Published in final edited form as:
Plant Biotechnol J. 2010 August ; 8(6): 719–733. doi:10.1111/j.1467-7652.2010.00503.x.

Production of therapeutic proteins in algae, analysis of expression of seven human proteins in the chloroplast of *Chlamydomonas reinhardtii*

Beth A Rasala[1,*], **Machiko Muto**[1,*], **Philip A Lee**[1], **Michal Jager**[1], **Rosa MF Cardoso**[2], **Craig A Behnke**[2], **Peter Kirk**[3], **Craig A Hokanson**[3], **Roberto Crea**[3], **Michael Mendez**[2,◆], and **Stephen P Mayfield**[1,§,◆]

[1] Department of Cell Biology and the Skaggs Institute for Chemical Biology, 10550 N. Torrey Pines Rd, La Jolla, CA92037, USA

[2] Sapphire Energy, 3115 Merryfield Rd., San Diego, CA 92121

[3] Protelica, 26118 Research Pl, Hayward, CA 94545

Summary

Recombinant proteins are widely used today in many industries, including the biopharmaceutical industry, and can be expressed in bacteria, yeasts, mammalian and insect cell cultures, or in transgenic plants and animals. In addition, transgenic algae have also been shown to support recombinant protein expression, both from the nuclear and chloroplast genomes. However, to date, there are only a few reports on recombinant proteins expressed in the algal chloroplast. It is unclear if this is due to few attempts or to limitations of the system that preclude expression of many proteins. Thus, we sought to assess the versatility of transgenic algae as a recombinant protein production platform. To do this, we tested whether the algal chloroplast could support the expression of a diverse set of current or potential human therapeutic proteins. Of the seven proteins chosen, greater than 50% expressed at levels sufficient for commercial production. Three expressed at 2% to 3% of total soluble protein, while a forth protein accumulated to similar levels when translationally fused to a well-expressed serum amyloid protein. All of the algal chloroplast-expressed proteins are soluble and showed biological activity comparable to that of the same proteins expressed using traditional production platforms. Thus, the success rate, expression levels, and bioactivty achieved demonstrate the utility of *C. reinhardtii* as a robust platform for human therapeutic protein production.

Introduction

Since the FDA approval of recombinant insulin over 25 years ago, the class of protein-based therapeutics has grown quickly. The majority of therapeutic proteins produced today are made in bacteria (*E. coli*), yeast (*S. cerevisiae*) or mammalian cell culture (Chinese hamster ovary cells, CHO) (Demain and Vaishnav, 2009; Walsh, 2003; Walsh, 2006). Other

§Corresponding author: SPM: smayfield@ucsd.edu.
*These authors contributed equally to this work
◆The work presented here was a collaborative effort between the laboratories of Michael Mendez at Sapphire Energy and Stephen Mayfield at the Scripps Research Institute

Authors' contributions

BR, PL and SPM prepared the manuscript. BR, PL, M Muto, MJ, RC, CB, PK, and CH developed reagents for and/or designed and executed the experiments in this work. SPM, M Mendez and R Crea conceived the study and participated in the study's design. All authors read and approved the manuscript.

production systems under development include the yeast *P. pastoris*, insect cell culture, and transgenic animals and plants.

In general, transgenic plants offer several advantages over other recombinant protein production platforms. The cost of protein production in plants is much lower than other production systems due to low cost of goods and capital expenses (Dove, 2002). Proteins purified from plants should be free from toxins and viral agents that may be present in preparations from bacteria or mammalian cell culture. Finally the ability to rapidly scale production in plants is difficult to achieve in other systems. Transgenic plants have been engineered to express recombinant genes from both the nuclear and plastid (chloroplast) genomes. Nuclear expression of transgenes enables regulated and tissue-specific expression, as well as post-translational modifications. However, nuclear expression has several drawbacks to protein production; for example transgene silencing, lower yields, and the potential risk of gene flow to surrounding food crops and other native plants (Daniell, 2006). Alternatively, plastid genomes have been successfully engineered to express recombinant proteins. Advantages of chloroplast bioreactors include absence of gene silencing, targeted transgene integration by homologous recombination, expression of multiple genes from polycistrons, transgene containment via maternal inheritance of the chloroplast genome, and robust expression (Bock, 2007; Chebolu and Daniell, 2009; Daniell, 2006). The chloroplast of higher plants has been shown to accumulate therapeutic proteins to 6–16% total soluble protein (TSP), vaccine antigens to as high as 31% TSP, and antimicrobial peptides to greater than 70% TSP (Chebolu and Daniell, 2009; Daniell, 2006; Oey et al., 2009). However, the possibility of transgene escape to surrounding food crops and native plants remains, although it is greatly reduced compared to nuclear-transformed plants. While the plastid genomes in most species are maternally inherited (Hagemann, 2004), several recent reports have demonstrated that transfer of the paternal plastid genome to pollen does occur at a low, but measurable frequency. For example, paternal inheritance has been estimated at 0.03%-0.0002% in *Setaria italica* (foxtail) (Shi et al., 2008; Wang et al., 2004), 0.01–0.00029% in tobacco (Ruf et al., 2007; Svab and Maliga, 2007) and 0.0039% in *Arabidopsis thaliana* (Azhagiri and Maliga, 2007).

Interest in eukaryotic microalgae as an alternative platform for recombinant protein production has been gaining in recent years. Protein production in transgenic algae could presumably offer many of the same advantages as transgenic plants, including cost, safety, and rapid scalability. addition, microalgae can be grown in containment in enclosed bioreactors (Pulz, 2001), thus reducing the possibility of gene flow. Expression of recombinant proteins in the chloroplast of the green algae *Chlamydomonas reinhardtii* is now well established (Mayfield et al., 2007). These include reporter proteins (Franklin et al., 2002; Mayfield and Schultz, 2004; Muto et al., 2009), a large complex mammalian single chain antibody (Mayfield et al., 2003), more traditional single chain antibodies (Franklin and Mayfield, 2005), a full length monoclonal antibody (Tran et al., 2009) and potential vaccine antigens (Surzycki et al., 2009). Thus far, the *psbA* promoter and untranslated regions (UTRs) has been shown to support the highest levels of recombinant protein accumulation in *C. reinhardtii*, but only in *psbA* deficient strains (Manuell et al., 2007; Surzycki et al., 2009). Indeed, VP28 protein of the White Spot Syndrome Virus accumulated to levels as high as 20.9% total cell protein (TCP) when placed under the control of the *psbA* promoter and UTRs in a *psbA* deficient strain (Surzycki et al., 2009). However, because the *psbA* gene product D1 of photosystem II is required for photosynthesis, these transgenic algae are non-photosynthetic. Recently, we achieved expression of a mammalian serum amyloid protein (M-SAA) in *C. reinhardtii* chloroplast to about 10% of TSP by using the *psbA* promoter and UTRs in a targeted *psbA* knockout strain (Manuell et al., 2007). Importantly, when the *psbA* gene was reintroduced elsewhere in the genome under the control of the *psbD* promoter, photosynthesis was restored while M-SAA protein levels were only slightly reduced,

showing that photosynthetic competent algae can produce high levels of recombinant proteins (Manuell et al., 2007). Furthermore, the purified protein was found to have bioactivity similar to the authentic protein, demonstrating the potential of the system as a robust platform for the production of recombinant proteins.

To examine the versatility of algal chloroplasts for the production of human protein therapeutics, we challenged the system with seven different recombinant genes, all encoding current or potential human protein therapeutics. Using three different expression vectors we were able to achieve production of four of the seven genes tested. Three proteins accumulated to above 2% of total soluble protein, levels sufficient for easy purification, when the genes were driven from the *psbA* promoter in a *psbA* deficient strain. The *atpA* promoter also drove expression of the same three proteins, but to significantly lower levels. A carboxy-terminal fusion of the seven therapeutic proteins to the M-SAA protein resulted in the accumulation of the same three proteins that expressed with the *psbA* promoter alone, as well as an additional recombinant protein that did not express on its own. Two of the proteins were purified and assayed for bioactivity using standard assays, and both were found to have similar activity to the same protein produced in a more traditional expression system. Together, the data demonstrate that the algal chloroplast is a viable platform for the expression of a diverse set of recombinant human therapeutic proteins.

Results

Recombinant genes used in this study

The proteins chosen for this study are a diverse group of proteins, some of which are already produced as therapeutics, and others that have the potential to become therapeutic proteins in the future (Table 1). All would benefit from being able to be produced in large quantities and at low cost. The first protein is human erythropoietin (EPO) without its signal peptide, a human hormone produced by both the liver and kidney that regulates red blood cell production and also plays an important role in the response of the brain to neural injury and wound healing (Haroon et al., 2003;Siren et al., 2001). Recombinant EPO is currently produced in mammalian cells and is used in the treatment of anemia (Eschbach et al., 1989;Jelkmann, 2007). The second and third proteins are domains ten and fourteen of human fibronectin, respectively. Fibronectin is an extracellular matrix glycoprotein that functions in cell adhesion, migration, growth and differentiation (Pankov and Yamada, 2002). Fibronectin is comprised of multiple domains and can bind to integrins as well as collagen, fibrin and heparan sulfate proteoglycans. The tenth human fibronectin type III domain (10FN3) is a stable 10 kDa beta-sandwich subunit that has potential to be an antibody mimic (monobody) (Garcia-Ibilcieta et al., 2008;Koide and Koide, 2007). The fourteenth human fibronectin type III domain (14FN3) is part of the heparin-II/VEGF binding domain (Wijelath et al., 2006) and is in development as a framework for antibody mimics. The fourth protein is human interferon β1. Interferons improve the integrity of the blood brain barrier and are used in the treatment of Multiple Sclerosis (MS) (Murdoch and Lyseng-Williamson, 2005). A one month supply of interferon β, Avonex (Biogen Idec) or Rebif (EMD Serono and Pfizer), can cost anywhere from $1,600 to more than $2,000 USD (McCormack and Scott, 2004). The fifth protein used in this study is human proinsulin (without its signal peptide), a hormone that regulates blood sugar level. Insulin is used in the treatment of type I diabetes, has a multi-billion dollar market dominated by Eli Lilly (e.g. Humulin) and Novo Nordisk, and was the first genetically engineered drug approved by the FDA. The sixth protein is human vascular endothelial growth factor (VEGF) isoform 121 (without its signal peptide). Patients suffering from pulmonary emphysema have decreased levels of VEGF in their pulmonary arteries. VEGF also has the potential to be a treatment for erectile dysfunction (Strong et al., 2008) and depression (Warner-Schmidt and Duman, 2008). The seventh and final protein is high mobility group protein B1 (HMGB1), which

mediates a number of important functions involved in wound healing including endothelial cell activation, stromagenesis, recruitment and activation of innate immune cells, and dendritic cell maturation (Sun and Chao, 2005). It has also been suggested that HMGB1 has the potential to enhance the effectiveness of some anti-cancer therapies if co-administered (Dong Xda et al., 2007;Krynetskaia et al., 2008).

