MDA 1, a nucleus-encoded factor involved in the stabilization and processing of the atpA transcript in the chloroplast of Chlamydomonas

Stefania Viola, Marina Cavaiuolo, Dominique Drapier, Stephan Eberhard, Olivier Vallon, Francis-andré Wollman, Yves Choquet

MDA1, a nucleus-encoded factor involved in the stabilization and processing of the *atpA* transcript in the chloroplast of Chlamydomonas

Stefania Viola, Marina Cavaiuolo, Dominique Drapier, Stephan Eberhard, Olivier Vallon, Francis-André Wollman, Yves Choquet*[1][2]

Laboratoire de Physiologie Membranaire et Moléculaire du Chloroplaste -UMR7141, IBPC, CNRS-Sorbonne Université, 13, rue Pierre et Marie Curie, 75005 Paris, France

*Author for correspondence: Yves Choquet (choquet@ibpc.fr)

Short title: Stabilization and processing of the *atpA mRNA*

FUNDING

This work was supported by Unité Mixte de Recherche 7141, CNRS/Sorbonne Université, by the Agence Nationale de la Recherche (ChloroRNP: ANR-13-BSV7-0001-01), and by the "Initiative d'Excellence" program (Grant "DYNAMO," ANR-11-LABX-0011-01).

No conflict of interest declared.

[1] **AUTHOR CONTRIBUTIONS:**

S V, F-A W and Y C designed research plans, S V, M C, D J, D D, S E, OV, and Y C performed experiments, O V analysed the sequencing data, S V, F-A W and Y C analysed the data and wrote the article. All authors read and approved the manuscript.

[2] AUTHOR FOR CORRESPONDENCE: Yves Choquet (choquet@ibpc.fr)
Phone: +33 1 5841 5075
Fax +33 1 5841 50200

ABSTRACT

In *Chlamydomonas reinhardtii*, chloroplast gene expression is tightly regulated post-transcriptionally by gene-specific *trans*-acting protein factors. Here we report the molecular identification of an OctotricoPeptide Repeat (OPR) protein, MDA1, which governs the maturation and accumulation of the *atpA* transcript, encoding subunit α of the chloroplast ATP synthase. As does TDA1, another OPR protein required for the translation of the *atpA* mRNA, MDA1 targets the *atpA* 5'UnTranslated Region (UTR). Unexpectedly, it binds within a region of ~100 nt in the middle of the *atpA* 5'UTR, at variance with the stabilisation factors characterised so far, which bind to the 5'end of their target mRNA to protect it from 5' → 3' exonucleases. It binds the same region as TDA1, with which it forms a high molecular weight complex that also comprises the *atpA* mRNA. This complex dissociates upon translation, promoting degradation of the *atpA* mRNA. We suggest that *atpA* transcripts, once translated, enter the degradation pathway because they cannot reassemble with MDA1 and TDA1, which preferentially bind to *de novo* transcribed mRNAs.

Significance statement： Two OctotricoPeptide Repeat (OPR) proteins, MDA1 and TDA1, respectively required for the .accumulation and for the translation of the *atpA* mRNA interact with a region in the middle of the *atpA* 5'UTR, instead of the usual interaction with its 5' end, to form a ternary complex. This complex dissociates upon translation, which promotes the degradation of the *atpA* mRNA.

Keywords: Chloroplast gene expression, RNA stability, OctotricoPeptide Repeat Protein, Nucleo-Chloroplastic interactions, *Chlamydomonas reinhardtii*

INTRODUCTION

Oxygenic photosynthesis evolved about 2.7 billion years ago (Holland, 2006), when cyanobacteria started to convert sunlight energy into chemical bonds and reducing power, then used to fix carbon into organic molecules, while oxidising water to molecular dioxygen. In photosynthetic eukaryotes, photosynthesis takes place in specialized organelles, the chloroplasts, which derive from an ancestral cyanobacterium engulfed by a heterotrophic eukaryote (Howe et al, 2003; Douglas and Raven, 2003; Ponce-Toledo et al., 2017). Following this initial event, most genes of the endosymbiont were lost or transferred to the nuclear genome of the host cell (Kleine et al, 2009). As a result, chloroplast protein complexes are mosaics of plastid- and nucleus-encoded subunits, and their biogenesis and function require an efficient bi-directional coordination of gene expression between the two genetic compartments (reviewed in Pesaresi et al, 2007; Woodson and Chory, 2008).

The unicellular green alga *Chlamydomonas reinhardtii* has been extensively used as a model organism to study the regulation of chloroplast gene expression. Remarkably, the genetic analysis of nuclear mutants affected in chloroplast gene expression has uncovered an extensive post-transcriptional regulation by nucleus-encoded proteins termed Organelle *Trans*-Acting Factors (OTAFs). These factors control the stability and maturation (M factors) and/or translation activation (T factors) of specific chloroplast mRNA targets and typically act on their 5' untranslated (Barneche et al., 2006; Dauvillee et al., 2003; Drager et al., 1998; Eberhard et al., 2011; Loiselay et al., 2008; Nickelsen et al., 1999; Rahire et al., 2012; Stampacchia et al., 1997; Vaistij et al., 2000b; Wostrikoff et al., 2001; Zerges and Rochaix, 1994; reviewed in Stern et al, 2010; Choquet and Wollman, 2002). In Chlamydomonas, the M factors characterised so far bind the very 5' end of their target mRNA to prevent its degradation by 5' → 3'exonucleases (Drager et al., 1998; Loiselay et al, 2008; Loizeau et al, 2014; Wang et al., 2015; Pfalz et a, 2009; Cavaiuolo et al, 2017). . They often generate "footprints", *i.e.* a cluster with a sharp 5'end of small RNAs (sRNAs) protected from ribonucleolytic degradation by their interaction with their cognate M factors (Ruwe and Schmitz-Linneweber, 2012; Zhelyazkova et al, 2012; Cavaiuolo et al 2017; Loizeau et al, 2014). Most of these footprints disappear in mutant strains defective for the corresponding M factor (Loizeau et al 2014; Ruwe et al 2016; Cavaiuolo et al, 2017). By contrast, how T factors activate translation remains poorly understood. They may recruit ribosomes at the appropriate translation start sites, possibly compensating the lack of *bona fide* Shine-Dalgarno sequences in most chloroplast genes (Schwarz et al, 2007; Fargo et al., 1998; Hirose and Sugiura, 2004; Lim et al., 2014; but see Scharff et al., 2017). RBP40, the

psbD translation activator, facilitates the recruitment of ribosomes on the *psbD* start codon by melting a secondary structure sequestering the initiation codon (Schwarz et al., 2007), as does TAB1, the *psaB* translation activator in Chlamydomonas (Stampacchia et al., 1997) or the PentatricoPeptide Repeat (PPR) PPR10 factor controlling *atpH* mRNA translation and stability in maize (Pfalz et al, 2009; Zhelyazkova et al, 2012).

We previously showed that the expression of the *atpA* chloroplast gene encoding the α subunit of chloroplast ATP synthase (CF1α) requires the OctotricoPeptide Repeat (OPR) protein TDA1 in *C. reinhardtii,* which activates its translation but does not control its stability (Drapier et al, 1992; Eberhard et al, 2011). TDA1 belongs to the OPR protein family, defined by tandem repeats of a degenerate motif of 38 residues. OPR proteins, abundant in green algae, bind mRNAs in a sequence-specific manner (Eberhard et al, 2011; Rahire et al, 2012; Boulouis et al, 2015), as do PPR proteins, another helical repeat protein family prevalent in land plants. Both families are almost exclusively dedicated to the post-transcriptional control of gene expression in chloroplasts and mitochondria (Pfalz et al, 2009; Hammani et al, 2016; Zhou et al, 2017; Loiselay, 2008).

Here we characterise another OPR protein, MDA1, required for the stable accumulation and 5'end maturation of the *atpA* transcript in Chlamydomonas. Together, MDA1 and TDA1 constitute one of the few couples of M and T factors acting on the same mRNA identified so far and, therefore, represent an ideal tool to study the molecular basis of the regulation of chloroplast gene expression.

RESULTS

The soluble OPR protein MDA1 is required for the stable accumulation of the *atpA* mRNA.

The *C. reinhardtii* mutant *mda1-1* was isolated as a non-phototrophic strain in an insertional mutant library (Dent, 2005), and its characterization pointed to a defect in chloroplast ATP synthase. It fails to accumulate the chloroplast *atpA* mRNA (Fig. 1A), at variance with *tda1* mutants, where the absence of the ATP synthase subunit α results from an impaired translation of the *atpA* mRNA (Drapier et al, 1992; Eberhard et al, 2011). Therefore, while *TDA1* encodes an *atpA* mRNA translation factor, the *MDA1* gene might encode a stability/maturation factor. Genetic analysis indicated that the *mda1-1* mutation was nuclear but not linked to the *ble* resistance marker used for insertional mutagenesis. Therefore, to identify the mutation responsible for the phenotype, the strain was crossed to S1D2, a wild isolate showing a high level of sequence polymorphism with the reference genome (Gross et al., 1988; Rymarquis et al., 2005). Upon genome sequencing of pooled wild-type and mutant progeny and mapping of the SNP frequency to the chromosomes, a 7 bp deletion in the third exon of the nuclear gene Cre16.g680850 (Fig. 1B) correlated with the defect in *atpA* mRNA accumulation. The Cre16.g680850 gene model is 4972 nucleotides long, encompasses 5 exons and 4 introns and encodes a predicted protein of 1207 residues (~120 kDa). The 7 bp deletion leads to a premature stop codon and translation abortion after 357 residues. Analysis with the FT-Rep software (Rahire et al., 2012) of the predicted gene product identified it as an OPR protein with 13 OPR repeats (Suppl. Fig. S1), while a chloroplast targeting peptide was predicted at the N-terminus by the PredAlgo software (Tardif et al, 2012). As other *trans*-acting factors (Eberhard et al., 2011; Loiselay et al., 2008; Raynaud et al., 2007), MDA1 has orthologs in closely related Chlamydomonadaceae algae, but evolved rapidly with multiple species-specific insertions and its phylogenetic signature is rapidly lost in more distant organisms (Suppl. Fig. S2).

To confirm that this deletion was responsible for the mutant phenotype, we complemented the *mda1-1* mutant with a BAC clone (clone 16F5 from the BAC library distributed by the Clemson Genomics Institute) spanning Cre16.g680850 and two other genes (Suppl. Fig. S3A). We recovered phototrophic transformants, which accumulated wild-type levels of CF1α, indicative of a restored accumulation of the *atpA* transcript (Suppl. Fig. S3B). Genotyping with primers specific for the wild-type or mutated *MDA1* alleles showed that both gene versions were present in transformants (Suppl. Fig. S3C).

