Enhanced chloroplast transgene expression in a nuclear mutant of Chlamydomonas

Michelet, Laure; Legendre Lefebvre, Linnka; Burr, Sarah E; Rochaix, Jean-David; Goldschmidt-Clermont, Michel P.

How to cite

MICHELET, Laure et al. Enhanced chloroplast transgene expression in a nuclear mutant of Chlamydomonas. In: Plant biotechnology journal, 2011, vol. 9, n° 5, p. 565–574. doi: 10.1111/j.1467-7652.2010.00564.x

This publication URL: https://archive-ouverte.unige.ch//unige:18377

Publication DOI: 10.1111/j.1467-7652.2010.00564.x

Enhanced chloroplast transgene expression in a nuclear mutant of Chlamydomonas.

Laure Michelet[1,+], Linnka Lefebvre-Legendre[1], Sarah E. Burr[2,*], Jean-David Rochaix[1] and Michel Goldschmidt-Clermont[1]

(1) Departments of Plant Biology and of Molecular Biology, University of Geneva, Quai E. Ansermet 30, 1211 Genève 4, Switzerland
(2) Institute for Veterinary Bacteriology, Universität Bern, Länggasstrasse 122, 3012 Berne, Switzerland.

(+) Present address: CEA Saclay, iBiTec-S, CNRS URA 2096, Service de Bioénergétique Biologie Structurales et Mécanismes, 91191 Gif-sur-Yvette Cedex, France.
(*) Present address: London School of Hygiene and Tropical Medicine, Medical Research Council Laboratories, P.O. Box 273, Atlantic Road, Fajara, Banjul, The Gambia, West Africa

Author for correspondence:
Michel Goldschmidt-Clermont
University of Geneva
Department of Molecular Biology
Sciences III
30, Quai E. Ansermet
1211 Genève 4
Switzerland

Phone : +41 22 379 6188
Fax : +41 22 379 6868
Email: michel.goldschmidt-clermont@unige.ch

Emails of other authors:
Laure Michelet : laure.michelet@cea.fr
Linnka Legendre : linnka.legendre@unige.ch
Sarah E. Burr : sburr@mrc.gm
Jean-David Rochaix : jean-david.rochaix@unige.ch

Running title: Enhanced chloroplast transgene expression

Keywords : plastid / chloroplast transformation / anterograde signaling / microalgae / *Chlamydomonas reinhardtii*

Word count : 8856 words, 58’260 characters (with spaces)

SUMMARY

Chloroplast transformation in microalgae offers great promise for the production of proteins of pharmaceutical interest or for the development of novel biofuels. For many applications, high level expression of transgenes is desirable. We have transformed the chloroplast of *Chlamydomonas reinhardtii* with two genes, *acrV* and *vapA*, which encode antigens from the fish pathogen *Aeromonas salmonicida*. The promoters and 5' untranslated regions of four chloroplast genes were compared for their ability to drive expression of the bacterial genes. The highest levels of expression were obtained when they were placed under the control of the *cis*-acting elements from the *psaA-exon1* gene. The expression of these chimeric genes was further increased when a nuclear mutation that affects a factor involved in *psaA* splicing was introduced in the genetic background of the chloroplast transformants. Accumulation of both the chimeric mRNAs and the recombinant proteins was dramatically increased, indicating that negative feedback loops limit the expression of chloroplast transgenes. Our results demonstrate the potential of manipulating anterograde signaling to alter negative regulatory feedback loops in the chloroplast and improve transgene expression.

INTRODUCTION

Microalgae are attracting increasing interest for biotechnological applications such as the expression of recombinant proteins with pharmaceutical value (pharming) or the production of biomass and biofuels (Mayfield et al., 2007; Posten and Schaub, 2009; Specht et al., 2010; Walker et al., 2005). As a model for microalgae, the green unicellular eukaryote *Chlamydomonas reinhardtii* offers sophisticated genetic and genomic tools, and the possibility of readily transforming the three genomes in the nucleus, chloroplast and mitochondria (Harris, 2009; Merchant et al., 2007). Chlamydomonas has been used for the expression of various antigens to develop recombinant vaccines, and for the production of human antibodies and other proteins (Daniell et al., 2009; Dreesen et al., 2009; Griesbeck et al., 2006; Rasala et al., 2010; Walker et al., 2005). Its versatile metabolism and ability to produce hydrogen under certain anaerobic conditions offers great promise for the development of biofuels (Grossman et al., 2007; Rupprecht, 2009).

Chloroplast transformation has many advantages for engineering of algae or plants (Bock and Warzecha, 2010; Chebolu and Daniell, 2009; Daniell et al., 2009; Mayfield et al., 2007). The organelle is of central importance in metabolism, as it is the site of photosynthesis and of many other important biochemical pathways, for example the biosynthesis of lipids and amino acids. Transformation of its comparatively small genome proceeds by homologous recombination, so that insertion of transgenes can be predictably designed. Chloroplast transgene expression is not subjected to position effects, transcriptional silencing or RNA interference, which can affect nuclear genes. Moreover, plastid genomes are inherited uniparentally in some species, allowing a certain degree of transgene containment.

Chloroplast gene expression is controlled at the level of transcription but also to a large degree at post-transcriptional steps. In Chlamydomonas, the latter include RNA processing from poly-cistronic precursors, RNA splicing, RNA degradation and translation, but not RNA editing. These processes are mediated by a large number of nucleus-encoded proteins that specifically target single chloroplast genes or small subsets of genes (Barkan and Goldschmidt-Clermont, 2000; Stern et al., 2010). These factors are the basis for anterograde signaling that enables the nucleo-cytoplasmic compartment to

control chloroplast gene expression. On the other hand, chloroplast gene expression is also regulated by chloroplast-derived signals (Pfannschmidt, 2003). The assembly of the photosynthetic complexes requires the coordinate synthesis of numerous subunits, pigments and cofactors. In a regulatory mechanism called CES (Control by Epistasy of Synthesis), unassembled subunits exert negative feed-back inhibition on the translation of their own mRNA (Choquet and Wollman, 2002).

The level of expression of transgenic proteins in plant chloroplasts is dependent on the particular gene of interest and can reach very high levels (Bock and Warzecha, 2010; Daniell et al., 2009), with for example 70 % total soluble protein in tobacco chloroplasts for a phage-derived endolysin (Bock and Warzecha, 2010; Oey et al., 2009) and 72 % total leaf protein for CTB-Pins (Cholera Toxin B-subunit - human Pro-insulin) (Ruhlman et al., 2010). However transgenic protein levels in the chloroplast of Chlamydomonas are generally an order of magnitude lower than in plants, suggesting that there is space for improvement (Manuell et al., 2007; Mayfield et al., 2007; Rasala et al., 2010; Surzycki et al., 2009). To maximize transgene expression in the chloroplast of this alga, many reports have shown the importance of *cis*-acting elements such as promoters and 5' untranslated regions (5'UTRs) which drive efficient transcription and translation (Barnes et al., 2005; Ishikura et al., 1999). In the Chlamydomonas chloroplast, the levels of expression of chimeric transgenes driven by the *psbA cis*-acting elements were higher when the endogenous *psbA* gene was deleted, an effect that was attributed to competition for activators and to negative feedback regulatory loops (Manuell et al., 2007; Minai et al., 2006; Rasala et al., 2010; Surzycki et al., 2009). *Trans*-acting factors that participate in anterograde signaling have been used to engineer inducible or repressible gene expression systems in the chloroplast (Kato et al., 2007; Surzycki et al., 2007). However, the possibility of manipulating such factors to maximize the expression of foreign genes in the plastids remains largely unexplored. Here we address the effects of both *cis*-acting elements and *trans*-acting factors to improve expression of foreign genes in the chloroplast of *C. reinhardtii*. We identify the *psaA-exon1* promoter and its 5'UTR as efficient *cis*-acting elements to drive the expression of chimeric genes. We further show that the expression of *psaA*-driven transgenes in the chloroplast is enhanced in a nuclear

mutant deficient for a *trans*-acting factor that is imported into the chloroplast where it is involved in *psaA* RNA splicing.

RESULTS

Chlamydomonas chloroplast transformation

We transformed the chloroplast of Chlamydomonas with the genes for two antigens, *vapA* and *acrV*, from the fish pathogen *Aeromonas salmonicida*. To facilitate their translation, the two genes were entirely re-synthesized with the codon usage of *C. reinhardtii* chloroplast genes. To drive the expression of the transgenes, highly expressed chloroplast genes were chosen as a source of *cis*-acting regulatory sequences. Chimeric genes were constructed with the promoter and 5' untranslated region (P/5'UTR) of *atpA*, *psbA* and *psbD*, which were previously reported to allow high levels of expression for various transgenes (Barnes et al., 2005; Goldschmidt-Clermont, 1991; Ishikura et al., 1999). The P/5'UTR of *psaA-exon1*, which was not previously compared to other P/5'UTRs for its ability to drive efficient transgene expression, was also included. The 3' UTR of *rbcL* was used in all the chimeric constructs because it features two redundant *cis*-acting elements that mediate efficient RNA processing and stability (Goldschmidt-Clermont et al., 2008). A previous comparative study has shown that the 3'UTR has only a minor influence on transgene expression (Barnes et al., 2005).

Two series of transformation vectors were constructed to target insertion of the transgene by homologous recombination either downstream of the *atpB* gene (atpB-int vectors; Fig 1A) or in the intergenic region between the ribosomal RNA genes and *psbA* (IR-int vectors; Fig 1B). In addition to the chimeric transgene, the atpB-int vectors also contain the wild-type *atpB* gene which allows selection for photoautotrophic growth after transformation of a host strain deleted for *atpB* (*FUD50*) (Goldschmidt-Clermont et al., 1991). The IR-int vectors target the duplicated part of the chloroplast genome known as the inverted repeat (IR), so that in the transformed lines the transgene is present in two copies per genome. They contain an *aadA* cassette which permits selection on media containing spectinomycin or streptomycin (Goldschmidt-Clermont, 1991). This "recyclable" marker is flanked by two direct repeats (483 bp of bacterial DNA + 3'UTR of *rbcL*) that allow a spontaneous excision of the antibiotic resistance gene by

homologous recombination when the selective pressure is released (Fischer et al., 1996). The IR-int vectors exist in two versions, (F) and (R), which only differ by the orientation of the insert (Fig 1B). Both the atpB-int and the IR-int vectors were used to generate transgenic lines that are free of selectable markers. The integration of the transgene and the homoplasmicity of all transformed lines were assessed by PCR.