Introduction of recombinant genes into the *C. reinhardtii* chloroplast genome

The *C. reinhardtii* chloroplast genome shows a high AT content and noted codon bias (Franklin et al., 2002; Mayfield and Schultz, 2004). To achieve protein expression, the genes of interest were first converted to match the codon usage of *C. reinhardtii* by synthesizing each of the seven genes in a codon-bias optimized for the *C. reinhardtii* chloroplast (Table 1). A codon bias threshold of greater than 10% of codons normally used for that amino acid was chosen and the genes were assembled via overlapping oligonucleotides. A FLAG-tag epitope was added to the C-terminal end of each protein sequence to allow detection by western blot and to facilitate protein purification.

A range of endogenous promoters and UTRs were previously examined for recombinant protein expression in wild-type *C. reinhardtii* 137c, with the *atpA* and *psbD* promoters and UTRs showing the best expression (Barnes et al., 2005). More recent work has shown that expression from the *psbA* promoter has very good potential when the endogenous *psbA* gene product, D1, is absent (Manuell et al., 2007; Surzycki et al., 2009), probably due to the interruption of a negative feedback loop (Minai et al., 2006). Expression from this promoter is also increased by an increase in light intensity (Manuell et al., 2007). Thus, both the *atpA* and the *psbA* promoters were chosen for the expression analysis of the recombinant genes described above. For *psbA* expression the genes were cloned into the transformation vector pD1-Kan under the control of the *psbA* promoter and 5′ and 3′ UTRs (Figure 1A). The pD1-Kan vector also contains a kanamycin resistance gene (*aphA6*) under the control of the *atpA* promoter and 5′UTR and the *rbcL* 3′UTR, which is cloned downstream of the *psbA* expression site, and is used for the selection of transformants (Figure 1). This expression cassette contains homology to the *psbA* region of the *C. reinhardtii* chloroplast genome and thus after transformation will replace the *psbA* locus (and gene) by homologous recombination (Manuell et al., 2007). The resulting transformants are resistant to the antibiotic kanamycin and are *psbA* deficient.

Expression of the seven genes was also tested using the *atpA* promoter and 5′ UTR and the *rbcL* 3′ UTR (Figure 1C). The genes were cloned into the p322 plasmid and therefore integrated into a silent site in the inverted repeat just downstream of the *psbA* locus (Franklin et al., 2002). These constructs were co-transformed with the p228 plasmid conferring resistance to spectinomycin (Franklin et al., 2002). All of the genes were also cloned into the *psbA*::SAA fusion plasmid, which was designed to fuse the protein of interest to the carboxy terminus of the well expressed mammalian protein M-SAA (Manuell et al., 2007). A protease cleavage site (Thrombin) between SAA and the protein of interest was engineered so that SAA could be removed during downstream processing (Figure 1B). As in the pD1-Kan vector, the SAA-fusion constructs are under the control of the *psbA* promoter and UTRs, replace the endogenous *psbA* locus, and contain the *atpA::aphA6* kanamycin resistance gene for selection.

All constructs were transformed by particle bombardment into *C. reinhardtii* wild type strain 137c (mt+). Primary transformants were selected on media containing either kanamycin (for pD1-Kan) or spectinomycin (p228) and screened for integration and homoplasmicity by PCR (Figure 2). Each of the seven genes, in all three constructs, was stably integrated into the chloroplast genome (G, Figure 2). Homoplasmic cell lines, in which all copies of the chloroplast genome contained the recombinant gene, were isolated through multiple rounds

of streaking for single colonies under antibiotic resistance selection. Colony PCR screening was used to confirm that strains were homoplasmic for the correct gene integration (H–I, Figure 2). Efficiency of transformation (number of gene positive colonies/number of colonies) with the construct containing the kanamycin cassette in *cis* with the gene of interest was much greater than that seen with co-transformation protocols used previously (data not shown).

Accumulation of recombinant proteins in transgenic chloroplast

Six homoplasmic cell lines were chosen at random for each of the recombinant genes and screened for recombinant protein expression to determine the variation in protein expression levels of primary transformants. Using anti-FLAG western blotting, we determined that all isolated VEGF and 14FN3 transgenic lines, and 4 out of the 6 HMGB1 lines, accumulated their respective recombinant protein (Supplemental Figure 1). One transgenic line for each protein was characterized in detail, and shown in Figure 3. Proteins 14FN3, VEGF, and HMGB1 show significant expression when the corresponding gene is expressed from the *psbA* promoter (Figure 3A). Moreover, all three proteins were soluble. Expression from this promoter was also induced by a shift from dark or dim light into bright light (data not shown), as has previously been reported for other recombinant genes expressed from the *psbA* promoter (Barnes et al., 2005; Manuell et al., 2007). Native VEGF is active as a dimer (Potgens et al., 1994) and even under the denaturing conditions of SDS-PAGE *Chlamydomonas* expressed VEGF appears to show dimerization (Figures 3 and 6, and Supplemental Figure 1), suggesting proper protein folding. 14FN3, VEGF, and HMGB1 accumulated to approximately 3%, 2% and 2.5%, respectively, of total soluble protein when expressed using the *psbA* promoter (Figure 4). This represents a level of expression high enough to allow for relatively easy purification of the proteins. As we have shown previously, when the *psbA* gene under the control of the *psbD* promoter was reintroduced by particle bombardment into the 3HB silent site (Manuell et al., 2007), protein levels were only slightly reduced while photosynthesis was completely restored (Supplemental Figure 2). In this way, recombinant protein expression is maintained under photosynthetic growth conditions.

To test whether the remaining proteins were expressed at levels below that detectable by western blotting of total soluble protein, we performed immunoprecipitations from 50 ml liquid cultures using anti-FLAG chromatography resin (Sigma). Using this technique we were able to observe low levels of expression of 10FN3 and proinsulin, but no detectable protein accumulation for interferon β or EPO (data not shown).

When each of the genes was expressed from the *atpA* promoter, the same three proteins (14FN3, VEGF, and HMGB1) accumulated, however to significantly lower levels than when expressed from the *psbA* promoter (Figure 3B). Under the *atpA* promoter, 14FN3, VEGF, and HMGB1 accumulated to approximately 0.15%, 0.1% and 1% of total soluble protein (data not shown). Thus, both promoters support expression of the same three proteins, but the *psbA* promoter and UTRs drives recombinant protein accumulation up to twenty times greater than that from the *atpA* promoter and 5′ UTR.

Fusion to M-SAA protein increases accumulation of fibronectin domain 10

One possible explanation for the lack of protein accumulation of 10FN3, proinsulin, interferon β, and EPO is protein instability. Fusion of poorly expressed proteins to a well-expressed and stable protein has been shown to increase the accumulation of the former in many expression systems, including bacteria (Butt et al., 2005; De Marco et al., 2004; Pryor and Leiting, 1997; Sachdev and Chirgwin, 2000; Wang et al., 1999) and plant and algal chloroplasts (Muto et al., 2009; Streatfield, 2007). Indeed, expression of human proinsulin in

E. coli and yeast was facilitated by the construction of fusion proteins (Chan et al., 1981; Stepien et al., 1983). Human proinsulin fusions have been expressed from the plant nuclear genomes of potato tubers at 0.1% TSP (Arakawa et al., 1998) and *Arabidopsis* at 0.1% total seed protein (Nykiforuk et al., 2006), and in the chloroplast of tobacco and lettuce at ~16% and ~2.5% TSP, respectively (Ruhlman et al., 2007). To test whether this same effect could work in algal chloroplasts, each of the recombinant genes was cloned as a fusion partner to the gene encoding the mammary-associated serum amyloid protein (M-SAA). We have previously shown that M-SAA can accumulate to about 10% of total soluble protein when driven by the *psbA* promoter and UTRs (Manuell et al., 2007). Therefore, M-SAA fusion constructs were placed under the control of the *psbA* promoter and UTRs (Figure 1B) and transformed as above, selecting for transformants by kanamycin resistance. Western blots of total soluble protein revealed that fusions of 14FN3, VEGF and HMGB1 to M-SAA led to significant protein accumulation (Figure 3C); the same three proteins that accumulated using the *psbA* and *atpA* promoters alone. In addition, the fusion of fibronectin 10FN3 to M-SAA enabled significant protein accumulation to occur, and expression levels similar to those achieved with the three other proteins was observed (Figure 3C). Interestingly, the expression of HMGB1, which was substantial without fusion to M-SAA, actually showed decreased accumulation when fused to the SAA protein, from 2.5% to approximately 1% of total soluble protein (data not shown).

Accumulation of recombinant mRNAs from different promoters

The regulation of endogenous chloroplast gene expression primarily occurs at the level of translation (Zerges, 2000). To address the relationship between translation and transcription in recombinant gene expression, reverse transcriptase quantitative PCR (RT-qPCR) was used to determine the level of recombinant mRNAs for each of the seven genes under the control of the *psbA* and *atpA* promoters (Figure 5). Total RNA was isolated from liquid cultures grown in 1000 lux of light illumination (Supplemental Figure 3). Following cDNA synthesis, SYBR green based qPCR was performed using gene-specific primers. qPCR signals were detectable for all 14 constructs, indicating that stable mRNA transcripts accumulated for all of the recombinant genes (Figure 5). In general, the *psbA* promoter yielded higher levels of mRNA transcript than the *atpA* promoter. Interestingly, while 14FN3, VEGF, and HMGB1 protein accumulated to approximately equal levels (3%, 2% and 2.5% respectively of total soluble protein; Figure 4), the *psbA-HMGB1* mRNA transcript was found to be approximately 75-fold less abundant than *psbA-14FN3* and *psbA-VEGF* mRNA transcripts (Figure 5). Overall there was a poor correlation between mRNA accumulation and protein accumulation, as has been reported for endogenous chloroplast genes (Eberhard et al., 2002; Nickelsen, 2003; Zerges, 2000). However in this study, in no case was the lack of protein accumulation caused by a lack of mRNA accumulation.

