 was digested with SacII and NheI to excise the *Actin* promoter and *PpYFP*. The inserts were amplified using oASH162 and oASH180 from genomic DNA (promoter) and oASH187 and oASH12 from pASH1 (PpYFP). The fragments were ligated using Gibson assembly.	oASH162 oASH187	oASH180 oASH12	*P. purpureum*
pASH43	pASH1	PpYFP under the control of the endogenous hypothetical PsbQ protein promoter (FVE85_3577)	pASH1 was digested with SacII and NheI to excise the *Actin* promoter and *PpYFP*. The inserts were amplified using oASH159 and oASH236 from genomic DNA (promoter) and oASH237 and oASH12 from pASH1 (PpYFP). The fragments were ligated using Gibson assembly.	oASH159 oASH237	oASH236 oASH12	*P. purpureum*
pASH45	pASH8	PpYFP targeted for secretion using the *carbonic anhydrase* signal peptide	pASH8 was digested with SpeI and NheI to excise *cat*. The inserts were amplified using oASH131 and oASH183 from genomic DNA (*CA* SP) and oASH184 and oASH12 from pASH1 (*YFP*). The fragments were ligated using Gibson assembly.	oASH131 oASH183	oASH184 oASH12	*P. purpureum*
pASH48	pASH8	PpYFP targeted for ER-residency using the *carbonic anhydrase* signal peptide and a C-terminal HDEL	pASH8 was digested with SpeI and NheI to excise *cat*. The inserts were amplified using oASH131 and oASH183 from genomic DNA (*CA* SP) and oASH184 and oASH223 from pASH1 (*YFP*). The fragments were ligated using Gibson assembly.	oASH131 oASH183	oASH184 oASH223	*P. purpureum*
pASH49	pASH1	PpYFP under the control of the endogenous *formate/nitrite transporter* promoter (FVE85_0107)	pASH1 was digested with SacII and NheI to excise the *Actin* promoter and *PpYFP*. The inserts were amplified using oASH224 and oASH225 from genomic DNA (promoter) and oASH226 and oASH12 from pASH1 (PpYFP). The fragments were ligated using Gibson assembly.	oASH224 oASH226	oASH226 oASH12	*P. purpureum*
pASH50	pASH1	PpYFP under the control of the endogenous PYP1 protein promoter (FVE85_6364)	pASH1 was digested with SacII and NheI to excise the *Actin* promoter and *PpYFP*. The inserts were amplified using oASH227 and oASH228 from genomic DNA (promoter) and oASH229 and oASH12 from pASH1 (PpYFP). The fragments were ligated using Gibson assembly.	oASH227 oASH229	oASH228 oASH12	*P. purpureum*
pASH52	pASH1	PpYFP under the control of the endogenous *Chlorophyll a binding protein* promoter (FVE85_6435)	pASH1 was digested with SacII and NheI to excise the *Actin* promoter and *PpYFP*. The inserts were amplified using oASH233 and oASH234 from genomic DNA (promoter) and oASH235 and oASH12 from pASH1 (PpYFP). The fragments were ligated using Gibson assembly.	oASH233 oASH235	oASH234 oASH12	*P. purpureum*
pASH54	pASH8	6xHis-tagged secreted HBV S/preS1	pASH8 was digested with SpeI and NheI to excise *cat*. The insert was amplified using oASH131 and oASH189 from a synthetic plasmid carrying the *Porphyridium* codon optimized S/preS1 protein. The fragments were ligated using Gibson assembly.	oASH131	oASH189	*P. purpureum*

pASH55	pASH8	6xHis-tagged secreted HCV E2$^{\Delta HVR1}$	pASH8 was digested with SpeI and NheI to excise *cat*. The insert was amplified using oASH131 and oASH188 from a synthetic plasmid carrying the *Porphyridium* codon optimized HCV E2$^{\Delta HVR1}$. The fragments were ligated using Gibson assembly.	oASH131	oASH188	*P. purpureum*
pASH59	pASH8	Untagged HBV S/preS1 protein without signal peptide	pASH8 was digested with SpeI and NheI to excise *cat*. The insert was amplified using oASH129 and oASH1254 from a synthetic plasmid carrying the *Porphyridium* codon optimized HCV E2$^{\Delta HVR1}$. The fragments were ligated using Gibson assembly.	oASH131	oASH188	*P. purpureum*
pASH60	pASH1	Secreted HCV E2$^{\Delta HVR1}$ under the control of the endogenous *ChlBP* promoter	pASH1 was digested with SacII and NheI to excise the *Actin* promoter and *PpYFP*. The inserts were amplified using oASH233 and oASH255 from genomic DNA (promoter) and oASH256 and oASH130 from a synthetic plasmid carrying codon optimized HCV E2$^{\Delta HVR1}$ The fragments were ligated using Gibson assembly.	oASH233 oASH256	oASH255 oASH130	*P. purpureum*

1) Li & Bock, 2018
2) Barahimipour et al., 2015
3) Barahimipour et al., 2016

2.12 Primers

Primer	Primer sequence 5'-3'	Usage
oASH24	AAGAGGTGGACGGCTTG	sequencing
oASH25	CAGTAAGTGGACAACAAACAAG	sequencing
oASH72	TGGAATTGTGAGCGGATAAC	sequencing
oASH73	CACGGTTCTGATCACCAGCCTAC	sequencing
oASH143	GCAAGGCGTACCGGAGAAG	northern probe
oASH144	CATGTTGCCTGCACTCCTCG	northern probe
oASH2	GGCACTCTGAAGCCTTGTGAGAGCTAGCGCGACGTTTACTTGATCAGCTCGTCCATGC	cloning
oASH4	GGCACTCTGAAGCCTTGTGAGAGCTAGCGCGACGTTTATTTAATTAATTCATCCATACC	cloning
oASH6	GGCACTCTGAAGCCTTGTGAGAGCTAGCGCGACGTTTACTTGATGAGCTCGTCCATCC	cloning
oASH8	GGCACTCTGAAGCCTTGTGAGAGCTAGCGCGACGTTTAGGCGCCGCCCTGCCA	cloning
oASH12	GGCACTCTGAAGCCTTGTGAGAGCTAGCGCGACGTCTACTTGATCAGCTCGTCCAT	cloning
oASH13	GAAGCAGATCGCAGCTCGAGGAGTGCAGGCAACACTAGTATGGAGAAGAAGATCACCGGCTAC	cloning
oASH14	GAAGCAGATCGCAGCTCGAGGAGTGCAGGCAACATGGTGAGCAAGGGCGA	cloning
oASH15	GAAGCAGATCGCAGCTCGAGGAGTGCAGGCAACATGGTTTCAAAAGGTGA	cloning
oASH16	GAAGCAGATCGCAGCTCGAGGAGTGCAGGCAACATGGTCTCGAAGGGGGAG	cloning
oASH19	GAAGCAGATCGCAGCTCGAGGAGTGCAGGCAACATGGTGTCGAAGGGCGAG	cloning
oASH129	CTCGAGGAGTGCAGGCAACACTAGTATGGAGAACATCACGTCGGGCTTCC	cloning
oASH130	CTGAAGCCTTGTGAGAGCTAGCGCGACGTCTAGGCGTAATCGGGGACATC	cloning
oASH131	CTCGAGGAGTGCAGGCAACACTAGTATGCGTAAGATGACGCTCACG	cloning
oASH132	CTGAAGCCTTGTGAGAGCTAGCGCGACGTCTACAGCTCGTCGTGCGAGCC	cloning
oASH136	GGAGTGCAGGCAACACTAGTATGCCGGGCGCGAAGC	cloning
oASH159	CACGGAGTCGGACAATCTCGCACCGCAGATTGATGAGAAATCGGAGTAGC	cloning
oASH162	GGCACGGAGTCGGACAATCTCGCACCGCGCCGTGATGCTTGAAAGGATAG	cloning
oASH170	CTTTAAGAAGGAGATATACCATGGCGTACCCATACGATGTTCCTGAC	cloning
oASH171	GTAAATAATTCTTCACCTTTTGAAACCATGCTGCCAGCGTAATCTG	cloning
oASH177	CCACTGCTACTCACAACAAGCCCATATGGAGAACATCACCAGCGGCTTC	cloning
oASH178	CCCCACTGCTACTCACAACAAGCCCATATGGCCCGCACCGGC	cloning
oASH179	CACTGCTACTCACAACAAGCCCATATGCCCGGCGCCAAGC	cloning
oASH175	ACAGGCGGTCCAGCTGCTGCCAGAATTCTTACGCGTAGTCCGGGACGTCG	cloning
oASH176	GGTACAGGCGGTCCAGCTGCTGCCAGAATTCTTACAGCTCGTCGTGGCTGC	cloning
oASH180	ACAGCTCCTCGCCCTTCGACACCATACTAGGGTTGCGTACGACGAACGTG	cloning
oASH183	CAGCTCCTCGCCCTTCGACACTGCCTGGACTGTTGCAATCAG	cloning
oASH184	CTGATTGCAACAGTCCAGGCAGTGTCGAAGGGCGAGGAGCTG	cloning
oASH187	GTTGTTAACAGGTAAGAAGTCTAGCCTAGTATGCGTAAGATGACGCTCACGGG	cloning
oASH188	CTTGTGAGAGCTAGCGCGACGTCTAGTGGTGGTGGTGGTGGTGGCCCGAGCCCTCCGAGC	cloning
oASH189	GAAGCCTTGTGAGAGCTAGCGCGACGTCTAGTGGTGGTGGTGGTGGTGCGAGCCGATGTACACCCACAG	cloning
oASH223	CTGAAGCCTTGTGAGAGCTAGCGCGACGTCTACAGCTCGTCGTGCTTGATCAGCTCGTCCATGCC	cloning

oASH224	CACGGAGTCGGACAATCTCGCACCGCTTTGCGCTCCACGATTCATG	cloning
oSAH225	ACAGCTCCTCGCCCTTCGACACCATTTTTCCTTGTTCCTTCTTGTCAAAC	cloning
oASH226	GGTTTGACAAGAAGGAACAAGGAAAAATGGTGTCGAAGGGCGAG	cloning
oASH227	CACGGAGTCGGACAATCTCGCACCGCGCGCGTTTCGTGCAAGC	cloning
oASH228	ACAGCTCCTCGCCCTTCGACACCATACCCCCGTCACTTCCTCTC	cloning
oASH229	GGGAGAGGAAGTGACGGGGGTATGGTGTCGAAGGGCGAG	cloning
oASH233	CACGGAGTCGGACAATCTCGCACCGCCGAGTCATGAAGACTCTGTGATG	cloning
oASH234	ACAGCTCCTCGCCCTTCGACACCATTTTGCTATCTGTGTGTGTCTGTTG	cloning
oASH235	CAACAGACACACACAGATAGCAAAATGGTGTCGAAGGGCGAG	cloning
oASH236	CAGCTCCTCGCCCTTCGACACCATGCTAGACTTCTTACCTGTTAACAACTC	cloning
oASH237	GTTGTTAACAGGTAAGAAGTCTAGCATGGTGTCGAAGGGCGAG	cloning
oASH255	CAAGCCCGTGAGCGTCATCTTACGCATTTTGCTATCTGTGTGTGTCTGTTG	cloning
oASH256	CAACAGACACACACAGATAGCAAAATGCGTAAGATGACGCTCACG	cloning

3 Methods

3.1 Cultivation and transformation of microorganisms

3.1.1 Cultivation of microalgae

Cultivation of *P. purpureum* strain SAG 1380-1d, obtained from Culture Collection of Algae at Goettingen University, was performed under photoautotrophic growth conditions in low salt artificial seawater (lsASW) or agar-solidified lsASW. For liquid culture cultivation, the desired volume of medium was inoculated with *P. purpureum* cells transferred from lsASW-agar plates or liquid culture and grown at 25°C on a rotary shaker (120 rpm) under continuous light (100 µmol photons m^{-2} s^{-1}). To ensure sufficient CO_2 supply in the photoautotrophic medium, cellulose or cotton plugs or punctured aluminum foil sealed with leukopor were used. For long-term storage, stationary cultures were washed once with lsASW medium, resuspended in lsASW medium supplemented with 7% DMSO and stored at - 80°C without snap freezing.

Cultivation of *C. reinhardtii* strain UVM11 was performed under photomixotrophic standard growth conditions in TAP medium as described elsewhere (Harris, 2009). Maintenance of *C. reinhardtii* UVM11 transformant strains took place on TAP plates without antibiotics under 10 µmol photons m^{-2} s^{-1} constant light at 22°C. Re-streaking was performed every two to three months.

3.1.2 Determination of microalgal cell number and cell diameter

Algal cell number and cell diameter were determined using the Z2 Coulter Particle Count and Size Analyzer (Beckman Coulter). 100 µl cell culture were mixed with 9.9 mL counting buffer (isoton II). Cell diameters from 3.5 µm to 12 µm were considered for cell number determination. Samples were counted twice and the average was taken.

3.1.3 Determination of dry biomass

10 mL of *P. purpureum* cell culture grown in bioreactors have been filtered through 0.2 µm pre-weight filters with the help of a vacuum. Filters were dried overnight in an oven set to 65°C and weighted. The difference of the pre-weighted to the dried, microalgae containing filters was the dried biomass. Biomass in milligram per liter per day (mg L^{-1} day^{-1}) productivity was calculated as

$$P_{Biomass}(mgL^{-1}day^{-1}) = (X_2 - X_1)/(t_2 - t_1)$$

With X_1 and X_2 being the dry weight on day t_1 and t_2.

3.1.4 Freeze drying of *P. purpureum* cells

Algal cultures, harvested after cultivation in the bioreactor, were washed once with lsASW by centrifugation at 5,000 *g*, 4°C for 10 min to remove exopolysaccharides. Then, cells were centrifuged for 10 min at 5,000 *g*, 4°C in 50 mL tubes and were frozen at – 80°C. Afterwards

cells have been freeze-dried in an open tube overnight in a freeze drier to completeness. Samples were stored at – 80°C or processed by bead milling two times for 1 min at 30 sec^{-1} frequency.

3.1.5 Plating of *P. purpureum* to obtain single colonies

The attempt to obtain individual colonies of *Porphyridium* by plating a dilution series directly on lsASW agar plates failed in our hand. Either no cells (high dilution), or a lawn of cells (low dilution) grew on those plates. We employed a plating protocol by Skeffington et al., 2019 to obtain single cells. In short, 200 µl of diluted liquid culture was mixed with 1 mL of 0.5 mg/mL wheat starch in lsASW, plated on lsASW agar plates and incubated for two weeks at 100 µmol photons m^{-2} s^{-1} light intensity until first colonies appeared.

3.1.6 Transformation of microalgae

Nuclear transformation of *P. purpureum* was performed using a biolistic approach with a DuPont PDS-1000/He biolistic gun (Bio-Rad). To this end, 5 x 10^7 cells from an exponential growth phase culture (1-2 x 10^6 cells/mL) were harvested by centrifugation for 5 min at 4,000 *g*, washed twice with lsASW medium and plated on an lsASW-agar plate without antibiotics. Bombardment was performed with 0.5 mg gold particles (0.6 µm diameter size) coated with 1 µg circular plasmid DNA using a 1350 psi rupture disc at a distance of 9 cm from the stopping screen to the target. After bombardment, the cells were washed off the plate with lsASW and washed twice in liquid medium to remove cell debris. Then, cells were plated on lsASW plates containing 25 mg/L Zeocin™, incubated in dark overnight and put in light at 100 µmol photons m^{-2} s^{-1} for 3 weeks. Colonies were picked and transferred to liquid lsASW medium with 25 mg/L Zeocin™ and the resistant lines were genotyped for the presence of the plasmid via PCR.
C. reinhardtii was transformed using the glass bead method as described in the literature (Harris, 2009).

3.1.7 *E. coli* cultivation for plasmid DNA cloning

The *E. coli* strain Top 10 F' (Invitrogen, genotype F' *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(araleu) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G) was used for cloning. Bacteria were cultivated in liquid or on solid LB medium containing appropriate antibiotics and incubated overnight at 37°C. Liquid cultures were shaken at 200 rpm in test tubes (for mini prep) or culture flasks (for midi prep). For cultivation on solid medium, LB medium was solidified with 1.5% (w/v) Micro agar.

3.1.8 *E. coli* cultivation and induction for protein overexpression

The *E. coli* expression strain BL21 DE3, genotype F– *ompT hsdS*B (rB– mB–) *gal dcm lacY1* was used for recombinant protein expression. 50 mL LB medium were inoculated in a flask with a single colony, as a pre-culture. After incubation overnight at 37°C and shaking at 200

rpm, 500 ml of the appropriate medium was inoculated with the pre-culture and the starting OD of this main culture was set to an OD_{600} of 0.1 followed by incubation at 37°C in a baffled shaking flask (200 rpm) until the culture reached an OD_{600} of approx. 0.8. Protein overexpression was induced by adding IPTG to a final concentration of 500 µM. Further incubation was performed at 30°C and 200 rpm for 4 h. Cell cultures were harvested by centrifugation at 13,000 *g* for 15 min and 4°C. Cell pellets were stored at -20°C or processed directly.

3.1.9 Preparation of *E. coli* heat-shock competent cells

E. coli heat-shock competent cells were prepared using a protocol described by Inoue et al., 1990. A single colony of *E.coli* Top 10 F' strain was inoculated into 4 mL LB medium and grown at 37°C overnight. The next morning, 250 mL of SOB medium with 20 mM $MgSO_4$ were inoculated with the over-night culture to set a starting OD of OD_{600} of 0.01. The bacterial culture was incubated under strong shaking (180 rpm) at 18°C until an OD_{600} of 0.5 – 0.75 was reached. The culture was chilled on ice for 10 min and cells were harvested by centrifugation at 2,500 *g* for 10 min at 4°C. The cells were resuspended in 80 mL ice-cold TB , incubated on ice for 15 min and spun at 2,500 *g* for 10 min at 4°C. Bacterial pellet was resuspended gently in 20 mL ice cold TB and DMSO was added to a final concentration of 7%. Finally, 200 µl of the cell suspension were aliquoted into pre-chilled tubes, snap frozen in liquid nitrogen and stored at -80°C prior to use.

3.1.10 Heat-shock transformation of *E. coli*

Chemo-competent *E. coli* cells were transformed with the desired plasmid via heat shock. Aliquoted cells were thawed on ice and 100 ng plasmid DNA or DNA ligation reaction mixture was added followed by incubation on ice for 30 min. The heat shock was performed at 42°C for 45 sec. After the heat shock, cells were incubated on ice for 3 min and 900 µl LB medium was added afterwards. The cells were put in a thermomixer set to 37°C and mixed at 800 rpm for 30 min. Finally, the bacteria were pelleted at 2,000 *g* for 1 min, then resuspended in the remaining medium and plated on a LB plate supplemented with the appropriate antibiotics. Cells were cultivated overnight at 37°C and resistant colonies were picked for further analysis.

3.2 Nucleic acid techniques

3.2.1 Isolation of total DNA from *Porphyridium*

Total DNA for PCR analysis was extracted by pelleting 1 mL of algal cell culture, followed by resuspension in 100 µl of 10 mM EDTA and boiling it 5 min at 99°C. The lysate was centrifuged for 3 min at 1,000 *g* and the DNA containing supernatant was used as template in a PCR.

Genomic DNA for all other purposes was extracted using a modified CTAB method (Li and Bock, 2018). In short, 5 mL of an exponential phase culture was harvested by centrifugation

for 5 min at 4,000 *g*. The cells were then washed once with lsASW before resuspension in 1 mL of CTAB extraction and incubated with shaking at 65°C for 30 min. Afterwards 1 mL of phenol/chloroform/isoamyl alcohol (25:24:1) was added to the sample and vortexed thoroughly. The mixture was then centrifuged at 10°C for 10 min at 12,000 *g* and the upper aqueous phase was transferred to a new tube. RNA was digested for 30 min at 37°C by adding 10 µg RNAse A, and the DNA was re-extracted with the same volume of chloroform/isoamyl (24:1). After centrifugation at 10°C for 10 min at 12,000 *g*, the DNA in the upper aqueous phase was precipitated with 0.7 volumes of ice-cold isopropanol and an incubation at -20°C for 1 h. The sample was then centrifuged at 4°C for 10 min, 12,500 *g*, and the isopropanol was discarded. The DNA pellet was washed with 70% ethanol (v/v) and air-dried before being resuspended in water.

3.2.2 Polymerase chain reaction (PCR)

PCR reaction for genotyping (including colony PCR) was performed in 25 µl reaction mixture using the DreamTaq DNA polymerase following the manufacturer's instructions. For cloning experiments, where higher fidelity is needed, Phusion DNA polymerase was used in a 50 µl reaction volume, following the manufacturer's instructions.

Dreamtaq reactions were performed by adding the DNA template to a reaction volume of 20-50 µL using the following reagents (in their final concentrations): Dreamtaq DNA polymerase at 0.04 U/µL, primers at 0.2 nM each, green Dream Taq buffer 1X, and 0.2 mM of each dNTP (dATP, dCTP, dTTP, dGTP), plus deionized water to complete the rest of volume.

The PCR was performed in a Mastercycler EPGradient thermocycler (Eppendorf). The thermocycler program used was: initial denaturation 3 minutes at 94°C, followed by 25-40 cycles of three steps: (1) denaturation of 20 s at 94°C; (2) annealing of 20 s at 51- 64°C; and (3) extension for 15-120 s at 72°C (30s per kb of amplicon). The reaction ends after an extension step of 3 min at 72°C.

For Phusion reactions, reactions were performed adding a DNA template to complete a 20-50 µL reaction using the following recipe (final concentrations). 0.1 U/µL of Phusion polymerase, primers at 0.2 nM each, green Phusion HF buffer 1X, and 0.2 mM of each dNTP (dATP, dCTP, dTTP, dGTP), plus deionized water to complete the rest of volume.

The 35 thermocycler program used was: initial denaturation 2 minutes at 98°C, followed by 25-40 cycles of three steps: (1) denaturation of 15 s at 98°C; (2) annealing of 20s at 57-68°C; and (3) extension for 15-60s at 72°C (15s per kb of amplicon). The reaction ends after an extension step of 3 min at 72°C.

3.2.3 DNA agarose gel electrophoresis

The separation by size and visualization of DNA was enabled by agarose gel electrophoresis. DNA samples were separated in 1 % agarose gels including 0.1 mg/L EtBr. Samples were mixed with 10 x DNA loading buffer (1x final concentration) and run at 80 V to 100 V for 1 h to 2 h. DNA was detected under UV light (λ=312 nm) in the UV documentation Quantum CX 5 (Vilber, Marne-la-Vallée, France).

3.2.4 Purification of DNA fragments from gel and PCR

DNA was purified using the NucleoSpin® Gel and PCR Clean-up kit, according to the manufacturer's instructions. Purified DNA were eluted in ddH_2O and stored at -20°C.

3.2.5 Measurement of DNA and RNA concentration and purity

DNA and RNA concentrations were quantified by optical density measurement using a NanoDrop 1000 spectrophotometer (PeqLab, Germany). Purity levels of RNA and DNA were accepted when having an A_{260}/A_{280} of 1.8-2.0 for DNA, A_{260}/A_{280} of 1.9-2.1 for RNA and A_{260}/A_{230} in the range of 2.0-2.2 for both nucleic acids.

3.2.6 Gibson assembly based cloning

DNA of interest was inserted into the transformation vector by Gibson assembly using either Gibson Assembly 2x Master Mix Kit (NEB) or 5X In-Fusion HD Enzyme Premix (Takara), following manufacturer's instructions. In short, the destination vector was linearized and mixed with DNA segments harboring 20 bp overlapping 5' and 3' ends generated by PCR. Master-mix in the appropriate dilution was added and the reaction was incubated at 50 °C for 15 min. Half of the ligation reaction was used for transformation of *E. coli*.

3.2.7 Plasmid DNA isolation

5 mL LB medium containing appropriate antibiotics were inoculated with a single colony grown on a selection plate or with frozen cell material from a DMSO cryo-stock and cultivated at 37°C on a rotary shaker (200 rpm). After overnight incubation, 2 to 4 mL were harvested through centrifugation and the supernatant was decanted. For the isolation of the plasmid DNA, the NucleoSpin® Plasmid EasyPure Kit (Macherey-Nagel, Düren, Germany) was used according to the instructions supplied by the manufacturer. When higher amounts of pure plasmid DNA were required, the plasmid DNA was isolated using the NucleoBond® Xtra Midi/Maxi Kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instruction with 200 mL of *E. coli* culture as input material.

3.2.8 Sanger sequencing of DNA

Purified PCR amplification products or plasmid DNA were sequenced using the service of Eurofins MWG Operon (Ebersberg, Germany), following the operator's instruction. Sequences were visualized using SeqMan software by DNAStar.

3.2.9 Isolation of total RNA from *Porphyridium* using the Zymo Research kit

10 mL of liquid cell culture at exponential growth phase were used for RNA isolation. Cells were harvested by centrifugation at 2,000 g and 4°C for 2 minutes. The pelleted cells were immediately frozen in liquid nitrogen. The cell pellet was resuspended in 1 mL TRIzol® (Thermo Fisher Scientific,) and vortexed at maximum speed until the cell pellet dissolved completely. The suspension was incubated at 65°C for up to 5 minutes, then centrifuged at 12,000 *g* at room temperature for 1 minute. The supernatant was transferred to a new tube containing an equal volume of 100% ethanol, mixed and loaded onto the RNA extraction column (Direct-zol™ RNA MiniPrep, Zymo Research, California, USA), followed by centrifugation at 12,000 *g* at room temperature for 30 sec. Columns were washed with 400 µl wash buffer at 12,000 *g* at room temperature for 30 seconds. RNA isolation was followed by in-column DNA digestion using DNAse I for 30 min at 37°C. The column was washed twice with 400 µl prewash buffer (Zymo Research, California, USA), followed by centrifugation at 12,000 *g* at room temperature for 30 seconds. Afterwards, the column was washed with 700 µl wash buffer and centrifuged at 12,000 *g* at room temperature for 1 minute. The column was dried by centrifugation at 12,000 *g* at room temperature for 2 minutes. RNA was eluted from the column with 30 µl RNase-free, DNase-free water and immediately transferred onto ice until further experiments, or was frozen and stored at -80°C.

3.2.10 Isolation of total RNA from *Porphyridium* using Trizol/chloroform

Total RNA was isolated following the standard Trizol/chloroform protocol with some modifications. In short, 10 mL of liquid cell culture at exponential growth phase was harvested by centrifugation at 2,000 *g*, 4°C for 2 min and washed once with lsASW. Cells were resuspended in 2 mL Trizol and incubated for 5 min at 65°C. Following homogenization, cells were centrifuged at 12,000 *g* at 4°C for 10 min and the pellet was discarded. 400 µl chloroform was added to the supernatant and sample was vigorously shaken and incubated for 3 min at room temperature. The sample was centrifuged at 12,000 *g*, 4°C for 15 min and the upper aqueous phase containing the RNA was transferred to a new tube. 1 mL of ice cold isopropanol was added and the RNA was precipitated at -20°C for 1 h. After a centrifugation at 12,000 *g*, 4°C for 15 min, the RNA pellet was washed with 70% EtOH, air dried and resuspended in water.

3.2.11 Blotting of RNA onto nylon membranes by capillary transfer

For Northern blotting, 4 µg total RNA was used and mixed with sample buffer composed of 1 x MOPS pH 7, 50% formamide, 7% formaldehyde, 1 mM EDTA, 10% glycerol, 0.1% EtBr, 0.015% bromphenol blue and 0.015% xylene cyanol. Samples were denatured at 75°C for 15 min, shortly cooled on ice and loaded on a denaturing 1% agarose gel containing 1x MOPS pH 7 and 6% formaldehyde. The gel electrophoresis was run for approximately 4 h at 80 V and 4°C and blotted onto a Hybond-N nylon membrane (GE Healthcare, Buckinghamshire, UK) using capillary blotting with 10x SSC buffer overnight. After blotting, the RNA was cross-linked to the membrane by exposure to UV light (0.120 Joules/cm^2) using a BLX-254 UV crosslinker. Nylon membranes with covalently bound RNA were used for nucleic acid detection with the appropriate radioactively labeled complementary probes.

3.2.12 Radioactive labeling of DNA probes using Klenow polymerase

Radioactive probes were synthesized using the Megaprime™ DNA labelling kit, according to the manufacturer's instruction. Briefly, a mixture of 25 ng DNA, 2.5 µL random primer in 16.5 µL was denatured for 8 min at 95ºC and cooled for 2 min on ice. Afterwards 2.5 µL dNTPs, 5 µL labelling buffer, 1 U Klenow DNA Polymerase and 2.5 µL [α-^{32}P] dCTP were added. Following incubation at 37ºC for 30 min and enzyme inactivation at 95ºC for 5 min, unincorporated nucleotides were removed using the Illustra MicroSpin G-50 column following the manufacturer's instructions.

3.2.13 Hybridization of immobilized nucleic acids with radioactive probes

The nylon membrane with the immobilized nucleic acids was pre-hybridized with 15 mL of Church buffer at 65°C for at least 1 hour. The radioactively labelled DNA probe was added to the Church buffer in the isotope lab and hybridized at 65°C over-night. The next day, membranes were washed for 15 min at 65°C with pre-heated RNA wash buffer I , followed by two washes at 65°C with wash buffer II. Membranes were wrapped in plastic foil and exposed to a storage phosphorus screen, overnight followed by radioactive signal detection using the Typhoon™ TRIO+ scanner.

3.3 Protein techniques

3.3.1 Total soluble protein extraction from microalgae

Total soluble protein (TSP) was extracted from microalgae by harvesting an appropriate amount of *Porphyridium* or *C. reinhardtii cell* culture. Pelleted cells were washed in the corresponding liquid medium and resuspended in 300 to 500 µl extraction buffer. The cells were disrupted by three rounds of sonication at 20% amplitude for 10 sec on ice. The lysate was cleaned by two rounds of centrifugation at 4°C at 15,000 *g* for 10 min and stored at -20°C

or -80°C for long time storage. Proteins were quantified using the BCA assay (Thermo Scientific).

3.3.2 Total protein extraction from microalgae

Total protein was extracted using a modified phenol/methanol method (Chatterjee et al., 2012). In short, cells were resuspended in extraction buffer, sonicated at 20% amplitude 3x 10 sec on ice and mixed with an equal volume of phenol pH 7.6. The sample was vortex vigorously and centrifuged at 12,000 *g*, 4°C for 10 min. The upper phenolic phase was taken and proteins were precipitated with 4 volumes of ice cold 0.1 M ammoniumacetate in 100% MeOH at -20°C overnight (or at least 1h). After centrifugation at 12,000 *g*, 4°C for 5 min the protein pellet was washed twice with ice cold methanol and resuspended in protein buffer. Proteins were quantified using the BCA assay (Thermo Scientific).

3.3.3 Trichloroacetic acid (TCA) protein precipitation from exopolysaccharide-rich liquid cultivation medium

Medium from a liquid alga culture was collected by centrifugation at 12,000 *g*, 4°C for 10 min to remove cell debris and transferred to a new tube. CTAB to a final concentration of 0.2 % was added and sample was vortexed rigorously. Precipitated polysaccharides were removed and 1 to 2 µg BSA per mL medium was added. TCA was added to a final concentration of 15 % and the sample was incubated on ice for 30 min to let the proteins precipitate. After centrifugation at 20,000 *g*, 4°C for 10 min, the pellet was washed twice with ice-cold methanol and 0.1 M NaAc. The protein pellet was air dried and resuspended in 1x SDS sample buffer.

3.3.4 Protein extraction from *E. coli* under native conditions

Protein containing *E. coli* pellet from 3.1.8 was resuspended in Ni-NTA lysis buffer at 5 mL per gram wet weight and 1 mg/mL lysozyme was added. After an incubation of 30 min on ice, cells were sonicated on ice for 6 x 10 sec at 50% amplitude with 10 sec cooling periods in-between. RNase A (10 µg/mL) and DNase I (5 µg/mL) were added and lysate was incubated for another 15 min on ice. The lysate was cleared by centrifugation at 10,000 *g*, 4°C for 20 min and cleared lysate was used for affinity purification of recombinant His-tagged proteins.

3.3.5 Affinity purification of recombinant His-tagged proteins

All steps of the purification protocol were performed at 4°C. An empty polypropylene column was loaded with 2 mL of Ni-NTA slurry and equilibrated with 2 column bed volumes (CV) Ni-NTA lysis buffer. The cleared lysate of *E. coli* was loaded onto the column und protein was allowed to bind to the Ni-NTA resins. When the lysate almost run out, column out-flow was closed, 3 CV washing buffer I was added to the column and incubated for 15 min at room temperature to release common contaminants of *E. coli*. Afterwards, gravity flow was re-instated and column was washed with consecutive washings of 3 CV washing buffer II and

washing buffer III. Proteins were eluted by adding 5 mL elution buffer in 1 mL steps and the fractions were collected. Eluate fractions were evaluated by SDS-PAGE followed by Coomassie staining. Further purification was performed via size exclusion chromatography.

3.3.6 Protein concentration and buffer exchange

After evaluation by SDS-PAGE and Coomassie staining, eluate fractions containing the proteins were pooled and concentrated using Amicon® Ultra Centrifugal Filters with 10 MWCO following manufacturer's instruction. At the same time, buffer was exchanged to PBS (before size exclusion chromatography) or PBS + 20% glycerol (after size exclusion chromatography).

3.3.7 Size exclusion chromatography

Concentrated samples obtained from 3.3.6 were subjected to size exclusion chromatography using the HiLoad 16/60 Superdex 75 prep grade column, connected to the FPLC system Äkta Explorer 100. Column was equilibrated and operated using manufacturer's instructions and samples were fractionated into 2 mL aliquots. Briefly, PBS equilibrated column was loaded with 0.2 mL concentrated protein sample and run at 0.3 mL/min. Sample collection started after omitting the dead volume of the column. Fractions containing YFP were confirmed by visualization of fluorescence under UV light. These were pooled and concentrated as described in 3.3.6, quantified by BCA assay and spectroscopy using the molecular weight and extinction coeficient of the recombinant protein. Proteins were stored in PBS and 20% glycerol at -80°C.

3.3.8 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Protein separation by molecular size was accomplished by SDS-PAGE according to (Laemmli, 1970). For the separation of the samples a 12% polyacrylamide gel with a 3% stacking gel was used, that were cast using the Mini-PROTEAN Tetra Cells (Bio-Rad). 5x SDS sample buffer was added and the samples were incubated at 99 °C for 5 min. Denatured samples were loaded on the gel and separated for 1.5 h at 100 V.

3.3.9 Colloidal Coomassie staining of proteins

To visualize separated proteins on the SDS gel a Coomassie staining was used. The special formulation of colloidal Coomassie improves the sensitivity at a high rate (Kang et al., 2002). After separation, the SDS gel was washed two times with destainer (40% ethanol, 10% acidic acid), rinsed with water and stained overnight in colloidal Coomassie (BioRad). After that, the gel was incubated in water until bands became visible.

For the Coomassie working solution, the stock solution was mixed with methanol to a final concentration of 20 % (v/v) prior to use. All concentrations for the stock solution refer to a final volume of 1 L working solution.

3.3.10 Western Blot

The western blot method is used for electrophoretic transfer of proteins from polyacrylamid gels to nitrocellulose membranes, allowing the immunological detection of the desired proteins with a primary and secondary antibody. Proteins were blotted onto nitrocellulose membranes (0.45 µm, Amersham) using PerfectBlueTM „Semi-Dry“-Electroblotter in a three buffer system (Kyhse-Andersen, 1984) at 0.8 mA/cm^2 for 75 min.

Following blotting, the membrane was rinsed with water, incubated 1 min in Ponceau S staining solution (Thermo scientific) and destained with water until protein bands became visible. Protein stain was documented using a standard scanner. Following this, the nitrocellulose membrane was blocked for 30 min in TBS-T with 5 % skimmed milk (w/v), rinsed with TBS-T and incubated for 1 h or overnight with the appropriate primary antibody. The antibody was then collected for further usage and the membrane was washed three times in TBS-T for 10 min. Afterwards, the secondary antibody was applied for 1 h followed by three washes of TBS-T for 10 min each. Finally, an equal mixture of ECL 1 and ECL 2 solution was applied onto the membrane and chemiluminescence was detected in the G:Box detector.

3.3.11 Tunicamycin experiments

Tunicamycin (TM) was diluted to a concentration of 1 mg/mL, aliquoted and stored at -20 °C. For dosage titration experiments, *Porphyridium* was inoculated into 5 mL lsASW to a concentration of 4.5 x10^6 cells/mL (mid log phase) and 0, 5, 25 and 50 µg tunicamycin was added to a final concentration of 0, 1, 5 and 10 µg/mL. Cells were incubated at standard conditions for 40 h and samples were taken for protein isolation and determination of cell number after 16 h and 40 h, respectively.

For time course experiments, algae were inoculated in 20 mL lsASW to a final concentration of 5 x 10^6 cells/mL and after addition of 1 µg/mL final TM concentration incubated for 24 h under standard conditions. Samples for protein isolation were taken after 5, 10, 15, 30, 60, 120, 240, 360 and 1440 min. Proteins were extracted using the phenol extraction method (see 3.3.2) and analyzed using standard Western blot method (see 3.3.10).

3.4 Microscopy

Fluorescence of YFP was detected in living cells with a confocal laser-scanning microscope (TCS SP5; Leica, Wetzlar) using an argon laser for excitation (514 nm), a 510–535 nm filter for detection of YFP fluorescence, and a 630–720 nm filter for detection of chlorophyll fluorescence.

For fluorescence ratio analysis, YFP and chlorophyll fluorescence were detected using the same filters with an epifluorescence BX-51 microscope (Olympus). Images were captured and automatically analyzed using Image J software with a macro script provided by (Grishagin,

2015). The count number of chlorophyll fluorescence was considered as the total cell number, while the count number of YFP fluorescence was considered as the fluorescent cell number. The ratio between the fluorescent and total cell numbers was then calculated for each sample.

4 Results

4.1 Optimization of *Porphyridium* cultivation and transformation

4.1.1 Porphyridium heavily relies on the availability of CO_2 during cultivation

As *Porphyridium purpureum* has yet to be established as a model organism in the algal research field, there is a lack of reliable literature, especially when it comes to standard cultivation conditions. Many growth protocols and their outcomes are contradictory in the literature, likely due to variations in strains and growth media employed. In light of these uncertainties, it was necessary to investigate and optimize growth parameters that are best suited to the laboratory environment and the available strain to ensure optimal growth and reproducibility.

To date, the majority of growth media described for *P. purpureum* are minimal media that only support photoautotrophic growth, with the artificial seawater medium (ASW) proposed by Jones et al. (1963) being the most widely employed. Although there exist a few reports on mixotrophic growth of *Porphyridium*, the majority of published experiments have utilized minimal media. The ASW medium was also utilized in our lab for growth of *P. purpureum*. However, due to the recent increase of interest in red microalgae in general and *Porphyridium* in particular, optimized medium compositions have been published. Kathiresan et al. (2007) investigated various ASW salt concentrations, resulting in the determination of optimal concentrations for biomass production, which included 15.2 g/L of sodium chloride (in contrast to the standard ASW concentration of 27 g/L), as well as minor adjustments in the levels of $MgSO_4$ (4.9 g/L compared to 6.6 g/L) and KNO_3 (2 g/L compared to 1 g/L). This newly formulated medium, designated as lsASW (low-salt ASW), was evaluated alongside the regular ASW medium (Figure 2B) in a growth experiment.

P. purpureum cells cultivated in lsASW medium reached a significantly higher cell number in comparison to cultivation in ASW medium. After seven days of photoautotrophic cultivation, the lsASW-grown culture reached a final cell concentration of 1.46×10^7 cells/mL compared to the culture in ASW medium that reached a final concentration of 1.18×10^7 cells/mL. The doubling time was calculated by taking the natural logarithm of the difference of cell number over time in the exponential phase. Cells grown in the half-salt medium exhibited a slightly faster doubling time (38.5 h) than those cultivated in the ASW medium (42.8 h). It is important to note that both media were tested using aluminum foil for flask sealing, which is suboptimal, because it does not allow sufficient air exchange.

When using a minimal medium, the CO_2 availability plays a crucial role for the growth rate of photosynthetic organisms in both liquid and solid media. Therefore, different closure systems

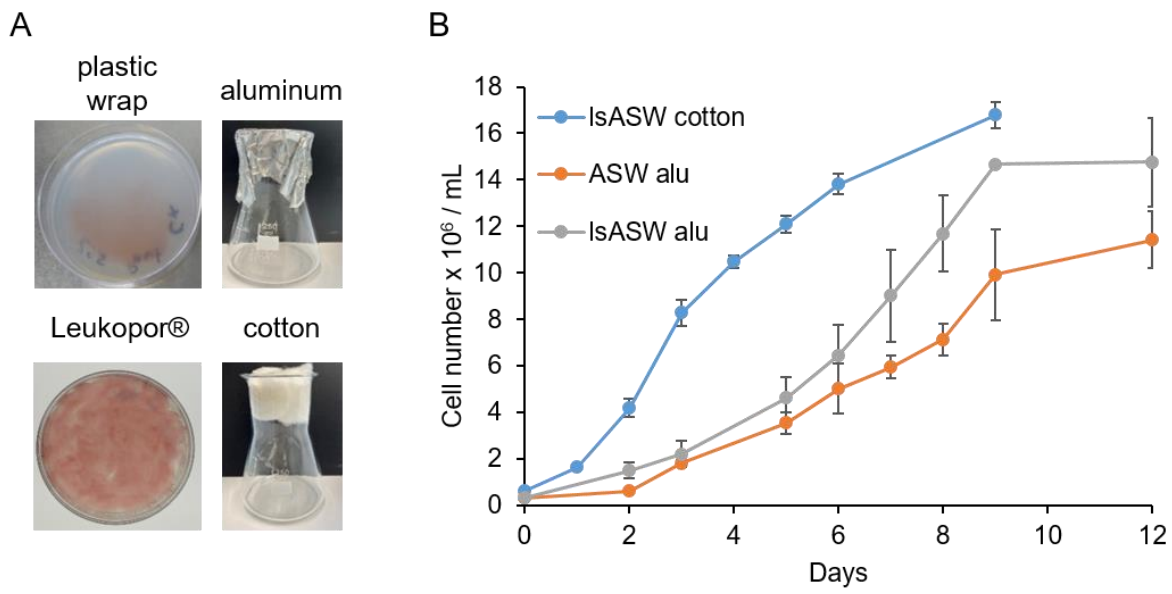


Figure 2: Cultivation of *Porphyridium purpureum* under standard conditions testing different seals, plugs and media.
A: Usage of different flask closures (aluminum or cotton plugs) and seals (plastic wrap or Leukopor® tape) for liquid cultures and cultures on plates, respectively. The higher the permeability (Leukopor® and cotton) of the sealing system, the better is the growth of *Porphyridium*. Pictures of plates were taken two weeks after spreading cells on solid medium. **B:** Growth curve of *Porphyridium* cultured in different media and using different closure systems. Liquid cultures were set to a starting cell density of 5 x 10^5 cells/mL and were grown photoautotrophically at 25°C under continuous light conditions (100 µmol photons m^{-1} s^{-1}). *Porphyridium* cells cultivated in lsASW show a faster growth and reached a higher final cell density compared to cells incubated in standard ASW medium. Cultures grown with the cotton plugs show a much faster growth, especially in the early exponential phase with a similar end concentrations as cultures grown in lsASW with aluminum seals.

with varying CO_2 permeability were tested for liquid cultures and cultures on solid medium as depicted in Figure 2A and B.

Liquid cultures grown in lsASW medium utilizing a cotton plug exhibited a shorter lag phase, a faster exponential doubling time of 19.0 h, and a higher maximal cell number on day nine (16.8 x 10^6 cells/mL) in contrast to the two other conditions tested with an aluminum foil cover. In comparison to the cultivation in lsASW with aluminium seal, this is a twofold increase of the growth rate. In addition to the increased growth rate, the variation in cell density between the three biological replicas was strongly reduced when cotton plugs were used instead of aluminum foil.

For the growth experiment on solid medium, 200 µl of a late-phase *Porphyridium* culture was spread onto lsASW agar plates, dried and sealed with either normal plastic wrap or Leukopor®, an air-permeable but microorganism-proof medical fabric band. After two weeks, growth was assessed visually. Plates sealed with plastic wrap showed only a sparse reddish cell lawn on a small part of the plate. In contrast, plates sealed with Leukopor® displayed a dense red cell lawn, indicating substantially increased growth as the result of the improved air exchange.

The comparison between cultures sealed with aluminum foil and plastic wrap versus those closed with cotton plugs or Leukopor® conclusively demonstrated *P. purpureum's* high dependence on efficient air exchange during cultivation. This was expected since both media tested were minimal media and the only carbon source that can be used for biomass production is CO_2. Due to this observation, lsASW medium in combination with the usage of cotton or cellulose plugs was selected as the standard cultivation method for *P. purpureum.*

Another advantage of the use of lsASW medium is that the activity of the selective agent Zeocin™ is reduced at high salt concentrations (see Zeocin™ selection reagent user guide 25-0078, Invitrogen). Low salt would therefore prolong Zeocin™ stability in the lsASW medium used for the maintenance of transgenic lines.

To further optimize laboratory cultivation of *Porphyridium*, the possibility of growth under mixotrophic or heterotrophic conditions was investigated using our laboratory strain.

4.1.2 *Porphyridium purpureum* strain SAG 1380-1d cannot grow mixo- or heterotrophically under standard laboratory conditions

For the use of *Porphyridium* as a host organism in biotechnology, growth rates and productivity play an important role. It was published that *P. purpureum* can grow mixotrophically, with glucose, glycerol or acetate as well as heterotrophically in the dark with glucose as carbon source (Jiao et al., 2018; Oh et al., 2009). The first published genome draft of this microalga revealed the presence of a sodium:glucose symporter and four putative sucrose transporters, which could play a role in mixotrophic and heterotrophic growth (Bhattacharya et al., 2013). Based on these findings, heterotrophic and mixotrophic growth was tried to be replicated with the in-house used *P. purpureum* SAG 1380-1d strain. For heterotrophic growth, SAG 1380-1d cells were incubated for one month in the dark in liquid medium supplemented with different carbon sources (acetate (NaAc), glucose, sucrose and glycerol) (figure 3A).

Liquid cultures with a starting cell density of 5 $x10^5$ cells/mL were directly placed in the dark at 25 °C and shaking (125 rpm) for one month. No detectable increase in cell density was observed under any of the conditions tested. Consequently, none of the supplemented carbon sources allowed detectable heterotrophic growth, indicating that the SAG 1380-1d strain is incapable of utilizing acetate, glucose, sucrose or glycerol as the sole carbon source. The cells started floating on top of the medium or formed clumps inside the medium, which made the determination of the cell number impossible. Additionally, heterotrophic growth on solid medium was tested by adding 0.5% glucose or 1% glycerol to agar-solidified lsASW medium. 1 million cells were spread on the plates and incubated in the dark for 1 month. After the incubation period, no growth could be detected, neither on the lsASW plates nor on the lsASW

plates supplemented with the respective carbon source (Suppl. Figure 1B). Overall, no conclusive evidence for heterotrophic growth in the dark was observed for this *Porphyridium* strain.