Demonstration that Cre16.g680850 actually corresponds to *MDA1* was obtained by complementation with a synthetic version of the gene (*MDA1-Fl*, see Fig. 1B), which contained the full-length CDS and the second intron, placed under the control of the AR promoter (Schroda et al, 2002) and of the *rpl17* 5'UTR (Takahashi et al., 2016). To detect the MDA1 protein by immunoblotting and perform pull-down experiments, we introduced a triple Flag tag (Fl) after residue 162, in a region that does not contain OPR repeats and where insertions are found in some MDA1 orthologs (Suppl. Fig. S2).

Seven phototrophic clones were analysed by immunoblot for the presence of MDA1-Fl and CF1α and by RNA blot for the recovery of the *atpA* mRNA (Fig. 1D). The MDA1 protein accumulated to contrasting levels in these strains, probably because of the random insertion of the transgene in the nuclear genome. None of the complemented strains however accumulated wild-type levels of the *atpA* mRNA, as the highest level of *atpA* mRNA accumulation, in clone #7, hereafter named *MDA1-Fl* and used for subsequent experiments, was only ~35% of the wild-type content. We suspect that the abundance of the tagged *MDA1* expressed from the synthetic gene remained below that of the endogenous factor. As expected from the requirement of MDA1 for *atpA* mRNA accumulation and expression, we observed some correlation between the levels of *atpA* mRNA and those of MDA1 (Fig. 2B), the clones accumulating less MDA1 (#1, #2 and #4) also accumulating lower levels of *atpA* mRNA.

In parallel, we also constructed tagged versions of the TDA1 protein by inserting either a triple Flag tag after codon 1115 or a triple HA tag at position 210 (Fig. 1C; see Materials and Methods for details). We first used these tagged versions of the proteins to assess their intracellular localisation. To this end, we generated by crosses a double mutant strain, *mt*, lacking both MDA1 and TDA1. We then recovered strains expressing either MDA1-Fl in the absence of TDA1 (*mt*M), or TDA1-HA in the absence of MDA1 (*mt*T) by transformation of this double mutant with the appropriate plasmid and selection of transformants on paromomycin. Lastly, we transformed the *mt*M strain with the tagged version of TDA1 and selected *mt*MT double complemented transformants for the restoration of phototrophy.

After ultracentrifugation of whole cell extract from strain *mt*MT, both proteins were mostly found in the soluble fractions (Fig. 1E). We note, however a limited but significant amount of TDA1 in the membrane, cushion and pellet fractions, suggesting that this factor either belong to very high molecular weight complexes or tends to form aggregates upon chloroplast disruption.

MDA1 and TDA1 target the same region of the *atpA* 5' UTR.

M and T factors characterized so far act on the 5'UTR of their target transcript , with M factors interacting more specifically with the very 5'end of the mature transcript (reviewed in Choquet and Wollman, 2002; Stern et al, 2010; Barkan, 2011; Germain et al., 2013). To determine if MDA1 also acts on the *atpA* 5'UTR, we transformed the *mda1-1* and *MDA1-Fl* strains with the chimeric construct *5'dAf* (Drapier et al, 2007; Boulouis et al, 2011), where the 5'UTR of the endogenous *petA* gene is replaced by that of *atpA* (Fig. 2A). RNA blots showed that the 5'*atpA-petA* chimeric transcript behaves as the endogenous *atpA* mRNA, failing to accumulate in the *mda1-1* nuclear context while being stabilized in the complemented strain (Fig. 2B). The *atpA* 5'UTR is thus sufficient to confer an MDA1-dependent stability to a downstream coding sequence. Using the same chimera, we previously showed that the *atpA* 5'UTR was also the target of the translation activator TDA1 (Eberhard et al, 2011).

In several instances, *trans*-acting factors have been shown to physically interact with their target transcript (Dauvillée et al, 2003; Pfalz et al, 2009; Rahire et al, 2012). To determine whether MDA1 and TDA1 also interact with the *atpA* 5'UTR, we immuno-precipitated Flag-tagged versions of the two proteins (in strains *MDA1-Fl* and *TDA1-Fl*, respectively) and performed a similar experiment on a wild-type cellular extract, hence devoid of Flag-tagged proteins, as a negative control (Fig. 2C). Using a probe covering most of the *atpA* 5'UTR or an unrelated probe derived from the *petD* chloroplast gene as a control, we analysed the RNA extracted from the immuno-precipitated pellets (P) by dot blot. Samples from both *MDA1-Fl* and *TDA1-Fl* strains, but not from the wild-type strain, were specifically enriched in *atpA* 5'UTR, but not in *petD* transcript, demonstrating the specific binding of the two proteins to the *atpA* 5'UTR *in vivo*. To further define the region of interaction, we repeated this experiment using a series of partially overlapping probes about 100 nucleotides long (Fig. 2D), that together span the whole *atpA* 5'UTR from the transcription start site (TSS) up to the beginning of the coding sequence. A probe covering the *atpA* 3'UTR was used as a control. Unexpectedly, for both MDA1-Fl and TDA1-Fl, the strongest enrichment was detected with the probe *5'atpA-3* (Fig. 2E), that covers nucleotides 135-249 of the 5'UTR. Thus, MDA1 and TDA1 interact with the same region in the middle of the *atpA* 5'UTR.

The MDA1-TDA1-*atpA* ternary complex.

The latter observation prompted us to assess their possible association in a same oligomeric complex. We thus performed RNA and protein co-immuno-precipitation experiments in soluble cellular extract from strain *mt*MT (Fig. 3). This input cellular extract was either untreated or RNase-treated, to investigate the possible dependence of a putative interaction between MDA1

and TDA1 on the presence of an intact *atpA* 5'UTR. Treated and untreated extracts were immuno-precipitated using either α-Flag or α-HA magnetic beads and the recovered RNA analysed by dot blot. Hybridization with the *5'atpA-3* probe confirmed the absence of RNA in RNase-treated pulled-down (P) fractions (Fig. 3A). The RNase-treated total extract (E) fraction saved contains partially digested *atpA* mRNAs that could be detected by dot blot using the *5'atpA-3* probe. Immunoblot analysis of total extract (E), supernatant (S) and pulled-down (P) fractions showed that a large fraction of MDA1 co-precipitated with TDA1 (Fig. 3B). However, only a minor -but significant- fraction of TDA1 was pulled down with MDA1. This asymmetrical behaviour suggests that the accumulation level of TDA1 might be higher than that of MDA1 in the double-complemented strain *mt*MT. An RNase treatment prevented most if not all of these co-IP signals. Therefore, we conclude that MDA1 and TDA1 can assemble in a ternary complex on the *atpA* 5'UTR. Moreover, the binding of MDA1 to the *atpA* 5'UTR does not depend on the presence of TDA1, as shown by RIP experiments performed on the *mt*M strain, the double *mda1-1*/*tda1–1* mutant complemented with the tagged version of MDA1 only, that lacks TDA1 (Suppl. Fig. S4). Whether TDA1 interacts with its target in absence of MDA1 could not be assessed, since the *atpA* mRNA does not accumulate in such a situation.

We then performed size exclusion chromatography using untreated and RNase-treated *mt*MT soluble extracts as well as untreated extracts from the *mt*M and *mt*T strains (Fig. 4). In untreated *mt*MT, MDA1 was mostly detected in fractions 3 to 8, with a peak in fractions 5 to 7 corresponding to a complex with an apparent molecular mass of ~550-670 kDa. The distribution of TDA1 partially overlapped that of MDA1, but most of it was found in a complex of lower molecular mass (~450-600 kDa, fractions 6-9). Upon RNase treatment the distribution of MDA1 shifted to lower molecular weight complexes (fractions 6-10), with another population now found in higher molecular weight (fraction 2), possibly corresponding to protein aggregates. The minor fraction of TDA1 found in fractions 3-5 was lost upon RNAse treatment, releasing an enrichment of TDA1 in fractions 8-9, in the ~450-520 kDa region. In the *mt*M sample the distribution of MDA1 was bimodal with a broad repartition in fractions 3-4 similar to that in the untreated sample and a sharp peak in fractions 6-7. In the *mt*T strain, where neither MDA1 nor the *atpA* mRNA are present, TDA1 was mainly found in the 520 kDa region (fractions 7-8). As a control, the distribution of the RuBisCO holoenzyme (~520 kDa) remained the same in all samples (lower panel). Taken together, the co-immuno-precipitation and size-exclusion chromatography experiments indicate that MDA1 and TDA1 interact with the same region in the middle of the *atpA* 5'UTR to form high molecular weight (HMW) complexes of

about 670 kDa and above. Owing to the uneven expression of TDA1 and MDA1 in *mt*MT, which results from the random insertion of their genes in the double *mt* mutant, the excess of TDA1 remains MDA1-free in a lower molecular weight complex. Note that in these experiments no special care was taken to preserve the integrity of RNAs: RNAs retained in ternary complexes were probably restricted to fragments closely surrounding the MDA1 or TDA1 binding sites, as shown in RIP experiments by the weak signal against probes *5'atpA-2* and *-4* despite a strong interaction with probe *5'atpA-3*. RNA should thus account for only a minor increase in molecular mass.

MDA1, but not TDA1, is required for the accumulation of the 5'end processed *atpA*-mRNA

In *C. reinhardtii,* most chloroplast mRNAs are processed after transcription. The *atpA* gene is transcribed as a tetracistronic primary transcript with a tri-phosphorylated 5'end (Cavaiuolo et al, 2017). It is then processed to position +36 relative to the TSS to a mono-phosphorylated monocistronic mRNA (Drapier et al, 1998; Cavaiuolo et al, 2017). Because MDA1 interacts with a region located in the middle of the *atpA* 5'UTR, we investigated whether it had any role in this 5' processing event. We used circular RT-PCR (cRT-PCR) on total RNA samples from wild-type and *mda1-1* strains, as well as the complemented *MDA1-Fl* strain (Fig. 5). To assess if 5'end maturation also depends on the presence of the translation activator TDA1, we also included the *tda1Δ* mutant (Eberhard et al, 2011).