Purification of bio-active recombinant proteins from *Chlamydomonas*

To become a viable protein production platform, algal chloroplasts must not only express recombinant proteins, but those proteins must be biologically active in a highly purified state. 14FN3, VEGF, and HMGB1 were affinity purified using FLAG affinity chromatography to the C-terminal FLAG epitope. Western blotting of samples taken throughout the purification processes indicates that all detectable protein was found in the soluble fractions of the cell lysates (TSP, Figure 6). Thus, most, if not all, of the recombinant protein is soluble. Furthermore, the FLAG-tagged proteins efficiently bound to the resin as indicated by little to no detectable protein in the column flow-through, allowing for ease of purification and good recovery (Flow, Figure 5). Finally, coomassie staining of purified 14FN3 and HMGB1 each revealed a single predominant band at approximately the predicted molecular weight. Coomassie staining of purified VEGF revealed a single predominant band around 16 kDa, the predicted size of the monomer, with a faint band at

approximately 30 kDa, the expected mass of the VEGF dimer. 14FN3 has a predicted molecular mass of 12,820 mass units. Algal-expressed 14FN3 has a mass average of 12,820 and appears predominantly as a single peak in matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis (Figure 6). HMGB1 aa13–169 has a predicted molecular mass of 24,036 mass units. The MALDI-TOF MS analysis of algal-expressed HMGB1 shows predominately a single peak at 24,059, just 23 mass units off the predicted value. Two peaks appear in the VEGF MALDI-TOF MS analysis, a mass average value of 16,985 and 33,901. The predicted value of a VEGF monomer is 17,064, and 34,128 for the dimer. Thus the two peaks mostly likely represent the monomer and dimer of VEGF. Together, these data suggest that the algal-expressed proteins accumulate in a soluble form, can be purified to high purity, and are largely unmodified.

To test for bioactivity, purified VEGF and HMGB1 were assayed using a VEGF receptor-binding assay and a fibroblast chemotaxis assay, respectively. As potential antibody mimics, there is no bioactivity assay for 10FN3 or 14FN3. The most important characteristic of these recombinant proteins is that they are soluble, which we have shown to be the case when expressed in the chloroplast (Figure 3 and Figure 6). To determine whether algal-produced VEGF is biologically active, a sandwich ELISA was first performed to demonstrate antigenic integrity, an indicator of correct folding, and was used to establish effective concentration by reference to commercially available VEGF produced in *E. coli* (Figure 7A). Algal-expressed VEGF was then compared to bacterial-expressed VEGF in a VEGF-receptor binding assay. Algal-expressed VEGF exhibited dose-dependent binding to the VEGF receptor, albeit with slightly lower affinity as compared to the control bacterial-expressed VEGF (Supplemental Figure 4). This may be due to the presence of a small proportion of misfolded or truncated VEGF in the protein preparation. To determine VEGF bioactivity, a VEGF-receptor binding competition assay was used. VEGF derived from bacteria was able to compete with VEGF derived from algae for VEGF receptor binding (Figure 7B). Bacterial VEGF displaced algal VEGF (200ng/ml) from VEGFR with an IC50 of ~40ng/ml, consistent with a shared binding-site and broadly similar affinity. Overall, the data demonstrates that algal chloroplasts have the capability to express bioactive VEGF.

To determine whether algal-produced HMGB1 is biologically active, purified HMGB1 (~1mg) was sent to BioQuant (San Diego, CA), an independent contract research organization for bioactivity analysis using a mouse (A) or pig (B) fibroblast chemotaxis assay (Figure 8). The algal expressed HMGB1 (Scripps) showed similar bioactivity to commercial bacterial-expressed HMGB1 (Bio3 HMGB1; HMGBiotech, Italy).

Taken together, these data indicate that high quantities of highly purified and bioactive human therapeutic proteins can be expressed in and purified from the chloroplast of *C. reinhardtii*.

Discussion

In this study we investigated the general versatility of algae as a platform for therapeutic protein production by examining the expression of seven recombinant human proteins in the chloroplast of *Chlamydomonas reinhardtii*. The seven proteins are either presently sold as therapeutics, or have the potential to become human therapeutics, and each protein was tested using three different expression strategies. Protein expression and mRNA levels were quantitatively compared for each protein. Some level of expression was observed for five of the seven proteins, and three of the proteins accumulated to levels above 1% of total soluble protein, levels sufficient for easy purification. Another protein accumulated to these same high levels when fused to a stable and highly expressed protein, M-SAA, and an additional protein could be detected by immunoprecipitation, but accumulated at very low levels.

Importantly, all four of the highly expressed proteins were soluble, with no evidence that any of them were found in insoluble aggregates or inclusion bodies. Although this is a small number of proteins, it is a diverse set of protein types, and this level of success is much greater then the 20% to 30% success rate of human and viral proteins expressed in bacteria (Alzari et al., 2006), and equivalent to the 45% (Banci et al., 2006) to 58% (Aricescu et al., 2006) success rate reported for recombinant protein expression in other eukaryotic systems.

While the highest levels of recombinant protein accumulation were found to occur in the non-photosynthetic *psbA* null background, restoration of photosynthesis by reintroduction of the *psbA* gene under the control of the *psbD* promoter (Manuell et al., 2007), resulted in photosynthetic strains with recombinant protein accumulation, albeit levels were variable and reduced compared to the non-photosynthetic strains. Thus, high levels of recombinant protein accumulation can be achieved under photosynthetic growth conditions, which is required for some of the economic advantages algae holds over other expression systems. The mechanism in which D1 protein levels affect *psbA* promoter and/or UTR function is currently under investigation in our lab.

VEGF and HMGB1 accumulated to 3% and 2.5% respectively, and these proteins were purified using affinity chromatography to the FLAG epitope added to the carboxy terminus of each protein. Using standard bioactivity assays both proteins were found to have similar activity as the same proteins expressed in bacteria, the system presently used for production of these two proteins for therapeutic use. The yields reported here are close to the yields reported for therapeutic proteins expressed from the chloroplast of higher plants. For example, human serum albumin was reported to accumulate to 11% TSP (Fernandez-San Millan et al., 2003), somatotropin to 7% TSP (Staub et al., 2000), interferon gamma to 6% TSP (Leelavathi, 2003), and a CTB-proinsulin fusion to 16% TSP (Ruhlman et al., 2007). These data confirm that algal chloroplasts are able to produce bioactive proteins, and that the proteins can be easily purified from algal extracts.

Algal biomass can presently be produced at about $3/kilogram at commercial scale (Chisti, 2007), and as 25% of biomass is soluble protein, we can project that at 2% of soluble protein, we can presently make recombinant protein at about $0.6/gram prior to purification. This is near the cost estimates for the least expensive protein expression systems presently available (Dove, 2002), and considerably cheaper than mammalian cell culture. With the expected improvements in expression levels above 2% total soluble protein, and the continued reduction in algal biomass cost associated with the large scale efforts to use algae for biofuel production, we anticipate at least a 10-fold reduction in these cost over the next few years, which should make algal protein production the least expensive platform available. This reduced cost of goods, coupled with an ability to rapidly scale production in inexpensive bioreactors, suggests that algae may become an economically superior platform for therapeutic protein production in the future.

Why do some proteins accumulate in algae while others do not? Protein expression is variable in all expression platforms and algae are not unique in that regard. The greater than 50% expression that we find in this study is actually much higher then the percent of protein expressed in bacterial systems (Alzari et al., 2006), and comparable with the best expression reported for other eukaryotic systems (Aricescu et al., 2006; Banci et al., 2006). RT-qPCR analysis revealed that mRNA transcripts accumulated for all recombinant genes tested, and there was a poor correlation between transcript level and protein accumulation, suggesting that transcription and mRNA accumulation do not determine the level of recombinant protein accumulation in algae. We also see that the proteins that expressed with the *atpA* promoter and UTR were the same proteins that expressed with the *psbA* promoter, suggesting that failure to accumulate these proteins is not determined by the promoter or

UTRs used. Thus, either the proteins that express poorly are highly unstable, or their coding regions somehow precluded translation of the chimeric mRNAs. Work in our laboratory is presently under way to determine if there are general rules that can be applied to predict if a protein will be expressed in algal chloroplasts or not.

Although accumulation of recombinant proteins in algae at 2% to 3% of total soluble protein is sufficient for economic production, higher levels of accumulation would obviously reduce cost even more, and would also likely improve protein purification efficiency. We are examining a variety of ways to increase protein production beyond our present accumulation levels, and as this platform is still relatively new, we expect additional improvements in this area over the next few years. The data also shows that the *psbA* promoter/UTR showed better protein expression for all the proteins tested compared to the *atpA* promoter/UTR, and that this increase does not correlate directly with increased mRNA accumulation. This data suggest that translation of chimeric mRNAs containing the *psbA* UTR is better than translation of mRNAs with the same coding but containing the *atpA* UTR. Thus it is possible that altering UTRs may further improve protein translation as a way to increase protein accumulation, and work in the lab to address this is now ongoing.

Experimental Procedures

Plasmid construction

Codon optimization for *C. reinhardtii* chloroplast expression was performed using software developed in-house, specifically designed for polymerase cycling assembly (PCA)-based de-novo gene synthesis. This program generates gene sequences by the simultaneous optimization of multiple parameters: normalization of the codon distribution to that of the *C. reinhardtii* chloroplast (data obtained from http://www.kazusa.or.jp/codon (Nakamura et al., 2000)); uniformity of physical properties of the output oligonucleotides (GC content, melting temperature, length); and avoidance of unfavorable mRNA structures. The seven genes were assembled by PCA using sense and antisense oligonucleotides ranging in length from 51 to 63 bases, sharing eighteen base pairs of overlapping sequence homology (Minshull et al., 2004). The number of oligonucleotides used ranged from eight oligos for proinsulin to sixteen oligos for HMGB1.

