Multigenic engineering of the chloroplast genome in the green alga Chlamydomonas reinhardtii

Marco Larrea-Alvarez† and Saul Purton*

Abstract

The chloroplast of microalgae such as Chlamydomonas reinhardtii represents an attractive chassis for light-driven production of novel recombinant proteins and metabolites. Methods for the introduction and expression of transgenes in the chloroplast genome (=plastome) of C. reinhardtii are well-established and over 100 different proteins have been successfully produced. However, in almost all reported cases the complexity of the genetic engineering is low, and typically involves introduction into the plastome of just a single transgene together with a selectable marker. In order to exploit fully the potential of the algal chassis it is necessary to establish methods for multigenic engineering in which many transgenes can be stably incorporated into the plastome. This would allow the synthesis of multi-subunit proteins and the introduction into the chloroplast of whole new metabolic pathways. In this short communication we report a proof-of-concept study involving both a combinatorial and serial approach, with the goal of synthesizing five different test proteins in the C. reinhardtii chloroplast. Analysis of the various transgenic lines confirmed the successful integration of the transgenes and accumulation of the gene products. However, the work also highlights an issue of genetic instability when using the same untranslated region for each of the transgenes. Our findings therefore help to define appropriate strategies for robust multigenic engineering of the algal chloroplast.

Current commercial production of recombinant molecules such as therapeutic proteins and bioactive metabolites relies almost exclusively on heterotrophic cell platforms such as bacteria, yeasts and mammalian cell lines [1]. However, there is increasing interest in the exploitation of eukaryotic microalgae as low-cost, phototrophic platforms [2, 3]. Green algal species such as Chlamydomonas reinhardtii, Chlorella vulgaris and Haematococcus pluvialis can be cultivated at scale in simple photobioreactor systems using a basic growth medium [4, 5]. Furthermore, these species have GRAS status and do not harbor harmful endotoxins or viral or prion contaminants, thereby simplifying the downstream processing steps [6]. In addition, plants and algae possess a unique biosynthetic and storage organelle – the chloroplast – that is not present in animal or fungal cells, and which is the site of synthesis of key metabolites such as fatty acids, terpenoids, carbohydrates and tetrapyrroles, as well as the major proteins of the photosynthetic apparatus [7]. As the chloroplast contains its own minimal genome (=plastome) and genetic system that is derived from its cyanobacterial ancestor [8], this organelle represents an attractive sub-cellular ‘chassis’ on to which can be bolted new metabolic pathways through chloroplast genetic engineering [9, 10].

DNA transformation of the algal chloroplast has been reported for a handful of species, but the technology is most advanced for C. reinhardtii [11]. Numerous studies have shown that insertion of foreign DNA into the plastome occurs exclusively via homologous recombination allowing precise and predictive targeting of transgenes into specific loci [12]. High-level expression of the transgenes is achieved by fusing them to cis elements (promoters and untranslated regions) of highly expressed chloroplast genes, and transgene expression is stable in the absence of selection since there are no gene silencing mechanisms in the chloroplast. Methods have been developed for generating marker-free transgenic lines [13] and for regulated expression of the transgenes [9], together with a codon-reassignment method that ensures their biocontainment [14]. In addition, we are now starting to see the application...
of synthetic biology strategies for the rapid and standardised assembly of expression cassettes [3].

Our survey of the literature has revealed that over 100 different foreign proteins have been produced successfully in the *C. reinhardtii* chloroplast. These include a wide range of therapeutic proteins such as vaccines, hormones and antibodies [11]; industrial enzymes [15, 16], and enzymes for synthesizing novel metabolites in the organelle [17–21]. However, in almost all cases the genetic engineering has involved the insertion of just one transgene (or a single transgene together with a bacterial gene such as *aadA* or *aphA-6* as the selectable marker [13]). In the few remaining reports, two transgenes have been employed – either as a dicistronic operon [22]; as two linked gene cassettes inserted into
one site of the plastome [18], or as independent gene cassettes introduced by co-transformation into two separate loci [23]. In order to fully exploit the algal chloroplast as an expression platform there is a need to advance the genetic engineering technology such that multiple transgenes can be introduced into the plastome. This would allow the biosynthesis of multi-subunit complexes such as the bacterial nitrogenase enzyme (a long-standing goal in chloroplast engineering [24]) and elaborate metabolic engineering involving multiple enzyme steps.