Next, we tested the possibility of SAG 1380-1d growing under mixotrophic conditions in light by supplementing the medium with the organic carbon source glucose, which is commonly described in the literature for *Porphyridium* (Lee et al., 2019). Triplicates of lsASW medium, supplemented with 0.5% (v/v) glucose were inoculated with a starting cell density of 5 x 10^5 cells/mL and grown side-by-side with cells grown in plain lsASW medium under standard growth conditions (figure 3B). No growth improvement was observed during the exponential growth phase, and in the case of the glucose-containing culture, a slight decrease in growth rate was even noticed between day 6 and day12 of the experiment. The final cell numbers reached 11.51 x 10^6 cells/mL for the glucose-supplemented cultures compared to 14.76 x 10^6 cells/mL in minimal lsASW medium and thus showed no significant difference. Additionally,

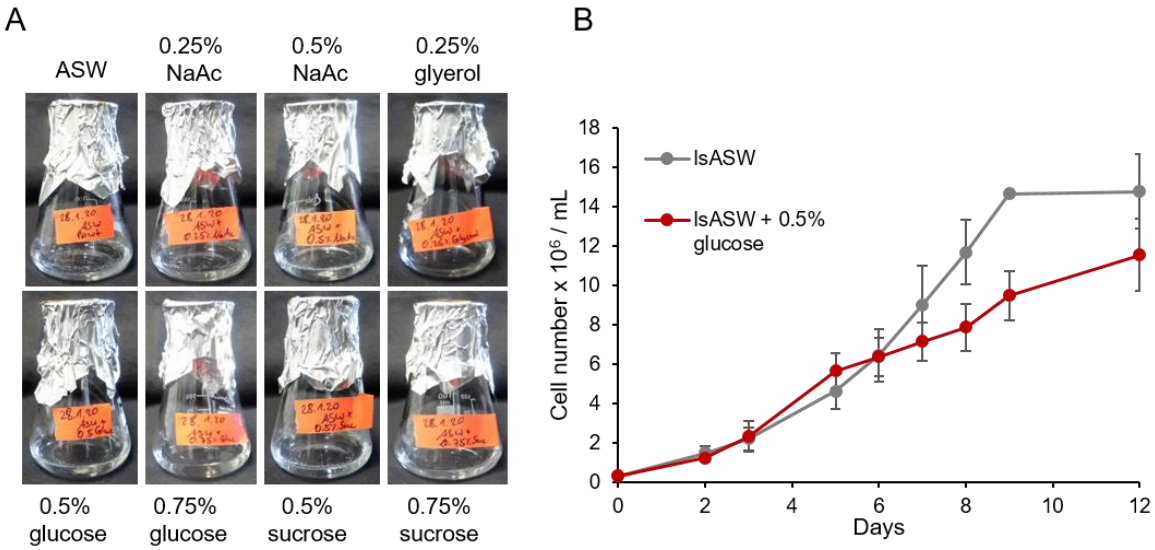


Figure 3: Heterotrophic and mixotrophic growth of *Porphyridium*.
A: *Porphyridium* grown in the dark for one month with addition of different potential carbon sources in liquid ASW medium. No visible growth could be detected.
B: Growth curve of cells grown in lsASW medium supplemented with 0.5% glucose. No significant difference could be detected between supplemented and non-supplemented medium at the end of cultivation, and at day 8 and 9, the non-supplemented culture reached even higher cell numbers.

other carbon sources (0.5% sucrose, 0.25% glycerol, 0.5% glucose, 0.5% NaAc, 0.25% NaAc and 1% LB) were tested in lsASW medium under standard conditions and showed no positive, or in the case of LB and 0.5% NaAc, even a detrimental effect on growth (suppl. figure 1A).

In conclusion, under the tested conditions, no mixotrophic growth could be observed for *P. purpureum* strain SAG 1380-1d. Although it would be desirable to identify conditions that allow for mixotrophic growth, no further efforts were made in this regard, as the use of lsASW medium and cotton plugs already improved growth and cultivation time significantly.

4.1.3 *P. purpureum* is able to reach high cell densities in an aerated photobioreactor

Being able to upscale cultivation is crucial to using a microorganism as a producer in biotechnology. A microalgal strain that expresses proteins to high level but does not grow in a high-volume bioreactor cannot compete with a microorganism that might produce only moderate amounts of recombinant protein but can be cultivated in an industrial-sized bioreactor. For instance, the cell wall-deficient *C. reinhardtii* strain UVM11 exhibits high expression levels of some heterologous proteins but encounters challenges when growing within a photobioreactor (PBR) due to the increased shearing forces associated with this cultivation method. In order to assess the possibility of *P. purpureum* SAG1380-1d to be up-scaled, cell number and biomass were monitored during growth in a PBR.

Figure 4A depicts the configuration of a custom-built PBR employed for cultivating of *P. purpureum*. The PBRs utilized in this study, designated as LWS-05, were constructed by the Institute für Getreideverarbeitung (IGV GmbH) in Nuthetal, Germany. Due to their customized nature, each bioreactor varied slightly in terms of filling volume, height, and diameter. The total volume ranged from 650 mL to 850 mL. These variations make it hard to compare growth experiments conducted across different PBRs. However, since the objective of this experiment was to demonstrate the viability of *P. purpureum* cultivation in a high-volume PBR, such discrepancies were acceptable, and only the end concentration of biomass after cultivation was assessed in duplicates (see Table 3). The growth of a culture cultivated in the PBR in

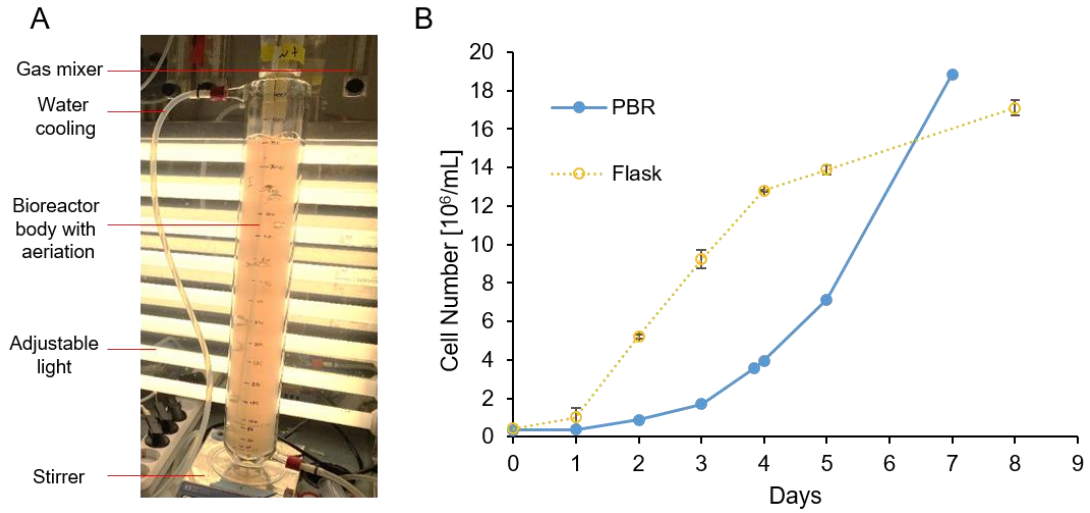


Figure 4: Cultivation of *P. purpureum* in a custom-made photobioreactor.
A: Setup of the custom-made photobioreactor cultivation system. The bioreactor body has a filling volume of 800 mL and was aerated with a 2% CO_2/air mix at an airflow of 50 L/min. The medium was mixed with a magnetic stirrer and by the convection of the air being released at the bottom of the PBR. The body was cooled to 25 °C using a water-cooling system and the light was adjusted to 150 µmol photons $m^{-2}s^{-1}$ light intensity.
B: Exemplary growth of *P. purpureum* wild type in the PBR (blue line) compared to growth in 200 mL flasks with 40 mL filling volume sealed with cotton. After a longer lag phase and initially slower growth, cells in the bioreactor showed a higher cell number at the end of cultivation, reaching 18.86 x 10^6 cells/mL.

figure 4B is shown as an exemplary growth curve from one run in comparison to cells grown in flasks.

The PBR setup encompassed a double-walled glass vessel used as a cooling mechanism through which the PBR inner body that is surrounded by it could be cooled to the desired temperature of 25 °C using water. Within the body, a pipe extends to the base of the PBR tube, facilitating aeration and ensuring mixing of the culture through the upward movement of the air bubbles generating agitation. A two-way input system connected to a flow meter enables precise regulation of both CO_2 and air concentration. In this particular experiment (Figure 4B and Table 2), the CO_2 concentration was set to 2% (v/v) with a total airflow of 50 L/min. Additionally, a magnetic stirrer operating at 150 rpm facilitates mixing of the lower part of the PBR body. The light intensity was adjusted to 150 µmol photons $m^{-2}s^{-1}$ to provide sufficient illumination even in dense cultures.

The growth experiment started with the inoculation of 800 mL of lsASW medium within the PBR at a starting cell density of the culture of 3 x 10^5 cells/mL. Simultaneously, 200 mL flasks sealed with cotton plugs were inoculated with 40 mL of culture at the same cell density and subjected to standard growth conditions. According to the findings in section 4.1.1, the cultures in the flasks showed the expected growth pattern, with the cotton-sealed culture exhibiting rapid growth and reaching a high cell concentration. In contrast, the cells cultivated in the bioreactor initially displayed a prolonged lag phase and only began to accelerate their growth on day three. When comparing the doubling time during the exponential phase, the bioreactor-grown cells exhibited a doubling time of 27.6 h, whereas the cotton-sealed culture had a doubling time of 19.6 h.

The cell concentration in the PBR on day seven reached 18.86 x 10^6 cells/mL, surpassing the final cell concentration of the flask-grown culture, which was 17.16 x 10^6 cells/mL, on day eight. The experiment in the PBR was terminated on day seven due to significant foam formation, which caused the culture to overflow from the bioreactor vessel. In addition to cell number, the dry biomass was monitored.

Table 3: Measurement of biomass from algal cultures grown in the PBR.

Day	2	3	4	5	7 (n=2)	Final concentration excl. exopolysaccharides (n=2)
Dry weight [mg/L]	790	1033	1143	1457	2504±362	1460 ±352
Biomass productivity [mg/L/d]	395	243	110	314	357.71±51.72	208.48±35.62

To assess the biomass, 10 mL from the PBR-grown culture was harvested daily following the protocol described in 3.1.3, and the dry biomass was determined. In short, the 10 mL culture

samples were filtered through a 0.2 µm membrane using vacuum, the retained biomass was dried overnight at 65°C and the weight difference of the membrane was measured. This value indicates dry biomass including exopolysaccharides that were collected on the membrane as well. At the end of the cultivation period on day seven, the remaining volume in the bioreactor was harvested, washed once with lsASW, freeze-dried, and the dry biomass of the culture was determined. In this case, the exopolysaccharides were washed off and the value reflects the dry cellular biomass. The results are presented in Table 3. The increase in dry weight accumulation, including the exopolysaccharides, was linear during the cultivation duration, reaching 2504 mg/L on day seven. Biomass productivity was calculated as described in method section 3.1.3, as the difference of daily-generated biomass divided by the days, yielding an overall average of 276.7 mg/L/day for biomass inclusive of exopolysaccharides. After the removal of exopolysaccharides through washing of the cells, which should leave only the cellular biomass, the dry weight reduced to 1460 mg/L leading to a productivity of 208.48 mg/L/day. These findings suggest that approximately 1 g/L can be attributed to exopolysaccharides.

4.1.4 Single colony formation on solid medium through matrix embedding

Being able to obtain single colonies from a culture after plating is a crucial step in handling microorganisms in the lab. This method is commonly employed to produce individual colonies, with the progeny of a colony usually being genetically uniform. Such isolation methods are valuable for eliminating unwanted phenotypes, achieving uniform expression levels, or isolating independent mutants. Unfortunately, the conventional plating technique of diluted *P. purpureum* cultures onto lsASW agar plates did not yield single colonies. When a low number of algal cells was plated, no growth, i.e., no colony formation on the plates, was observed. Plating a higher number of algal cells, resulted in the formation of a lawn of algal cells instead of the growth of distinct colonies.

In order to achieve single colony formation, two different methods were tested: the starch plating method as described for *Emiliania huxleyi* by Skeffington et al., 2019, and a method where cells are embedded in a low-concentration agar phase.

For the starch plating method, a sample of 114 cells in 200 µl lsASW was either directly plated on lsASW agar plates or mixed with 1 mL of 0.5 g/mL wheat starch in lsASW before plating on lsASW agar plates (Figure 5A). The plated suspensions were dried in a sterile bench, plates were closed with Leukopor® and incubated for two weeks under 100 µmol photons $m^{-2}s^{-1}$ continuous light. In theory, such a diluted sample should result in the growth of 114 colonies. However, for the solid cultures, where cells were spread on plates without the addition of starch, no growth of any colony could be observed (referred to as plain plating). In contrast, 99 algal colonies formed on the plate where cells were plated using the starch method

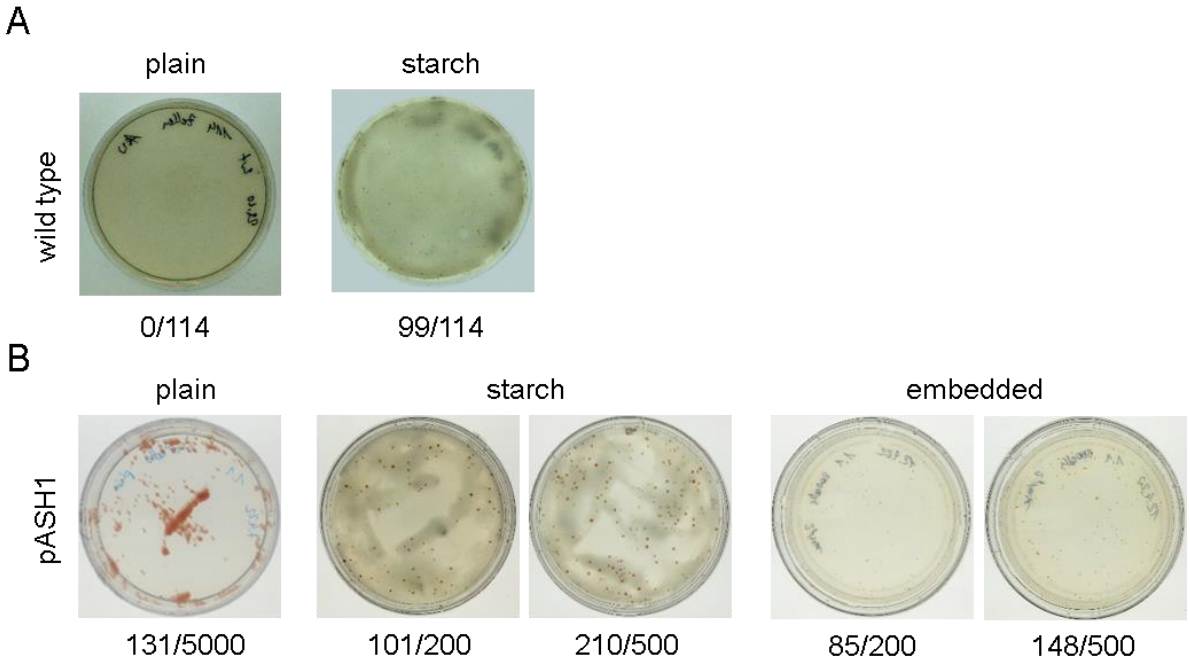


Figure 5: Efficiency of colony formation from single cells with different plating methods.
A: A dilution equivalent to 114 wild-type cells was either plated directly on lsASW agar plates (plain) or mixed with 1.2 mL starch suspension (0.5 g/mL) and then plated on lsASW agar plates (starch). No formation of colonies was visible after two weeks of incubation under continuous illumination (100 µmol photons $m^{-2}s^{-}1$) upon direct plating, whereas approximately 86 % of the cells were recovered with the starch plating method.
B: A transgenic strain was plated with different methods on (zeocin-containing) lsASW agar plates, and the efficacy of colony recovery was calculated. The directly plated cells (plain) formed a local lawn (presumably due to cell aggregation) and showed only a low number of distinct colonies. The starch method yielded an efficiency of colony formation of approximately 46%, and the embedding method, where plated cells are embedded in a thin film of 0.2% agar on top of the lsASW agar yielded an efficiency of colony formation of around 36%

(referred to as starch), resulting in a plating efficacy of 86 %. The experiment was repeated with a transgenic line that was spread on lsASW plates containing 25 mg/l Zeocin™ as a selective agent (Figure 5B). In this case, 5000 cells were used for the plain plating method, while 200 and 500 cells were used for the starch and the gel-embedding methods, respectively. Plating such a high number of cells using the plain method resulted in the formation of regions with a cell lawn and indistinct colonies that were impossible to be counted and picked. Excluding the cells in the lawn, approximately 131 colonies formed from 5000 cells using this method. In contrast, the starch-embedding plates exhibited the growth of well-defined and easy-to-pick colonies, with 101 colonies out of 200 cells and 210 colonies out of 500 cells spread, resulting in a colony formation efficacy of 46%. For the agar-embedding method, 200 µl of diluted culture containing the cell numbers as indicated in Figure 5B were mixed with 1.5 mL of 0.4% agar solidified lsASW medium and spread on lsASW agar plates. After the agar solidified, the plates were sealed with Leukopor® and incubated for two weeks under standard light conditions. This method also facilitated the formation of distinct colonies,

although they were not as easy to pick as those obtained with the starch method. A total of 85 colonies from 200 cells and 148 colonies from 500 cells could grow, yielding a colony formation efficacy of approximately 36%.

In conclusion, the addition of a matrix enclosing the algal cells to the culture sample promoted the formation of single colonies and was especially effective in the case of starch. This method was subsequently used to generate single-cell-derived colonies from transformed strains exhibiting heterogeneous transgene expression within a population, where some cells expressed the transgene and some cells did not.

4.1.5 Transformation efficacy is improved by washing cells before plating on selection media

The development of genetic tools for red algae, specifically *P. purpureum*, is still in its early stages. While numerous methods are readily available for the widely studied microalga *C. reinhardtii*, they cannot be directly applied and need to be adapted for use in *P. purpureum*. Two transformation methods have been reported in the literature, with the method described by Li and Bock (2018) being the most reliable. However, in my own experiments, this episomal transformation method using particle bombardment did not produce satisfactory results, often resulting in only one or two positive transformants per transformation experiment. The limited transformation efficiency of *P. purpureum* became a significant bottleneck for subsequent experiments. Various optimization strategies were explored to enhance transformation efficacy, and the most promising approach is presented here.

The initial transformation protocol, based on the method described by Li and Bock, involved plating $1x10^8$ cells from an exponential-stage culture on lsASW plates supplemented with 25 mg/L Zeocin™ as selective antibiotic. Gold particles coated with 1 µg of circular transformation plasmid were then bombarded onto the selection plates at a distance of 9 cm and a pressure of 1350 psi. Following bombardment, the transformation plates were incubated overnight in darkness and subsequently exposed to continuous light of 100 µmol photons $m^{-1}s^{-1}$ for three weeks until the first colonies appeared. To ensure an adequate number of transformants, for each plasmid construct, five transformations were performed. However, there were instances where no or only a single resistant colony were obtained. Consequently, the transformation protocol was modified, as illustrated in Figure 6, in an attempt to enhance resistant colony formation.

In the modified protocol, Zeocin™ was not added to the bombardment plates. After bombardment, the cells were washed off the plate and subjected to two washing rounds with lsASW medium to remove cell debris. The resulting supernatant, depicted in Figure 6C,

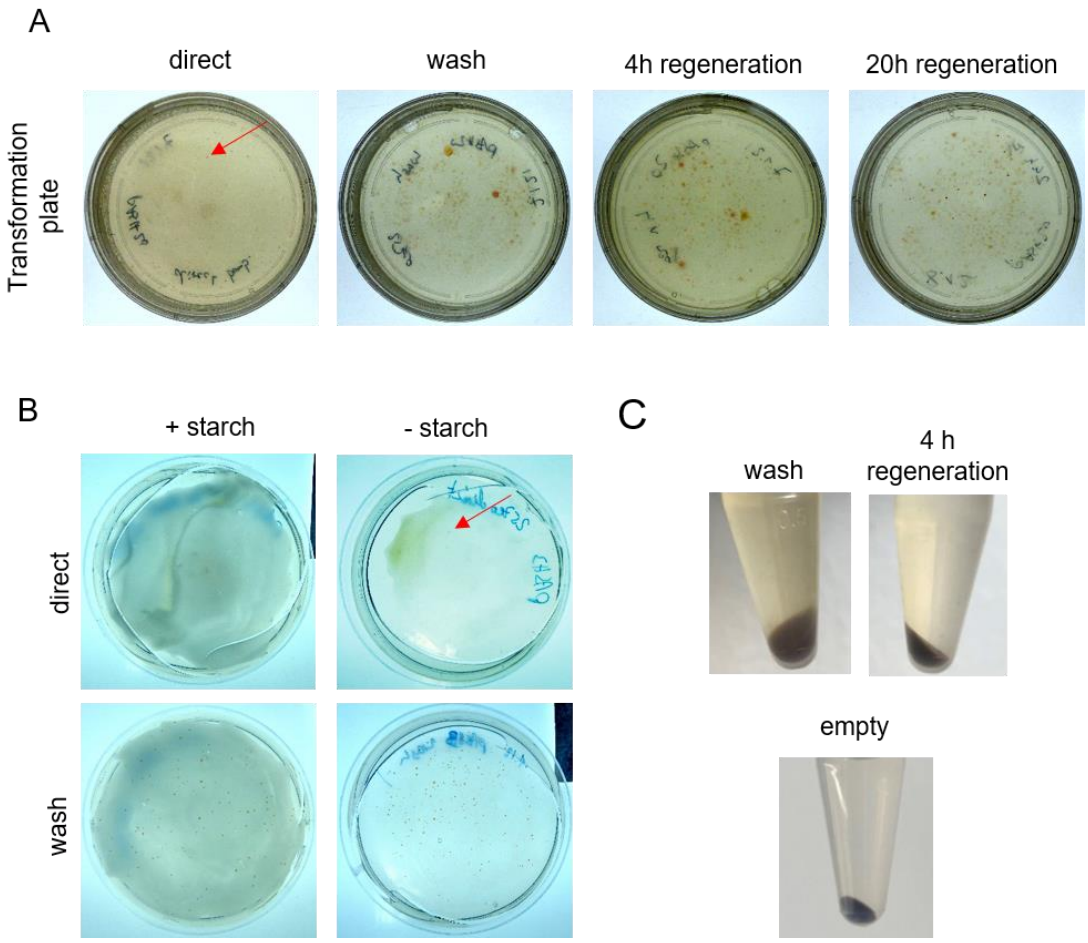


Figure 6: Development of an improved transformation protocol for *Porphyridium purpureum*
A: Comparison of the previously used transformation protocol (Li and Bock, 2018) with the new protocol. In the previously established protocol, *Porphyridium* cells are directly plated on zeocin-containing lsASW plates, bombarded with plasmid vector-coated gold particles and regenerated on the same plate (labeled "direct"). In the new protocol, cells are bombarded on lsASW plates without antibiotics, then washed off the plates, washed twice with lsASW, and then plated either directly on zeocin-containing plates (wash), or after a recovery period in liquid medium of 4 or 20 hours ("4 h regeneration" and "20 h regeneration", respectively). Note that the wash alone is sufficient to greatly increase the number of colonies forming after 3 weeks.
B: The same protocol was also tested for cells plated with starch, to possibly further support colony formation. While the starch had no significant influence on colony formation after transformation, the wash substantially increased the number of transformants recovered.
C: Supernatant after the first wash with lsASW following biolistic bombardment without a recovery phase ("wash") and after a recovery phase of 4 h ("4 h regeneration"). The greenish color of the supernatant indicates broken cells that are washed away by the procedure. As a control, pelleted unbombarded cells are shown ("empty").

exhibited a greenish color, indicating the release of chlorophyll and cell rupture, in contrast to the supernatant of the control where cells were 'bombarded' without gold particles (empty). Following the wash, different regeneration strategies were employed. In the first approach, the washed cells were directly plated on Zeocin™-containing selection plates without any recovery phase, incubated overnight in darkness, and then subjected to standard light conditions for three weeks (Figure 6A, wash). In the second approach, the washed cells were regenerated in liquid lsASW medium under standard conditions for either 4 hours or 20 hours.

After the regeneration round, the cells were collected and spread onto Zeocin™-containing selection plates followed by incubation for three weeks under standard conditions. As shown in Figure 6A, direct plating of bombarded cells on the selection plates yielded only one small resistant colony, while the cells that were washed, or washed and regenerated led to the growth of numerous resistant colonies. No difference between washing only, or washing and regeneration could be observed in terms of resistant colony formation.
The colonies varied in size. Further analysis by PCR and picking into Zeocin™-containing lsASW medium revealed that all large colonies were true transformants, while some small colonies died after being transferred to liquid medium containing Zeocin™, suggesting that these colonies were no true transformants.
Another approach aimed at enhancing transformation efficiency at the level of colony formation in dependence of the plating technique, as described in section 3.1.5. To this end, cells were mixed with starch and plated on Zeocin™-containing transformation plates before bombardment (Figure 6B, direct starch). The results were compared with the original transformation protocol without starch (direct, no starch). Alternatively, transformed cells underwent the washing method as described previously, but were mixed with starch before plating on Zeocin™-containing plates (Figure 6B, wash starch). The results showed that starch did not improve transformation efficiency in either method. Only the washing of the cells after bombardment led to colony formation, and there was no difference observed between starch and no starch after washing. Moreover, no improvement was observed after direct plating, where only one colony appeared on the plate without starch, compared with several dozen on the plates with the washed cells.

The results clearly demonstrated that bombardment on Zeocin™-free plates and washing off debris after transformation significantly improved transformation efficacy and led to an increased number of colonies on the plates. This improvement facilitated the isolation of an adequate number of transformants, and thereby eliminated the transformation step as a bottleneck for subsequent investigations. Comparison of different optimization approaches revealed that an additional regeneration step in liquid medium did not enhance colony formation compared to the washing of the cells after bombardment alone. Consequently, the use of non-selective plates for the bombardment followed by transferring washed cells to selection plates became the standard method for nuclear transformation of *P. purpureum*. An additional regeneration step after cell washing was omitted in further transformation experiments.

4.2 Expression analysis of different *YFP* gene variants in *Porphyridium*

4.2.1 Determining the codon usage of *P. purpureum* and design of *YFP gene* variants

Codon optimization of heterologous transgenic DNA has been identified as one of the most important steps in order to express recombinant proteins successfully in various organisms. A comprehensive analysis of the impact of codon usage on transgene expression in *Chlamydomonas* was conducted by Barahimipour et al., 2015, dissecting the impact of codon usage and GC content. To assess if the same principles apply to *P. purpureum*, a similar approach was pursued.

In this study, gene variants encoding the identical amino acid sequence of a yellow fluorescent protein (YFP) while differing in their synonymous codon use, were cloned under the control of the promoter and terminator sequence of the endogenous A*ctin* gene. The expression levels of the different *YFP* gene variants was systematically determined and then compared in the different *Porphyridium* transformants (Figure 7A). Since no information on codon usage was available for *P. purpureum*, an analysis of the coding sequence of the top 100 most highly expressed genes in this alga was performed (see Table 2 and Supplementary Tables 1 and 2, and 4.2.4). To this end, existing RNAseq data was re-analyzed (for details refer to 4.2.4). From those top 100 most highly expressed genes, 92 genes do not contain an intron sequence and those were chosen for further analysis. The codon usage was determined based on a total of 81,120 bases (27,040 codons) using the "codon usage" webtool of the "Sequence Manipulation Suite" (bioinformatics.org/sms2/codon_usage). The resulting codon usage table (Supplementary Table 1) was imported into the CodonWorkbench software (developed by Dr. Mark Lohse, www.buba-basis.de/software/cwb/cwb.html) for further analysis and codon optimization of transgenes. To ensure full codon adaptation of the transgenic *YFP* (referred to as *PpYFP*) to the nuclear genome of *P. purpureum*, the optimal codon for each amino acid in the transgene was selected (see Table 4). Hence, the *PpYFP* gene variant has 100% relative codon adaptation (RCA) to the nuclear genome.

Table 4: Overview of the most frequently used codons for each amino acid in *Porphyridium* and *Chlamydomonas.* Bold amino acids differ between the two algae.

Amino acid	Best codon *Porphyridium*	Best codon *Chlamydomonas*
Ala	**GCG**	**GCC**
Arg	CGC	CGC
Asn	AAC	AAC
Asp	GAC	CAC
Cys	UGC	UGC
Glu	GAG	GAG
Gln	CAG	CAG
Gly	GGC	GGC

Amino acid	Best codon *Porphyridium*	Best codon *Chlamydomonas*
His	CAC	CAC
Ile	AUC	AUC
Leu	CUG	CUG
Lys	AAG	AAG
Met	AUG	AUG
Phe	UUC	UUC
Pro	**CCG**	**CCC**
Ser	**UCG**	**AGC**
Thr	**ACG**	**ACC**
Trp	UGG	UGG
Tyr	UAC	UAC
Val	GUG	GUG
Stop	UAA	UAA

Additionally, the *YFP* variant sequences, namely *cpYFP*, *laYFP*, *vYFP*, and *CrYFP*, as described in the study of Barahimipour et al., 2015, were subjected to in silico analysis to determine their RCA to the nuclear genome of *P. purpureum* (Figure 7B). The *cpYFP* variant represents a codon-optimized version for the *Chlamydomonas* chloroplast genome, exhibiting an average RCA of 41.2% and a minimal RCA of 6.7% to the nuclear genome of *Porphyridium*. The *laYFP* variant designed as a low-adaptation *YFP* gene with the same GC content as the fully adapted gene, but with poor adaptation to the nuclear codon usage of *Chlamydomonas*, had an average RCA of 84.3% and a minimal RCA of 28.0%. The *vYFP* variant was initially designed for YFP expression in mammalian cells and is GC-rich but harbors three lowly adapted codons resulting in an average RCA of 89.9% and a minimal RCA of 14.5%. Finally, *CrYFP* is the fully codon-optimized gene variant to the codon usage of the *Chlamydomonas* nuclear genome and has an average RCA of 94.3% and a minimal RCA of 50.5% to the nuclear genome of *Porphyridium*.

Those variants were introduced into *P. purpureum* and the expression levels of the different constructs were compared.

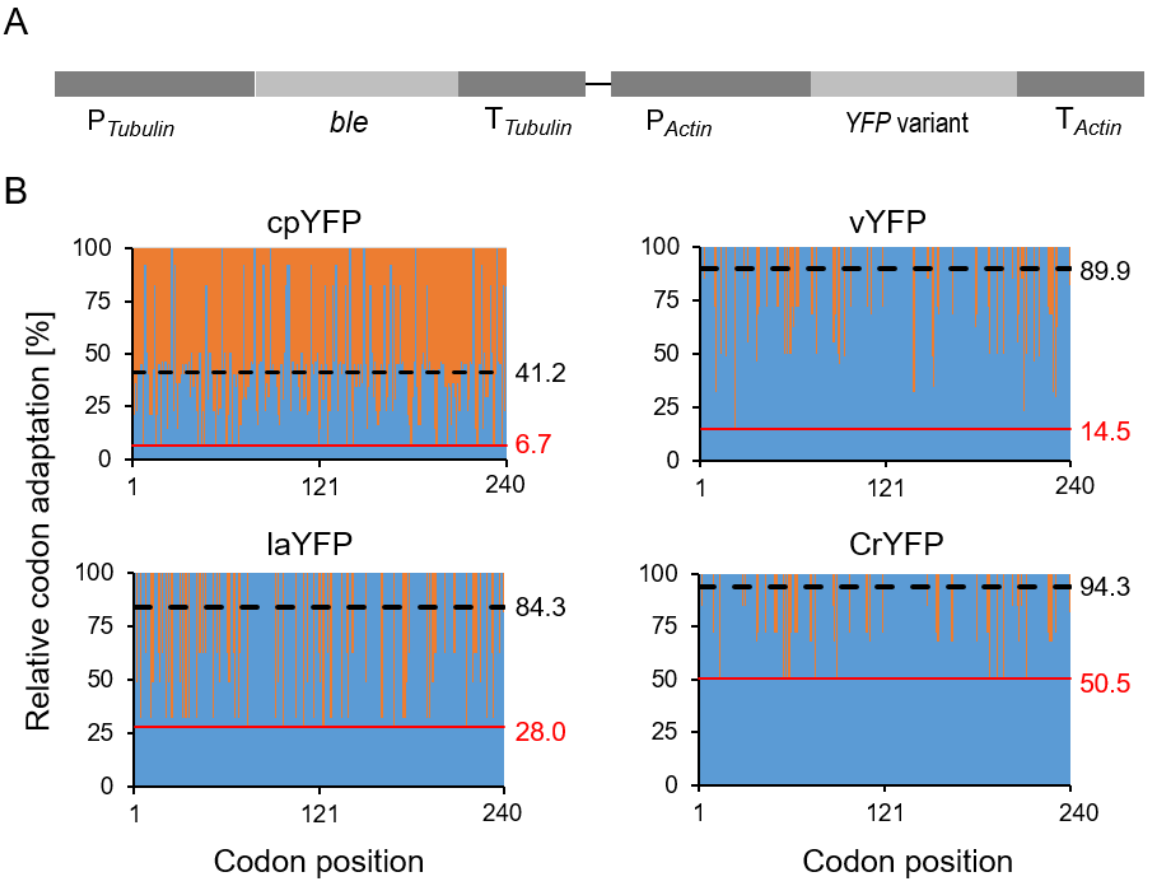


Figure 7: Physical map of the basic expression vector used for transformation of *P. purpureum* and codon adaptation of the *YFP* gene variants tested for nuclear expression.
A: Physical map of the transformation vector used for expression of the *YFP* gene variants in *Porphyridium.* The zeocin resistance gene *ble* was placed under the control of the tubulin gene promoter and 5' UTR ($P_{Tubulin}$) and the *Tubulin* gene terminator ($T_{Tubulin}$). The different *YFP* gene variants were expressed under the control of the *Actin* gene promoter and 5' UTR (P_{Actin}) and the *Actin* gene terminator (T_{Actin}).
B: *YFP* gene variants used in this study and their relative codon adaptation (RCA) compared to the codon usage in the nuclear genome of *Porphyridium* (see Supplementary Table 1). The RCA was plotted against the codon position, and blue bars indicate the relative adaptation of each codon (in %). The black number at the right indicates the average adaptation and the red number the minimal adaptation of each gene variant. The fully codon-optimized variant for *Porphyridium* (PpYFP) is not shown here, because, in this gene, all RCAs are at 100%.

4.2.2 Expression analysis of *YFP* variants in *P. purpureum* and comparison to expression in *C. reinhardtii*

After transformation of *P. purpureum* with the *YFP* gene variants, two independent transformed strains per gene variant were isolated and further analyzed. All lines analyzed here are uniformly expressing transformants, meaning that all cells within the cell population exhibited YFP fluorescence. For some of the transformants, however, inhomogeneous YFP expression pattern within the cell population had been observed when they were initially isolated after transformation. *vYFP*, *laYFP* and *CrYFP* transformants showed a heterogeneous expression pattern, where often only around 50% of the cells within a

population showed fluorescence. As this hampers the comparative expression analysis, it was necessary to first investigate what was the cause of this inhomogeneous expression pattern within certain transformed strains. As an approach to homogenize the expression level within

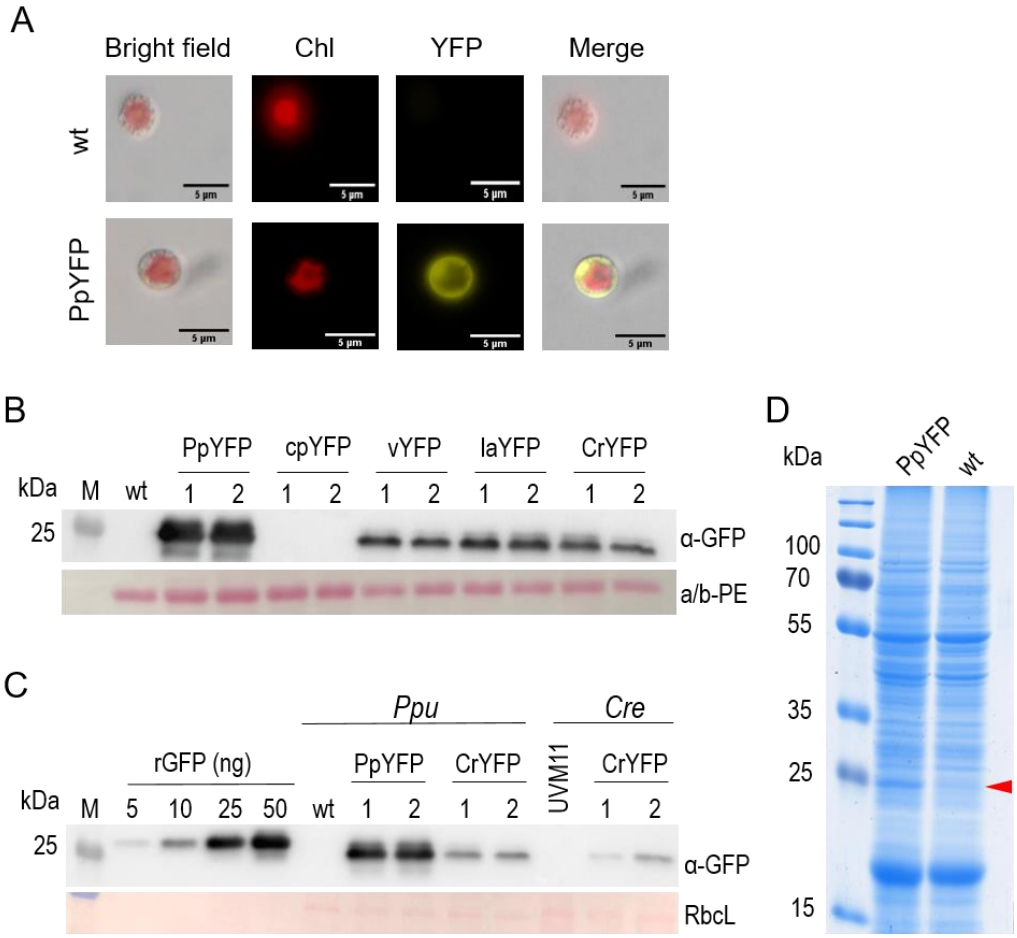


Figure 8: Protein expression from the different *YFP* variants and comparison to the green microalga *C. reinhardtii*

A: Subcellular localization of the cytosolic PpYFP in a *P. purpureum* cell. The red chlorophyll fluorescence and the merged image show the single plastid present in P. purpureum cells. A wild-type cell is shown as negative control.

B: Immunoblot analysis to compare YFP accumulation in transgenic algal strains. Samples of 5 µg total soluble protein (TSP) extracted from transformed *Porphyridium* clones expressing the different gene variants were electrophoretically separated, blotted and immunodecorated with anti-GFP antibodies. Note that the perfectly codonoptimized gene version (PpYFP) results in the highest protein accumulation levels, whereas the poorly adapted AT-rich variant (cpYFP) shows no detectable expression. Ponceau staining of the α/β-phycoerythrin band (a/b-PE) on the blotted membrane served as loading control. M, molecular weight marker; wt, wild-type *P. purpureum*.

C: Semi-quantitative immunoblot analysis to determine YFP expression levels from fully codon-optimized gene variants in *P. purpureum* (PpYFP) and *C. reinhardtii* (CrYFP).1 µg TSP was loaded for *P. purpureum* and *C. reinhardtii* and compared to a dilution series of recombinant GFP (Roche). YFP accumulates in Porphyridium to approximately 5% of TSP from the gene fully codon optimized for *P. purpureum* (PpYFP) lines, and to approximately 1.2% from the gene codon optimized for *C. reinhardtii* (CrYFP) lines. *Chlamydomonas* expression strain UVM11 accumulates the CrYFP to approximately 0.8% of TSP. Ponceau staining of the region of the blotted membrane containing the large subunit of Rubisco (RbcL) served as loading control.

D: Visualization of PpYFP accumulation in P. purpureum by Coomassie staining of 10 µg electrophoretically separated total soluble protein. The YFP can be seen as a distinct band at 25 kDa in the PpYFP transgenic strain.

a transformant, single cells were generated as described in 3.1.5 and results for this are described in 4.2.5. Uniformly expressing lines were then selected for further analysis, which is described in this section.

Figure 8A depicts the subcellular localization of the fully codon-optimized *PpYFP* variant. The YFP fluorescence is mainly located in the cytosol, surrounding the star-shaped chloroplast (red).

Clear differences in expression strength from the various *YFP* gene variants could be seen by immunodetection of YFP (Figure 8B). The AT-rich *Chlamydomonas* plastid codon optimized *cpYFP* showed no detectable accumulation of YFP, while the *P. purpureum* fully codon-optimized variant *PpYFP* showed the strongest expression, followed by *laYFP*, *CrYFP* and *vYFP* with lower but similar expression values. For side-by-side comparison of expression levels of the fully-codon optimized *YFP* variants from *Porphyridium* and *Chlamydomonas,* protein samples from *Porphyridium* expressing PpYFP and from *Chlamydomonas* strain UVM11 expressing CrYFP were loaded next to each other and next to a GFP standard to allow absolute quantification. When comparing the YFP accumulation levels of the perfectly codon-optimized *YFP* variants in the two algal species *Porphyridium* (*PpYFP*) and *Chlamydomonas* (*CrYFP*), it turned out that the red alga expressed up to six times more YFP than the green alga (Figure 8C) with YFP accumulation levels being 5% (*Porphyridium*) and 0.8% (*Chlamydomonas*) of the total soluble protein (TSP). Even the imperfectly codon-optimized *CrYFP* variant introduced in *P. purpureum* led to higher YFP accumulation levels reaching 1.2% TSP compared to *CrYFP* transformants of *Chlamydomonas*. It is notable that YFP expression in *C. reinhardtii* UVM11 strain was driven by the strong *PsaD* promoter, whereas the expression in *P. purpureum* was driven by the endogenous *Actin* promoter. Further evidence of the strong expression capacity of *Porphyridium* is shown in Figure 8D, where 10 µg TSP of the wild type and a *PpYFP* transformant were separated in an SDS gel, and where the YFP protein could be detected as a clearly visible band at a size of 25 kDa after Coomassie-staining of the gel in the *P. purpureum* transgenic strain.

This experiment clearly demonstrated that the capacity of *P. purpureum* for expression of YFP is much higher than that of *C. reinhardtii*. Not only the strong expressor strain UVM11, but also the strong *PsaD* promoter was used to ensure the highest possible expression in the green alga. Nevertheless, *C. reinhardtii* accumulated 6 times less YFP than *P. purpureum*, utilizing the endogenous *Actin* promoter.

4.2.3 Transcripts from the non-codon optimized gene variants accumulate to much lower levels and seem to be partially degraded

Besides YFP protein accumulation, also the *YFP* transcript levels were analyzed in the strains expressing the different *YFP* variants. For RNA gel blot analysis, total RNA was isolated from the same biological samples that were used for western blot analysis and northern blot analysis was performed. To facilitate simultaneous detection of all five *YFP* transcript variants, a probe targeting the *Actin* 5' untranslated region (UTR) was generated, as this region was present in all transformation vectors (see Figure 9A). To ensure equal loading and RNA transfer to the membrane, the RNA gel was stained with ethidium bromide, while the membrane was stained with methylene blue. Following hybridization, two bands at 1.5 kb and 1.1 kb were detected, corresponding to the endogenous Actin mRNA (calculated to be 1473 nt) and the transgenic *YFP* mRNA (calculated to be 1078 nt). The calculation was performed from the putative translation start site to the putative poly-adenylation signal on the terminator. In general, the abundance of *YFP* mRNA mirrored that of the YFP protein, with *PpYFP* mRNA

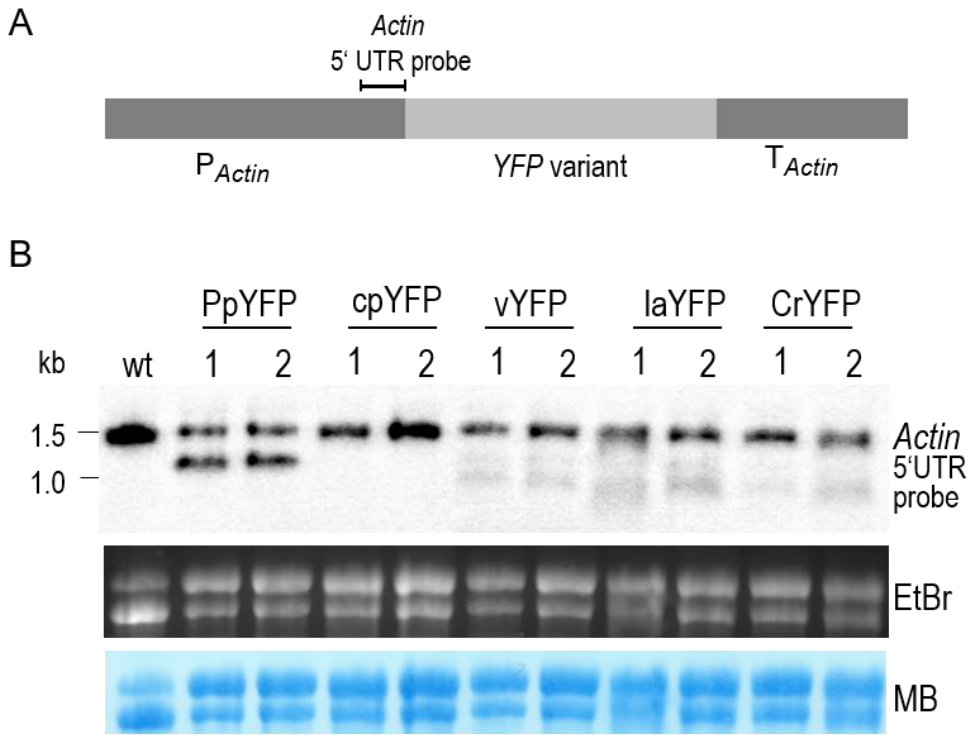


Figure 9: Northern blot analysis of *YFP* mRNA accumulation from the various gene variants expressed in *P. purpureum*.
A: Position of the *Actin* 5' UTR probe designed to (i) simultaneously detect all five *YFP* variants at equal sensitivity, and (ii) detect the endogenous actin mRNA as internal control.
B: RNA gel blot analysis of *YFP* transcript accumulation. Samples of 4 µg of total cellular RNA extracted from the same transgenic strains analyzed by western blotting (Figure 7) were electrophoretically separated in a 1 % denaturing agarose gel, blotted and hybridized to the radioactively labeled actin 5' UTR probe indicated in panel (a). While the fully codon optimized *YFP* variant shows mRNA accumulation (~1.1 kb signal) to levels comparable to those of the endogenous actin transcript (~1.5 kb signal present also in the wild type; wt), all other variants show low transcript accumulation and evidence of partial mRNA degradation. Equal loading was verified by UV light detection of rRNA bands using ethidium bromide (EtBr), and blotting efficacy was checked by membrane staining with methylene blue (MB).

being most abundant, and no detectable mRNA for *cpYFP*. Notably, all other transformed lines exhibited a truncated version of *YFP* mRNA in addition to the full-length *YFP* mRNA, as indicated by an additional band at approximately 900 nt. This shorter mRNA variant appeared to be the predominant form in the non-fully optimized *YFP* lines, as evidenced by its higher band intensity, while the full-length mRNA exhibited relatively low accumulation and indicates impaired mRNA stability.

4.2.4 Identification of new endogenous promoters to drive strong transgene expression in *P. purpureum*

An important feature of biotechnological platforms is the availability of different promoters and 5'UTRs that can drive transgene expression and confer different expression strengths. Until now, in *P. purpureum* only two promoters from the endogenous genes encoding Actin and Tubulin, respectively, were used for transgene expression. In most organisms, the promoters of the housekeeping genes A*ctin* and *Tubulin* are medium-strength promoters and not used for recombinant protein production. In order to identify new strong endogenous promoter elements in *P. purpureum*, an analysis of RNAseq data was performed based on published data (Ferrari et al., 2019). For this, existing RNAseq datasets (SRA Accession ERP111278**,** un-stranded single end data) were analyzed to determine highly transcribed genes in *P. purpureum* using the Galaxy (Afgan et al., 2018) platform and a custom-made pipeline. For generation of this RNAseq dataset, cells were grown in ASW medium at 25°C in a 12 hour light/12 hour dark photoperiod at 100 µmol photons $m^{-2}s^{-1}$. The cells were collected every two hours in triplicates, with the first sample collected one hour after the onset of the light. For the analysis in this work, samples taken after 5 h, 7 h, and 9 h of light were pooled to a total of 9 replicates. The updated genome of *Porphyridium* strain CCMP1328 was used as reference genome for mapping (NCBI Accession VRMN01) (Lee et al., 2019). All programs used for RNAseq analysis were part of the Galaxy (see Material and Method section).