Each gene was constructed with an *NdeI* restriction site at the 5′ end and a *XbaI* site at the 3′ end of the coding region, along with a C-terminal TEV protease recognition site and a FLAG-tag. The genes were directionally cloned into the pD1-Kan vector, constructed by the addition of the Kanamycin resistance gene *aphA*6 (*Acinetobacter baumannii*) into the unique *BamHI* site in the *psbA* vector described previously (Manuell et al., 2007). The coding sequence of *aphA6* was ordered in *C. reinhardtii* chloroplast codon bias from DNA2.0 (www.dna20.com) flanked by *atp*A 5′ promoter and UTR and *rbc*L 3′ UTR sequences (Barnes et al., 2005). For constructs containing the *atpA* promoter, the genes of interest were cloned into the unique *NdeI/XbaI* restriction sites in p3222 vector (Franklin et al., 2002).

CAI values were determined with the CAI calculator (http://genomes.urv.cat/CAIcal/) using the *C. reinhardtii* chloroplast codon usage table (http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=3055.chloroplast). CAI values range from 0 to 1, with 1 being if a gene always uses the most frequently used codon of a reference set (Puigbo et al., 2008).

C. reinhardtii strains, transformations and growth conditions

For chloroplast transformations, *C. reinhardtii* wt strain 137c (mt+) was grown to late logarithmic phase in the presence of 40 mM 5-fluorodeoxyuridine in TAP (Tris-acetate-phosphate) medium (Gorman and Levine, 1965) at 23°C under constant illumination of 5000

lux on a rotary shaker. Cells were harvested by centrifugation and resuspended in TAP medium. Approximately 5×10^7 cells/plate were spread onto TAP/agar plates containing the appropriate antibiotic (see below). Chloroplast transformations were performed by particle bombardment (Boynton et al., 1988) using DNA-coated gold particles (S550d, Seashell Technologies, San Diego, CA). The following bombardment parameters were used with the PDS-1000/He system (BioRad, Hercules, CA): chamber vacuum of 27–28 inches Hg, target distance of 9 cm, helium pressure of 1350 psi, and approximately 1 mg of 0.55 μm gold coated with 3 μg of DNA per transformation plate. Transformations with the *psbA* promoter were carried out using kanamycin selection (100μg/ml in TAP agar, 150μg/ml for propagation after transformation), for which resistance was conferred by the kanamycin resistance gene *aphA6* expressed from the same construct under the control of the *atpA* promoter/5′ UTR and *rbcl* 3′ UTR. The recombinant genes under the control of the *atpA* promoter were co-transformed with the plasmid p228 (Chlamydomonas Stock Center, Duke University, Durham, NC, USA), which contains a point mutation in the 16S rRNA gene that confers spectinomycin resistance. Chloroplast transgenic lines were identified by growth on media containing spectinomycin (150μg/ml TAP agar).

PCR screening of transformants

Primary transformants were screened for the presence of the gene of interest using promoter specific forward primers and gene specific reverse primers. Cells were resuspended in Tris-EDTA solution and heated to 95°C for 10 minutes. The cell lysate was then used as a template for PCR under standard conditions for 40 cycles using *Taq* polymerase (NEB, Ipswich, MA). For assessing homoplasmicity of clones, the PCR primers 5′-GGAAGGGGAGGACGTAGGTACATAAA-3′ and 5′-TTAGAACGTGTTTTGTTCCCAAT-3′ were used for constructs incorporated at the *psbA* site and 5′-CCCATAAATAGTTTCAATTGG-3′ and 5 ′-CGGTGGTTATTCCAGGCCAAACTTATG-3′ for constructs incorporated at the p322 site. Each set of primers anneal to regions that are disrupted upon gene insertion through homologous recombination. Thus, the loss of a PCR product indicates proper gene integration into all copies of the chloroplast genome. For these reactions, a second set of control primers were used (5′-CCGAACTGAGGTTGGGTTTA-3′ and 5′-GGGGGAGCGAATAGGATTAG-3′). These primers amplify a region of the genome away from the recombinant gene insertion site and serve as a positive control in the multiplex PCR.

Western blotting

Whole cell samples were resuspended in lysis buffer (Franklin et al., 2002) and lysed by sonication. Total soluble protein was isolated by centrifugation and denatured by the addition of SDS-PAGE loading buffer (Laemelli) followed by incubation at 60°C for 10 minutes. When protein determination was required, a sample was taken prior to the addition of SDS-PAGE sample buffer and protein concentration determined using the Bio-Rad DC protein assay as per the manufacturer's instructions (Bio-Rad). Proteins were separated on 12% or 16% SDS-page gels at 120–150 volts unless otherwise stated and transferred to nitrocellulose membrane at 200 mAmps for 1.5 hours. After blocking with 5% milk, membranes were probed with an anti-FLAG monoclonal antibody conjugated to HRP (A8592, Sigma, St. Louis, MO) or to alkaline phosphatase (A9469, Sigma).

Protein purification

1–2 liters of algal cells grown to late log phase in TAP media were collected by centrifugation at 5000×G for 10 minutes. The cell paste was resuspended to a volume of 40–100ml per liter of culture in lysis buffer, 50 mM Tris pH 8.0, 400 mM NaCl, 0.1% Tween 20, 1 mM phenylmethylsulfonylfluoride (PMSF), and lysed by sonication. The lysate was

clarified by centrifugation at 30,000×G for 20 minutes, and the supernatant was collected. One-two mls of anti-FLAG M2 resin (Sigma) was added to the clarified lysate, and rotated end over end at 4°C for four hours. The anti-FLAG beads were collected by filtration in a Bio-Rad Econo-pac column, and washed extensively with lysis buffer. The protein was eluted from the resin using lysis buffer containing 100 micrograms per ml FLAG peptide (Sigma) or with 100 mM glycine pH 3.5, 400 mM NaCl and neutralized with Tris pH 7.9 to final a concentration of 50 mM. After adding five column volumes of elution buffer, the column was incubated at 4°C overnight, under rotation. Fractions were collected and assayed by SDS-PAGE followed by western blot and coomassie staining to determine the fractions that contained the recombinant protein. Fractions containing recombinant protein were pooled and concentrated using an Amicon Ultra centrifugal filter with a molecular weight cut-off of 5 kDa (Millipore, Billerica, MA). After concentration, samples were checked again by SDS-PAGE, and concentrations were determined using the BCA protein assay (Bio-Rad) with bovine serum albumin as a standard.

Activity assays

VEGF activity assay—VEGF concentration was determined by ELISA. Maxisorp plates were coated with monoclonal anti-human VEGF (R&D Systems, Minneapolis, MN). After blocking with BSA, serial dilutions of VEGF purified from algae and commercial bacteria-derived VEGF (R&D Systems) were applied. After washing, bound VEGF was detected using a biotinylated polyclonal anti-human VEGF antibody (R&D Systems), alkaline-phosphatase-conjugated streptavidin and pNPP substrate (Sigma). Readings from uncoated wells were subtracted to give specific binding.

Binding of VEGF to receptor was assessed in a similar way, by coating Maxisorp plates with a human VEGF-R2(KDR):Fc fusion protein (R&D Systems), applying VEGF and detecting bound protein using biotinylated anti-VEGF. For competition assays a serial dilution of bacteria-derived VEGF was applied along with a constant concentration of algae-derived VEGF. Bound algae-derived VEGF was detected using HRP-conjugated anti-FLAG antibody (Sigma).

HMGB1 activity assay

HMGB1 bioactivity was assessed using an *in vitro* analysis of the chemotactic effect of algal-expressed HMGB1. The relative ability of human and pig fibroblasts to migrate toward NIH3T3 conditioned media complemented with 10ng/ml VEGF, PDGF or HMGB1 was assessed using a modified Boyden chamber (NeuroProbe, Inc., Gaithersburg, MD). The cells were placed in the apical chambers of the apparatus and the media containing the chemotactic factors were placed in the basal chambers. A PVP membrane with 8 micron pores coated with 1mg/ml collagen IV separated the apical and basal chambers. Cells were incubated for 24 hours at 37°C. The cells that migrated onto the basal surface of the membrane were stained and quantified using a microscope.

RT-quantitative PCR analysis of mRNA accumulation

Cells were grown in 50 ml liquid TAP under 1000 lux of light illumination until mid to late log phase. 10 mls of cells were harvested by centrifugation and total RNA was purified using the Plant RNA Reagent (Invitrogen, Carlsbad, CA) as per the manufacturer's instructions for small scale purification. RNA integrity was monitored by agarose gel electrophoresis (Supplemental Figure 3). 10 μg of total RNA was treated with DNase to remove any contaminating genomic DNA (Ambion DNA-free, Austin, TX). 400 ng of DNase-treated total RNA was then used for first strand cDNA synthesis using Bio-Rad's iScript cDNA Synthesis kit (following the manufacturer's instructions). The thermocycling conditions used are as follows: 5 minutes at 25°C, 30 minutes at 42°C, 5 minutes at 85°C,

hold at 4°C. Reactions were also carried out in the absence of reverse transcriptase as a control for genomic DNA contamination. cDNAs were either diluted 1:10 for the qPCR experimental reactions, or diluted 1:4 and then subjected to a 4-fold serial dilution to determine PCR efficiencies for each primer pair. 6.5 μl of diluted cDNA was used in a 25 μl qPCR reaction using Bio-Rad SYBR Green Supermix and 0.5 μM oligonucleotides. Real-time qPCR was carried out in triplicate for each sample in a Bio-Rad My iQ thermocycler performing 40 cycles of a two-step protocol with an annealing/extension temperature of 55°C, followed by a melt curve to monitor for primer dimers. For all qPCR reactions, *rbcL* was used as the control gene. Relative mRNA levels were determined using the Pfaffl method (ratio = E^-dCt $_{target}$/E^-dCt $_{control}$) (Pfaffl, 2001).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant to Stephen P. Mayfield from The National Institutes of Health (A1059614), a grant to Beth A. Rasala from the San Diego Foundation (C-2008-00296), and by Sapphire Energy. We thank Guido Cappuccilli, who performed the codon optimization, and Miller Tran for careful reading of the manuscript.