Comparison of chloroplast promoters and 5'UTRs

We determined the levels of expression of the transgenes by preparing total protein extracts which were analyzed by SDS-PAGE and immunoblotting. The transformants obtained with the IR-int vectors, which insert in the two copies of the inverted repeat, generally gave higher levels of expression than those obtained with the atpB-int vectors, which insert in the single-copy *atpB* locus (Table S1). For both VapA and AcrV, the highest levels of expression were obtained with the P/5'UTR of *psaA-exon1*. Somewhat lower levels were found with the P/5'UTR of *atpA* and *psbD*, but only very low levels were obtained with the P/5'UTR of *psbA* (Fig 2). For reasons detailed in Supporting Information (Fig S1), some *atpA::vapA* lines lost the transgene (Fig 2A, line #8Δ) and no homoplasmic *atpA::acrV* lines were obtained.

It has been reported that higher levels of expression can be reached if the transgene is fused to a small part of the coding region of a chloroplast gene rather than to its AUG translation start codon (Kasai et al., 2003). In the *atpA::vapA* lines the transgene is fused to the first 26 codons of the *atpA* gene and is expressed at comparatively high levels (Fig2A; Table S1). We tested this more systematically for the *psaA::vapA* and *psaA::acrV* constructs. We included the first 20 or 30 codons of *psaA-exon1* in the constructs and compared them to the corresponding fusions at the translation start codon. However, the resulting transformed lines showed lower levels of expression than lines where the transgenes were fused directly to the start codon (Fig S2). Thus the advantageous effect of using a fusion of the transgene to the coding region of a chloroplast gene is not always realised and appears to depend on the particular genes that are used.

Integration by homologous recombination gives reproducible expression

In independent transformants obtained with a given construct, we observed very homogeneous levels of expression, as exemplified for *psaA::vapA* (Fig 3A). This is in contrast with some previous reports, where highly variable levels were observed in different transformants (Dreesen et al., 2009; Mayfield and Schultz, 2004; Surzycki et al., 2009). To try to reconcile this discrepancy, we used the IR-int vector with *psaA::vapA* to transform a mutant of Chlamydomonas with a deletion of the *psbA* gene in both copies of the inverted repeat (*FUD7*, Fig 1B), as described in the previous work (Mayfield and Schultz, 2004; Surzycki et al., 2009). In this case, we did observe large variations in transgene expression between different transformed lines (Fig 3B, Table S1). Mapping of *FUD7* by PCR and sequencing (U. Johanningmeier, personal communication and our unpublished data) showed that the deletion extends across the insertion site of the IR-int vector so that homologous recombination is only possible on one side of the transgene (Fig 1B, *FUD7*, HR4). Thus illegitimate recombination events may be required on the other side of the transgene to allow its stable integration. In contrast, when the same vector was used to transform the wild-type strain, homologous recombination on both sides of the transgene (Fig 1B, HR3 and HR4) led to a predictable genomic structure and to homogeneous expression levels (Fig 3A). The highest level of expression was obtained in one of the *FUD7* transformants (line #8). A similar comparison with *psaA::acrV* also showed reproducible levels of expression in the wild-type transformants and large variations in the *FUD7* context (Fig S3).

Increased chloroplast *trans*-gene expression in a nuclear mutant

In *C. reinhardtii*, the *psaA* mRNA is spliced in *trans* from three separate precursors. In mutants defective in *psaA trans*-splicing, the precursor of *psaA-exon1* over-accumulates (Choquet et al., 1988). Thus we reasoned that a chimeric mRNA with the P/5'UTR of *psaA-exon1* might also be present at higher levels. Furthermore, unassembled PsaA subunits inhibit the translation of their own mRNA, a regulatory process known as CES (Control by Epistasy of Synthesis) (Wostrikoff et al., 2004). Hence in a *trans*-splicing deficient mutant where PsaA cannot be produced, this feedback inhibition (CES) is circumvented and the translation of a reporter expressed under the *psaA-exon1* P/5'UTR is expected to be higher.

We tested whether these effects would apply to *vapA* and *acrV* transgenes driven by the *psaA-exon1* P/5'UTR in the context of a nuclear mutant deficient for *psaA trans*-splicing, *raa-L121G* (Goldschmidt-Clermont et al., 1990). A transgenic line expressing *psaA::vapA* in the background of the *raa-L121G* mutant (mating type plus, mt+) was back-crossed to the wild type (mt-) (see Experimental procedures). In such a cross, the chloroplast transgene is inherited uniparentally from the mt+ parent in all the progeny, while the nuclear mutation segregates 2:2 in the tetrad. We observed the expected co-segregation of the *psaA* deficient phenotype with increased levels of VapA in six independent tetrads. In a representative tetrad (Fig 4A), two spores (DB and DD) showed a lack of PsaA and higher amounts of VapA. Compared to the original wild-type *psaA::vapA* line (#2), the PSI-deficient *raa-L121G* mutants (DB and DD) accumulated approximately 15 to 30 fold higher levels of VapA. RNA blot hybridization showed that the PSI-deficient progeny had the expected defect in *trans*-splicing (Fig 4B). In the wild type, *psaA-exon3* was part of the mature spliced mRNA, but in the *raa-L121G* mutant progeny, only the free *psaA-exon3* precursor was observed and the mature mRNA was not formed. The *psaA::vapA* chimeric mRNA, which had the *psaA-exon1* P/5'UTR, strongly over-accumulated in the mutant progeny, to levels that were approximately 7 times higher than in the wild-type background (see Experimental Procedures). Thus both higher levels of mRNA accumulation and higher levels of translation may account for the increase in VapA expression in the mutant background.

C. reinhardtii mutants with defects in PSI, such as *raa-L121G*, are sensitive to photo-oxidative damage and have to be grown in the dark, which might be disadvantageous for efficient gene expression. We compared the expression of *psaA::vapA* and *psaA::acrV* in the wild-type background, grown either in the dark or in the light, with the same constructs in the PSI-deficient *raa-L121G* mutant background, grown in the dark (Fig 5). The maximal expression levels were obtained when the cultures were in exponential phase (2 x 10^6 cells / mL), although the mutant cultures in the dark grew more slowly than the wild-type cultures in the light. In the wild-type strains, the expression of *psaA::vapA* and *psaA:acrV* was higher in the light than in the dark. In the *raa-L121G* mutant background, the expression of *psaA::vapA* (lines D, DB and DD) and of *psaA::acrV* (lines T6, T6B and T6C) was higher than in the wild-type

strains grown in the light. Thus the enhancing effect of the mutation on transgene expression was stronger than the stimulation by light. In this mutant background, VapA accumulated to approximately 0.3 % of total protein and AcrV to approximately 0.8 %.

Proteolysis limits the accumulation of VapA protein

There is more than a two-fold difference in the accumulation of VapA and AcrV expressed from the same regulatory sequences (Figs 2 and 5, Table S1). This could be due to a difference in gene expression, but also to differences in protein degradation which is known to influence transgene expression in higher plants and in Chlamydomonas (Birch-Machin et al., 2004; Doran, 2006; Leelavathi and Reddy, 2003; Lenzi et al., 2008; Stevens et al., 2000; Surzycki et al., 2009). To test this possibility, the stability of the proteins was investigated. Cultures of transgenic Chlamydomonas were split in three parts (Fig 6): one was used as a control, the second was treated with chloramphenicol to specifically inhibit chloroplast translation, and the third was treated with a cocktail of protease inhibitors. In the transgenic line expressing VapA, the recombinant protein disappeared rapidly when translation was arrested, but accumulated more than in the control when protein degradation was inhibited (Fig 6A). In contrast, the treatments had little effect on the accumulation of AcrV (Fig 6B). These results indicate that the expression of VapA is limited by proteolytic degradation in the transgenic lines.

DISCUSSION

In this work, we have developed vectors for chloroplast transformation that allow the generation of marker-free transgenic lines. The importance of *cis*-acting elements, such as promoters and 5'UTRs, for efficient transplastomic gene expression is well documented in vascular plants (Herz et al., 2005; Ruhlman et al., 2010; Ye et al., 2001; Zou et al., 2003) and in Chlamydomonas (Barnes et al., 2005; Ishikura et al., 1999). These cis-acting elements can exert their effects at different levels such as transcription, mRNA stability, or translation (Ruhlman et al., 2010; Zou et al., 2003). Of the four P/5'UTRs that we tested in Chlamydomonas, *psaA-exon1* gave the highest levels of

expression for the two bacterial transgenes that were used (*vapA* and *acrV*). The P/5'UTR of *psaA-exon1* was better than the P/5'UTRs of *atpA*, *psbA* or *psbD*, which in previous studies gave the highest levels of expression of transgenes in the chloroplast of *C. reinhardtii*. The amounts of either protein that we obtained with the P/5'UTR of *psbA* were relatively low compared to those that were reported for other transgenes with the same P/5'UTR, suggesting that *vapA* and *acrV* are particularly difficult to express (Manuell et al., 2007; Surzycki et al., 2009).

It was previously reported that inclusion of the beginning of the coding sequence of a native chloroplast gene in a chimeric construct can improve expression in Chlamydomonas (Kasai et al., 2003) or in higher plants (Lenzi et al., 2008; Maliga, 2002). To test whether this would apply to our transgenes, we compared *psaA::vapA* or *psaA::acrV*, which were fused at the AUG start codon of *psaA*, to constructs where the transgenes were fused to the coding region of *psaA* after codon 20 or codon 30. We observed that inclusion of a part of the coding region actually reduced expression of the transgenes. Thus the beneficial effects that were reported apply to particular chloroplast gene / transgene pairs, but are not universally applicable. In our constructs with the P/5'UTR of *atpA*, the fusion is at position 26 of the coding sequence, similar to what was reported to be optimal (Kasai et al., 2003). However this construct was not investigated further because expression was lower than with the fusion to the AUG start codon of *psaA-exon1*.

In several studies, it was observed that different transformed lines obtained with the same construct can exhibit different levels of transgene expression (Dreesen et al., 2009; Mayfield and Schultz, 2004; Surzycki et al., 2009). This may seem unexpected because all lines should be genetically identical if chloroplast transformation proceeds by homologous recombination. It was argued that concomitant vector integration in the nuclear genome could cause insertional mutations that influence chloroplast transgene expression (Surzycki et al., 2009). However it is unlikely that such mutations would occur at a frequency sufficient to explain the variation that was observed. Our comparison of chloroplast transformation with the IR-int vector using as a host either the wild type or the *FUD7* deletion strain favors a different explanation. In the wild type, we observed consistent amounts of foreign protein in different lines, while in *FUD7*, we

found large variations (Figure 3). This correlates with the fact than in the wild type, the site of vector integration is flanked by two regions of homology that allow specific transgene insertion, by a mechanism that can be formally described as involving crossovers on either side (HR3 and HR4 in figure 1B). In contrast in *FUD7*, one of these regions of homology is deleted, so that integration of the vector by a single crossover on one side (HR4) would be readily reversible, and that stable integration of the transgene would require non-homologous recombination on the other side. We propose that such non-homologous recombination events can differ from line to line, and that they may create different genomic contexts that can lead to different levels of expression for the transgene. A possibility which should be considered is that in such cases a segment of chloroplast DNA containing the transgene and the marker could undergo amplification under the selective pressure applied for selection of transformants (Kindle et al., 1994). Whatever the mechanism, the observed variation can be exploited to screen for lines that have high levels of expression. Hosts bearing chloroplast deletions similar to *FUD7* were used in some reports, but the extent of the deletion was not precisely described so that it is possible that non-homologous recombination may have been involved in these cases (Mayfield and Schultz, 2004; Surzycki et al., 2009). In other instances, there was actually only a single region of homology on one side of the transgene in the vector construct (Dreesen et al., 2009; Goldschmidt-Clermont, 1991), again precluding stable integration by homologous recombination. The Chlamydomonas chloroplast genome contains large numbers of small dispersed repeats (Maul et al., 2002), so that the postulated non-homologous recombination events could actually represent homologous recombination between short repetitive elements.