The first step in our analysis pipeline was the quality control of the FASTQ data using the FastQC tool. This was followed by mapping of RNAseq reads with RNAStar using standard settings. The mapped reads were counted with FeatureCount and finally normalized over all replicates to produce the final count using DESeq2. Transcripts per million (TPM) were calculated as described previously (Wagner et al., 2012). As in our analysis, the differential expression of genes was not of interest, but the most highly accumulating transcripts over all nine replicates were to be identified, genes showing a significant difference in expression over those 3 sampling time points were omitted. Genes with the top counts after normalization were blasted against the non-redundant protein sequence database and, if no hit was obtained, a conserved domain search was performed. Of the top 100 most highly transcribed genes from the final table, 92 intron-less genes were used to determine the codon usage in *Porphyridium*

Table 5: The top five most highly transcribed genes in *P. purpureum*, whose promoters and 5'UTRs were used in a subsequent *YFP* expression analysis. The gene products are ranked according to their transcript accumulation (given in transcripts per million reads). Actin and Tubulin are shown as reference values.

Gene product	Transcripts per million (TPM)
Formate/nitrite transporter	14407.3
Hypothetical (PsbQ domain) containing protein	1483.6
PYP1	13274.2
Carbonic anhydrase	10894.8
Chlorophyll a binding protein, chloroplast targeted	9631.1
Actin	1190
Tubulin	320

in order to perform codon optimization of transgenes as described in the previous chapter (Suppl. Table 2, 4.2.1). Furthermore, the five highest-ranked genes were selected to experimentally evaluate their promoter strength (Table 5). However, the gene encoding a hypothetical protein FVE85_6371, which occupied the fifth rank, could not be cloned and was therefore replaced with the gene encoding the chlorophyll a binding protein, ranked sixth. To provide a reference value, the TPM values of Actin (1190) and Tubulin (320) were included. Interestingly, the TPM values of the most highly transcribed genes were up to 12 times higher than that of Actin.

The promoters and 5'UTRs of the genes listed in Table 3 were cloned upstream of the codon optimized *PpYFP* with the *Actin* terminator sequence downstream. Depending on gene

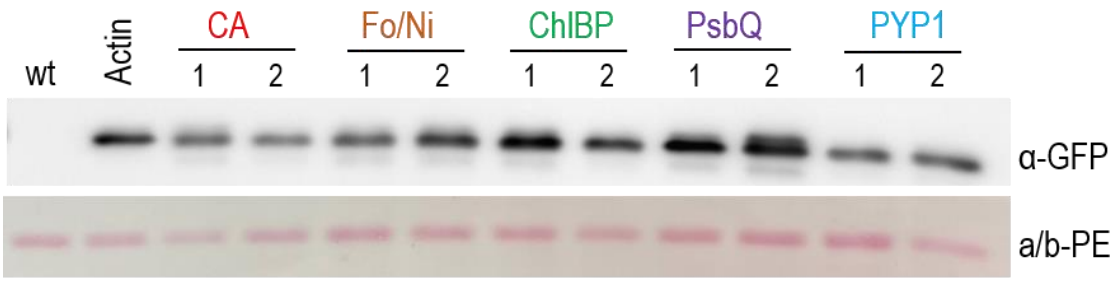


Figure 10: Analysis of the expression capacity of the five endogenous promoters that confer the highest mRNA accumulation levels in *P. purpureum*.
Accumulation of YFP under the control of the five strong endogenous promoters in comparison to the *Actin* promoter. Samples of 2 µg total protein were separated by gel electrophoresis, and YFP accumulation was examined using the identical YFP transgene controlled by the *Actin* promoter as reference. Small increases in YFP accumulation are observed in the transgenic strains expression YFP under the control of the chlorophyll a/b-binding protein promoter (ChlBP) or the hypothetical PsbQ domain protein promoter (PsbQ). Ponceau staining of the α/β-phycoerythrin band (a/b-PE) on the blotted membrane served as loading control.

density, 500 bp (PsbQ), 1000 bp (Fo/Ni, ChlBP and PYP1) or 1500 bp (CA) before the transcription start site were used. Accumulation levels of YFP were assessed in the different promoter constructs using western blot analysis (Figure 10).

One transformed line of the *Formate/nitrite transporter* (*Fo/Ni*) promoter construct and one transformant line of the *PYP1* promoter construct showed inhomogeneous expression levels within the cell population, and single cell-derived subcultivation of those lines was necessary as described in 3.2.5 in order to obtain cultures with uniform fluorescence levels.

For immunoblot analysis, 2 µg of total protein from the newly generated promoter lines were loaded alongside an *Actin* promoter-driven YFP line as control. The analysis revealed a decrease in YFP accumulation in the *carbonic anhydrase* (*CA*) promoter lines as well as the *PYP1* promoter lines, while a slight increase in YFP accumulation levels was observed in the *ChlBP* promoter lines and the *PsbQ* promoter lines compared to the control line (*Actin* promoter). The *Fo/Ni* promoter lines exhibited similar YFP accumulation levels as the *Actin* promoter line.

4.2.5 Variation in transgene expression levels among independent transgenic strains of the same construct can be cured by single colony-derived subcultivation rounds under selection

The analysis of the codon usage with *YFP* variants and the new endogenous promoter lines showed that independent transformed lines generated by transformation with the same plasmid exhibited varying YFP accumulation levels. Microscopic analysis of individual transformed lines revealed that, within a cell population of a transgenic line, not all cells showed YFP fluorescence. The proportion of non-fluorescing cells to fluorescing cells was variable among the independent transformed lines, leading to the varying YFP expression values when comparing the transformants by western blot analyses (Figure 11A). It was hypothesized that a contamination with residual wild-type cells could be the reason for this observation. In the following, an in-depth analysis of the *Fo/Ni* promoter lines is exemplarily shown, and the strategy for obtaining transformed strains with uniform YFP fluorescence is explained. Three independent transformants carrying the *YFP* reporter gene under transcriptional control of the *Fo/Ni* promoter exhibited varying levels of YFP accumulation, which correlated with heterogeneous fluorescence patterns observed under the microscope (Figure 11A and B). Fo/Ni line 3 displayed 100% fluorescence in its population, while Line 1 exhibited 40% fluorescence and line 2 displayed less than 10% fluorescence (microscopic data only shown for Fo/Ni 1), matching the YFP accumulation seen in the western blot. The goal was to eliminate all non-fluorescing and probably wild-type cells to obtain a homogeneous culture and expression. Initially, Fo/Ni line 1 was subjected to fluorescence-activated cell

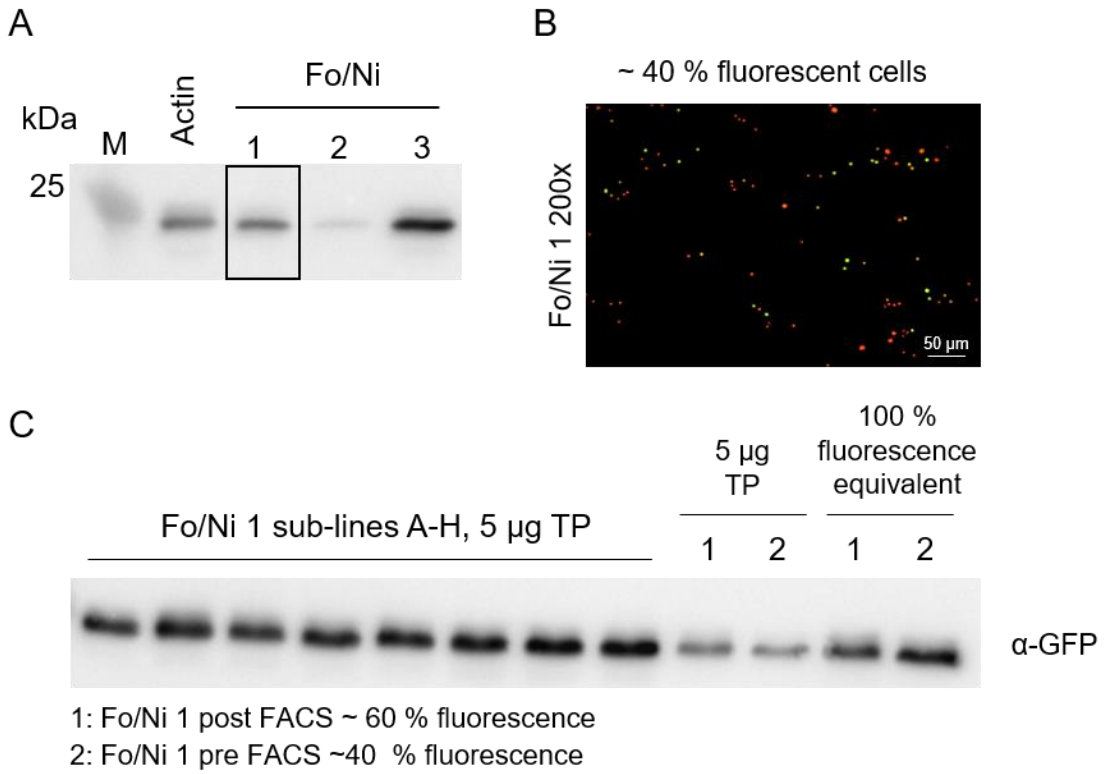


Figure 11: Analysis of the variation of YFP fluorescence in the Fo/Ni promoter lines
A: Immunoblot analysis of YFP accumulation in the Fo/Ni promoter lines compared to an *Actin* promoter line. The three independent transformants show varying YFP accumulation that correlates with the number of fluorescing cells in the population. Line 1 was selected for further analysis.
B: Fluorescence microscopy picture of Fo/Ni line 1. Green cells exhibit YFP fluorescence, while orange cells only exhibit auto fluorescence. Only 40% of the cells showed YFP fluorescence.
C: Fo/Ni line 1 was subjected to FACS sorting to sort fluorescent from non-fluorescent cells. By this, the proportion of fluorescing cells was increased to 60%. Single colony-derived subcultures of the sorted cells were generated and analyzed via immunoblotting for YFP expression. These subcultivated lines (A-H) show increased amounts of YFP accumulation, which corresponds to the equivalent protein amount of 100% fluorescent cells (5 µg TP 1 and 2: Fo/Ni 1 before and after FACS sorting; 100% fluorescence equivalent 1 and 2: increased total protein loading to match the hypothetical 100% fluorescence).

sorting (FACS) in an attempt to enrich fluorescent cells. However, this approach only resulted in an increase in the number of fluorescing cells to approximately 60%.

To achieve uniform YFP fluorescence of all cells within a population, single cells were generated using the starch plating method (4.1.4), and eight colonies were isolated (hereafter referred to as sub-lines) and analyzed by microscopy and immunoblotting. All eight *Fo/Ni* 1 sub-lines displayed 100% fluorescence and showed elevated YFP accumulation in the immunoblot compared to the progenitor line (Figure 11C). This indicates that, after subculturing on plates containing Zeocin™, the cell population within a line was purged of potential non-fluorescing wild-type cells left in the population. The progenitor line before and after FACS sorting was included in the analysis (5 µg TP, 1 = after FACS and 2 = before FACS), along with protein from the same sample loaded with adjusted amounts to mimic the signal intensity from a sample with100% fluorescent cells (i.e., 12.5 µg TP for pre-FACS cells

and 8.3 µg TP for post-FACS cells). The YFP accumulation level corresponds to the fluorescence ratio observed under the microscope and was comparable to *Fo/Ni* line 3 that exhibited 100% fluorescence. The post-FACS sorted *Fo/Ni* line displayed slightly higher YFP accumulation compared to the line prior to sorting.

Having demonstrated that the varying expression levels are mostly due to contamination with non-fluorescent cells, which possibly are wild-type cells having survived the selective agent Zeocin™or transformed cell that lost the plasmid during cultivation, the lines expressing the codon variants *vYFP*, *laYFP*, *CrYFP* and the *PYP1* promoter line were also subjected to the single colony-derived subcultivation. According to microscopic analysis, the liquid cultures of those lines initially showed a proportion of non-YFP fluorescing cells that varied from 15% (*PYP1* line 1) to 80% (*laYFP* line 2) before FACS sorting.

The lines were sorted by flow cytometry and the obtained cells were subjected to microscopic analysis. After FACS sorting, the percentage of fluorescent cells in the population was slightly increased, but still was not uniform. Afterwards, sub-lines were generated using the starch plating method and eight colonies per line were picked and analyzed under the microscope as described in 1.3.4 to determine the percentage of fluorescent cells (Figure 12). The total number of cells analyzed per transgenic line is given in

Suppl. Table **3**. The red horizontal bars in Figure 12 indicate the percentage of fluorescent cells of the progenitor lines used for single cell isolation, and the colored bars show the fluorescence percentages of the respective sub-lines. Only the *laYFP* progenitor lines 1 and 2 reached almost homogeneity after FACS (96.1% and 97.95%, respectively) and continued to exhibit a high fluorescence percentage throughout all sublines with some reaching 100%. All other lines showed an increased average fluorescence percentage after the first round of subculturing. Especially *CrYFP2* sub-line 8 and *PYP1* sub line 3 showed almost 100% fluorescence after the first subcultivation round. Sublines indicated with black bars were chosen to undergo a second round of single colony-derived subcultivation to further increase the ratio of fluorescing cells. This analysis showed that lines that already had a high fluorescent cell proportion, like *laYFP* 1 and 2, *CrYFP* 2_8 and *PYP1* 1_3, consistently displayed high ratios in all following sublines, indicating efficient elimination of wild-type cells from the population. The fluorescence ratio of all lines could be elevated to at least 90% fluorescence (*CrYFP*1_5_2) except for *PYP1* line 2, where the fluorescent percentage even decreased in the second round. This might have been caused by instability of the plasmid in this line in that it was diluted out during subculturing or the selective agent could mistakenly have been forgotten during subcultivation, resulting in enhanced growth of the wild-type cells in the population.

The lines exhibiting the highest fluorescent cell ratios were chosen for the analyses described in 4.2.2 and 4.2.4, namely *vYFP1*_6_7, *laYFP1*_6_3, *laYFP2*_2_5, *CrYFP2*_8_2 and *PYP1*

2_3_8 with 100% fluorescence, respectively. *vYFP2*_6_5 reached 96% fluorescence and *CrYFP1*_5_8 reached 90% fluorescence. *PYP1* 2 was omitted and replaced by another *PYP1*

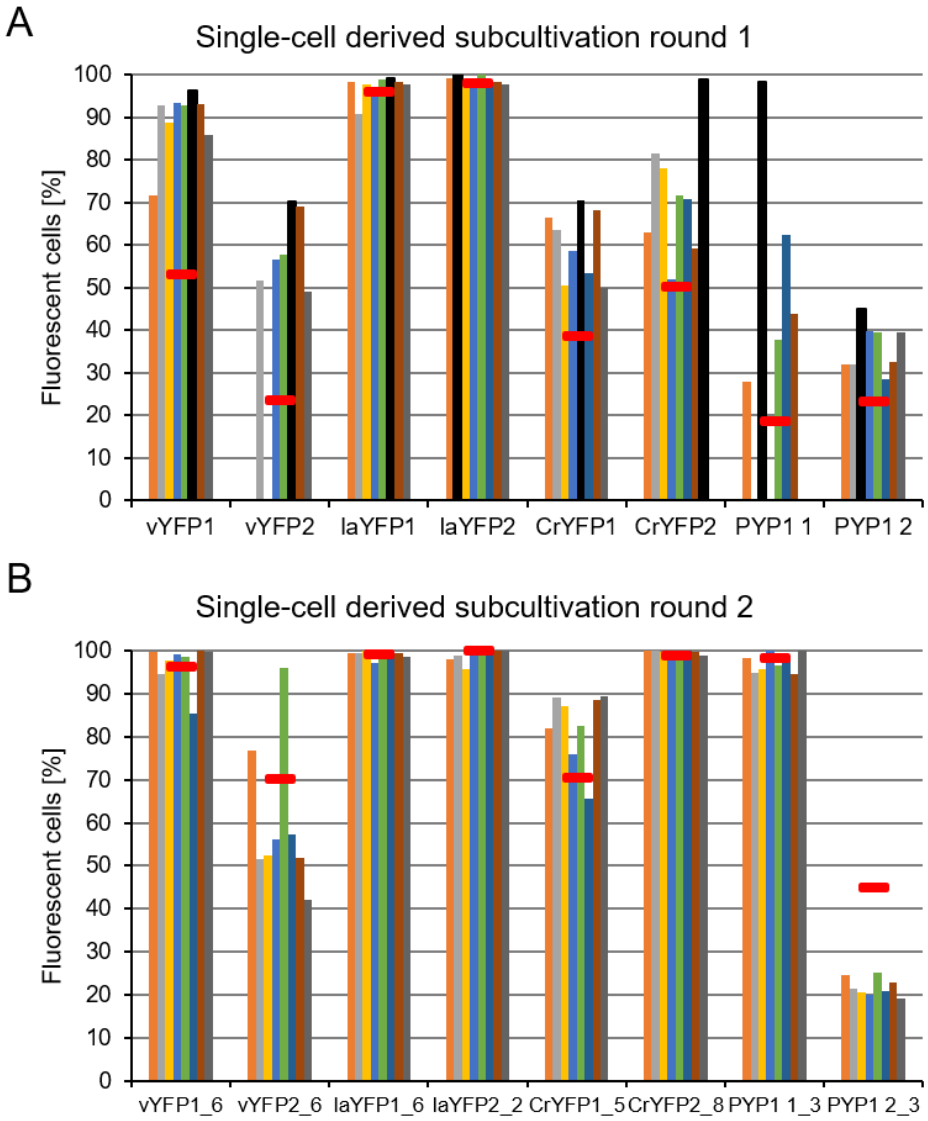


Figure 12: Purification of transgenic strains by single colony isolation.
A: First round of purification. Cells of the progenitor lines (with transgenic clone numbers indicated on the x-axis) growing in liquid culture and showing the proportions of fluorescing cells indicated by the red horizontal bars were plated on lsASW agar plates using the starch method to obtain single colonies. Eight colonies per strain were randomly picked, and the ratio of fluorescent to non-fluorescent cells in liquid culture was determined by microscopy. After the first round of purification, all ratios were higher than those of the progenitor cultures (and reached 100% in several cultures)
B: Second round of purification. The subcultures indicated by the black bars in (a) were picked, and used for another round of single-colony purification. A large number of the subcultures is now homogeneously fluorescent.

line that had been generated by transformation and exhibited 100% fluorescing cells.

4.3 Efficient secretion of YFP into the medium by *P. purpureum*

4.3.1 Immunoblot analysis of secreted and ER-targeted YFP

An important feature of a biotechnological host organism is the capacity to efficiently secrete recombinant proteins into the medium. The accumulation of recombinant protein in the medium offers many advantages, for instance easier purification at lower cost. Additionally, upon expression of toxic proteins, the toxicity for the cell is reduced, resulting in higher protein yields. The protein yields in the medium are not only a matter of expression and secretion capacity, but also of the stability of the protein in the culture medium, which is composed of salts, secreted polysaccharides and possibly organism-specific proteases.

To assess the ability of *P. purpureum* to secrete recombinant proteins, two *YFP* variants were designed based on the codon optimized *PpYFP* gene variant from section 4.2.1. First, the secreted YFP variant named *secYFP* that harbors the signal peptide of *Carbonic anhydrase* (CA) FVE85_3079 was constructed. Carbonic anhydrase is used by many microalgae to increase the intracellular CO_2 concentration by many ways, depending on their subcellular localization (Moroney et al., 2011). *P. purpureum* has nine annotated carbonic anhydrases or carbonic anhydrase-like proteins, but only one is predicted to be localized at the cell membrane, where it converts bicarbonate to CO_2.

The amino acid sequence of carbonic anhydrase was analyzed with ‘TargetP 2.0' and the first 21 amino acids were predicted to act as signal peptide, targeting the protein for secretion. The DNA sequence encoding these amino acids was cloned upstream of the *YFP* reading frame under the control of the *Actin* promoter. For the second variant, an additional HDEL-motif, acting as an ER retention signal, was introduced upstream of the stop codon to obtain an ER-targeted YFP variant, termed *ER-YFP*. *P. purpureum* was transformed with both variants and the obtained transformants were analyzed with respect to YFP localization and accumulation.

To determine the amount and stability of the secreted YFP, an immunoblot analysis of *secYFP* and *ER-YFP* transgenic lines was performed. Proteins in the medium needed to be precipitated without co-precipitation of the exopolysaccharides that accumulate to high amounts during *P. purpureum* cultivation and would impair western blot analysis. To achieve this, a new protocol for protein precipitation had to be established. This was accomplished by the addition of small amounts of the positively charged detergent CTAB, which binds the negatively charged sulphated polysaccharides and precipitates them after vortexing, while keeping the proteins in solution. The cleared supernatant was then used for TCA precipitation to obtain the total secreted proteins (see Material and Methods, section 3.3.3). An equivalent of 1 mL culture medium was loaded on an SDS gel next to 10 µl of total protein obtained from

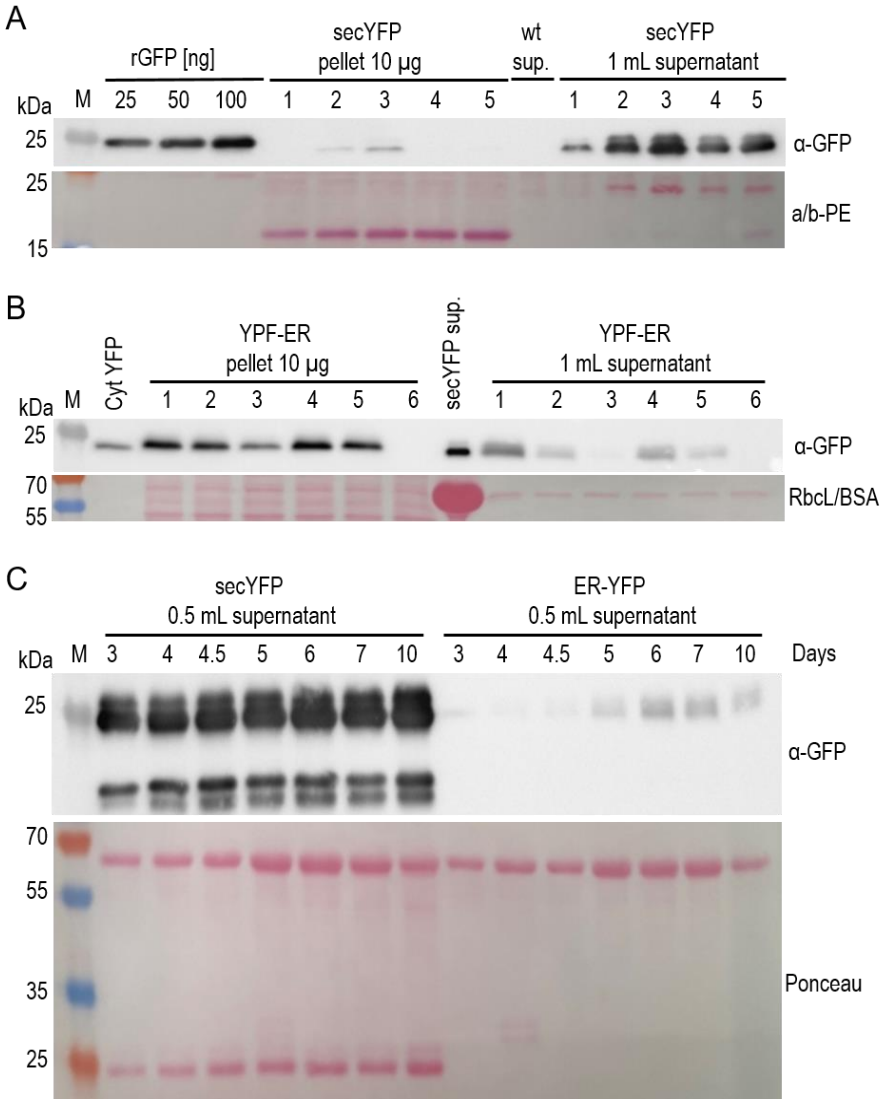


Figure 13: Secretion and ER retention of YFP in *Porphyridium*.
A: Western blot analysis to detect YFP in *Porphyridium* strains engineered to secrete the protein into the culture medium. Samples of 10 µg total protein extracted from cells washed and pelleted were electrophoretically separated by SDS-PAGE, blotted and immunodecorated with anti-GFP antibodies (α-YFP). Proteins precipitated with TCA from aliquots of 1 mL culture medium (supernatant; sup.) were loaded to assess YFP secretion. Five independently generated secYFP transformants were analyzed. For semiquantitative analysis of YFP accumulation, a dilution series of recombinant GFP was loaded. While high levels of YFP accumulate in the medium, only very low amounts of protein are detectable in the washed cell pellet. Clone secYFP-3 was used in all further experiments. Ponceau staining of the α/β-phycoerythrin band (a/b-PE) on the blotted membrane served as loading control. M: molecular weight marker; wt: wild type.
B: Western blot detection of YFP in transgenic *Porphyridium* cultures targeting the protein to the ER. Total protein extracted from pelleted cells was analyzed and compared to 1 mL of the culture medium. As positive controls, 2 µg total protein extracted from an algal strain expression the cytosolic PpYFP (Cyt YFP) positive control, and the equivalent supernatant (60 µL) of a secYFP strain were included. As expected, the YFP is mainly present in the cell pellet, and only small amounts are found in the medium. YFP-ER strain 4 was used in all subsequent experiments. Ponceau staining of the region of the blotted membrane containing the large subunit of Rubisco (RbcL) served as loading control. BSA was used as precipitation carrier for the secYFP control and gives rise to a strong band in the gel.
C: Accumulation of YFP in the culture medium of strains expressing secYFP or YFP-ER over 10 days. secYFP accumulation in the medium increased with cultivation time. The double band of YFP (migrating at ~ 25 kDa) is likely explained by incomplete cleavage of the N-terminal secretion signal. The two bands of lower molecular weight represent putative degradation intermediates. Ponceau staining of the blotted membrane was conducted as control for equal loading, and is shown below the blot. Note that YFP accumulates as the only prominent protein present in the culture medium. The other strong protein band (at ~66 kDa) is BSA that was added as precipitation carrier.

washed and pelleted cells, and analyzed via immunoblotting (Figure 13A). Of five analyzed transformed lines, all lines showed YFP accumulation almost exclusively in the medium, with

only the two strongest lines *secYFP* 2 and *secYFP* 3 exhibiting slight accumulation in the cell pellet fraction as well. The YFP accumulation in the medium derives from secretion and not from cell rupture. This can be deduced from the fact that there is no phycoerythrin band visible on the Ponceau-stained membrane for the culture medium equivalent in lines 2 and 3. Phycoerythrin is part of the phycobilisome, the light-harvesting complex of red algae. This protein is soluble in water and highly abundant in *P. purpureum.* Therefore, a band corresponding to this protein is a good indication of cell rupture. Only in the supernatant sample of *secYFP* line 5, a faint Ponceau-stained phycoerythrin band was detected.
All lines analyzed showed a double band for YFP, with the smaller version being the predominant one. This could indicate either a partial cleavage of the signal peptide partial degradation of the protein. The highly expressing line *secYFP* 3 has been used for all further analyses.
Another immunoblot analysis was performed with six *ER-YFP* lines, of which five accumulated YFP mostly inside the cell (Figure 13B), but a slight accumulation of YFP was detectable in the culture medium for *ER-YFP* lines 1, 2, 4 and 5. YFP that accumulated inside the cells showed no double band, whereas YFP that was secreted to the culture medium showed a double band just as in Figure 13A. This again might indicate partial cleavage of the signal peptide or slight YFP degradation in the culture medium. YFP was either secreted to small amounts, or released during cell rupture, even though no phycoerythrin band was detected on the Ponceau stain. A shift to a higher molecular weight of YFP was seen in the case of the ER-targeted YFP in comparison to the cytosolic YFP (Figure 13B). This is due to the HDEL sequence attached to the C-terminus of the YFP, which accounts for approximately 0.5 kDa.

When comparing YFP accumulation levels in the supernatant of the *secYFP* 3 and *ER-YFP* 4 lines (Figure 13C), it is unsurprising that the strongest YFP secretion was achieved from the *secYFP* construct. Samples for protein extraction were taken at specific time points during the course of the cultivation, as depicted in Figure 15B. Over time, the YFP concentration in the culture medium increased for both constructs, with the highest concentration in the supernatant being reached at the end of the cultivation at day 10. This is particularly well seen after the Ponceau staining of the blotted membranes, where a strong signal at the size of 25 kDa was visible for the secYFP samples, but not the YFP-ER samples, which shows that there is strong YFP accumulation in the culture medium. According to the PAGE, YFP degradation products running at a lower molecular weight were present in the culture medium and their concentration increased over time as well. This is most likely caused by proteases that are co-

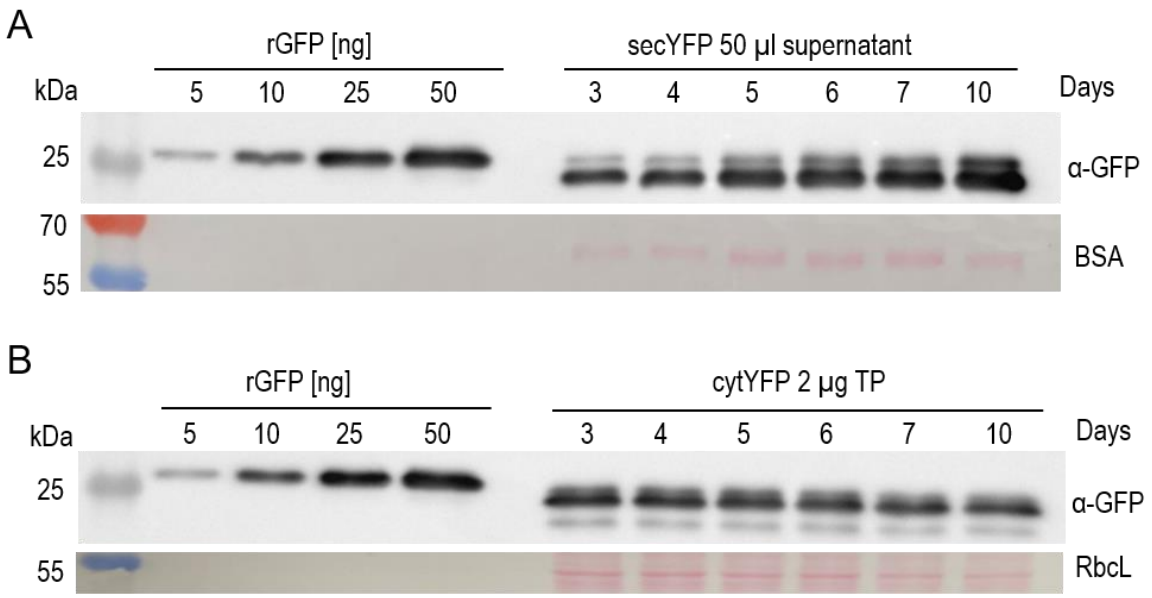


Figure 14: Quantification of YFP accumulation in transgenic *Porphyridium* strains expressing the secreted or the cytosolic version of YFP
A: Time course of secYFP accumulation in the culture medium of *Porphyridium* over 10 days. The initial cell density was 2.73 x 105 cells/mL, and the samples were collected during the growth experiment shown in Figure 15B. Aliquots of 50 µL medium were loaded and compared to a dilution series of a recombinant GFP standard. At day 10, YFP accumulation in the medium reached approximately 1.6 mg/L. BSA was added to the collected supernatant as precipitation carrier, and the region containing the BSA band is shown in the Ponceau stain below the blot. M: molecular weight marker.
B: Time course of cytYFP accumulation. Samples of 2 µg total protein were loaded and compared to the GFP standard. YFP accumulation in the cytosol remains largely constant during cultivation, with the expression level being approximately 2% of the total soluble protein. Ponceau staining of the region of the blotted membrane containing the large subunit of Rubisco (RbcL) served as loading control. Note that the total soluble protein accounts for approximately half of the total cellular protein in *Porphyridium*

secreted into the culture medium, and that degrade YFP as no protease inhibitor was added to the medium during cultivation.

The secretion capacity was determined using semi-quantitative immunoblotting by loading a recombinant YFP standard next to 50 µl of medium equivalent (protein from 1 ml medium was precipitated and resuspended in a defined amount of 1x SDS buffer, the equivalent of 50 µl medium was then used for immunoblot analysis). Samples from a liquid culture were taken each day in order to monitor the accumulation of YFP during a whole cultivation period (Figure 14A, Figure 15B). The accumulation of YFP in the medium increased with cultivation time and reached 1.62 mg/l on day 10. This correlates with the rising cell density in the culture. At the same time, the degradation products increased slightly over time. The average secretion productivity was calculated to be 0.16 mg YFP per liter per day. The accumulation of cytosolic YFP was stable throughout the cultivation and slightly decreased in the stationary phase on day 7 and 10. Total protein (TP) was loaded for this immunoblot and it was calculated that the YFP accumulation was on average 2% of TP, which is approximately half of the amount that

was calculated for the accumulation levels relative to total soluble protein (see section 4.2.2). This indicates that the total soluble protein accounts for approximately half the total protein in a *P. purpureum* cell.

4.3.2 Microscopic analysis and growth of secreted and ER-targeted YFP expressing strains

The subcellular localization of YFP was determined by confocal laser-scanning microscopy using the cytosolic *PpYFP* line from section 4.2.1 as a control (Figure 15A). For the secreted and ER-targeted YFP, two cultures of the same line in the exponential and the stationary growth phase respectively were analyzed. As expected, cytosolic YFP was visible as a thin circle surrounding the chloroplast, corresponding to the cytosol of the cell. The ER-targeted YFP in the exponential phase culture accumulated similarly in an even thinner circle surrounding the chloroplast and the nucleus, but compared to the cytosolic version, it appeared to be closer to the cell membrane. In the stationary phase culture, the ER-targeted YFP appeared to be concentrated in the area surrounding the nucleus and two distinct points close to the cell membrane. The same was true for the stationary phase culture expressing the secreted YFP, where the YFP fluorescence signal concentrated around the nucleus and some dots close to the membrane. The cells in the exponential phase expressing the secreted YFP showed only very little fluorescence in an area close to the cell membrane, indicating

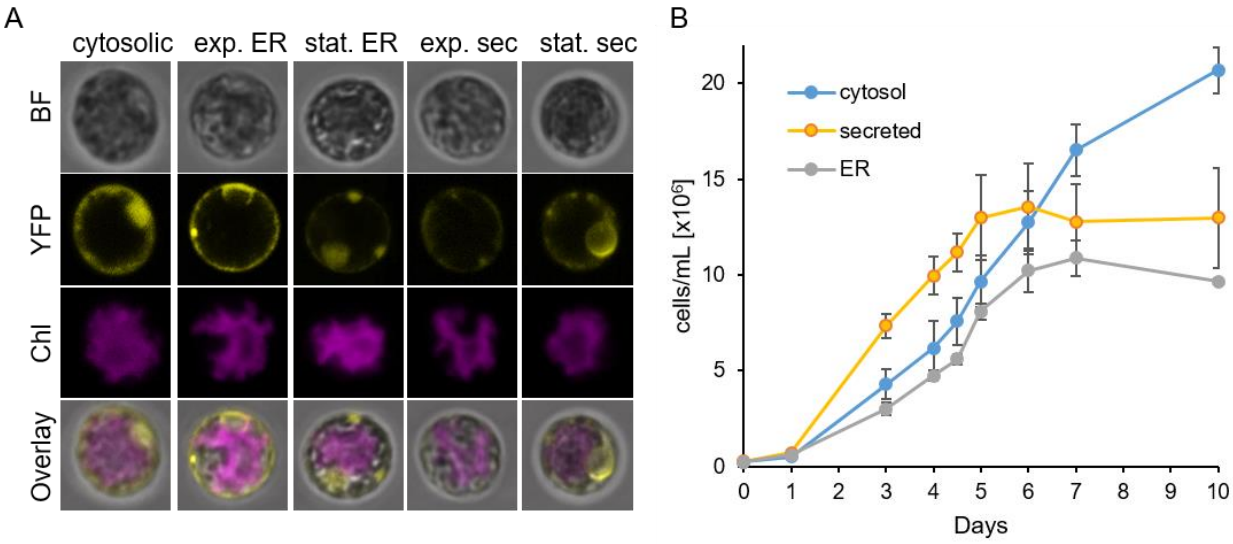


Figure 15: Microscopy and growth analysis of secreted and ER targeted YFP expression lines
A: Confocal microscopy images of *Porphyridium* cells expressing cytosolic, ER-targeted or secreted (sec) YFP versions in the exponential (exp.) and stationary (stat.) growth phase. The bright field images (BF), the YFP fluorescence (YFP), the chlorophyll fluorescence (Chl) in the single star-shaped chloroplast and the overlay of the three images are shown. Note virtually complete absence of YFP fluorescence from cells expressing the secreted version in the exponential growth phase, but substantial YFP retention (presumably in the ER and/or Golgi apparatus) in the stationary phase.
B: Growth curves of *Porphyridium* cell cultures expressing cytosolic, secreted and ER-targeted YFP versions. Slightly lower cell numbers of the secYFP and YFP-ER strains in the stationary phase suggests a small growth penalty, presumably resulting from overloading the secretory pathway with large amounts of recombinant protein. Samples for microscopy were taken on days 3 and 10.

effective secretion of YFP out of the cell. The decrease of YFP fluorescence in the stationary phase of the *ER-YFP* line may indicate ER stress induction and thus decrease of YFP accumulation.

Analysis of the growth of the *secYFP* 3, *ER-YFP* 4 and cytosolic *PpYFP* cell lines, revealed that the YFP-secreting line was growing faster in the early stage of cultivation compared to *PpYFP* and *ER-YFP* (Figure 15B), until it reached the stationary phase on day five. The doubling time in the exponential phase between day 1 and day 5 was calculated to be 28.0 h for the line expressing the cytosolic YFP, 25.6 h for the *secYFP* line 3 and 31.9 h for *ER-YFP* line 4. At the beginning of the stationary phase, the *secYFP* line was outperformed by the line expressing the cytosolic variant, which reached a higher cell density of 20.7 x 10^6 cells/mL at the end of the experiment, compared to 12.98 x 10^6 and 9.65 x 10^6 cells/mL for the *secYFP* and the *ER-YFP* lines, respectively. The ER-YFP expressing line showed lower cell numbers compared to the other two lines throughout the whole cultivation period.

In summary, in the course of this study, it could be shown that *P. purpureum* effectively secretes proteins into the medium using the *carbonic anhydrase* signal peptide. It was furthermore demonstrated that, prior to secretion, the protein was targeted to the ER and can be retained there by the addition of the HDEL motif. This is an important finding, because glycosylated proteins, which can be expressed in *P. purpureum* as shown in the next section, rely on trafficking trough the ER and the Golgi apparatus in order to obtain the N-glycosylation. YFP in the medium was susceptible to degradation most likely due to secreted proteases in the medium.

4.4 Expression of biotechnologically valuable Hepatitis B and Hepatitis C virus antigens in *P. purpureum*

Using *YFP* as a reporter gene, the capability of *P. purpureum* for high transgene expression and secretion could be demonstrated. Additionally, *P. purpureum* strain SAG 1380-1d was able to growth to high cell density in a photobioreactor, which is important for later biotechnological applications. The following section will focus on the suitability of *P. purpureum* for biotechnological applications. Therefore, the expression of two antigens that could be used as potential subunit vaccines against the Hepatitis B virus (HBV) and the Hepatitis C virus (HCV) was examined. First, the general expression of different variants of the HBV and HCV antigens (see Introduction 1.4) in *P. purpureum* was shown and compared to the expression in the *C. reinhardtii* UVM11 expressoin strain. Next, the obtained antigens were biochemically characterized, and the expression was optimized using the results obtained in section 3.1 and 3.2. Lastly, dry mass of the optimized expression lines was generated in order to test the immunogenicity of the antigens by mice immunization trials. The mice trials were divided in

two parts. In the first part, the antigen was administered orally to test if feeding of dry algal biomass can cause humoral and cellular immune responses. The second part was an injection-based administration, where the antibodies were purified, injected into mice, and compared to antigens produced in human cell culture (HCV) or yeast (HBV). This part of the work was performed in close collaboration with our partners from the Romanian National Academy of Science in the laboratory of Prof. Dr. Norica Nichita. They provided the initial design of the novel subunit vaccine candidates, executed some biochemical analyses and performed the mice immunization trials. Their work will be briefly discussed in this thesis and clearly marked as not performed by me.

4.4.1 Expression and purification of 3xHA-tagged YFP-6xHis in *E. coli*

To facilitate easy and cheap semi-quantitative immunoblot analysis of the expressed HBV and HCV antigens, a 3xHA-tagged YFP-6xHis was designed, expressed in *E.* coli and purified. The quantified YFP recombinant protein could then be detected together with the 3xHA-tagged antigen variants using an anti-HA tag antibody.

A plasmid carrying a C-terminally His-tagged *YFP* was used as template and the 3xHA-tag was added to the N-terminus, yielding plasmid pASH41. To simplify reading, the protein will be named 3xHA-YFP, although it additionally contains the 6xHis tag. This plasmid was used

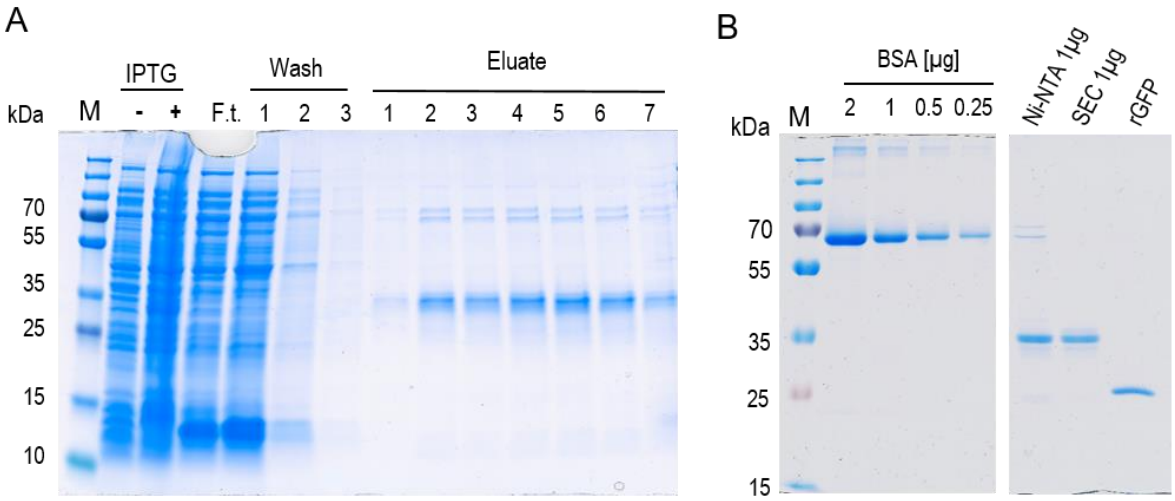


Figure 16: Purification of 3xHA-tagged YFP
A: Coomassie stained SDS-PAA gel showing affinity purification of 3xHA-tagged YFP-6xHis using Ni-NTA resin. The *E. coli* total protein sample after IPTG induced expression (IPTG +) of the 3xHA-YFP-6xHis -tagged YFP showed an intensified band at approximately 35 kDa compared to the control before induction (IPTG -). The 3xHA-tagged YFP-6xHis runs at approximately 35 kDa and is mainly visible in the eluted fractions number 2 to 6. Some residual contaminations at 70 kDa were still present after purification, which is common due to histidine-rich proteins in *E. coli*.
B: Coomassie stained gel of different purifications of 3xHA-YFP-6xHis (preparative PAGE, affinity purification via Ni-NTA beads and SEC) next to a BSA standard and commercial recombinant GFP (Roche). The respective YFP preparations were quantified using the BCA assay and NanoDrop® analysis. Samples that were loaded on this gel and are unrelated to this experiment were omitted from this image (which, therefore, is split).

for transformation of *E. coli* strain BL21 DE3, and protein expression was induced as described in 3.1.8. The purification of the recombinant YFP included a His-tag affinity purification step as described in 3.3.5, followed by size exclusion chromatography (SEC) as described in 3.3.7 to increase the purity of the protein. During the affinity purification, 20 µl samples of each step were taken, separated on a 12% polyacrylamide gel and stained with Coomassie to follow the fate of the recombinant protein accumulation throughout the purification (Figure 16A, B). The 3xHA-YFP has a calculated molecular weight of 31.65 kDa, and, after induction of its expression with 0.5 mM IPTG, *E. coli* transformants showed an increased amount of a protein at approx. 35 kDa, compared to the un-induced control. The recombinant protein bound effectively to the Ni-NTA resins, as no protein was seen in the flow through and the wash at the size of 35 kDa. The 3xHA-YFP was then released by elution with imidazole as seen by a prominent protein band at a size of around 35 kDa mainly in the elution fractions number two to six. These fractions were pooled, concentrated using Amicon® ultrafiltration units and the buffer was exchanged as described in 3.3.6. The concentrated protein was then used as an input for the SEC run in an ÄKTA instrument as described in 3.3.7. After the SEC run, fractions containing YFP were detected using UV light, pooled and concentrated using Amicon® ultracentrifugation units, and the buffer was exchanged to PBS. Finally, the protein obtained from the different purification steps was quantified by the BCA assay and by spectroscopy using the calculated molecular weight and the calculated extinction coefficient $\varepsilon = 35425\ M^{-1}\ cm^{-1}$. The extinction coefficient was calculated via the 'ProtParam' web tool by providing the protein amino acid sequence. The concentration was adjusted to 1 µg/µl and all samples were loaded on a 12% SDS-PAA gel next to a BSA standard and 1 µg of commercial recombinant GFP, and after separation the gel was stained with Coomassie (Figure 16B). The concentration of the Ni-NTA resin-purified and the SEC-purified 3xHA-YFP was similar to 1 µg of the commercial recombinant GFP and the BSA loaded on the same gel, indicating correct quantification via NanoDrop®. By applying an additional SEC after the affinity purification, the purity of the protein could be increased and no visible contamination at 70 kDa could be seen anymore, in contrast to the only Ni-NTA-purified 3xHA-YFP.

This purified 3xHA-YFP was used in all subsequent quantification experiments as a standard.

4.4.2 Design and expression of recombinant HBV and HCV antigen variants in *P. purpureum* and *C. reinhardtii*

To assess the suitability of *P. purpureum* as a production platform in biotechnology, antigens of HBV and HCV that can act as subunit vaccines were expressed in *P. purpureum*. The expression in the red alga was compared to the "gold standard" C*. reinhardtii* UVM11. These new antigens were rationally designed by our collaboration partners in the lab of Prof. Dr. Norica Nichita from the Institute of Biochemistry of the Romanian Academy, and the design is

described in the Introduction. In short, for the HBV antigen, the small protein of the envelope (S-protein) was modified by inserting the amino acids 16 to 42 of the immunogenic preS1 region into the major antigenic loop of the S-protein. This new chimeric protein contains four transmembrane domains and one N-glycosylation site, and was named HBV-S/preS1$^{16-42}$. For simplicity, all variants of this antigen will be referred to as HBV antigen (HBV-Ag). The trafficking of this protein through the ER is crucial for its correct folding and subviral particle (SVP) formation, which in turn is crucial for the immunogenicity of this antigen (refer to Introduction section 1.4.2).
For the HCV antigen, a modified version of the ectodomain of the E2 surface glycoprotein without the hypervariable region 1 (HVR1) was chosen and named HCV$^{\Delta HVR1}$. In the following, this antigen and all its variants will be referred to as HCV antigen (HCV-Ag). This protein has eleven potential N-glycosylation sites and nine disulfide bonds and relies on trafficking through the ER and Golgi apparatus to acquire the N-glycans and the correctly folded conformation.