References

Alzari PM, Berglund H, Berrow NS, Blagova E, Busso D, Cambillau C, Campanacci V, Christodoulou E, Eiler S, Fogg MJ, Folkers G, Geerlof A, Hart D, Haouz A, Herman MD, Macieira S, Nordlund P, Perrakis A, Quevillon-Cheruel S, Tarandeau F, van Tilbeurgh H, Unger T, Luna-Vargas MP, Velarde M, Willmanns M, Owens RJ. Implementation of semi-automated cloning and prokaryotic expression screening: the impact of SPINE. Acta Crystallogr D Biol Crystallogr. 2006; 62:1103–1113. [PubMed: 17001088]

Arakawa T, Yu J, Chong DK, Hough J, Engen PC, Langridge WH. A plant-based cholera toxin B subunit-insulin fusion protein protects against the development of autoimmune diabetes. Nat Biotechnol. 1998; 16:934–938. [PubMed: 9788349]

Aricescu AR, Assenberg R, Bill RM, Busso D, Chang VT, Davis SJ, Dubrovsky A, Gustafsson L, Hedfalk K, Heinemann U, Jones IM, Ksiazek D, Lang C, Maskos K, Messerschmidt A, Macieira S, Peleg Y, Perrakis A, Poterszman A, Schneider G, Sixma TK, Sussman JL, Sutton G, Tarboureich N, Zeev-Ben-Mordehai T, Jones EY. Eukaryotic expression: developments for structural proteomics. Acta Crystallogr D Biol Crystallogr. 2006; 62:1114–1124. [PubMed: 17001089]

Azhagiri AK, Maliga P. Exceptional paternal inheritance of plastids in Arabidopsis suggests that low-frequency leakage of plastids via pollen may be universal in plants. Plant J. 2007; 52:817–823. [PubMed: 17931353]

Banci L, Bertini I, Cusack S, de Jong RN, Heinemann U, Jones EY, Kozielski F, Maskos K, Messerschmidt A, Owens R, Perrakis A, Poterszman A, Schneider G, Siebold C, Silman I, Sixma T, Stewart-Jones G, Sussman JL, Thierry JC, Moras D. First steps towards effective methods in exploiting high-throughput technologies for the determination of human protein structures of high biomedical value. Acta Crystallogr D Biol Crystallogr. 2006; 62:1208–1217. [PubMed: 17001097]

Barnes D, Franklin S, Schultz J, Henry R, Brown E, Coragliotti A, Mayfield SP. Contribution of 5′- and 3′-untranslated regions of plastid mRNAs to the expression of Chlamydomonas reinhardtii chloroplast genes. Mol Genet Genomics. 2005; 274:625–636. [PubMed: 16231149]

Bock R. Plastid biotechnology: prospects for herbicide and insect resistance, metabolic engineering and molecular farming. Curr Opin Biotechnol. 2007; 18:100–106. [PubMed: 17169550]

Boynton JE, Gillham NW, Harris EH, Hosler JP, Johnson AM, Jones AR, Randolph-Anderson BL, Robertson D, Klein TM, Shark KB, et al. Chloroplast transformation in Chlamydomonas with high velocity microprojectiles. Science. 1988; 240:1534–1538. [PubMed: 2897716]

Butt TR, Edavettal SC, Hall JP, Mattern MR. SUMO fusion technology for difficult-to-express proteins. Protein Expr Purif. 2005; 43:1–9. [PubMed: 16084395]

Chan SJ, Weiss J, Konrad M, White T, Bahl C, Yu SD, Marks D, Steiner DF. Biosynthesis and periplasmic segregation of human proinsulin in Escherichia coli. Proc Natl Acad Sci U S A. 1981; 78:5401–5405. [PubMed: 7029534]

Chebolu S, Daniell H. Chloroplast-derived vaccine antigens and biopharmaceuticals: expression, folding, assembly and functionality. Curr Top Microbiol Immunol. 2009; 332:33–54. [PubMed: 19401820]

Chisti Y. Biodiesel from microalgae. Biotechnol Adv. 2007; 25:294–306. [PubMed: 17350212]

Daniell H. Production of biopharmaceuticals and vaccines in plants via the chloroplast genome. Biotechnol J. 2006; 1:1071–1079. [PubMed: 17004305]

De Marco V, Stier G, Blandin S, de Marco A. The solubility and stability of recombinant proteins are increased by their fusion to NusA. Biochem Biophys Res Commun. 2004; 322:766–771. [PubMed: 15336530]

Demain AL, Vaishnav P. Production of recombinant proteins by microbes and higher organisms. Biotechnol Adv. 2009; 27:297–306. [PubMed: 19500547]

Dong Xda E, Ito N, Lotze MT, Demarco RA, Popovic P, Shand SH, Watkins S, Winikoff S, Brown CK, Bartlett DL, Zeh HJ 3rd. High mobility group box I (HMGB1) release from tumor cells after treatment: implications for development of targeted chemoimmunotherapy. J Immunother. 2007; 30:596–606. [PubMed: 17667523]

Dove A. Uncorking the biomanufacturing bottleneck. Nat Biotechnol. 2002; 20:777–779. [PubMed: 12148000]

Eberhard S, Drapier D, Wollman FA. Searching limiting steps in the expression of chloroplast-encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of Chlamydomonas reinhardtii. Plant J. 2002; 31:149–160. [PubMed: 12121445]

Eschbach JW, Abdulhadi MH, Browne JK, Delano BG, Downing MR, Egrie JC, Evans RW, Friedman EA, Graber SE, Haley NR, et al. Recombinant human erythropoietin in anemic patients with end-stage renal disease. Results of a phase III multicenter clinical trial. Ann Intern Med. 1989; 111:992–1000. [PubMed: 2688507]

Fernandez-San Millan A, Mingo-Castel A, Miller M, Daniell H. A chloroplast transgenic approach to hyper-express and purify Human Serum Albumin, a protein highly susceptible to proteolytic degradation. Plant Biotechnol J. 2003; 1:71–79. [PubMed: 17147744]

Franklin S, Ngo B, Efuet E, Mayfield SP. Development of a GFP reporter gene for Chlamydomonas reinhardtii chloroplast. Plant J. 2002; 30:733–744. [PubMed: 12061904]

Franklin SE, Mayfield SP. Recent developments in the production of human therapeutic proteins in eukaryotic algae. Expert Opin Biol Ther. 2005; 5:225–235. [PubMed: 15757384]

Garcia-Ibilcieta D, Bokov M, Cherkasov V, Sveshnikov P, Hanson SF. Simple method for production of randomized human tenth fibronectin domain III libraries for use in combinatorial screening procedures. Biotechniques. 2008; 44:559–562. [PubMed: 18476822]

Gorman DS, Levine RP. Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of Chlamydomonas reinhardi. Proc Natl Acad Sci U S A. 1965; 54:1665–1669. [PubMed: 4379719]

Hagemann, R. Molecular Biology and Biotechnology of Plant Organelles. Netherlands: Springer Netherlands; 2004. The Sexual Inheritance of Plant Organelles; p. 93-114.

Haroon ZA, Amin K, Jiang X, Arcasoy MO. A novel role for erythropoietin during fibrin-induced wound-healing response. Am J Pathol. 2003; 163:993–1000. [PubMed: 12937140]

Jelkmann W. Erythropoietin after a century of research: younger than ever. Eur J Haematol. 2007; 78:183–205. [PubMed: 17253966]

Koide A, Koide S. Monobodies: antibody mimics based on the scaffold of the fibronectin type III domain. Methods Mol Biol. 2007; 352:95–109. [PubMed: 17041261]

Krynetskaia N, Xie H, Vucetic S, Obradovic Z, Krynetskiy E. High mobility group protein B1 is an activator of apoptotic response to antimetabolite drugs. Mol Pharmacol. 2008; 73:260–269. [PubMed: 17951356]

Leelavathi, SaRVS. Chloroplast expression of His-tagged GUS-fusions: a general strategy to overproduce and purify foreign proteins using transplastomic plants as bioreactors. Molecular Breeding. 2003; 11:49–58.

Manuell AL, Beligni MV, Elder JH, Siefker DT, Tran M, Weber A, McDonald TL, Mayfield SP. Robust expression of a bioactive mammalian protein in Chlamydomonas chloroplast. Plant Biotechnol J. 2007; 5:402–412. [PubMed: 17359495]

Mayfield SP, Franklin SE, Lerner RA. Expression and assembly of a fully active antibody in algae. Proc Natl Acad Sci U S A. 2003; 100:438–442. [PubMed: 12522151]

Mayfield SP, Manuell AL, Chen S, Wu J, Tran M, Siefker D, Muto M, Marin-Navarro J. Chlamydomonas reinhardtii chloroplasts as protein factories. Curr Opin Biotechnol. 2007; 18:126–133. [PubMed: 17317144]

Mayfield SP, Schultz J. Development of a luciferase reporter gene, luxCt, for Chlamydomonas reinhardtii chloroplast. Plant J. 2004; 37:449–458. [PubMed: 14731263]

McCormack PL, Scott LJ. Interferon-beta-1b: a review of its use in relapsing-remitting and secondary progressive multiple sclerosis. CNS Drugs. 2004; 18:521–546. [PubMed: 15182221]

Minai L, Wostrikoff K, Wollman FA, Choquet Y. Chloroplast biogenesis of photosystem II cores involves a series of assembly-controlled steps that regulate translation. Plant Cell. 2006; 18:159–175. [PubMed: 16339851]

Minshull J, Govindarajan S, Cox T, Ness JE, Gustafsson C. Engineered protein function by selective amino acid diversification. Methods. 2004; 32:416–427. [PubMed: 15003604]

Murdoch D, Lyseng-Williamson KA. Subcutaneous recombinant interferon-beta-1a (Rebif): a review of its use in relapsing-remitting multiple sclerosis. Drugs. 2005; 65:1295–1312. [PubMed: 15916455]

Muto M, Henry RE, Mayfield SP. Accumulation and processing of a recombinant protein designed as a cleavable fusion to the endogenous Rubisco LSU protein in Chlamydomonas chloroplast. BMC Biotechnol. 2009; 9:26. [PubMed: 19323825]

Nakamura Y, Gojobori T, Ikemura T. Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res. 2000; 28:292. [PubMed: 10592250]

Nickelsen J. Chloroplast RNA-binding proteins. Curr Genet. 2003; 43:392–399. [PubMed: 12955455]

Nykiforuk CL, Boothe JG, Murray EW, Keon RG, Goren HJ, Markley NA, Moloney MM. Transgenic expression and recovery of biologically active recombinant human insulin from Arabidopsis thaliana seeds. Plant Biotechnol J. 2006; 4:77–85. [PubMed: 17177787]