In this study we have explored multigenic engineering by combining two approaches: namely, the integration of three gene cassettes into a single neutral locus, followed by a further round of transformation to integrate a fourth gene cassette plus the \( \textit{aadA} \) marker cassette at a second locus. The four genes were selected from previous studies in our group where each coding sequence was codon-optimised for expression in the chloroplast, and each was modified to encode the haemagglutinin epitope tag (HA-tag), YPYDVPDYA at the C terminus [25]. The gene products are unrelated, and were chosen based on their different sizes and the different levels of accumulation observed previously in chloroplast transformants expressing just the single gene. The four genes are: (i) \( \textit{splB} \) encoding a serine protease from the bacterium \textit{Staphylococcus aureus} that cleaves at the recognition sequence W-E-L-X [26]. SplB accumulates to a high level in the chloroplast (>1% total soluble protein) and has application as an oral enzyme for \( \textit{in vivo} \) processing of recombinant proteins [27]; (ii) \( \textit{codA} \) encoding a variant of \textit{E. coli} cysteine deaminase that can be exploited as a negative marker in the \textit{C. reinhardtii} chloroplast catalyzing the conversion of 5-fluorocytosine into the toxic 5-fluorouracil [25]; (iii) \( \textit{ibv-ctb} \), designed to encode a chimeric protein that could serve as an oral vaccine for the poultry pathogen, infectious bronchitis virus (IBV). The protein comprises three antigenic sections of the viral N protein [28] fused to the \( \beta \)-subunit of cholera toxin (CTB), which can serve as an immunological adjuvant [11]; (iv) \( \textit{cpl-1} \) encoding a bacteriophage endolysin that targets the bacterial pathogen \textit{Streptococcus pneumoniae} and has been shown to be active when synthesized in the \textit{C. reinhardtii} chloroplast [29].

As shown in Fig. 1a, the first three genes were assembled into a single plasmid (pICS) that is based on the chloroplast expression vector pASapI [30]. Details of this vector and the cloning strategy used are given in the supplementary data (available in the online version of this article). A key feature of pASapI is that it carries a wild–type copy of the essential photosystem II gene, \( \textit{psbH} \) on the right homology arm. Consequently, \( \textit{psbH} \) can be used as a selectable marker for phototrophic restoration of \( \Delta \textit{psbH} \) strains such as TN72 in which \( \textit{psbH} \) has been deleted using the \( \textit{aadA} \) cassette. Transformant colonies are selected on minimal medium and scored for loss of the spectinomycin resistance conferred by \( \textit{aadA} \) [30]. When building each of the cassettes, the three coding sequences were fused to a promoter/3′UTR element from a different endogenous gene (namely, \( \textit{psaA-1} \), \( \textit{petB} \) and \( \textit{atpA} \)) in order to minimise the competition for \textit{trans}-acting factors required for mRNA stability and translation initiation that bind to the 5′UTR [31]. Conversely, the same 3′UTR element from \( \textit{rbcL} \) was used in each case since the choice of 3′UTR has little effect on the level of transgene expression [32]. However, it is important to note (as discussed below) that the \( \textit{codA} \) cassette used in the making of pICS contained a newer and shorter version of the \( \textit{rbcL} \) 3′UTR element (i.e. 258 bp versus 407 bp) in which

---

**Fig. 2.** Western blot analysis demonstrates accumulation of the three recombinant proteins. Cell extracts of the four multigenic lines (ICS1–4) were analysed using primary antibodies to the HA epitope tag on each recombinant protein as described [25]. Antibodies to the chloroplast protein RbcL were used as a loading control. For comparison, a negative control transformant (rTN72) and three ‘single transgene’ transformant lines of TN72 were included in the analysis (labelled ‘I only’, etc.). The recombinant proteins and their predicted molecular weights are indicated together with an unknown endogenous protein (*) that is detected by the \( \alpha\)-HA antibodies.
unnecessary sequence spanning the end of the \textit{rbcL} coding region was removed [25].