Negative feedback loops may operate at two levels in the control of *psaA* expression in *C. reinhardtii*. One acts on the translation of *psaA* mRNA which is negatively regulated by unassembled PsaA subunits (Wostrikoff et al., 2004). The other seems to affect the levels of the *psaA-exon1* precursor RNA, as is apparent from its over-accumulation in mutants defective in *psaA trans*-splicing (Choquet et al., 1988). It was not determined whether the accumulation of the *psaA-exon1* precursor is controlled at the level of transcription or of RNA turnover. Since the details of these regulatory mechanisms are not known, and since translation can affect RNA stability, it remains

possible that translation and RNA accumulation are coupled (Herrin and Nickelsen, 2004). The effects that we observe on both mRNA levels and protein accumulation suggest that it may be fruitful to further investigate the effects of other natural or synthetic promoters and 5'UTRs. By placing transgenes driven by the *psaA-exon 1* P/5'UTR in the background of a nuclear mutant deficient in *trans*-splicing of *psaA* (*raa-L121G*), we have been able to circumvent these negative regulatory feedback loops and to strongly augment the expression of the desired proteins. For this purpose it was advantageous to use a mutation that affects splicing because it acts through an intron rather than through the 5'UTR, leaving the chimeric mRNA unaffected, even though it shares the same 5'UTR with the resident *psaA-exon1* gene. The use of a nuclear mutation which is inherited in Mendelian manner, as opposed to chloroplast mutations or transgenes which are inherited uniparentally, facilitates the recovery of the desired genotype in genetic crosses. The mutation clearly had a marked positive effect both on the levels of the chimeric mRNAs and on the amounts of protein that were produced. While the release of negative feedback by PsaA on its own translation (CES regulation) may have played a part, the higher amounts of chimeric mRNA in the *raa-L121G* background suggests that the effect was in part mediated at the level of transcript accumulation. It is interesting to compare our observations on the negative feedback regulation of transgenes driven by *psaA-exon1* with results that were previously obtained for chimeric genes placed under the control of the *psbA* promoter/5'UTR. These chimeric genes were expressed to much higher levels when the *psbA* gene was deleted from the chloroplast genome than in the wild type (Manuell et al., 2007; Rasala et al., 2010). Because expression was still high when PsbA protein (D1) was restored by introducing a modified *psbA* gene driven by the P/5'UTR of *psbD*, this effect was mainly ascribed to competition for activators between the P/5'UTRs of the transgene and of the native gene, rather than to negative feedback. However a partial reduction of transgene expression indicated that autoinhibitory feedback also played a role (Manuell et al., 2007; Rasala et al., 2010).

In this work, we have investigated *cis*-acting elements that promote efficient transgene expression in the chloroplast. We have also explored the potential of manipulating anterograde signaling by *trans*-acting factors to alter chloroplast regulatory

feed-back loops and to increase expression. Our strategy relied on a nuclear mutation that affects *psaA trans*-splicing and had a beneficial effect on transgene expression, but also caused loss of photoautotrophy, which can be seen as a drawback because the mutant cultures grow more slowly in the dark than the wild type in the light. However, large volumes of the transgenic algae could be grown in conventional fermentors without light sources, using acetate as a source of reduced carbon (data not shown). The mutation can also be considered as a means of biological containment for the transgenic lines. Clearly, a better understanding of the detailed mechanism of regulatory feed-back loops will be required to engineer algal strains that show enhanced transgene expression while retaining photoautotrophy.

ACKNOWLEDGEMENTS

We thank Pascal Damay for help with fermentor cultures, Dr. Bernd Köllner for providing antibodies, Nicolas Roggli for preparing the figures and Dr. Luis Lopez-Molina for useful comments on the manuscript. This work was supported by grants from the Swiss NRP59 program (Benefits and Risks of the Deliberate Release of Genetically Modified Plants), and from the Swiss National Science Foundation (3100AO-117712) to JDR and MGC.

EXPERIMENTAL PROCEDURES

Chlamydomonas Cultures

Chlamydomonas reinhardtii wild type, *FUD50* (deletion of *atpB*) and *FUD7* (deletion of *psbA*) strains were grown in Tris acetate phosphate medium (TAP) (Rochaix et al., 1988) to densities of 1-2.10^6 cells/ml in the dark or under fluorescent lights (60 μE.m^{-2}.sec^{-1}) at 25°C.

After helium gun bombardment with atpB-int vectors, *FUD50* transformants were selected on HSM plates under illumination (60 μE.m^{-2}.sec^{-1}) and subcultured several times to obtain homoplasmic strains. With IR-int vectors, wild type or *FUD7* transformants were selected on TAP plates supplemented with 100 μg/ml spectinomycin and repeatedly subcultured to obtain strains homoplasmic for the transgene. They were then repeatedly grown on TAP plates without antibiotic to allow complete loss of the *aadA* cassette. WT and *FUD7* transformants were grown in the light (60 μE.m^{-2}.sec^{-1}) or in the dark, respectively.

Genetic crosses were performed as described (Harris, 2009). A transgenic line expressing *psaA::vapA* (mt+) was crossed to *raa-L121G* (mt-). The pattern of segregation in the progeny indicated that two mutations were present in the *raa-L121G* parental line that both influenced VapA expression. A progeny showing high expression and a *psaA*-deficient phenotype (line D: *psaA::vapA*, *raa-L121G*, mt+) was used for a backcross to the wild type. The progeny showed co-segregation in 6 tetrads of high VapA expression

with the *psaA*-deficient phenotype (*raa-L121G*), showing that line D had lost the second mutation.

Chloroplast transformation vectors

The *vapA* and *acrV* genes from *Aeromonas salmonicida* strain JF2267 were resynthesized using the preferred chloroplast codon usage of Chlamydomonas and integrated in the pGA4 vector (GeneArt AG, Germany).

The atpB-int plasmids were derived from cg12 (Nickelsen et al., 1994) by replacing the *aadA* coding sequence with the transgenes (*vapA* or *acrV*) excised from the pGA4 vector with *Nco*I and *Sph*I sites engineered in the synthetic DNA. The inserts were transferred as *Nco*I – *Sma*I fragments, which also included the 3'UTR of *rbcL*, to derivatives of cg12 containing other promoter / 5'UTR fragments [*atpA* (656 bp) , *psbA* (483 bp), *psaA-exon1* (285 bp)]. To express the transgenes fused at the N-terminus to the first 20 or 30 amino acids of *psaA*, the PCR products obtained with primers 5'-XbaI-ClaI-5'-psaA and either 3'-NcoI-5'-psaA-ex1-20 or 3'-NcoI-5'-psaA-ex1-30 using pEX1 as a template (Choquet et al., 1988) were digested with *Xba*I and *Nco*I and inserted in the atpB-int::*psaA-exon1::vapA* or atpB-int::*psaA-exon1::acrV* plasmids digested with the same enzymes.

The IR-int vectors were obtained by cloning the *Eco*RI-*Xho*I of the chloroplast genome fragment that includes the region between the 5S gene and *psbA* in the pUC19 vector digested with *Eco*RI and *Sal*I (pLM7). A PCR product was obtained with the primers 5'-Bgl-T3 and 3'-Bam-T7 using as a template the recyclable *aadA* cassette cloned in pBluescript KS (Fischer et al., 1996). It was digested with *Bgl*II and *Bam*HI and inserted in pLM7 digested with *Bam*HI in either orientation [IR-int (F) and IR-int (R)]. The *acrV* and *vapA* transgenes, together with the different promoters / 5'UTRs and *rbcL* 3'UTR, were excised from the respective atpB-int derivatives with *Cla*I and *Sma*I, and inserted in the IR-int(F) or IR-int(R) vectors digested with the same enzymes.

Genotyping.

Homoplasmicity of the insertions was determined by PCR on total DNA extracts (Cao et al., 2009) using the following protocol (5 min at 95°C / 25 or 35 cycles: 1 min at

95°C, 1 min at 50°C or 54°C, 1 min at 72°C / 7°C) with the number of cycles and annealing temperatures indicated below. The sequences of the oligonucleotide primers are listed in Table S2.

Insertion of *vapA* and *acrV* in the transformants was revealed by the presence of a 242 bp fragment amplified with 5'-VapA and 3'-VapA or a 302 bp fragment with 5'-AcrV and 3'-AcrV, respectively (25 cycles, 50°C).

For atpB-int transformants, lack of parental *FUD50* genomes was ascertained by the absence of a 1254 bp fragment amplified from the *atpB* deletion locus with EcoRI-AtpB-for and AtpB-rev4 (35 cycles, 50°C).

For IR-int transformants of the wild type, homoplasmic lines were determined from the absence of a 102 bp fragment amplified with 5'-cpIR and 3'-cpIR, and the presence of a 1 kb amplification product with rbcL-3884 and 5' cpIR for transformants obtained with IR-int(F) or rbcL-3884 and 3' cpIR for those obtained with IR-int(R) (25 cycles, 50°C).

For IR-int transformants of *FUD7*, homoplasmic lines were determined from the absence of a 1.8 kb amplification of the parental locus with 5S-fud7 and 5'-psbA (25 cycles, 54°C), and the presence of a 1 kb amplification with rbcL-3884 and 5'-cpIR for transformants obtained with IR-int(F) or rbcL-3884 and 3'-cpIR for those obtained with IR-int(R) (25 cycles, 50°C).

RNA analysis

Chlamydomonas cells from 50 ml cultures (1-2 x 10^6 cells / mL) were harvested by centrifugation, washed once with 5 ml of 20 mM Tris pH 7.9 and the pellets were stored in liquid nitrogen. One ml of Tri Reagent® (SigmaAldrich) and 500 µl of glass beads (G8772, SigmaAldrich) were added to the frozen pellets and the suspensions were vigorously agitated for 2 min. The homogenates were transferred to a Phase Lock Gel™ tube (5 PRIME). After incubation for 1 min at room temperature, 200 µl of chloroform were added before thorough mixing. After 10 min centrifugation at 14000 × g at 4°C, the RNA was precipitated from the aqueous phase with 0.5 ml isopropanol, collected by centrifugation and washed with 70% ethanol. Total RNA (2 µg) was analyzed by agarose gel electrophoresis, transfer to Nylon membranes and hybridization using probes labelled with ^{32}P-dATP by random priming (Ausubel et al., 1998). The *vapA* probe was obtained

by digestion of pGA4::*vapA* with *Nco*I and *Sph*I, the different *psaA* probes and the *psbD* probe were obtained as described previously (Choquet et al., 1988). The expression of the mRNAs was quantified by phosphorimaging (BioRad) and normalized to *psbD* mRNA. The mRNA for VapA increased 7 fold in the *raa-L121G* mutant background (from 3.4 (+/- 0.4) to 24.4 (+/- 3.5), arbitrary units, n=3).