Oey M, Lohse M, Kreikemeyer B, Bock R. Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. Plant J. 2009; 57:436–445. [PubMed: 18939966]

Pankov R, Yamada KM. Fibronectin at a glance. J Cell Sci. 2002; 115:3861–3863. [PubMed: 12244123]

Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001; 29:e45. [PubMed: 11328886]

Potgens AJ, Lubsen NH, van Altena MC, Vermeulen R, Bakker A, Schoenmakers JG, Ruiter DJ, de Waal RM. Covalent dimerization of vascular permeability factor/vascular endothelial growth factor is essential for its biological activity. Evidence from Cys to Ser mutations. J Biol Chem. 1994; 269:32879–32885. [PubMed: 7806514]

Pryor KD, Leiting B. High-level expression of soluble protein in Escherichia coli using a His6-tag and maltose-binding-protein double-affinity fusion system. Protein Expr Purif. 1997; 10:309–319. [PubMed: 9268677]

Puigbo P, Bravo IG, Garcia-Vallve S. E-CAI: a novel server to estimate an expected value of Codon Adaptation Index (eCAI). BMC Bioinformatics. 2008; 9:65. [PubMed: 18230160]

Pulz O. Photobioreactors: production systems for phototrophic microorganisms. Appl Microbiol Biotechnol. 2001; 57:287–293. [PubMed: 11759675]

Ruf S, Karcher D, Bock R. Determining the transgene containment level provided by chloroplast transformation. Proc Natl Acad Sci U S A. 2007; 104:6998–7002. [PubMed: 17420459]

Ruhlman T, Ahangari R, Devine A, Samsam M, Daniell H. Expression of cholera toxin B-proinsulin fusion protein in lettuce and tobacco chloroplasts--oral administration protects against

development of insulitis in non-obese diabetic mice. Plant Biotechnol J. 2007; 5:495–510. [PubMed: 17490448]

Sachdev D, Chirgwin JM. Fusions to maltose-binding protein: control of folding and solubility in protein purification. Methods Enzymol. 2000; 326:312–321. [PubMed: 11036650]

Shi Y, Wang T, Li Y, Darmency H. Impact of transgene inheritance on the mitigation of gene flow between crops and their wild relatives: the example of foxtail millet. Genetics. 2008; 180:969–975. [PubMed: 18780732]

Siren AL, Knerlich F, Poser W, Gleiter CH, Bruck W, Ehrenreich H. Erythropoietin and erythropoietin receptor in human ischemic/hypoxic brain. Acta Neuropathol. 2001; 101:271–276. [PubMed: 11307627]

Staub JM, Garcia B, Graves J, Hajdukiewicz PT, Hunter P, Nehra N, Paradkar V, Schlittler M, Carroll JA, Spatola L, Ward D, Ye G, Russell DA. High-yield production of a human therapeutic protein in tobacco chloroplasts. Nat Biotechnol. 2000; 18:333–338. [PubMed: 10700152]

Stepien PP, Brousseau R, Wu R, Narang S, Thomas DY. Synthesis of a human insulin gene. VI. Expression of the synthetic proinsulin gene in yeast. Gene. 1983; 24:289–297. [PubMed: 6357951]

Streatfield SJ. Approaches to achieve high-level heterologous protein production in plants. Plant Biotechnol J. 2007; 5:2–15. [PubMed: 17207252]

Strong TD, Gebska MA, Burnett AL, Champion HC, Bivalacqua TJ. Endothelium-specific gene and stem cell-based therapy for erectile dysfunction. Asian J Androl. 2008; 10:14–22. [PubMed: 18087639]

Sun NK, Chao CC. The cytokine activity of HMGB1--extracellular escape of the nuclear protein. Chang Gung Med J. 2005; 28:673–682. [PubMed: 16382751]

Surzycki R, Greenham K, Kitayama K, Dibal F, Wagner R, Rochaix JD, Ajam T, Surzycki S. Factors effecting expression of vaccines in microalgae. Biologicals. 2009; 37:133–138. [PubMed: 19467445]

Svab Z, Maliga P. Exceptional transmission of plastids and mitochondria from the transplastomic pollen parent and its impact on transgene containment. Proc Natl Acad Sci U S A. 2007; 104:7003–7008. [PubMed: 17420457]

Tran M, Zhou B, Pettersson PL, Gonzalez MJ, Mayfield SP. Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts. Biotechnol Bioeng. 2009

Walsh G. Biopharmaceutical benchmarks--2003. Nat Biotechnol. 2003; 21:865–870. [PubMed: 12894198]

Walsh G. Biopharmaceutical benchmarks 2006. Nat Biotechnol. 2006; 24:769–776. [PubMed: 16841057]

Wang C, Castro AF, Wilkes DM, Altenberg GA. Expression and purification of the first nucleotide-binding domain and linker region of human multidrug resistance gene product: comparison of fusions to glutathione S-transferase, thioredoxin and maltose-binding protein. Biochem J. 1999; 338 (Pt 1):77–81. [PubMed: 9931301]

Wang T, Li Y, Shi Y, Reboud X, Darmency H, Gressel J. Low frequency transmission of a plastid-encoded trait in Setaria italica. Theor Appl Genet. 2004; 108:315–320. [PubMed: 14513219]

Warner-Schmidt JL, Duman RS. VEGF as a potential target for therapeutic intervention in depression. Curr Opin Pharmacol. 2008; 8:14–19. [PubMed: 18061540]

Wijelath ES, Rahman S, Namekata M, Murray J, Nishimura T, Mostafavi-Pour Z, Patel Y, Suda Y, Humphries MJ, Sobel M. Heparin-II domain of fibronectin is a vascular endothelial growth factor-binding domain: enhancement of VEGF biological activity by a singular growth factor/matrix protein synergism. Circ Res. 2006; 99:853–860. [PubMed: 17008606]

Zerges W. Translation in chloroplasts. Biochimie. 2000; 82:583–601. [PubMed: 10946109]

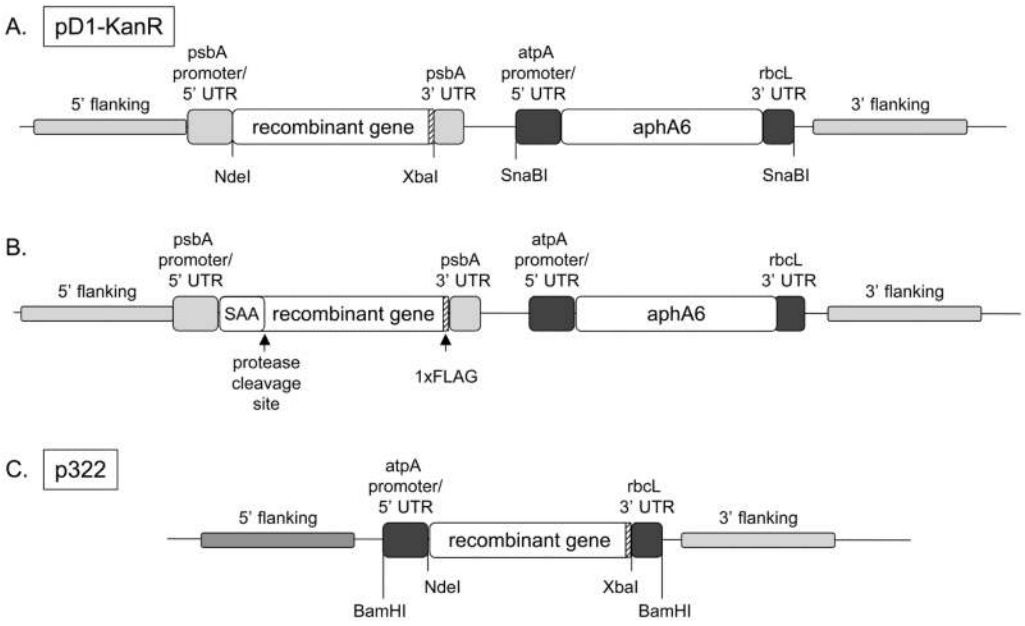


Figure 1. Introduction of the recombinant genes into the *Chlamydomonas reinhardtii* chloroplast genome

Schematic diagram of transformation vectors used, including relevant restriction sites. (A and B) pD1-Kan: Replacement of the endogenous *psbA* gene with the gene of interest (A), or with the gene of interest fused to the C-terminus of M-SAA (Manuell et al., 2007) (B). The kanamycin resistance gene *aphA6* under the control of the *atpA* promoter and 5′ UTR is genetically linked to the gene of interest. Grey regions flanking the gene of interest and resistance gene corresponds to regions of the chloroplast genome used for homologous recombination between the insertion plasmid and the *C. reinhardtii* chloroplast genome. (C) Schematic diagram of p322 (Franklin et al., 2002) used to transform the genes of interest under the control of the *atpA* promoter and 5′ UTR and the *rbcL* 3′ UTR into the BamHI silent site near the *psbA* gene (Barnes et al., 2005). (A-C) All recombinant proteins were C-terminally fused to the 1 x FLAG-tag sequence (DYKDDDDKS) for western blotting and purification.