Three different versions of both the HBV and HCV antigens were designed and a schematic representation of the design is shown in Figure 17A. The antigen sequences were codon optimized for the *P. purpureum* and *C. reinhardtii* nucleus, respectively, to ensure optimal expression. Cytosolic (*cytHBV* and *cytHCV*), secreted (*secHBV* and *secHCV*) and ER-targeted (*ER-HBV* and *ER-HCV*) versions were constructed. The signal peptide sequence of the respective endogenous carbonic anhydrase was incorporated upstream of the respective codon-optimized antigen sequence in the secreted and ER targeted antigen versions (refer to 4.3). A C-terminal HDEL motif was additionally incorporated in the ER-targeted antigens. The cytosolic antigens do not contain ER translocation signal sequences. Expression of all the antigen variants in *Porphyridium* was controlled by the endogenous *Actin* promoter and terminator sequences. For expression in *C. reinhardtii* UVM11 strain, the *Chlamydomonas PsaD* promoter and terminator sequences were used. All versions carry a 3xHA-tag for easy detection via immunoblotting. In total, 12 different antigen variants were cloned: the cytosolic, the ER-retained, and the secreted versions of the HBV and the HCV antigen codon optimized sequences for *Porphyridium*, as well as the cytosolic, the ER-retained and the secreted versions of both antigens that were codon optimized for *Chlamydomonas*. Each alga was transformed with the constructs optimized for its specific codon usage.

C. reinhardtii UVM11 strain failed to express all three HBV antigen variants and only expressed the HCV antigen variants to detectable levels (see Suppl. Figure 2). In this immunoblot analysis, 40 µg of total *C. reinhardtii* protein was loaded next to a positive control of *P. purpureum* expressing the respective antigen. CytHCV was expressed in three out of 12 tested lines to detectable levels. The secreted and ER-targeted versions were expressed to higher levels and could be detected in five and eight out of 12 lines analyzed, respectively.

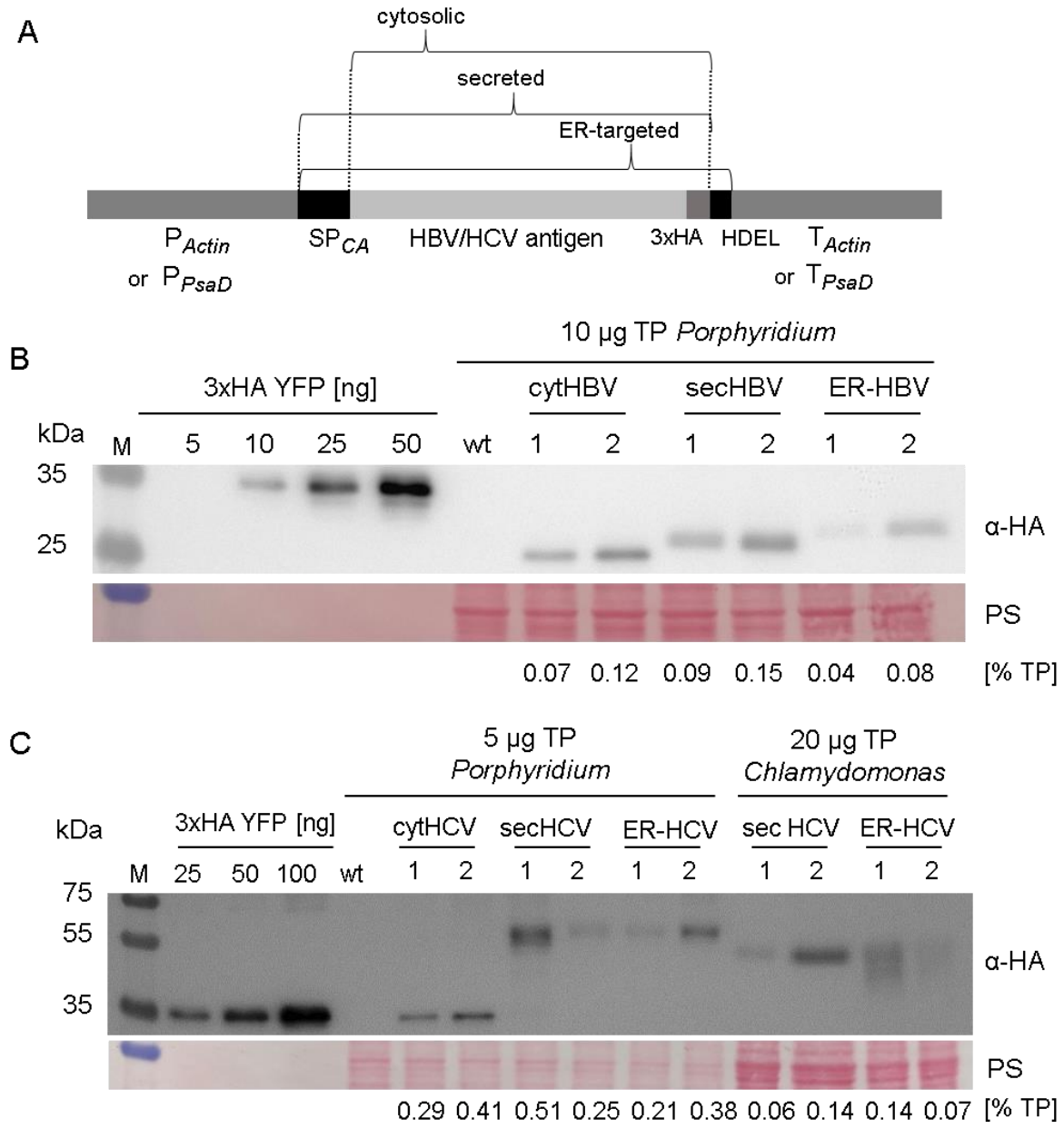


Figure 17: Design, expression and quantification of different HBV and HCV antigen variants in *P. purpureum* and *C. reinhardtii*.

A: Schematic representation of the design of the HBV and HCV antigens. A cytosolic (*cytHBV* and *cytHCV*), a secreted (*secHBV* and *secHCV*) and an ER-targeted (*ER-HBV* and *ER-HCV*) version were designed by incorporating the endogenous signal peptide of the carbonic anhydrase upstream of the codon-optimized antigen-coding sequence for the secreted and the ER-targeted antigen of the respective algal species. A C-terminal HDEL motif was incorporated only in the ER-targeted antigen. The cytosolic antigen does not contain ER translocation signal sequences. All antigens carry a 3xHA-tag for easy detection via immunoblotting. For expression in *P. purpureum,* the promoter (and 5'UTR) and terminator of the endogenous *Actin* was utilized. For expression in *C. reinhardtii*, the *Chlamydomonas* endogenous *PsaD* promoter and terminator sequences were used.

B: Immunoblot analyses of samples of 10 µg total protein of *P. purpureum* lines expressing different HBV variants using the anti-HA antibody. Recombinant 3xHA-YFP was loaded as a standard to enable quantification. The membrane was stained with Ponceau (PS) to check for equal loading. Quantification values of the antigen are shown below each lane (given in percent of total protein).

C: Immunoblot analyses ofsamples of 5 µg total protein of *P. purpureum* lines expressing the different variants of the HCV antigen and 20 µg of total protein of *C. reinhardtii* lines expressing two variants of the HCV antigen. The Ponceau-stained membrane clearly shows the difference in the loading between the two algae. The recombinant 3xHA-YFP was used as standard, and accumulation values of the antigen are shown below each lane (in percent of total protein).

For the secHCV and the ER-HCV antigen constructs, two *Chlamydomonas* transformed lines each, identified as the strongest expressors, were isolated and compared to the expression levels of the strongest *Porphyridium* expressors carrying the corresponding constructs as shown in Figure 17B and C.

In *P. purpureum,* all HBV antigen variants could be expressed and accumulated to levels of 0.07% to 0.12% of total protein (TP) in case of the cytHBV-Ag variant, to a level of 0.09% to 0.15% of TP for the secHBV-Ag variant, and for the ER-HBV-antigen variant to a level of 0.04% to 0.08% of TP. However, no secreted secHBV-Ag was detected in the medium. The theoretical mass shift resulting from one $Man_9GlcNAc2$-glycan, which is the glycan attached to an N-glycosylated protein at the end of ER-trafficking, and the final conserved N-glycosylation step in yeast, plants and mammals is approximately +1.86 kDa (Dobrica et al., 2021). This is in line with the observations made in the immunoblot analysis, when comparing the band sizes of the cytosolic and the secretory pathway-targeted or ER-retarded versions.

For the HCV antigen, all three variants could be successfully expressed in both microalgae and accumulation levels could be compared. For this, 5 µg of *P. purpureum* and 20 µg of *C. reinhardtii* total protein of the lines expressing the respective variants were loaded next to a 3xHA-YFP standard and analyzed by immunoblotting using the HA-antibody (Figure 17C). The highest level of antigen accumulation was reached in the *secHCV*_1 line (0.51% of total protein), and the lowest level of antigen accumulation was observed for the *ER-HCV* construct with 0.21% of total protein. The strongest expressor in *C. reinhardtii* was the *secHCV* 2 line with only 0.14% of TP. Overall, *P. purpureum* showed an up to 3.5 times higher accumulation of the HCV antigen compared to *C. reinhardtii*.

The 11 glycosylation sites (at the Asn-X-Ser/Thr consensus sequence) of the HCV antigen caused a theoretical mass shift of +20.5 kDa assuming that the high mannose $Man_9GlcNAc_2$-glycan is attached to every potential glycosylation site. This size shift was observed in the western blot analysis for the constructs *secHCV* and *ER-HCV*. One striking observation is the apparent difference in the glycosylation pattern between the two algae, as evidenced by the different masses of the glycosylated HCV antigens. For both algae, no secreted HCV-Ag could be detected in the cultivation medium, which is surprising, as the ectodomain should readily be secreted into the medium, as it is in many other organisms expressing the E2 ectodomain (see Introduction XXX). Another notable point was the high variation in antigen accumulation between transformed lines of the same construct. This variation was not only seen for the secHCV construct, but also for the ER-HCV and the ER-HBV construct. This variation in expression was studied in more detail for two of the ER-HCV transformants, one with a low expression (ER-HCV1) and another line with a higher expression of ER-HCV (ER-HCV2).

In summary, it was shown that *P. purpureum* was capable of expressing complex proteins like the highly glycosylated HCV antigen and the transmembrane domain-containing HBV antigen to relatively high levels of up to 0.5% and 0.15% of the total protein, respectively. *C. reinhardtii* on the other hand was not capable of expressing the HBV antigen to detectable levels. It was furthermore demonstrated that *P. purpureum* exhibited a stronger expression of the HCV antigens compared to *C. reinhardtii*.

4.4.3 Investigation of the variance of antigen accumulation between expression lines of the same construct

P. purpureum transformation relies on the episomal maintenance of the transformation plasmid. Assuming similar plasmid copy numbers in the independently transformed lines of the same construct, as it was shown by Li and Bock (2018), the expression of the transgene should be similarly high. Variation in transgene expression that may be caused by position effects can be excluded, as *Porphyridium* does not integrate the plasmid DNA into the genome. A possible reason for the observed variation in the expression of the viral antigens could be growth of wild-type cells in the background of the population as it was seen for YFP in 4.2.5. Different ratios of wild-type to transgenic cells between the antigen-expressing lines could cause the variation of expression among lines of the same construct. Another reason could be that the different lines were harvested at different stages of the cell cultures.

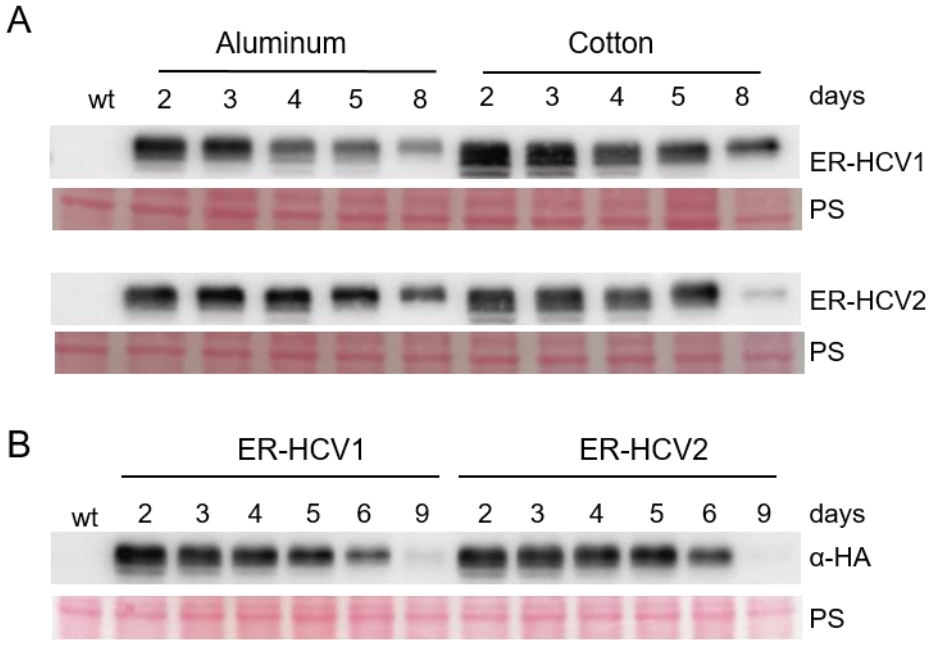


Figure 18: Immunoblot analysis of *Porphyridium* ER-HCV transformed lines, sampled during a growth experiment.
A: 10 µg total protein were loaded and separated on a 12 % SDS gel and the HCV antigen was detected using an anti-HA antibody. Equal loading was confirmed by Ponceau staining of the blotted membrane. The samples were taken during a growth experiment on specific days after inoculation, and the impact of different seals on antigen accumulation was determined.
B: Cotton-sealed cultures were grown until the stationary phase at day nine and samples were taken on the indicated days. 10 µg TP of both ER-HCV-Ag expressing lines were loaded and antigen accumulation was detected using the HA antibody.

Cells sampled in the exponential growth phase usually show different recombinant protein accumulation levels compared to cells sampled in the stationary phase. Especially proteins expressed in the ER show susceptibility to the growth state of the culture, presumably due to ER-stress occurring in those cells. It can happen that the protein folding machinery in the ER is at maximum capacity, and therefore, misfolded proteins start to accumulate. This triggers the unfolded protein response (UPR) in the ER, resulting in decreased protein synthesis and onset of protein degradation, and thus decreased protein accumulation (Howell, 2021).

In order to investigate, if the culture stage has an influence on the strength of transgene expression, ER-HCV lines were inoculated in triplicates and grown until they reached the stationary phase. The experiment was performed with both the aluminum seal and the cotton seal, to test if CO_2 availability has any influence on antigen accumulation. The growth curve of this cultivation experiment is shown in Suppl. Figure 3 and shows the expected growth behavior, with cotton-sealed cultures growing much faster than aluminum-sealed cultures. Cell samples for protein isolation were taken during the course of cultivation and analyzed by immunoblotting (Figure 18). The immunoblot clearly shows a strong reduction of ER-HCV-Ag accumulation from day 4 on for the aluminum-sealed cultures, and from day 5 on for the cotton-sealed cultures. The decline in antigen level correlated with the entrance into the stationary phase. The CO_2 availability only played a minor role in antigen accumulation, since the antigen accumulation is similar between cotton-sealed and aluminum-sealed cultures. Figure 18B shows a second growth experiment performed with cotton seals only, where *ER-HCV* line 1 and 2 are loaded side by side on a single gel. The data clearly show that the accumulation level of the HCV antigen was strongly dependent on the sampling time and that, if samples are taken at the same growth state, no variation of accumulation strength is observed. In late stationary phase cultures (i.e., at day 9), the antigen accumulation levels decrease to almost zero.

The two *secHCV* lines showed variations in antigen accumulation as well. The accumulation of the secHCV antigen was not as strongly dependent on the growth phase as that of the ER-targeted antigen (Figure 24A).

To explain the differences in protein levels between *secHCV*-Ag line 1 and 2, an inhomogeneous population of antigen-expressing and wild-type cells was assumed, as it was already shown for the YFP-expressing lines in section 4.2.5. If not all cells of a transgenic line express the antigen, there will be a mixed population of wild-type and antigen-expressing cells, whose ratio will be different between the lines, leading to varying antigen accumulation.

To test this hypothesis, single colonies from the *secHCV* line 2, which has lower expression compared to line 1, were generated and nine single cell-derived sublines (*secHCV* 2 A to I) were analysed via immunoblotting (Figure 19).

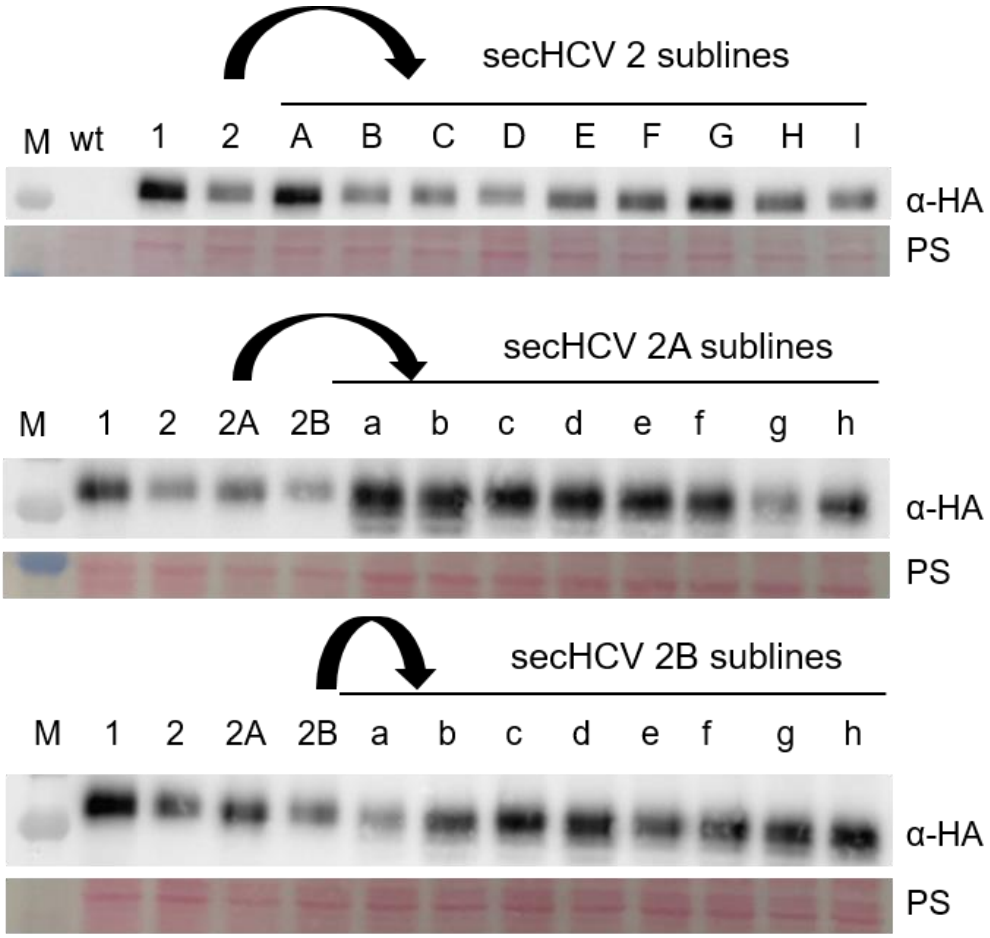


Figure 19: Generation of single colony-derived lines from secHCV2
Single colonies of line secHCV2 were generated using the starch plating method and nine colonies were picked and analyzed by immunoblotting. 5 µg total protein of secHCV-Ag expressing lines 1 and 2 were loaded next to 5 µg total protein of the sublines from secHCV2. Out of these sublines, a high (2A) and medium (2B) expressing line were chosen and single colony lines were generated again and analyzed via immunoblotting.

The sublines *secHCV* 2 A to I showed varying antigen levels ranging from high levels (comparable to *secHCV* 1) in sublines A and G to low level in subline D, which was even lower than in the progenitor line. Assuming a varying ratio of wild-type to antigen-expressing cells in the sublines, the sublines of the high accumulation line secHCV2 A should also exhibit a stable high accumulation in their respective sublines, because a higher proportion of cells are not wild-type. On the other hand, sublines derived from lower expressing sublines should again show variation of expression. This would be comparable to the results obtained in the *YFP* lines subcultivation experiments, where lines exhibiting a higher percentage of fluorescent cells showed higher fluorescent percentages in their respective sub-lines as well.

To test this, single colony-derived sublines of *secHCV2A* (high level of antigen accumulation) and of *secHCV2B* (medium level of antigen accumulation) were generated, and antigen levels were analysed via immunoblotting. Those sublines were named *secHCV2A*_a to _h (middle panel in Figure 19) or *secHCV2B*_a to _h (bottom panel in Figure 19), respectively. All sublines

of *secHCV2A* except for line _g showed an even higher antigen accumulation than the level of the progenitor line and the *secHCV* line 1 that was not subjected to single colony generation. This indicates that *secHCV1* most likely was a mixed culture of expressing and wild-type cells as well. The sublines of *secHCV2B,* except for subline _a, also showed higher HCV protein accumulation levels compared to their progenitor line, and were in between the levels of *secHCV2A* and *secHCV1*. In general, HCV protein expression in the *secHCV2B* sublines was lower than in the *secHCV2A* sublines.

In conclusion, it was shown that the variation in antigen accumulation of different lines of the same construct was most likely caused by a combination of two factors. First, the different sampling time point, where entrance in the stationary phase caused ER-targeted proteins to be degraded and thus appearing differently expressed. Second, as already demonstrated in section 4.2.5, the presence of a heterogeneous culture with a mix of expressing and wild-type cells. If this heterogeneity is different between independently transformed lines, then the antigen accumulation appears to be different between two lines.

The *secHCV2A_a* line that produces high amounts of the secreted HCV antigen was used in further experiments such as the biochemical assays and the bioreactor cultivation experiments. From an immunological perspective, the secHCV antigen is of utmost interest, as it closely resembles the antigen produced in mammalian cells.

4.4.4 Investigation of the glycosylation state of the HBV and HCV antigens by employing the N-glycosylation inhibitor tunicamycin

Both viral antigens contain potential N-glycosylation sites in their sequences. The chimeric HBV S/preS1 protein has one N-glycosylation site and the HCV E2 ectodomain carries eleven N-glycosylation sites. A first strong evidence for successful glycosylation of both proteins was shown in Figure 17, where a molecular weight shift was detected for the sec and ER variants, which did not occur in the cytosolic version. Because the first steps of N-glycosylation take place in the ER, this indicated successful translocation into the ER and attachment of N-glycans. So far, not much is known about N-glycosylation in *P. purpureum*. A bioinformatics study described the presence of all key enzymes in the ER of *Porphyridium* to achieve glycosylation of the high-mannose type $Man_9GlcNAc_2$, which represents the conserved eukaryotic glycan core trafficking through the ER (Levy-Ontman et al., 2014). This N-glycan has a molecular weight of 1.86 kDa and would cause a theoretical mass shift of 20.5 kDa for the HCV antigen, if all eleven sites were glycosylated.

To test if the observed molecular weight shift is caused by N-glycans, the N-glycosylation inhibitor tunicamycin (TM) was added during growth of the respective antigen-expressing lines. Tunicamycin acts as an inhibitor of the enzyme N-glycosylamine phosphotransferase

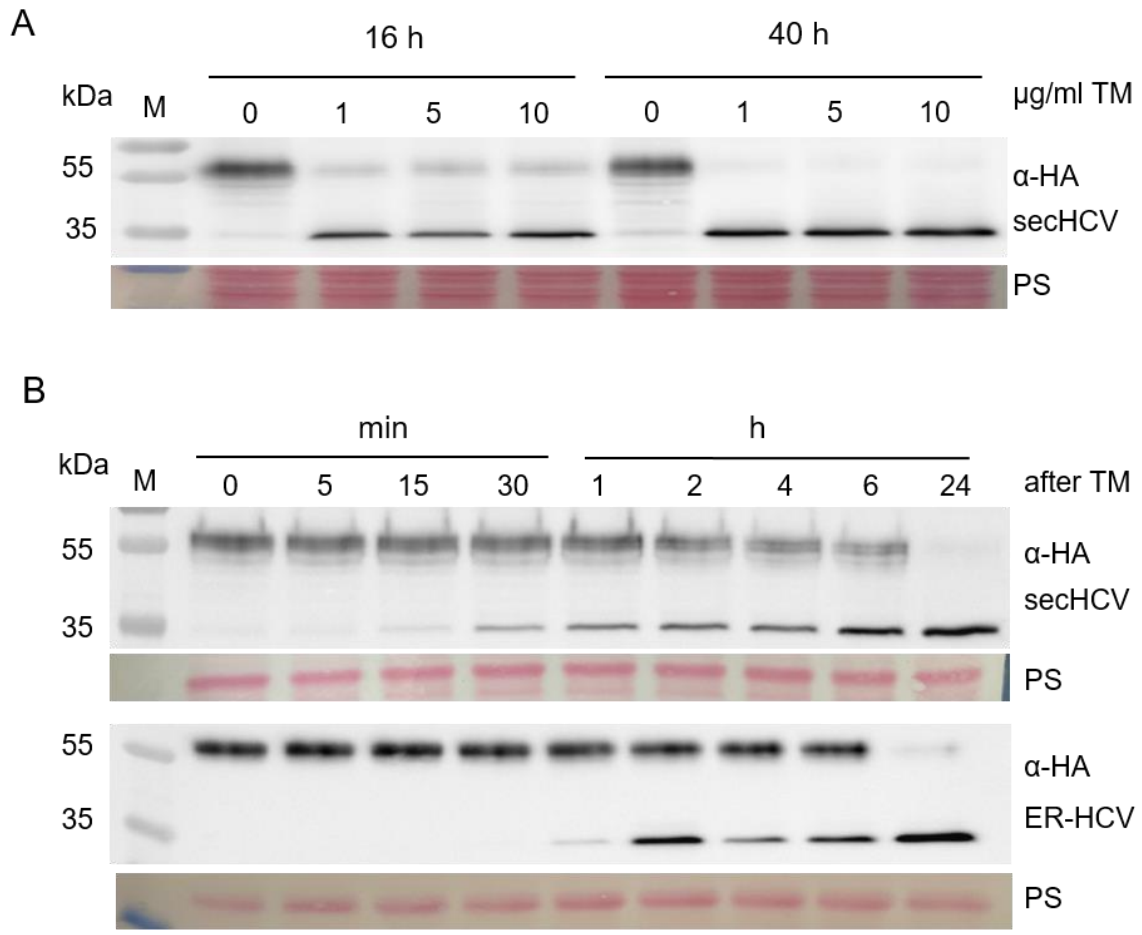


Figure 20: Impact of tunicamycin addition on secHCV and ER-HCV antigen glycosylation
A: Test of tunicamycin (TM) effectiveness in *P. purpureum* cells expressing the secHCV antigen. Samples of cells grown in medium with different TM concentrations were taken after 16 h and 40 h. 20 µg of total protein were loaded on a 12% SDS gel and analysed via immunoblotting with the HA antibody. Equal loading was confirmed via Ponceau staining (PS) of the membrane .
B: 1 µg/mL TM was added to a culture with a cell density of 4.5 x 10^6 cells/mL of algal strains expressing secHCV or ER-HCV. The cultures were grown for 24 h and samples for immunoblot analysis were taken at the indicated time points. 20 µg of total protein were analysed via immunoblotting with the HA antibody, and equal loading was confirmed by Ponceau staining.

(GPT), which is the key enzyme in the first step of glycosylation. This enzyme transfers N-glycosylamine (GlcNAc) to a dolichol phosphate anchor embedded in the ER membrane that is later transferred onto the proteins. When this step is inhibited, the whole N-glycosylation machinery is disrupted, leading to non-glycosylated proteins.

First, the toxicity and effectiveness of TM on *P. purpureum* cells was tested by adding different amounts ranging from 1 µg/mL to 10 µg/mL TM to a liquid culture of exponentially growing *P. purpureum* cells (4.5 x 10^6 cells/mL). The growth was monitored for two days and samples were taken after 16 h and 40 h (see Suppl. Figure 4). The addition of 1 µg/mL TM had negative effect on the growth rate, whereas the addition of higher concentrations led to arrested growth and even to a decrease in cell number. A culture without the addition of tunicamycin was grown side by side as a control. At the protein level, the addition of 1 µg/mL TM was sufficient

to inhibit N-glycosylation to the same extent as higher concentrations, as evident from the immunoblot analysis (Figure 20A). After 16 h, all samples grown in the presence of different TM concentrations showed a shift of the antigen signal from 55 kDa to 35 kDa, corresponding to the non-glycosylated HCV antigen (see cytHCV in Figure 17C). Slight traces of glycosylated HCV antigen at 55 kDa were still visible, which however were almost completely gone after 40 h of TM incubation.

To investigate the dynamics of the inhibition of glycosylation, a time course experiment with lines expressing secHCV-Ag and ER-HCV-Ag was performed. 1 µg/mL TM was added to a culture with a density of 4.5 x 10^6 cells/mL and samples for protein isolation were taken at the indicated time points (Figure 20B). The earliest appearance of non-glycosylated HCV antigen at a size of 35 kDa was seen after 15 min of TM treatment for the secHCV-Ag expressing line, and after 1 h for the ER-HCV-Ag expressing line. The accumulation of the non-glycosylated form of the antigen increased over time and no glycosylated form could be detected after 24 h. At the same time, the signal for the glycosylated HCV antigen decreased, starting two hours after TM addition for the secHCV-Ag (77% after 2 h, 65% after 4 h and 59% signal intensity of the 0 h time point after 6 h of TM treatment). The decrease in glycosylation for the ER-HCV antigen is not as pronounced, and was 90% at 2 h, 87% at 4 h and 84% at 6 h respectively, compared to T0. Another interesting point is the varying signal sharpness between secHCV and ER-HCV, indicating different antigen sizes in the secHCV sample.

This result conclusively demonstrated that the secHCV and ER-HCV antigens pass the ER and obtained N-glycosylation. Further glycosylation analysis employing EndoH and PNGaseF digestion was performed by our collaboration partners and confirmed N-glycosylation that was similar to mammalian type N-glycosylation (Hammel et al., manuscript in preparation). Having demonstrated that the HCV can be strongly expressed and glycosylated, a better strategy for antigen purification was developed. To this end, the secHCV antigen construct was chosen and the 3xHA tag was replaced by a 6xHis tag.

The tunicamycin analysis was performed with the lines expressing the HBV antigen variants as well. In mammals, the first transmembrane domain of this protein acts as signal peptide for secretion, and it is possible that this holds true for *P. purpureum* as well. Because of this, the *cytHBV* version was also subjected to the tunicamycin experiments. Figure 21A shows an immunoblot analysis of the different HBV antigen variant-expressing lines grown with or without the addition of tunicamycin. Samples were taken 16 h after TM addition and analysed by immunoblotting with the HA antibody. In all samples that were taken from TM grown cultures, the antigen accumulation was greatly decreased, indicating that this antigen is highly susceptible to missing glycosylation and/or indication of ER stress. As the non-glycosylated

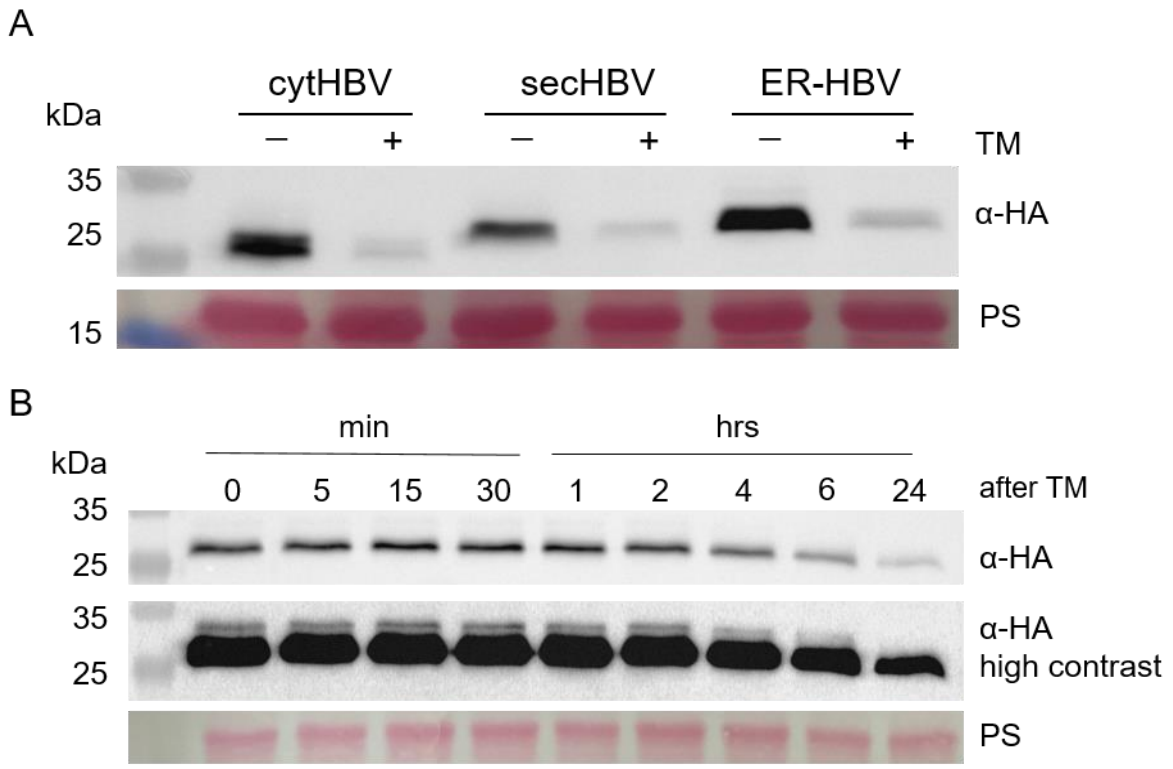


Figure 21: Impact of tunicamycin addition on the HBV antigen variants.
A: Immunoblot analysis of cyt, sec and ER-HBV expressing lines grown for 16 h with (+) or without (-) the addition of 1 µg/mL TM. 50 µg total protein were separated and analyzed using the HA-antibody. PS: Ponceau staining of the phycoerythrin.
B: 1 µg/mL TM was added to a culture with a cell density of 4.5 x 10^6 cells/mL of cells expressing cytHBV. The cultures were grown for 24 h and samples for immunoblot analysis were taken at the indicated time points. 20 µg of total protein were analysed via immunoblot with the HA-antibody. The lower immunoblot shows the same detected membrane with increased contrast. PS: Ponceau staining of the phycoerythrin.

form is easily produced in the ER of yeast, it is more likely that TM is inducing ER stress, which in turn leads to degradation of ER-resident HBV antigen.

Interestingly, this effect was also seen with the cytHBV antigen, which should not be affected by TM. When looking closely, cytHBV showed a double band in the non-treated control, indicating partial glycosylation. This, together with the decrease in antigen accumulation upon TM treatment, suggests that the cytosolic HBV antigen is in fact not cytosolic, but instead the first transmembrane domain may function as signal peptide targeting the antigen to the ER. To further investigate this phenomenon, a time course experiment with the cytHBV-expressing line was conducted and samples were taken at the indicated time points after TM addition (Figure 21B). The high contrast image of the immunoblot showed a distinct double band and the more slowly migrating band may correspond to the glycosylated version of the HBV antigen. The intensity of this band is decreasing over time, similar to the glycosylated HCV version, before completely disappearing after 24 h. Additionally, the accumulation of the non-glycosylated version starts decreasing after 2 h before being greatly reduced after 24 h fitting

to the results obtained in Figure 21A. The band intensity after 2 h was at 90%, after 4 h at 69%, after 6 h at 41% and after 24 h at 15% of time point 0.

This experiment demonstrated that all three variants are targeted into the ER, due to the clear reduction of accumulation in response to TM treatment, and that the cytHBV antigen obtained an N-glycan. Digestion with the glucosidases EndoH and PNGase was performed by our collaboration partners for secHBV and ER-HBV, and showed that both antigens were present in a mixture of glycosylated and non-glycosylated forms. The ratio of glycosylated to non-glycosylated protein was higher for the secHBV protein and much lower for the ER-targeted HBV antigen.

Further experiments conducted by our collaboration partners supported the finding that the cytosolic HBV antigen was actually imported into the ER. An analysis conducted with conformation-specific antibodies verified the native HBV S-protein conformation that can only be reached when the protein maturated in the ER. Another analysis performed by our partners revealed that the cytHBV and secHBV antigens were unable to form SVPs, which are crucial for immunogenicity, and instead only formed dimers. This is most likely due to the addition of the 3xHA tag at the C-terminus of the antigen. This behaviour was observed in mammalian and yeast cells expressing the HBV S-protein as well.

4.4.5 Expression of a new, un-fused HBV S/preS1 variant in *P. purpureum*

Taking all the results of section 4.4.4 and from our collaboration partners into account, a new HBV S/preS1 protein version was constructed, omitting the 3xHA tag and the signal peptide, as the first TMD already confers import into the ER. Even though the antigen with the additional signal peptide of *P. purpureum* showed much higher glycosylation rates (discussed in section 5.5.2), it was decided to omit the signal peptide to increase to probability that the antigen folds into the native conformation in order to allow SVP formation. The molecular weight shift between the cytHBV and secHBV variants, as also seen in the TM treated samples in Figure 21A, indicates that the size shift in secHBV is caused by the additional signal peptide sequence. If the signal peptide were only partially cleaved off, this would also interfere with SVP formation. The glycosylation state of the HBV S-protein plays only a minor role in immunogenicity compared to SVP occurrence (see Introduction). In case of SVP formation, the newly constructed tag-free antigen variant was planned to be used for immunization experiments in mice. SVP formation was tested by ultracentrifugation of natively lysed *P. purpureum* cells expressing the antigen on a sucrose gradient in comparison with mammalian cell derived HBV S/preS1 protein SVPs. Dimers of the chimeric S/preS1 protein are found in higher fractions compared to the SVPs that consist of up to 100 chimeric S/preS1 proteins and thus are present in lower fractions of the gradient.

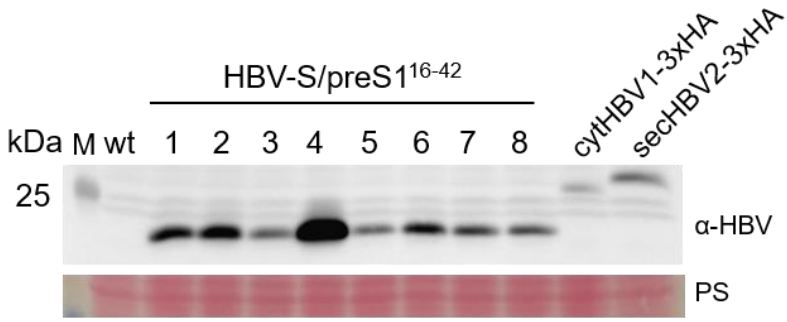


Figure 22: Immunoblot analysis of *P. purpureum* lines expressing native, untagged HBV-S/preS1$^{16-42}$
20 µg of total protein derived from HBV-S/preS1 transformants were separated on a 12% SDS gel and analyzed by immunoblotting with an HBV-S protein-specific antibody. Protein samples from 3xHA tagged cytHBV and secHCV transformants were loaded as controls. M: molecular weight marker; PS: Ponceau staining; wt; wild-type control; PS: Ponceau staining of RbcL.

The codon optimized HBV-S/preS1$^{16-42}$ sequence was cloned under the control of the native A*ctin* promoter and terminator, and *P. purpureum* was transformed with this construct. The construct was named HBV-S/preS1. Eight Zeocin™-resistant colonies were isolated and screened by immunoblot analysis using the HBV S-protein antibody next to a wild-type control and the previously expressed 3xHA tagged cytHBV and secHBV antigens (Figure 22). All eight lines expressed HBV-S/preS1 to different amounts with line number four being the strongest expressor. Compared to the cytHBV and secHBV antigen control, the accumulation of the new antigen variant is much higher, possibly indicating increased stability of the protein. The molecular weight difference between the new and the tagged variants is caused by the absence of the 3xHA tag, which adds a molecular weight of approximately 5 kDa to the tagged variants. *P. purpureum* HBV-S line 4 was chosen for all further experiments and freeze-dried cell material was sent to our collaboration partner, where SVP formation was in fact confirmed (see Discussion section 5.5.2).

To generate enough material for mice immunization trials, HBV-S line 4 was grown in a bioreactor. The results of the experiment are described in section 4.4.7.

4.4.6 Expression of the 6xHis-tagged secHCV antigen in *P. purpureum*

In the previous sections, it was demonstrated that the 3xHA tagged secHCV antigen was expressed to relatively high levels of up to 0.5% of TP in *P. purpureum*, and that faithful glycosylation took place in the ER. Furthermore, our collaboration partner could confirm the binding of the recombinant HCV antigen to conformation-specific antibodies, indicating correct folding of the HCV E2 protein ectodomain. In the next step, a new secHCV construct with a 6xHis-tag instead of the HA-tag was cloned in order to facilitate easy purification of this antigen. High purity and yield are key for effective immunization trials, especially when the

HCV antigen is administered parenterally. The attachment of a 6xHis-tag to the C-terminus of the antigen offers the possibility for easy and cheap affinity purification.

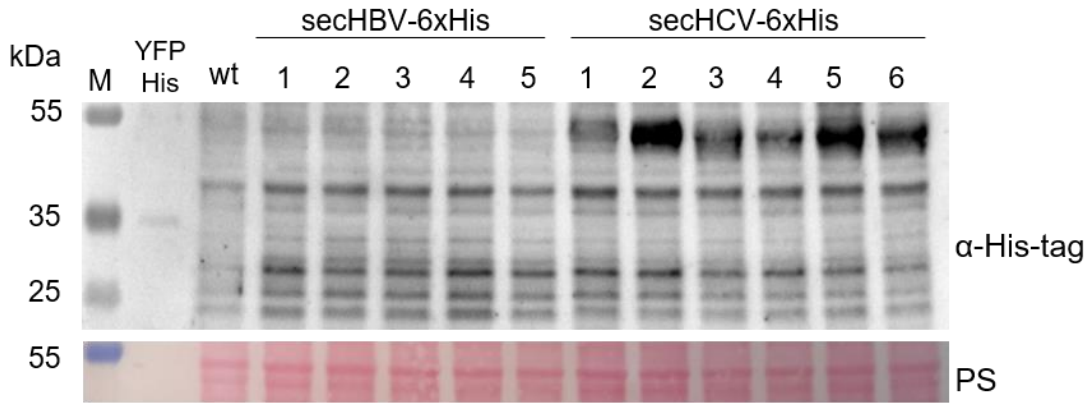


Figure 23: Immunoblot analysis of *P. purpureum* lines expressing 6xHis-tagged secHBV and secHCV antigens
Samples of 10 µg of total protein derived from the transformed lines were loaded next to a wild-type control and a positive control of 10 ng His-tagged YFP. The immunoblot was analyzed using an anti-His-tag antibody. secHCV-6xHis accumulates slightly below 55 kDa PS: Ponceau staining. M: molecular weight marker; wt: wild-type control.

Initially, the same approach was applied for the secHBV antigen (Figure 23), but later our collaboration partners informed us that the native confirmation of the HBV S/preS1 antigen in the SVPs does not allow for effective His-tag purification. Nevertheless, *P. purpureum* was transformed with the secHBV-6xHis construct and analyzed by immunoblotting, but no expressing transformant could be identified and this project was not pursued further (Figure 23).

For the secHCV-6xHis construct, the 3xHA tag sequence was exchanged with the sequence for the 6xHis tag, with no other changes introduced in the construct. Six Zeocin™-resistant secHCV-6xHis transformants were analyzed by immunoblotting using an anti-His-tag antibody. The 6xHis-tagged secHCV antigen was detected as running at approximately 55 kDa and was strongly expressed in line 2 and 5. The 6xHis-tag antibody caused many cross-reactions, making it difficult to detect lowly expressed proteins. This might be an explanation why no specific signal could be seen for the secHBV-6xHis antigen lines.

P. purpureum secHCV-6xHis line 5 was used for growth in a bioreactor in order to obtain enough cell material. This material was later used by our collaboration partner to isolate and purify the HCV antigen for immunization trials by parenteral administration in mice.

4.4.7 Photobioreactor-based upscaling of antigen production in *P. purpureum*

All HBV and HCV antigen variants were shown to be successfully expressed in *P. purpureum*. The next step was the accumulation of enough algal material to be able to perform mice immunization trials. As mentioned earlier, the immunization trials were split into an orally

administration and a parenterally administration experiment. Both trials have different requirements for the antigen that is used. Upon oral administration, the transgenic algal biomass is fed to the mice. For this, it is crucial to have an estimate of how much total antigen is fed to the mice. Also, the stability of the antigen plays an important role. The 3xHA-tagged secHCV2A_a line and the untagged HBV S/preS1 line 4 were used for feeding experiments. For the parenteral administration, the antigen has to be purified and quantified before injection into mice. It is important to have a good purification scheme for isolating the HBV and HCV antigens. The purification was performed by our collaboration partners directly before administration into mice. They employed the His-tag purification approach for the HCV antigen, and isolated the SVPs via ultracentrifugation for the HBV antigen.

To achieve the above stated requirements, it was necessary to have a constant supply of antigen-producing *P. purpureum* biomass. An initial harvest of algal biomass was performed after growing the red algae in liquid culture in 250 mL Erlenmeyer flasks and harvest of the cells before the stationary phase. This was done for initial mice feeding experiments using the secHCV2A_a expressing line. As *P. purpureum* is strictly dependent on good aeriation, as demonstrated in section 4.1.1, the total volume in the flask had to be as small as possible to ensure a satisfactory surface to volume ratio. Due to the employment of many flask, this experiment was not infinitely scalable, and in the end, 260 mg of dry biomass were obtained from 800 mL culture. Overall, 12.825 µg of secHCV antigen were fed three times to five mice, which was by far not sufficient to trigger an immune response.

To scale up the production, the 3xHA tagged secHCV2A_a line, the secHCV-6xHis line 5 and the untagged HBV S/preS1 line 4 were grown in a PBR. The secHCV2A_a-expressing *P. purpureum* strain was grown for the mice feeding experiments, whereas the secHCV-6xHis line was grown for the purification and subsequent injection-based mice trials. The HBV S/preS1 expressing line 4 was grown for both feeding and injection experiments. All PBR experiments were performed without the addition of Zeocin™.

Below, the results for the secHCV antigen are described. Two independent PBR runs were performed to accumulate biomass for strains expressing secHCV-3xHA and secHCV-6xHis. The first run was a cultivation of the wild type and the first batch (PBR 1) of secHCV-3xHA expressing *P. purpureum* employing two PBRs per line. The second run was a cultivation of the second batch (PBR 2) of the secHCV-3xHA expressing *P. purpureum* line using two PBRs and a cultivation of secHCV-6xHis expressing *P. purpureum* line using one PBR. An exemplary growth curve of the secHCV2A_a line is shown in Suppl. Figure 4. It shows a slower growth and an extended lag phase compared to the PBR-grown wild-type culture, indicating potential ER stress, induced by continuous high expression of the antigen.

