Plant Molecular Biology (2006) 60:405–422
DOI 10.1007/s11103-005-4477-z
© Springer 2006

Regulatory sequences of orthologous *petD* chloroplast mRNAs are highly specific among *Chlamydomonas* species

Lynn M. Kramzar, Toby Mueller, Brian Erickson and David C. Higgs*
*Department of Biological Sciences, University of Wisconsin-Parkside, Kenosha, WI 53141, USA (*author for correspondence; e-mail higgs@uwp.edu)*

Received 11 June 2005; accepted in revised form 24 October 2005

Key words: organelles, photosynthesis, phylogenetics, plastids, post-transcription

Abstract

The 5′ untranslated regions (UTR) of chloroplast mRNAs often contain regulatory sequences that control RNA stability and/or translation. The *petD* chloroplast mRNA in *Chlamydomonas reinhardtii* has three such essential regulatory elements in its 362-nt long 5′ UTR. To further analyze these elements, we compared 5′ UTR sequences from four *Chlamydomonas* species (*C. reinhardtii*, *C. incerta*, *C. moewusii* and *C. eugametos*) and five independent strains of *C. reinhardtii*. Overall, these *petD* 5′ UTRs have relatively low sequence conservation across these species. In contrast, sequences of the three regulatory elements and their relative positions appear partially conserved. Functionality of the 5′ UTRs was tested in *C. reinhardtii* chloroplasts using β-glucuronidase reporter genes, and the nearly identical *C. incerta petD* functioned for mRNA stability and translation in *C. reinhardtii* chloroplasts while the more divergent *C. eugametos petD* did not. This identified what may be key features in these elements. We conclude that these *petD* regulatory elements, and possibly the corresponding *trans*-acting factors, function via mechanisms highly specific and surprisingly sensitive to minor sequence changes. This provides a new and broader perspective of these important regulatory sequences that affect photosynthesis in these algae.

Introduction

The 5′ untranslated regions (UTRs) of many chloroplast mRNAs contain important *cis*-acting regulatory sequences (elements) that control chloroplast gene expression. Such regulatory elements have been identified in chloroplast mRNAs of vascular plants and green algae (recently reviewed in Barkan and Goldschmidt-Clermont (2000) and Monde *et al*. (2000)), as well as in Euglena (Betts and Spremulli, 1994; Koo and Spremulli, 1994). Some chloroplast elements promote RNA processing and stability (Monde *et al*., 2000), while others promote translation (Zerges, 2000). A few appear to control both RNA stability and translation (Drager *et al*., 1999; Nickelsen *et al*., 1999). Often, these elements function by interacting with nucleus-encoded *trans*-acting factors, and many work through gene-specific processes (Barkan and Goldschmidt-Clermont, 2000; Erickson *et al*., 2005; Murakami *et al*., 2005; Zerges, 2000).

From previous studies, the *petD* chloroplast mRNA from the green alga *Chlamydomonas reinhardtii* was shown to have three essential RNA regulatory elements (I, II, and III) located in the 362-nt long 5′ UTR (Sakamoto *et al*., 1994a; Higgs *et al*., 1999). Deletions and substitutions combined with chloroplast transformation were used to delimit these elements and determine their function. The 5′ terminal element I sequence promotes RNA stability while elements II and III are internal to the 5′ UTR and promote

translation. Sequences and location of these elements are shown in Figure 1A. The *petD* 5′ UTR is autonomous and can promote RNA stability and translation in chimeric reporter genes transformed into *C. reinhardtii* chloroplasts (Higgs *et al.*, 1999; Sakamoto *et al.*, 1993).

In *C. reinhardtii*, *petD* mRNA accumulation is dependent on the nuclear gene *Mcd1*, and *mcd1* mutants fail to accumulate *petD* mRNA despite normal transcription (Drager *et al.*, 1998). These mutants are non-photosynthetic because the *petD*-encoded subunit IV protein (SUIV) of the cytochrome b_6/f complex is absent. Mcd1 stabilizes the transcript by interacting directly or indirectly with the 5′ terminal element I and blocking a 5′-3′ exoribonucleolytic activity (Drager *et al.*, 1998, 1999). *Mcd1* specifically interacts with *petD* mRNA, as we showed with organellar microarrays (Erickson *et al.*, 2005). The *Mcd1* gene sequence was recently reported for *C. reinhardtii* and the predicted protein is novel (Murakami *et al.*, 2005). Although the *cis*-acting element I sequence and predicted *trans*-acting Mcd1 factor are specific to *petD*, they appear to function by a more general mechanism. For example, the *C. reinhardtii psbD* and *psbB* chloroplast mRNAs parallel *petD* in having 5′ terminal RNA stability elements and dependence on nucleus-encoded regulatory factors (Nac2 and Mbb1, respectively) that stabilize these transcripts against a 5′-3′ exoribonucleolytic degradation (Nickelsen *et al.*, 1999; Vaistij *et al.*, 2000a, b). However, the RNA stability sequences for these three mRNAs are different, and in *mcd1* mutants *psbD* and *psbB* mRNAs accumulate normally (Erickson *et al.*, 2005). The *C. reinhardtii* chloroplast *psbC* mRNA also has a 5′ terminal stability sequence (Zerges *et al.*, 2003). Together, these suggest a general mechanism of chloroplast RNA stability that relies, in part, on gene-specific regulatory factors that interact with specific transcripts via unique *cis* elements.

Homologous non-coding regions typically have lower sequence conservation, as compared to coding regions, a feature that can be used to identify conserved, functional elements in generally divergent non-coding regions (Shaw and Kamen, 1986; Newman *et al.*, 1993; Erlitzki *et al.*, 2002; Shabalina *et al.*, 2004). Among *Chlamydomonas* species, low conservation of non-coding sequences between orthologous chloroplast genes occurs (Turmel and Otis, 1994). Comparative phylogenetic studies that investigate *cis*-element function can be useful in identifying important nucleotides and structures in RNAs. To date, only one report has used *in vivo* testing of orthologous chloroplast 5′ UTRs in heterologous transformed chloroplasts to determine functionality (Nickelsen, 1999). The spinach *psbB* and wheat *psbA* 5′ UTRs were combined with the *aadA* chimeric reporter gene expression tested in *C. reinhardtii* chloroplasts. Despite being transcribed, the reporter mRNAs failed to accumulate, indicating that these vascular plant 5′ UTRs failed to stabilize mRNAs in *C. reinhardtii* chloroplasts. Considering the phylogenetic distance between vascular plants and this green alga, these results are not surprising.

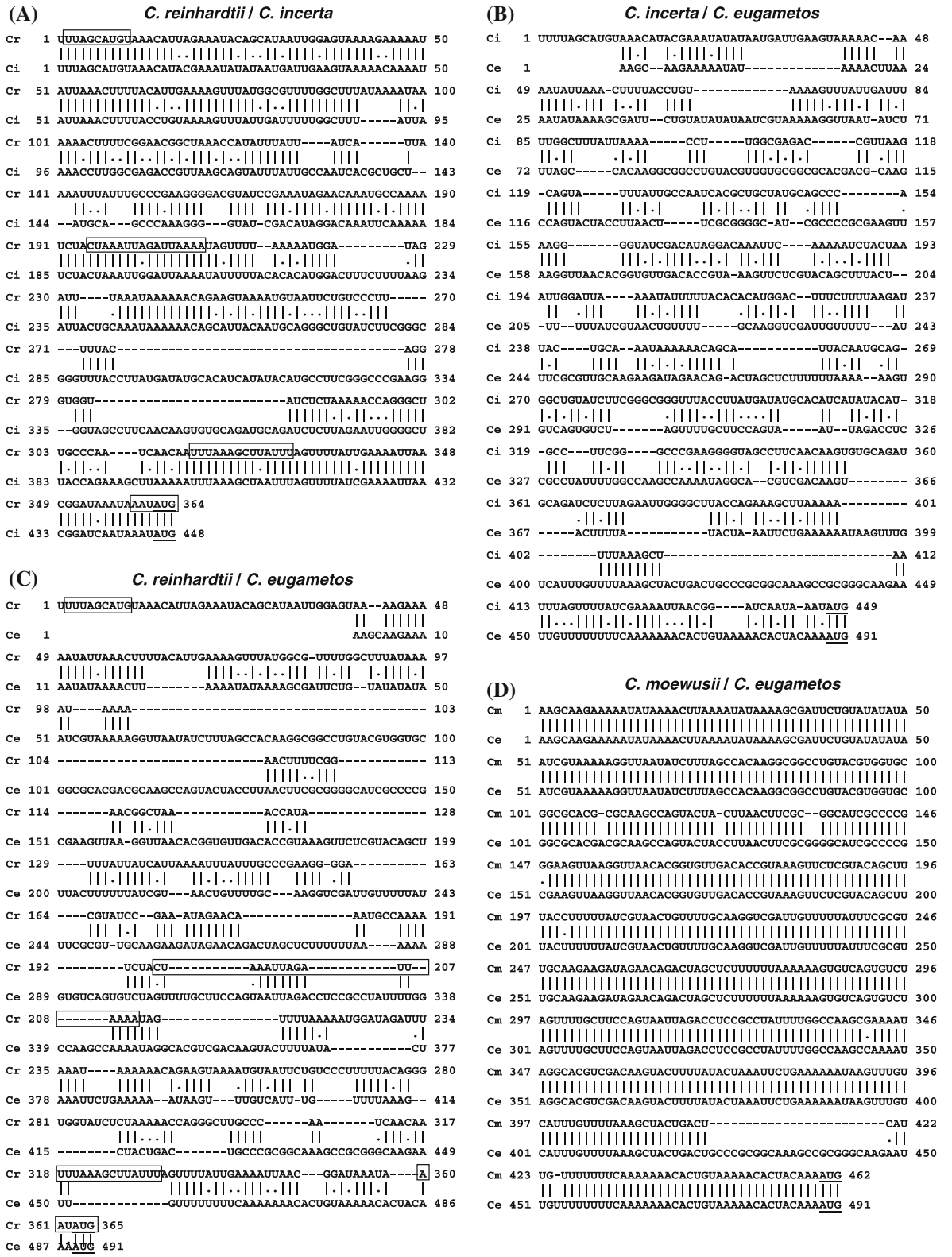
In this report, we have compared the 5′ UTR sequences of orthologous *petD* mRNAs among four *Chlamydomonas* species. Small sequences within these divergent 5′ UTRs have relatively high identity to the three known *petD* regulatory elements (I, II and III) and are in similar relative positions. Chimeric GUS reporter genes were used to test if these divergent *petD* 5′ UTRs function for RNA stability and translation in transgenic *C. reinhardtii* chloroplasts. We conclude that the *C. reinhardtii* chloroplast RNA stability and translational components (*trans*-acting factors) are highly specific and only function with orthologous regulatory sequences that are nearly identical to *C. reinhardtii petD*.