Protein analysis

Recombinant 6xHis-tagged VapA and AcrV were produced using the pETHIS1/BL21 expression system (Novagen) and purified by Ni-NTA chromatography (QIAgen) under denaturing conditions (6 M GuHCl).

Chlamydomonas cells (15 mL, 1-2 × 10^6 cells/ml) were collected by centrifugation, washed with 30mM Tris-HCl pH 7.9, 1 mM EDTA, and resuspended in lysis buffer [100 mM Tris-HCl pH 6.8, 4% SDS, 20 mM EDTA, protease inhibitor cocktail (Sigma-Aldrich)]. After 30 minutes at room temperature, cell debris were removed by centrifugation and 100 µg of total proteins were analyzed by SDS-PAGE (12% acrylamide) and immunoblotting. To quantify transgene expression, different amounts of recombinant protein were mixed with 100 µg of protein extract from the respective host strains. Proteins were blotted onto nitrocellulose filters, incubated with anti-VapA or AcrV primary antibody at 4°C overnight in Elisa buffer (10 mM Tris pH 8.0, 100 mM NaCl, 0.5% Tween). Labeling of the membranes with anti-PsaA, Cytf or RbcL antisera was carried out at room temperature in 1 × TBS (50 mM Tris-HCl pH 7.6, 150 mM NaCl), 0.1% Tween 20 and 1% non-fat powder milk. After washing the membranes, the antibodies were revealed with a peroxidase-linked secondary antibody (Promega) and visualized by enhanced chemiluminescence.

Protein degradation was assayed using Chlamydomonas cultures at 2 × 10^6 cells/ml, untreated, supplemented with 200 µg/ml chloramphenicol, or with protease inhibitors [1 mM AEBSF (4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride (SigmaAldrich)) plus protease inhibitor cocktail tablets (2 × Complete mini, EDTA-free (Roche)]. Aliquots were collected at different times for protein extraction and western blot analysis.

REFERENCES

Ausubel, F.A., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1998) *Current Protocols in Molecular Biology*. New York, NY: John Wiley & sons.

Barkan, A. and Goldschmidt-Clermont, M. (2000) Participation of nuclear genes in chloroplast gene expression. *Biochimie*, **82**, 559-572.

Barnes, D., Franklin, S., Schultz, J., Henry, R., Brown, E., Coragliotti, A. and Mayfield, S.P. (2005) Contribution of 5'- and 3'-untranslated regions of plastid mRNAs to the expression of Chlamydomonas reinhardtii chloroplast genes. *Mol Genet Genomics*, **274**, 625-636.

Birch-Machin, I., Newell, C.A., Hibberd, J.M. and Gray, J.C. (2004) Accumulation of rotavirus VP6 protein in chloroplasts of transplastomic tobacco is limited by protein stability. *Plant Biotechnol J*, **2**, 261-270.

Bock, R. and Warzecha, H. (2010) Solar-powered factories for new vaccines and antibiotics. *Trends Biotechnol*, **28**, 246-252.

Cao, M., Fu, Y., Guo, Y. and Pan, J. (2009) Chlamydomonas (Chlorophyceae) colony PCR. *Protoplasma*, **235**, 107-110.

Chebolu, S. and Daniell, H. (2009) Chloroplast-derived vaccine antigens and biopharmaceuticals: expression, folding, assembly and functionality. *Curr Top Microbiol Immunol*, **332**, 33-54.

Choquet, Y., Goldschmidt-Clermont, M., Girard-Bascou, J., Kuck, U., Bennoun, P. and Rochaix, J.D. (1988) Mutant phenotypes support a trans-splicing mechanism for the expression of the tripartite psaA gene in the C. reinhardtii chloroplast. *Cell*, **52**, 903-913.

Choquet, Y. and Wollman, F.A. (2002) Translational regulations as specific traits of chloroplast gene expression. *FEBS Lett*, **529**, 39-42.

Daniell, H., Singh, N.D., Mason, H. and Streatfield, S.J. (2009) Plant-made vaccine antigens and biopharmaceuticals. *Trends Plant Sci*, **14**, 669-679.

Doran, P.M. (2006) Foreign protein degradation and instability in plants and plant tissue cultures. *Trends Biotechnol*, **24**, 426-432.

Dreesen, I.A., Charpin-El Hamri, G. and Fussenegger, M. (2009) Heat-stable oral alga-based vaccine protects mice from Staphylococcus aureus infection. *J Biotechnol*, **145**, 273-280.

Fischer, N., Stampacchia, O., Redding, K. and Rochaix, J.D. (1996) Selectable marker recycling in the chloroplast. *Mol Gen Genet*, **251**, 373-380.

Goldschmidt-Clermont, M. (1991) Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker of site-directed transformation of chlamydomonas. *Nucleic Acids Res*, **19**, 4083-4089.

Goldschmidt-Clermont, M., Choquet, Y., Girard-Bascou, J., Michel, F., Schirmer-Rahire, M. and Rochaix, J.D. (1991) A small chloroplast RNA may be required for trans-splicing in Chlamydomonas reinhardtii. *Cell*, **65**, 135-143.

Goldschmidt-Clermont, M., Girard-Bascou, J., Choquet, Y. and Rochaix, J.D. (1990) Trans-splicing mutants of Chlamydomonas reinhardtii. *Mol Gen Genet*, **223**, 417-425.

Goldschmidt-Clermont, M., Rahire, M. and Rochaix, J.D. (2008) Redundant cis-acting determinants of 3' processing and RNA stability in the chloroplast rbcL mRNA of Chlamydomonas. *Plant J*, **53**, 566-577.

Griesbeck, C., Kobl, I. and Heitzer, M. (2006) Chlamydomonas reinhardtii: a protein expression system for pharmaceutical and biotechnological proteins. *Mol Biotechnol*, **34**, 213-223.

Grossman, A.R., Croft, M., Gladyshev, V.N., Merchant, S.S., Posewitz, M.C., Prochnik, S. and Spalding, M.H. (2007) Novel metabolism in Chlamydomonas through the lens of genomics. *Curr Opin Plant Biol*, **10**, 190-198.

Harris, E.H. (2009) *The Chlamydomonas sourcebook*. Oxford, GB: Academic Press, Elsevier.

Herrin, D.L. and Nickelsen, J. (2004) Chloroplast RNA processing and stability. *Photosynth Res*, **82**, 301-314.

Herz, S., Fussl, M., Steiger, S. and Koop, H.U. (2005) Development of novel types of plastid transformation vectors and evaluation of factors controlling expression. *Transgenic Res*, **14**, 969-982.

Ishikura, K., Takaoka, Y., Kato, K., Sekine, M., Yoshida, K. and Shinmyo, A. (1999) Expression of a foreign gene in Chlamydomonas reinhardtii chloroplast. *J Biosci Bioeng*, **87**, 307-314.

Kasai, S., Yoshimura, S., Ishikura, K., Takaoka, Y., Kobayashi, K., Kato, K. and Shinmyo, A. (2003) Effect of coding regions on chloroplast gene expression in Chlamydomonas reinhardtii. *J Biosci Bioeng*, **95**, 276-282.

Kato, K., Marui, T., Kasai, S. and Shinmyo, A. (2007) Artificial control of transgene expression in Chlamydomonas reinhardtii chloroplast using the lac regulation system from Escherichia coli. *J Biosci Bioeng*, **104**, 207-213.

Kindle, K.L., Suzuki, H. and Stern, D.B. (1994) Gene Amplification Can Correct a Photosynthetic Growth Defect Caused by mRNA Instability in Chlamydomonas Chloroplasts. *Plant Cell*, **6**, 187-200.

Leelavathi, S. and Reddy, V.S. (2003) Chloroplast expression of His-tagged GUS-fusions: a general strategy to overproduce and purify proteins using transplastomic plants as bioreactors. *Mol Breed*, **11**, 49-58.

Lenzi, P., Scotti, N., Alagna, F., Tornesello, M.L., Pompa, A., Vitale, A., De Stradis, A., Monti, L., Grillo, S., Buonaguro, F.M., Maliga, P. and Cardi, T. (2008) Translational fusion of chloroplast-expressed human papillomavirus type 16 L1 capsid protein enhances antigen accumulation in transplastomic tobacco. *Transgenic Res*, **17**, 1091-1102.

Maliga, P. (2002) Engineering the plastid genome of higher plants. *Curr Opin Plant Biol*, **5**, 164-172.

Manuell, A.L., Beligni, M.V., Elder, J.H., Siefker, D.T., Tran, M., Weber, A., McDonald, T.L. and Mayfield, S.P. (2007) Robust expression of a bioactive mammalian protein in Chlamydomonas chloroplast. *Plant Biotechnol J*, **5**, 402-412.

Maul, J.E., Lilly, J.W., Cui, L., dePamphilis, C.W., Miller, W., Harris, E.H. and Stern, D.B. (2002) The Chlamydomonas reinhardtii plastid chromosome: islands of genes in a sea of repeats. *Plant Cell*, **14**, 2659-2679.

Mayfield, S.P., Manuell, A.L., Chen, S., Wu, J., Tran, M., Siefker, D., Muto, M. and Marin-Navarro, J. (2007) Chlamydomonas reinhardtii chloroplasts as protein factories. *Curr Opin Biotechnol*, **18**, 126-133.

Mayfield, S.P. and Schultz, J. (2004) Development of a luciferase reporter gene, luxCt, for Chlamydomonas reinhardtii chloroplast. *Plant J*, **37**, 449-458.

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Marechal-Drouard, L., Marshall, W.F., Qu, L.H., Nelson, D.R., Sanderfoot, A.A., Spalding, M.H., Kapitonov, V.V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S.M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.L., Cognat, V., Croft, M.T., Dent, R., Dutcher, S., Fernandez, E., Fukuzawa, H., Gonzalez-Ballester, D., Gonzalez-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P.A., Lemaire, S.D., Lobanov, A.V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J.V., Moseley, J., Napoli, C., Nedelcu, A.M., Niyogi, K., Novoselov, S.V., Paulsen, I.T., Pazour, G., Purton, S., Ral, J.P., Riano-Pachon, D.M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S.L., Allmer, J., Balk, J., Bisova, K., Chen, C.J., Elias, M., Gendler, K., Hauser, C., Lamb, M.R., Ledford, H., Long, J.C., Minagawa, J., Page, M.D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A.M., Yang, P., Ball, S., Bowler, C., Dieckmann, C.L., Gladyshev, V.N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y.W., Jhaveri, J., Luo, Y., Martinez, D., Ngau, W.C., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I.V., Rokhsar, D.S. and Grossman, A.R. (2007) The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science*, **318**, 245-250.