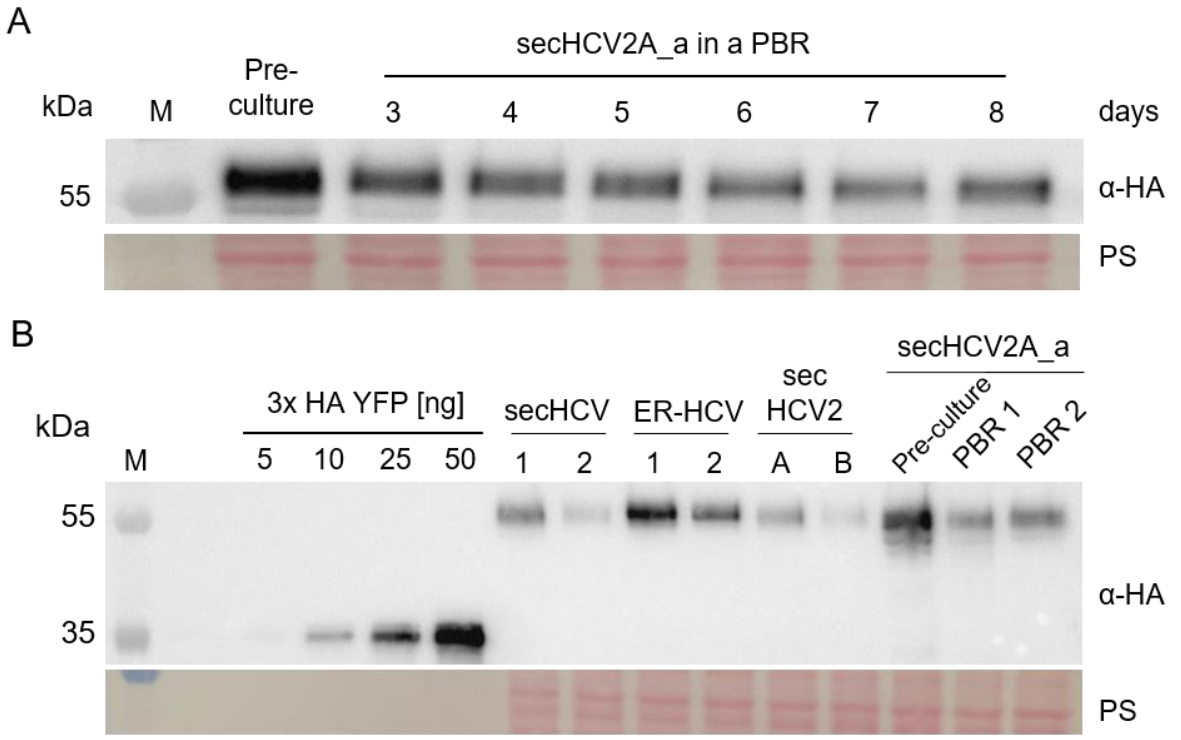


Figure 24: Quantification of secHCV antigen accumulation in a secHCV2A_a-expressing *P. purpureum* line grown in a PBR
A: Immunoblot analysis of secHCV antigen accumulation throughout the cultivation in a PBR. 10 µg of total protein obtained from the pre-culture was loaded as a control together with 10 µg of total protein sampled on the indicated days after inoculation.
B: Quantification of different HCV antigen variants using a 3xHA YFP standard. 10 µg of total protein obtained from cultures grown either in the bioreactor (PBR1 and PBR2, indicating two independent batches) or in flasks (all other samples) were analyzed using the anti-HA antibody. M: marker; PS: Ponceau staining.

To follow antigen accumulation in the PBR-grown algal expressor lines, samples were taken on the indicated days, and the secHCV-3xHA antigen accumulation was compared to the antigen accumulation of the pre-culture grown in a flask (Figure 24A). The secHCV-3xHA antigen was stably expressed in the PBR throughout the cultivation time and only decreased slightly over time. However, the antigen accumulation in the PBR-grown culture compared to the pre-culture, when sampled in the exponential growth phase, was decreased by approximately 50%. This reduction is also seen in Figure 24B where the pre-culture accumulated secHCV-3xHA to approximately 0.39% of total protein, whereas in the PBR runs 1 and 2, secHCV-3xHA accumulated to 0.21% and 0.22%, respectively. It is noteworthy that the protein levels of the HCV antigen seem to differ between this quantification and the quantification made in section 4.4.2. An important difference was that the cultures used here were quantified almost 1.5 years after transformation, whereas the cultures tested in section 4.4.2 were sampled freshly after transformation. It seems that over the course of 1.5 years, the antigen expression level halved, although the selective agent Zeocin™ was always added to the growth medium of the maintained cultures.

The results of the PBR runs for biomass accumulation and antigen production are summarized in

Table **6**. In total, the first PBR run yielded 2025 mg dry weight of wild type *P. purpureum* and 1400 mg of secHCV-3xHA-expressing *P. purpureum*. The difference in weight is most likely due to differences in cell density at the end of the cultivation (see Suppl. Figure 5). The

Table 6: Summary of the PBR runs for biomass generation for mice immunization trials

DW: Dry weight; Ag: antigen; Numbers in parentheses indicate the number of feedings, and the number of mice used for the feeding trials; n.a: not applicable

	Wild type	secHCV 3xHA 1st batch	secHCV 3xHA 2nd batch	secHCV-6xHis
Dry weight [mg]	2025	1400	1040	850
Protein of DW [%]	39.12	42.48	22.1	19.43
Total protein [mg]	792.18	594.72	228.80	161.50
DW feeding /mouse [mg]	140 (2x7)	100 (2x7)	40 (2x7)	n.a
Ag concentration [%]	n.a	0.21	0.22	n.a
Ag per DW [µg/mg]	n.a	0.892	0.484	n.a
Ag/mouse [µg]	n.a	89.21	19.36	n.a

second PBR run yielded 1040 mg of dry weight for secHCV-3xHA expressing cells and 850 mg of secHCV-6xHis expressing cells.

The dry weight values indicate the final weight sent for mice trials after washing. Notably, the total protein content differs by a factor of two between the runs. In run one, a total protein content of 39.12% for the wild type and 42.48% for the first batch of secHCV-3xHA was achieved. In the second run, a protein content of only 22.1% for the 2nd batch of secHCV-3xHA and 19.43% for secHCV-6xHis was achieved. As protein contents have been measured by the same method, the difference in absolute protein yield is most likely due to differences in cultivation conditions between the two independent runs (e.g. slightly differing aeriation conditions, varying cell densities, differences in light intensities). This wide range of total protein content in the dry mass of *P. purpureum* was also described in the literature and reached from 10% to 56% of dry weight (Li et al., 2019; Safi et al., 2013). Notably, the HCV antigen accumulation per µg of total protein was similar in both runs, and was 0.21% for the first batch and 0.22% for the second batch (see Figure 24B). In the end, a total of 0.892 µg antigen per mg dry weight could be obtained for the 1st batch and 0.484 µg antigen per mg dry weight for the 2nd batch.

Because the first mice immunization trials with the flask-grown cultures did not yield positive results, our collaboration partners questioned the digestibility of *P. purpureum* cells by mice, and thus, the availability of the antigen to intestinal immune cells. Indeed, the thick cell wall composed of polysaccharides makes *P. purpureum* a hard to digest microalga (Niccolai et al., 2019). To investigate this in more detail and to potentially overcome this problem, freeze-dried biomass, as it was sent to the collaboration partners for the first mice immunization trials, was analyzed by microscopy (Suppl. Figure 6). After suspension in lsASW medium, the cells seemed intact, with almost no broken cells being visible. To make the antigen more available to the mice, the freeze-dried cells were mechanically broken up by using a bead mill. The processed cells were then analyzed in the microscope and many half-broken and lysed cells were detected. Therefore, the bead-milling method was chosen for processing of all samples in all subsequent mouse-feeding experiments.

Seven mice were fed twice with either the wild-type algal biomass or the biomass from the secHCV-3xHA-expressing line for oral immunization experiments. The 2025 mg of wild-type dry biomass were sufficient to feed 140 mg per serving, limiting the amount of the biomass of the antigen-expressing line that could be fed to the mice to exactly this number. Due to the higher antigen per dry weight ratio of the first batch, 100 mg of this dry mass were combined with 40 mg of the second batch to a total of 140 mg. This means that 108.57 µg of HCV antigen (tagged with 3xHA) was fed per feeding (89.21 µg + 19.36 µg).

The biomass of the secHCV-6xHis-expressing line was used for affinity purification of the antigen by our collaboration partners. Approximately half of the dry biomass was used for optimization experiments, but 400 mg could be used for final purification of the HCV antigen. In the end, 200 µg of pure HCV antigen could be obtained, amounting to a concentration of 0.5 µg antigen per mg dry weight. This value is comparable to the levels obtained in the second batch of the 3xHA-tagged antigen. Taking the total protein amount of secHCV-6xHis into account, the HCV antigen level was approximately 0.25% of the total protein. The true values are probably even higher, due to some loss of antigen during purification. The purified HCV antigen from the secHCV-6xHis line was used for injection-based immunization trials in mice.

For the untagged chimeric HBV S/preS1 protein expressing line, a total of 2940 mg dry biomass was obtained. 990 mg were processed by bead milling and, together with 990 mg of dry wild type biomass were used for feeding-based mouse immunization experiments. The remaining 1950 mg were sent to our collaboration partners for SVP purification and subsequent injection-based mouse immunization experiments.

To estimate the antigen accumulation of the untagged S/preS1 antigen, the antigen amount of the 3xHA-tagged secHBV was determined using a 3xHA-tagged YFP as a standard and

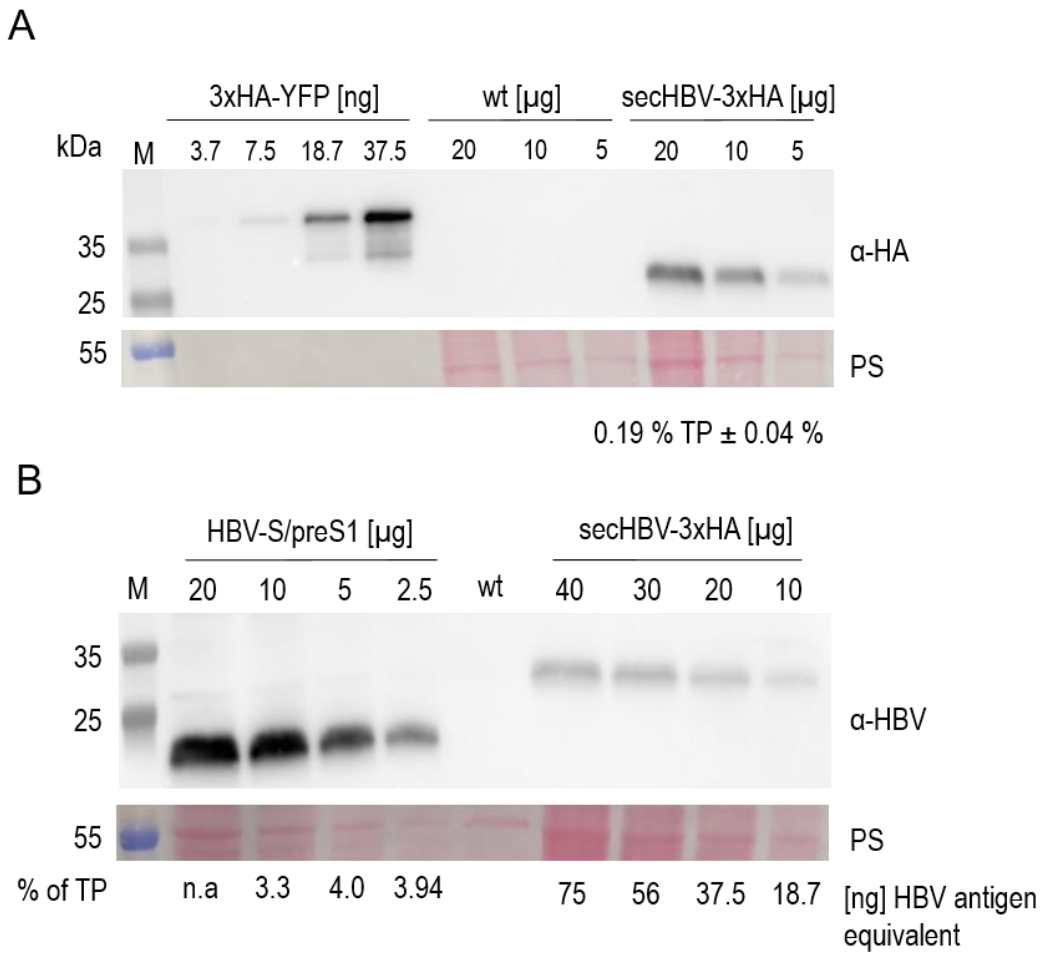


Figure 25: Quantification of untagged HBV S/PreS1 protein accumulation.
A: 3xHA tagged secHBV was quantified by loading samples of 20, 10 and 5 µg of total protein from the antigen-expressing line next to a standard of 3xHA-tagged YFP, and detecting the proteins with an anti 3xHA antibody. Numbers below the western blot indicate the calculated yield of antigen in the expression line. PS: Ponceau staining
B: Untagged HBV S/preS1 protein quantification using the tagged secHBV antigen as a standard. Different amounts of total protein from the respective antigen-expressing line were loaded next to each other, and protein was detected using an anti HBV-S protein antibody. Numbers below the western blot indicate the antigen equivalent loaded in the case of cytHBV-3xHA (calculated from panel A), and the calculated yield in the case of the untagged HBV-S/preS1 antigen. PS: Ponceau staining

detecting it with a 3xHA antibody. Afterwards, the quantified 3xHA-tagged secHBV antigen was used as a standard for quantification of the untagged HBV antigen using an HBV S protein antibody (Figure 25). Because both antigens have the same amino acid sequence and only differ in the presence versus absence of the 3xHA tag and the SP, the HBV S protein antibody should exhibit equal binding affinities to both antigens after denaturing SDS-PAGE. A total yield of approximately 0.19% of 3xHA tagged secHBV antigen per total protein was quantified (Figure 25A), comparable to the quantification performed in Figure 17C. This value was then used to calculate the antigen equivalent that was loaded as a standard to quantify the untagged HBV-S/preS1 antigen. For this, 40, 30, 20 and 10 µg total protein derived from the

exact same sample as in Figure 25A were loaded, corresponding to 75, 56, 37.5 and 18.75 ng HBV antigen, respectively. A total average yield of 3.8% of the total protein was obtained for the untagged HBV-S/preS1 antigen expressed in *P. purpureum* (Figure 25B). The concentration in the lane with 20 µg protein loaded from the HBV-S/preS1-expressing line could not be quantified due to oversaturation of the signal.

After quantification of the antigen yield, an estimate of antigen content per mg dry weight could be performed, which is particularly important for the feeding-based immunization experiments. BCA-based total protein quantification of the dried biomass gave a total protein concentration of approximately 22.25%, which is comparable to the second batch of 3xHA-tagged secHCV antigen and the secHCV-6xHis antigen (see Table 6). Taking this into account, 1 mg of dried biomass derived from untagged HBV-S/preS1 expressing cells contains approximately 8.45 µg HBV-S/preS1 antigen. This is around 20 times more than for the secHCV antigen. This means that, in total, up to 8.43 mg antigen could be used for feeding-based immunization experiments (8.45 µg times 990 mg dry mass), and that up to 16.4 mg could theoretically be purified and used for injection-based immunization experiments (8.45 times 1950 mg dry biomass).

The results of the immunization trials for the HCV antigen are discussed in section 5.5.3. The immunization trials for the HBV antigen are not yet completed.

4.4.8 Employment of a strong endogenous promoter to enhance secHCV antigen accumulation and secretion

Although the expression values of the secHCV antigen were initially already at approximately 0.5% of the total protein, attempts were made to increase expression even further. New endogenous promoters were identified as described in section 4.2.4 using *YFP* as a reporter system, and it was shown that compared to the *Actin* promoter-driven lines, the chlorophyll binding protein (*ChlBP*) promoter-driven lines accumulated YFP to higher levels (Figure 10). To test if this promoter can also boost the expression of the secHCV antigen, the A*ctin* promoter of the secHCV-3xHA construct was exchanged with the *ChlBP* promoter, and *P. purpureum* was transformed with the new construct. Figure 26A shows the immunoblot analysis of the selected transformants, where the ChIBP::secHCV lines were compared to the lines expressing the *Actin* promoter driven secHCV. Protein derived from cellular pellets showed a band at approximately 55 kDa corresponding to secHCV in both the *Actin* and the *ChlBP* promoter-driven lines. The new expressing lines *ChlBP*::secHCV 5 and 7 expressed the antigen to high levels, comparable to the *Actin* promoter-driven antigen expression. At the same time, the culture medium of the transformation lines was harvested and an equivalent of 2 mL was analyzed via immunoblotting. Here, a distinct antigen accumulation at 55 kDa,

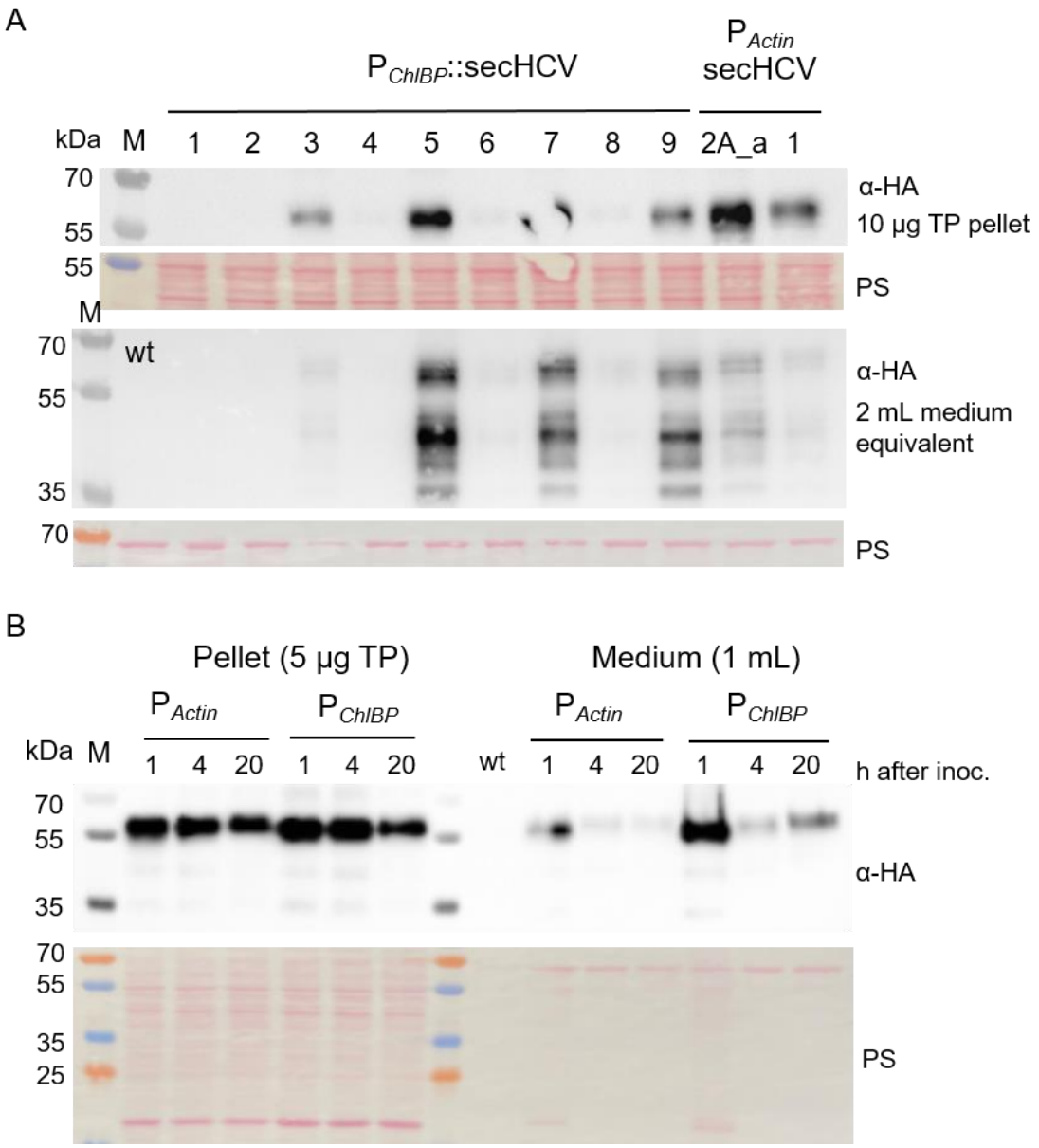


Figure 26: Immunoblot analysis of lines expressing secHCV-3xHA under the control of the strong endogenous *ChlBP* promoter
A: Samples of 10 µg of total protein derived from nine independent Zeocin™-resistant lines were analysed by immunoblotting using an anti-HA antibody, and compared to the secHCV 2A_a and the secHCV 1-expressing line that are under transcriptional control of the *Actin* promoter. In addition, protein from the supernatant of these lines was collected, and a medium equivalent of 2 mL was analysed by immunoblotting. A wild type supernatant sample was loaded as control. Ponceau staining was used as loading control.
B: Growth medium supplemented with protease inhibitor was inoculated with an initial cell density of 5 x 10^6 cell/mL of the P_{Actin}::secHCV2A_a line and the P_{ChlBP}::secHCV 5 line. Protein samples were taken at the indicated time after inoculation and analyzed via immunoblotting. Ponceau staining served as loading control and as a control for lysed cells in the medium.

40/45 kDa and 35 kDa was visible, indicating that secHCV was in fact secreted into the medium. The secretion in the newly constructed lines was much stronger than the secretion from the previously generated, *Actin* promoter-driven secHCV-expressing lines. The bands of lower molecular weight that were detected in the culture medium could be either degradation

products of the antigen or differentially glycosylated antigen proteins that occur after passage through the Golgi apparatus, where they obtained complex glycans.

To further investigate this, an experiment was designed, where cultures of the *ChlBP*::secHCV5 and *Actin*::secHCV2A_a line were inoculated with a density of 5 x 10^6 cells/mL and protease inhibitor was added to the medium (Figure 26B). Samples of the cell pellet and the medium were taken at 1 h, 4 h and 20 h after inoculation and addition of protease inhibitor and analyzed by immunoblotting. The intracellular accumulation of the antigens driven by the two promoters showed similar secHCV accumulation that decreased after 20 h. Antigen accumulation in the medium, on the other hand, increased during the cultivation period, especially for the *ChlBP* promoter-driven construct, which showed much higher accumulation of the secHCV antigen compared to the A*ctin* promoter-driven construct after 20 h. The high antigen levels in both lines at 1 h is most likely an artifact from the presence of intracellular secHCV due to lysed cells in the medium. This can be seen in the Ponceau staining, where at 1 h, a band at a size of approximately 17 kDa is visible (Figure 26B). This band corresponds to phycoerythrin, which is a water-soluble protein of the phycobilisome, and can serve as marker for cell lysis. The addition of protease inhibitor to the liquid culture inhibited the degradation of secHCV and only a single band at 55 kDa was visible, strongly suggesting that the previously seen bands at 40/45 and 35 kDa are in fact degradation products.

This experiment demonstrates that secHCV accumulation can be boosted by using one of the newly identified strong endogenous promoters. Furthermore, we could confirm that secHCV was indeed secreted into the medium. Although intracellular antigen levels generally correlated with the levels detected in the medium, antigen accumulation in the medium was much higher for the new *ChlBP* promoter constructs. Lower secretion, together with protein degradation in the medium, may be the reason why the accumulation of the secreted antigens could not be detected in earlier experiments.

In summary, *P. purpureum* exhibits significant potential for expressing complex and biotechnologically valuable proteins to high levels. Additionally, this microalga displays the ability to secrete antigens, thus facilitating easy purification for future applications, despite challenges posed by proteases affecting the stability of the secreted protein. The utilization of strong endogenous promoters offers the potential to enhance antigen expression, thus establishing *P. purpureum* as a promising biotechnological host organism.

5 Discussion

5.1 *P. purpureum* strain SAG 1380-1d shows normal microalgal growth characteristics but is unable to utilize organic carbon sources

The red microalga *P. purpureum* is already being utilized for biotechnological purposes due to its ability to produce valuable natural products such as PUFAs, sulphated polysaccharides, and phycoerythrin (Li et al., 2019). Previous research has primarily focused on optimizing growth conditions to enhance the production of these natural compounds. However, due to its status as a non-model organism, limited information is available regarding the physiological growth conditions and the underlying molecular mechanisms in different strains of *P. purpureum*. This includes factors like optimal medium composition, additives, and the strains' ability to grow under mixotrophic or heterotrophic conditions.

In the three major microalgal strain collections—CCAP (www.ccap.ac.uk), SAG (sagdb.uni-goettingen.de), and UTEX (utex.org)—16 strains of *P. purpureum* are deposited that were collected from diverse habitats around the globe (e.g., saltwater, brackish water, fresh water). These strains are likely to exhibit significant genetic diversity, but only *P. purpureum* strain CCMP 1328 has been sequenced (Bhattacharya et al., 2013; Lee et al., 2019). During the work conducted for this thesis, it was observed that the cloned promoter sequences for the new promoter-YFP constructs partly differed from the sequences in the genome draft. In fact, one sequence (associated with the gene for the hypothetical protein FVE85_6371) could not be amplified at all. A recent study comparing the mitochondrial genomes of two *P. purpureum* strains showed a large size difference (Kim et al., 2022). This suggests the existence of substantial genetic variation among strains, underscoring the importance of carefully investigating the SAG1380-1d strain used in this thesis.

In this study, attempts were made to promote the growth of *P. purpureum* strain SAG 1380-1d. This was approached by varying the quantities of macronutrients and micronutrients, and supplementing the medium with organic carbon sources to enable mixotrophic or heterotrophic growth. While commonly used photoautotrophic media for *Porphyridium* growth include f/2, ASW, KOCH, and Pringshein media (Li et al., 2019), Su et al. (2016) found ASW medium to be the most effective in terms of achieving high biomass. Further research by Kathiresan et al. (2007) investigated different salt compositions of the ASW medium to enhance phycoerythrin production in *P. purpureum*. They discovered that halving the sodium chloride content was advantageous for biomass generation. This "low salt ASW medium" (lsASW) was applied in this study using strain SAG 1380-1d, and resulted in improved growth.

The new medium composition yielded a higher final cell density in the stationary phase (1.46 x 10^6 cells/mL) compared to the standard ASW medium (1.18 x 10^6 cells/mL; Figure 2B).

P. purpureum is a mesophilic red alga with considerable salt tolerance, capable of growth in sodium chloride concentrations ranging from 0 to 46 g/L (Golueke and Oswald, 1962; Lu et al., 2020). Previous studies have shown that salinity significantly affects growth, phycoerythrin content, exopolysaccharide production, and lipid content, albeit with conflicting results. A general consensus suggests that optimal growth occurs at salinities between 10 to 20 g/L, while lower or higher salt concentrations hinder growth (Kathiresan et al., 2007; Li et al., 2023; Li et al., 2019). The contradicting results between the studies are most likely due to the employment of different strains or different media compositions.

The results for SAG 1380-1d are in line with the existing literature, and demonstrated a positive impact of reduced salinity. A future area of investigation could be to determine the ideal medium composition for increasing total protein accumulation per cell. While the primary aim of this thesis was to express recombinant proteins in *P. purpureum*, redirecting metabolism towards protein synthesis—rather than fatty acids or polysaccharides—could enhance the yield of recombinant proteins per dry weight in this microalga.

Another parameter that was investigated is the availability of CO_2 in the culture. Since the minimal lsASW medium lacks an organic carbon source, *P. purpureum* strictly relies on the photosynthetic conversion of CO_2 to generate biomass. As expected, a higher availability of CO_2, achieved through the use of cotton plugs for liquid media or Leukopor® for solid media, significantly enhanced the growth of *P. purpureum* (Figure 2).

The initial determination of the doubling time in standard ASW medium at a lower temperature (21°C) was provided by Jones et al. (1963), reporting a doubling time of 20-28 h, depending on the culture volume. This aligns well with the 19 h doubling time observed in cotton-sealed cultures. The average ambient CO_2 concentration is approximately 0.038%, and it has been demonstrated that providing a higher CO_2 concentration—by bubbling either ambient air or air enriched with CO_2 into the culture—is beneficial for the growth of all photoautotrophic microalgal species (Singh and Singh, 2014). Another source of inorganic carbon is sodium bicarbonate, which can be added to the medium. Research has shown that the addition of 0.8 g/L of sodium bicarbonate promotes both *P. purpureum* growth and PUFA accumulation (Su et al., 2016).

Next, it was tested whether *P. purpureum* SAG 1380-1d is able to grow mixo- or heterotrophically by utilizing different organic carbon sources. Mixotrophic growth, which involves the simultaneous utilization of organic carbon and fixation of inorganic carbon by photosynthesis, could substantially shorten the generation time of the microalga, potentially

streamlining laboratory processes. The green alga *C. reinhardtii*, for instance, can achieve growth rates up to 1.5 times higher by supplying 2 g/L acetate into the medium (TAP medium) compared to growth under photoautotrophic conditions at 30 µmol photons m^{-2} s^{-1}, and it reaches the stationary phase several days earlier (Young et al., 2022). Furthermore, achieving heterotrophic growth could facilitate the study of photosynthetic mutants in *P. purpureum*. The initial draft genome of *P. purpureum* CCMP 1328 indicated the presence of several hexose transporters, suggesting a potential role of those in mixotrophic or heterotrophic growth (Bhattacharya et al., 2013).

To assess whether the *P. purpureum* strain SAG 1380-1d can grow under heterotrophic growth conditions, varying concentrations of sodium acetate, glucose, sucrose, and glycerin were added to the lsASW medium, and the cultures were incubated in the dark for one month. Following this period, no notable differences were observed between the non-supplemented control and the supplemented cultures (Figure 3A). The only publication that describes heterotrophic growth in the dark was published by Oh et al. (2009) and described slow but steady growth of *P. purpureum* in the dark through the addition of 5 g/L glucose. This growth rate was further improved with higher glucose concentrations. When 10 g/L glucose was added, the culture exhibited a doubling time of around 4.5 days (in comparison to 18 h light / 6 h dark conditions, with a doubling time of 23.8 hours). Several parameters differed in this publication compared to the experiments done in this work. This group used a strain from the 'Korea Marine Microalgae Culture Center' sampled from the north Atlantic that is not further specified. Other differences were the usage of f/2 medium that partially uses real seawater in the recipe, the constant aeriation with 2% CO_2, even in non-illuminated cultures, and the growth inside 5 L fermenters. It could be worthwhile to replicate all the conditions of this experiment to check whether SAG 1380-1d would grow in the dark under these conditions.

Another publication described mixotrophic growth of *P. purpureum* by addition of glucose, glycerol or sodium acetate to the minimal growth media (Jiao et al., 2018). The addition of 0.5% glucose to the medium led to a significant increase in the final biomass harvested after 18 days of mixotrophic cultivation (in the stationary phase), from 9.58 g/L to 15.85 g/L, corresponding to a 1.5-fold increase. Similar promoting effects on biomass accumulation were observed with glycerol and sodium acetate supplementation. This study utilized ASW medium under conditions similar to those employed in this work. However, when the supplementation of lsASW medium with 0.5% glucose was tested under continuous light conditions of 100 µmol photons $m^{-2}s^{-1}$, no improvement in growth was detected during the exponential growth phase (Figure 3B). The key difference between the published experiment and the work conducted in this thesis is the strain used. The authors stated that they specifically screened for the employed strain CoE1 and maintained it in the working group. No details about their screening

approach was given, but in case the CoE1 strains was specifically screened for the utilization of organic carbon sources, it is obvious that growth was promoted under mixotrophic conditions. The strain SAG 1380-1d showed no mixotrophic growth with the addition of 0.5% glucose, but might exhibit mixotrophic growth under other conditions that were not tested in this thesis.

The last parameter tested for optimized growth of SAG 1380-1 was the possibility of cultivation inside a photobioreactor (PBR). PBRs with a total volume of up to 850 mL were employed that are most comparable to an airlift column photobioreactor. *Porphyridium* was already successfully grown in this kind of bioreactor (Luo and Al-Dahhan, 2012). The benefits of large-scale cultivation are described in 4.4.7., where several grams of biomass had to be collected in order to conduct immunization experiments. Some microalgal species can only be cultivated in a PBR with significant efforts. For instance, the cell wall-less UVM11 mutant of *Chlamydomonas* exhibits high expression levels, but its lack of a cell wall makes it vulnerable to shear forces within a PBR (Wang and Lan, 2018).

P. purpureum strain SAG 1380-1d was able to grow in a PBR, aeriated with 2% CO_2 at an airflow of 50 L/min. The initial lag phase observed in PBR-grown cultures after inoculation from flasks is a common phenomenon, likely attributable to the adaptation of cells to new growth conditions within the PBR, including higher shear forces, CO_2 bubbling, increased light intensity at the periphery of the PBR, and changes in cell density during inoculation (Luo and Al-Dahhan, 2012).
The cell density at the end of cultivation was most likely not the maximum reachable density, but the cultivation had to be terminated due to significant foam formation. The maximum cell densities described in the literature differ significantly ranging from 4.11×10^6 cells/mL to 1.37×10^8 cells/mL, depending on culture conditions, sampling time and the medium used (Li et al., 2019). Comparisons between microalgal growth in different PBR systems and media are difficult, as evident from the high variation in reported cell densities. The biomass that was obtained in this thesis, which is approximately 2.5 g/L including polysaccharides and 1.46 g/L excluding polysaccharides, aligns with literature values ranging from 0.95 g/L to 18.81 g/L. The variation is largely due to differences in culture conditions and, notably, harvesting times (Li et al., 2019).

The polysaccharide content of *P. purpureum* ranging from 15% to 50% of the dry biomass, is influenced by culture media and growth conditions (Fuentes-Grünewald et al., 2015; Su et al., 2016; Li et al., 2019). The literature values correspond well with the reduction in biomass from 2.5 g/L to 1.46 g/L after removing the polysaccharides by washing (Table 2), which amounts to a 42% exopolysaccharide content. Exopolysaccharides begin to accumulate to high levels

in the beginning of the stationary phase, most likely induced by the limitation of nitrogen in the medium (Arad et al.,1988). At the same time, a high nitrogen content seems to channel the metabolism towards protein production rather than carbohydrates, increasing total protein amounts per biomass (Li et al., 2019). Thus, increasing the nitrogen content of the culture media for PBR growth can be beneficial for recombinant protein production.

When comparing *Porphyridium* to *C. reinhardtii* cultivated in a similar PBR under photoautotrophic conditions, the green alga reached similar biomass accumulation of approximately 1.8 g/L after 160 hours (6.6 days), but at a higher light intensity of 300 µmol photons $m^{-2}s^{-1}$ (Tebbani et al., 2013). It is worth noting that the PBR system used in this thesis is relatively basic, leaving room for optimization to improve growth conditions of *P. purpureum*. Particularly in the context of recombinant protein production, further research is needed to establish an optimized growth protocol for *P. purpureum*.

5.2 Improved methods for transformation and single colony formation of *P. purpureum*

5.2.1 Optimization of the biolistic transformation protocol

Two different nuclear transformation protocols were independently established for *P. purpureum* in the recent years. One protocol described *Agrobacterium*-mediated transformation to express a mutated phytoene desaturase gene that confers resistance to the herbicide norflurazon by stable integration into the genome (Prasad et al., 2018). This transformation gave rise to an average of 15 to 25 transformants per plate, depending on cultivation conditions. The DNA integration was checked using Southern blot analysis on genomic DNA digested with two different restriction enzymes and showed random integration of multiple copies of transgenic DNA into the *P. purpureum* genome. Although resistant colonies were obtained, this publication failed to demonstrate expression of other transgenes, and lacks any data at the protein level.

The second method involved biolistic transformation of *P. purpureum* using a circular transformation vector containing the *ble* gene, which provides resistance to Zeocin™ (Li and Bock, 2018) as well as the prokaryotic resistance marker for ampicillin resistance in the backbone of the plasmid. The Zeocin™-resistant colonies maintained the transformation plasmid episomally, with up to 20 copies per cell. All transgenic strains generated with the same vector displayed identical hybridization patterns in Southern blot analysis when compared to the transformation plasmid. Importantly, the presence of plasmids was confirmed directly via plasmid rescue experiments. These experiments involved the transformation of *E. coli* with total DNA isolated from the transgenic strains, resulting in the recovery of ampicillin resistant colonies containing plasmids that are identical to the algal transformation vector.

Moreover, protein expression was verified through microscopic and immunoblot analyses, demonstrating the presence of GFP within *P. purpureum* cells.

However, in my own experiments, this transformation method using particle bombardment did not produce satisfactory results, often resulting in no or only one positive transformed line per plate. This posed a significant challenge in generating a sufficient number of transgenic lines and hindered research progress. Consequently, substantial efforts were dedicated to enhancing the efficiency of biolistic transformation in *P. purpureum*.

The modified protocol consisted of bombardment of the *P. purpureum* cell lawn on plates without the addition of Zeocin™, subsequent washing of cells with lsASW medium followed by plating of the washed cells onto Zeocin™-containing plates. The *ble* gene encodes a protein that binds the DNA double-strand break (DSB)-inducing Zeocin™ in a 1:1 stoichiometric ratio and thus deactivates it (Stevens et al., 1996). To test whether an incubation period in liquid medium without Zeocin™ has any positive impact on formation of transformed colonies, a regeneration time of 4 h and 20 h in lsASW was tested. During the regeneration time, potential transformed cells could start accumulating the ble protein while not being exposed to Zeocin™.

Figure 6A demonstrated a clear positive impact of the modified protocol on colony formation. One reason for the improvement could be the omission of Zeocin™ from the medium plates used for bombardment. It was demonstrated in *C. reinhardtii,* that double-strand breaks are induced after only a 1-minute incubation with Zeocin™ (Chankova et al., 2007). In the original transformation protocol, the cells came into contact with Zeocin™ for approx. 30 min before the bombardement with the transformation plasmid, which may give the selective agent enough time to already induce harmful DNA double-strand breaks. Together with the stress of the bombardment itself, this could have led to a reduced transformation efficiency, leaving only few transformaed cells surviving. Bombardment of cells prior to Zeocin™ exposure is beneficial for cell survival and colony formation later on the Zeocin™-containing selection plate. During the regeneration phase the transformed cells have time to produce at least small amounts of the ble protein. In *E. coli,* it takes approximately 1 min to transcribe a 1 kb gene into mRNA, followed by 1 min for the translation of a 300 amino acid long protein (Shamir et al., 2016). Assuming similar synthesis rates in *P. purpureum*, since the *ble* gene does not contain any introns, has a size of only 330 bp and is located on a plasmid present in several copies in the cell, many copies of the ble protein can be produced within 30 min before exposure to Zeocin™. Under this assumption, the transformed cells should have accumulated enough ble protein that then binds to the Zeocin which otherwise is detrimental to the non-transformed algal cells.

Another reason for the increased efficiency of the new transformation method could be the washing of the cell debris derived from broken cells after bombardment. Figure 6C shows that the supernatant of washed cells after bombardment is green, indicating the release of chlorophyll from broken cells. The cell debris could interfere with regeneration of transformed cells by releasing toxic substances or by releasing negative growth signals for the surrounding cells. While Zeocin™ also leads to cell death and eventually cell debris, the negative impact of cell debris accumulation in this case likely would need some time. Zeocin usually first induces DSBs that accumulate over time. Zeocin-induced DSB will block cell division, leading to growth arrest, followed by cell death and finally cell lysis after long-term exposure (Čížková et al., 2019). This effect of delayed cell death and lysis would give the transformed cell enough time to grow before toxic cell debris from non-transformed cells are released into the medium.

Biolistic transformation is generally used for transformation of cell wall-containing microalgae and diatoms (León and Fernández, 2007). While *C. reinhardtii* could easily be transformed biolistically without the need of washing and changing selection plates (Ramesh et al., 2011), transformation of the diatom *P. tricornutum* employed a similar approach as the newly modified *P. purpureum* transformation protocol. For diatom transformation, bombardment takes place on Zeocin™-free plates, before transferring the cells to antibiotic containing plates (Apt et al., 1996). Many microalgal species like *Chlamydomonas* and *P. tricornutum* were first successfully transformed using the biolistic approach (Mayfield and Kindle, 1990; Apt et al., 1996) before more advanced and effective nuclear transformation methods like glass bead agitation or electroporation were developed (Kindle, 1990; Brown et al., 1991). Using glass bead agitation is usually only suitable for cell wall-less microalgae and did not yield any transformants when tested on *P. purpureum* in the course of this work. Electroporation is generally used for cell wall-less strains as well, but the development of square-wave electroporation enabled efficient electroporation-based transformation of cell-walled *C. reinhardtii* strains (Wang et al., 2019). It would be worthwhile to develop an efficient square-wave electroporation-based transformation protocol for *P. purpureum*, as it could greatly simplify and accelerate the generation of transgenic lines.

5.2.2 Variation of expression levels between transformants can be eliminated by colony isolation from single cells

In the course of this thesis, many transgenic lines of *P. purpureum* were generated of which some showed variable expression strength between lines of the same construct, namely the HCV and HBV antigen-expressing lines and some of the lines testing new promoters (Figure 11A, Figure 17B, C). As the transformation vector is maintained episomally and is not integrated into the genome, position effects on expression strength resulting from random integration into transcriptionally differentially active or inactive regions can be excluded.

Transgene silencing due to position effects and epigenetic silencing are usually detected in transformants of *C. reinhardtii,* making it obligatory to screen many transformants to obtain a strong expressor (Zhang et al., 2014). In *P. purpureum,* the variation has to arise from a different source. Another possibility would be variable numbers of plasmid copies, but Li & Bock, (2018) demonstrated that the maintained episomal plasmids are present at a constant number of 20 copies per cell in independent transformants of the same construct. This finding is also supported by Southern blot analyses performed in the course of this thesis, where hybridization signals for the plasmid DNA exhibited the same intensity in different strains (data not shown).

Microscopic analysis of *YFP* transformants revealed that, within a population of cells derived from a single primary transgenic colony, not all cells were showing fluorescence (Figure 11B). The most likely explanation was hypothesized to be a contamination with wild-type *P. purpureum* cells. This would explain the variation between lines of the same construct, if the wild-type to transgenic cell ratio differed accordingly (Figure 11A, C). The un-transformed cells could come directly from inoculation with primary transgenic algal colonies, or some cells lost their plasmid in the course of sub-cultivation and became wild-type cells.

It was shown that the transformation plasmid can be slowly diluted out, if *P. purpureum* transformants are not maintained under selection pressure (Li and Bock, 2018). The selective agent was always present in the culture medium of transformants at a concentration of 25 mg/L, but Zeocin™ is a salt-sensitive compound that is inhibited by high ionic strengths as present in lsASW medium (Zeocin™ selection reagent user guide MAN0000019, Invitrogen). In fact, the manufacturer recommends decreasing salt concentration to under 5 g/L to ensure optimal selection with this antibiotic. It was shown that *P. purpureum* cells can survive Zeocin™ concentrations of up to 10 mg/L (Li and Bock, 2018), and it is likely that the effective concentration of the selective agent fell below the selection threshold of 10 mg/L.

In light of this, it is not surprising that some wild-type cells survived and grew within a population of transformed cells, especially in cultures maintained over a longer period. This is especially the case for colonies derived directly from the primary transformation plate that has to be incubated for more than three weeks before the first colonies appear.

To be able to investigate the precise expression values of a transgene, homogeneous cultures of transgenic algae need to be obtained. FACS was performed to try to isolate only YFP-fluorescent cells but resulted in only a small increase in the percentage of fluorescent cells (Figure 11). This is most likely due to cell clumping caused by the cell wall polysaccharides. A subsequent isolation of single colonies on plates had to be performed and is described in 4.2.5.

P. purpureum cells readily grow on solid media when plated at a high cell density, but do not grow when the cell density during plating is too low. Plating at low cell density, however, is required for the isolation of single colonies.

To allow single colony formation in *P. purpureum*, a special plating protocol involving the embedding of cells in a starch matrix was applied and had a success rate of around 86% colony formation. This protocol was successfully applied by Skeffington et al. (2019) for plating of *E. huxleyi,* a coccolithophore that also did not show growth on solid medium in prior experiments. The authors argued that the starch matrix, in which the cells were embedded, creates a microenvironment that preserves moisture around the cells and supports cell growth and colony formation. This finding is also supported by the fact that plating of higher cell densities supports algal growth on plates, whereas low densities do not. The requirement for an extracellular matrix around the cells is additionally supported by the fact that transformants appear as defined colonies on Zeocin™-containing selection plates. The selective agent leads to death of all untransformed cells on the selection plate, which may effectively build a matrix consisting of polysaccharides and cell debris. The debris may support formation of colonies. This might also be a reason why additional starch plating on selective plates after transformation did not lead to any increase of transformation efficacy (Figure 6B).

The newly established method was used for the purification of transgenic cell lines to homogeneity by passing them through several rounds of single colony formation on plates. This was performed with the transgenic lines expressing the *YFP* codon variants (except for *PpYFP* that was already homogeneous), the *PYP1* and *Fo/Ni* promoter lines and the secHCV antigen expressing lines that showed strong variation in transgene expression. The lines expressing YFP were screened by microscopy for their ratio of YFP fluorescencing cells to non-fluorescing cells (Figure 12), whereas the secHCV antigen expressing lines had to be screened by western blotting (Figure 19).

A good example of wild-type contamination is seen in Figure 11 depicting different lines expressing the *Fo/Ni* promoter-driven YFP. The YFP accumulation detected by western blot analysis correlates with the number of non-fluorescent cells detected by microscopy (compare Figure 11A and 10B). After purification of Fo/Ni line 2 through single colony generation and sub-culturing, all sub-lines showed uniform YFP accumulation levels that were equivalent to the adjusted protein levels of the progenitor lines to 100% fluorescent cells.

Indeed, almost all lines expressing YFP could be purified to homogeneity after two rounds of sub-cultivation, except for line PYP1_2 (Figure 12). The percentage of fluorescent cells in a line increased with each sub-cultivation round derived from single colonies. When progenitor cell lines that showed a uniform fluorescent pattern (e.g., high percentage of fluorescence)

were sub-cultured again, the sub-lines showed uniform fluorescence as well, indicating the elimination of wild-type cells from those lines.

For the secHCV-expressing lines, a microscopic evaluation could not be performed and the effectiveness of the single colony-derived sub-cultivation had to be assessed by immunoblotting (Figure 19). To eliminate the variation between secHCV line 1 and 2, the lower expressing line 2 was purified by two rounds of single cell selection. Antigen accumulation was indeed increased to levels even surpassing that in secHCV line 1, indicating that line 1 was also contaminated with wild-type cells. The sublines of secHCV 2A in round two of the purification showed uniform transgene expression, similar to the sublines of the YFP expressing lines after purification. Additionally, in this experiment, a line showing intermediate expression was chosen for sub-cultivation after the first round (secHCV 2B) and showed variation of expression, indicating residual wild-type contamination. This line probably could have been purified to homogeneity by a third round of single colony generation and sub-cultivation.

In conclusion, the recommendation is to pass all primary transformants through at least two rounds of single cell isolation and sub-cultivation to ensure homogeneous transgenic cultures that (i) can be compared between each other, and (ii) accumulate the maximal possible amount of the recombinant protein.
Another recommendation is to increase the initial Zeocin™ concentration in the transformation plates. Preliminary results performed in this work showed that transgenic *P. purpureum* cells can survive Zeocin™ concentrations of 100 µg/mL. Such a high concentration would most likely ensure that the effective antibiotic concentration would not drop under the threshold allowing survival of the cells, even in high-salt medium. This would reduce, or even eliminate, wild-type contamination in primary transformants.