Results

Phylogenetic comparison of petD *5′ UTRs from different Chlamydomonas species and strains*

Total genomic DNA was isolated from five independent strains of *C. reinhardtii* [P17, Minnesota

Figure 1. Pairwise alignments of *petD* 5′ UTR sequences. 5′ UTR sequences include the translation start codon (underlined). 5′ ends were determined from data in Figures 3 and 4. In the *C. reinhardtii* sequence, the known elements I, II, III are indicated with boxes. Abbreviations are Cr = *C. reinhardtii*, Ci = *C. incerta*, Ce = *C. eugametos*, and Cm = *C. moewusii*. (A) *C. reinhardtii–C. incerta* alignment (B) *C. incerta–C. eugametos* alignment (C) *C. reinhardtii–C. eugametos* alignment (D) *C. moewusii–C. eugametos* alignment.


(A)
C. reinhardtii / C. incerta
(B)
C. incerta / C. eugametos
(C)
C. reinhardtii / C. eugametos
(D)
C. moewusii / C. eugametos

(MN), North Carolina (NC), Quebec (Q) and "*C. smithii*" (interfertile with *C. reinhardtii* and believed to be the same species)], and four additional species from the genus *Chlamydomonas* [*C. incerta*, *C. applanta*, *C. moewusii*, and *C. eugametos*]. These representative species from the two major *Chlamydomonas* lineages were selected because they are easily cultured and should provide a phylogenetic gradient of sequence divergence (Harris, 1989; Turmel *et al*., 1993). Independently isolated strains of *C. reinhardtii* were tested for intra-species variation of *petD* 5′ UTR regulatory elements. DNA samples were filter hybridized with a *C. reinhardtii petD* probe to confirm cross hybridization and identify gene fragments (data not shown). The *petD* 5′ UTR sequences from these species and strains were obtained and submitted to GenBank (see Methods for Accession numbers).

Among the *C. reinhardtii* strains (including "*C. smithii*") the *petD* 5′ UTR sequences are identical. Observed DNA fragment polymorphisms among these strains (data not shown) were due to insertions/deletions in the upstream intergenic region between *petA* and *petD*. For *C. reinhardtii* P17 and *C. reinhardtii* NC (strains for which larger amounts of intergenic sequence data are available) the insertions/deletions occurred more than 110 bp upstream of the mapped mRNA 5′ ends (see below) and is consistent with the *petD* promoter being within 100 bp of the 5′ end (Sturm *et al*., 1994). This conserved region includes the prokaryotic "−10" like TATATT sequence 12 bp upstream of the +1 site that was previously proposed as the *petD* promoter (Sakamoto *et al*., 1993). However, there are several other "−10" like sequences in this 110-bp region.

The *C. eugametos* and *C. moewusii* 5′ UTR sequences are ~93% identical to each other but differ significantly from *C. reinhardtii* and *C. incerta petD* 5′ UTRs (Figure 1). *C. reinhardtii* and *C. eugametos petD* have an overall sequence identity, factoring in gaps, of 35%. *C. incerta petD* is 61% identical to that of *C. reinhardtii* and 44% identical to that of *C. eugametos*. The element I, II and III-like sequences in *C. incerta* have high identity to *C. reinhardtii* elements (Figure 2B). Based on GC content of *C. incerta petD* 5′ UTR, the chance occurrence of the *C. incerta* element I-like sequence is once in 2.9×10^6 bp while the element II and III-like sequences are once in 3.7×10^8 bp and once in 2.1×10^7 bp, respectively. Pairwise alignment data are consistent with *C. incerta* and *C. reinhardtii* being members of the Euchlamydomonas morphological group and *C. eugametos* and *C. moewusii* being of the Chlamydella group (Harris, 1989). Finally, *C. incerta petD* has what appears to be homologs of elements I, II and III and in the same relative positions as *C. reinhardtii petD* (Figures 1A and 2A).

Pairwise alignments between *C. reinhardtii* and *C. eugametos* or *C. moewusii* did not reveal element I, II or III-like sequences. This could be due to the absence of these elements or because extensive substitutions and insertions/deletions make accurate global alignment more difficult. As an alternative approach, these 5′ UTRs were searched with the individual *C. reinhardtii petD* element sequences, and this identified *C. eugametos* and *C. moewusii* element II and III-like sequences with relative positions that parallel *C. reinhardtii* (Figure 2). Based on the GC content, chance occurrence of these element II and III-like sequences are 6.8×10^8 and 9.9×10^7 bp, respectively. Although these low chance occurrences suggest element conservation, this will need to be tested to verify function of these sequences.

To correlate sequence variation with potential RNA structures, element I and III-like sequences of *C. incerta* and *C. eugametos* were overlaid onto the *in vivo*-predicted *petD* RNA structures from *C. reinhardtii* (Higgs *et al*., 1999). Figure 2C shows these models and highlights nucleotide substitutions. *C. incerta* and *C. eugametos* were included in this analysis based on the functional data presented in Figures 5 and 6.

Transcript length and cross-hybridization of *petD* mRNAs from *Chlamydomonas* species and strains were determined with RNA filter blots (data not shown). Consistent with the high sequence variation, species-specific 5′ UTR probes had weak cross-hybridization to *petD* mRNAs from divergent species. Transcript length also varied some, and *C. incerta* (~1 kb), *C. moewusii* (~1.1 kb) and *C. eugametos* (~1.1 kb) *petD* mRNA bands were of higher molecular weight than *C. reinhardtii* (~0.9 kb). In *C. eugametos*, an additional smaller *petD* band (~0.7 kb) was detected (data not shown).

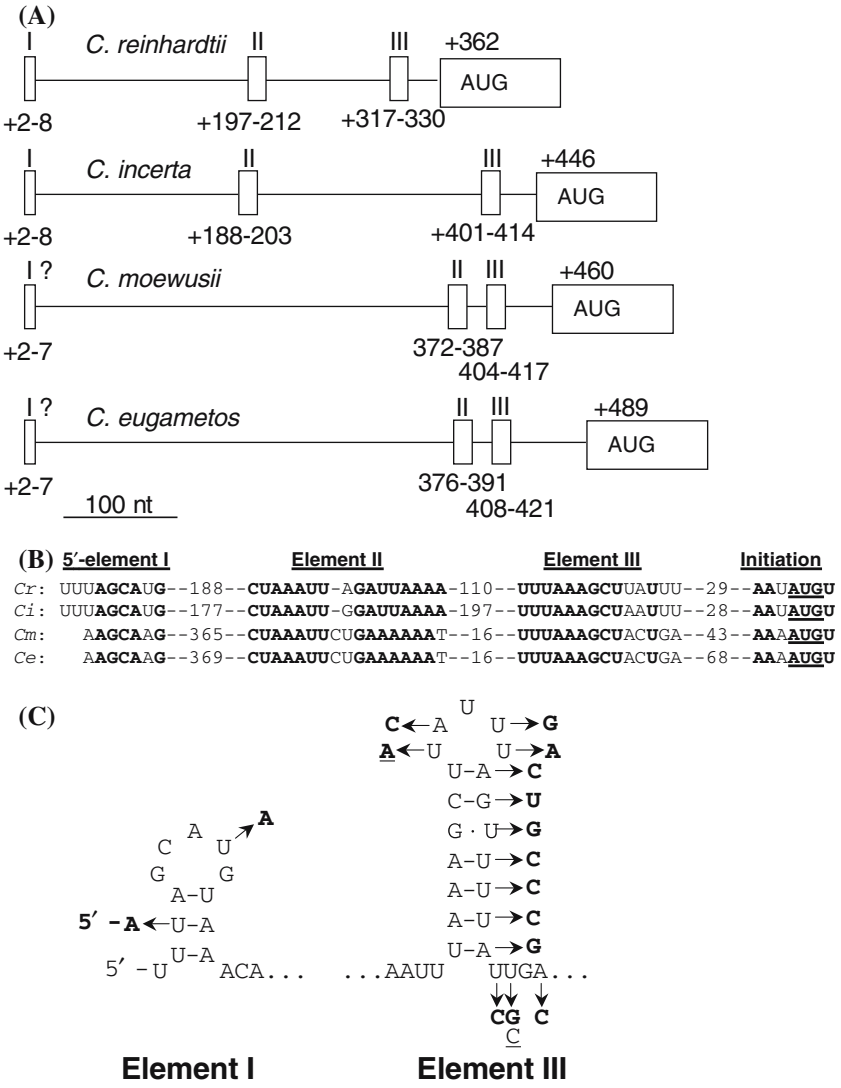


Figure 2. Alignment of *petD* 5′ UTR regulatory sequences from four *Chlamydomonas* species. (A) Diagrams of the different 5′ UTRs are drawn to scale. Known and putative RNA regulatory elements (I, II and III) are shown as open boxes labeled with nucleotide positions with respect to the mapped RNA 5′ ends (Figure 3). Translation start site (AUG) and its position are shown. (B) Known and putative *petD* regulatory sequences from the four species are shown, including the 5′ termini-element I region, elements II and III, and the translation initiation site (underlined). Conserved nucleotides are shown in bold. The number of nucleotides between each of the elements (spacers) is included. (C) Predicted *petD* RNA structures for elements I and III. Structures are modeled from the *C. reinhardtii* sequences (Higgs *et al*., 1999) and are shown in plain font. Substitutions in *C. incerta* are shown as underlined nucleotides and *C. eugametos* are shown as bold nucleotides. The substitutions present in *C. incerta* and *C. eugametos* is shown as a bold, underlined nucleotides.