Minai, L., Wostrikoff, K., Wollman, F.A. and Choquet, Y. (2006) Chloroplast biogenesis of photosystem II cores involves a series of assembly-controlled steps that regulate translation. *Plant Cell*, **18**, 159-175.

Nickelsen, J., van Dillewijn, J., Rahire, M. and Rochaix, J.D. (1994) Determinants for stability of the chloroplast psbD RNA are located within its short leader region in Chlamydomonas reinhardtii. *Embo J*, **13**, 3182-3191.

Oey, M., Lohse, M., Kreikemeyer, B. and Bock, R. (2009) Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. *Plant J*, **57**, 436-445.

Pfannschmidt, T. (2003) Chloroplast redox signals: how photosynthesis controls its own genes. *Trends Plant Sci*, **8**, 33-41.

Posten, C. and Schaub, G. (2009) Microalgae and terrestrial biomass as source for fuels--a process view. *J Biotechnol*, **142**, 64-69.

Rasala, B.A., Muto, M., Lee, P.A., Jager, M., Cardoso, R.M.F., Behnke, C.A., Kirk, P., Hokanson, C.A., Crea, R., Mendez, M. and Mayfield, S.P. (2010) Production of therapeutic proteins in algae, analysis of expression of seven human proteins in the chloroplast of Chlamydomonas reinhardtii. *Plant Biotechnol J*, **8**, 1-15.

Ruhlman, T., Verma, D., Samson, N. and Daniell, H. (2010) The role of heterologous chloroplast sequence elements in transgene integration and expression. *Plant Physiol*, **152**, 2088-2104.

Rupprecht, J. (2009) From systems biology to fuel--Chlamydomonas reinhardtii as a model for a systems biology approach to improve biohydrogen production. *J Biotechnol*, **142**, 10-20.

Specht, E., Miyake-Stoner, S. and Mayfield, S. (2010) Micro-algae come of age as a platform for recombinant protein production. *Biotechnol Lett*.

Stern, D., Goldschmidt-Clermont, M. and Hanson, M.R. (2010) Chloroplast RNA metabolism. *Annu Rev Plant Biol*, **61**, in press.

Stevens, L.H., Stoopen, G.M., Elbers, I.J., Molthoff, J.W., Bakker, H.A., Lommen, A., Bosch, D. and Jordi, W. (2000) Effect of climate conditions and plant developmental stage on the stability of antibodies expressed in transgenic tobacco. *Plant Physiol*, **124**, 173-182.

Surzycki, R., Cournac, L., Peltier, G. and Rochaix, J.D. (2007) Potential for hydrogen production with inducible chloroplast gene expression in Chlamydomonas. *Proc Natl Acad Sci U S A*, **104**, 17548-17553.

Surzycki, R., Greenham, K., Kitayama, K., Dibal, F., Wagner, R., Rochaix, J.D., Ajam, T. and Surzycki, S. (2009) Factors effecting expression of vaccines in microalgae. *Biologicals*, **37**, 133-138.

Walker, T.L., Purton, S., Becker, D.K. and Collet, C. (2005) Microalgae as bioreactors. *Plant Cell Rep*, **24**, 629-641.

Wostrikoff, K., Girard-Bascou, J., Wollman, F.A. and Choquet, Y. (2004) Biogenesis of PSI involves a cascade of translational autoregulation in the chloroplast of Chlamydomonas. *Embo J*, **23**, 2696-2705.

Ye, G.N., Hajdukiewicz, P.T., Broyles, D., Rodriguez, D., Xu, C.W., Nehra, N. and Staub, J.M. (2001) Plastid-expressed 5-enolpyruvylshikimate-3-phosphate synthase genes provide high level glyphosate tolerance in tobacco. *Plant J*, **25**, 261-270.

Zou, Z., Eibl, C. and Koop, H.U. (2003) The stem-loop region of the tobacco psbA 5'UTR is an important determinant of mRNA stability and translation efficiency. *Mol Genet Genomics*, **269**, 340-349.

FIGURE LEGENDS

Figure 1.

Chloroplast transformation vectors and hosts

(A) ***FUD50***: map of the insertion site in the host strain *FUD50*. The inverted repeat of the chloroplast genome is shown as a black bar, the single copy region as a white bar, the 16 S rRNA and *atpB* genes as grey boxes. The extent of the deletion in *FUD50* is delineated by brackets. Two regions that allow homologous recombination of the vector are shown hatched (HR1 and HR2).

atpB-int vectors: the transgene and the *atpB* gene are shown as grey arrows. Four different promoters and 5'UTRs (*atpA*, *psaA*, *psbA*, *psbD*) were used in combination with either *acrV* or *vapA* and with the 3'UTR of *rbcL*. Restriction sites used for construction of the vectors are shown below. The two regions that allow for homologous recombination with the host chloroplast DNA are indicated by hatched bars (HR1 and HR2). For reasons that are not relevant to this work, the 3' UTR of *atpB* was replaced with the 3'UTR of *psbD*.

(B) **WT**: map of the insertion site in the inverted repeat of the wild type. The *psbA* and the ribosomal RNA genes (5S, 23S, 3S) are shown as boxes with exons in dark grey and introns in light grey. The site of insertion at a *Bam*HI site is indicated with an arrow. The two regions that allow homologous recombination of the vector are represented with hatched bars (HR3 and HR4).

IR-int vectors: the structure of the chimeric transgene is shown as in (A) with a grey arrow. The selection marker *aadA* with the 5'UTR of *atpA* and the 3'UTR of *rbcL* is flanked by two direct repeats of bacterial DNA (483 bp). Together with the duplicated 3'UTR of *rbcL*, they allow excision of the selectable marker. Two regions that allow homologous recombination with the host chloroplast DNA are hatched (HR3 and HR4). The forward orientation (F) of the vector is represented; in the reverse (R) orientation, the entire fragment containing the *aadA* marker and the transgene is inserted in the opposite orientation at the same *Bam*HI site.

FUD7: brackets delineate the extent of the deletion in this mutant.

Figure 2.

Comparison of promoters / 5'UTRs

Expression of VapA **(A)** or AcrV **(B)** in Chlamydomonas lines obtained by transformation of the wild type with IR-int vectors carrying chimeric transgenes with the P/5'UTRs indicated at the top. Two independent transgenic lines are shown for each construct with their numbers (#) indicated above each lane. Total proteins (100 μg) from light-grown cultures were subjected to SDS-PAGE and immuno-blotting with anti-VapA monoclonal antibody, anti-AcrV serum or anti-RbcL serum (large subunit of Rubisco), as indicated on the right. As a reference, different amounts of recombinant VapA or AcrV protein as indicated on the top left were mixed with total protein extract from non-transformed Chlamydomonas. The migration of the recombinant proteins is slightly retarded due to the presence of a poly-histidine tag. (F) and (R) indicate the orientation of the IR-int vectors in the transgenic strains (see Fig. 1B legend). One of the *atpA::vapA* lines lost the transgene (Fig 2A, lane #8Δ) as discussed in Supporting Information (Fig S1).

Figure 3.

Comparison of *psaA::vapA* expression in the wild type and *FUD7*

Expression of VapA in Chlamydomonas lines obtained by transformation of the wild type **(A)** or *FUD7* **(B)** with an IR-int vector carrying the chimeric *psaA::vapA* transgene. Several independent lines are shown, designated with their numbers (#). Total proteins (100 μg) from light-grown cultures of the WT transformants and from dark-grown cultures of *FUD7* transformants were subjected to SDS-PAGE and immuno-blotting with sera against VapA, RbcL or Cyt f, as indicated on the right. As a reference, different amounts of recombinant VapA protein were included as in Fig 2.

Figure 4.

Enhanced transgene expression in a PSI deficient nuclear mutant

Analysis of a representative tetrad (DA, DB, DC, DD) from a cross of line D (*psaA::vapA*; *L121G*; *mt+*) to the wild type (WT; *mt-*). The original *psaA::vapA* line #2 and the *raa-L121G* mutant are included for comparison. All cultures were grown in the dark.

(A) Total proteins were analyzed by SDS-PAGE and immunoblotting with antisera against VapA, PsaA or RbcL.

(B) Total RNA was subjected to agarose gel electrophoresis, blotting to a membrane and hybridization with probes specific for *vapA*, *psaA-exon3*, *psaA-exon1* or *psbD* as indicated on the right. The mature *psaA* mRNA (exons 1-2-3) and the unspliced precursors of exon 3 (pre-exon 3) and of exon 1 (pre-exon 1) are labeled on the left.

Figure 5.

Comparison of transgene expression in wild-type and mutant backgrounds

Total proteins were analyzed by SDS-PAGE and immunoblotting with antisera against VapA, AcrV, PsaA or RbcL as indicated on the right. The cultures were grown either in the dark (Dk) or in the light (L; 60 $\mu E.m^{-2}.s^{-1}$) as indicated above each lane. Independent lines are shown, designated with their numbers (#). (F) and (R) denote the orientation of the IR-int vectors in the transgenic strains (see Fig.1B legend).

(A) Comparison of *psaA::vapA* expression in the wild type (WT) or in the *raa-L121G* mutant defective in *psaA trans*-splicing (D, DB, DD; see figure 4).

(B) Comparison of *psaA::acrV* expression in the wild type (WT) or in the *raa-L121G* mutant. T6 was obtained from a cross of *psaA::acrV* to *raa-L121G*; T6B and T6D are progeny of a backcross of T6 to the wild type.

Figure 6.

Effect of proteolysis on transgene expression

Dark-grown cultures of Chlamydomonas transgenic line D (*psaA::vapA*; *raa-L121G*) (**panel A**) or line T6 (*psaA::acrV*; *raa-L121G*) (**panel B**) were split in three parts and either untreated, supplemented with chloramphenicol or with protease inhibitor cocktail. Samples of total protein were extracted at the times indicated above the lanes and subjected to SDS-PAGE and immunoblotting with antisera against VapA, AcrV or RbcL as indicated on the right. As a reference, different amounts of recombinant VapA or AcrV protein were mixed with total extract from non-transformed Chlamydomonas as indicated on the top left. The cytotoxicity of the protease inhibitors became apparent after 6 – 9 hours of treatment, and probably led to the decline in RbcL observed at 9 hours in panel B.

SUPPORTING INFORMATION

Figure S1.

Repeated sequences cause transgene instability.