5.3 Nuclear transgene expression in *P. purpureum* can be increased by codon optimization and with newly identified promoters

5.3.1 Codon optimization is a main determinant for efficient transgene expression

After obtaining stable and uniformly expressing lines for all *YFP* codon variants, a quantitative analysis of these lines was performed. The comparative analysis of five different *YFP* gene variants with varying RCA to the nuclear genome revealed that, similar to *C. reinhardtii*, codon usage is an important parameter for effective transgene expression. AT-rich genes, as represented by the *cpYFP* variant, are difficult to express and showed no detectable accumulation, neither at the protein nor at the mRNA level, whereas all other variants showed a strong expression of the YFP protein. A positive correlation between mRNA levels and protein accumulation could be detected for all *YFP* variants (Figures 8 and 9). The fully codon-

optimized *PpYFP* gene variant accumulated YFP to the highest level and showed the highest mRNA accumulation of all tested lines. All other lines with detectable YFP (CrYFP, vYFP and laYFP) showed similar protein accumulation and exhibited a truncated mRNA entity that was approximately 200 bp shorter, indicating mRNA instability of non-codon optimized transcripts. These findings suggest that high rates of translation lead to increased mRNA stability. This is in agreement with results obtained in yeast and *C. reinhardtii* where a strong positive influence of codon usage on mRNA stability was seen (Barahimipour et al., 2015; Presnyak et al., 2015). This is most likely caused by translating ribosomes shielding the mRNA from enzymatic degradation, and thus prolonging half-life.

The difference in expression strength between the best variant (PpYFP, 100% RCA) and the second-best variant (CrYFP, 94.3% RCA) in *P. purpureum* was comparable with the difference in expression strength in *C. reinhardtii* between the respective best variant (CrYFP, 100% RCA) and second-best variant (vYFP, 93.9% RCA) (Barahimipour et al., 2015). The rather low difference in expression strength between the lines vYFP, laYFP and CrYFP in *P. purpureum* were most likely due to the minor differences in RCA of only 10% between those variants. The differences between the RCA of the different *YFP* variants in C*. reinhardtii* were substantially larger (around 30%), which could be an explanation why differences in YFP accumulation were seen in the green alga.

The employment of a probe hybridizing to the 5'UTR of the actin mRNA enabled the simultaneous detection of the transgenic mRNA and the endogenous actin mRNA that can serve as an internal standard. The mRNA levels of actin were comparable with the transgenic mRNA levels in the lines expressing the fully codon-optimized *PpYFP* gene variant. This indicates that the promoter is the main determinant of mRNA accumulation, since both genes were driven by the *Actin* promoter.

Regarding the quantification of YFP in *P. purpureum* in comparison to *C. reinhardtii*, the red alga clearly showed superior expression capacity. PpYFP accumulated to up to 5% of TSP, which was easily visible in a Coomassie-stained gel, and even the not perfectly codon-optimized CrYFP accumulated to 1.2% of TSP. This is a six times higher accumulation than in *C. reinhardtii*, where the perfectly codon optimized CrYFP accumulated to only 0.8% of TSP. This expression level is impressive considering that the strong *Chlamydomonas* expressor strain UVM11 was used to express the codon-optimized YFP driven by the strong endogenous *PSAD* promoter. This combination ensured high expression in this green alga. The *C. reinhardtii* strains generated by Barahimipour et al. (2015) were used that had been quantified to express YFP to up to 1% TSP. Even after seven years of maintenance, these strains still express YFP to comparable levels, indicating that (i) the UVM11 strain stably

expresses transgenes over years without silencing, and (ii) the quantification was reliably performed in the current work.

The accumulation of a recombinant protein to up to 5% of the total soluble protein in an algal system using expression from the nuclear genome is one of the highest expression values achieved in a microalgal system.

5.3.2 The use of novel endogenous promoters results in modest improvement of transgene expression

Promoters are important determinants of transgene expression and so far, only the *Actin* and *Tubulin* promoters were used to drive gene expression in *P. purpureum*. To further enhance recombinant protein accumulation, a set of new promoters was tested. Those promoters were derived from genes that are highly transcribed in standard growth conditions in this red microalga. A published mRNA dataset was re-analyzed and the top 100 highly transcribed genes were identified, from which the top five genes were picked for promoter analysis (Table 3). The promoters (including the 5'UTRs of these genes) were cloned to drive *PpYFP* expression. The analysis revealed relatively minor changes in YFP accumulation compared to the *Actin* promoter-driven *PpYFP* (Figure 10). An increase of around 30% was seen for the *ChlBP* and *PsbQ* promoter-driven YFP, whereas all other promoters conferred no increase in YFP accumulation.

The calculated mRNA abundance (based on transcripts per million reads) of the actin mRNA is up to 12 times lower than that of the genes employed in this analysis. Assuming that, in *P. purpureum* the promoter is the main determinant of the accumulation level of transcripts encoded by fully codon-optimized genes, this implies a high discrepancy between mRNA and protein abundances, indicating strong post-transcriptional regulation of gene expression, possibly at the level of mRNA stability and / or translation. To follow up on this, a northern blot analysis could be performed with probes hybridizing to (i) the respective 5'UTR and (ii) the PpYFP sequence. The 5'UTR probe would answer the questions if both the transgenic and the endogenous mRNA accumulate to similar levels, and if both mRNAs exhibit similar stability. The probe hybridizing to the coding sequence can be used to compare all transgenic lines with each other, and determine, if the difference observed at the protein level is present at the mRNA level as well.

Another possibility for the relatively weak influence of the promoter could be the fact that the transgene is transcribed from a plasmid. Nothing is known about the maintenance of the plasmid in *P. purpureum* and its possible association with histones and other DNA-binding proteins. The packing of DNA in nucleosomes and the accessibility of it to transcription factors are main determinants of transcription efficacy in *C. reinhardtii*. Chromatin remodeling factors

make chromatin more accessible to factors initiating transcription, consequently leading to increased transcript levels (reviewed in Schroda, 2019). Promoters are the main targets of the regulation of transcription initiation via binding of transcription factors to *cis*-regulatory elements, DNA-modifying enzymes as well as nucleosome organization which itself is influenced by chromatin modifying enzymes and chromatin-remodeling factors. Most of the actively transcribed genes have a nucleosome-depleted promoter region at the transcriptional start site (TSS). Nucleosomes can occlude *cis*-regulatory sequences that are important for transcription factor binding and transcription initiation at the transcriptional start site. Access to promoters is established, for example, by recruiting chromatin-remodeling factors of the SWI/SNF-type that actively eject nucleosomes, thus making the *cis*-regulatory elements or the TSS accessible (Cairns, 2009; Müller and Tora, 2014). However, if plasmid DNA is only loosely associated with nucleosomes in *P. purpureum*, promoter strength may not play an important role, because a rather loose chromatin would allow the transcription machinery to have similar access to all promoter copies. Whether or not this is the case remains to be investigated. Isolation of the plasmids together with the associated proteins would shed light on the organization and the episomal maintenance of the introduced vectors.

Finally, promoters are regulatory elements that are differentially active under different growth conditions. In contrast to housekeeping genes, which are constitutively active, many other genes need to be tightly regulated depending on the need for their gene product. The promoters tested here drive genes that should be tightly regulated, especially carbonic anhydrase (CA), the PsbQ domain-containing gene product and the chlorophyll-binding protein. CA is a secreted protein usually up-regulated under limited CO_2 availability, helping the cell to increase intracellular CO_2 concentrations (Moroney et al., 2011). PsbQ and ChlBP are proteins involved in the structure and function of photosystems in the chloroplast. The hypothetical PsbQ domain-containing protein has no sequence similarities with PsbQ proteins from other organisms, but a recent cryo-EM study found this protein associated with photosystem II in *P. purpureum* (You et al., 2023). Photosynthetic proteins are usually upregulated with increasing light intensities (reviewed in Tyagi and Gaur, 2003). The YFP accumulation conferred by the new promoter constructs was tested under standard conditions (100 µmol photons $m^{-2}s^{-1}$ light and cotton plugs), and higher light intensities might lead to upregulation of the *PsbQ* and *ChlBP* promoter and lead to higher YFP accumulation. Preliminary experiments with the *PsbQ* promoter driving cytHCV expression showed that HCV antigen accumulation is lower under low light (30 µmol photons $m^{-2}s^{-1}$) compared to normal light (100 µmol photons $m^{-2}s^{-1}$) (data not shown). If the right growth conditions can be found, transgene accumulation could be boosted with the new promoters to a greater extent than it was demonstrated in this thesis.

Although the promoters tested here do not seem to have a strong influence on YFP accumulation, it may still be worth testing other transgenes with different promoters. It was shown in *C. reinhardtii* that the strength of a promoter in a transgenic context can differ from the endogenous gene and may vary strongly depending on the transgene to be expressed. The *IFT25* promoter did not confer detectable expression of GFP, but showed very high expression values for its own gene and transcriptional GFP fusions (Dong et al., 2017). The performance of a promoter can depend on its transgenic context, and can exhibit unexpected synergies that boost expression even further. As demonstrated for the secHCV antigen, compared to the *Actin* promoter, the newly employed *ChlBP* promoter boosted expression and secretion of the antigen to levels higher than those seen for YFP (Figure 26).

5.4 Efficient protein secretion renders *P. purpureum* a promising and cost-effective production organism

The downstream processing costs of biopharmaceuticals are approximately 80% of the total product costs (Rosales-Mendoza et al., 2020). The possibility to purify a protein from the culture medium, omitting cell lysis and extraction from highly complex lysates, is a prerequisite for cost-effective recombinant protein production. Employment of the signal peptide (SP) of the endogenous carbonic anhydrase demonstrated the successful translocation of YFP into the ER and subsequent secretion into the medium (Figure 13). Furthermore, efficient ER-retention was accomplished by inclusion of a C-terminal HDEL signal. Besides the relatively stable YFP reporter protein, also the complex, glycosylated HCV antigen was successfully secreted into the medium. While secHCV-Ag secretion could not be readily detected in the *Actin* promoter strains, clear secretion of the 3xHA-tagged secHCV antigen was detected in the *ChlBP* promoter stains (Figure 26).

The occurrence of a double band in the secYFP lines suggested incomplete cleavage of the SP from the reporter protein. Usually, the SP is directly cleaved by signal peptidase I upon co-translational translocation into the ER (Auclair et al., 2012). This cleavage is dependent on precise recognition of the cleavage site located in the C-terminal region of the SP. The recognition might be less efficient, when the SP is fused in a non-native context to the transgenic protein, leading to miscleavage. This was observed with the SP of the BIP1 protein in *C. reinhardtii* when fused to a GFP reporter construct to facilitate ER targeting, resulting in occurrence of a double band as well (Niemeyer et al., 2021). Maintenance of the SP's native sequence context might alleviate this problem, and further increasing translocation and subsequent secretion efficacy.

Sulphated cell wall-associated polysaccharides are released by *P. purpureum* into the medium in high abundance. These exopolysaccharides lead to high viscosity of the medium, which can impair efficient protein purification, e.g., by clogging of the purification column. Fortunately, these exopolysaccharides are negatively charged due to the sulphate group, and can easily be eliminated by adding a positively charged detergent like CTAB. This detergent forms water insoluble complexes with long negatively charged polymers, which precipitate from the medium, while proteins are kept in solution (Beri and Gandhi, 2021). The use of this surfactant represents a safe and cost-effective method for clearing the medium without affecting protein quality.

YFP accumulated to a yield of 1.6 mg/L in the medium after a cultivation time of 10 days (Figure 14). The secretome of *P. purpureum* seems to have a low complexity, as judged by the absence of abundant proteins on the Ponceau-stained membrane. This is another advantage of this microalga, as it further simplifies recombinant protein purification. Unfortunately, *P. purpureum* seems to co-secrete proteases leading to degradation of the recombinant protein, which was observed for YFP and for the secHCV antigen. Distinct degradation products of YFP were observed 3 days after cultivation, and increased proportionally with the secreted recombinant protein. The secHCV antigen also showed a degradation pattern, which could be largely suppressed when protease inhibitor was added to the cultivation medium. The effect of the protease inhibitor seemed to be stable for at least 20 h, as seen in Figure 26, when the culture medium was harvested 20 h after addition of the protease inhibitor and no degradation was observed.

The oserved degradation, together with the generally lower secretion of the *Actin* promoter-driven secHCV antigen was the reason for the initial assumption that the secHCV antigen was not secreted into the medium. After establishing the CTAB precipitation protocol, high-quality medium samples could be obtained, which enabled us to detect the secreted antigen by western blot analysis in the *Actin* promoter-driven strains. It is notable that the intracellular accumulation of the HCV antigen expressed from the *secHCV* gene is comparable between the different promoter strains, whereas the extracellular accumulation of the antigen in the *ChlBP*-secHCV strains is much higher than in the *Actin*-secHCV strains. This might suggest fast traffic of the antigen through the ER, thus only allowing a “snapshot” quantification of the secretory pathway-resident antigen, which may be similar in both strains. The precise accumulation of the HCV antigen in the medium remains to be determined.

A constant addition of protease inhibitor is costly and would decrease the value of *P. purpureum* as a production organism for molecular farming. One possibility to alleviate this problem is to survey the genome of *P. purpureum* for secreted proteins with protease activity, possibly by mass spectrometric analysis of the secretome. Because of the low secretome

complexity, chances are high to identify candidate proteins, which then could be knocked out using CRISPR/Cas9 that is currently under development for *P. purpureum* in our group. This could dramatically increase the stability of the secreted recombinant protein and probably increase the yield while lowering production costs.

The microscopic analysis of the YFP secretion lines revealed recombinant protein accumulation in the ER or Golgi apparatus of the cells in the stationary phase, indicating decreasing secretion efficacy (Figure 15A). Whether this is due to ER overloading with misfolded protein or the increased formation of the thick, viscous cell wall prevalent in the stationary phase remains to be investigated.

Microalgal secretion of recombinant proteins is still an under-studied field with high potential. The highest secretion yield obtained so far in the microalga *C. reinhardtii* was 15 mg/L reached by employing a fused synthetic glycomodule to the venus YFP reporter protein. Secreted venus YFP without the glycomodule, which is comparable to the experiment in *P. purpureum*, accumulated to 1.3 mg/L (Ramos-Martinez et al., 2017). The disadvantage of this approach lies in the fusion to a highly glycosylated module that can cause unwanted immunogenic reactions when applied to antigens and other biopharmaceuticals. The module could be enzymatically cleaved off, but this would substantially increase the downstream processing costs. This SP20 glycomodule was first developed to facilitate efficient secretion in plant cell culture, before it was successfully applied to *C. reinhardtii* (Shpak et al., 1999). It would be worth testing this glycomodule in combination with secreted proteins expressed in *P. purpureum* to check whether this can improve secretion.

In this study, only the signal peptide of the carbonic anhydrase was evaluated, suggesting that there could be significant potential for further enhancing secretion in *Porphyridium*. The signal peptide is one on the main determinants of efficient secretion, and a difference in secretion efficiency of up to 10-fold was demonstrated when employing different SPs in *C. reinhardtii* (Molino et al., 2018). Additionally, a positive synergistic effect was recently demonstrated between the *PSAD* promoter (and 5'UTR) and its corresponding chloroplast targeting signal, probably due to motifs that enable more efficient transcription or translation (Einhaus et al., 2021). A similar synergy could occur when combining a promoter with its genes' native signal peptide, which could further boost secretion efficiency. The CA promoter may be highly regulated, but another promoter driving a highly expressed gene that encodes a secreted protein could also be chosen. In the list of the top 100 highly transcribed genes, there are six genes whose products are predicted to be targeted to the secretory pathway. It would be worthwhile to analyze the promoters of these genes together with their respective signal peptides and compare the secretion efficiency of a reporter protein.

In summary, *P. purpureum* proves to be a suitable producer of recombinant proteins. Compared to the commonly used alga *Chlamydomonas*, it exhibits a significantly higher protein expression capacity of up to 5% TSP, which most probably can be further enhanced using endogenous promoters. Additionally, this red alga efficiently secretes proteins and is characterized by a remarkably protein-scarce secretome, thus simplifying the purification of the recombinant protein. This study evaluated the biotechnological potential of *P. purpureum*, and established a foundation for future research in this interesting organism. The results of this section were published (Hammel et al., under review).

To demonstrate the true potential of *P. purpureum*, biotechnologically valuable and complex proteins were expressed in this red alga, including variants of the S-protein of HBV and the E2 surface protein of HCV.

5.5 The expression of viral proteins demonstrates the potential of *P. purpureum* as a production platform for complex biopharmaceuticals

Having achieved expression and secretion of YFP at high levels, complex viral proteins derived from the hepatitis B and C viruses, suitable as subunit vaccines, were produced in *P. purpureum*. Both proteins are complex in that they require disulphide bond formation to reach their final conformation, have one and eleven glycosylation sites, respectively, and HBV S-protein contains four TMDs. In order to produce conformationally correctly folded antigens that can act as vaccines, the proteins had to be targeted to the secretory pathway to undergo the posttranslational modifications. To test for this, three versions were constructed for each protein. The version without an endogenous signal peptide was designed for localization in the cytosol. Another version with the carbonic anhydrase signal peptide was targeted to the secretory pathway to test for possible secretion, and finally, a version designed as ER-residual protein included an HDEL-motif. In parallel, the same versions were constructed for expression in *C. reinhardtii* to be able to compare expression values between both organisms. In the following, the expression and accumulation of all different HBV and HCV antigen variants will be discussed. Afterwards, the biochemical characterization involving findings from our collaboration partners is going to be discussed. Lastly, the mice immunization trials performed by our collaboration partners will be mentioned to give a full picture of the potential of *P. purpureum* as a producer organism for recombinant proteins.

5.5.1 High-yield recombinant protein production in *P. purpureum*

All of the differently targeted versions of the HBV and HCV antigens were successfully expressed in the red algal system, whereas only the HCV antigen versions were expressed to detectable levels in *C. reinhardtii* (Figure 17 and Suppl. Figure 2). As demonstrated with YFP, the expression capacity of the red alga was much higher, and it accumulated on average up

to four times more HCV antigen than the green alga. *P. purpureum* primary transformants accumulated the secHCV antigen to up to 0.5% of the total protein and the secHBV-3xHA antigen to 0.15% of TP. Additionally, a 6xHis-tagged secHCV version was expressed in *P. purpureum* to facilitate affinity purification of this antigen. For the HBV antigen, an untagged version was expressed to remedy problems with sub-viral particle (SVP) formation of the 3xHA-tagged antigen. Furthermore, the endogenous SP was omitted, because the first TMD of the S-protein is able to facilitate translocation into the ER of *P. purpureum* (Figure 21). After cultivation of the respective strains in a PBR, the final yields of the antigens were determined and those antigens were later used for immunization trials. Table 6 summarizes the expression yields of all recombinant antigens determined in this work and compares them with other biopharmaceuticals expressed in various eukaryotic expression systems.

Table 7: Overview of the biopharmaceuticals produced in various eukaryotic expression systems.

Species	Biopharmaceutical	Yield	Reference
P. purpureum	HCV $E2^{\Delta HVR1}$ (3xHA-tagged)	0.5% TP, primary transformants 0.22% TP, PBR-grown 0.89 mg/g DW, PBR-grown 0.97 mg/L PBR culture	This work
P. purpureum	HCV $E2^{\Delta HVR1}$ (6xHis-tagged)	0.25% TP, PBR grown 0.5 mg/g DW	This work
P. purpureum	HBV S/preS1 (3xHA-tagged)	0.15% TP, primary transformants	This work
P. purpureum	HBV S/preS1 (untagged)	3.8% TP, PBR grown 8.45 mg/g DW 3.38 mg/L PBR culture	This work
Microalgae			
C. reinhardtii	Fusion protein of luciferase and human butyryl cholinesterase	0.4% TSP, nuclear	(Kumar et al., 2013)
C. reinhardtii	HIV, p24	0.25% TSP, nuclear	(Barahimipour, et al., 2016)
C. reinhardtii	Chimeric HBV nucleocapsid + angiotensin II	0.05% TSP, nuclear	(Soria-Guerra et al., 2014b)
C. reinhardtii	Sars-CoV-2 spike receptor-binding domain	31 µg/g FW, nuclear	(Berndt et al., 2021)
C. reinhardtii	SARS-CoV-2 spike protein	11 µg/L, nuclear, secreted	(Kiefer et al., 2022)
C. reinhardtii	mammary-associated serum amyloid (M-SAA)	5% TSP, chloroplast	(Manuell et al., 2007)

Species	Biopharmaceutical	Yield	Reference
C. reinhardtii	*Plasmodium falciparum* surface protein 25 and 28	0.5% TSP and 0.2% TSP, respectively; chloroplast	(Gregory et al., 2012)
C. reinhardtii	Human diabetes associated antigen hGAD65	0.3% TSP, chloroplast	(Wang et al., 2008)
Dunaliella salina	HBV S-protein	0.00031% of TSP, nuclear 3.11 µg/g TSP	(Geng et al., 2003)
Phaeodactylum tricornutum	HBV S-protein	0.7% TSP, nuclear, ER -targeted	(Hempel et al., 2011)
Phaeodactylum tricornutum	Human IgG against HBV S-protein	8.6% TSP, nuclear, ER-targeted 21 mg/g DW	(Hempel et al., 2011)
***Schizochytrium* sp. strain 20888**	GP1 from *Zaire ebolavirus*	1.25 mg/g FW, transient	(Bañuelos-Hernández et al., 2017)
Plants			
Lactuca sativa	HBV S/preS1 HBV S	155 ng/g FW, transient 233 ng/g FW, transient	(Dobrica et al., 2018)
Nicotiana benthamiana	HBV S/preS1	200 µg/g FW, transient	(Pantazica et al., 2023)
Nicotiana benthamiana	HCV $E2^{\Delta HVR1}$	20 µg/g FW, transient	(Dobrica et al., 2021)
Nicotiana tobacum	HIV, p24	3.5 mg/g TSP 0.35% TSP	(Zhang et al., 2002)
Miscellaneous			
Chinese hamster ovary cells	HBV S-protein	983 mg/L, intracellular 3.4 mg/L, secreted	(Li et al., 2006)
Pichia pastoris	HBV S-protein	6.6 mg/g FW, intracellular	(Cregg et al., 1987)
S. cerevisiae	HBV S-protein	18.31 mg/L culture 9.7 mg/g TP 0.97% TP	(Hadiji-Abbes et al., 2009)

TP: total protein; TSP: total soluble protein; DW: dry weight; FW: fresh weight; nuclear: expressed from the nucleus; chloroplast: expressed from the chloroplast; transient: viral or Agrobacterium-mediated; PBR: photobioreactor

This table can only give an overview of some biopharmaceuticals produced mainly in microalgal systems, and, for comparison, other eukaryotic systems that produced the same antigens as in this work. For a more complete list of biopharmaceuticals produced in microalgae, please refer to Rosales-Mendoza et al., (2020) and Arias et al., (2023).

The majority of studies on the microalgal production of biopharmaceuticals were performed in *C. reinhardtii,* because it is the microalga with the most advanced repertoire of molecular tools. A striking difference in accumulation of recombinant proteins is seen between nuclear and chloroplast-expressed transgenes. While biopharmaceuticals expressed in the chloroplast can reach values of up to 5% TSP, nuclear expressed transgenes are usually tenfold less abundant in Chlamydomonas (reviewed in Schroda, 2019). To compare the values of *P. purpureum* given in Table 6, it has to be considered that the quantification was performed based on total cellular protein, rather than total soluble protein. As demonstrated based on YFP accumulation of the PpYFP strain when loading TSP (Figure 8C) or TP (Figure 14), the soluble protein accounts for approximately half of the total protein in a *P. purpureum* cell. Considering this, quantification values must be doubled for the soluble HCV antigen to convert them to percentage of TSP. This would place *P. purpureum* in the top rank of microalgal producer organism, together with *P. tricornutum*. Regarding the absolute yield, the HCV antigen accumulated to approximately 0.89 mg/g of DW in PBR-grown cultures, which represents approximately 0.97 mg/L culture volume. 1400 mg DW was obtained from 2 times 650 mL PBR-grown culture, resulting in a DW yield of 1.1 g/L, which is comparable to the total biomass yield obtained with the wild-type strain (Table 1). The E2 protein was not expressed in any other microalgal system to date, but it was transiently expressed in *N. benthamiana* to levels of 20 µg/g fresh weight, approximately 40 times lower than in *P. purpureum*.

Several antigens derived from HBV have been produced in microalgae and plants. The expression values achieved ranged from 0.00031% of TSP in *D. salina* to 0.7% of TSP in *P. tricornutum*., HBV S/preS1 was also expressed in *N. benthamiana* and *L. sativa* to levels of 200 µg and 155 ng per gram fresh weight, respectively. The untagged HBV S/preS1 antigen version was expressed in *P. purpureum* to levels of 3.8% of TP or 8.45 mg/g of DW, several orders of magnitude higher than in plants. For the HBV S/preS1 strain, 0.4 g/L DW was obtained in PBR-grown cultures. This is less than for the secHCV strain, because the culture was only grown to the mid exponential phase with an average of 4.5×10^6 cells/mL. Related to the total yield per liter PBR-grown culture, this represents 3.38 mg/L. This high value has to be taken with caution, as the quantification was indirect and it needs to be confirmed with an HBV S-protein standard.

Nevertheless, higher expression values of the untagged antigen compared to the tagged antigen are conceivable, because the untagged version is able to form SVPs, whereas the 3xHA containing antigen only forms dimers (see 1.4.2). HBV-S protein dimers accumulate inside the ER and evoke ER stress that upregulates components of the ER-associated protein degradation pathway, leading to decreased yields of recombinant protein. The formation of SVPs is correlated with the release of these particles from the ER and trafficking along the

secretory pathway, alleviating the ER-stress and possibly leading to higher accumulation (reviewed in Lazar et al., 2014). The same argument can be made for the ER-targeted HCV antigen that showed decreased accumulation in the stationary phase (Figure 18), indicating increased ER stress in the stationary phase. Our collaboration partner could show that expression of the HCV antigen in *P. purpureum* led to upregulation of BiP, a key regulator of the unfolded protein response (Hammel et al., manuscript in preparation).

Interestingly, the yield of HBV S/preS1 in *P. purpureum* seems to be comparable to *P. pastoris*, a yeast that is used for industrial production of the HBV S-protein for vaccination. Compared to *S. cerevisiae*, another industrial producer of the HBV S-protein, *P. purpureum* has a higher accumulation in regard to TP, but a lower accumulation of antigen per liter culture. This is due to a higher cell density of yeast cultures and a higher total protein content per gram dry weight, which can be up to 60% (Yamada and Sgarbieri, 2005).

It needs to be evaluated in more detail, if *P. purpureum* is able to be a competitive producer of the HBV S antigen compared to other producers. The comparably low cell density requires high-volume cultivation of microalgae to produce a competitive amount of recombinant protein. Recently, a new formulation for a high-density culture medium in *C. reinhardtii* was described, allowing biomass accumulation of up to 20 g/L (Freudenberg et al., 2021). Finding a similar media formulation for *P. purpureum*, that might additionally also increase total protein content (see 5.1), would greatly increase the value of this alga. Techno-economic analysis of a large-scale microalgal cultivation showed that 1 kg of DW costs approximately 53 € for stacked tubular PBR cultivation in 15,000 L volume (Vázquez-Romero et al., 2022). This evaluation is only valid for the growth of microalgae and does not include upstream and downstream processing costs of recombinant protein production. A techno-economical evaluation of microalgae producing high value recombinant proteins is still pending.

A notable finding was the apparent decrease of antigen accumulation over time in the HCV antigen-expressing strains. When comparing the HCV accumulation levels quantified in the primary transformants with those of the transformants that were maintained for 1.5 years, the accumulation decreased by approximately 2 fold (Figure 17C and 24B). This might be due to improper maintenance conditions. As discussed in 5.2.2, contamination with wild-type cells regularly occurs in primary transformants. In addition, according to the data obtained with the single colony-derived secHCV strains (Figure 19), the primary transformants were inhomogeneous. When maintaining a culture over a long period with the salt-sensitive antibiotic Zeocin™ in high-salt media, the occurrence of wild-type cells and their increase in number within the population seems likely. This is supported by the fact that, when inoculating a cryo-conserved stock of primary transformants after years, they show the exact same

antigen accumulation as shortly after transformation. This observation supports the necessity to generate single colony-derived, homogeneous transformants as early as possible after transformation.

Trafficking of the antigens beyond the ER enables N-glycan processing by glycosidases and glycosyltransferases along the secretory pathway, with potential impact on immunogenic properties. Therefore, the secreted version of both antigens was selected for growth in a bioreactor and further characterization.

5.5.2 *P. purpureum* provides a proper environment for correct folding and glycosylation of the HBV and HCV antigens

After successful expression of all antigens in the red microalga, the recombinant proteins were biochemically characterized. The N-glycosylation inhibitor tunicamycin (TM) was employed (Figure 20 and 21) to confirm the assumed glycosylation seen by a shift of the molecular weight between the cytosolic and the secretory pathway-targeted antigens (Figure 17B and C). TM interferes with the first step of glycosylation by inhibiting the formation of the N-glycosylamine (GlcNAc) dolichol phosphate anchor. Indeed, the addition of TM abolished the molecular weight shift in the secHCV and HCV-ER antigens. The time course experiment performed demonstrated a rapid accumulation of non-glycosylated antigen already after 15 min of TM addition for the secHCV antigen construct that was coupled with a decline in the glycosylated protein after 2 h. This indicates a fast passage of the antigen through the ER. In contrast, the ER-targeted version accumulated non-glycosylated antigen only 1 h after addition of TM and the glycosylated form was more stable throughout the experiment, indicating ER residency.

Another striking difference between the two HCV variants is the occurrence of several bands at a size of around 55 kDa in the secHCV construct. These size differences could be caused by different glycans attached to the antigen. The basic N-glycosylation process in the ER ($Man_9GlcNAc_2$-glycan) is highly conserved among most eukaryotes and is completed at the transition from ER to Golgi, where maturation of the N-glycans takes place (Schoberer et al., 2018). This maturation is highly organism specific. A bioinformatics analysis of glycosylation patterns in different microalgae revealed that *P. purpureum* possibly obtains a $Man_{8-9}Xyl_{1-2}Me_3GlcNAc_2$ glycosylation (Liu et al., 2021). Slight differences in N-glycan composition can explain the small apparent size differences seen in the secHCV antigen. A preliminary mass spectrometry analysis of the purified 6xHis-tagged secHCV antigen showed the occurrence of a high mannose type glycosylation, fitting with the bioinformatic analysis (Hammel et al., manuscript in preparation).

The addition of TM to strains expressing the three variants of the HBV antigen led to decreased protein accumulation in all variants. Since TM is inducing ER-stress, which in turn leads to decreased protein accumulation in the ER and the secretory pathway, this indicates that all three variants were in fact targeted to the ER. The first TMD acted as a signal peptide targeting the “cytosolic” version into the ER. The molecular weight shift, visible between the cytHBV, the sec-HBV and the ER-HBV versions, was most likely due to a miscleavage of the endogenous carbonic anhydrase SP (that was seen in the secYFP as well, Figure 21A).

Time course experiments with the addition of TM to the cytHBV lines revealed that this antigen was in fact glycosylated, but only to minor amounts. This glycosylated form, detected as as band running approximately 2 kDa higher, had completely disappeared after 24 h TM treatment (Figure 21B). Here again, the accumulation of the antigen decreased to 15% after 24 h of TM-induced ER stress, supporting the hypothesis that the 3xHA-tagged HBV antigen is not stable upon ER stress. The diminished glycosylation of the HBV S/preS1 protein when only relying on the first TMD for translocation (cytHBV) is not problematic from an immunological point of view. S-protein for vaccination, currently produced in yeast, is not glycosylated at all, without any detrimental effect on immunization success (Dobrica et al., 2020).

The secretory pathway-targeted HCV antigen was subjected to detection with conformational antibodies by our collaboration partners, only detecting correctly folded antigen. Both, the 3xHA-tagged antigen and the 6xHis-tagged antigen were successfully detected with this antibody, confirming correct folding and correct formation of the nine disulphide bridges (Hammel et al., manuscript in preparation). This is important for immunization, because it was shown that different conformations of the E2 protein elicit immune reactions of different strength, with the correctly folded antigen leading to stronger reactions than a misfolded antigen (Pierce et al., 2020). Furthermore, the *Porphyridium*-produced E2 antigen was able to bind to the HCV receptor CD81 as tested by ELISA (Hammel et al., manuscript in preparation).

The HBV antigen-expressing strains were subjected to native lysis and density gradient centrifugation by our collaboration partners followed by ELISA detection of the fractions containing the HBV antigen. This analysis revealed that none of the 3xHA-tagged HBV antigen variants were able to form SVPs, and instead they all formed dimers. These dimers react to conformation-dependent antibodies, indicating correct folding and disulphide bridge formation. The inability to form SVPs is most likely due to the presence of the 3xHA tag, which hinders the formation of higher-order oligomers. As stated in the Introduction, SVP formation is strictly conformation dependent and many of the structural HBV S-protein features are involved. S-protein dimers are not immunogenic enough to evoke a proper immune reaction. For the

production of SVPs in *P. purpureum*, an untagged version of the S/preS1 protein was designed, omitting the endogenous signal peptide and the 3xHA tag, because it was shown that the first TMD acts as intrinsic translocation signal and the endogenous SP might be miscleaved. This miscleavage might also hinder SVP formation.

The untagged S/preS1 protein expressing strain was subjected to the above-mentioned density gradient analysis and SVP formation in this strain was confirmed.

In conclusion, it was demonstrated that *P. purpureum* was able to successfully express and process topologically complex antigens, leading to the formation of their native conformation. Glycosylation was confirmed for both antigens and it seems to be of a high mannose type. The pending MS analysis will shed further light on the N-glycans of this red alga. Compared to the yeast production system for HBV, *P. purpureum* seems to be superior in that it glycosylates and correctly folds the S/preS1 antigen. Whether SVP formation already takes place inside the algal cell or, as in yeast, during cell lysis with subsequent self-assembly, still needs to be determined. The much higher accumulation levels of the untagged HBV protein suggest trafficking beyond the ER-Golgi transition and intracellular SVP formation, but this has to be confirmed by electron microscopy.

After sufficient expression and characterization of the antigens, their use as vaccine was tested in mice immunization trials.

5.5.3 Recombinant antigens expressed in *P. purpureum* are able to elicit a strong immune response in mice

To complete the investigation of recombinant antigens produced in *P. purpureum*, the immunogenicity of the secHCV $E2^{\Delta HVR1}$ and the HBV S/preS1 SVPs was tested by our collaboration partners in Dr. Norica Nichitas lab from the Romanian Academy of Science. For this, two different approaches were pursued, first intramuscular administration of purified antigen and second, feeding of transgenic algal biomass.

In the following only, the results for the HCV antigen are discussed, since the immunization trials for HBV S/preS1 SVPs are still pending.

The 6xHis-tagged secHCV antigen was affinity purified and parenterally administered to a group of 7 mice three times in 14-day intervals at a concentration of 10 µg per dosage (Hammel at al., manuscript in preparation). The antigen was able to induce a strong HCV-specific immune response in the vaccinated mice. Mice sera were further tested for the presence of neutralizing antibodies against the homologous HCV isolate by using a luciferase-based HCV neutralization assay (Dobrica et al., 2021). Indeed, sera from mice receiving the *P. purpureum* antigen showed significant inhibition of HCV infection compared to the adjuvant

only group. Additionally, injection of the antigen did not show any signs of toxicity in the mice, indicating compatibility of the glycosylation pattern of *P. purpureum* with the mammalian immune system.

The next administration mode was oral administration to test whether sole feeding with transgenic algal can induce a humoral immune response. Algal cells are considered a suitable delivery system for subunit vaccines (Specht and Mayfield, 2014; Vela Ramirez et al., 2017). The safe enclosing of the antigen inside the alga ensures the safe transition of the protein into the intestine where it can induce a mucosal and cell-mediated immune response (reviewed in Pantazica *et al.*, 2021). *P. purpureum* biomass was shown to be non-toxic for rats, chickens and fish, and even promotes animal health, indicating a general non-toxicity of biomass of this red alga (Díaz-Rosales et al., 2008; Dvir et al., 2000; Šefcová et al., 2021).
A group of 7 mice was fed twice with 140 mg transgenic *P. purpureum*, equivalent to approximately 108 µg of antigen. Unfortunately, no detectable immune reaction against HCV was detected in the sera of the mice. Whether this is due to poor presentation of the antigen, or its degradation, or the absence of an adjuvant from the feeding trials is not known and needs further investigation.
At the same time, another experiment was performed where mice were primed with human cell line-derived antigen and then fed with an equivalent of 13 µg antigen twice. Compared to the control (wild-type *P. purpureum* biomass), the oral administration was able to boost the immune response against HCV (Hammel et al., manuscript in preparation). This is comparable to results obtained by feeding transgenic lettuce, where the immune response could be boosted after primal injection with the E1/E2 heterodimer (Clarke et al., 2017).

In conclusion, antigens produced in *P. purpureum* show no sign of toxicity when administered by feeding or injection, and can elicit a strong and specific immune response when injected. This is the first time that a red algal-produced recombinant protein induced an immune response and provides proof of concept for *P. purpureum* being suitable to produce safe and potent subunit vaccines.

5.6 General recommendations

Working with a non-model organism is always challenging, since the methodologies describes and the tools available are limited. In the course of this work, several new methods for working with *P. purpureum* were established and new molecular tools were created.

The most important recommendation for working with *P. purpureum* is the generation of homogeneous transgenic strains as early as possible, before wild-type contamination takes over. For this, the concentration of Zeocin™ on primary selection plates can be increased to

50 µg/mL to further suppress growth of wild-type cells. Furthermore, generation of single colony-derived lines has to take place as early as possible in order to obtain a stable homogeneous strain. To avoid the decrease of the effective antibiotic concentration, the sub-culturing of transgenic *Porphyridium* should be performed at least every two weeks and the maintenance of cells should take place at low light intensities to avoid cell death caused by a dense culture.

The red alga *P. purpureum* represents a more efficient expression host for recombinant protein production than the green alga *C. reinhardtii* for several reasons. First, this work shows that *P. purpureum* can achieve notably higher expression levels using the same transgenes. Intriguingly, even a gene variant optimized for *C. reinhardtii* yields higher expression in transgenic *P. purpureum* compared to transgenic *C. reinhardtii*. The higher expression levels may be attributable to the high copy number of the episomally maintained transgenes in *P. purpureum*. Furthermore, transgenes in *P. purpureum* do not integrate into nuclear chromosomes, making them impervious to position effects and reducing the risk of epigenetic transgene inactivation. Consistently, there is little variation in transgene expression among different transgenic clones, except the variation coming from wild-type contamination or the growth phase. The occurrence of episomally maintained transformation plasmids that can slowly segregate out if selection pressure is released can be an advantage for creating CRISPR/Cas-edited strains in the future, in that the transformation plasmid carrying the transgenes can be eliminated. This would result in a non-transgenic edited strain. The intron-poor genome of *P. purpureum* simplifies the design of transformation constructs, compared to *C. reinhardtii* that relies on introns to express long (trans)genes. Similar to the antigens produced in the green microalga, the antigens produced in *P. purpureum* are correctly folded and glycosylated, and are able to elicit a strong, specific and non-toxic immune response in mice.

Together, these findings provide new opportunities for algal biotechnology and synthetic biology, and suggest red microalgae as a highly competitive production system for molecular farming.

5.7 Outlook

The next step to finalize the ongoing experiments is to quantify the secreted HCV antigen driven by the *ChlBP* promoter using protease inhibitor-supplemented media. Additionally, the same construct including a 6xHis-tag instead of a 3xHA-tag should be expressed in *P. purpureum* to facilitate easy purification, and to establish the maximum obtainable yield in this microalga using a secreted antigen.

Another urgently missing experiment is the more accurate quantification of the untagged HBV antigen, to confirm the high accumulation value. This can be accomplished by using a commercially available HBV S-protein as a standard, omitting the indirect quantification through the 3xHA antibody and instead performing direct quantitation via immunoblotting or ELISA.

To understand, why stronger promoters do not confer a much stronger increase in the accumulation of recombinant proteins, a first step would be to separately analyze transcription and translation rates. For this, a northern blot analysis on the different promoter strains should be performed, and the transcript accumulation from transgenes in relation to the corresponding endogenous gene (from which the promoter was taken) should be compared

The expression of transgenes can potentially be enhanced in various ways. One possibility would be to increase the plasmid copy number. To accomplish this, one could evaluate plasmids with different origins of replication. Interestingly, upon isolation of genomic DNA of wild-type *P. purpureum* and separation by agarose gel electrophoresis, several bands ranging from 500 bp to 7,000 bp can be detected. Since both the mitochondrial and the plastid genomes are larger than 150 kbp (Kim et al., 2022; Tajima et al., 2014) it is unlikely that these bands represent DNA derived from those organellar genomes. Mitochondrial genomes can occur in subgenomic circles, fitting to the observed lengths, but they usually occur in much lower concentration than genomic DNA, and are not visible, for example, in DNA preparations of *Chlamydomonas.* Thus, the presence of autonomously replicating DNA in *P. purpureum* seems possible, and it would be worthwhile to isolate and sequence those DNA bands. Potentially, new regulatory elements can be found that further enhance plasmid copy number and thus expression.

Another important question is the mechanism underlying the maintenance of transformed plasmids in *P. purpureum*. To address this, the isolation of native plasmids, along with all associated proteins and other factors, should be conducted. This will help to identify candidate proteins that play a role in plasmid maintenance. Such an analysis can be executed using techniques like pull-down with DNA-binding proteins of a plasmid that contains the corresponding binding motifs in its DNA sequence.

In the near future, genome editing via CRISPR/Cas should be established in *P. purpureum*. Once this becomes feasible, a plethora of opportunities in basic research and biotechnology will emerge for *P. purpureum*. Given that *P. purpureum* directs a significant portion of its biomass into polysaccharide production, knocking out genes responsible for the corresponding biosynthetic pathway could redirect metabolism towards growth and protein production. This alteration could also address the viscosity issue that arises due to the release

of exopolysaccharides into the medium during the stationary phase, thus potentially simplifying the purification of secreted proteins. To identify candidate proteins that might play a role in polysaccharide biosynthesis, an RNAseq analysis during both the exponential and the stationary growth phases might be helpful. Since polysaccharide production is induced in the stationary growth phase, any carbohydrate-related gene that is up-regulated during this stage could be involved in polysaccharide production. Preliminary work was done by Ji et al., (2021). In this work, the authors conducted a transcriptome analysis of *P. purpureum* cultures grown in ASW medium or nitrogen-deprived ASW medium in the exponential growth phase and the stationary phase. These datasets can be used for candidate gene identification.

5.8 Summary

Microalgae have been recognized as new green biofactories for recombinant protein production, but their full potential has not nearly been realized due to the sparsity of genetic tools that currently are only available for a few selected species. This work has provided new molecular tools for the red microalga *Porphyridium purpureum*, a promising new candidate host for microalgal recombinant protein production. Major parameters for high-level protein expression, including GC content, codon usage and promoters were determined, and YFP accumulation of up to 5% of the total soluble protein was demonstrated.

P. purpureum's ability to efficiently secrete not only reporter proteins, but also viral antigens into the medium opens up new opportunities for this red alga as host organism for biotechnology. The complex viral glycoproteins here tested were successfully glycosylated and folded into their native conformation, demonstrating that *P. purpureum* is suitable to produce complex high-value biopharmaceutics to high levels. The antigens were used for immunization experiments in mice and elicited a strong and specific immune response without any indication of toxicity.

In sum, this work demonstrates the great potential of *P. purpureum* in biotechnology and has established this red microalga as a competitive production system for recombinant proteins.