Prior to RNA self-ligation, half of the RNA was treated with RNA 5'polyphosphatase (RPP) which converts tri-phosphorylated 5'ends into mono-phosphorylated forms. Comparison between RPP and mock samples allows distinguishing primary transcripts from processed mRNA, as only RNAs with mono-phosphorylated 5'ends can be self-ligated. Circular RNA was retro-transcribed with an *atpA* 5'UTR-specific primer (*A-5'* 2 RV) and the single-stranded cDNA was used as a template to amplify the junction between transcript 5' and 3'ends using outward directed primers to the *atpA* 3' (*A-3'* FW) and 5' (*A-5'* 1 RV) UTRs. We observed two major products: A1 with the expected size of 240 bp and A2 with a size of ~450 bp (Fig. 5A and B). The A1 amplicon was detected in all, but not in the mock-treated *mda1-1,* samples. The A2 amplicon, readily detectable in the RPP-treated sample from strain *mda1-1* and in both RPP- and mock-treated samples from strain *MDA1-Fl* was hardly detectable in wild-type and *tda1*Δ samples. Gel-purified amplicons A1 and A2 were sequenced with the primers used for the amplification to map the major 5' and 3'end of the *atpA* transcript (Suppl. Table ST2). In the wild-type and *tda1*Δ samples, either mock- or RPP-treated, only the mature 5'end could be

detected, while the 3'end was found at the position (+162), identified in (Drapier et al, 1998) for the mono-cistronic *atpA* mRNA. In contrast, the *mda1-1* strain showed only the primary 5'end, as suggested by the absence of A1 amplicon in the mock-treated sample. In the complemented *MDA1-Fl* strain the correctly processed 5'end was found in the mock sample but the primary 5'end predominated in the RPP-treated sample, in both the A2 and A1 bands. Sequencing the PCR products from the 3'ends showed that A2 bands extend variously into *psbI*. We also analysed the 5'-3' junctions of the *atpA-psbI* and *atpA-psbI-cemA* di- and tri-cistronic mRNAs and found the same dependency on the presence of MDA1 for 5'end maturation (Suppl. Fig. S5 and Table ST3). Altogether, these results demonstrate that the *atpA-psbI-cemA-atpH* gene cluster is transcribed in the *mda1-1* mutant, the lack of *atpA* mRNA in this strain being due to a stability defect: *atpA* transcripts are still produced even if most of them are rapidly degraded.

Many M factors generate "footprints" on the target mRNA. A sRNA of 34-nt mapping the 5'end of the processed *atpA* transcript has been identified in the sRNA sequencing datasets of the wild-type strain (Cavaiuolo et al, 2017). We similarly performed sRNA-seq on the *mda1-1* and *tda1Δ* mutant strains (Fig. 5C). The *atpA* 5'end-associated sRNA was missing in the *mda1-1* mutant and hardly detectable in strain *tda1*Δ. However, since MDA1 and TDA1 bind at a distance from this 5'end, we cannot exclude that this footprint is generated by another protein, whose stable binding to the *atpA* 5'end would require the presence of MDA1.

Thus, MDA1, even though not binding the 5'end of the *atpA* mRNA, is required for the stabilization of the *atpA* mRNA and for the maturation of its 5'end.

Once translated, the *atpA* mRNA is targeted for degradation

To assess if MDA1 and TDA1 remain associated with the *atpA* mRNA once translation has started, we analysed their localisation in polysomes gradients (Fig. 6A). None of the two proteins were found in polysomal fractions 1-5, where *atpA* mRNA are steadily detected. Even if we cannot exclude that the affinity of MDA1 or TDA1 for their mRNA target is too low to resist our polysome purification procedure, this observation suggests a release of both factors from their 5'UTR target once the transcript associates with the translation machinery. A similar situation has been described for other *trans*-acting factors in Chlamydomonas chloroplasts (Boudreau et al, 2000; Schwarz et al, 2007; Dauvillée et al, 2003). It should be noted that no ribosomal components could be immuno-precipitated with either MDA1 or TDA1 (Fig. 3B), even in presence of Chloramphenicol, which stalls the ribosomes on transcripts, thereby stabilising polysomes (Fig. 6B and C).

We looked for a possible post-translational degradation of the *atpA* mRNA, that could result from the absence of rebinding of MDA1, or prevent its re-binding. Indeed, the MDA1 and TDA1 factors are stable proteins with a half-life longer than 6 hr, as shown by immuno-chase experiments (Suppl. Fig. S6). Thus, they should remain potentially available for rebinding the *atpA* transcript, once translated. To study the stability of the *atpA* mRNA upon translation, we inhibited chloroplast transcription with rifampicin, in absence or presence of lincomycin, an inhibitor of chloroplast translation initiation (Fig. 6D). A control treatment with lincomycin alone did not show any effect on *atpA* mRNA accumulation, as previously reported (Boulouis et al, 2015; Cavaiuolo et al, 2017) and in agreement with the phenotype of the *tda1*Δ strain that does not show a decreased accumulation of the *atpA* transcript despite its impaired translation. Upon transcription inhibition, the half-life of the *atpA* mRNA was of about 3 hours, but it increased considerably (over 6 hours) when chloroplast translation was also inhibited, which shows that translation promotes its degradation.

DISCUSSION

MDA1 is an OPR protein required for the accumulation of the *atpA* mRNA

In the chloroplast of *C. reinhardtii*, the post-transcriptional control of *atpA* gene expression was previously shown to rely on the *trans*-acting protein factor TDA1, a nucleus-encoded OPR protein which targets the *atpA* 5'UTR to activate its translation (Drapier et al, 1992; Eberhard et al, 2011). Here, we identified another *trans*-acting factor, MDA1, specifically controlling the maturation and stability of the *atpA* mRNA. MDA1 is also an OPR protein, which further illustrates how, in *C. reinhardtii*, this family of nucleus-encoded helical repeat proteins is extensively involved in the post-transcriptional control of chloroplast gene expression. By complementing an *mda1* mutant with a tagged version of the protein, we retrieved a series of independent phototrophic clones expressing various levels of MDA1. As observed with other M factor, in mutants complemented with a tagged version of the mutated gene none of the complemented strains restored wild-type levels of the *atpA* mRNAs. The expression level of transgenes in Chlamydomonas rarely reaches that of resident genes, as observed here with our MDA1-transformants. We cannot exclude that the tag, although not abolishing the function of the protein, decreases its affinity for the *atpA* mRNA.

MDA1 and TDA1 physically interact with the *atpA* 5'UTR *in vivo*.

MDA1 and TDA1 genetically target the *atpA* 5'UTR, as shown using chimeric constructs (this work and Eberhard et al, 2011, respectively). Our RIP experiments provided further evidence for a physical interaction of MDA1 and TDA1 with the *atpA* 5'UTR *in vivo*. Whether these interactions are direct or involve additional and still unidentified protein factors remains to be determined, but OPR proteins typically interact directly with RNA (Rahire et al, 2012; Kleinknecht et al., 2014). Unexpectedly, a fine mapping of these interactions showed that both MDA1 and TDA1 target the middle (+136 to +250 from the TSS) of the 425 nucleotides long *atpA* 5'UTR. This situation is unprecedented and implies that MDA1 does not bind the very 5'end of the *atpA* mRNA, but stabilises it from within. TDA1 also binds the *atpA* 5'UTR at least 170 nucleotides away from the start codon. The molecular mechanism by which it activates translation remains to be studied and could involve the remodelling of *cis*-acting structural elements in the *atpA* 5'UTR, an interaction with additional *trans*-acting factors or ribosome recruitment.

Neither MDA1 nor TDA1, generate footprints in the middle of the *atpA* 5'UTR, even though their binding to the target RNA is stable enough for the interaction to resist our

experimental conditions for RIP analysis. Remarkably, the sRNA-Seq analysis reveals a 34 nt long sRNA, which maps at the very 5'end of the mature *atpA* transcript and is no longer observed in absence of MDA1. It is unlikely to represent the footprint of MDA1, which binds in the middle of the *atpA* 5'UTR. We cannot exclude, however, that the most upstream OPR repeats of MDA1 can bind the 5'end of the transcript and would generate the footprint in our wild-type strain. In the *MDA1-Fl* complemented strain, this interaction may be weakened by the insertion of the triple Flag tag, preventing us to observe the footprint in our RIP experiments. But this would not explain the absence of the footprint in the TDA1-Fl complemented strain, which expresses a wild-type MDA1 protein. We rather consider that this footprint results from a conformational change of the *atpA* mRNA upon MDA1 binding to its distant target or is the footprint of another "canonical", yet unknown, M factor (MDA2?) bound to the 5'end of the transcript, whose association to the *atpA* mRNA would depend on the presence of MDA1. While the binding of the hypothetical MDA2 factor would be cooperative with that of MDA1, it should be independent of the presence of TDA1, the *atpA* transcript being stable in *tda1* mutants. Strikingly, however, this sRNAs is also almost absent in the absence of TDA1 (see below).

MDA1 and TDA1 form high molecular weight complexes

Many *trans*-acting factors for chloroplast gene expression form high molecular weight mRNA-protein complexes (HMW-RNP: Auchincloss et al., 2002; Boudreau et al, 2000; Boulouis et al, 2011; Dauvillée et al, 2003; Johnson et al 2010; Perron et al, 2004; Schwarz et al, 2007; Stoppel et al, 2011; Vaistij et al., 2000). MDA1 and TDA1 also both belong to HMW-RNP. Here, we observed that a noticeable proportion of MDA1 co-precipitated with TDA1, while only a limited fraction of TDA1 co-precipitated with MDA1. However this experiment was performed on an MDA1/TDA1 double complemented strain that does not express a wild-type amount of the two proteins. Thus, owing to the co-IP profiles, this strain is likely to express TDA1 in excess of MDA1. This is consistent with the size-exclusion chromatography experiments where MDA1 was mostly found, when the *atpA* mRNA was present, in a complex of ~550-670 kDa, where a significant but minor fraction of TDA1 is also detected. Indeed, the bulk of TDA1 belonged to a lower MW complex in the 450-500 kDa region, where MDA1 was almost absent. The fractions that contain both proteins likely correspond to a ternary complex: i) they are the most affected by an RNase treatment ii) the presence of MDA1 in the 670 kDa region is decreased in the absence of TDA1 and iii) TDA1 is not found any more in these fractions when MDA1, and therefore the *atpA* mRNA, are not present. The apparent size of this

ternary complex, though, does not allow estimating the factor stoichiometry, mostly because we cannot exclude the presence of additional components.

Maturation of the *atpA* mRNA requires MDA1

In *C. reinhardtii*, *atpA* is the first cistron of a polycistronic unit also comprising *psbI, cemA* and *atpH*. In chloroplasts, polycistronic transcripts are often processed to monocistronic mRNAs by intercistronic cleavages followed by 5' and 3' exonucleolytic trimming. Stem loop structures in 3'UTRs protect the mRNA from 3'→5' exonucleases (Stern and Kindle, 1993; Goldschmidt-Clermont et al, 2008; Blowers et al, 1993), while binding of protein factors defines the position of the mature 5'ends by hampering 5'→3' exonuclease activity (Drager et al, 1998; reviewed in Barkan, 2011). Primary mRNAs possess tri-phosphorylated 5' ends, while their processed forms have mono-phosphorylated 5' ends. In agreement, the *atpA* gene is transcribed as a tri-phosphorylated tetracistronic mRNA and processed to a monocistronic mRNA with a mature mono-phosphorylated 5'end at position +36 with respect to the TSS (Cavaiuolo et al, 2017; Drapier et al, 1998).