In the special case where the P/5'UTR of *atpA* is fused to the transgene in the IR-int vectors, the presence of the *atpA::aadA* cassette creates a direct repeat of the *atpA* P/5'UTR. Homologous recombination between the two copies of this direct repeat can lead to the loss of the transgene (red arrows). If this recombination event has not yet occurred when the selective pressure for spectinomycin resistance is released to remove the *aadA* marker, homologous recombination between the 483 bp repeats can lead to the excision of the *atpA::aadA* cassette (green arrows), resulting in transgene stabilization. For this reason, only a few *atpA::vapA* lines and no *atpA::acrV* lines could be recovered with the IR-int vector, while with the atpB-int vector the corresponding lines were readily obtained.

Figure S2.

Fusions to the coding region of *psaA-exon1*

(**A**) The open reading frame of *vapA* was fused to a fragment containing the P/5'UTR and coding sequence of *psaA-exon1* either after codon 20 (*psaA-psaA*$_{20}$*::vapA*) or after codon 30 (*psaA-psaA*$_{30}$*::vapA*) in the atpB-int vector. Several independent lines are shown, designated with their numbers (#). For comparison, a transformant with the fusion at the AUG start codon of *psaA-exon1* (*psaA::vapA*) is also included. Total proteins (100 µg) from light-grown cultures were analyzed by SDS-PAGE and immunoblotting with antisera against VapA, or RbcL (large subunit of Rubisco) as indicated on the right. As a reference, 35 ng recombinant VapA protein was mixed with total extract from non-transformed Chlamydomonas (leftmost lane).

(**B**) Similar constructs with *acrV* were prepared and analyzed as in (A).

These transgenic strains contain the IR-int vectors in the forward (F) orientation (see Fig 1B legend)

Figure S3.

Comparison of *psaA::acrV* expression in the wild type and *FUD7*

Expression of AcrV in Chlamydomonas lines obtained by transformation of the wild type **(A)** or *FUD7* **(B)** with an IR-int vector carrying the chimeric *psaA::acrV* transgene. Several independent lines are shown, designated with their numbers (#). Total proteins (100 μg) from light-grown cultures for the WT transformants and from dark-grown cultures for *FUD7* transformants were subjected to SDS-PAGE and immuno-blotting with anti-AcrV serum, anti- RbcL serum (Rubisco large subunit) or anti-Cytf, as indicated on the right. For calibration, different amounts of recombinant AcrV protein were mixed with total extract from non-transformed Chlamydomonas as indicated on the top left. The migration of the recombinant protein is slightly retarded due to the presence of a poly-histidine tag. These transgenic strains contain the IR-int vectors in the forward (F) orientation (see Fig 1B legend).

Table S1. Expression levels of VapA and AcrV.

The amounts of transgenic proteins obtained with different constructs and in different genetic backgrounds (*ΔatpB*: *FUD50*; WT: wild type; *ΔpsbA*: *FUD7*) were estimated by immunoblotting, in comparison to known amounts of recombinant proteins (Figure 2, Figure 3 and data not shown). (F) and (R) indicate the orientation of the IR-int vectors in the transgenic strains (see Fig 1B legend).
The symbols correspond to the following approximate amounts (% total cell protein).

(+)	< 0.010 %
+	0.010 - 0.020 %
++	0.030 - 0.050 %
+++	0.075 - 0.150 %
++++	0.200 - 0.300 %
~	indicates the range of expression that was observed in different lines
nd	not determined

	promoter & 5'UTR	Estimated transgenic protein expression				
		light grown			dark grown	
		atpB-int vector in *ΔatpB*	IR-int vector in WT		IR-int vector in *ΔpsbA*	
			(F)	(R)	(F)	(R)
VapA	*atpA*	(+)	nd	++	nd	nd
	psbD	(+)	++	++	nd	nd
	psbA	(+)	(+)	(+)	++	++
	psaA	++	+++	+++	+ ~ ++++	+
	psaA + 20 a.a.	(+)	++	nd	(+)	nd
	psaA + 30 a.a.	(+)	++	nd	(+)	nd
AcrV	*atpA*	+++	nd	nd	nd	nd
	psbD	+	++	++	+++	++ ~ +++
	psbA	(+)	(+)	nd	++ ~ +++	nd
	psaA	+	+++	nd	+ ~ +++	nd
	psaA + 20 a.a.	(+)	++	nd	(+)	nd
	psaA + 30 a.a.	(+)	++	nd	(+)	nd

Table S2. Oligonucleotides used in this work.

Name	Sequence (5’-3’)
5'-VapA	GCTGCTATTGTTGTTGGTTC’
3'-VapA	CATTGACGAATTTCAGCATC
5'-AcrV	CCAGGTGCTTTTGTTGGTTG
3'-AcrV	CAATGTGTAACTGGTGGTGT
EcoRI-AtpB-for	CTCTTCAACGCTATATTCCACAAA
AtpB-rev4	AAGTAAACTTAGGGATTTTAATGCAATGCAATAAA
5'-cpIR	TGGAATTGGATATGGACTAG
3'-cpIR	ACTTATGAAATGCAAGTACC
rbcL-3884	TAAATGGTCTCCAGAACTTGC
5S-fud7	ACTTTACGGGTCGCCGTCTG
5'-psbA	TAAGGGGAAGGGGACGTAGG
5’-XbaI-ClaI-5’-psaA	GTTACCAATCTAGAATCGATAAGCTTTCTTAATTCAACATT
3’-NcoI-5’-psaA-ex1-20	CTTATTTCACCATGGCCGCCCATTTTTCAAAACTTGT
3’-NcoI-5’-psaA-ex1-30	ACTTGTTTCCATGGCATTACGATCAACCGCAATCTT
5’-Bgl-T3	AGTAGATCTTAATACGACTCACTATA
3’-Bam-T7	AATGGATCCGATATCGATAAGCCCGGGATTAACCCTCACTAAAG

Figure 1

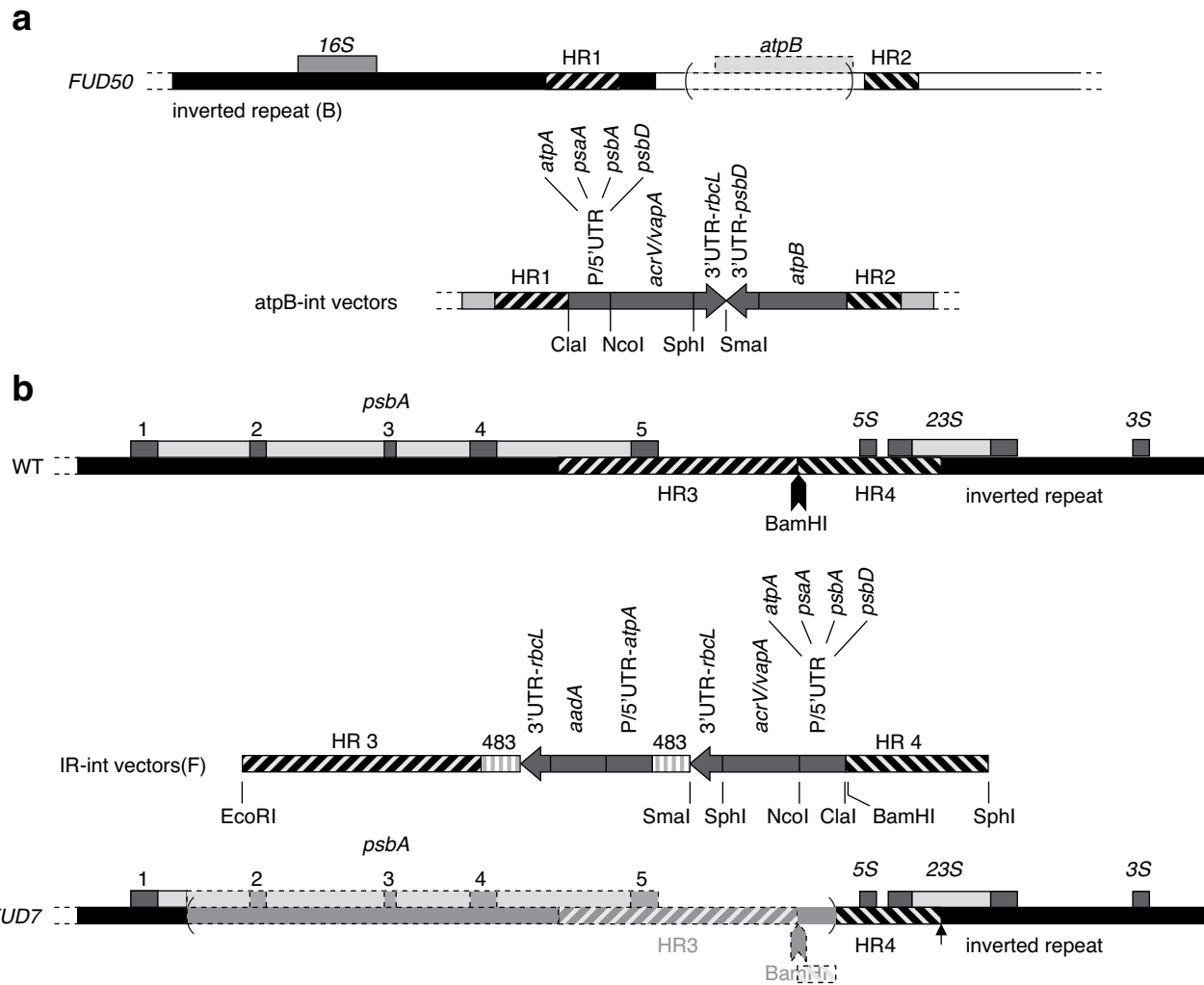

a
16S
HR1
atpB
HR2
FUD50
inverted repeat (B)
atpA
psaA
psbA
psbD
P/5'UTR
acrV/vapA
3'UTR-rbcL
3'UTR-psbD
atpB
HR1
HR2
atpB-int vectors
ClaI
NcoI
SphI
SmaI
b
psbA
1
2
3
4
5
5S
23S
3S
WT
HR3
HR4
inverted repeat
BamHI
3'UTR-rbcL
aadA
P/5'UTR-atpA
3'UTR-rbcL
acrV/vapA
atpA
psaA
psbA
psbD
P/5'UTR
HR 3
483
483
HR 4
IR-int vectors(F)
EcoRI
SmaI
SphI
NcoI
ClaI
BamHI
SphI
psbA
1
2
3
4
5
5S
23S
3S
FUD7
HR3
HR4
inverted repeat
BamHI