5′ end mapping of petD *mRNAs*

Because of the close proximity of element I to the 5′ end and the need to delimit 5′ UTRs, RNA 5′ ends were mapped using primer extensions with species-specific ^{32}P-labeled reverse primers. WS5 reverse primer (Table 1) was used for the *C. reinhardtii* strains (Figure 3A). The same 5′ end (+1) was observed in all five, and this location is consistent with previous reports (Sakamoto *et al*., 1993; Higgs *et al*., 1999). For *C. incerta petD*, one abundant primer extension band (+1) of high molecular weight was observed along with two smaller and far less

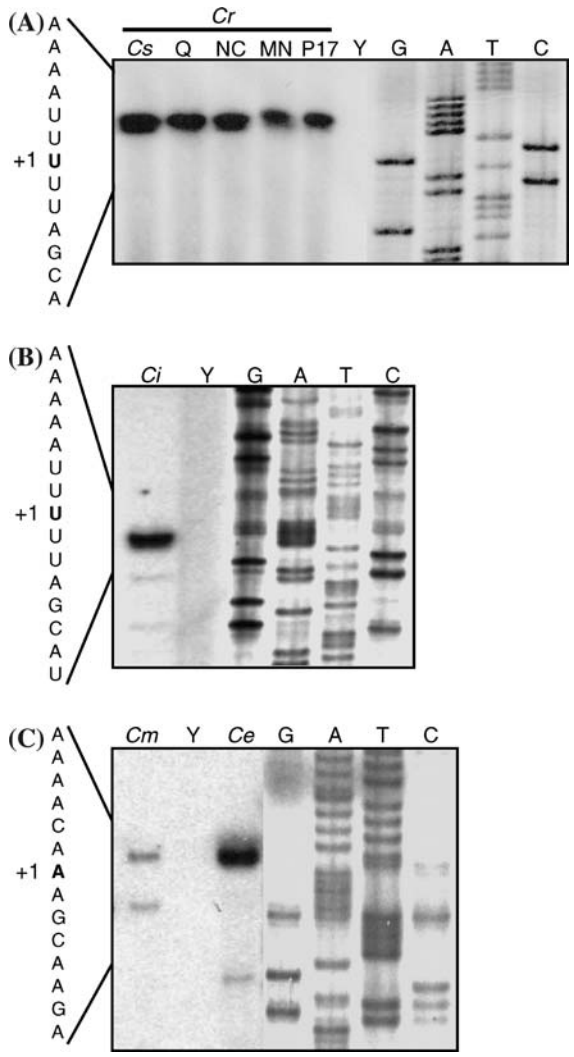


Figure 3. Primer extensions to map *petD* mRNA 5′ ends in the *Chlamydomonas* species. (A) Total RNA (5 μg) from the known *C. reinhardtii* (*Cr*) strains, including "*C. smithii*" (*Cs*), were extended with the WS5 primer. Yeast (Y) RNA was included as a negative control. Four lanes (G, A, T, C) of sequencing ladder using the same primer and plasmid DNA for *petD C. reinhardtii* was included to size extension products. Mapped 5′ ends (+1) and RNA sequences are shown on left. (B) Total RNA (5 μg) from *C. incerta* (*Ci*) was extended with the Cinc1 primer, and yeast (Y) total RNA was included as a negative control. Four lanes (G, A, T, C) of sequencing ladder using the same primer and plasmid DNA for *petD C. incerta* was included to size extension products. The mapped 5′ end (+1) and RNA sequence are shown on left. (C) Total RNA (5 μg) from *C. eugametos* (*Ce*) and *C. moewusii* (*Cm*) were extended with the LK11 primer, and yeast (Y) total RNA was included as a negative control. Four lanes (G, A, T, C) of sequencing ladder using a different primer and plasmid DNA for *petD C. reinhardtii* was included to size extension products. Note, a different sequence ladder was used due to repeatedly poor sequencing ladder from the LK11 primer. Mapped 5′ ends (+1) and RNA sequences are shown on left.

abundant bands at +9 and +18 (Figure 3B). The *C. incerta* +1 band corresponds identically to the *C. reinhardtii* 5′ end sequence, but with a longer 5′ UTR. Despite identical terminal sequences between these two species, the minor bands at +9 nt and +18 nt were not observed in *C. reinhardtii*, suggesting a possible difference in *trans*-acting factors that affect RNA processing and/or degradation.

For *C. moewusii* and *C. eugametos* 5′ end mapping, the largest and most prominent extension product was identical in length and terminal sequence in these two species. These *petD* 5′ ends differ in sequence and position from those

in *C. reinhardtii* and *C. incerta*. Additional, smaller primer extension products from *C. moewusii* and *C. eugametos* were detected and may correspond to 5′ ends of less abundant and smaller *petD* mRNAs (Figure 3C). However, these could be artifacts resulting from early reverse transcriptase termination. To distinguish these possibilities, an RNase protection was used as an independent 5′-end mapping method with *C. eugametos* RNA. A ^{32}P-labeled antisense RNA probe was made with the upstream *C. eugametos* sequence starting at −302 (*Cla*I site) with respect to the mapped 5′ end and continuing 96 nt downstream of the translation start site at +489 (Figure 4B). The largest and most abundant protected band (Figure 4A) corresponded to the prominent +1 primer extension-mapped 5′ end (Figure 3C). Two additional bands of medium intensity were detected with 5′ ends at +55 and +225. The 5′ end at +225 corresponds in size with the 0.7-kb *C. eugametos petD* mRNA detected by RNA filter hybridization (data not shown). In contrast, a transcript with 5′ end at +55 would be difficult to distinguish by filter hybridization from the 1.1-kb full-length mRNA. Another faint RNase protected band at +25 was reproducibly observed, and this is consistent with the low-intensity primer extension band at this size (Figure 3C). Additional faint bands were variably observed, and these were seen in the probe-alone lane with longer exposure.

Taken together, these data suggest that *C. eugametos*, and likely C. *moewusii* and *C. incerta*, differ from *C. reinhardtii* in accumulating multiple *petD* mRNAs with heterogenous 5′ ends. It is not known how many or which of these multiple *petD* mRNAs are translated. To test if the multiple 5′ ends of these *C. eugametos* and *C. moewusii* transcripts are conserved, the terminal sequences were aligned (Figure 4C). Although the +1 terminal sequences are identical, there is not a clear conserved sequence among the 5′ ends of the less abundant, small transcripts. Finally, the mapped *petD* 5′ ends (+1) from the different species are consistent with the size of *petD* mRNAs detected by filter hybridization (data not shown).

GUS reporter genes to test functionality of divergent petD5′ *UTRs in* C. reinhardtii

DG2 is a chimeric GUS reporter gene previously used to study *petD* 5′ UTR structure and function

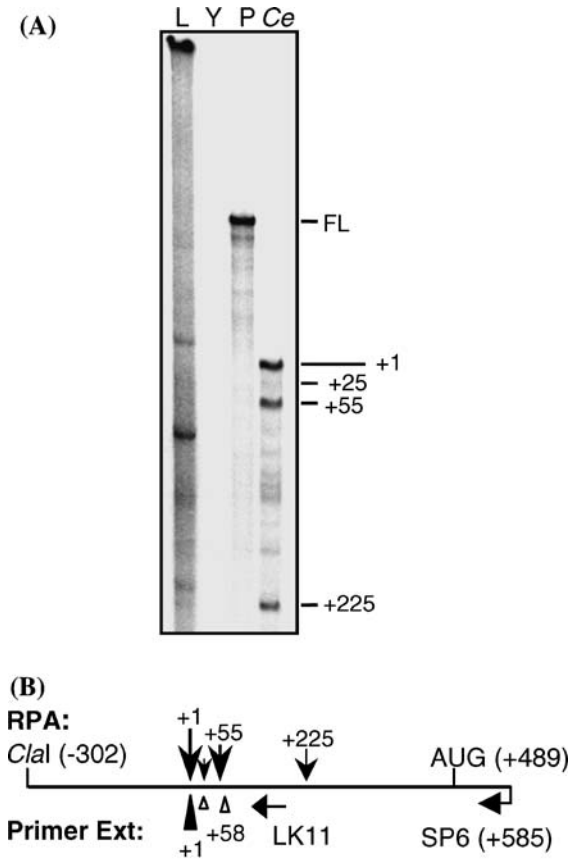


Figure 4. RNase protection assay to size the multiple *C. eugametos petD* mRNAs. (A) Total RNA (10 μg) from *C. eugametos* (*Ce*) was analyzed along with a non-digested probe control (P), yeast negative control (Y), and end-labeled RNA size ladder (L). The probe full-length (FL) and predicted nucleotide positions (+1, +25, +55 and +225) for the prominent protected products are shown on the right. Numbers indicate size of protected products relative to +1 site from the *Ce* primer extension data (Figure 3C). (B) Diagram summarizing the RNase protection assay (RPA) and primer extension data (Figure 3C). Arrows above the line indicate 5′ ends mapped by RPA, and size of arrow represents relative abundance of each. Arrowheads below line indicate 5′ ends mapped by primer extensions, and size of arrowhead represents relative abundance. Reverse horizontal arrow (LK11) shows reverse primer used for extensions. Nucleotide numbers are based on +1 site mapped by primer extension. (C) Alignment of different *C. eugametos* and *C. moewusii* 5′ end sequences. Numbers indicate position of 5′ ends.

in *C. reinhardtii* chloroplasts (Sakamoto *et al*., 1993; Higgs *et al*., 1999). It contains the *C. reinhardtii petD* promoter and 5′ UTR (including *petD* translation start site) fused in-frame to the *Escherichia coli uidA* coding region (encoding β-glucuronidase, GUS) and the *C. reinhardtii* chloroplast *rbcL* 3′ UTR (Sakamoto *et al*., 1993). To test if putative regulatory elements in the orthologous *petD* 5′ UTRs can functionally interact with *trans* factors and promote RNA stability and translation in *C. reinhardtii* chloroplasts, a series of chimeric GUS reporter genes were tested (Figure 5A). These had combinations of divergent *petD* 5′ UTR sequences from *C. reinhardtii*, *C. incerta*, and *C. eugametos* added to the DG2 plasmid. These three species were used since their sequences span the phylogenetic range. Because *C. eugametos* has multiple *petD* mRNAs with different 5′ ends