Although MDA1 does not bind the *atpA* mRNA 5'end, it is required for its stabilisation and processing to its mature 5'end, as shown by circular RT-PCR experiments. Compared to primer extension analysis, cRT-PCR allows selecting a specific 5'-3' junction and it also allows assessing specifically the 5' maturation of each of the processing intermediates of the primary transcript. In the *mda1-1* mutant, the only 5'end detected for the mono-, di- and tri-cistronic transcripts is the tri-phosphorylated TSS. This TSS was not detected in the wild-type and *tda1Δ* strains, where only the mature monophosphorylated 5'end was observed. Therefore, while MDA1 is required for the proper processing and stable accumulation of the *atpA* transcript, TDA1 is not, as the mature *atpA* mRNA accumulates to wild-type levels in the *tda1Δ* mutant. The MDA1-complemented strain shows both the primary and processed ends, respectively in the RPP and mock-treated samples. In that strain, *atpA* transcript accumulates to only 35% of the wild-type levels and hence at least 65% of the newly transcribed *atpA* mRNA would fail to bind MDA1 and undergo degradation. Interestingly, the tri-phosphorylated *atpA* transcripts detected in the *mda1-1* and *MDA1-Fl* strains only (A2 amplicon) show heterogeneous 3'ends, extending into the downstream genes, which suggests that this technique captures *atpA* transcripts very early after transcription, which are too short-lived to be properly processed, before their 5' and 3' maturation is completed and before binding of MDA1. These processing intermediates should also exist in the wild-type and *tda1*Δ strains, but would only represent a tiny and undetectable fraction of the total transcripts. Thus 3' processing and 5'end maturation

events are not mechanistically coupled or even temporally ordered, but rather happen stochastically until the final mature mono-cistronic form is produced.

We conclude that, although not directly binding the mature *atpA* 5'end, MDA1 appears pivotal in its formation, by recruiting or assisting additional factor(s) that would protect the mature 5'end.

The subset of translated *atpA* mRNAs is subsequently targeted for degradation

The interaction of MDA1 and TDA1 with the *atpA* transcript appears stable enough to be detected in co-IP and RIP experiments. However, their interaction with the *atpA* transcript is disrupted upon translation initiation: neither MDA1 nor TDA1 were found in polysomal fractions, and co-IPs failed to show a stable interaction with ribosomal components. Similarly, the NAC2 (MBD1) factor, required for *psbD* mRNA stabilisation (Nickelsen et al., 1994; Boudreau et al, 2000), recruits the *psbD* translational activator, RBP40, with which it forms a HMW-complex. This complex is not associated with polysomes, suggesting that the proteins dissociate from their target mRNA once it is engaged in translation (Schwarz et al, 2007). Also the translation activators of the *psbC* and *psaB* genes, respectively TBC2 and TAB2, were not found in polysomal fractions (Auchincloss et al, 2002; Dauvillée et al, 2003).

The dissociation of both TDA1 and MDA1 from the *atpA* 5'UTR upon translation initiation should deliver an unprotected transcript that would be prone to degradation. However, mutations in T factors, here the deletion of *TDA1* (Eberhard et al, 2011), do not affect mRNA accumulation (Drapier et al, 1992; Stampacchia et al, 1997; Schwarz et al, 2007; Dauvillee et al, 2003; Rahire et al, 2012; Girard-Bascou et al., 1992). Similarly, inhibition of translation initiation by lincomycin does not significantly modify the steady-state levels of most chloroplast mRNAs, including *atpA* (Boulouis et al, 2015; Cavaiuolo et al, 2017). Either translation does not interfere with mRNA degradation pathways or, if it does, on-going chloroplast transcription compensates for degradation. Here we addressed this question by examining the effect of translation inhibition by lincomycin on mRNA decay upon rifampicin treatment. We showed that the *atpA* mRNA is degraded faster when translated: in absence of *de novo* transcription, its half-life of ~3 hr approximately doubles when translation is also prevented. This is probably because MDA1 remains bound to its target, thus protecting it from degradation. In physiological conditions, transcription compensates for transcript degradation and the size of the pool of *atpA* transcript is determined by the amount of MDA1. Translation-induced degradation of the *atpA* mRNA is consistent with its over-accumulation in *tda1* mutants (Drapier et al, 1992) and with its high instability in cells grown in autotrophic conditions under

high light, *i.e.* when an active translation of chloroplast genes is required (Eberhard et al, 2002). One may argue that in the presence of rifampicin alone, MDA1 and TDA1 dissociated from the *atpA* mRNA upon translation remain available - both being long-lived enough (Suppl. Fig. S6) - to rebind and stabilise *atpA* transcripts once released from ribosomes. However we showed in this study that such is not the case. They do not rebind at a rate high enough to prevent the post-translational degradation of the *atpA* transcript. This observation can be interpreted in two ways (Fig. 7). Translation of the *atpA* mRNA, prone to translational pauses (Stollar et al., 1994), may last long enough to allow some co-translational degradation of the *atpA* 5'end by the 5' → 3'exonucleases particularly active in the chloroplast, thereby preventing the transcript to be rebound by an hypothetical 5' MDA2 factor or antisense sRNA. Assuming that its binding is cooperative with that of MDA1, as suggested by the loss of the *atpA* mRNA in *mda1* mutant, the two factors will be directed to *de novo* transcribed mRNA with intact 5'ends, rather than to already translated and 5' trimmed *atpA* transcripts, explaining why newly transcribed *atpA* mRNAs are favoured over those that have already been translated.

The increased stability of the *atpA* mRNA upon translation inhibition may also explain the puzzling observation of the loss of the 5' sRNA footprint in strain *tda1*Δ, despite a preserved or even increased stability of the *atpA* transcript. Accumulation levels of footprint sRNA probably result from a complex interplay between the abundance of the mRNA complexed to its cognate M factor, the rate of endonucleolytic cleavages followed by 3' → 5' trimming up to the region protected by the M factor and the rate of dissociation of the M factor and its footprint. These levels appear extremely variable, even when comparing genes with similar accumulation: the footprint at the 5'end of the *rbcL* mRNA is almost below detection threshold, while that at the 5'end of the *atpH* mRNA is highly abundant (Cavaiuolo et al, 2017~~; Ozawa et al, 2018~~). The increased stability of *atp*A transcripts upon inhibition of translation, whether by antibiotics treatment or deletion of *TDA1*, may also be due to a lower rate of endonucleolytic cleavages associated with ribosome stalling or pausing (Stollar et al, 1994). The rate of production of the footprint may be lowered up to a level where its dissociation from the M factor and its subsequent degradation take over.

The selective advantage of such a "translate and destroy" mechanism, compared to a less wasteful "translate and recycle" strategy remains to be studied. *atpA* is one of those transcripts that accumulate in large excess over what is required for translation, since as little as 5% of the wild-type *atpA* mRNA levels still ensure a 50% accumulation of the α subunit (Boulouis et al, 2015; Drapier et al, 2002), while less than 50 % of the transcript are loaded on polysomes at a

given time (Eberhard et al, 2011). For many genes, transcription rates exceed what can be stabilised by M factors (Loiselay et al, 2008), while transcript accumulation exceeds what is required for translation (Eberhard et al, 2002; Boulouis et al, 2015). As a consequence, the rate of production of a protein can be adjusted simply by modulating translation initiation. Would biosynthetic processes being adjusted at the level of downstream steps, more complex regulatory events would be required to also increase mRNA accumulation. In bioenergetics organelles, the key to evolutionary success probably resides more in fast responses to environmental changes than in minimizing energy expenditure. The physiological conditions in which the "translate and destroy" system would provide an improved fitness remain to be determined. We note however that translation of *atpA* is subject to CES regulation, being stimulated by unassembled CF1β subunits, whose level in turn depends on CF1α availability (Drapier et al, 2007). Furthermore, we performed the present experiments in acetate-supplemented medium under continuous light and we cannot rule out that a recycling of *atpA* mRNA after translation would be favoured under photoautotrophic, circadian or high-light stress conditions, where the levels of the *atpA* OTAFs may drastically change.

MATERIALS AND METHODS

Strains and growth conditions

Wild-type (t222+: CC5001, derived from 137c: *nit1 nit2 mt+*; Gallaher et al, 2015), mutant, and transformed strains of *C. reinhardtii* were grown at 25°C in Tris-acetate-phosphate (TAP) medium (pH 7.2) (Harris, 1989) under continuous low light (5 to 10 µE m^{-2} s^{-1}). The original *mda1* mutant strain CAL018.01.04*,* obtained by insertional mutagenesis (Dent, 2005) was crossed twice with our reference strain t222+ to generate the zeocin-sensitive strain *mda1-zeo^S mt+* (hereafter called *mda1-1*) used in that study. The nuclear mutants *tda1*Δ and *tda1-F54* have been described in (Eberhard et al, 2011). The *mt* double mutant was obtained in a cross between strains *mda1-1* and *tda1-1*, performed according to Harris, 1989.

Mapping of the *mda1-1* mutation

Strain *mda1-1* was crossed to S1D2 (CC-2290, mt-) and 32 individual zygotes were dissected, from which acetate requiring (ac-) and acetate independent (ac+) progeny were isolated and grown individually in 15 ml cultures. DNA was isolated (DNeasy Plant maxi kit, Qiagen) from cell pellets of pooled cultures of ac+ or ac-progeny. The two DNA samples were sequenced using Illumina technology (PE125) by Eurofins Genomics, in parallel with DNA from S1D2 and t222+. SHOREmap (Schneeberger et al, 2009) identified a region around Chromosome_16:5,300,000 (genome version 5.5) where the ac- pool showed the lowest proportion of S1D2 polymorphisms, and the ac+ pool the highest proportion (Fig. S7). Within this region, SHOREmap 'annotate' identified candidate mutations that were individually examined. A 7-bp deletion (pos. 5478830-5478836) in gene Cre16.g680850 was identified as possible cause of the phenotype.

Construction of plasmids:

Standard nucleic acid manipulations were performed according to (Sambrook et al., 1989). Primers used in this study are listed in Supplemental Table ST1. All DNA constructs were sequenced before transformation in Chlamydomonas.

i) pMDA1-Fl was purchased from the GenScript Company. The *MDA1* coding sequence containing the first intron and the 3xFlag tag was generated by DNA synthesis (Fig. 1B), digested with *Eco*RI and *Bam*HI and cloned into the pPEARL plasmid (Takahashi et al, 2016) digested with the same enzymes.

ii) To construct pTDA1-Fl, the first 2359 bp of the *TDA1* genomic sequence were amplified using the primers NTDA1 FW and NTDA1 RV and cloned into the pPEARL plasmid

into restriction sites *Eco*RI and *Bam*HI. A 3xFlag tag was inserted by overlapping PCR (Higuchi, 1990) in the last 4448 bp of the *TDA1* genomic sequence, using the external primers CTDA1 FW and CTDA1 RV and the mutagenic primers TDA1-Flag-Rev and TDA1-Flag-Cod. The obtained fragment was then cloned into the previous construct using the restriction sites SgrAI and NdeI.

iii) To construct pTDA1-HA, a 3xHA tag was first inserted by overlapping PCR in the first 2359 bp of the *TDA1* genomic sequence, using primers NTDA1 FW and NTDA1 RV (external) and TDA1-HA-Rev and TDA1-HA-Cod (internal). The amplified fragment was cloned into pPAR in restriction sites *Eco*RI and *Bam*HI. The last 4448 bp of the TDA1 genomic sequence were then amplified using primers CTDA1 FW and CTDA1 RV and cloned into the previous construct into *Sgr*AI and *Nde*I sites.