Four phototrophic colonies (ICS1 – ICS4) were selected following glass-bead mediated transformation of TN72 [33] with plasmid pICS, and were restreaked twice to single colonies to obtain homoplasmic transformants. As shown in Fig. 1b, PCR analysis confirmed the integration of the three-cassette cluster into the plastome. The four transformant lines were then subjected to Western blot analysis in which cell lysates were fractionated by SDS polyacrylamide gel electrophoresis and probed using an antibody against the HA epitope present on each of the three recombinant proteins (Fig. 2). Comparison of the four lines to control lines containing just one of the transgene cassettes or an empty cassette lacking any coding sequence (strain rTN72) confirmed that all four are producing the three recombinant proteins, with high levels of CodA and SplB and a much lower level of IBV-CTB. This low level probably reflects a higher rate of protein turnover rather than a lower rate of expression, since the \textit{psaA-1} element used to drive \textit{ibv-ctb} expression is actually stronger than the \textit{atpA} element driving \textit{splB} expression [30]. The artificial and composite structure of IBV-CTB (three epitope regions from the N protein together with the CTB protein fused into a single polypeptide) probably limits folding into a stable, protease-resistant form, unlike CodA and SplB which are two natural bacterial enzymes. The observed reduction in RbcL levels in the ICS transformants compared to the single transformants (Fig. 2) is possibly due to the presence of multiple copies of the \textit{rbcL} 3′UTR resulting in the titrating out of \textit{trans}-acting factors required for 3′ processing and stability of the \textit{rbcL} transcript [34]. This reduced level is also reflected in a small reduction in growth performance of the ICS transformant compared to the rTN72 control (Fig. S1).
Having demonstrated that three foreign proteins could be synthesized in a single transgenic line, we then sought to further engineer this line by addition of two more transgenes at a second locus. We therefore built a second transformation plasmid (named pAP) that was designed to target the *aadA* cassette that confers spectinomycin-resistance [35] and a cassette for expression of the endolysin gene *cpl-1* [29] into the middle of the non-essential chloroplast gene, *chlL*. This gene encodes a subunit of the dark-operative protochlorophyllide reductase and a knockout of *chlL* gives rise to a ‘yellow-in-the-dark’ phenotype [36]. Consequently, the insertion of transgenes into *chlL* can be used as a simple screen to distinguish genuine *aadA* transformants from spectinomycin-resistant clones arising from spontaneous mutation [13].