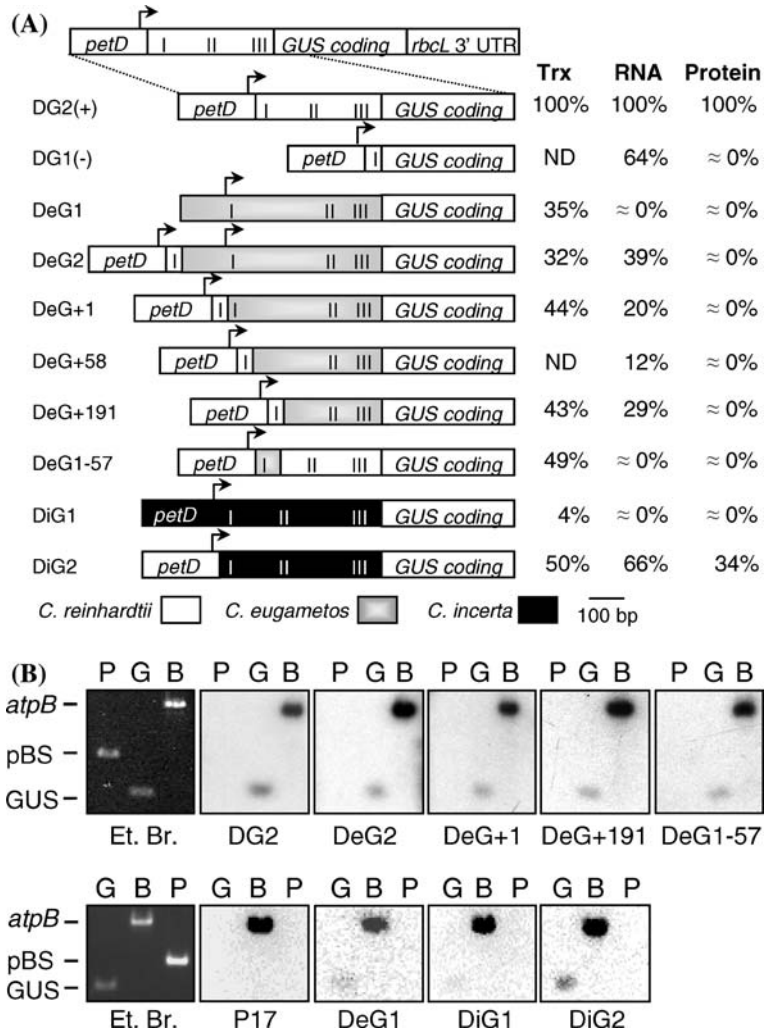


Figure 5. Expression of chimeric GUS reporter genes tested in transformed *C. reinhardtii* chloroplasts. (A) Gene diagrams of the GUS reporter genes, names are shown at left of each diagram. Top diagram shows the full-length DG2 gene with *C. reinhardtii petD* promoter and 5′ UTR, GUS coding region, and *C. reinhardtii rbcL* 3′ UTR. Remaining diagrams show only promoter, 5′ UTR and 5′ portion of coding region. Bent arrows represent promoters. Roman numbers (I, II and III) indicate known or putative regulatory elements. *petD* 5′ UTR sequences from the three species are indicated by shading, as shown in key at bottom. Indicated amounts of transcription (Trx), RNA accumulation (RNA) and GUS protein are based on data in (B), Figure 6A, B, respectively. (B) *In vivo* labeling directly measures transcription from reporter genes. Representative ethidium bromide stained gel is shown at left of each row. DNA fragments are chloroplast *atpB* (B), GUS coding region (G), and pBluescript (P) negative control. ND indicates data not determined.

Table 1. Primers.

Name	Gene	Strand	Source	Position[a]	Sequence[b]
Cinc1	*petD*	–	*C. incerta*	+90	GCTTAACGGTCTCGCC
Cinc2	*petD*	+	*C. incerta*	–377	AACTCGAGCTGCAGCCCCGGAAAG
Cinc3	*petD*	+	*C. incerta*	+25	AAAGATCTAATGATTGAAGTAAAAACAA
Cinc4	*petD*	–	*C. incerta*	+432	TTCCCGGGACTGACATATTTATTGATCCG
GUS2	*UidA*	–	*E. coli*	+113 (ATG)	TGCCCGGCTTTCTTGTAACG
LK1	*petD*	+	*C. eugametos*	–366	TACTCGAGAGATCTAAGCTTTTGCGCAACTAAAACAAA
Lk2	*petD*	–	*C. eugametos*	+473	TTCCCGGGACAGACATTTTGTAGTGTTTTTAC
LK4	*petD*	–	*C. eugametos*	+273	GCTAGTCTGTTCTATCTTC
LK5	*petD*	+	*C. eugametos*	+151	TAAGATCTGAAGTTAAGGTTAACACGGTG
LK7	*petD*	+	*C. eugametos*	+191	TAAGATCTGTACAGCTTTACTTTTTTATCGTAACT
LK8	*petD*	–	*C. eugametos*	+545	GTGTTATGACCAAAACCTTTTGC
LK9	*petD*	+	*C. eugametos*	+127	TGCGGCGCACGACGCAAG
LK10	*petD*	–	*C. eugametos*	+174	ACAGGCACTGCAGCCATGGCAA
LK11	*petD*	–	*C. eugametos*	+100	CGCCGCACCACGTAC
LK12	*petD*	+	*C. eugametos*	+1	TTAGATCTAAGCAAGAAAAATATAAAACTTAAAA
LK13	*petD*	+	*C. eugametos*	+58	TTAGATCTAAAGGTTAATATCTTTAGCCAC
LK14	*petD*	+	*C. reinhardtii* + *C. eugametos*	–20 to +57	GGCCGCATAATATATATTAAAA ATTAAGCAAGAAAAATATAAAACTTAAAATATAA AAGCGATTCTGTATATATAATCGTAAA
LK15	*petD*	–	*C. reinhardtii* + *C. eugametos*	–20 to +57	GATCTTTACGATTATATATACAG AATCGCTTTTATATTTTAAGTTTTATATTTTTCTTG CTTAATTTTTAATATATATTATGC
LK16	*petD*	–	*C. eugametos*	–179	ACGTTCAGGTTTAGTCAAAGCG
LK17	*petD*	–	*C. eugametos*	–261	TTATAACTTTTAGAAAAAACAAGG
WS5	*petD*	–	*C. reinhardtii*	+67	TTAGATCTGCCAAAACGCCATAAACTTTTC
WS7	*petD*	+	*C. reinhardtii*	+166	CCAGATCTGAACAAATGCCAAAATCTACT
WS10	*petD*	–	*C. reinhardtii*	+380	TTAGATCTTGGATCGCTTAAATCAGG
WS12	*petD*	+	*C. reinhardtii*	–120	CACTCGAGTGCCACTGACGTCCCGTCA
WS13	*petD*	+	*C. reinhardtii*	–2	CCAGATCTTTTTTAGCATGTAAACATTAGAAATA

[a] Position is the 5′ most nucleotide on + strand that corresponds to 5′ nt of + stranded primers or nt that pairs with 3′ end of – stranded primers. Numbering is with respect to RNA 5′ end, except GUS2 primer which is with respect to translation start site.
[b] All primers shown in 5′-3′ orientation.

(Figures 3C and 4B) and the uncertainty of which is translated, multiple *C. eugametos* constructs with different 5′ UTR lengths were tested (Figure 5A). The 5′ UTR length may indeed be important because of the link between 5′-processing of chloroplast mRNAs and a higher frequency of translation from the smaller, processed transcripts for some genes (Barkan *et al.*, 1994; Sturm *et al.*, 1994; Shapira *et al.*, 1997; Bruick and Mayfield, 1998). To test transcription activity from the *C. eugametos* and *C. incerta petD* upstream/promoter sequences, the DeG1 and DiG1 constructs were made with the *C. reinhardtii petD* promoter removed (Figure 5A).

The GUS reporter gene plasmids also have a wild-type copy of the *C. reinhardtii* chloroplast *atpB* gene and flanking chloroplast genomic DNA. These were biolistically transformed into chloroplasts of the non-photosynthetic *C. reinhardtii* CC-373 (ΔatpB) mutant with a deletion in the chloroplast *atpB* gene (Sakamoto *et al.*, 1993). The plasmid *atpB* gene and flanking DNA directs homologous recombination that inserts *atpB* and GUS genes at the *atpB* locus (Sakamoto *et al.*, 1993). Transformed cells were selected for photoautotrophic growth on minimal medium. Homoplasmic transformants, in which all chloroplast genomes contain the GUS reporter gene, were obtained after several repeated rounds of single-colony subcloning and selection on a minimal medium that requires *atpB* expression. Transformants were tested by PCR and DNA filter hybridization to confirm insertion of the GUS reporter gene and homoplasmicity (data not shown).