Transformation experiments

Chloroplast transformations were performed by tungsten particle bombardment (Boynton and Gillham, 1993) as described by Kuras and Wollman (1994). Transformants were selected on TAP medium supplemented with spectinomycin (100 µg ml^{-1}) under low light (5 to 10 µE m^{-2} s^{-1}) and further sub-cloned under darkness on selective medium until they reached homoplasmy, assessed by RFLP analysis of specific PCR fragments. At least three independent transformants were analysed and proved identical.

Nuclear transformations were performed by electroporation, as described by Shimogawara et al. (1998) and Raynaud et al, (2007), using the following parameters: 10 µF/1000 to 1200 V cm^{-1}, depending on the recipient strains. Plasmids were linearized with *Psi*I prior to transformation. Transformants were either selected for phototrophy under high light (200 µE m^{-2} s^{-1}) on minimum medium (Harris, 1989) or for resistance to paromomycin (10 µg ml^{-1}) under low light on TAP medium.

Protein isolation and analysis

Protein isolation, separation, and immunoblot analyses were performed on exponentially growing cells ($2x10^{-6}$ cells.ml^{-1}) as described by Kuras and Wollman (1994). Cell extracts were loaded on an equal chlorophyll basis. Protein samples from co-immuno-precipitation, size-exclusion chromatography and polysome gradients were loaded on an equal volume basis. At least three biological replicas were performed for each experiment. Proteins were detected by ECL using the ChemiDoc Imager (BioRad) and quantified, as described in (Choquet et al, 2003), using the ImageLab software (BioRad). Primary antibodies, diluted

100,000-fold (cyt. *f*), 50,000-fold (CF1β, RuBisCO), 10,000-fold (CF1α, RPS12), were revealed by horseradish peroxidase–conjugated antibodies against rabbit IgG (Promega). Antibodies against α and β subunits of CF1, and cytochrome *f* have been previously described (Lemaire et al, 1989; Kuras and Wollman, 1994). HA- and Flag-tagged proteins were detected by ECL using monoclonal antibodies anti HA.11 (Covance) and anti-Flag M2 (Sigma-Aldrich), respectively, and revealed with horseradish peroxidase-conjugated antibody against mouse IgG (Promega).

Subcellular protein localisation was assessed as described in Boulouis et al, 2011.

RNA isolation and analysis

Total RNAs were extracted and analysed as described in (Drapier et al, 1998) except that probes amplified with primers listed in Supplemental Table ST1 were Digoxigenin-labelled, using DIG-dUTP, the anti-Digoxigenin Fab fragment and CDP Star reagent (Roche). Signal was aquired in ChemiDoc Imager (BioRad) and analysed with the ImageLab software (BioRad). Ribosomal RNAs were visualized by Methylene blue staining of filters. Circular RT-PCR was performed as described in (Pfalz et al., 2009). PCR products were gel-extracted and sequenced. sRNA-Seq datasets (SRA BioProject PRJNA379963) were produced and analysed as described in (Cavaiuolo et al, 2017).

RNA chase was performed on exponentially growing cells in TAP medium, supplemented at t = 0 with rifampicin (350 µg ml^{-1}), lincomycin (500 µg ml^{-1}), or both. At the indicated time points, aliquots were removed for RNA isolation.

RNA and protein co-immuno-precipitations

Immuno-precipitations and RNase treatment of soluble extracts were done as described by Boulouis et al (2015), with minor modifications since cells were broken by sonication and the immuno-precipitation was performed on 1mL fraction at 4°C for 2 hr. For RNA immunoprecipitation (RIP), beads, resuspended in 250 µL AE buffer (50 mM Na-acetate pH 5.2, 10 mM EDTA), were extracted once with phenol/ chloroform/IsoAmyl Alcohol (25:24:1) before ethanol precipitation in the presence of 2 µL GlycoBlue (Invitrogen, USA). For CAP treatment, chloramphenicol (100 µg.mL^{-1}) was added to the cells 10 mn before sonication.

Size-exclusion chromatography

Size-exclusion chromatography was done according to Boulouis et al (2015), with minor modifications since cells were broken by sonication and protein separation was

performed on Superose 6 Increase 10/300 GL column (GE Healthcare) column, with elution performed at 4°C at a rate of 0.2 $ml \cdot min^{-1}$.

Polysome analysis

Polysomes were purified as described by Minai et al. (2006) under $MgCl_2$ conditions. Initial solutions to form the gradients were prepared at 15% or 55% w/v sucrose, but the real sucrose concentration of the collected fractions was determined by refractrometry. Both RNA and proteins were purified from each collected fraction.

Figure legends:

Figure 1. The *mda1-1* mutant strain lacks stable accumulation of *atpA* mRNA.

A) Accumulation of *atpA* mRNA (left) and ATPase α subunit (CF1α, right) in wild-type, *tda1Δ* and *mda1-1* strains. Accumulations of *petD* mRNA and cytochrome *f* provide the respective loading controls.

B) Structures of the *MDA1* gene (upper) and of the synthetic gene (lower) used for complementation. UTRs are shown in light grey and exons in dark grey, the 7 bp deleted in the *mda1-1* strain and the positions of the sequences encoding the tags and the OPR repeats in the synthetic gene are indicated. The nucleotide sequence encompassing the deletion (Δ, underlined) and its translation in the wild-type (black) and mutant (red) alleles are shown. Grey nucleotides correspond to the frameshift-induced STOP codon.

C) Schematic representation of the MDA1 and TDA1 proteins, showing the localisation of the predicted chloroplast targeting peptide (cTP, arrow), the tags insertion points (triangles) and of the OPR repeats (boxes). The position of the *mda1-1* mutation is also shown.

D) Accumulation of *atpA* and *Cβlp2* (loading control) mRNAs in wild-type, *mda1-1* and seven independent clones complemented with the *MDA1-Fl* construct. Accumulation of MDA1 (Flag), ATPase α subunit (CF1α) and cytochrome *f* (loading control) proteins in the same strains is shown below.

E) MDA1 and TDA1 are soluble proteins.

Whole cell extracts (E) from the *mt*MT strain, overlaid on a 1.5 M sucrose cushion, were separated by ultracentrifugation into the supernatant (S), membrane (M), cushion (C) and pellet (P) fractions. Equal volumes of each fraction were loaded on gel and probed with the indicated antibodies. Hsp70 and cytochrome *f* provide markers of the soluble and membrane fractions, respectively.

Figure 2. Both MDA1 and TDA1 physically interact with the *atpA* 5'UTR.

A) Schematic representation of the 5'*dAf* chimera, used to replace the *petA* endogenous gene in the *mda1-1 {5'dAF}* and *MDA1-Fl {5'dAF}* strains.

B) Transcript accumulation detected using probes specific for the *atpA* 5'UTR (shown in panel D), *petA* coding sequence and *Cβlp2* (nuclear control) in *mda1-1* and *MDA1-Fl* recipient strains and in the respective transformants containing the 5'*dAf* chimeric construct (indicated by the stars).

C) RNA immuno-precipitation with anti-Flag magnetic beads identifying RNAs associated with TDA1-Fl and MDA1-Fl in the strains respectively expressing them (wild-type used as control strain). Dot blot of input (E) and beads pellet (P) RNA samples for each strain was hybridised against *atpA-5'* and *petD* (control) probes.
D) Schematic representation of the *atpA* gene model with a zoom on the 5'UTR showing the positions of the probes (black lines) used in dot blot experiments. The *atpA* transcription start site (TSS) and the mature 5'end are indicated. The black rectangle symbolises a sRNA mapping at the 5'end of the mature transcript (see below).
E) Immuno-precipitated RNA samples from panel C were probed with the partially overlapping probes 5'*atpA* 1 to 5 with the control probe *atpA-3'*.

Figure 3. The *atpA* 5'UTR mediates the *in-vivo* interaction of MDA1 and TDA1.
A) RNA immuno-precipitation with anti-Flag and anti-HA magnetic beads identifying RNAs associated with MDA1-Fl and TDA1-HA in untreated and RNAse-treated *mt*MT samples. Dot blot of input (E) and beads pellet (P) RNA fractions for each sample analysed with 5'*atpA-3* (shown in Figure 3D) and *petD* (control) probes.
B) Immuno-detection of MDA1-Fl and TDA1-HA in the input (E), supernatant (S) and pellet (P) samples from the same immuno-precipitations as in (A). The antibody against the ribosomal protein Rps12 was used as control.

Figure 4. Interaction of MDA1, TDA1 and *atpA* mRNA probed by size-exclusion chromatography.
Soluble cell extract from strains *mt*MT, either untreated or RNase-treated, *mt*M or *mt*T, respectively expressing both MDA1-Fl and MDA1-HA, MDA1-Fl alone or MDA1-HA alone were fractionated on a Superose 6 column. After gel electrophoresis, fractions were immunodecorated with anti-Flag or HA antibodies. Molecular masses of the complexes found in each fraction were estimated by comparison with protein standards of the MHW gel filtration calibration kit and with the elution profile of the RuBisCO holoenzyme (represented by the large subunit, LSU).

Figure 5. Mapping of *atpA* mRNA 5'ends and sRNA accumulation in presence or absence of MDA1 and TDA1.

A) Schematic representation of the *atpA-psbI-cemA-atpH* gene cluster. Coding sequences are represented by grey boxes, promoters by bent arrows. Positions of the primers used for retro-transcription (*A-5'* 2 RV; grey) and cRT-PCR (*A-5'* 1 RV and *A-3'* FW; black) are shown.
B) Agarose gel electrophoresis of the cRT-PCR amplicons from RPP- and mock-treated samples from the WT, *mda1-1*, *MDA1-Fl* and *tda1Δ* strains.
C) Distribution of sRNAs in mutant strains *mda1-1* (red) and *tda1*Δ (blue) along the *atpA* 5'UTR compared to that in the wild-type (black). Coverage is normalized as Reads Per Million (RPM) and averaged over two biological replicates. The lower scheme depicts the *atpA* 5'UTR region (same convention than in Fig. 2D).

Figure 6. MDA1 and TDA1 are not associated with the *atpA* mRNAs engaged in translation, which are destabilized.