The pAP plasmid was used to transform line ICS3 to spectinomycin resistance as previously described [33]. Three colonies (ICS3-AP #1–3) were selected; confirmed to be transformants by their yellow-in-the-dark phenotype, and checked for the correct insertion of the two gene cassettes by similar PCR analysis to that carried out in Fig. 1. Since the *cpl-1* gene was designed to encode an HA-tagged version of Cpl-1 [29], we anticipated that Western analysis of the ICS3-AP transformants using the anti-HA antibodies would demonstrate the synthesis of four different recombinant proteins: namely, Cpl-1, together with *IBV-CTB*, *CodA* and *SplB*. However, only Cpl-1 and *IBV-CTB* could be detected in the three transformant lines despite *CodA* and *SplB* being readily detectable in the ICS3 recipient line (Fig. 3b). We therefore reasoned that the cluster of three cassettes was genetically unstable, and that selection and re-streaking of the *aadA* transformants resulted in the loss or rearrangement of the *codA* and *splB* cassettes. This hypothesis was confirmed by PCR analysis and sequencing of the PCR products, which showed that homologous recombination between the two longer copies of the *rbcL* 3’UTR within the *ibv-ctb* and *splB* cassettes had resulted in the loss of both *codA* and *splB* (Fig. 3a). As seen in Fig. 3c, PCR analysis with primer F2 and R5 gave rise to a 1.0 kb band rather than the 5.0 kb band expected for the intact gene cluster. Furthermore, no 2.4 kb band was seen when an alternative reverse primer (R4) from within the *splB* gene was used with F2, suggesting that the transformants were homoplasmic with all plastome copies carrying the *codA-splB* deletion. For the ICS3 recipient line, the 2.4 kb band was obtained but so was the 1.0 kb band. This indicates that recombination between the *rbcL* direct repeats is occurring in the ICS3 strain such that a heteroplasmic population of plastomes exists, and that the selection for individual transformants drives this to a homoplasmic state since there is no counter-selection to retain the intact cluster. Interestingly, despite the reduced ploidy of the intact cluster in ICS3 due to this heteroplasmicity, we still achieve levels of *SplB* and *CodA* accumulation in ICS3 (and the other ICS lines) that are approximately 50–90% of that seen for the homoplasmic transformant lines ‘C only’ and ‘S only’ (Fig. 2). This observation supports the findings of Eberhard *et al.* [37] that chloroplast protein synthesis is relatively insensitive to changes in gene copy number, and a marked reduction of plastome ploidy has only a small effect on protein synthesis rates. Finally, our PCR analysis did not reveal any evidence of recombination between either of the two longer (407 bp) copies of the *rbcL* 3’UTR and the shorter (258 bp) copy on *codA*. This suggests that there is a lower size limit, somewhere between 0.25 kb and 0.4 kb, for efficient intramolecular homologous recombination in the *C. reinhardtii* chloroplast. This range correlates well with an earlier study of different sized direct repeats in the algal plastome where Fischer *et al.* [38] found that a 230 bp repeat did not give rise to recombination whereas repeats of 483 bp and 832 bp did.

In conclusion, this short report demonstrates that multiple transgene cassettes can be inserted into the *C. reinhardtii* plastome to allow ever-more complex genetic engineering such as that required for introduction of novel metabolic pathways. A cluster of cassettes can be targeted to a chosen locus in a single transformation event, and then further clusters can be introduced into additional loci by serial transformation. However, a clear caveat is that the endogenous *cis* elements (promoters, 5’UTRs and 3’UTRs) used to drive expression of the transgenes should not be used multiple times, and/or should be smaller than ~0.2 kb to limit recombination between each other and between the endogenous gene itself. Alternately, heterologous elements from other green algae or from cyanobacteria could be used where the primary DNA sequence is sufficiently different to prevent recombination. Several studies have demonstrated that heterologous promoters and 3’UTRs are functional in the algal chloroplast, whereas 5’UTRs are much more species-specific [39, 40]. Nevertheless, directed mutagenesis of *C. reinhardtii* 5’UTRs could be employed as an alternative strategy to both reduce their homology and improve their performance [41]. With these considerations in mind, we are currently developing a DNA parts library for rapid and standardized assembly of transgene cassettes based on the Start-Stop Assembly method described recently by Taylor *et al.* [42].

**Funding information**

This study was funded by grants BB/L002957/1 and BB/R016534/1 from the UK’s Biotechnology and Biological Sciences Research Council. ML-A was supported by Ecuador’s SENESCYT (Secretariat of Higher Education, Science, Technology and Innovation) under the programme ‘Top World Universities’.

**Conflicts of interest**

The authors declare that there are no conflicts of interest.

**References**

1. Sanchez-Garcia L, Martin L, Mangues R, Ferrer-Miralles N, Vázquez E *et al.* Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microb Cell Fact* 2016;15:33.
2. Hempel F, Maier UG. Microalgae as solar-powered protein factories. *Adv Exp Med Biol* 2016;896:241–262.
3. Taunt HN, Stoffels L, Purton S. Green biologics: the algal chloroplast as a platform for making biopharmaceuticals. *Bioengineered* 2018;9:48–54.
4. Gimpe J, Hyun JS, Schoepp NG, Mayfield SP. Production of recombinant proteins in microalgae at pilot greenhouse scale. *Biotechnol Bioeng* 2015;112:329–346.
5. Borowitcka MA, Vonshak A. Scaling up microalgal cultures to commercial scale. *Eur J Phycol* 2017;52:407–418.