Transformants were analyzed for reporter gene transcription (Figure 5B), RNA accumulation (Figure 6A) and translation (Figure 6B). Figure 5A summarizes these data. To test GUS

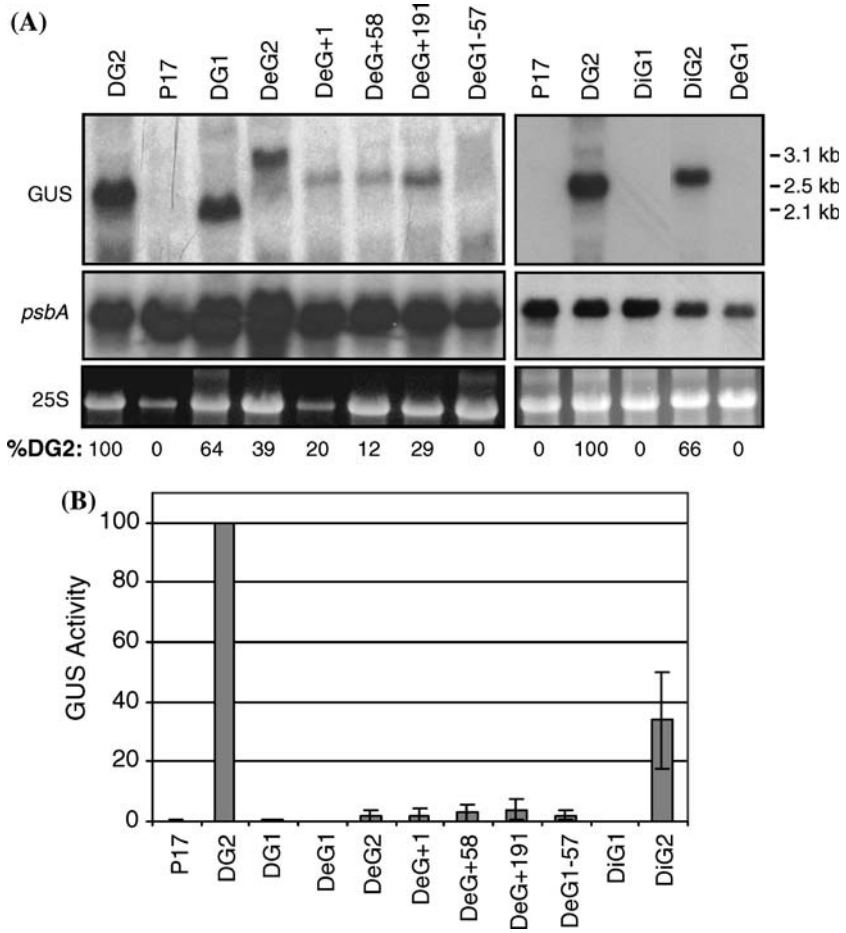


Figure 6. RNA and protein expression from GUS reporter genes. (A) RNA Filter hybridization with total RNA (10 μg) from each strain hybridized to antisense GUS RNA probe. Wild-type (P17) control is RNA from non-transformed *C. reinhardtii* P17 cells. Sizes of RNAs in kb are shown on right. Hybridized chloroplast *psbA* mRNA and ethidium bromide stained 25S rRNA are shown as loading controls. Average quantified relative GUS mRNA (%DG2) for each is shown at bottom. (B) GUS enzyme activity was determined with fluorometric assay. Activities were standardized to mg of total protein and finally made relative to the DG2 positive control. Data are mean and error bars show standard deviation of the mean.

transcription, *in vivo* RNA labeling experiments were conducted. Permeabilized, frozen cells were thawed in the presence of ^{32}P-rUTP nucleotides and transcription was continued for 15 min. Radioactive total RNA was extracted and hybridized to excess amounts of DNA fragments of GUS, *atpB* (reference gene), and pBS plasmid (negative control). For each transformant, intensity of the hybridized GUS RNA was quantified and compared to hybridized *atpB* reference RNA to determine relative GUS transcription (Figure 5B). As a negative control, wild-type P17 cells were tested, and as expected no signal was detected. DG2 (positive control), DeG1, DeG2, DeG + 1, DeG + 191, DeG1-57, and DiG2 were tested and had significant amounts of GUS transcription. DiG1, in contrast, had near background (4% DG2) transcription. This may be due to DiG1 lacking the *C. incerta petD* promoter in the 377 bp of upstream sequence. The 377-bp *C. incerta* upstream region in DiG1 has five prokaryotic "−10" (TATAAT) promoter-like sequences, the nearest is 13 nt upstream of the +1 site with the sequence TATAT. However, it does not have the TATATT sequence at −55 found in *C. reinhardtii*, and perhaps this sequence is necessary in *C. reinhardtii*. Instead, the *C. incerta petD* promoter may be upstream of the 377-bp sequence. Related to this, wild-type levels of *C. reinhardtii petD* can be transcribed by the

endogenous upstream *petA* promoter (3.5 kb from *petD*), allowing the *petD* promoter to be deleted without reducing transcription (Sturm *et al*., 1994). Perhaps, *C. incerta* parallels this situation with a far-upstream promoter. Alternatively, the DiG1 377-bp fragment may contain the *C. incerta* promoter, but with sufficient sequence divergence that it does not function in *C. reinhardtii*. Current data do not distinguish between these two possibilities.

The promoter from the more distantly related *C. eugametos* did function in *C. reinhardtii* (Figure 5B). DeG1 had 35% transcription as compared to DG2. Because the DeG reporter mRNA was very unstable (see below), the mRNA 5′ end could not be determined. Interesting, this promoter has a partial −10 promoter sequence (TATAAA) at −55 that is similar to *C. reinhardtii*, but different from *C. incerta*.

C. reinhardtii transformants were tested for GUS mRNA accumulation and transcript size using RNA filter hybridization (Figure 6A). The non-transformed *C. reinhardtii* P17 strain was tested, and as expected, no GUS mRNA was detected, while DG2 accumulated high levels of GUS mRNA (Figure 6A; Sakamoto *et al*., 1993). Each of the DeG2, DeG + 1, DeG + 58, DeG + 191 and DiG2 transformants accumulated varying amounts of GUS mRNA of expected sizes. As loading controls, the hybridized *psbA* chloroplast mRNA and ethidium bromide stained 25S rRNAs are shown (Figure 6A). Transformants DeG1, DeG1-57, and DiG1 had no detectable GUS mRNA. This lack of mRNA in DeG1 and DeG1-57 is likely due to absence of RNA stabilizing sequences at the 5′ ends. In DiG1, however, lack of mRNA is due primarily to the little to no transcription (Figure 5B).

Additional *C. eugametos* reporter genes (DeG2, DeG + 1, DeG + 58, and DeG + 191) were made to have varying lengths of *C. eugametos* 5′ UTR sequences plus the *C. reinhardtii* element I sequence to promote RNA stability. These strains all accumulated GUS mRNA of expected lengths. To confirm that these transcripts had the expected *C. reinhardtii* element I sequence, RT-PCR detection of DeG2 and DeG + 1 mRNAs was done using an element I-specific forward primer (WS13) and *C. eugametos*-specific reverse primers (LK11 or LK16). RT-PCR products of the correct size were detected, showing the presence of element I (data not shown).

To test the RNA stabilizing function of the *C. eugametos* 5′ end, DeG1-57 gene was made by replacing the RNA stability element I sequence (position +1 −25) in the original DG2 with the *C. eugametos petD* terminal sequence (position +1 −57). Despite the relatively high level of transcription (49% DG2), DeG1-57 mRNA failed to accumulate. This demonstrates directly that the *C. eugametos* terminal sequence (+1 −57), including the partial element I sequence, is not capable of stabilizing GUS mRNA in *C. reinhardtii* chloroplasts. These data combined with those for DeG1 show that the *C. eugametos petD* terminal sequence is sufficiently different from *C. reinhardtii* that it does not function to stabilize mRNA in *C. reinhardtii*. One possible explanation for this difference is that the *petD* mRNA stability mechanisms are in general similar between these species, but the *C. eugametos* 1-57 sequence differences prevent it from interacting with *C. reinhardtii* stability factors, such as Mcd1 (Erickson *et al*., 2005; Murakami *et al*., 2005). A second possibility is that *C. eugametos petD* relies on an entirely different mechanism with non-homologous RNA *cis* elements and *trans* factors.

To test translation from these GUS transcripts, the quantitative fluorometric GUS assay was performed as previously reported (Sakamoto *et al*., 1993). To control for variability between assays, enzyme activities were made relative to the DG2 positive control (Figure 6B). Although most strains accumulated significant amounts of GUS mRNA, only DG2 and DiG2 (34 ± 16% DG2) had GUS activities significantly greater than the non-transformed P17 and the non-translated DG1 GUS mRNA (Sakamoto *et al*., 1993). In DiG2, the 34% level of GUS protein was translated from the 66% level of GUS mRNA (Figure 6A). This suggests that DiG2 mRNA with the *C. incerta petD* 5′ UTR sequence is translated at about one-half the frequency of DG2.

The absence of translation from the four accumulating *C. eugametos* GUS mRNAs in different transformants is interesting with respect to chloroplast translation initiation mechanisms. These mRNAs all have correct in-frame translation start sites and the element II and III-like sequences, and despite these features, no GUS protein activity was detected. We conclude that

C. eugametos petD 5′ UTR sequences failed to promote translation initiation in *C. reinhardtii* chloroplasts, but the underlying reason is not known.

Methods

Strains and culture conditions

The *C. reinhardtii* strains Minnesota (CC-1952), North Carolina (CC-2931), Quebec (CC-2935) and "*C. smithii*" (CC-1373) and the species *C. incerta* (CC-3347), *C. applanta* (CC-2988), *C. moewusii* (CC-55), and *C. eugametos* (CC-1419) were obtained from the *Chlamydomonas* Genetics Center, Duke University. The non-photosynthetic strain CC-373 (Shepherd *et al.*, 1979) has a deletion in the chloroplast *atpB* gene and was used to transform GUS reporter genes. The *C. reinhardtii* P17 wild-type strain (CC-125) was derived from CC-373 by complementing it with the wild-type *atpB* gene (Stern *et al.*, 1991). The GUS reporter strains DG1 and DG2 were previously reported (Sakamoto *et al.*, 1993). Unless otherwise stated, all strains and species were grown in TAP medium (Harris, 1989) under constant low-level fluorescent light.

DNA isolation, sequencing and PCR

Genomic DNA for PCR was prepared by a small-scale isolation procedure as previously described (Drager *et al.*, 1996). The *petD* 5′ UTRs from *C. reinhardtii* MN, Q, "*C. smithii*", and *C. moewusii* were obtained by PCR amplification, as described below, with the following primers: *C. reinhardtii* MN, Q, and "*C. smithii*" (WS12–WS10), and *C. moewusii* (LK1–LK8), see Table 1 for details. The amplified 5′ UTR fragments that included variable amounts of both upstream and coding regions were sequenced. To control for possible mutations arising during PCR, two independently amplified PCR fragments were sequenced for each fragment and shown to be identical.