6 References

Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D., Ech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B.A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., Blankenberg, D., 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. 46, W537–W544. https://doi.org/10.1093/nar/gky379

Alves, C., Silva, J., Pinteus, S., Gaspar, H., Alpoim, M.C., Botana, L.M., Pedrosa, R., 2018. From marine origin to therapeutics: The antitumor potential of marine algae-derived compounds. Front. Pharmacol. 9, 1–24. https://doi.org/10.3389/fphar.2018.00777

Amack, S.C., Antunes, M.S., 2020. CaMV35S promoter – A plant biology and biotechnology workhorse in the era of synthetic biology. Curr. Plant Biol. 24, 100179. https://doi.org/10.1016/J.CPB.2020.100179

Apt, K.E., Grossman, A.R., Kroth-Pancic, P.G., 1996. Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*. Mol. Gen. Genet. MGG 252, 572–579. https://doi.org/10.1007/BF02172403

Arad, S. (Malis), Friedman, O. (Dahan), Rotem, A., 1988. Effect of Nitrogen on Polysaccharide Production in a *Porphyridium* sp. Appl. Environ. Microbiol. 54, 2411–2414. https://doi.org/10.1128/aem.54.10.2411-2414.1988

Arad, S., Levy-Ontman, O., 2010. Red microalgal cell-wall polysaccharides: biotechnological aspects. Curr. Opin. Biotechnol. 21, 358–364. https://doi.org/10.1016/J.COPBIO.2010.02.008

Arad, S., Richmond, A., 2003. Industrial Production of Microalgal Cell-Mass and Secondary Products - Species of High Potential: *Porphyridium* Sp., in: Handbook of Microalgal Culture. Blackwell Publishing Ltd, Oxford, UK, pp. 289–297. https://doi.org/10.1002/9780470995280.ch15

Arias, C.A.D., Oliveira, C.F.M. de, Molino, J.V.D., Ferreira-Camargo, L.S., Matsudo, M.C., Carvalho, J.C.M. de, 2023. Production of Recombinant Biopharmaceuticals in *Chlamydomonas reinhardtii*. Int. J. Plant Biol. 14, 39–52. https://doi.org/10.3390/ijpb14010004

Arteni, A.A., Liu, L.N., Aartsma, T.J., Zhang, Y.Z., Zhou, B.C., Boekema, E.J., 2008. Structure and organization of phycobilisomes on membranes of the red alga *Porphyridium cruentum*. Photosynth. Res. 95, 169–174. https://doi.org/10.1007/S11120-007-9264-Z/FIGURES/3

Auclair, S.M., Bhanu, M.K., Kendall, D.A., 2012. Signal peptidase I: Cleaving the way to mature proteins. Protein Sci. 21, 13–25. https://doi.org/10.1002/pro.757

Bae, H., Coller, J., 2022. Codon Optimality-Mediated mRNA degradation (COMD): Linking Translational Elongation to mRNA Stability. Mol. Cell 82, 1467. https://doi.org/10.1016/J.MOLCEL.2022.03.032

Baier, T., Kros, D., Feiner, R.C., Lauersen, K.J., Müller, K.M., Kruse, O., 2018a. Engineered Fusion Proteins for Efficient Protein Secretion and Purification of a Human Growth Factor from the Green Microalga *Chlamydomonas reinhardtii*. ACS Synth. Biol. 7, 2547–2557. https://doi.org/10.1021/ACSSYNBIO.8B00226/ASSET/IMAGES/LARGE/SB-2018-00226M_0005.JPEG

Baier, T., Wichmann, J., Kruse, O., Lauersen, K.J., 2018b. Intron-containing algal transgenes mediate efficient recombinant gene expression in the green microalga *Chlamydomonas reinhardtii*. Nucleic Acids Res. 46, 6909–6919. https://doi.org/10.1093/nar/gky532

Banerjee, A., Ward, V., 2022. Production of recombinant and therapeutic proteins in microalgae. Curr. Opin. Biotechnol. 78, 102784. https://doi.org/10.1016/j.copbio.2022.102784

Bañuelos-Hernández, B., Monreal-Escalante, E., González-Ortega, O., Angulo, C., Rosales-Mendoza, S., 2017. Algevir: An expression system for microalgae based on viral vectors. Front. Microbiol. 8, 1100. https://doi.org/10.3389/FMICB.2017.01100/FULL

Barahimipour, R., Neupert, J., Bock, R., 2016. Efficient expression of nuclear transgenes in the green alga *Chlamydomonas* : synthesis of an HIV antigen and development of a new selectable marker. Plant Mol. Biol. 90, 403–418. https://doi.org/10.1007/s11103-015-0425-8

Barahimipour, R., Strenkert, D., Neupert, J., Schroda, M., Merchant, S.S., Bock, R., 2015. Dissecting the contributions of GC content and codon usage to gene expression in the model alga Chlamydomonas reinhardtii. Plant J. 84, 704–717. https://doi.org/10.1111/tpj.13033

Barbosa, M.J., Janssen, M., Südfeld, C., D'Adamo, S., Wijffels, R.H., 2023. Hypes, hopes, and the way forward for microalgal biotechnology. Trends Biotechnol. 41, 452–471. https://doi.org/10.1016/j.tibtech.2022.12.017

Barta, D.G., Coman, V., Vodnar, D.C., 2021. Microalgae as sources of omega-3 polyunsaturated fatty acids: Biotechnological aspects. Algal Res. 58, 102410. https://doi.org/10.1016/J.ALGAL.2021.102410

Bayne, A.C. V., Boltz, D., Owen, C., Betz, Y., Maia, G., Azadi, P., Archer-Hartmann, S., Zirkle, R., Lippmeier, J.C., 2013. Vaccination against Influenza with Recombinant Hemagglutinin Expressed by *Schizochytrium* sp. Confers Protective Immunity. PLoS One 8, e61790. https://doi.org/10.1371/JOURNAL.PONE.0061790

Beri, S., Gandhi, D., 2021. Quantification of residual cetyltrimethylammonium bromide (CTAB) and sodium deoxycholate (DOC) in Haemophilus influenzae type b (Hib) polysaccharide using NMR. Biologicals 70, 22–27. https://doi.org/10.1016/J.BIOLOGICALS.2021.02.001

Bermejo, R., Ruiz, E., Acien, F.G., 2007. Recovery of B-phycoerythrin using expanded bed adsorption chromatography: Scale-up of the process. Enzyme Microb. Technol. 40, 927–933. https://doi.org/10.1016/J.ENZMICTEC.2006.07.027

Bernaerts, T.M.M., Kyomugasho, C., Van Looveren, N., Gheysen, L., Foubert, I., Hendrickx, M.E., Van Loey, A.M., 2018. Molecular and rheological characterization of different cell wall fractions of *Porphyridium cruentum*. Carbohydr. Polym. 195, 542–550. https://doi.org/10.1016/J.CARBPOL.2018.05.001

Berndt, A.J., Smalley, T.N., Ren, B., Simkovsky, R., Badary, A., Sproles, A.E., Fields, F.J., Torres-Tiji, Y., Heredia, V., Mayfield, S.P., 2021. Recombinant production of a functional SARS-CoV-2 spike receptor binding domain in the green algae *Chlamydomonas reinhardtii*. PLoS One 16, e0257089. https://doi.org/10.1371/JOURNAL.PONE.0257089

Bhatt, A., Khanchandani, M., Rana, M.S., Prajapati, S.K., 2022. Techno-economic analysis of microalgae cultivation for commercial sustainability: A state-of-the-art review. J. Clean. Prod. 370, 133456. https://doi.org/10.1016/J.JCLEPRO.2022.133456

Bhattacharya, D., Price, D.C., Xin Chan, C., Qiu, H., Rose, N., Ball, S., Weber, A.P.M., Cecilia Arias, M., Henrissat, B., Coutinho, P.M., Krishnan, A., Zäuner, S., Morath, S., Hilliou, F., Egizi, A., Perrineau, M.M., Yoon, H.S., 2013. Genome of the red alga Porphyridium purpureum. Nat. Commun. 4, 1941. https://doi.org/10.1038/ncomms2931

Bhola, V., Swalaha, F., Ranjith Kumar, R., Singh, M., Bux, F., 2014. Overview of the potential of microalgae for CO2 sequestration. Int. J. Environ. Sci. Technol. 11, 2103–2118. https://doi.org/10.1007/S13762-013-0487-6/TABLES/3

Blobel, G., Sabatini, D.D., 1971. Ribosome-Membrane Interaction in Eukaryotic Cells. Biomembranes 193–195. https://doi.org/10.1007/978-1-4684-3330-2_16

Bock, R., 2015. Engineering plastid genomes: Methods, tools, and applications in basic research and biotechnology. Annu. Rev. Plant Biol. 66, 211–241. https://doi.org/10.1146/ANNUREV-ARPLANT-050213-040212

Brown, L.E., Sprecher, S.L., Keller, L.R., 1991. Introduction of exogenous DNA into *Chlamydomonas reinhardtii* by electroporation. Mol. Cell. Biol. 11, 2328–2332. https://doi.org/10.1128/MCB.11.4.2328

Bucci, M., 2020. First recombinant DNA vaccine for HBV. Nat. Res. 2021.

Cairns, B.R., 2009. The logic of chromatin architecture and remodelling at promoters. Nature 461, 193–198. https://doi.org/10.1038/nature08450

Campo, D.S., Dimitrova, Z., Yokosawa, J., Hoang, D., Perez, N.O., Ramachandran, S., Khudyakov, Y., 2012. Hepatitis C Virus Antigenic Convergence. Sci. Rep. 2, 267. https://doi.org/10.1038/srep00267

Chan, C.X., Yang, E.C., Banerjee, T., Yoon, H.S., Martone, P.T., Estevez, J.M., Bhattacharya, D., 2011. Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. Curr. Biol. 21, 328–333. https://doi.org/10.1016/j.cub.2011.01.037

Chankova, S.G., Dimova, E., Dimitrova, M., Bryant, P.E., 2007. Induction of DNA double-strand breaks by zeocin in *Chlamydomonas reinhardtii* and the role of increased DNA double-strand breaks rejoining in the formation of an adaptive response. Radiat. Environ. Biophys. 46, 409–416. https://doi.org/10.1007/s00411-007-0123-2

Chatterjee, M., Gupta, S., Bhar, A., Das, S., 2012. Optimization of an Efficient Protein Extraction Protocol Compatible with Two-Dimensional Electrophoresis and Mass Spectrometry from Recalcitrant Phenolic Rich Roots of Chickpea (Cicer arietinum L.) . Int. J. Proteomics 2012, 1–10. https://doi.org/10.1155/2012/536963

Chen, S.L., Lee, W., Hottes, A.K., Shapiro, L., McAdams, H.H., 2004. Codon usage between genomes is constrained by genome-wide mutational processes. Proc. Natl. Acad. Sci. U. S. A. 101, 3480–3485. https://doi.org/10.1073/PNAS.0307827100

Čížková, M., Slavková, M., Vítová, M., Zachleder, V., Bišová, K., 2019. Response of the Green Alga *Chlamydomonas reinhardtii* to the DNA Damaging Agent Zeocin. Cells 8, 735. https://doi.org/10.3390/cells8070735

Clarke, J.L., Paruch, L., Dobrica, M.-O., Caras, I., Tucureanu, C., Onu, A., Ciulean, S., Stavaru, C., Eerde, A., Wang, Y., Steen, H., Haugslien, S., Petrareanu, C., Lazar, C., Popescu, C.-I., Bock, R., Dubuisson, J., Branza-Nichita, N., 2017. Lettuce-produced hepatitis C virus E1E2 heterodimer triggers immune responses in mice and antibody production after oral vaccination. Plant Biotechnol. J. 15, 1611–1621. https://doi.org/10.1111/pbi.12743

Cohen, S.N., Chang, A.C.Y., Boyer, H.W., Helling, R.B., 1973. Construction of Biologically Functional Bacterial Plasmids In Vitro. Proc. Natl. Acad. Sci. U. S. A. 70, 3240. https://doi.org/10.1073/PNAS.70.11.3240

Cregg, J.M., Tschopp, J.F., Stillman, C., Siegel, R., Akong, M., Craig, W.S., Buckholz, R.G., Madden, K.R., Kellaris, P.A., Davis, G.R., Smiley, B.L., Cruze, J., Torregrossa, R., Velicçlebi, G., Thill, G.P., 1987. High–Level Expression and Efficient Assembly of Hepatitis B Surface Antigen in the Methylotrophic Yeast, *Pichia Pastoris*. Nat. Biotechnol. 5, 479–485. https://doi.org/10.1038/nbt0587-479

Crick, F., Watson, J., 1953. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature 737–738.

Crozet, P., Navarro, F.J., Willmund, F., Mehrshahi, P., Bakowski, K., Lauersen, K.J., Pérez-Pérez, M.-E., Auroy, P., Gorchs Rovira, A., Sauret-Gueto, S., Niemeyer, J., Spaniol, B., Theis, J., Trösch, R., Westrich, L.-D., Vavitsas, K., Baier, T., Hübner, W., de Carpentier, F., Cassarini, M., Danon, A., Henri, J., Marchand, C.H., de Mia, M., Sarkissian, K., Baulcombe, D.C., Peltier, G., Crespo, J.-L., Kruse, O., Jensen, P.-E., Schroda, M., Smith, A.G., Lemaire, S.D., 2018. Birth of a Photosynthetic Chassis: A MoClo Toolkit Enabling Synthetic Biology in the Microalga *Chlamydomonas reinhardtii*. ACS Synth. Biol. 7, 2074–2086. https://doi.org/10.1021/acssynbio.8b00251

Das, D., Pandya, M., 2018. Recent Advancement of Direct-acting Antiviral Agents (DAAs) in Hepatitis C Therapy. Mini-Reviews Med. Chem. 18, 584–596. https://doi.org/10.2174/1389557517666170913111930

Datta, S., Chatterjee, S., Veer, V., Chakravarty, R., 2012. Molecular Biology of the Hepatitis B Virus for Clinicians. J. Clin. Exp. Hepatol. 2, 353. https://doi.org/10.1016/J.JCEH.2012.10.003

Debuchy, R., Purton, S., Rochaix, J.D., 1989. The argininosuccinate lyase gene of *Chlamydomonas reinhardtii*: an important tool for nuclear transformation and for correlating the genetic and molecular maps of the ARG7 locus. EMBO J. 8, 2803–2809. https://doi.org/10.1002/J.1460-2075.1989.TB08426.X

Dehghani, J., Movafeghi, A., Mathieu-Rivet, E., Mati-Baouche, N., Calbo, S., Lerouge, P., Bardor, M., 2022. Microalgae as an Efficient Vehicle for the Production and Targeted Delivery of Therapeutic Glycoproteins against SARS-CoV-2 Variants. Mar. Drugs 2022, Vol. 20, Page 657 20, 657. https://doi.org/10.3390/MD20110657

Demain, A.L., Vaishnav, P., 2009. Production of recombinant proteins by microbes and higher

organisms. Biotechnol. Adv. 27, 297–306. https://doi.org/10.1016/J.BIOTECHADV.2009.01.008

Devi, A., Verma, M., Saratale, G.D., Saratale, R.G., Ferreira, L.F.R., Mulla, S.I., Bharagava, R.N., 2023. Microalgae: A green eco-friendly agents for bioremediation of tannery wastewater with simultaneous production of value-added products. Chemosphere 336, 139192. https://doi.org/10.1016/J.CHEMOSPHERE.2023.139192

Díaz-Rosales, P., Chabrillón, M., Abdala, R.T., Figueroa, F.L., Balebona, M.C., Moriñigo, M.A., 2008. Effect of dietary administration of *Porphyridium cruentum* on the respiratory burst activity of sole, Solea senegalensis (Kaup), phagocytes. J. Fish Dis. 31, 489–495. https://doi.org/10.1111/j.1365-2761.2008.00923.x

Díaz-Santos, E., de la Vega, M., Vila, M., Vigara, J., León, R., 2013. Efficiency of different heterologous promoters in the unicellular microalga *Chlamydomonas reinhardtii*. Biotechnol. Prog. 29, 319–328. https://doi.org/10.1002/BTPR.1690

Dobrica, M.-O., Lazar, C., Paruch, L., van Eerde, A., Clarke, J.L., Tucureanu, C., Caras, I., Ciulean, S., Onu, A., Tofan, V., Branzan, A., Urban, S., Stavaru, C., Branza-Nichita, N., 2018. Oral administration of a chimeric Hepatitis B Virus S/preS1 antigen produced in lettuce triggers infection neutralizing antibodies in mice. Vaccine 36, 5789–5795. https://doi.org/10.1016/j.vaccine.2018.07.072

Dobrica, M.O., Lazar, C., Branza-Nichita, N., 2020. N-Glycosylation and N-Glycan Processing in HBV Biology and Pathogenesis. Cells 9. https://doi.org/10.3390/CELLS9061404

Dobrica, M.O., Lazar, C., Paruch, L., Skomedal, H., Steen, H., Haugslien, S., Tucureanu, C., Caras, I., Onu, A., Ciulean, S., Branzan, A., Clarke, J.L., Stavaru, C., Branza-Nichita, N., 2017. A novel chimeric Hepatitis B virus S/preS1 antigen produced in mammalian and plant cells elicits stronger humoral and cellular immune response than the standard vaccine-constituent, S protein. Antiviral Res. 144, 256–265. https://doi.org/10.1016/j.antiviral.2017.06.017

Dobrica, M.O., van Eerde, A., Tucureanu, C., Onu, A., Paruch, L., Caras, I., Vlase, E., Steen, H., Haugslien, S., Alonzi, D., Zitzmann, N., Bock, R., Dubuisson, J., Popescu, C.I., Stavaru, C., Liu Clarke, J., Branza-Nichita, N., 2021. Hepatitis C virus E2 envelope glycoprotein produced in *Nicotiana benthamiana* triggers humoral response with virus-neutralizing activity in vaccinated mice. Plant Biotechnol. J. 19, 2027. https://doi.org/10.1111/PBI.13631

Dong, B., Hu, H.-H., Li, Z.-F., Cheng, R.-Q., Meng, D.-M., Wang, J., Fan, Z.-C., 2017. A novel

bicistronic expression system composed of the intraflagellar transport protein gene ift25 and FMDV 2A sequence directs robust nuclear gene expression in *Chlamydomonas reinhardtii*. Appl. Microbiol. Biotechnol. 101, 4227–4245. https://doi.org/10.1007/s00253-017-8177-9

Doron, L., Segal, N., Shapira, M., 2016. Transgene expression in microalgae—from tools to applications. Front. Plant Sci. 7, 177417. https://doi.org/10.3389/FPLS.2016.00505/BIBTEX

Dubuisson, J., Fabiani, F.L., 2007. Hepatitis C virus proteins. World J Gastroenterol 13, 2406–2415.

Dvir, I., Chayoth, R., Sod-Moriah, U., Shany, S., Nyska, A., Stark, A.H., Madar, Z., Arad, S.M., 2000. Soluble polysaccharide and biomass of red microalga *Porphyridium* sp. alter intestinal morphology and reduce serum cholesterol in rats. Br. J. Nutr. 84, 469–476. https://doi.org/10.1017/S000711450000177X

Einhaus, A., Baier, T., Rosenstengel, M., Freudenberg, R.A., Kruse, O., 2021. Rational Promoter Engineering Enables Robust Terpene Production in Microalgae. ACS Synth. Biol. 10, 847–856. https://doi.org/10.1021/ACSSYNBIO.0C00632/ASSET/IMAGES/LARGE/SB0C00632_0005.JPEG

Fabris, M., Abbriano, R.M., Pernice, M., Sutherland, D.L., Commault, A.S., Hall, C.C., Labeeuw, L., McCauley, J.I., Kuzhiuparambil, U., Ray, P., Kahlke, T., Ralph, P.J., 2020. Emerging Technologies in Algal Biotechnology: Toward the Establishment of a Sustainable, Algae-Based Bioeconomy. Front. Plant Sci. 11, 515802. https://doi.org/10.3389/FPLS.2020.00279/BIBTEX

Ferrari, C., Proost, S., Janowski, M., Becker, J., Nikoloski, Z., Bhattacharya, D., Price, D., Tohge, T., Bar-Even, A., Fernie, A., Stitt, M., Mutwil, M., 2019. Kingdom-wide comparison reveals the evolution of diurnal gene expression in Archaeplastida. Nat. Commun. 10, 1–13. https://doi.org/10.1038/s41467-019-08703-2

Fischer, N., Rochaix, J.D., 2001. The flanking regions of PsaD drive efficient gene expression in the nucleus of the green alga *Chlamydomonas reinhardtii* 265, 888–894.

Fischer, R., Buyel, J.F., 2020. Molecular farming – The slope of enlightenment. Biotechnol. Adv. 40, 107519. https://doi.org/10.1016/J.BIOTECHADV.2020.107519

Fox, J.M., Erill, I., 2010. Relative Codon Adaptation: A Generic Codon Bias Index for Prediction of Gene Expression. DNA Res. An Int. J. Rapid Publ. Reports Genes

Genomes 17, 185. https://doi.org/10.1093/DNARES/DSQ012

Freudenberg, R.A., Baier, T., Einhaus, A., Wobbe, L., Kruse, O., 2021. High cell density cultivation enables efficient and sustainable recombinant polyamine production in the microalga *Chlamydomonas reinhardtii*. Bioresour. Technol. 323, 124542. https://doi.org/10.1016/j.biortech.2020.124542

Fuentes-Grünewald, C., Bayliss, C., Zanain, M., Pooley, C., Scolamacchia, M., Silkina, A., 2015. Evaluation of batch and semi-continuous culture of *Porphyridium purpureum* in a photobioreactor in high latitudes using Fourier Transform Infrared spectroscopy for monitoring biomass composition and metabolites production. Bioresour. Technol. 189, 357–363. https://doi.org/10.1016/j.biortech.2015.04.042

Gaignard, C., Gargouch, N., Dubessay, P., Delattre, C., Pierre, G., Laroche, C., Fendri, I., Abdelkafi, S., Michaud, P., 2019. New horizons in culture and valorization of red microalgae. Biotechnol. Adv. 37, 193–222. https://doi.org/10.1016/j.biotechadv.2018.11.014

Gantt, E., Conti, S.F., 1965. THE ULTRASTRUCTURE OF *PORPHYRIDIUM CRUENTUM*. J. Cell Biol. 26, 365–381. https://doi.org/10.1083/jcb.26.2.365

Gaynes, R., 2017. The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use. Emerg. Infect. Dis. 23, 849. https://doi.org/10.3201/EID2305.161556

Geng, D., Wang, Y., Wang, P., Li, W., Sun, Y., 2003. Stable expression of hepatitis B surface antigen gene in *Dunaliella salina* (Chlorophyta). J. Appl. Phycol. 15, 451–456. https://doi.org/10.1023/B:JAPH.0000004298.89183.e5

Gerlich, W.H., 2017. Do we need better hepatitis B vaccines? Indian J. Med. Res. 145, 414. https://doi.org/10.4103/IJMR.IJMR_1852_16

Goeddel, D. V, Kleid, D.G., Bolivar, F., Heyneker, H.L., Yansura, D.G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K., Riggs, A.D., 1979. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. Proc. Natl. Acad. Sci. 76, 106–110. https://doi.org/10.1073/pnas.76.1.106

Goff, L.J., Coleman, A.W., 1990. Red algal plasmids. Curr. Genet. 18, 557–565. https://doi.org/10.1007/BF00327028

Golueke, C.G., Oswald, W.J., 1962. The Mass Culture of *Porphyridium cruentum*. Appl. Microbiol. 10, 102–107. https://doi.org/10.1128/AEM.10.2.102-107.1962

Gomez-Escobar, E., Roingeard, P., Beaumont, E., 2023. Current Hepatitis C Vaccine

Candidates Based on the Induction of Neutralizing Antibodies. Viruses 15. https://doi.org/10.3390/V15051151

Gould, S.B., Waller, R.F., McFadden, G.I., 2008. Plastid evolution. Annu. Rev. Plant Biol. https://doi.org/10.1146/annurev.arplant.59.032607.092915

Gregory, J.A., Li, F., Tomosada, L.M., Cox, C.J., Topol, A.B., Vinetz, J.M., Mayfield, S., 2012. Algae-Produced Pfs25 Elicits Antibodies That Inhibit Malaria Transmission. PLoS One 7, e37179. https://doi.org/10.1371/journal.pone.0037179

Grishagin, I. V., 2015. Automatic cell counting with ImageJ. Anal. Biochem. 473, 63–65. https://doi.org/10.1016/j.ab.2014.12.007

Guiry, M.D., 2012. HOW MANY SPECIES OF ALGAE ARE THERE? J. Phycol. 48, 1057–1063. https://doi.org/10.1111/j.1529-8817.2012.01222.x

Hadiji-Abbes, N., Borchani-Chabchoub, I., Triki, H., Ellouz, R., Gargouri, A., Mokdad-Gargouri, R., 2009. Expression of HBsAg and preS2-S protein in different yeast based system: A comparative analysis. Protein Expr. Purif. 66, 131–137. https://doi.org/10.1016/j.pep.2009.03.006

Han, S., Zhao, J., Liu, P., Wang, K., Qin, S., Zhao, Z., Cui, Y., 2022. Two Foreign Antimicrobial Peptides Expressed in the Chloroplast of *Porphyridium purpureum* Possessed Antibacterial Properties. Mar. Drugs 20. https://doi.org/10.3390/MD20080484

Hanson, G., Coller, J., 2018. Codon optimality, bias and usage in translation and mRNA decay. Nat. Rev. Mol. Cell Biol. 19, 20. https://doi.org/10.1038/NRM.2017.91

Harris, E.H., 2009. The Chlamydomonas Sourcebook: Introduction to Chlamydomonas and Its Laboratory Use: Volume 1. Elsevier Science.

Helliwell, K.E., Scaife, M.A., Sasso, S., Araujo, A.P.U., Purton, S., Smith, A.G., 2014. Unraveling Vitamin B 12 -Responsive Gene Regulation in Algae. Plant Physiol. 165, 388–397. https://doi.org/10.1104/pp.113.234369

Hempel, F., Lau, J., Klingl, A., Maier, U.G., 2011. Algae as protein factories: Expression of a human antibody and the respective antigen in the diatom phaeodactylum tricornutum. PLoS One 6. https://doi.org/10.1371/journal.pone.0028424

Herrscher, C., Roingeard, P., Blanchard, E., 2020. Hepatitis B Virus Entry into Cells. Cells 9, 1486. https://doi.org/10.3390/cells9061486

Hiatt, A., Cafferkey, R., Bowdish, K., 1989. Production of antibodies in transgenic plants. Nat. 1989 3426245 342, 76–78. https://doi.org/10.1038/342076a0

Ho, J.K.T., Jeevan-Raj, B., Netter, H.J., 2020. Hepatitis B Virus (HBV) Subviral Particles as Protective Vaccines and Vaccine Platforms. Viruses 12. https://doi.org/10.3390/V12020126

Howell, S.H., 2021. Evolution of the unfolded protein response in plants. Plant. Cell Environ. 44, 2625–2635. https://doi.org/10.1111/PCE.14063

Hu, J., Liu, K., 2017. Complete and incomplete hepatitis B virus particles: Formation, function, and application. Viruses. https://doi.org/10.3390/v9030056

Huheihel, M., Ishanu, V., Tal, J., Arad, S., 2002. Activity of *Porphyridium* sp. polysaccharide against herpes simplex viruses in vitro and in vivo. J. Biochem. Biophys. Methods 50, 189–200. https://doi.org/10.1016/S0165-022X(01)00186-5

Iriarte, A., Lamolle, G., Musto, H., 2021. Codon Usage Bias: An Endless Tale. J. Mol. Evol. 89, 589–593. https://doi.org/10.1007/S00239-021-10027-Z/FIGURES/1

Ji, L., Li, S., Chen, C., Jin, H., Wu, H., Fan, J., 2021. Physiological and transcriptome analysis elucidates the metabolic mechanism of versatile *Porphyridium purpureum* under nitrogen deprivation for exopolysaccharides accumulation. Bioresour. Bioprocess. 8, 73. https://doi.org/10.1186/s40643-021-00426-x

Jiao, K., Xiao, W., Xu, Y., Zeng, X., Ho, S.H., Laws, E.A., Lu, Y., Ling, X., Shi, T., Sun, Y., Tang, X., Lin, L., 2018. Using a trait-based approach to optimize mixotrophic growth of the red microalga Porphyridium purpureum towards fatty acid production. Biotechnol. Biofuels 11, 273. https://doi.org/10.1186/s13068-018-1277-7

Johnson, N., Powis, K., High, S., 2013. Post-translational translocation into the endoplasmic reticulum. Biochim. Biophys. Acta - Mol. Cell Res. 1833, 2403–2409. https://doi.org/10.1016/J.BBAMCR.2012.12.008

Joshi, C.J., Ke, W., Drangowska-Way, A., O'Rourke, E.J., Lewis, N.E., 2022. What are housekeeping genes? PLoS Comput. Biol. 18. https://doi.org/10.1371/JOURNAL.PCBI.1010295

Kanaya, S., Yamada, Y., Kinouchi, M., Kudo, Y., Ikemura, T., 2001. Codon usage and tRNA genes in eukaryotes: correlation of codon usage diversity with translation efficiency and with CG-dinucleotide usage as assessed by multivariate analysis. J. Mol. Evol. 53, 290–298. https://doi.org/10.1007/S002390010219

Kang, D., Gho, Y.S., Suh, M., Kang, C., 2002. Highly sensitive and fast protein detection with Coomassie brilliant blue in sodium dodecyl sulfate-polyacrylamide gel electrophoresis

[5]. Bull. Korean Chem. Soc. 23, 1511–1512. https://doi.org/10.5012/bkcs.2002.23.11.1511

Kathiresan, S., Sarada, R., Bhattacharya, S., Ravishankar, G.A., 2007. Culture media optimization for growth and phycoerythrin production from Porphyridium purpureum. Biotechnol. Bioeng. 96, 456–463. https://doi.org/10.1002/bit.21138

Khan, A.H., Bayat, H., Rajabibazl, M., Sabri, S., Rahimpour, A., 2016. Humanizing glycosylation pathways in eukaryotic expression systems. World J. Microbiol. Biotechnol. 2016 331 33, 1–12. https://doi.org/10.1007/S11274-016-2172-7

Kiefer, A.M., Niemeyer, J., Probst, A., Erkel, G., Schroda, M., 2022. Production and secretion of functional SARS-CoV-2 spike protein in *Chlamydomonas reinhardtii*. Front. Plant Sci. 13. https://doi.org/10.3389/FPLS.2022.988870

Kim, D., Lee, J., Cho, C.H., Kim, E.J., Bhattacharya, D., Yoon, H.S., 2022. Group II intron and repeat-rich red algal mitochondrial genomes demonstrate the dynamic recent history of autocatalytic RNAs. BMC Biol. 20, 2. https://doi.org/10.1186/s12915-021-01200-3

Kindle, K.L., 1990. High-frequency nuclear transformation of *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. U. S. A. 87, 1228–32.

Kindle, K.L., Schnell, R.A., Fernández, E., Lefebvre, P.A., 1989. Stable nuclear transformation of Chlamydomonas using the Chlamydomonas gene for nitrate reductase. J. Cell Biol. 109, 2589–601.

Kober, L., Zehe, C., Bode, J., 2013. Optimized signal peptides for the development of high expressing CHO cell lines. Biotechnol. Bioeng. 110, 1164–1173. https://doi.org/10.1002/BIT.24776

Kucho, K., Ohyama, K., Fukuzawa, H., 1999. CO2-Responsive Transcriptional Regulation of CAH1 Encoding Carbonic Anhydrase Is Mediated by Enhancer and Silencer Regions in *Chlamydomonas reinhardtii*. Plant Physiol. 121, 1329–1337. https://doi.org/10.1104/pp.121.4.1329

Kumar, A., Falcao, V.R., Sayre, R.T., 2013. Evaluating nuclear transgene expression systems in *Chlamydomonas reinhardtii*. Algal Res. 2, 321–332. https://doi.org/10.1016/J.ALGAL.2013.09.002

Kyhse-Andersen, J., 1984. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. J. Biochem. Biopltvsical Methods 10, 203–209.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685. https://doi.org/10.1038/227680a0

Lapidot, M., Raveh, D., Sivan, A., Arad, S., Shapira, M., 2002. Stable chloroplast transformation of the unicellular red alga Porphyridium species. Plant Physiol. 129, 7–12. https://doi.org/10.1104/pp.011023

Lauersen, K.J., Berger, H., Mussgnug, J.H., Kruse, O., 2013. Efficient recombinant protein production and secretion from nuclear transgenes in *Chlamydomonas reinhardtii*. J. Biotechnol. 167, 101–110. https://doi.org/10.1016/J.JBIOTEC.2012.10.010

Lazar, C., Uta, M., Branza-Nichita, N., 2014. Modulation of the unfolded protein response by the human hepatitis B virus. Front. Microbiol. 5, 99882. https://doi.org/10.3389/FMICB.2014.00433/BIBTEX

Lee, J., Kim, K.M., Yang, E.C., Miller, K.A., Boo, S.M., Bhattacharya, D., Yoon, H.S., 2016. Reconstructing the complex evolutionary history of mobile plasmids in red algal genomes. Sci. Rep. 6, 23744. https://doi.org/10.1038/srep23744

Lee, J.M., Kim, D., Bhattacharya, D., Yoon, H.S., 2019. Expansion of phycobilisome linker gene families in mesophilic red algae. Nat. Commun. 10. https://doi.org/10.1038/s41467-019-12779-1

Lee, Y., Rio, D.C., 2015. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. Annu. Rev. Biochem. 84, 291–323. https://doi.org/10.1146/ANNUREV-BIOCHEM-060614-034316

León, R., Fernández, E., 2007. Nuclear Transformation of Eukaryotic Microalgae, in: Transgenic Microalgae as Green Cell Factories. Springer New York, New York, NY, pp. 1–11. https://doi.org/10.1007/978-0-387-75532-8_1

Levy-Ontman, O., Fisher, M., Shotland, Y., Weinstein, Y., Tekoah, Y., Arad, S.M., 2014. Genes involved in the endoplasmic reticulum N-Glycosylation pathway of the red microalga *Porphyridium* sp.: A bioinformatic study. Int. J. Mol. Sci. 15, 2305–2326. https://doi.org/10.3390/ijms15022305

Li, H., d'Anjou, M., 2009. Pharmacological significance of glycosylation in therapeutic proteins. Curr. Opin. Biotechnol. 20, 678–684. https://doi.org/10.1016/J.COPBIO.2009.10.009

Li, J., Sun, X., Zhang, Y., 2006. Improvement of hepatitis B surface antigen expression by dimethyl sulfoxide in the culture of recombinant Chinese hamster ovary cells. Process Biochem. 41, 317–322. https://doi.org/10.1016/j.procbio.2005.08.017

Li, Q., Chen, Y., Liu, X., Li, Y., Xu, J., Li, T., Xiang, W., Li, A., 2023. Effect of salinity on the biochemical characteristics and antioxidant activity of exopolysaccharide of *Porphyridium purpureum* FACHB 806. Front. Mar. Sci. 9. https://doi.org/10.3389/fmars.2022.1097200

Li, S., Ji, L., Shi, Q., Wu, H., Fan, J., 2019. Advances in the production of bioactive substances from marine unicellular microalgae Porphyridium spp. Bioresour. Technol. 292, 122048. https://doi.org/10.1016/j.biortech.2019.122048

Li, T., Xu, J., Wu, H., Jiang, P., Chen, Z., Xiang, W., 2019. Growth and Biochemical Composition of *Porphyridium purpureum* SCS-02 under Different Nitrogen Concentrations. Mar. Drugs 17, 124. https://doi.org/10.3390/md17020124

Li, X., Zhang, R., Patena, W., Gang, S.S., Blum, S.R., Ivanova, N., Yue, R., Robertson, J.M., Lefebvre, P.A., Fitz-Gibbon, S.T., Grossman, A.R., Jonikasa, M.C., 2016. An Indexed, Mapped Mutant Library Enables Reverse Genetics Studies of Biological Processes in *Chlamydomonas reinhardtii*. Plant Cell 28, 367–387. https://doi.org/10.1105/TPC.15.00465

Li, Z., Bock, R., 2018. Replication of bacterial plasmids in the nucleus of the red alga Porphyridium purpureum. Nat. Commun. 9. https://doi.org/10.1038/s41467-018-05651-1

Liang, T.J., 2009. Hepatitis B: The Virus and Disease. Hepatology 49, S13. https://doi.org/10.1002/HEP.22881

Liu, X., Xie, X., Du, H., Sanganyado, E., Wang, W., Aslam, M., Chen, J., Chen, W., Liang, H., 2021. Bioinformatic analysis and genetic engineering approaches for recombinant biopharmaceutical glycoproteins production in microalgae. Algal Res. 55, 102276. https://doi.org/10.1016/J.ALGAL.2021.102276

López-Paz, C., Liu, D., Geng, S., Umen, J.G., 2017. Identification of *Chlamydomonas reinhardtii* endogenous genic flanking sequences for improved transgene expression. Plant J. 92, 1232–1244. https://doi.org/10.1111/TPJ.13731

Lu, X., Nan, F., Feng, J., Lv, J., Liu, Q., Liu, X., Xie, S., 2020. Effects of Different Environmental Factors on the Growth and Bioactive Substance Accumulation of *Porphyridium purpureum*. Int. J. Environ. Res. Public Health 17, 2221. https://doi.org/10.3390/ijerph17072221

Luo, H.-P., Al-Dahhan, M.H., 2012. Airlift column photobioreactors for *Porphyridium* sp. culturing: Part I. effects of hydrodynamics and reactor geometry. Biotechnol. Bioeng. 109, 932–941. https://doi.org/10.1002/bit.24361

Ma, K., Deng, L., Wu, H., Fan, J., 2022. Towards green biomanufacturing of high-value recombinant proteins using promising cell factory: *Chlamydomonas reinhardtii* chloroplast. Bioresour. Bioprocess. 2022 91 9, 1–14. https://doi.org/10.1186/S40643-022-00568-6

Malla, A., Rosales-Mendoza, S., Phoolcharoen, W., Vimolmangkang, S., 2021. Efficient Transient Expression of Recombinant Proteins Using DNA Viral Vectors in Freshwater Microalgal Species. Front. Plant Sci. 12. https://doi.org/10.3389/FPLS.2021.650820

Manuell, A.L., Beligni, M.V., Elder, J.H., Siefker, D.T., Tran, M., Weber, A., McDonald, T.L., Mayfield, S.P., 2007. Robust expression of a bioactive mammalian protein in *Chlamydomonas* chloroplast. Plant Biotechnol. J. 5, 402–412. https://doi.org/10.1111/j.1467-7652.2007.00249.x

Marcati, A., Ursu, A.V., Laroche, C., Soanen, N., Marchal, L., Jubeau, S., Djelveh, G., Michaud, P., 2014. Extraction and fractionation of polysaccharides and B-phycoerythrin from the microalga *Porphyridium cruentum* by membrane technology. Algal Res. 5, 258–263. https://doi.org/10.1016/J.ALGAL.2014.03.006

Mayfield, S.P., Franklin, S.E., Lerner, R.A., 2003. Expression and assembly of a fully active antibody in algae. Proc. Natl. Acad. Sci. U. S. A. 100, 438–442. https://doi.org/10.1073/PNAS.0237108100

Mayfield, S.P., Kindle, K.L., 1990. Stable nuclear transformation of *Chlamydomonas reinhardtii* by using a C. reinhardtii gene as the selectable marker. Proc. Natl. Acad. Sci. 87, 2087–2091. https://doi.org/10.1073/pnas.87.6.2087

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., Maréchal-Drouard, L., Marshall, W.F., Qu, L.-H., Nelson, D.R., Sanderfoot, A.A., Spalding, M.H., Kapitonov, V. V, Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S.M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.-L., Cognat, V., Croft, M.T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P.A., Lemaire, S.D., Lobanov, A. V, Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J. V, Moseley, J., Napoli, C., Nedelcu, A.M., Niyogi, K., Novoselov, S. V, Paulsen, I.T., Pazour, G., Purton, S., Ral, J.-P., Riaño-Pachón, D.M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S.L., Allmer, J., Balk, J., Bisova, K., Chen, C.-J., Elias, M., Gendler, K., Hauser, C., Lamb, M.R., Ledford, H., Long, J.C., Minagawa, J., Page, M.D., Pan, J., Pootakham,

W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A.M., Yang, P., Ball, S., Bowler, C., Dieckmann, C.L., Gladyshev, V.N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y.W., Jhaveri, J., Luo, Y., Martínez, D., Ngau, W.C.A., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I. V, Rokhsar, D.S., Grossman, A.R., Grossman, A.R., 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. Science 318, 245–50. https://doi.org/10.1126/science.1143609

Milito, A., Aschern, M., McQuillan, J.L., Yang, J.-S., 2023. Challenges and advances towards the rational design of microalgal synthetic promoters in *Chlamydomonas reinhardtii*. J. Exp. Bot. https://doi.org/10.1093/JXB/ERAD100

Mohanta, T.K., Mishra, A.K., Khan, A., Hashem, A., Abd_Allah, E.F., Al-Harrasi, A., 2020. Gene Loss and Evolution of the Plastome. Genes (Basel). 11, 1133. https://doi.org/10.3390/genes11101133

Molino, J.V.D., de Carvalho, J.C.M., Mayfield, S.P., 2018. Comparison of secretory signal peptides for heterologous protein expression in microalgae: Expanding the secretion portfolio for *Chlamydomonas reinhardtii*. PLoS One 13. https://doi.org/10.1371/JOURNAL.PONE.0192433

Moroney, J. V., Ma, Y., Frey, W.D., Fusilier, K.A., Pham, T.T., Simms, T.A., DiMario, R.J., Yang, J., Mukherjee, B., 2011. The carbonic anhydrase isoforms of *Chlamydomonas reinhardtii*: intracellular location, expression, and physiological roles. Photosynth. Res. 109, 133–149. https://doi.org/10.1007/S11120-011-9635-3

Moulin, M., Nguyen, G.T.D.T., Scaife, M.A., Smith, A.G., Fitzpatrick, T.B., 2013. Analysis of *Chlamydomonas* thiamin metabolism in vivo reveals riboswitch plasticity. Proc. Natl. Acad. Sci. U. S. A. 110, 14622–14627. https://doi.org/10.1073/PNAS.1307741110/-/DCSUPPLEMENTAL/PNAS.201307741SI.PDF

Müller, F., Tora, L., 2014. Chromatin and DNA sequences in defining promoters for transcription initiation. Biochim. Biophys. Acta - Gene Regul. Mech. 1839, 118–128. https://doi.org/10.1016/j.bbagrm.2013.11.003

Nägeli, C., 1849. Gattungen einzelliger Algen, physiologisch und systematisch bearbeitet. Neue Denkschriften der Allg. Schweizerischen Gesellschaft für die Gesammten Naturwissenschaften 10, i–viii, 1–139.

Neupert, J., Gallaher, S.D., Lu, Y., Strenkert, D., Segal, N., Barahimipour, R., Fitz-Gibbon,

S.T., Schroda, M., Merchant, S.S., Bock, R., 2020. An epigenetic gene silencing pathway selectively acting on transgenic DNA in the green alga *Chlamydomonas*. Nat. Commun. 11. https://doi.org/10.1038/S41467-020-19983-4

Neupert, J., Karcher, D., Bock, R., 2009. Generation of Chlamydomonas strains that efficiently express nuclear transgenes. Plant J. 57, 1140–1150. https://doi.org/10.1111/j.1365-313X.2008.03746.x

Niccolai, A., Chini Zittelli, G., Rodolfi, L., Biondi, N., Tredici, M.R., 2019. Microalgae of interest as food source: Biochemical composition and digestibility. Algal Res. 42. https://doi.org/10.1016/j.algal.2019.101617

Niemeyer, J., Scheuring, D., Oestreicher, J., Morgan, B., Schroda, M., 2021. Real-time monitoring of subcellular H2O2 distribution in *Chlamydomonas reinhardtii*. Plant Cell 33, 2935–2949. https://doi.org/10.1093/plcell/koab176

Novoa, E.M., Ribas de Pouplana, L., 2012. Speeding with control: codon usage, tRNAs, and ribosomes. Trends Genet. 28, 574–581. https://doi.org/10.1016/J.TIG.2012.07.006

Nyathi, Y., Wilkinson, B.M., Pool, M.R., 2013. Co-translational targeting and translocation of proteins to the endoplasmic reticulum. Biochim. Biophys. Acta 1833, 2392–2402. https://doi.org/10.1016/J.BBAMCR.2013.02.021

Oh, S.H., Han, J.G., Kim, Y., Ha, J.H., Kim, S.S., Jeong, M.H., Jeong, H.S., Kim, N.Y., Cho, J.S., Yoon, W.B., Lee, S.Y., Kang, D.H., Lee, H.Y., 2009. Lipid production in Porphyridium cruentum grown under different culture conditions. J. Biosci. Bioeng. 108, 429–434. https://doi.org/10.1016/j.jbiosc.2009.05.020

Ohresser, M., Matagne, R.F., Loppes, R., 1997. Expression of the arylsulphatase reporter gene under the control of the nit1 promoter in *Chlamydomonas reinhardtii*. Curr. Genet. 31, 264–271. https://doi.org/10.1007/s002940050204

Orlova, O. V., Drutsa, V.L., Spirin, P. V., Prasolov, V.S., Rubtsov, P.M., Kochetkov, S.N., Beljelarskaya, S.N., 2015. The Role of HCV E2 Protein Glycosylation in Functioning of Virus Envelope Proteins in Insect and Mammalian Cells. Acta Naturae 7, 87. https://doi.org/10.32607/20758251-2015-7-1-87-97

Owczarek, B., Gerszberg, A., Hnatuszko-Konka, K., 2019. A Brief Reminder of Systems of Production and Chromatography-Based Recovery of Recombinant Protein Biopharmaceuticals. Biomed Res. Int. 2019. https://doi.org/10.1155/2019/4216060

Owji, H., Nezafat, N., Negahdaripour, M., Hajiebrahimi, A., Ghasemi, Y., 2018. A

comprehensive review of signal peptides: Structure, roles, and applications. Eur. J. Cell Biol. 97, 422–441. https://doi.org/10.1016/J.EJCB.2018.06.003

Pantazica, A.-M.M., Cucos, L.-M., Stavaru, C., Clarke, J.-L., Branza-Nichita, N., 2021. Challenges and Prospects of Plant-Derived Oral Vaccines against Hepatitis B and C Viruses. Plants 10, 2037. https://doi.org/10.3390/plants10102037

Pantazica, A., van Eerde, A., Dobrica, M., Caras, I., Ionescu, I., Costache, A., Tucureanu, C., Steen, H., Lazar, C., Heldal, I., Haugslien, S., Onu, A., Stavaru, C., Branza- Nichita, N., Liu Clarke, J., 2023. The “humanized” N- glycosylation pathway in CRISPR /Cas9-edited *Nicotiana benthamiana* significantly enhances the immunogenicity of a S/preS1 Hepatitis B Virus antigen and the virus- neutralizing antibody response in vaccinated. Plant Biotechnol. J. 21, 1176–1190. https://doi.org/10.1111/pbi.14028

Pantazica, A.M., Dobrica, M.O., Lazar, C., Scurtu, C., Tucureanu, C., Caras, I., Ionescu, I., Costache, A., Onu, A., Clarke, J.L., Stavaru, C., Branza-Nichita, N., 2022. Efficient cellular and humoral immune response and production of virus-neutralizing antibodies by the Hepatitis B Virus S/preS116-42 antigen. Front. Immunol. 13. https://doi.org/10.3389/FIMMU.2022.941243

Patient, R., Hourioux, C., Sizaret, P.-Y., Trassard, S., Sureau, C., Roingeard, P., 2007. Hepatitis B Virus Subviral Envelope Particle Morphogenesis and Intracellular Trafficking. J. Virol. 81, 3842–3851. https://doi.org/10.1128/jvi.02741-06

Petruzziello, A., Marigliano, S., Loquercio, G., Cacciapuoti, C., 2016. Hepatitis C virus (HCV) genotypes distribution: An epidemiological up-date in Europe. Infect. Agent. Cancer 11, 1–9. https://doi.org/10.1186/S13027-016-0099-0/FIGURES/3

Pierce, B.G., Keck, Z.-Y., Wang, R., Lau, P., Garagusi, K., Elkholy, K., Toth, E.A., Urbanowicz, R.A., Guest, J.D., Agnihotri, P., Kerzic, M.C., Marin, A., Andrianov, A.K., Ball, J.K., Mariuzza, R.A., Fuerst, T.R., Foung, S.K.H., 2020. Structure-Based Design of Hepatitis C Virus E2 Glycoprotein Improves Serum Binding and Cross-Neutralization. J. Virol. 94. https://doi.org/10.1128/JVI.00704-20

Pileri, P., Uematsu, Y., Campagnoli, S., Galli, G., Falugi, F., Petracca, R., Weiner, A.J., Houghton, M., Rosa, D., Grandi, G., Abrignani, S., 1998. Binding of Hepatitis C Virus to CD81. Science (80-.). 282, 938–941. https://doi.org/10.1126/science.282.5390.938

Prasad, B., Lein, W., Thiyam, G., Lindenberger, C.P., Buchholz, R., Vadakedath, N., 2018. Stable nuclear transformation of rhodophyte species Porphyridium purpureum: advanced molecular tools and an optimized method. Photosynth. Res. 140, 173–188.

https://doi.org/10.1007/s11120-018-0587-8

Prentoe, J., Bukh, J., 2018. Hypervariable Region 1 in Envelope Protein 2 of Hepatitis C Virus: A Linchpin in Neutralizing Antibody Evasion and Viral Entry. Front. Immunol. https://doi.org/10.3389/fimmu.2018.02146

Prentoe, J., Verhoye, L., Velázquez Moctezuma, R., Buysschaert, C., Farhoudi, A., Wang, R., Alter, H., Meuleman, P., Bukh, J., 2016. HVR1-mediated antibody evasion of highly infectious in vivo adapted HCV in humanised mice. Gut 65, 1988–1997. https://doi.org/10.1136/gutjnl-2015-310300

Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R., Coller, J., 2015. Codon optimality is a major determinant of mRNA stability. Cell 160, 1111–1124. https://doi.org/10.1016/j.cell.2015.02.029

Qiu, H., Yoon, H.S., Bhattacharya, D., 2016. Red Algal Phylogenomics Provides a Robust Framework for Inferring Evolution of Key Metabolic Pathways. PLoS Curr. 8. https://doi.org/10.1371/CURRENTS.TOL.7B037376E6D84A1BE34AF756A4D90846

Ramesh, V.M., Bingham, S.E., Webber, A.N., 2011. A Simple Method for Chloroplast Transformation in *Chlamydomonas reinhardtii*. pp. 313–320. https://doi.org/10.1007/978-1-60761-925-3_23

Ramos-Martinez, E.M., Fimognari, L., Sakuragi, Y., 2017. High- yield secretion of recombinant proteins from the microalga *Chlamydomonas reinhardtii*. Plant Biotechnol. J. 15, 1214. https://doi.org/10.1111/PBI.12710

Ramos-Vega, A., Angulo, C., Bañuelos-Hernández, B., Monreal-Escalante, E., 2021. Microalgae-made vaccines against infectious diseases. Algal Res. 58, 102408. https://doi.org/10.1016/J.ALGAL.2021.102408

Rasala, B.A., Muto, M., Lee, P.A., Jager, M., Cardoso, R.M.F., Behnke, C.A., Kirk, P., Hokanson, C.A., Crea, R., Mendez, M., Mayfield, S.P., 2010. Production of therapeutic proteins in algae, analysis of expression of seven human proteins in the chloroplast of *Chlamydomonas reinhardtii*. Plant Biotechnol. J. 8, 719–733. https://doi.org/10.1111/J.1467-7652.2010.00503.X