Figure 2

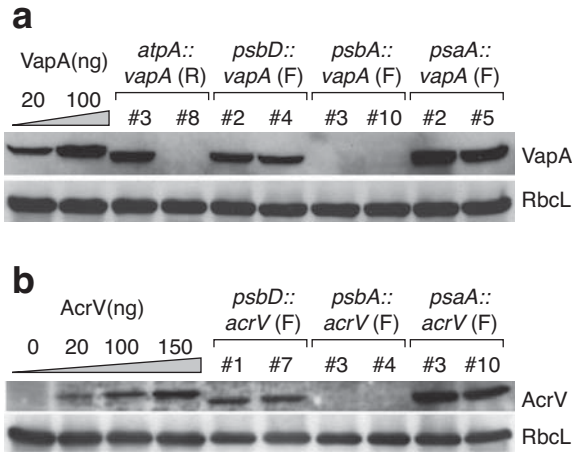

a
VapA(ng)
20 100
atpA:: vapA (R)
#3 #8
psbD:: vapA (F)
#2 #4
psbA:: vapA (F)
#3 #10
psaA:: vapA (F)
#2 #5
VapA
RbcL
b
AcrV(ng)
0 20 100 150
psbD:: acrV (F)
#1 #7
psbA:: acrV (F)
#3 #4
psaA:: acrV (F)
#3 #10
AcrV
RbcL

Figure 3

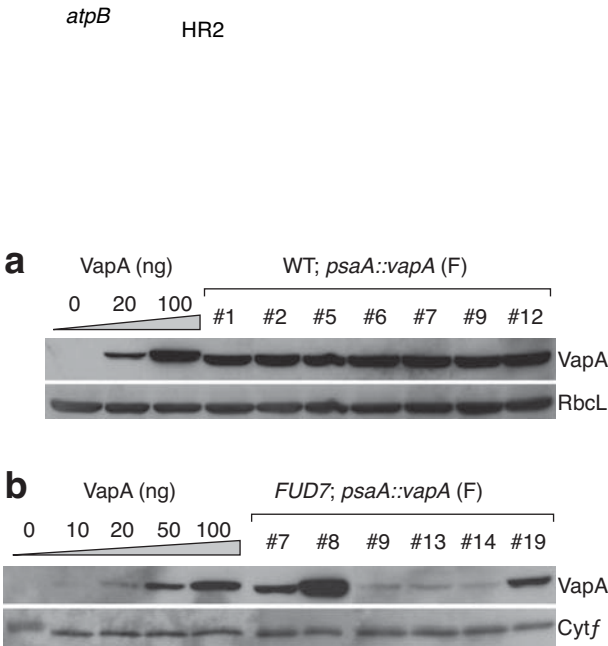

atpB
HR2
a
VapA (ng)
WT; psaA::vapA (F)
0 20 100
#1 #2 #5 #6 #7 #9 #12
VapA
RbcL
b
VapA (ng)
FUD7; psaA::vapA (F)
0 10 20 50 100
#7 #8 #9 #13 #14 #19
VapA
Cytf

Figure 4

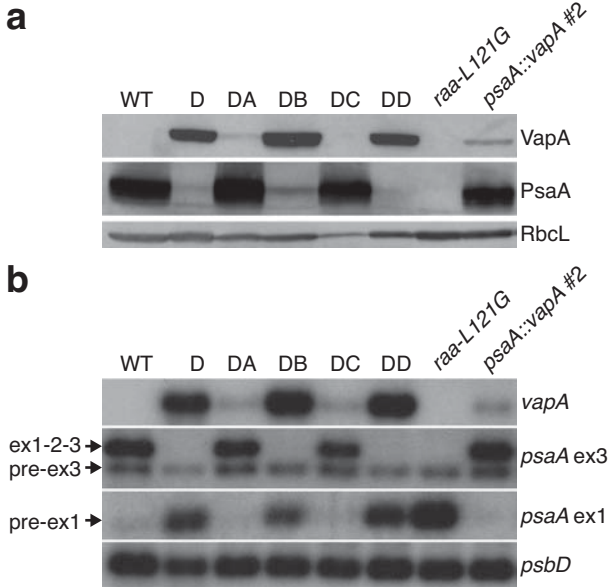

a
WT
D
DA
DB
DC
DD
raa-L121G
psaA::vapA #2
VapA
PsaA
RbcL
b
WT
D
DA
DB
DC
DD
raa-L121G
psaA::vapA #2
vapA
ex1-2-3
pre-ex3
psaA ex3
pre-ex1
psaA ex1
psbD

Figure 5

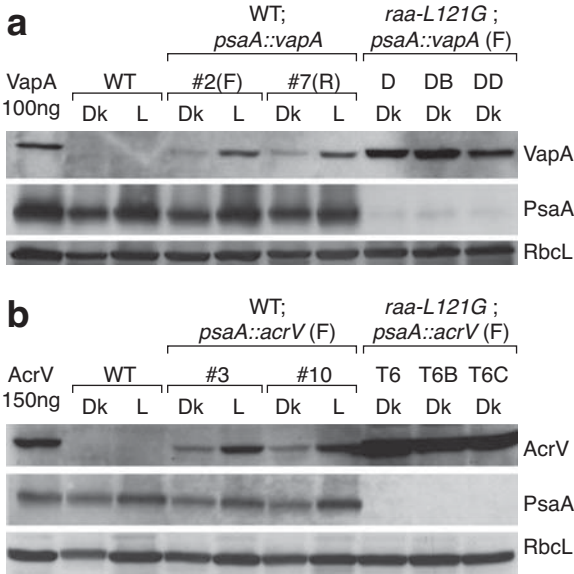
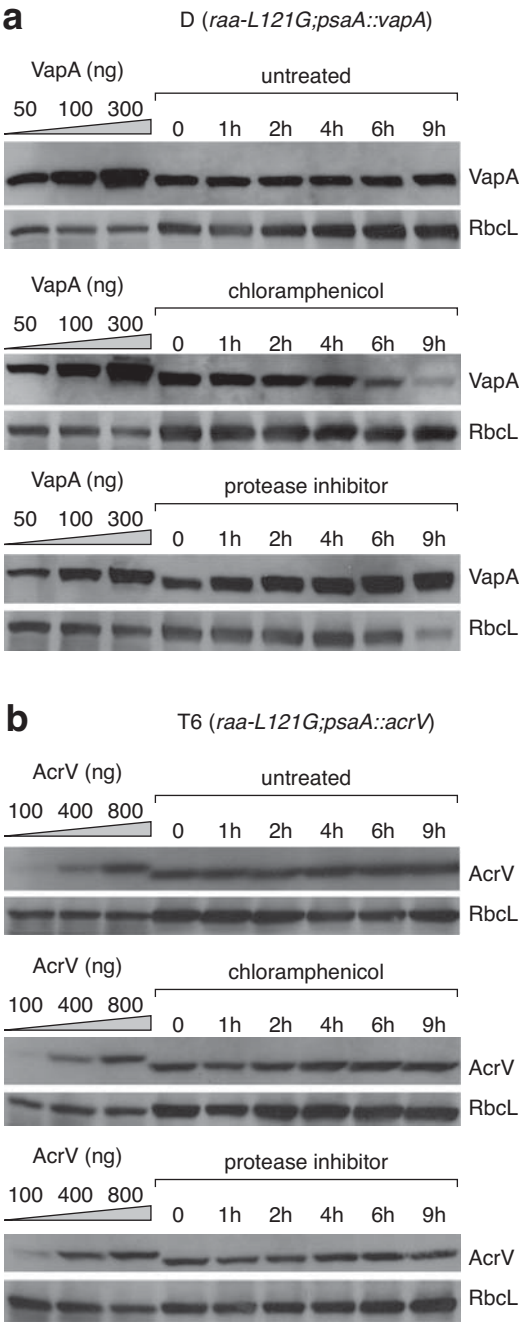

a
WT; psaA::vapA
raa-L121G ; psaA::vapA (F)
VapA 100ng
WT
#2(F)
#7(R)
D
DB
DD
Dk
L
VapA
PsaA
RbcL
b
WT; psaA::acrV (F)
raa-L121G ; psaA::acrV (F)
AcrV 150ng
WT
#3
#10
T6
T6B
T6C
AcrV
PsaA
RbcL

Figure 6


a
D (raa-L121G;psaA::vapA)
VapA (ng)
50 100 300
untreated
0 1h 2h 4h 6h 9h
VapA
RbcL
VapA (ng)
50 100 300
chloramphenicol
0 1h 2h 4h 6h 9h
VapA
RbcL
VapA (ng)
50 100 300
protease inhibitor
0 1h 2h 4h 6h 9h
VapA
RbcL
b
T6 (raa-L121G;psaA::acrV)
AcrV (ng)
100 400 800
untreated
0 1h 2h 4h 6h 9h
AcrV
RbcL
AcrV (ng)
100 400 800
chloramphenicol
0 1h 2h 4h 6h 9h
AcrV
RbcL
AcrV (ng)
100 400 800
protease inhibitor
0 1h 2h 4h 6h 9h
AcrV
RbcL

Enhanced chloroplast transgene expression in a nuclear mutant of Chlamydomonas.

Laure Michelet, Linnka Lefebvre-Legendre, Sarah E. Burr, Jean-David Rochaix and Michel Goldschmidt-Clermont

SUPPORTING INFORMATION

Figure S1.

Repeated sequences cause transgene instability.

In the special case where the P/5'UTR of *atpA* is fused to the transgene in the IR-int vectors, the presence of the *atpA::aadA* cassette creates a direct repeat of the *atpA* P/5'UTR. Homologous recombination between the two copies of this direct repeat can lead to the loss of the transgene (red arrows). If this recombination event has not yet occurred when the selective pressure for spectinomycin resistance is released to remove the *aadA* marker, homologous recombination between the 483 bp repeats can lead to the excision of the *atpA::aadA* cassette (green arrows), resulting in transgene stabilization. For this reason, only a few *atpA::vapA* lines and no *atpA::acrV* lines could be recovered with the IR-int vector, while with the atpB-int vector the corresponding lines were readily obtained.

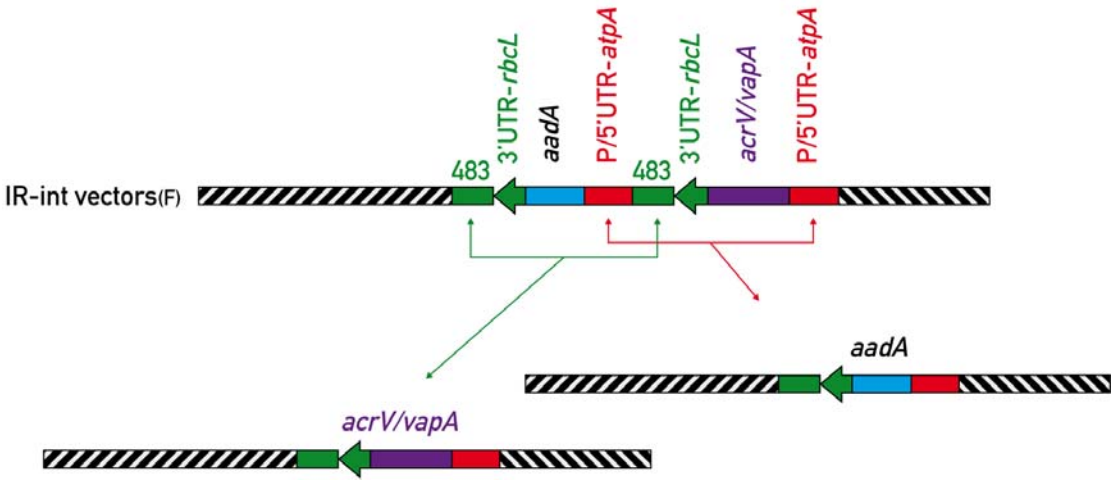


Figure S2.