The remaining *C. reinhardtii* NC and *C. incerta petD* fragments were sub-cloned from isolated chloroplast DNA. First, large-scale total DNA samples were prepared (Herrin and Worley, 1990); from this, chloroplast DNA was purified by CsCl-gradient centrifugation and visualized with *bis*-benzimide as previously described (Roffey *et al.*, 1991). The *C. reinhardtii* NC and *C. incerta* *Eco*RV–*Pst*I *petD* fragments were digested from chloroplast DNA, agarose gel-purified, and ligated into *Eco*RV–*Pst*I digested pBluescript II KS+ (pBSKSII+; Stratagene, La Jolla, CA) plasmid. The *petD* inserts were sequenced, and the 5′ ends of the coding regions were used to confirm *petD* identity. Sequence identities were determined by pairwise alignment with the European Bioinformatics Institute's EMBOSS (www.ebi.ac.uk) application using the Needleman and Wunsch algorithm (Needleman and Wunsch, 1970). Previously reported *petD* sequences for *C. reinhardtii* P17 strain (Sakamoto *et al.*, 1993; L05506) and *C. eugametos* (Turmel *et al.*, 1989; X14503) were used.

For PCR, 50 μl reactions contained 250 ng of each primer, 3 mM MgCl, *Taq* DNA polymerase (Promega, Madison, WI USA) and appropriate DNA templates. Depending on the reaction, either 1 ng of plasmid DNA or 1 μl of a 1/10 dilution of small-scale isolated genomic DNA was used as template. Reactions were run for 35 cycles with an annealing temperature of 55 °C.

RNA isolation, primer extension and RNase protection

Total RNA was isolated with Tri-Reagent (Sigma Chemicals, St. Louis, MO) from 10 ml of cells grown to 2×10^6 cells/ml as previously described (Drager *et al.*, 1999). To map 5′ ends of the different *petD* mRNAs, primer extensions with total RNA (5 μg) were performed as previously described (Sakamoto *et al.*, 1994b), and species-specific ^{32}P-end-labeled primers were used (Figure 3). For sequencing size ladders, plasmids with corresponding *petD* genes were sequenced with the same primer used for the primer extension. Labeled, denatured extension products were fractionated in a 6% denaturing polyacrylamide gel (Ausubel *et al.*, 1988), dried, and visualized by exposure to X-ray film at −75 °C with an intensifying screen (Fisher Scientific). Two or more independent extensions were conducted for each species and strain.

As a second method to map the 5′ ends of the *C. eugametos petD* transcripts the RNase Protection Assay Kit III from Ambion, Inc. (Austin, TX,

USA) was performed as described by manufacturer. For probe synthesis, a 837-bp DNA fragment, with the 5′ UTR and adjacent sequences, was PCR amplified from *C. eugametos* genomic DNA using LK1 and LK2 primers (Table 1). This PCR product was ligated into the *Eco*RV-digested, TA-tailed pBluescript II SK (Stratagene) to produce the pCeD plasmid. pCeD was digested at the *Cla*I site 302 bp upstream of the predicted RNA 5′ end, and the radiolabeled, antisense RNA was synthesized with 100 U SP6 RNA polymerase (Promega) and 40 μCi [α^{32}P]-UTP (PerkinElmer), according to manufacturer's protocol (Promega). The full-length 887-nt RNA probe was gel purified and hybridized to 10 μg total *C. eugametos* RNA for 16 h at 42 °C in the provided hybridization buffer. A 1/500 dilution of RNaseA/T_1 stock solution was made with RNase Digestion III Buffer (Ambion, Inc.) and used to digest single-stranded RNA for 30 min at 37 °C. Low-molecular weight (0.16–1.77 kb) RNA size ladder (Invitrogen, Carlsbad, CA) was end labeled with [γ^{32}P]-ATP and T4 polynucleotide kinase, according to the manufacturer's protocol. RNase protected fragments and RNA size ladder were simultaneously electrophoresed in a denaturing 5% polyacrylamide gel (Ausubel *et al.*, 1988) and exposed to X-ray film with an intensifying screen at −75 °C.

GUS reporter gene constructs and Chlamydomonas *chloroplast transformation*

Different combinations of *petD* 5′ UTRs from *C. reinhardtii*, *C. incerta* or *C. eugametos* were ligated into the DG2 reporter gene, or its derivatives, so that the *petD* translation start site was in-frame with the bacterial *uidA* coding region that encodes β-glucuronidase (GUS) (Sakamoto *et al.*, 1993). For DeG1 and DeG2 reporter genes, *C. eugametos petD* 5′ UTR was PCR amplified from *C. eugametos* genomic DNA with LK1 and LK2 primers (Table 1). The LK1 primer added a unique *Bgl*II site at −367 (with respect to the RNA 5′ end) and the LK2 primer added an in-frame *Sma*I site at +501, just 3′ of the translation start site. This *C. eugametos Bgl*II–*Sma*I fragment, containing the putative promoter, 5′ UTR and translation start site, was ligated into the DG2 plasmid replacing the *C. reinhardtii petD* 5′ UTR (*Bgl*II–*Sma*I), forming DeG2 reporter gene. DeG2 plasmid was *Xho*I–*Bgl*II digested, ends were blunted, and plasmid was re-circularized to form the DeG1 reporter gene with the upstream *C. reinhardtii petD* promoter and element I removed. To make the DeG + 1, DeG + 58, and DeG + 191 reporter genes, the *C. eugametos* LK12 (DeG + 1), LK13 (DeG + 58), and LK7 (DeG + 191) forward primers were combined with the LK2 reverse primer to PCR amplify different lengths of 5′ UTR from *C. eugametos* genomic DNA. All three PCR fragments were *Bgl*II–*Sma*I digested and ligated into DG2 to replace the *C. reinhardtii petD* 5′ UTR, leaving the *C. reinhardtii petD* promoter and element I. To make DeG1-57, DG2 plasmid was first modified to form the intermediate plasmid named DG2A by replacing the DG2 *Xho*I–*Bgl*II upstream fragment with the LS-20 *C. reinhardtii petD Xho*I–*Bgl*II upstream fragment [that has the engineered neutral *Not*I site at −20 with respect to the RNA 5′ end (Drager *et al.*, 1999)]. Complementary oligonucleotides LK14 and LK15 (Table 1) with *C. reinhardtii petD* sequence from −20 to −1 and *C. eugametos petD* sequence from +1 to +57 were annealed to generate a chimeric 78-bp double-stranded fragment with compatible *Not*I and *Bgl*II ends. This was ligated into the *Not*I–*Bgl*II digested DG2A plasmid forming DeG1-57.

To make DiG1, Cinc2 forward primer (−377; with an *Xho*I site) and Cinc4 reverse primer (with a *Sma*I site just downstream of the translation start) were used to PCR amplify the *petD* promoter region and 5′ UTR from *C. incerta* genomic DNA. This was *Xho*I–*Sma*I digested and ligated into *Xho*I–*Sma*I digested DG2, replacing *C. reinhardtii petD* promoter and 5′ UTR. This insertion placed the *C. incerta* translation start site in-frame with the GUS coding region. To make DiG2, Cinc3 forward primer (+25; with a *Bgl*II site) and Cinc4 reverse primer (with a *Sma*I site) were used to PCR amplify the *petD* 5′ UTR from *C. incerta* genomic DNA. This was *Bgl*II–*Sma*I digested and ligated into *Bgl*II–*Sma*I digested DG2. Because the first 25 bp of *C. incerta petD* is identical to *C. reinhardtii petD*, the DiG2 reporter gene has *C. reinhardtii petD* promoter sequence up to +1 followed by *C. incerta* 5′ UTR from +1 to just 3′ of the translation start site.

Plasmids with GUS reporter genes were isolated with Qiagen's Midi Plasmid 100 kit (Qiagen Inc., Valencia, CA, USA), according to manufacturer's protocol. All were sequenced to confirm

that the reporter gene sequences and the translation fusions were correct. GUS reporter genes were introduced into the chloroplast of the non-photosynthetic *Chlamydomonas reinhardtii* CC-373 (Δ*atpB*) strain by particle bombardment and homologous recombination (Kindle *et al*., 1991; Higgs *et al*., 1999).

In vivo *RNA labeling, GUS mRNA analysis and GUS fluorometric assay*

To directly test transcription from the different *petD* promoters in these GUS reporter genes, *in vivo* labeled (transcription run-off) GUS mRNA was quantified and compared to the labeled chloroplast *atpB* reference mRNA as previously described (Gagne and Guertin, 1992). Total, labeled RNA was isolated with Tri-Reagent (Sigma, St. Louis, MO) as described above, and total cpm (1×10^6) of RNA was hybridized at 65 °C (Church and Gilbert, 1984) to excess amounts of GUS, *atpB* and pBluescript (negative control) DNA immobilized on GeneScreen nylon membranes (PerkinElmer-NEN). For GUS DNA, the 210-bp *uidA Eco*RV internal coding region fragment was digested from the pBGEV GUS plasmid (Higgs and Colbert, 1993). For *atpB* DNA, the 900-bp *atpB Eco*RV–*Pst*I DNA fragment was digested from the DG2 plasmid (Sakamoto *et al*., 1993). For pBS negative control DNA, the 450-bp polylinker fragment was digested from pBSKS II + (Stratagene) with *Pvu*II. To detect hybridized transcripts, membranes were exposed to X-ray film with an intensifying screen at −75 °C. Two independent *in vivo* labeling experiments were conducted for each tested reporter gene. RNA bands were quantified by densitometry of digitized images using NIH Image version 1.62 software. The average background-subtracted GUS/*atpB* values are presented.

To analyze GUS mRNA accumulation in transformants, total RNA was isolated as described above. RNA (10 μg) was gel blotted as previously described (Higgs *et al*., 1999). A ^{32}P-labeled antisense GUS RNA probe was synthesized and hybridized as previously described (Higgs and Colbert, 1993). For the *psbA* probe (loading control), a plasmid containing the chloroplast DNA Eco15 fragment with the *psbA* gene (Harris, 1989) was used as previously described (Higgs *et al*., 1999). RNA bands were visualized by exposure to X-ray film at −75 °C with an intensifying screen. Bands from two independent experiments were quantified by densitometry of digitized images using NIH Image version 1.62 software. The average background-subtracted GUS/*psbA* values are presented.