A) Solubilized whole-cell extracts (T) from *TDA1-HA* and *MDA1-Fl* complemented strains collected in presence of chloramphenicol (CAP, 100 µg ml^{-1}) were loaded on sucrose gradients. After ultracentrifugation, ten fractions were collected and used for both RNA and protein extraction. The distribution of *atpA* along the gradient was analysed by RNA blotting (ribosomal RNAs were detected by methylene blue staining of the filters). The distributions of TDA1, MDA1 and of ribosomal protein Rps12 were analysed by immunoblotting. Sucrose concentrations are indicated at the bottom.
B) RNA co-immuno-precipitated with TDA1-HA and MDA1-Fl in untreated and CAP-pre-treated *mt*MT samples analysed by dot-blot with *16S* rRNA and *petD* (control) probes.
C) MDA1, TDA1 and ribosomal protein Rps12 immunodetection in the input (E), supernatant (S) and pellet (P) fractions from the CAP-treated samples shown in B.
D) Wild-type cells were treated with rifampicin (Rif, 350 µg ml^{-1}), lincomycin (Lin, 500 µg ml^{-1}) or both (Rif+Lin). RNA was extracted just before or 3 and 6 hr after inhibitors addition. Accumulation of *atpA* and *Cβlp2* (nuclear control) were analysed by RNA blot, ribosomal RNAs were detected by methylene blue staining of the filters. For each condition, the percentage of *atpA* mRNA amounts relative to time 0 is shown on the right (n=3).

Figure 7: Working model for the post-transcriptional control of the *atpA* gene expression.

De novo transcribed *atpA* mRNA molecules (represented in the monocistronic form for sake of simplicity), possessing a primary 5'end (in blue), are stabilized by the binding of MDA1 (M) to a region located in the middle of the 5'UTR. The binding of MDA1 is likely cooperative

with that of another factor (question mark) that binds the 5'end of the transcript, protecting it from exonucleases and defining the maturation site. MDA1 can bind *atpA* in absence of TDA1 and its level determines the size of the stable *atpA* mRNA pool. TDA1 (T) interacts with the 5'UTR of a fraction of this pool in close proximity with MDA1. The binding of TDA1 activates the translation of the *atpA* mRNAs. When translation is initiated, MDA1 and TDA1 dissociate from the 5'UTR of ribosome-bound *atpA* mRNAs, leaving them unprotected. After or during translation, the mRNAs are degraded by nuclease and are not re-bound by MDA1, which is recycled to bind newly transcribed *atpA* mRNAs; TDA1 is probably also recycled.

SUPPLEMENTAL DATA:

Suppl. Figure S1: Annotated sequence of the MDA1 protein.
Suppl. Figure S2: Conservation of MDA1 proteins among Chlamydomonadales algae.
Suppl. Figure S3: Complementation of *mda1-1* with BAC 16F5.
Suppl. Figure S4: MDA1 binds the *atpA* 5'UTR in absence of TDA1.
Suppl. Figure S5: Mapping of di- and tri-cistronic *atpA* mRNA 5'ends in presence or absence of MDA1.
Suppl. Figure S6: Stability of TDA1-Fl and MDA1-Fl.
Suppl. Figure S7: Mapping of the *mda1-1* mutation.

Suppl. Table ST1: Oligonucleotides used in this study.
Suppl. Table ST2: Mapping 5' and 3'ends of the monocistronic *atpA* mRNA in presence or absence of MDA1 and TDA1.
Suppl. Table ST3: Mapping 5'ends of the di- and tri-cistronic *atpA* mRNAs in presence or absence of MDA1.

ACKNOWLEDGEMENTS

We thank Dr Rachel Dent and Dr Silvia Raimundo for their kind gifts of the strain CAL018.01.04 (*mda1-1*) and of the antibody against Rps12, respectively, Nicolas Tourasse for setting up the pipeline used to identify the *mda1-1* mutation and all members of UMR7141 for stimulating discussions.

FUNDING

This work was supported by Unité Mixte de Recherche 7141, CNRS/Sorbonne Université, by the Agence Nationale de la Recherche (ChloroRNP: ANR-13-BSV7-0001-01), and by the "Initiative d'Excellence" program (Grant "DYNAMO," ANR-11-LABX-0011-01).

No conflict of interest declared.

Accession numbers: Sequence data from this article can be found in the UniProtKB/Swiss-Prot or GenBank databases under the following accession numbers: cytochrome f: CAA51422.1; RbcL: J01399.1; Rps12: AAC16329 ; AtpA: P26526 ; *cblp2*: X53574 ; *petA*: X72919 ; *TDA1* : CCA62914/FR850045 ; *MDA1*: Cre16.g680850.t1.1 (https://phytozome.jgi.doe.gov/pz/portal.html#!gene?search=1&detail=1&method=4614&searchText=transcriptid:30777176

REFERENCES

Auchincloss, A.H., Zerges, W., Perron, K., Girard-Bascou, J., and Rochaix, J.D. (2002). Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast *psbC* mRNA in *Chlamydomonas reinhardtii*. J. Cell Biol. **157**:953-962.

Barkan, A. (2011). Expression of Plastid Genes: Organelle-Specific Elaborations on a Prokaryotic Scaffold. Plant Physiol. **155**:1520–1532.

Barneche, F., Winter, V., Crèvecœur, M. and Rochaix, J. (2006). *ATAB2* is a novel factor in the signalling pathway of light-controlled synthesis of photosystem proteins. EMBO J. **25**: 5907-5918.

Blowers, A. D., Klein, U., Ellmore, G. S., and Bogorad, L. (1993). Functional in vivo analyses of the 3′ flanking sequences of the Chlamydomonas chloroplast *rbcL* and *psaB* genes. Mol. Gen. Genet. **238**:339–349.

Boudreau, E., Nickelsen, J., Lemaire, S. D., Ossenbühl, F., and Rochaix, J. D. (2000). The Nac2 gene of Chlamydomonas encodes a chloroplast TPR-like protein involved in *psbD* mRNA stability. EMBO J. **19**:3366–76.

Boulouis, A., Drapier, D., Razafimanantsoa, H., Wostrikoff, K., Tourasse, N. J., Pascal, K., Girard-Bascou, J., Vallon, O., Wollman, F.-A., and Choquet, Y. (2015). Spontaneous dominant mutations in Chlamydomonas highlight ongoing evolution by gene diversification. Plant Cell **27**:984–1001.

Boulouis, A., Raynaud, C., Bujaldon, S., Aznar, A., Wollman, F.-A., and Choquet, Y. (2011). The nucleus-encoded trans -acting factor MCA1 plays a critical role in the regulation of cytochrome f synthesis in Chlamydomonas chloroplasts. Plant Cell **23**:333–349.

Boynton, J. E., and Gillham, N. W. (1993). Chloroplast transformation in Chlamydomonas. Methods Enzymol. **217**:510–536.

Cavaiuolo, M., Kuras, R., Wollman, F. A., Choquet, Y., and Vallon, O. (2017). Small RNA profiling in Chlamydomonas: Insights into chloroplast RNA metabolism. Nucleic Acids Res. **45**:10783–10799.

Choquet, Y., Stern, D. B., Wostrikoff, K., Kuras, R., Girard-Bascou, J., and Wollman, F.-A. (1998). Translation of cytochrome f is autoregulated through the 5′ untranslated region of *petA* mRNA in Chlamydomonas chloroplasts. Proc. Natl. Acad. Sci. **95**:4380–4385.

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

Choquet Y., Zito F., Wostrikoff K., Wollman F.A. (2003). Cytochrome *f* translation in Chlamydomonas chloroplast is autoregulated by its carboxyl-terminal domain. Plant Cell **15**:1443-1454.

Dauvillée, D., Stampacchia, O., Girard-Bascou, J., and Rochaix, J. D. (2003). Tab2 is a novel conserved RNA binding protein required for translation of the chloroplast *psaB* mRNA. EMBO J. **22**:6378–6388.

Dent, R. M. (2005). Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of *Chlamydomonas reinhardtii*. Plant Physiol. **137**:545–556.

Douglas, A. E., and Raven, J. A. (2003). Genomes at the interface between bacteria and organelles. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **358**:5–18.

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

Drapier, D., Girard-Bascou, J., and Wollman, F.-A. (1992). Evidence for nuclear control of the expression of the *atpA* and *atpB* chloroplast genes in Chlamydomonas. Plant Cell **4**:283–295.

Drapier, D., Suzuki, H., Levy, H., Rimbault, B., Kindle, K. L., Stern, D. B., and Wollman, F. A. (1998). The chloroplast *atpA* gene cluster in *Chlamydomonas reinhardtii*. Functional analysis of a polycistronic transcription unit. Plant Physiol. **117**:629–641.

Drapier, D., Girard-Bascou, J., Stern, D.B., and Wollman, F.A. (2002). A dominant nuclear mutation in Chlamydomonas identifies a factor controlling chloroplast mRNA stability by acting on the coding region of the *atpA* transcript. Plant J **31,** 687-697.

Drapier, D., Rimbault, B., Vallon, O., Wollman, F. A., and Choquet, Y. (2007). Intertwined translational regulations set uneven stoichiometry of chloroplast ATP synthase subunits. Embo J. **26**:3581–3591.

Eberhard, S., Drapier, D., and Wollman, F. A. (2002). Searching limiting steps in the expression of chloroplast-encoded proteins: Relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. Plant J. **31**:149–160.

Eberhard, S., Loiselay, C., Drapier, D., Bujaldon, S., Girard-Bascou, J., Kuras, R., Choquet, Y., and Wollman, F. A. (2011). Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of *atpA* transcripts in *Chlamydomonas reinhardtii* chloroplasts. Plant J. **67**:1055–1066.

Fargo, D.C., Zhang, M., Gillham, N.W., and Boynton, J.E. (1998). Shine-Dalgarno-like sequences are not required for translation of chloroplast mRNAs in *Chlamydomonas reinhardtii* chloroplasts or in *Escherichia coli*. Mol. Gen. Genet. **257**:271-282.

Gallaher S.D., Fitz-Gibbon S.T., Glaesener A.G., Pellegrini M., and Merchant S.S. (2015). Chlamydomonas genome resource for laboratory strains reveals a mosaic of sequence variation, identifies true strain histories, and enables strain-specific studies. Plant Cell **27**:2335-2352

Germain, A. , Hotto, A. M., Barkan, A. and Stern, D. B. (2013), RNA processing and decay in plastids. WIREs RNA, **4**:295-316.

Girard-Bascou, J., Pierre, Y., and Drapier, D. (1992). A nuclear mutation affects the synthesis of the chloroplast *psbA* gene production *Chlamydomonas reinhardtii*. Curr.Genet. **22**:47-52.

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.

Gross, C.H., Ranum, L.P., and Lefebvre, P.A. (1988). Extensive restriction fragment length polymorphisms in a new isolate of *Chlamydomonas reinhardtii*. Curr. Genet. **13**:503-508.

Hammani, K., Takenaka, M., Miranda, R., and Barkan, A. (2016). A PPR protein in the PLS subfamily stabilizes the 5′-end of processed *rpl16* mRNAs in maize chloroplasts. Nucleic Acids Res. **44**:4278–4288.

Harris, E.H. (1989). The Chlamydomonas source book: a comprehensive guide to biology and laboratory use. San Diego: Academic Press.