6. Stoffels F, Finlan A, Mannall G, Purton S, Parker B. Downstream processing of *Chlamydomonas reinhardtii* for recombinant endolysin production. *Front Bioeng Biotechnol* 2019.

7. Rolland N, Curien G, Finazzi G, Kuntz M, Marestal E et al. The biosynthetic capacities of the plastids and integration between cytoplasmic and chloroplast processes. *Annu Rev Genet* 2012;46:233–264.

8. Green BR. Chloroplast genomes of photosynthetic eukaryotes. *Plant J* 2011;66:34–44.

9. Purton S, Szaub JB, Wannathong T, Young R, Economou CK. Genetic engineering of algal chloroplasts: progress and prospects. *Russ J Plant Physiol* 2013;60:491–499.

10. Siddiqui A, Wei Z, Boehm M, Ahmad N. Engineering microalgae through chloroplast transformation to produce high-value industrial products. *Biotechnol Appl Biochem* 2019;1.

11. Dyo YM, Purton S. The algal chloroplast as a synthetic biology platform for production of therapeutic proteins. *Microbiology* 2018;164:113–121.

12. Purton S. Tools and techniques for chloroplast transformation of *Chlamydomonas*. *Adv Exp Med Biol* 2007;616:34–45.

13. Esland L, Larrea-Alvarez M, Purton S. Selectable markers and reporter genes for engineering the chloroplast of *Chlamydomonas reinhardtii*. *Biotechnology* 2018;7:E46.

14. Young REB, Purton S. Codon reassignment to facilitate genetic engineering and biobtainment in the chloroplast of *Chlamydomonas reinhardtii*. *Plant Biotechnol J* 2016;14:1251–1260.

15. Faë M, Accossato S, Celli R, Fontana F, Goldschmidt-Clermont M et al. Comparison of transplastomic *Chlamydomonas reinhardtii* and *Nicotiana tabacum* expression systems for the production of a bacterial endoglucanase. *Appl Microbiol Biotechnol* 2017;101:4085–4092.

16. Richter LV, Yang H, Yazdani M, Hanson MR, Ahner BA. A downstream box fusion allows stable accumulation of a bacterial cellulase in *Chlamydomonas reinhardtii* chloroplasts. *Biotechnol Biofuels* 2018;11:133.

17. Fukusaki E-I, Nishikawa T, Kato K, Shinmyo A, Hemmi H et al. Introduction of the archaeobacterial geranylgeranyl pyrophosphate synthase gene into *Chlamydomonas reinhardtii* chloroplast. *J Biosci Bioeng* 2003;95:283–287.

18. Papaethimious D, Diretto G, Demurtas OC, Mini P, Ferrante P et al. Heterologous production of labdane-type diterpenes in the green alga *Chlamydomonas reinhardtii*. *Phytochemistry* 2019;167:112082.

19. Tevatia R, Payne S, Allen J, White D, Clemente TE et al. A synthetic codon/strand taurine pathway in the green unicellular alga *Chlamydomonas reinhardtii*. *Algal Research* 2019;40:101491.

20. Gangi D, Zedler JAZ, Wlodarczyk A, Jensen PE, Purton S et al. Expression and membrane-targeting of an active plant cytochrome P450 in the chloroplast of the green alga *Chlamydomonas reinhardtii*. *Phytochemistry* 2015;110:22–28.

21. Paik S-M, Kim J, Jin E, Jeon NL. Overproduction of recombinant E. coli malate synthase enhances *Chlamydomonas reinhardtii* biomass by upregulating heterotrophic metabolism. *Bioresour Technol* 2019;272:594–598.

22. Su Z-L, Qian K-X, Tai C-P, Meng C-X, Qin S. Recombination and heterologous expression of allophyocyanin gene in the chloroplast of *Chlamydomonas reinhardtii*. *Acta Biochim Biophys Sin* 2005;37:709–712.

23. 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;104:663–673.

24. Yang J, Xie X, Yang M, Dixon R, Wang Y-P. Modular electron-transport chains from eukaryotic organelles function to support nitrogenase activity. *Proc Natl Acad Sci U S A* 2017;114:E2460–E2465.