Fluorometric GUS activity assays were conducted on cells from 10 ml of liquid cultures of homoplasmic transformants as previously described (Sakamoto *et al*., 1993). GUS activities (pmoles 4-methylumbelliferone/min/mg protein) were standardized to total protein, as determined by the Bradford protein assay (Bradford, 1976). Standardized GUS activities were made relative to DG2 positive control activity. At least two independent GUS activity assays were performed for each transformed cell line, and the average percentage of DG2 activity ± standard deviation is presented for each reporter gene.

Sequence data from this article have been deposited with the EMBL/GenBank data libraries under the following accession numbers: *C. reinhardtii* MN *petD* (# AY940198), *C. reinhardtii* NC *petD* (#AY940197), *C. reinhardtii* Q *petD* (# AY940199), *C. smithii petD* (# AY940200), *C. incerta petD* (# AY940201), and *C. moewusii petD* (# AY940202).

Discussion

Chlamydomonas is a diverse genus of single-celled green algae with nearly 500 species that is reported to be polyphyletic (Harris, 1989; Boudreau *et al*., 1994; Boudreau and Turmel, 1996). The high divergence of non-coding *petD* sequences in this report is consistent with this conclusion. These data reveal an overall low level of *petD* 5′ UTR sequence identity common to these species. Despite the overall low identity, element I, II and III-like sequences are present in most of these species. The exception to this is the lack of a clear element I-like sequence in *C. eugametos* and *C. moewusii*. Pairwise alignment between *C. reinhardtii* and *C. incerta petD* revealed additional shared 5′ UTR sequences outside of the three known elements and translation initiation region. These additional sequences were not conserved in *C. eugametos* and *C. moewusii*. Previously reported deletions and substitutions in *C. reinhardtii* showed that many of these additional shared

sequences have no regulatory function (Sakamoto *et al*., 1994a; Higgs *et al*., 1999). For example, the region between elements I and II can be deleted without affecting RNA stability or translation (Sakamoto *et al*., 1994a). Also, the 24 nt region 3′ of element II and the 8 nt region 5′ of element III were shown to have no function (Higgs *et al*., 1999), leaving an untested 74-nt region between nt 235 and 309. For *C. reinhardtii* and *C. incerta*, this 74-nt region includes a high number of substitutions and insertions/deletions between these species, and has only a few, small sections with high sequence identity (Figure 1A). Because this region has not been directly tested, we cannot rule out that an uncharacterized element exists here.

The *petD* mRNA 5′ ends were mapped to assess the relation between 5′ termini and element I-like sequences. Although the five *C. reinhardtii* strains have DNA polymorphisms, the 5′ ends and UTR sequences are identical. In contrast, the 5′ ends in *C. eugametos* and *C. moewusii* are identical to each other but differ considerably from those of *C. reinhardtii* (Figure 3C). These differences included the terminal sequences, length of 5′ UTRs, and accumulation of multiple transcripts with different 5′ ends in *C. eugametos* and *C. moewusii* (Figures 3C and 4). We do not know how many or which of these transcripts are translated.

Chimeric GUS reporter genes and chloroplast transformation were used to test function of these divergent *petD* 5′ UTRs in *C. reinhardtii*. The *C. incerta*-based DiG2 mRNA was both stable and translated. We conclude that the few substitutions (Figure 2B) in the *C. incerta* elements II and III and element spacing differences did not interfere with expression in *C. reinhardtii*. In contrast, the more divergent *C. eugametos petD* 5′ UTR was not competent for expression in *C. reinhardtii*. Despite testing many *C. eugametos*-based GUS mRNAs, none were translated. Energy-minimization RNA structure predictions (Zuker, 1989) of the chimeric *C. eugametos* GUS mRNAs revealed no structures that would occlude the translation start site and block initiation (data not shown). The high sequence identity of these putative elements compared to the low overall 5′ UTR identity suggests that these sequences may be important in *C. eugametos*. To confirm this, the elements would have to be tested directly in *C. eugametos* chloroplast, but transformation is not currently possible in this species.

Nucleotides present in elements from *C. reinhardtii* and *C. incerta* but absent from *C. eugametos* are likely necessary in *C. reinhardtii* chloroplasts. Figure 2C correlates substitutions with RNA structure and functionality of elements I and III-like sequences. For element I, the different 5′ ends and two U to A substitutions may prevent *C. eugametos* RNA accumulation in *C. reinhardtii* chloroplasts. These substitutions would likely destabilize the structure at element I. For element II, no clear structure was predicted (Higgs *et al*., 1999). Position 8 of element II appears less important because of the A to G change in *C. incerta* did not block translation (Figure 2B). For element III, the U to A transversion in *C. incerta* at position 10 is in the loop region. This would not destabilize the structure and is consistent with this stem-loop being important (Higgs *et al*., 1999). Although the *C. eugametos* element III-like sequence has only two substitutions, including the U to A also present in *C. incerta*, there are extensive changes 3′ of the element III. These 3′ changes would destabilize this structure and may contribute to the lack of translation in *C. eugametos* reporter genes.

We tested the stabilizing function of the *C. eugametos* 1–57 terminal sequence with the DeG1-57 reporter gene, and despite significant transcription the mRNA failed to accumulate. Related to this, the *C. reinhardtii* FUD6 spontaneous photosynthetic revertants have functional chimeric *petD* 5′ UTRs (*psbZ*:*petD*) that promote RNA accumulation and translation (Higgs *et al*., 1998). The *psbZ*:*petD* chimeric 5′ UTR relies on *psbZ* 5′ RNA stabilizing sequence(s), instead of element I. From this, we can rule out that the instability of DeG1-57 mRNA was simply due to replacing element I with a heterologous stability sequence. We conclude that the *C. eugametos* 1–57 sequence does not function for RNA stability in *C. reinhardtii* either because it lacks a stability sequence all together or it has one that does not function in *C. reinhardtii*. The *C. reinhardtii* nucleus-encoded Mcd1 interacts with element I to stabilize *petD* mRNA through an unknown mechanism (Drager *et al*., 1999; Murakami *et al*., 2005). It would be interesting to determine if a *Mcd1* homolog exists in *C. eugametos*. To test this, the *C. reinhardtii Mcd1* probe was hybridized to *C. eugametos* genomic DNA on a filter blot and no band was detected in high and low stringency

conditions (data not shown). On the same blot, the expected band in *C. reinhardtii* DNA was detected. This lack of hybridization could be due to a low nucleic acid identity and/or GC content differences, so we cannot rule out the existence of a *C. eugametos Mcd1* homolog.

Spacing between the *petD* elements varies among these species. Most notably *C. eugametos* elements II and III are only 16 nt apart, while in *C. reinhardtii* they are 110 nt apart. This could affect how these elements function by altering binding of *trans*-acting factors. Previously, a non-translated element III substitution was shown to form an interfering RNA secondary structure immediately upstream of this element (Higgs *et al.*, 1999). Subsequent suppressor mutations, in-turn, prevented this alternative inhibitory structure from forming. Together, these suggest that either regulatory sequences or structures immediately upstream of element III might block translation. Perhaps this region needs to be free to allow proper binding of protein(s) and or ribosomes to element III.

Substitutions in *C. reinhardtii petD* and *petA* translation initiation context sequences can affect translation when combined with non-canonical initiation codons (Chen *et al.*, 1995; Esposito *et al.*, 2003). In *petA*, a U to A substitution one nt upstream of the initiation codon had no detectable affect on translation with the canonical AUG (Chen *et al.*, 1995). When combined with the non-canonical AUU initiation codon the same U to A substitution still supported translation, but at a rate that was reduced by about fivefold (Esposito *et al.*, 2003). Based on this, the U to A substitution in *C. eugametos* one nt upstream of the canonical AUG initiation codon by itself is likely to have little effect on translation in *C. reinhardtii*. From this, we infer that the lack of translation from the *C. eugametos*-containing GUS transcripts is not due to the variant initiation context sequence. It is possible, however, that this substitution partly contributes to the lack of translation for *C. eugametos* containing GUS genes.

In summary, these data provide a broader perspective of *petD* 5′ UTR regulatory sequences among this diverse genus. Expression data from the chimeric GUS reporter genes revealed the extent to which these regulatory elements are highly sequence specific and sensitive to nucleotide substitutions and/or spacing changes. We suggest that the corresponding *trans*-acting factors are sensitive to changes in these RNA elements. This is consistent with reports of nucleus-encoded chloroplast RNA stability and translation factors being specific to one chloroplast mRNA (Barkan and Goldschmidt-Clermont, 2000; Zerges, 2000), including the specificity of *Mcd1* to *petD* (Erickson *et al.*, 2005). These data go a step further and show that specificity for chloroplast RNA regulatory elements have species barriers and in some cases prevent orthologous mRNAs from functioning in closely related species. Once *trans* factors and *cis* elements are fully known, such sequence-specific regulation may provide a portable gene-specific regulatory system useful for controlling chloroplast-expressed transgenes in heterologous plant species.

Acknowledgements

We thank Dr. Elizabeth Harris for suggestions of species, strains and nomenclature, to Dr. David Stern for the gift of the *Mcd1* gene probe and critical review of manuscript, and to members of the Higgs laboratory for helpful suggestions. This work was supported by USDA-NRI (#2000-01475) award to DCH.

References

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G. and Struhl, K. 1988. Current Protocols in Molecular Biology. John Wiley and Sons Inc, New York.

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

Barkan, A., Walker, M., Nolasco, M. and Johnson, D. 1994. A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. EMBO J. 13: 3170–3181.

Betts, L. and Spremulli, L.L. 1994. Analysis of the role of the Shine-Dalgarno sequence and mRNA secondary structure on the efficiency of translational initiation in the *Euglena gracilis* chloroplast *atpH* mRNA. J. Biol. Chem. 269: 26456–26463.