Fusions to the coding region of *psaA-exon1*

(**A**) The open reading frame of *vapA* was fused to a fragment containing the P/5'UTR and coding sequence of *psaA-exon1* either after codon 20 (*psaA-psaA$_{20}$::vapA*) or after codon 30 (*psaA-psaA$_{30}$::vapA*) in the atpB-int vector. Several independent lines are shown, designated with their numbers (#). For comparison, a transformant with the fusion at the AUG start codon of *psaA-exon1* (*psaA::vapA*) is also included. Total proteins (100 µg) from light-grown cultures were analyzed by SDS-PAGE and immunoblotting with antisera against VapA, or RbcL (large subunit of Rubisco) as indicated on the right. As a reference, 35 ng recombinant VapA protein was mixed with total extract from non-transformed Chlamydomonas (leftmost lane).

(**B**) Similar constructs with *acrV* were prepared and analyzed as in (A).

These transgenic strains contain the IR-int vectors in the forward (F) orientation (see Fig 1B legend)

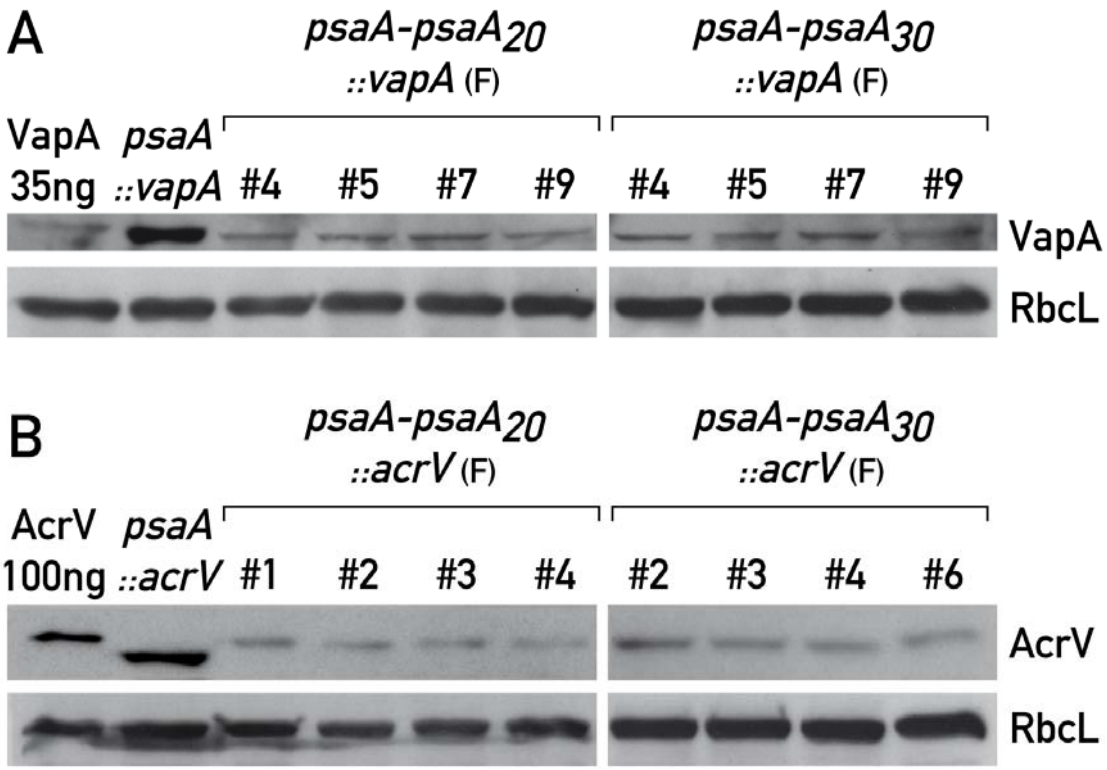


Figure S3.

Comparison of *psaA::acrV* expression in the wild-type and *FUD7*

Expression of AcrV in Chlamydomonas lines obtained by transformation of the wild-type **(A)** or *FUD7* **(B)** with an IR-int vector carrying the chimeric *psaA::acrV* transgene. Several independent lines are shown, designated with their numbers (#). Total proteins (100 µg) from light-grown cultures for the WT transformants and from dark-grown cultures for *FUD7* transformants were subjected to SDS-PAGE and immuno-blotting with anti-AcrV serum, anti- RbcL serum (Rubisco large subunit) or anti-Cytf, as indicated on the right. For calibration, different amounts of recombinant AcrV protein were mixed with total extract from non-transformed Chlamydomonas as indicated on the top left. The migration of the recombinant protein is slightly retarded due to the presence of a poly-histidine tag. These transgenic strains contain the IR-int vectors in the forward (F) orientation (see Fig 1B legend).

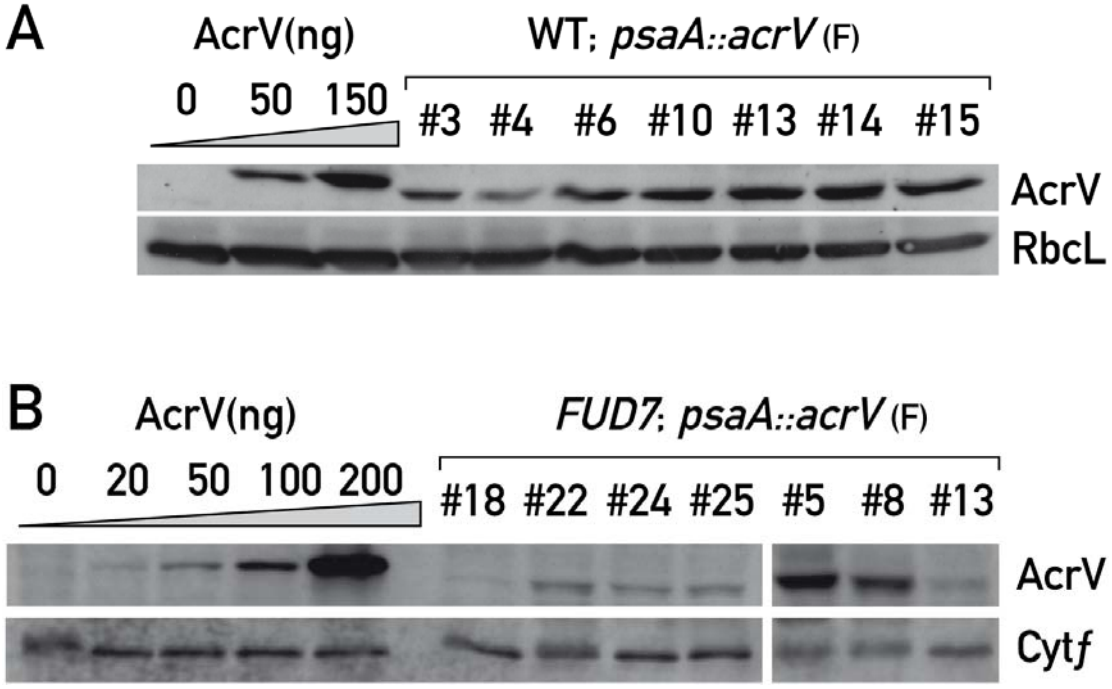


Table S1. Expression levels of VapA and AcrV.

The amounts of transgenic proteins obtained with different constructs and in different genetic backgrounds (*ΔatpB*: *FUD50*; WT: wild type; *ΔpsbA*: *FUD7*) were estimated by immunoblotting, in comparison to known amounts of recombinant proteins (Figure 2, Figure 3 and data not shown). (F) and (R) indicate the orientation of the IR-int vectors in the transgenic strains (see Fig 1B legend).
The symbols correspond to the following approximate amounts (% total cell protein).

(+)	< 0.010 %
+	0.010 - 0.020 %
++	0.030 - 0.050 %
+++	0.075 - 0.150 %
++++	0.200 - 0.300 %
~	indicates the range of expression that was observed in different lines
nd	not determined

	promoter & 5'UTR	Estimated transgenic protein expression				
		light grown			dark grown	
		atpB-int vector in *ΔatpB*	IR-int vector in WT		IR-int vector in *ΔpsbA*	
			(F)	(R)	(F)	(R)
VapA	*atpA*	(+)	nd	++	nd	nd
	psbD	(+)	++	++	nd	nd
	psbA	(+)	(+)	(+)	++	++
	psaA	++	+++	+++	+ ~ ++++	+
	psaA + 20 a.a.	(+)	++	nd	(+)	nd
	psaA + 30 a.a.	(+)	++	nd	(+)	nd
AcrV	*atpA*	+++	nd	nd	nd	nd
	psbD	+	++	++	+++	++ ~ +++
	psbA	(+)	(+)	nd	++ ~ +++	nd
	psaA	+	+++	nd	+ ~ +++	nd
	psaA + 20 a.a.	(+)	++	nd	(+)	nd
	psaA + 30 a.a.	(+)	++	nd	(+)	nd

Table S2. Oligonucleotides used in this work.

Name	Sequence (5’-3’)
5'-VapA	GCTGCTATTGTTGTTGGTTC’
3'-VapA	CATTGACGAATTTCAGCATC
5'-AcrV	CCAGGTGCTTTTGTTGGTTG
3'-AcrV	CAATGTGTAACTGGTGGTGT
EcoRI-AtpB-for	CTCTTCAACGCTATATTCCACAAA
AtpB-rev4	AAGTAAACTTAGGGATTTTAATGCAATGCAATAAA
5'-cpIR	TGGAATTGGATATGGACTAG
3'-cpIR	ACTTATGAAATGCAAGTACC
rbcL-3884	TAAATGGTCTCCAGAACTTGC
5S-fud7	ACTTTACGGGTCGCCGTCTG
5'-psbA	TAAGGGGAAGGGGACGTAGG
5’-XbaI-ClaI-5’-psaA	GTTACCAATCTAGAATCGATAAGCTTTCTTAATTCAACATT
3’-NcoI-5’-psaA-ex1-20	CTTATTTCACCATGGCCGCCCATTTTTCAAAACTTGT
3’-NcoI-5’-psaA-ex1-30	ACTTGTTTCCATGGCATTACGATCAACCGCAATCTT
5’-Bgl-T3	AGTAGATCTTAATACGACTCACTATA
3’-Bam-T7	AATGGATCCGATATCGATAAGCCCGGGATTAACCCTCACTAAAG