25. Young REB, Purton S. Cytosine deaminase as a negative selectable marker for the microalgal chloroplast: a strategy for the isolation of nuclear mutations that affect chloroplast gene expression. *Plant J* 2014;80:915–925.

26. Dubin G, Stec-Niemczyk J, Kisielewska M, Pustelny K, Popowicz GM et al. Enzymatic activity of the *Staphylococcus aureus* SpIB serine protease is induced by substrates containing the sequence Trp-Glu-Leu-Gln. *J Mol Biol* 2008;379:343–356.

27. Larrea-Alvarez M. Molecular tools and approaches for increasing complexity of transplastomic engineering in *Chlamydomonas reinhardtii*. PhD thesis, UCL (University College London). 2018. https://discovery.ucl.ac.uk/id/eprint/10047617/.

28. Cao H-P, Wang H-N, Zhang A-Y, Ding M-D, Liu S-T et al. Expression of avian infectious bronchitis virus multi-epitope based peptide EPIC in *Lactococcus lactis* for oral immunization of chickens. *Biosci Biotechnol Biochem* 2012;76:1871–1876.

29. Stoffels L, Taunt HN, Charalambous B, Purton S. Synthesis of bacteriophage lytic proteins against *Streptococcus pneumoniae* in the chloroplast of *Chlamydomonas reinhardtii*. *Plant Biotechnol J* 2017;15:1130–1140.

30. Wannathong T, Waterhouse JC, Young REB, Economou CK, Purton S. New tools for chloroplast genetic engineering allow the synthesis of human growth hormone in the green alga *Chlamydomonas reinhardtii*. *Appl Microbiol Biotechnol* 2016;100:5467–5477.

31. Choquet Y, Wollman FA. Translational regulations as specific traits of chloroplast gene expression. *FEBS Lett* 2002;529:39–42.

32. Barnes D, Franklin S, Schultz J, Henry R, Brown E et al. 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.

33. Economou C, Wannathong T, Szaub J, Purton S. A simple, low-cost method for chloroplast transformation of the green alga *Chlamydomonas reinhardtii*. *Methods Mol Biol* 2014;1132:401–411.

34. Goldschmidt-Clermont M, Rahire M, Rochaih J-D. Redundant cis-acting determinants of 3′ processing and RNA stability in the chloroplast *rbcL* mRNA of *Chlamydomonas*. *Plant J* 2008;53:566–577.

35. Goldschmidt-Clermont M. Transgenic expression of aminoglyco-side adenine transferase in the chloroplast: a selectable marker of site-directed transformation of *Chlamydomonas*. *Nucleic Acids Res* 1991;19:4083–4089.

36. Cheng Q, Day A, Dowson-Day M, Shen G-F, Dixon R. The Klebsiella pneumoniae nitrogenase Fe protein gene (*nifH*) functionally substitutes for the *chiL* gene in *Chlamydomonas reinhardtii*. *Biochem Biophys Res Commun* 2005;329:966–975.

37. Eberhard S, Drapier D, Wollman F-A. 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.

38. Fischer N, Stampacchia O, Redding K, Rochaix JD. Selectable marker recycling in the chloroplast. *Mol Genet Gen* 1996;251:373–380.

39. Gimpel JA, Mayfield SP. Analysis of heterologous regulatory and coding regions in algal chloroplasts. *Appl Microbiol Biotechnol* 2013;97:4499–4510.

40. Kramzar LM, Mueller T, Erickson B, Higgs DC. Regulatory sequences of orthologous petD chloroplast mRNAs are highly specific among *Chlamydomonas* species. *Plant Mol Biol* 2006;60:405–422.

41. Specht EA, Mayfield SP. Synthetic oligonucleotide libraries reveal novel regulatory elements in *Chlamydomonas* chloroplast mRNAs. *ACS Synth Biol* 2013;2:34–46.

42. Taylor GM, Mordaka PM, Heap JT. Start-Stop assembly: a functionally scarless DNA assembly system optimized for metabolic engineering. *Nucleic Acids Res* 2019;47:e17.