Boudreau, E., Otis, C. and Turmel, M. 1994. Conserved gene clusters in the highly rearranged chloroplast genomes of *Chlamydomonas moewusii* and *Chlamydomonas reinhardtii*. Plant Mol. Biol. 24: 585–602.

Boudreau, E. and Turmel, M. 1996. Extensive gene rearrangements in the chloroplast DNAs of *Chlamydomonas* species

featuring multiple dispersed repeats. Mol. Biol. Evol. 13: 233–243.

Bradford, M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. Anal. Biochem. 72: 248–254.

Bruick, R.K. and Mayfield, S.P. 1998. Processing of the *psbA* 5′ untranslated region in *Chlamydomonas reinhardtii* depends upon factors mediating ribosome association. J. Cell Biol. 143: 1145–1153.

Chen, X., Kindle, K.L. and Stern, D.B. 1995. The initiation codon determines the efficiency but not the site of translation initiation in *Chlamydomonas* chloroplasts. Plant Cell 7: 1295–1305.

Church, G.M. and Gilbert, W. 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA 81: 1991–1995.

Drager, R.G., Girard-Bascou, J., Choquet, Y., Kindle, K.L. and Stern, D.B. 1998. *In vivo* evidence for 5′ ϕ 3′ exoribonuclease degradation of an unstable chloroplast mRNA. Plant J. 13: 85–96.

Drager, R.G., Higgs, D.C., Kindle, K.L. and Stern, D.B. 1999. 5′ to 3′ exoribonucleolytic activity is a normal component of chloroplast mRNA decay pathways. Plant J. 19: 521–532.

Drager, R.G., Zeidler, M., Simpson, C.L. and Stern, D.B. 1996. A chloroplast transcript lacking the 3′ inverted repeat is degraded by 3′ ϕ 5′ exoribonuclease activity. RNA 2: 652–663.

Erickson, B., Stern, D.B. and Higgs, D.C. 2005. Microarray analysis confirms the specificity of a *Chlamydomonas reinhardtii* chloroplast RNA stability mutant. Plant Physiol. 137: 534–544.

Erlitzki, R., Long, J.C. and Theil, E.C. 2002. Multiple, conserved iron-response elements in the 3′-untranslated region of transferrin receptor mRNA enhance binding of iron regulatory protein 2. J. Biol. Chem. 277: 42579–42587.

Esposito, D., Fey, J.P., Eberhard, S., Hicks, A.J. and Stern, D.B. 2003. *In vivo* evidence for the prokaryotic model of extended codon–anticodon interaction in translation initiation. EMBO J. 22: 651–656.

Gagne, G. and Guertin, M. 1992. The early genetic response to light in the green unicellular alga *Chlamydomonas eugametos* grown under light/dark cycles involves genes that represent direct responses to light and photosynthesis. Plant Mol. Biol. 18: 429–445.

Harris, E.H. 1989. The Chlamydomonas Sourcebook: A Comprehensive Guide to Biology and Laboratory Use. Academic Press, Inc., San Diego, California.

Herrin, D. and Worley, T. 1990. A rapid procedure for the isolation of chloroplast DNA from *Chlamydomonas* using the TL-100 ultracentrifuge. Plant Mol. Biol. Rep. 8: 292–296.

Higgs, D.C. and Colbert, J.T. 1993. β-glucuronidase gene expression and mRNA stability in oat protoplasts. Plant Cell Rep. 12: 445–452.

Higgs, D.C., Kuras, R., Kindle, K.L., Wollman, F.-A. and Stern, D.B. 1998. Inversions in the *Chlamydomonas* chloroplast genome suppress a *petD* 5′ untranslated region deletion by creating functional chimeric mRNAs. Plant J. 14: 663–671.

Higgs, D.C., Shapiro, R.S., Kindle, K.L. and Stern, D.B. 1999. Small *cis*-acting sequences that specify secondary structures in a chloroplast mRNA are essential for RNA stability and translation. Mol. Cell. Biol. 19: 8479–8491.

Kindle, K.L., Richards, K.L. and Stern, D.B. 1991. Engineering the chloroplast genome: new techniques and capabilities for chloroplast transformation in *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. USA 88: 1721–1725.

Koo, J.S. and Spremulli, L.L. 1994. Analysis of the translational initiation region on the *Euglena gracilis* chloroplast ribulose-bisphosphate carboxylase/oxygenase, (*rbcL*) messenger RNA. J. Biol. Chem. 269: 7494–7500.

Monde, R.A., Schuster, G. and Stern, D.B. 2000. Processing and degradation of chloroplast mRNA. Biochimie 82: 573–582.

Murakami, S., Kuehnle, K. and Stern, D.B. 2005. A spontaneous RNA suppressor of a mutation in the *Chlamydomonas reinhardtii* nuclear *MCD1* gene required for stability of the chloroplast *petD* mRNA. Nuclei Acid Res. 33: 3372–3380.

Needleman, S.B. and Wunsch, C.D. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 48: 443–453.

Newman, T.C., Ohme-Takagi, M., Taylor, C.B. and Green, P.J. 1993. DST sequences, highly conserved among plant *SAUR* genes, target reporter transcripts for rapid decay in tobacco. Plant Cell 5: 701–714.

Nickelsen, J. 1999. Transcripts containing the 5′ untranslated regions of the plastid genes *psbA* and *psbB* from higher plants are unstable in *Chlamydomonas reihardtii* chloroplasts. Mol. Gen. Genet. 262: 768–771.

Nickelsen, J., Fleischmann, M., Boudreau, E., Rahire, M. and Rochaix, J.-D. 1999. Identification of *cis*-acting RNA leader elements required for chloroplast *psbD* gene expression in *Chlamydomonas*. Plant Cell 11: 957–970.

Roffey, R.A., Golbeck, J.H., Hille, C.R. and Sayre, R.T. 1991. Photosynthetic electron transport in genetically altered photosystem II reaction centers of chloroplasts. Proc. Natl. Acad. Sci. 88: 9122–9126.

Sakamoto, W., Chen, X., Kindle, K.L. and Stern, D.B. 1994a. Function of the *Chlamydomonas reinhardtii petD* 5′ untranslated region in regulating the accumulation of subunit IV of the cytochrome *b6/f* complex. Plant J. 6: 503–512.

Sakamoto, W., Kindle, K.L. and Stern, D.B. 1993. *In vivo* analysis of *Chlamydomonas* chloroplast *petD* gene expression using stable transformation of ß-glucuronidase translational fusions. Proc. Natl. Acad. Sci. USA 90: 497–501.

Sakamoto, W., Sturm, N.R., Kindle, K.L. and Stern, D.B. 1994b. *petD* mRNA maturation in *Chlamydomonas reinhardtii* chloroplasts: Role of 5′ endonucleolytic processing. Mol. Cell. Biol. 14: 6180–6186.

Shabalina, S.A., Ogurtsov, A.Y., Rogozin, I.B., Koonin, E.V. and Lipman, D.J. 2004. Comparative analysis of orthologous eukaryotic mRNAs: potential hidden functional signals. Nuclei Acid Res. 32: 1774–1782.

Shapira, M., Lers, A., Heifetz, P.B., Irihimovitz, V., Osmond, C.B., Gillham, N.W. and Boynton, J.E. 1997. Differential regulation of chloroplast gene expression in *Chlamydomonas reinhardtii* during photoacclimation: light stress transiently suppresses synthesis of the Rubisco LSU protein while enhancing synthesis of the PS II D1 protein. Plant Mol. Biol. 33: 1001–1011.

Shaw, G. and Kamen, R. 1986. A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 46: 659–667.

Shepherd, H.S., Boynton, J.E. and Gillham, N.W. 1979. Mutations in nine chloroplast loci of *Chlamydomonas* affecting photosynthetic functions. Proc. Natl. Acad. Sci. USA 76: 1353–1357.

Stern, D.B., Radwanski, E.R. and Kindle, K.L. 1991. A 3′ stem/loop structure of the *Chlamydomonas* chloroplast *atpB*

gene regulates mRNA accumulation *in vivo*. Plant Cell 3: 285–297.

Sturm, N.R., Kuras, R., Büschlen, S., Sakamoto, W., Kindle, K.L., Stern, D.B. and Wollman, F.-A. 1994. The *petD* gene is transcribed by functionally redundant promoters in *Chlamydomonas reinhardtii* chloroplasts. Mol. Cell. Biol. 14: 6171–6179.

Turmel, M., Boulanger, J. and Bergeron, A. 1989. Nucleotide sequence of the chloroplast petD gene of *Chlamydomonas eugametos* Nucleic Acids Res. 17: 3593.

Turmel, M., Gutell, R.R., Mercier, J.-P., Otis, C. and Lemieux, C. 1993. Analysis of the chloroplast large subunit ribosomal RNA gene from 17 *Chlamydomonas* taxa. Three internal transcribed spacers and 12 group I intron insertion sites. J. Mol. Biol. 232: 446–467.

Turmel, M. and Otis, C. 1994. The chloroplast gene cluster containing *psbF*, *psbL*, *petG* and *rps3* is conserved in *Chlamydomonas*. Curr. Genet. 27: 54–61.

Vaistij, F.E., Boudreau, E., Lemaire, S., Goldschmidt-Clermont, M. and Rochaix, J.-D. 2000a. Characterization of Mbb1: a nucleus-encoded tetratricopeptide-like repeat protein required for expression of the chloroplast *psbB*/*psbT*/*psbH* gene cluster in *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. USA 97: 14813–14818.

Vaistij, F.E., Goldschmidt-Clermont, M., Wostrikoff, K. and Rochaix, J.-D. 2000b. Stability determinants in the chloroplast *psbB*/*T*/*H* mRNAs of *Chlamydomonas reinhardtii*. Plant J. 21: 469–482.

Zerges, W. 2000. Translation in chloroplasts. Biochimie 82: 583–601.

Zerges, W., Auchincloss, A.H. and Rochaix, J.D. 2003. Multiple translational control sequences in the 5′ leader of the chloroplast *psbC* mRNA interact with nuclear gene products in *Chlamydomonas reinhardtii*. Genetics 163: 895–904.

Zuker, M. 1989. On finding all suboptimal foldings of an RNA molecule. Science 244: 48–52.