Higuchi, R. (1990). Recombinant PCR. In: PCR protocols: a guide to methods and application. London, New York: Academic press.

Hirose, T., and Sugiura, M. (2004). Functional Shine-Dalgarno-like sequences for translational initiation of chloroplast mRNAs. Plant & cell physiology **45,** 114-117.

Holland, H. D. (2006). The Oxygenation of the Atmosphere and Oceans. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **361**:903–915.

Howe, C. J., Barbrook, A. C., Koumandou, V. L., Nisbet, R. E. R., Symington, H. A., and Wightman, T. F. (2003). Evolution of the chloroplast genome. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **358**:99-106-107.

Johnson, X., Wostrikoff, K., Finazzi, G., Kuras, R., Schwarz, C., Bujaldon, S., Nickelsen, J., Stern, D.B., Wollman, F.A., and Vallon, O. (2010). MRL1, a conserved Pentatricopeptide repeat

protein, is required for stabilization of *rbcL* mRNA in Chlamydomonas and Arabidopsis. Plant Cell **22,** 234-248.

Kleine, T., Maier, U. G., and Leister, D. (2009). DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. Annu. Rev. Plant Biol. **60**:115–138.

Kleinknecht, L., Wang, F., Stube, R., Philippar, K., Nickelsen, J., and Bohne, A.V. (2014). RAP, the sole octotricopeptide repeat protein in Arabidopsis, is required for chloroplast 16S rRNA maturation. Plant Cell **26**:777-787.

Kuras, R., and Wollman, F. A. (1994). The assembly of cytochrome b6/f complexes: an approach using genetic transformation of the green alga *Chlamydomonas reinhardtii*. EMBO J. **13**:1019-1027.

Lemaire, C. and Wollman, F. A. (1989). The chloroplast ATP synthase in *Chlamydomonas reinhardtii*. J. Biol. Chem. **264**: 10235-10242.

Lim, K., Kobayashi, I., and Nakai, K. (2014). Alterations in rRNA-mRNA interaction during plastid evolution. Mol. Biol. Evol. **31**:1728-1740.

Loiselay, C., Gumpel, N. J., Girard-Bascou, J., Watson, A. T., Purton, S., Wollman, F.-A., and Choquet, Y. (2008). Molecular identification and function of cis- and trans-acting determinants for *petA* transcript stability in *Chlamydomonas reinhardtii* chloroplasts. Mol. Cell. Biol. **28**:5529-5542.

Loizeau, K., Qu, Y., Depp, S., Fiechter, V., Ruwe, H., Lefebvre-Legendre, L., Schmitz-Linneweber, C., and Goldschmidt-Clermont, M. (2014). Small RNAs reveal two target sites of the RNA-maturation factor Mbb1 in the chloroplast of Chlamydomonas. Nucleic Acids Res. **42**:3286–3297.

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**:159–75.

Nickelsen, J., van Dillewijn, J., Rahire, M., 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**:3182–91.

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.

Perron, K., Goldschmidt-Clermont, M., and Rochaix, J.D. (2004). A multiprotein complex involved in chloroplast group II intron splicing. RNA **10**:704-711.

Pesaresi, P., Schneider, A., Kleine, T., and Leister, D. (2007). Interorganellar communication. Curr. Opin. Plant Biol. **10**:600–606.

Pfalz, J., Ali Bayraktar, O., Prikryl, J., and Barkan, A. (2009). Site-specific binding of a PPR protein defines and stabilizes 5′ and 3′ mRNA termini in chloroplasts. EMBO J. **28**:2042–52.

Ponce-Toledo, R.I., Deschamps, P., Lopez-Garcia, P., Zivanovic, Y., Benzerara, K., and Moreira, D. (2017). An early-branching freshwater cyanobacterium at the origin of plastids. Current biology : CB **27**:386-391.

Rahire, M., Laroche, F., Cerutti, L., and Rochaix, J. D. (2012). Identification of an OPR protein involved in the translation initiation of the PsaB subunit of photosystem I. Plant J. **72**:652–661.

Raynaud, C., Loiselay, C., Wostrikoff, K., Kuras, R., Girard-Bascou, J., Wollman, F.A., and Choquet, Y. (2007). Evidence for regulatory function of nucleus-encoded factors on mRNA stabilization and translation in the chloroplast. Proc. Natl. Acad. Sci. **104**:9093-9098.

Ruwe, H., and Schmitz-Linneweber, C. (2012). Short non-coding RNA fragments accumulating in chloroplasts: Footprints of RNA binding proteins? Nucleic Acids Res. **40**:3106–3116.

Ruwe, H., Wang, G., Gusewski, S., and Schmitz-Linneweber, C. (2016). Systematic analysis of plant mitochondrial and chloroplast small RNAs suggests organelle-specific mRNA stabilization mechanisms. Nucleic Acids Res. **44**:7406–7417.

Rymarquis, L.A., Handley, J.M., Thomas, M., and Stern, D.B. (2005). Beyond complementation. Map-based cloning in *Chlamydomonas reinhardtii*. Plant Physiol. **137**:557-566.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: Cold Spring Harbor Laboratory Press.

Scharff, L.B., Ehrnthaler, M., Janowski, M., Childs, L.H., Hasse, C., Gremmels, J., Ruf, S., Zoschke, R., and Bock, R. (2017). Shine-Dalgarno sequences play an essential role in the translation of plastid mRNAs in Tobacco. Plant Cell **29**:3085-3101.

Schneeberger, K., Ossowski, S., Lanz, C., Juul, T., Petersen, A. H., Nielsen, K. L., Jørgensen, J. E., Weigel, D., and Andersen, S. U. (2009). SHOREmap: Simultaneous mapping and mutation identification by deep sequencing. Nat. Methods **6**:550–551.

Schroda, M., Beck, C. F., and Vallon, O. (2002). Sequence elements within an HSP70 promoter counteract transcriptional transgene silencing in Chlamydomonas. Plant J. **31**:445–455.

Schwarz, C., Elles, I., Kortmann, J., Piotrowski, M., and Nickelsen, J. (2007). Synthesis of the D2 protein of photosystem II in Chlamydomonas is controlled by a high molecular mass complex containing the RNA stabilization factor Nac2 and the translational activator RBP40. Plant Cell **19**:3627–3639.

Shimogawara, K., Fujiwara, S., Grossman, A., and Usuda, H. (1998). High-efficiency transformation of *Chlamydomonas reinhardtii* by electroporation. Genetics **148**:1821–1828.

Stampacchia, O., Girard-Bascou, J., Zanasco, J.L., Zerges, W., Bennoun, P., and Rochaix, J.D. (1997). A nuclear-encoded function essential for translation of the chloroplast *psaB* mRNA in Chlamydomonas. Plant Cell **9**:773-782.

Stern, D. B., and Kindle, K. L. (1993). 3' end maturation of the *Chlamydomonas reinhardtii* chloroplast *atpB* mRNA is a two-step process. Mol. Cell. Biol. **13**:2277–2285.

Stern, D. B., Goldschmidt-Clermont, M., and Hanson, M. R. (2010). Chloroplast RNA metabolism. Annu. Rev. Plant Biol. **61**:125–155.

Stollar, N.E., Kim, J.K., and Hollingsworth, M.J. (1994). Ribosomes pause during the expression of the large ATP synthase gene cluster in spinach chloroplasts. Plant Physiol. **105**:1167-1177.

Stoppel, R., Lezhneva, L., Schwenkert, S., Torabi, S., Felder, S., Meierhoff, K., Westhoff, P., and Meurer, J. (2011). Recruitment of a ribosomal release factor for light- and stress-dependent regulation of *petB* transcript stability in Arabidopsis chloroplasts. Plant Cell **23**:2680–2695.

Takahashi, H., Schmollinger, S., Lee, J.H., Schroda, M., Rappaport, F., Wollman, F.A., and Vallon, O. (2016). peto interacts with other effectors of cyclic electron flow in Chlamydomonas. Mol. Plant **9**:558-568.

Tardif, M., Atteia, A., Specht, M., Cogne, G., Rolland, N., Brugière, S., Hippler, M., Ferro, M., Bruley, C., Peltier, G., et al. (2012). Predalgo: A new subcellular localization prediction tool dedicated to green algae. Mol. Biol. Evol. **29**:3625–3639.

Uniacke, J., and Zerges, W. (2007). Photosystem II Assembly and Repair Are Differentially Localized in Chlamydomonas. Plant Cell **19**:3640–3654.

Vaistij, F.E., Boudreau, E., Lemaire, S.D., Goldschmidt-Clermont, M., and Rochaix, J.D. (2000). 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. **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**:469-82.

Wang, F., Johnson, X., Cavaiuolo, M., Bohne, A.V., Nickelsen, J., and Vallon, O. (2015). Two Chlamydomonas OPR proteins stabilize chloroplast mRNAs encoding small subunits of photosystem II and cytochrome b6 f. Plant J. **82**:861-873.

Woodson, J. D., and Chory, J. (2008). Coordination of gene expression between organellar and nuclear genomes. Nat. Rev. Genet. **9**:383–395.

Wostrikoff, K., Choquet, Y., Wollman, F.A., and Girard-Bascou, J. (2001). TCA1, a single nuclear-encoded translational activator specific for *petA* mRNA in *Chlamydomonas reinhardtii* chloroplast. Genetics **159**:119-132.

Zerges, W., and Rochaix, J.D. (1994). The 5' leader of a chloroplast mRNA mediates the translational requirements for two nucleus-encoded functions in *Chlamydomonas reinhardtii*. Mol.Cell Biol. **14**:5268-5277.

Zhelyazkova, P., Hammani, K., Rojas, M., Voelker, R., Vargas-Suárez, M., Börner, T., and Barkan, A. (2012). Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. Nucleic Acids Res. **40**:3092–3105.

Zhou, W., Lu, Q., Li, Q., Wang, L., Ding, S., Zhang, A., Wen, X., Zhang, L., and Lu, C. (2017). PPR-SMR protein SOT1 has RNA endonuclease activity. Proc. Natl. Acad. Sci. **114**:E1554–E1563.