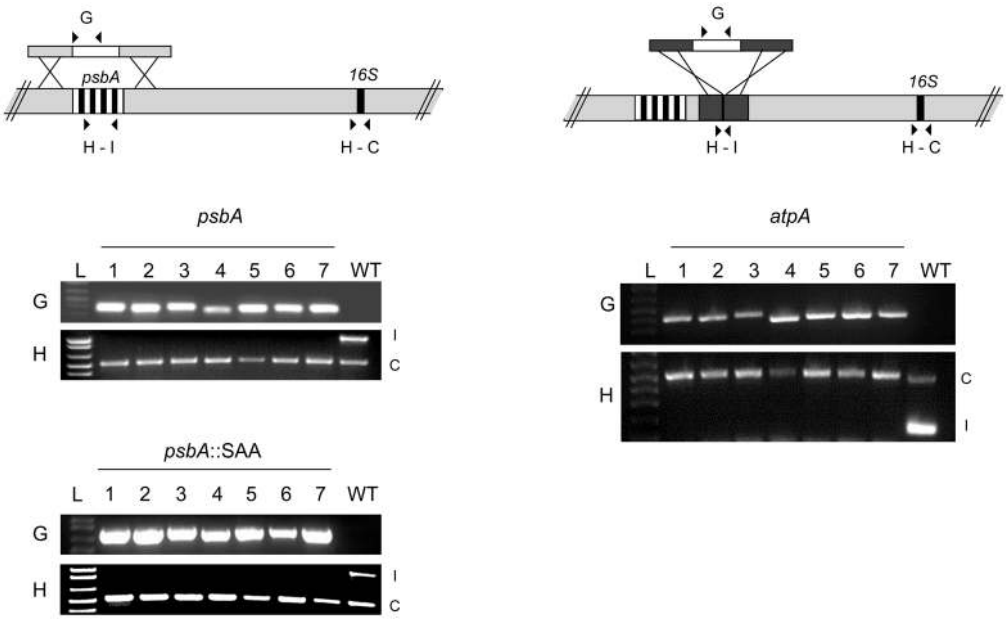


Figure 2. Identification of gene integration and isolation of homoplasmic strains

PCR using whole cell lysates as described in the Experimental Procedures. G: Gene specific PCR to show the presence of the corresponding recombinant gene in the transformants. H: PCR to show homoplasmicity of the clones. Each reaction contains two sets of primers (▶ ◀), one that amplifies an internal control gene (*16S rRNA*) to demonstrate that the PCR reactions worked, and can be seen in all lanes (H-C), and the other primer set amplifies the region of the genome that was targeted for integration and thus the parent strain shows a band whereas homoplasmic transformants do not (H-I).

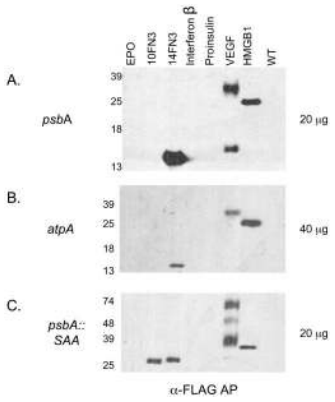


Figure 3. Accumulation of recombinant proteins in transgenic lines

Strains were grown under the same conditions and harvested for western blot as described in the Experimental Procedures. Equal amounts of total protein were loaded in each lane (20 μg for A and C, 40 μg for B). Western blots were probed with anti-FLAG antibody conjugated to Horse Radish Peroxidase. (A) Protein accumulation in strains when the corresponding genes were expressed from the *psbA* promoter and UTRs. (B) Protein accumulation when the corresponding genes were expressed from the *atpA* promoter and UTRs. (C) Protein accumulation of the SAA fusion proteins from the *psbA* promoter. Note, short exposure times were routinely used for (A) and (C), while much longer exposure times (several hours) were required to visualize the bands in (B).

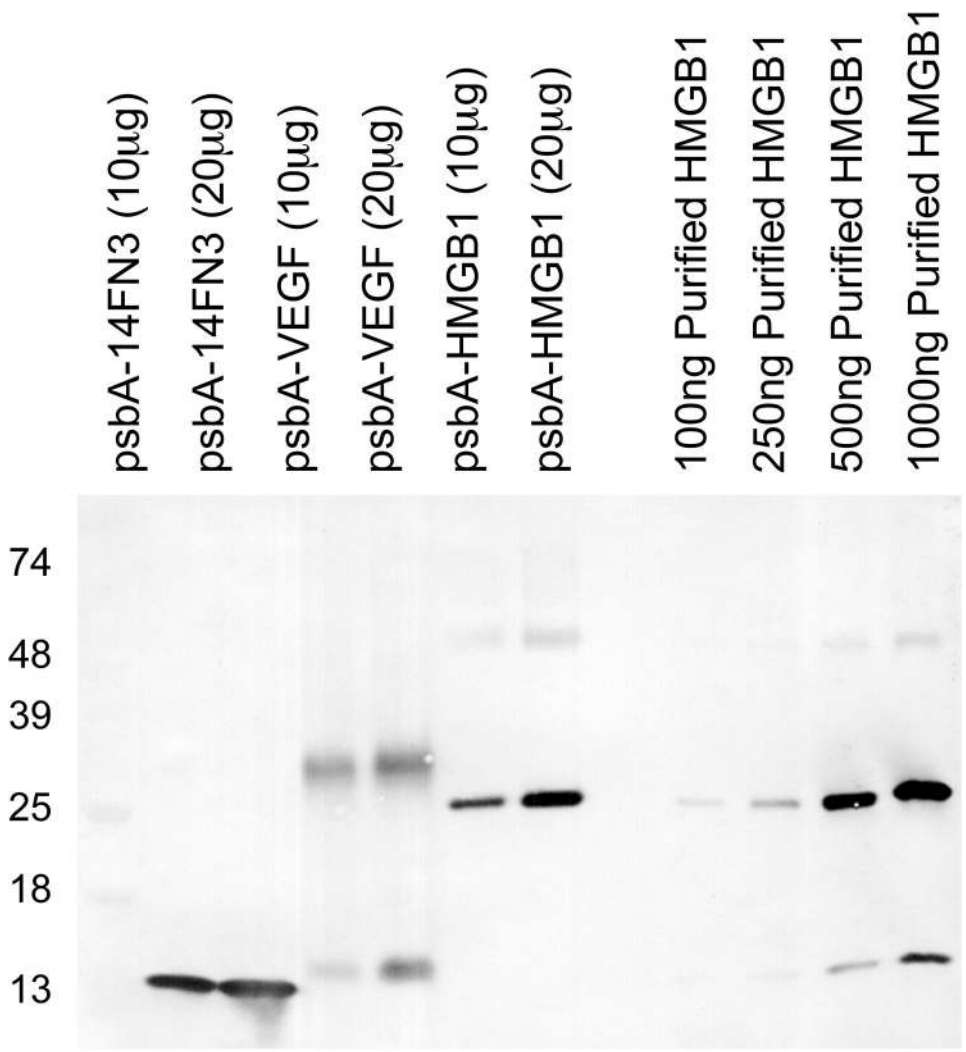
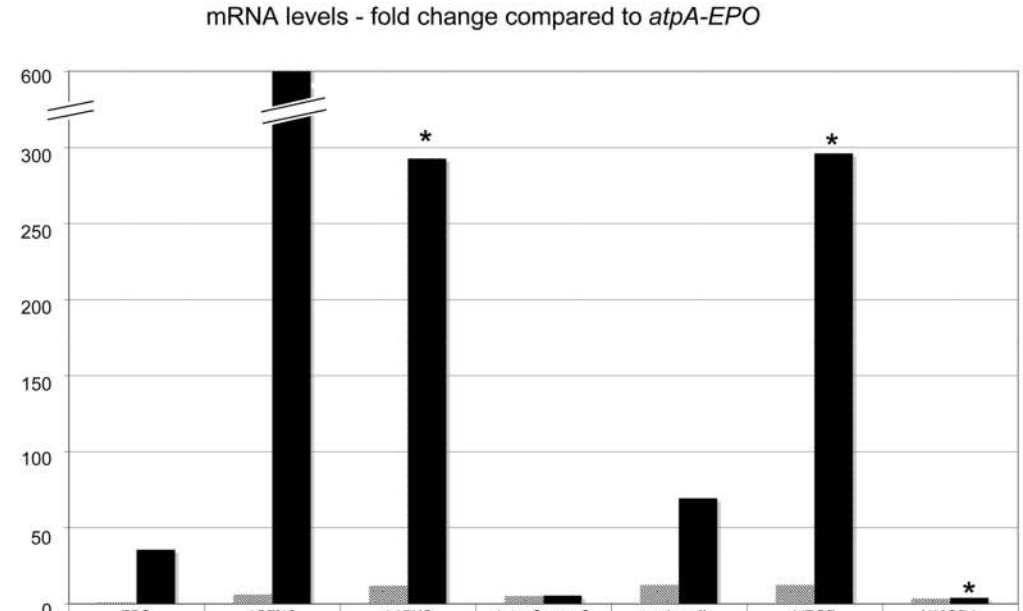


Figure 4. Quantitation of protein accumulation

Percent total soluble protein was determined for 14FN3, VEGF, and HMGB1 by loading 10 or 20 μg of soluble lysate from expression strains onto a SDS-PAGE gel alongside a serial dilution of highly pure HMGB1. Western blots were performed using anti-FLAG-HRP antibody.



	EPO	10FN3	14FN3	interferon-β	proinsulin	VEGF	HMGB1
atpA	1.0	6.1	12	5	12.5	12.3	3.4
psbA	35.4	577.4	292.5	5.4	69.3	296	3.9

Figure 5. Analysis of mRNA levels for *psbA* and *atpA* constructs

mRNA levels of the seven recombinant genes under the control of the *psbA* and *atpA* promoters. Fold change determined using the Pfaffl method (Pfaffl, 2001) to take into account differing PCR efficiencies with the different gene-specific primer pairs used. Note, *atpA-EPO* yielded the lowest level of mRNA, so all mRNA levels were calculated as fold change relative to *atpA-EPO*. The recombinant protein expressing lines are indicated (*).