Rosales-Mendoza, S., Solís-Andrade, K.I., Márquez-Escobar, V.A., González-Ortega, O., Bañuelos-Hernandez, B., 2020. Current advances in the algae-made biopharmaceuticals field. Expert Opin. Biol. Ther. 20, 751–766. https://doi.org/10.1080/14712598.2020.1739643

Rosano, G.L., Ceccarelli, E.A., 2014. Recombinant protein expression in *Escherichia coli*: Advances and challenges. Front. Microbiol. 5, 79503. https://doi.org/10.3389/FMICB.2014.00172/BIBTEX

Rose, A.B., 2019. Introns as Gene Regulators: A Brick on the Accelerator. Front. Genet. 9. https://doi.org/10.3389/fgene.2018.00672

Rubin, L.G., Levin, M.J., Ljungman, P., Davies, E.G., Avery, R., Tomblyn, M., Bousvaros, A., Dhanireddy, S., Sung, L., Keyserling, H., Kang, I., 2014. 2013 IDSA clinical practice guideline for vaccination of the immunocompromised host. Clin. Infect. Dis. 58, 309–318. https://doi.org/10.1093/CID/CIT684

Safi, C., Charton, M., Pignolet, O., Pontalier, P.Y., Vaca-Garcia, C., 2013. Evaluation of the protein quality of *Porphyridium cruentum*. J. Appl. Phycol. 25, 497–501. https://doi.org/10.1007/S10811-012-9883-4/TABLES/2

Sahana, H. V., Sarala, N., Prasad, S.R., 2017. Decrease in Anti-HBs Antibodies over Time in Medical Students and Healthcare Workers after Hepatitis B Vaccination. Biomed Res. Int. 2017, 1–5. https://doi.org/10.1155/2017/1327492

Saunders, R., Deane, C.M., 2010. Synonymous codon usage influences the local protein structure observed. Nucleic Acids Res. 38, 6719–6728. https://doi.org/10.1093/NAR/GKQ495

Schillberg, S., Raven, N., Spiegel, H., Rasche, S., Buntru, M., 2019. Critical analysis of the commercial potential of plants for the production of recombinant proteins. Front. Plant Sci. 10, 453222. https://doi.org/10.3389/FPLS.2019.00720/BIBTEX

Schoberer, J., Shin, Y.J., Vavra, U., Veit, C., Strasser, R., 2018. Analysis of protein glycosylation in the ER. Methods Mol. Biol. 1691, 205–222. https://doi.org/10.1007/978-1-4939-7389-7_16

Schroda, M., 2019. Good News for Nuclear Transgene Expression in *Chlamydomonas*. Cells 8. https://doi.org/10.3390/CELLS8121534

Schroda, M., Blocker, D., Beck, C.F., 2000a. The HSP70A promoter as a tool for the improved expression of transgenes in *Chlamydomonas*. Plant J. 21, 121–131. https://doi.org/10.1046/j.1365-313x.2000.00652.x

Schroda, M., Blöcker, D., Beck, C.F., 2000b. The HSP70A promoter as a tool for the improved expression of transgenes in Chlamydomonas. Plant J. 21, 121–131. https://doi.org/10.1046/j.1365-313X.2000.00652.x

Šefcová, M.A., Santacruz, F., Larrea-Álvarez, C.M., Vinueza-Burgos, C., Ortega-Paredes, D., Molina-Cuasapaz, G., Rodríguez, J., Calero-Cáceres, W., Revajová, V., Fernández-Moreira, E., Larrea-Álvarez, M., 2021. Administration of Dietary Microalgae Ameliorates Intestinal Parameters, Improves Body Weight, and Reduces Thawing Loss of Fillets in Broiler Chickens: A Pilot Study. Animals 11, 3601. https://doi.org/10.3390/ani11123601

Seitz, S., Habjanič, J., Schütz, A.K., Bartenschlager, R., 2020. The Hepatitis B Virus Envelope Proteins: Molecular Gymnastics Throughout the Viral Life Cycle. https://doi.org/10.1146/annurev-virology-092818

Shamir, M., Bar-On, Y., Phillips, R., Milo, R., 2016. SnapShot: Timescales in Cell Biology. Cell 164, 1302-1302.e1. https://doi.org/10.1016/j.cell.2016.02.058

Shaul, O., 2017. How introns enhance gene expression. Int. J. Biochem. Cell Biol. 91, 145–155. https://doi.org/10.1016/j.biocel.2017.06.016

Shouval, D., Roggendorf, H., Roggendorf, M., 2015. Enhanced immune response to hepatitis B vaccination through immunization with a Pre-S1/Pre-S2/S Vaccine. Med. Microbiol. Immunol. 204, 57–68. https://doi.org/10.1007/s00430-014-0374-x

Shpak, E., Leykam, J.F., Kieliszewski, M.J., 1999. Synthetic genes for glycoprotein design and the elucidation of hydroxyproline- O -glycosylation codes. Proc. Natl. Acad. Sci. 96, 14736–14741. https://doi.org/10.1073/pnas.96.26.14736

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E. V., Zdobnov, E.M., 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Singh, S.P., Singh, P., 2014. Effect of CO2 concentration on algal growth: A review. Renew. Sustain. Energy Rev. 38, 172–179. https://doi.org/10.1016/j.rser.2014.05.043

Sivan, A., Arad, S. (Malis), 1995. A mutant of the red microalga *Porphyridium* sp. (Rhodophyceae) resistant to DCMU and atrazine. Phycologia 34, 299–305. https://doi.org/10.2216/i0031-8884-34-4-299.1

Skeffington, A.W., Grimm, A., Schönefeld, S., Petersen, K., Scheffel, A., 2019. An Efficient Method for the Plating of Haploid and Diploid *Emiliania huxleyi* on Solid Medium1. J. Phycol. 56, 238–242. https://doi.org/10.1111/JPY.12942

Slattery, S.S., Giguere, D.J., Stuckless, E.E., Shrestha, A., Briere, L.A.K., Galbraith, A., Reaume, S., Boyko, X., Say, H.H., Browne, T.S., Frederick, M.I., Lant, J.T., Heinemann, I.U., O'Donoghue, P., Dsouza, L., Martin, S., Howard, P., Jedeszko, C., Ali, K., Styba, G.,

Flatley, M., Karas, B.J., Gloor, G.B., Edgell, D.R., 2022. Phosphate-regulated expression of the SARS-CoV-2 receptor-binding domain in the diatom *Phaeodactylum tricornutum* for pandemic diagnostics. Sci. Rep. 12. https://doi.org/10.1038/S41598-022-11053-7

Sommer, M.S., Schleiff, E., 2014. Protein Targeting and Transport as a Necessary Consequence of Increased Cellular Complexity. Cold Spring Harb. Perspect. Biol. 6. https://doi.org/10.1101/CSHPERSPECT.A016055

Sørensen, M.A., Kurland, C.G., Pedersen, S., 1989. Codon usage determines translation rate in *Escherichia coli*. J. Mol. Biol. 207, 365–377. https://doi.org/10.1016/0022-2836(89)90260-X

Soria-Guerra, R.E., Ramírez-Alonso, J.I., Ibáñez-Salazar, A., Govea-Alonso, D.O., Paz-Maldonado, L.M.T., Bañuelos-Hernández, B., Korban, S.S., Rosales-Mendoza, S., 2014a. Expression of an HBcAg-based antigen carrying angiotensin II in *Chlamydomonas reinhardtii* as a candidate hypertension vaccine. Plant Cell. Tissue Organ Cult. 116, 133–139. https://doi.org/10.1007/s11240-013-0388-x

Soria-Guerra, R.E., Ramírez-Alonso, J.I., Ibáñez-Salazar, A., Govea-Alonso, D.O., Paz-Maldonado, L.M.T., Bañuelos-Hernández, B., Korban, S.S., Rosales-Mendoza, S., 2014b. Expression of an HBcAg-based antigen carrying angiotensin II in *Chlamydomonas reinhardtii* as a candidate hypertension vaccine. Plant Cell. Tissue Organ Cult. 116, 133–139. https://doi.org/10.1007/S11240-013-0388-X/FIGURES/4

Specht, E.A., Mayfield, S.P., 2014. Algae-based oral recombinant vaccines. Front. Microbiol. 5, 1–7. https://doi.org/10.3389/fmicb.2014.00060

Stevens, D.R., Roehaix, J.-D., Purton, S., 1996. The bacterial phleomycin resistance geneble as a dominant selectable marker in *Chlamydomonas*. Mol. Gen. Genet. MGG 1996 2511 251, 23–30. https://doi.org/10.1007/BF02174340

Su, G., Jiao, K., Li, Z., Guo, X., Chang, J., Ndikubwimana, T., Sun, Y., Zeng, X., Lu, Y., Lin, L., 2016. Phosphate limitation promotes unsaturated fatty acids and arachidonic acid biosynthesis by microalgae *Porphyridium purpureum*. Bioprocess Biosyst. Eng. 39, 1129–1136. https://doi.org/10.1007/s00449-016-1589-6

Su, H., van Eerde, A., Rimstad, E., Bock, R., Branza-Nichita, N., Yakovlev, I.A., Clarke, J.L., 2023. Plant-made vaccines against viral diseases in humans and farm animals. Front. Plant Sci. 14. https://doi.org/10.3389/FPLS.2023.1170815

Suffner, S., Gerstenberg, N., Patra, M., Ruibal, P., Orabi, A., Schindler, M., Bruss, V., 2018. Domains of the Hepatitis B Virus Small Surface Protein S Mediating Oligomerization. J.

Virol. 92. https://doi.org/10.1128/JVI.02232-17

Tajima, N., Sato, S., Maruyama, F., Kurokawa, K., Ohta, H., Tabata, S., Sekine, K., Moriyama, T., Sato, N., 2014. Analysis of the complete plastid genome of the unicellular red alga *Porphyridium purpureum*. J. Plant Res. 127, 389–397. https://doi.org/10.1007/S10265-014-0627-1

Tebbani, S., Titica, M., Caraman, S., Boillereaux, L., 2013. Estimation of *Chlamydomonas reinhardtii* Growth in a Torus Photobioreactor. IFAC Proc. Vol. 46, 72–77. https://doi.org/10.3182/20131216-3-IN-2044.00053

Torres-Tiji, Y., Fields, F.J., Mayfield, S.P., 2020. Microalgae as a future food source. Biotechnol. Adv. 41, 107536. https://doi.org/10.1016/J.BIOTECHADV.2020.107536

Tyagi, A.K., Gaur, T., 2003. Light Regulation of Nuclear Photosynthetic Genes in Higher Plants. CRC. Crit. Rev. Plant Sci. 22, 417–452. https://doi.org/10.1080/07352680390243503

Valenzuela, P., Medina, A., Rutter, W.J., Ammerer, G., Hall, B.D., 1982. Synthesis and assembly of hepatitis B virus surface antigen particles in yeast. Nature 298, 347–350. https://doi.org/10.1038/298347A0

Vázquez-Romero, B., Perales, J.A., Pereira, H., Barbosa, M., Ruiz, J., 2022. Techno-economic assessment of microalgae production, harvesting and drying for food, feed, cosmetics, and agriculture. Sci. Total Environ. 837, 155742. https://doi.org/10.1016/J.SCITOTENV.2022.155742

Vela Ramirez, J.E., Sharpe, L.A., Peppas, N.A., 2017. Current state and challenges in developing oral vaccines. Adv. Drug Deliv. Rev. 114, 116. https://doi.org/10.1016/J.ADDR.2017.04.008

Vesikari, T., Langley, J.M., Segall, N., Ward, B.J., Cooper, C., Poliquin, G., Smith, B., Gantt, S., McElhaney, J.E., Dionne, M., van Damme, P., Leroux-Roels, I., Leroux-Roels, G., Machluf, N., Spaans, J.N., Yassin-Rajkumar, B., Anderson, D.E., Popovic, V., Diaz-Mitoma, F., Vesikari, T., Langley, J.M., Gantt, S., Dionne, M., McElhaney, J., Rankin, B., Griffin, C., Turner, M., Kirstein, J., Rizzardi, B.E., Williams, H., Ahonen, A., Henriksson, O., Ukkonen, B., Paassilta, M., Segall, N., Ward, B.J., Cooper, C., Poliquin, G., Smith, B., Van Damme, P., Leroux-Roels, I., Leroux-Roels, G., Machluf, N., Spaans, J.N., Yassin-Rajkumar, B., Anderson, D.E., Popovic, V., Diaz-Mitoma, F., 2021. Immunogenicity and safety of a tri-antigenic versus a mono-antigenic hepatitis B vaccine in adults (PROTECT): a randomised, double-blind, phase 3 trial. Lancet Infect. Dis. 21,

1271–1281. https://doi.org/10.1016/S1473-3099(20)30780-5

Visioli, F., Artaria, C., 2017. Astaxanthin in cardiovascular health and disease: Mechanisms of action, therapeutic merits, and knowledge gaps. Food Funct. 8, 39–63. https://doi.org/10.1039/c6fo01721e

von Heijne, G., 1983. Patterns of Amino Acids near Signal-Sequence Cleavage Sites. Eur. J. Biochem. 133, 17–21. https://doi.org/10.1111/J.1432-1033.1983.TB07424.X

Wagner, G.P., Kin, K., Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. 131, 281–285. https://doi.org/10.1007/s12064-012-0162-3

Walker, M.R., Leung, P., Eltahla, A.A., Underwood, A., Abayasingam, A., Brasher, N.A., Li, H., Wu, B.R., Maher, L., Luciani, F., Lloyd, A.R., Bull, R.A., 2019. Clearance of hepatitis C virus is associated with early and potent but narrowly-directed, Envelope-specific antibodies. Sci. Rep. 9. https://doi.org/10.1038/S41598-019-49454-W

Walsh, G., Walsh, E., 2022. Biopharmaceutical benchmarks 2022. Nat. Biotechnol. 40, 1722. https://doi.org/10.1038/S41587-022-01582-X

Wang, C., Lan, C.Q., 2018. Effects of shear stress on microalgae – A review. Biotechnol. Adv. 36, 986–1002. https://doi.org/10.1016/j.biotechadv.2018.03.001

Wang, L., Yang, L., Wen, X., Chen, Z., Liang, Q., Li, J., Wang, W., 2019. Rapid and high efficiency transformation of *Chlamydomonas reinhardtii* by square-wave electroporation. Biosci. Rep. 39. https://doi.org/10.1042/BSR20181210

Wang, X., Brandsma, M., Tremblay, R., Maxwell, D., Jevnikar, A.M., Huner, N., Ma, S., 2008. A novel expression platform for the production of diabetes-associated autoantigen human glutamic acid decarboxylase (hGAD65). BMC Biotechnol. 8, 87. https://doi.org/10.1186/1472-6750-8-87

Wong, D.K., Stark, M.S., Rader, S.D., Fast, N.M., 2021. Characterization of Pre-mRNA Splicing and Spliceosomal Machinery in *Porphyridium purpureum* and Evolutionary Implications for Red Algae. J. Eukaryot. Microbiol. 68, e12844. https://doi.org/10.1111/JEU.12844

Xia, D., Qiu, W., Wang, X., Liu, J., 2021. Recent advancements and future perspectives of microalgae-derived pharmaceuticals. Mar. Drugs 19, 1–23. https://doi.org/10.3390/md19120703

Yamada, E.A., Sgarbieri, V.C., 2005. Yeast (*Saccharomyces cerevisiae*) Protein Concentrate:

Preparation, Chemical Composition, and Nutritional and Functional Properties. J. Agric. Food Chem. 53, 3931–3936. https://doi.org/10.1021/jf0400821

Yan, H., Zhong, G., Xu, G., He, W., Jing, Z., Gao, Z., Huang, Y., Qi, Y., Peng, B., Wang, H., Fu, L., Song, M., Chen, P., Gao, W., Ren, B., Sun, Y., Cai, T., Feng, X., Sui, J., Li, W., 2012. Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. Elife 1. https://doi.org/10.7554/eLife.00049

Yan, N., Fan, C., Chen, Y., Hu, Z., 2016. The Potential for Microalgae as Bioreactors to Produce Pharmaceuticals. Int. J. Mol. Sci. 2016, Vol. 17, Page 962 17, 962. https://doi.org/10.3390/IJMS17060962

Yang, S., Wan, H., Wang, R., Hao, D., 2019. Sulfated polysaccharides from *Phaeodactylum tricornutum*: Isolation, structural characteristics, and inhibiting HepG2 growth activity in vitro. PeerJ 2019. https://doi.org/10.7717/peerj.6409

You, X., Zhang, Xing, Cheng, J., Xiao, Y., Ma, J., Sun, S., Zhang, Xinzheng, Wang, H.W., Sui, S.F., 2023. In situ structure of the red algal phycobilisome–PSII–PSI–LHC megacomplex. Nature 616, 199–206. https://doi.org/10.1038/s41586-023-05831-0

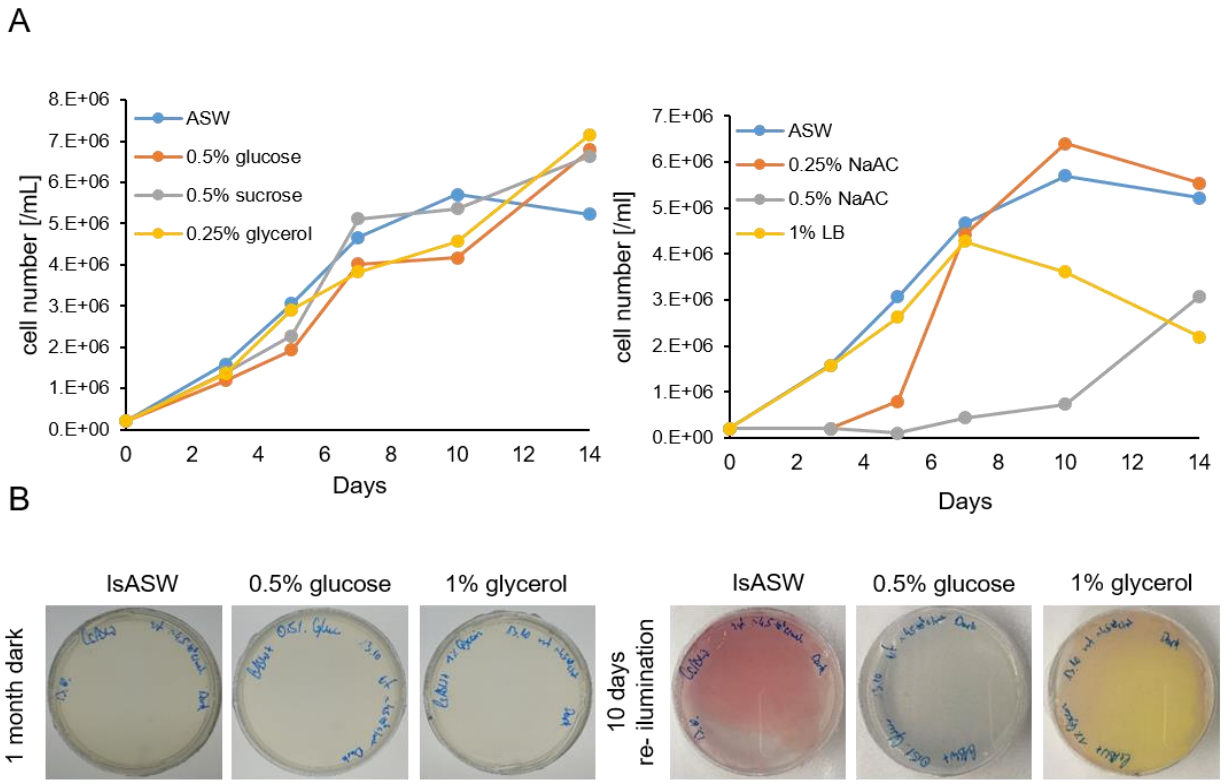
Young, E.B., Reed, L., Berges, J.A., 2022. Growth parameters and responses of green algae across a gradient of phototrophic, mixotrophic and heterotrophic conditions. PeerJ 10, e13776. https://doi.org/10.7717/peerj.13776

Zhang, G.G., Rodrigues, L., Rovinski, B., White, K.A., 2002. Production of HIV-1 p24 Protein in Transgenic Tobacco Plants. Mol. Biotechnol. 20, 131–136. https://doi.org/10.1385/MB:20:2:131

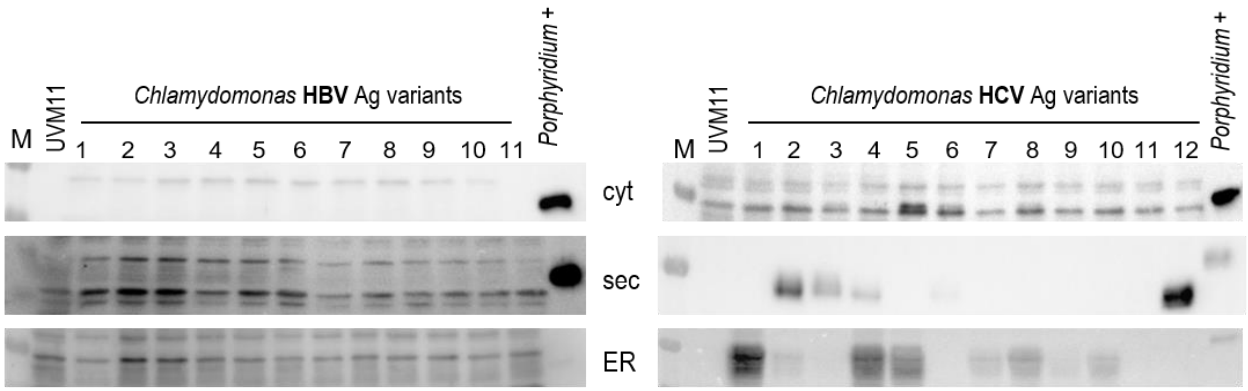
Zhang, R., Patena, W., Armbruster, U., Gang, S.S., Blum, S.R., Jonikas, M.C., 2014. High-Throughput Genotyping of Green Algal Mutants Reveals Random Distribution of Mutagenic Insertion Sites and Endonucleolytic Cleavage of Transforming DNA. Plant Cell 26, 1398–1409. https://doi.org/10.1105/tpc.114.124099

Zhao, L., Chen, F., Quitt, O., Festag, M., Ringelhan, M., Wisskirchen, K., Festag, J., Yakovleva, L., Sureau, C., Bohne, F., Aichler, M., Bruss, V., Shevtsov, M., van de Klundert, M., Momburg, F., Möhl, B.S., Protzer, U., 2021. Hepatitis B virus envelope proteins can serve as therapeutic targets embedded in the host cell plasma membrane. Cell. Microbiol. 23, e13399. https://doi.org/10.1111/CMI.13399

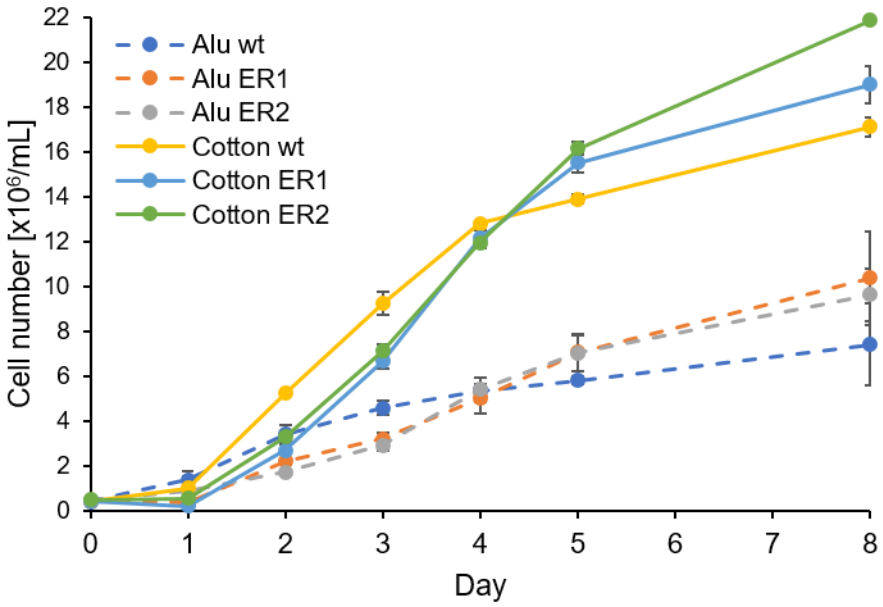
7 Supplements



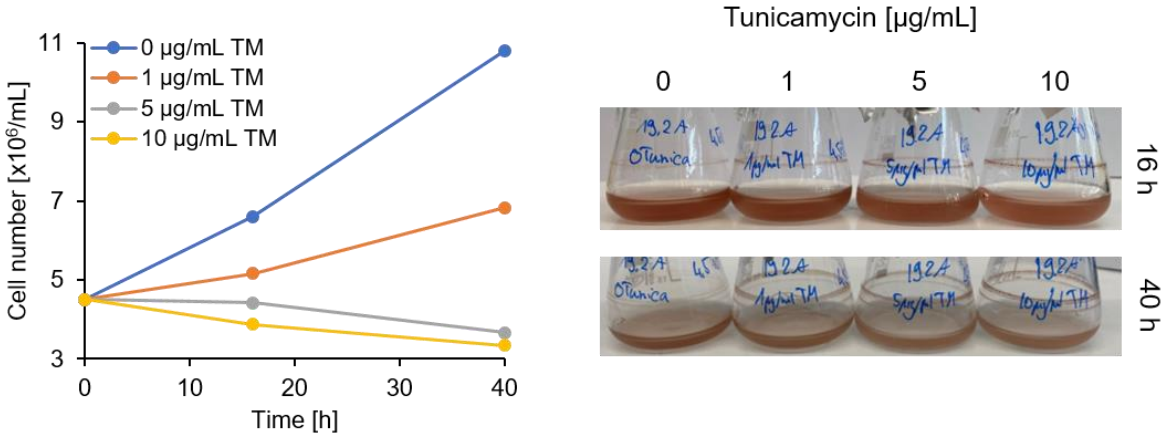
Suppl. Figure 1: Growth of *Porphyridium* in ASW medium supplemented with different carbon sources. **A**:Cells grown in medium supplemented with 0.5% glucose, 0.5% sucrose or 0.25% glycerol showed no improved growth compared to ASW-grown cells. Cells grown in 0.25% sodium acetate (NaAc) show a prolonged lag phase, and cells grown in 0.5% sodium acetate or 1% LB show a detrimental effect on growth. **B:** Cells grown on solid media supplemented with either 0.5% glucose or 1% glycerol in the dark for 1 month. After the dark incubation phase cells were re illuminated for 10 days and pictures were taken.



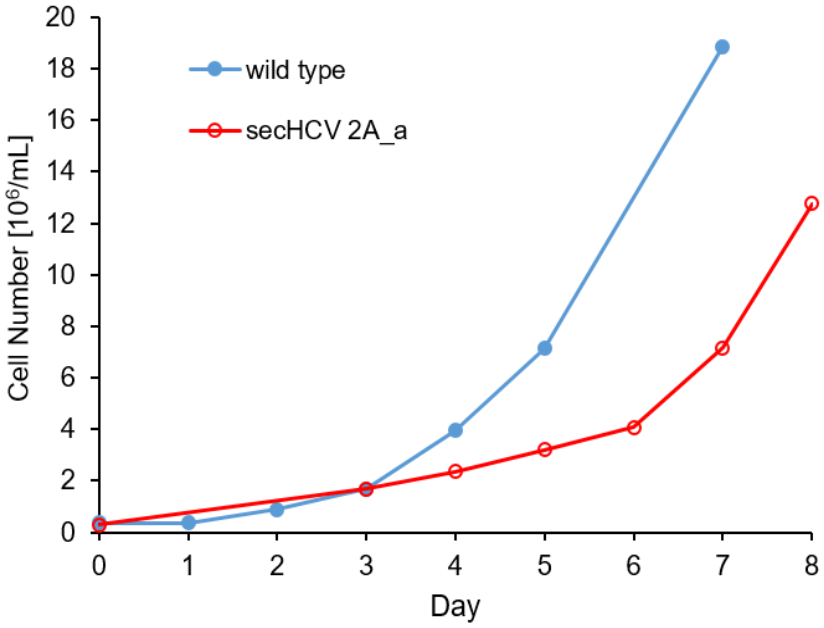
Suppl. Figure 2: Immunoblot analysis of HBV and HCV antigen variants expressed in *C. reinhardtii*. 40 µg total protein of the respective transformation lines were loaded on an SDS-PAA gel and, after blotting, the antigen was detected with the HA antibody. Total protein (3 µg) of transgenic *P. purpureum* was loaded as positive control in all blots (*Porphyridium* +). Protein derived from ER-HBV-expressing *P. purpureum* was likely degraded and does not show a positive signal. No accumulation of neither HBV variant was detectable in *C. reinhardtii*, whereas some accumulation for cytHCV was detected in *C. reinhardtii* transformed line 5 and 6. Background bands visible in the HBV-Ag blot are due to non-specific binding of the HA antibody to *Porphyridium*-derived proteins and long exposure time of over 10 minutes. The strongest antigen accumulation was detected in the secHCV (2 and 12) and ER-HCV (1 and 4) transformants. Those were used for comparison with *P. purpureum* HCV antigen expressing lines. UVM11: recipient strain control. M: molecular weight marker.



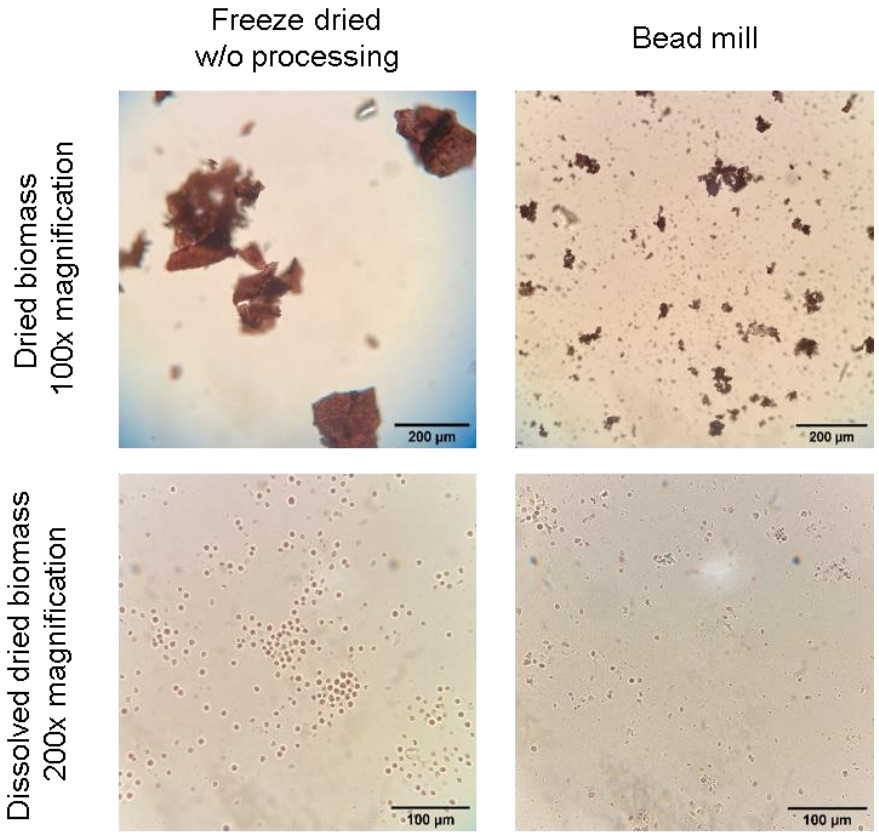
Suppl. Figure 3: Growth curve of the ER-HCV-Ag lines 1 and 2 (ER1 and ER2) and the wild type (wt). Cultures were inoculated in triplicates with a cell number of 3 x 10^5 cells/mL and grown for eight days under standard growth conditions. Flasks were sealed with aluminum foil (Alu) or cotton plugs (cotton) and growth was compared. Beginning with day two, protein samples were taken for immunoblot analysis.



Suppl. Figure 4: Determination of tunicamycin toxicity on *P. purpureum* cells
Different concentrations of tunicamycin (TM), ranging from 1 to 10 µg/mL were added to the *P. purpureum* culture, and growth was monitored for two days. Protein samples were taken after 16 h and 40 h. The addition of 1 µg/mL TM decreased growth rate and addition of higher concentration led to arrested growth or cell death. This is also visible in the right panel, where the culture flasks are depicted after 16 h and 40 h of TM addition. At 16 h, cultures with 5 or 10 µg/mL TM appear lighter than cultures with 0 and 1 µg/mL TM.



Suppl. Figure 5: Growth curve of the secHCV 2A_a expressing *P. purpureum* line grown in a PBR, compared to wild type (wt) *P. purpureum* grown in a PBR.
800 ml of lsASW medium were inoculated with 3 x 10^5 cells/mL and grown at 150 µmol photons $m^{-1}s^{-1}$ continuous light, 25°C and 2% CO_2. The transgenic algae culture shows a slower growth rate and longer lag time.



Suppl. Figure 6: Microscopy images of freeze-dried *P. purpureum* cells grown in a PBR.
The upper panel shows unprocessed cells directly after freeze-drying at 100 times magnification. The lower panel shows the same cells dissolved in lsASW medium at 200 times magnification. The right panel shows freeze-dried cells processed by bead milling.

Suppl. Table 1: Codon usage table of *P. purpureum* determined with 27,040 codons of the top 92 highly transcribed genes (omitting genes with introns).

AAcid	Codon	Count	/1000	Usage	AAcid	Codon	Count	/1000	Usage
Ala	GCG	1314	48.59	0.43	Leu	TTG	334	12.35	0.16
Ala	GCA	446	16.49	0.15	Leu	TTA	13	0.48	0.01
Ala	GCT	390	14.42	0.13	Leu	CTG	865	31.99	0.42
Ala	GCC	922	34.1	0.3	Leu	CTA	29	1.07	0.01
Arg	AGG	154	5.7	0.11	Leu	CTT	250	9.25	0.12
Arg	AGA	67	2.48	0.05	Leu	CTC	561	20.75	0.27
Arg	CGG	102	3.77	0.07	Lys	AAG	1284	47.49	0.89
Arg	CGA	142	5.25	0.1	Lys	AAA	163	6.03	0.11
Arg	CGT	201	7.43	0.15	Met	ATG	891	32.95	1
Arg	CGC	707	26.15	0.51	Phe	TTT	473	17.49	0.4
Asn	AAT	198	7.32	0.2	Phe	TTC	709	26.22	0.6
Asn	AAC	777	28.74	0.8	Pro	CCG	667	24.67	0.51
Asp	GAT	412	15.24	0.31	Pro	CCA	279	10.32	0.21
Asp	GAC	931	34.43	0.69	Pro	CCT	119	4.4	0.09
Cys	TGT	70	2.59	0.16	Pro	CCC	247	9.13	0.19
Cys	TGC	380	14.05	0.84	Ser	AGT	118	4.36	0.06
End	TGA	19	0.7	0.21	Ser	AGC	470	17.38	0.24
End	TAG	30	1.11	0.33	Ser	TCG	710	26.26	0.36
End	TAA	43	1.59	0.47	Ser	TCA	145	5.36	0.07
Gln	CAG	803	29.7	0.84	Ser	TCT	210	7.77	0.11
Gln	CAA	156	5.77	0.16	Ser	TCC	343	12.68	0.17
Glu	GAG	1196	44.23	0.77	Thr	ACG	501	18.53	0.38
Glu	GAA	365	13.5	0.23	Thr	ACA	181	6.69	0.14
Gly	GGG	273	10.1	0.12	Thr	ACT	176	6.51	0.13
Gly	GGA	454	16.79	0.2	Thr	ACC	465	17.2	0.35
Gly	GGT	440	16.27	0.19	Trp	TGG	312	11.54	1
Gly	GGC	1120	41.42	0.49	Tyr	TAT	174	6.43	0.22
His	CAT	128	4.73	0.27	Tyr	TAC	601	22.23	0.78
His	CAC	344	12.72	0.73	Val	GTG	1145	42.34	0.55
Ile	ATA	16	0.59	0.01	Val	GTA	114	4.22	0.05
Ile	ATT	440	16.27	0.41	Val	GTT	259	9.58	0.12
Ile	ATC	612	22.63	0.57	Val	GTC	580	21.45	0.28

AmAcid: Amino acid; /1000: Frequency per thousand codons; Usage: usage of this codon to encode the respective amino acid

Suppl. Table 2: Top 100 highly expressed endogenous genes in *Porphyridium purpureum*. Genes in italics contain an intron and were not considered for codon usage determination.

Locus tag	Gene product	TPM
FVE85_0107	Formate/Nitrite transporter YrhG	14407.33
FVE85_3577	hypothetical protein (PsbQ domain)	14183.62
FVE85_6364	Protein PYP1	13274.17
FVE85_3079	Carbonic anhydrase	10894.83
FVE85_6371	hypothetical protein FVE85_6371	10333.3
FVE85_6435	Chlorophyll a-b binding protein, chloroplastic	9631.395
FVE85_0111	D-glycerate 3-kinase, chloroplastic	8261.663
FVE85_6090	hypothetical protein FVE85_6090	7928.787
FVE85_2042	R-phycoerythrin gamma chain, chloroplastic	7258.901
FVE85_2041	R-phycoerythrin gamma chain, chloroplastic	6801.489
FVE85_7927	Chlorophyll a-b binding protein 1B-21, chloroplastic	6426.702
FVE85_5651	Chlorophyll a-b binding protein of LHCII type III, chloroplastic	6018.439
FVE85_2505	Fucoxanthin-chlorophyll a-c binding protein, chloroplastic	5984.899
FVE85_0338	Chlorophyll a-b binding protein, chloroplastic	5965.422
FVE85_7382	Chlorophyll a-b binding protein 1B-21, chloroplastic	5828.118
FVE85_3955	Chlorophyll a-b binding protein 1B-21, chloroplastic	5557.657
FVE85_3924	hypothetical protein FVE85_3924	5276.498
FVE85_8023	Phycobilisome 31.8 kDa linker polypeptide, phycoerythrin-associated, rod	5070.958
FVE85_2781	Photosystem I subunit O	4762.725
FVE85_8322	Photosystem II 12 kDa extrinsic protein, chloroplastic	4672.513
FVE85_6079	hypothetical protein FVE85_6079	4565.703
FVE85_6536	hypothetical protein FVE85_6536	4481.354
FVE85_1413	hypothetical protein FVE85_1413	4469.18
FVE85_9678	Uncharacterized protein FVE85_9678	4457.982
FVE85_7666	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	4414.724
FVE85_0106	Carbonic anhydrase 2	4311.254
FVE85_9240	R-phycoerythrin gamma chain, chloroplastic	3977.435
FVE85_2942	hypothetical protein FVE85_2942	3897.133
FVE85_8746	*hypothetical protein FVE85_8746*	*3749.734*
FVE85_1414	hypothetical protein FVE85_1414	3667.125
FVE85_7594	hypothetical protein FVE85_7594	3582.766
FVE85_2309	R-phycoerythrin gamma chain, chloroplastic	3562.359
FVE85_2293	hypothetical protein FVE85_2293	3546.824
FVE85_5340	hypothetical protein FVE85_5340	3426.4
FVE85_5355	Oxygen-evolving enhancer protein 1, chloroplastic	3295.069
FVE85_8316	Cytochrome b6-f complex iron-sulfur subunit 1, cyanelle	3231.383
FVE85_0135	Cold shock domain-containing protein 4	3201.967
FVE85_4028	*Cofilin*	*2917.666*
FVE85_6962	*hypothetical protein FVE85_6962*	*2899.28*
FVE85_6702	Superoxide dismutase Mn, mitochondrial	2844.892
FVE85_6005	hypothetical protein FVE85_6005	2765.933

Locus tag	Gene product	TPM
FVE85_7812	Methanesulfonate monooxygenase ferredoxin subunit	2664.067
FVE85_8365	Phycobilisome 31.8 kDa linker polypeptide, phycoerythrin-associated, rod	2554.758
FVE85_5468	Ammonium transporter 1 member 2	2546.475
FVE85_5305	hypothetical protein FVE85_5305	2535.308
FVE85_6170	Zinc transporter ZupT	2295.755
FVE85_2375	Ferrous iron permease EfeU	2250.688
FVE85_1592	hypothetical protein FVE85_1592	2242.989
FVE85_8359	hypothetical protein FVE85_8359	2175.146
FVE85_2640	hypothetical protein FVE85_2640	2105.432
FVE85_2854	*hypothetical protein FVE85_2854*	*2027.673*
FVE85_5140	Linker RC5	2024.174
FVE85_2471	hypothetical protein FVE85_2471	2021.889
FVE85_1715	hypothetical protein FVE85_1715	1982.415
FVE85_6254	*SNF1-related protein kinase regulatory subunit beta-2*	*1965.711*
FVE85_7087	hypothetical protein FVE85_7087	1955.946
FVE85_5222	hypothetical protein FVE85_5222	1940.605
FVE85_2240	Aspartate--ammonia ligase	1936.67
FVE85_6016	hypothetical protein FVE85_6016	1920.607
FVE85_2199	*hypothetical protein FVE85_2199*	*1909.823*
FVE85_7221	Phycobilisome 31.8 kDa linker polypeptide, phycoerythrin-associated, rod	1865.696
FVE85_2982	hypothetical protein FVE85_2982	1863.984
FVE85_0729	Phycobilisome 27.9 kDa linker polypeptide, phycoerythrin-associated, rod	1863.821
FVE85_3889	hypothetical protein FVE85_3889	1860.406
FVE85_2465	hypothetical protein FVE85_2465	1842.204
FVE85_2818	Phycobilisome 7.8 kDa linker polypeptide, allophycocyanin-associated, core	1812.344
FVE85_4479	hypothetical protein FVE85_4479	1809.635
FVE85_5869	6-phosphogluconolactonase	1807.669
FVE85_0134	Cold shock domain-containing protein 4	1779.967
FVE85_3229	Elongation factor 1-alpha	1731.754
FVE85_2468	hypothetical protein FVE85_2468	1728.796
FVE85_1011	Glycine-rich RNA-binding protein	1723.927
FVE85_5793	hypothetical protein FVE85_5793	1718.107
FVE85_9475	hypothetical protein FVE85_9475	1701.586
FVE85_6320	Peroxisomal membrane protein 2	1688.202
FVE85_4247	Linker RC4	1675.684
FVE85_8317	Protein sym-1	1669.745
FVE85_6037	Zinc transporter ZIP1	1567.177
FVE85_3578	hypothetical protein FVE85_3578	1560.506
FVE85_7258	Linker RC6	1542.948
FVE85_2314	Phycobilisome 32.1 kDa linker polypeptide, phycocyanin-associated, rod	1540.393
FVE85_8246	hypothetical protein FVE85_8246	1517.096
FVE85_1265	Fructose-bisphosphate aldolase 1, chloroplastic	1513.443
FVE85_1484	hypothetical protein FVE85_1484	1507.37

Locus tag	Gene product	TPM
FVE85_8240	Axial regulator YABBY 5	1440.153
FVE85_6163	hypothetical protein FVE85_6163	1423.781
FVE85_3879	Phycobilisome 31.8 kDa linker polypeptide, phycoerythrin-associated, rod	1415.86
FVE85_4096	14-3-3 protein epsilon	1406.856
FVE85_2264	Bifunctional enolase 2/transcriptional activator	1406.089
FVE85_1945	Ferredoxin--NADP reductase, cyanelle	1400.638
FVE85_8260	Peroxiredoxin-6	1394.547
FVE85_0385	Carbonic anhydrase 2	1393.334
FVE85_7126	hypothetical protein FVE85_7126	1389.001
FVE85_7722	*Calvin cycle protein CP12, chloroplastic*	*1383.908*
FVE85_7633	hypothetical protein FVE85_7633	1381.935
FVE85_0816	Phycobilisome 27.9 kDa linker polypeptide, phycoerythrin-associated, rod	1369.163
FVE85_7446	Fructose-bisphosphate aldolase	1353.193
FVE85_1309	Profilin-2	1323.515
FVE85_8054	hypothetical protein FVE85_8054	1322.545
FVE85_5088	Peptidyl-prolyl cis-trans isomerase	1303.259

Suppl. Table 3: Number of cells analyzed by fluorescence microscopy to determine the percentage of fluorescent cells in a transgenic population.

		Progenitor cell lines							
		vYFP1	**vYFP2**	**laYFP1**	**laYFP2**	**CrYFP1**	**CrYFP2**	**PYP1 1**	**PYP1 2**
	n=	409	691	720	733	767	657	665	471
		Sub-cultivation round 1							
	Subline	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Progenitor line	**vYFP1**	365	419	496	167	181	387	405	107
	vYFP2	-	275	-	405	118	252	188	184
	laYFP1	257	97	308	223	522	444	262	371
	laYFP2	549	608	445	495	147	283	396	275
	CrYFP1	668	170	352	297	511	271	337	264
	CrYFP2	410	223	237	289	282	423	317	319
	PYP1 1	108	-	647	-	162	93	148	-
	PYP1 2	131	182	120	168	246	307	117	238

		Sub-cultivation round 2							
	Subline	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Progenitor line	**vYFP1_6**	429	344	192	138	317	99	286	155
	vYFP2_6	508	281	552	433	623	605	473	904
	laYFP1_6	251	267	205	177	193	113	482	161
	laYFP2_2	255	211	71	198	144	161	324	70
	CrYFP1_5	1007	595	563	883	1153	122	915	913
	CrYFP2_8	308	807	625	557	797	159	610	716
	PYP1 1_3	174	269	924	248	600	60	151	104
	PYP1 2_3	418	358	452	215	210	187	254	422

8 Acknowledgments

I would like to thank all people that made this work possible, either by being directly involved in it, by being supportive, or just by their emotional support.

First, I'd like to thank Ralph Bock for the opportunity work on such an interesting topic. His guidance as my professor and supervisor and his ideas and scientific advice have been instrumental in shaping this research.

I owe a great debt of gratitude to Juliane Neupert, my direct supervisor in the lab, for her unwavering support, guidance, and for imparting her knowledge on numerous methods that were crucial for this work. Also, thank you for all your time in reading and correcting my reports, presentations and the thesis!

I am grateful to my PCA, Dr. Andre Scheffel, and Prof. Dr. Müller-Röber for their insightful advice and guidance on future research directions. Also, I am thankful for all the administrative support from Ina Talke. Without her, research would fall short to bureaucracy.

I extend my sincere thanks to my collaboration partner, Norica Nichita from Romania, whose support and discussions on topics related to HBV and HCV were invaluable. Without you, the project would have gone in a completely different direction, and would for sure not have been that exciting.

I want to acknowledge and appreciate the entire team of AG Bock for their openness in addressing my questions and engaging in fruitful discussions. Their willingness to share their lab techniques was truly appreciated.

A special mention goes to my colleagues who became friends during this journey, the BeerPIMPs group, including Thekla von Biskmarck, Zac Taylor, David Scharf, Jonathan Huc, Vittoria Clappero, and many others. Your friendship enriched this experience. Without the many bar/club night with you, the Work-Life balance would be pretty imbalanced… !

A very big thanks goes to Vanessa, the crazy person I very quickly became friends with during numerous cigarettes and beers. Without you, the MPIMP would be a pretty boring place.

Beyond the MPI, there a lots and lots of people I would like to thank for their support in the difficult last years. Especially Jonas and Laura, that were the sweetest new flatmates one could wish for and who introduced me to all the wonderful people I know now. Without you, moving to Berlin would have been much much harder.

And of course, I would like to thank my partner Felix. During the challenging and stressful time between the pandemics, the war and writing, you were always there for me and supported me, no matter what.

This work is a testament to the collective effort and support of these individuals. Thank you for being a part of this journey.