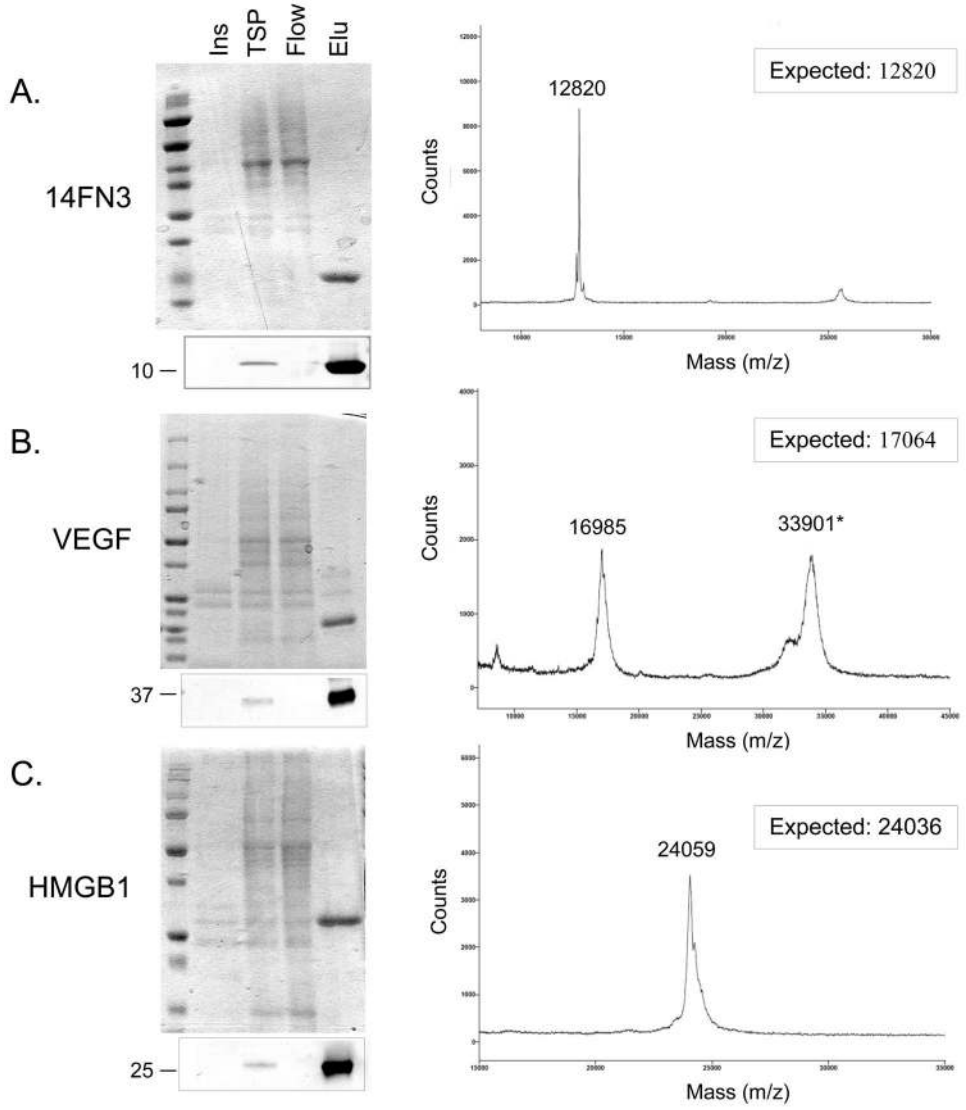


Figure 6. Affinity purification of algal-expressed therapeutic proteins

Coomassie staining (top left panel) and western blotting (boxed bottom panel) of (A) 14FN3, (B) VEGF, and (C) HMGB1 purifications are shown. Lanes from left to right are the following fractions: insoluble fraction (Ins), total soluble protein (TSP), column flow through (Flow), and the eluate (Elu). Equal volumes of Ins, TSP and Flow are loaded per lane. 3 μg of purified protein (Elu) is loaded on each coomassie gel, while 500 ng of Elu is loaded on each western blot. Right panels correspond to the MADLI-TOF MS results for each purified protein. Asterisk (*) in B indicates predicted dimer form of VEGF.

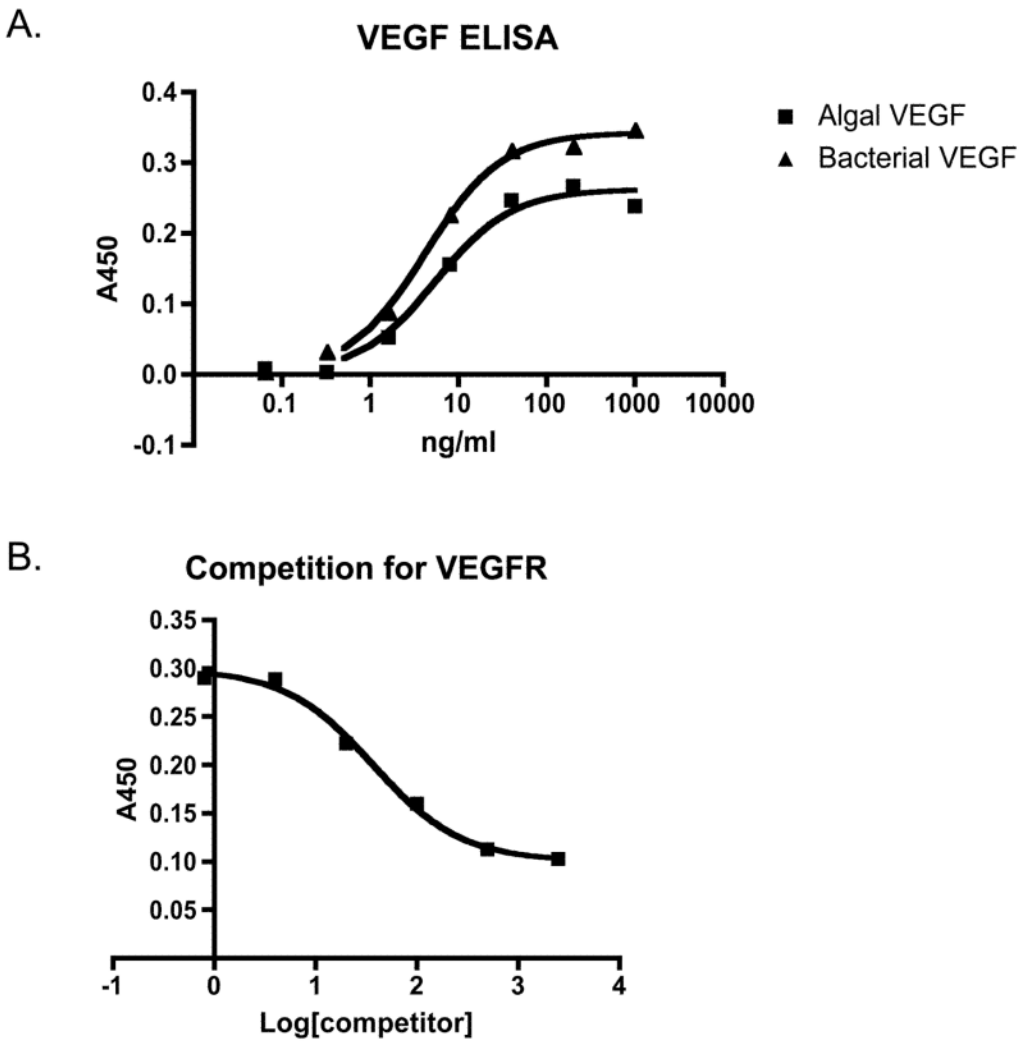


Figure 7. Bioactivity of VEGF

(A) Concentration of intact VEGF in the purified protein was assessed by comparison with bacteria-derived VEGF in a sandwich ELISA. (B) Competitive binding to the VEGF receptor was assayed by detecting binding of a fixed concentration of algal VEGF to VEGFR-coated wells in the presence of varying concentrations of bacteria-derived VEGF.

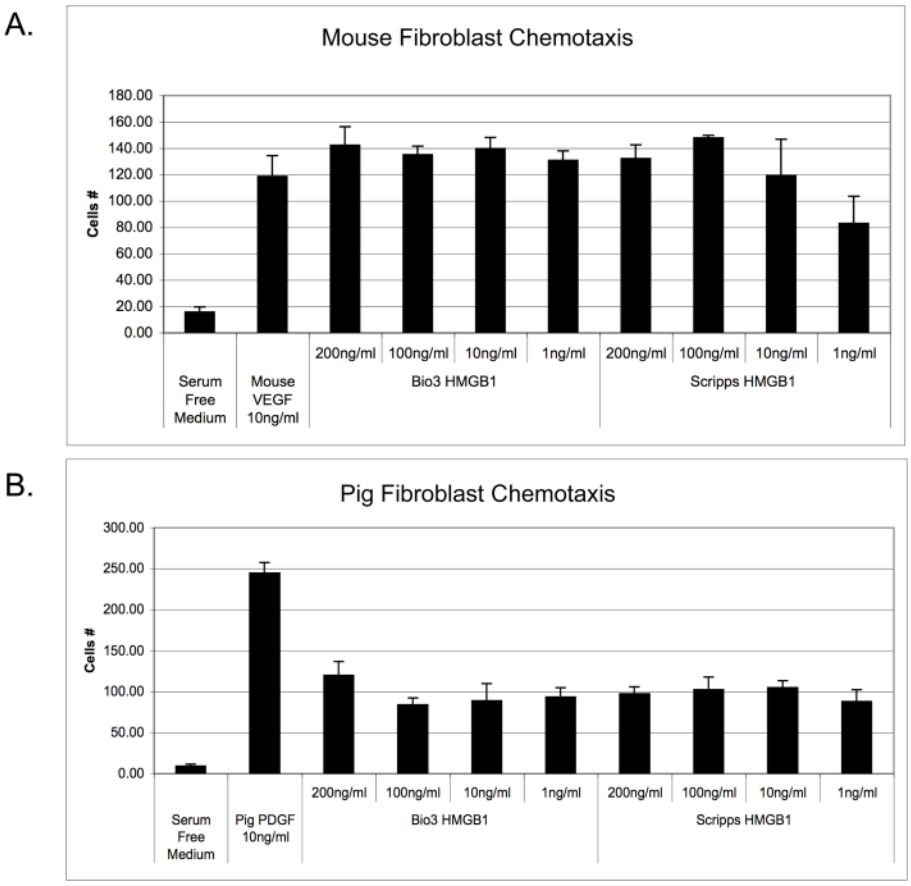


Figure 8. Bioactivity of HMGB1

Graph of the results from the fibroblast chemotaxis assay, as measured by the number of mouse (A) or pig (B) fibroblasts migrating towards the indicated chemokine is shown. Bioactivity of algal-expressed HMGB1 (Scripps) compared to commercial HMGB1 (Bio3), and to the controls mouse VEGF (A) or pig PDGF (B). Data represents mean and standard deviations of each treatment condition.

Table 1

Codon Adaptive Index (CAI) values for the native human sequences and the codon optimized sequences compared against the *C. reinhardtii* chloroplast codon usage table.

Gene	Corresponding amino acids	CAI native sequence	CAI codon optimized
EPO	aa 28–193	0.25	0.83
10FN3	aa 1447–1540	0.39	0.80
14FN3	aa 1723–1811	0.35	0.81
Inf β	aa 23–187	0.33	0.84
proinsulin	aa 25–147	0.24	0.77
VEGF isoform121	aa 27–147	0.30	0.79
HMGB1	aa 13–169	0.40	0.82