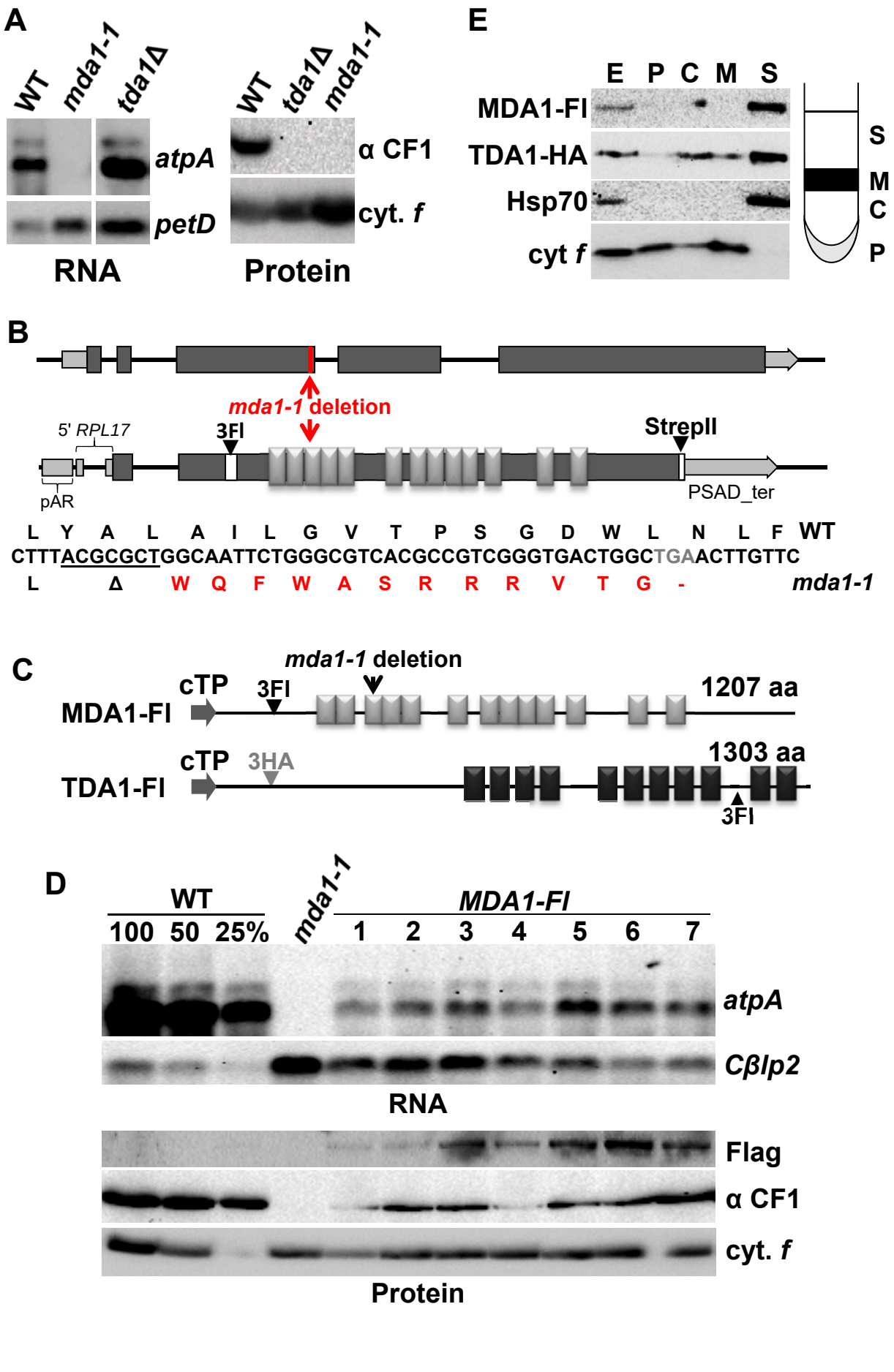

A
WT
mda1-1
tda1Δ
atpA
petD
RNA
WT
tda1Δ
mda1-1
α CF1
cyt. f
Protein
E
E P C M S
MDA1-FI
TDA1-HA
Hsp70
cyt f
S
M
C
P
B
mda1-1 deletion
5' RPL17
3FI
StrepII
pAR
PSAD_ter
L Y A L A I L G V T P S G D W L N L F WT
CTTTACGCGCTGGCAATTCTGGGCGTCACGCCGTCGGGTGACTGGCTGAACTTGTTC
L Δ W Q F W A S R R R V T G - mda1-1
C
mda1-1 deletion
cTP
3FI
MDA1-FI
1207 aa
cTP
3HA
TDA1-FI
1303 aa
3FI
D
WT
100 50 25%
mda1-1
MDA1-FI
1 2 3 4 5 6 7
atpA
Cβlp2
RNA
Flag
α CF1
cyt. f
Protein

Figure 1. The *mda1-1* mutant strain lacks stable accumulation of *atpA* mRNA.

A) Accumulation of *atpA* mRNA (left) and ATPase α subunit (α CF1, right) in wild-type, *tda1Δ* and *mda1-1* strains. Accumulations of *petD* mRNA and cytochrome *f* provide the respective loading controls.

B) Structures of the *MDA1* gene (upper) and of the synthetic gene (lower) used for complementation. UTRs are shown in light grey and exons in dark grey, the 7 bp deleted in the *mda1-1* strain are indicated and the positions of the sequences encoding the tags and the OPR repeats in the synthetic gene. The nucleotide sequence encompassing the deletion (Δ, underlined) and its translation in the wild-type (black) and mutant (red) alleles are shown. Grey nucleotides correspond to the frameshift-induced STOP codon.

C) Schematic representation of the MDA1 and TDA1 proteins, showing the localisation of the predicted chloroplast targeting peptide (cTP, arrow), the tags insertion points (triangles) and of the OPR repeats (boxes). The position of the *mda1-1* mutation is also shown.

D) Accumulation of *atpA* and *Cβlp2* (loading control) mRNAs in wild-type, *mda1-1* and seven independent clones complemented with the *MDA1-Fl* construct. Accumulation of MDA1 (Flag), ATPase α subunit (α CF1) and cytochrome *f* (loading control) proteins in the same strains is shown below.

E) MDA1 and TDA1 are soluble proteins.

Whole cell extracts (E) from the *mt*MT strain, overlaid on a 1.5 M sucrose cushion, were separated by ultracentrifugation into the supernatant (S), membrane (M), cushion (C) and pellet (P) fractions. Equal volumes of each fraction were loaded on gel and probed with the indicated antibodies. Hsp70 and cytochrome *f* provide markers of the soluble and membrane fractions, respectively.

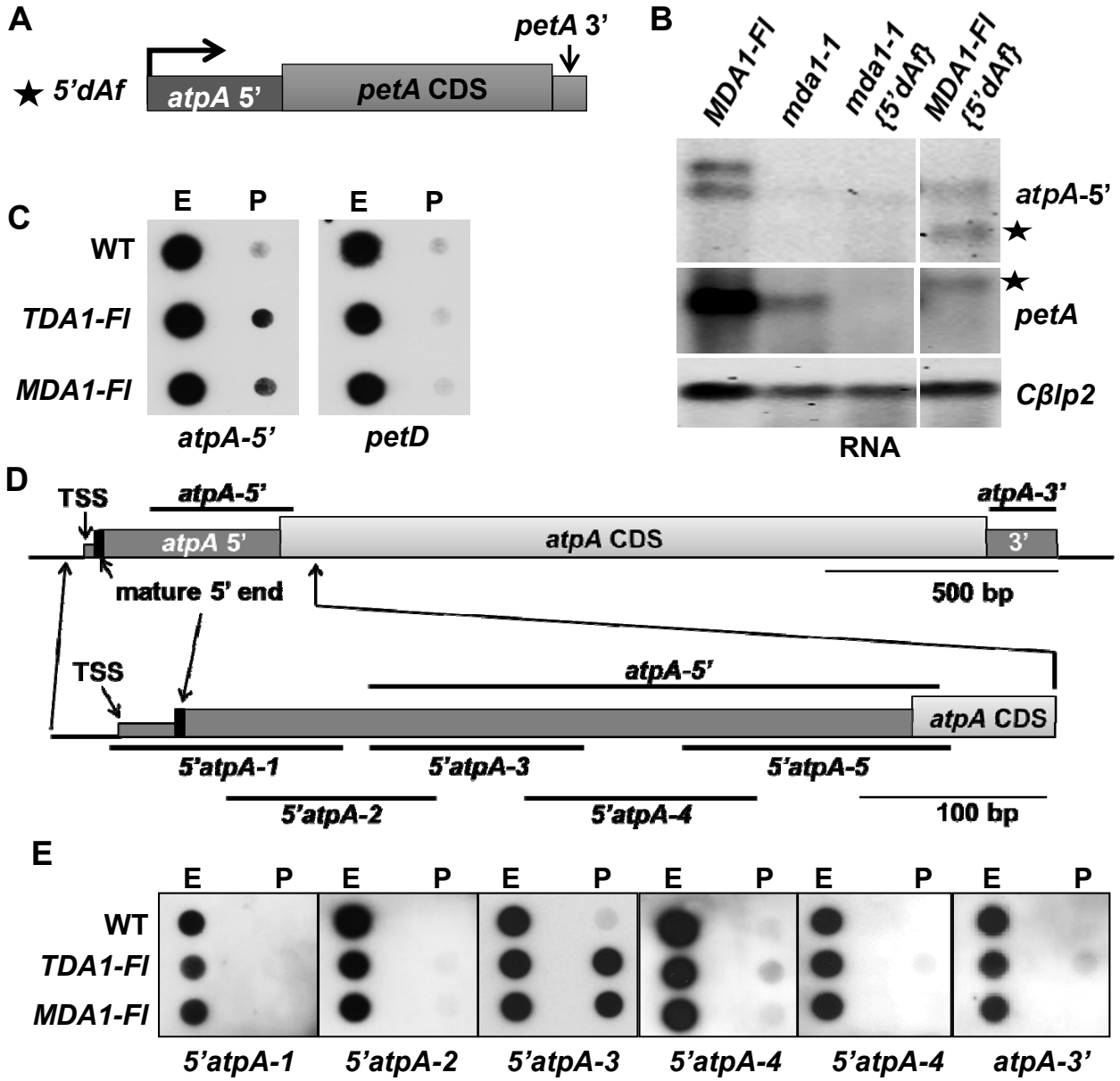


Figure 2. Both MDA1 and TDA1 physically interact with the *atpA* 5'UTR.

A) Schematic representation of the 5'*dAf* chimera, used to replace the *petA* endogenous gene in the *mda1-1 {5'dAF}* and *MDA1-Fl {5'dAF}* strains.

B) Transcript accumulation detected using probes specific for the *atpA* 5'UTR (shown in panel D), *petA* coding sequence and *Cβlp2* (nuclear control) in *mda1-1* and *MDA1-Fl* recipient strains and in the respective transformants containing the 5'*dAf* chimeric construct (indicated by the stars).

C) RNA immuno-precipitation with anti-Flag magnetic beads identifying RNAs associated with TDA1-Fl and MDA1-Fl in the strains respectively expressing them (wild-type used as control strain). Dot blot of input (E) and beads pellet (P) RNA samples for each strain was hybridised against *atpA-5'* and *petD* (control) probes.

D) Schematic representation of the *atpA* gene model with a zoom on the 5'UTR showing the positions of the probes (black lines) used in dot blot experiments. The *atpA* transcription start site (TSS) and the mature 5'end are indicated. The black rectangle symbolises a sRNA mapping at the 5'end of the mature transcript (see below).

E) Immuno-precipitated RNA samples from panel C were probed with the partially overlapping probes 5'*atpA* 1 to 5 with the control probe *atpA-3'*.

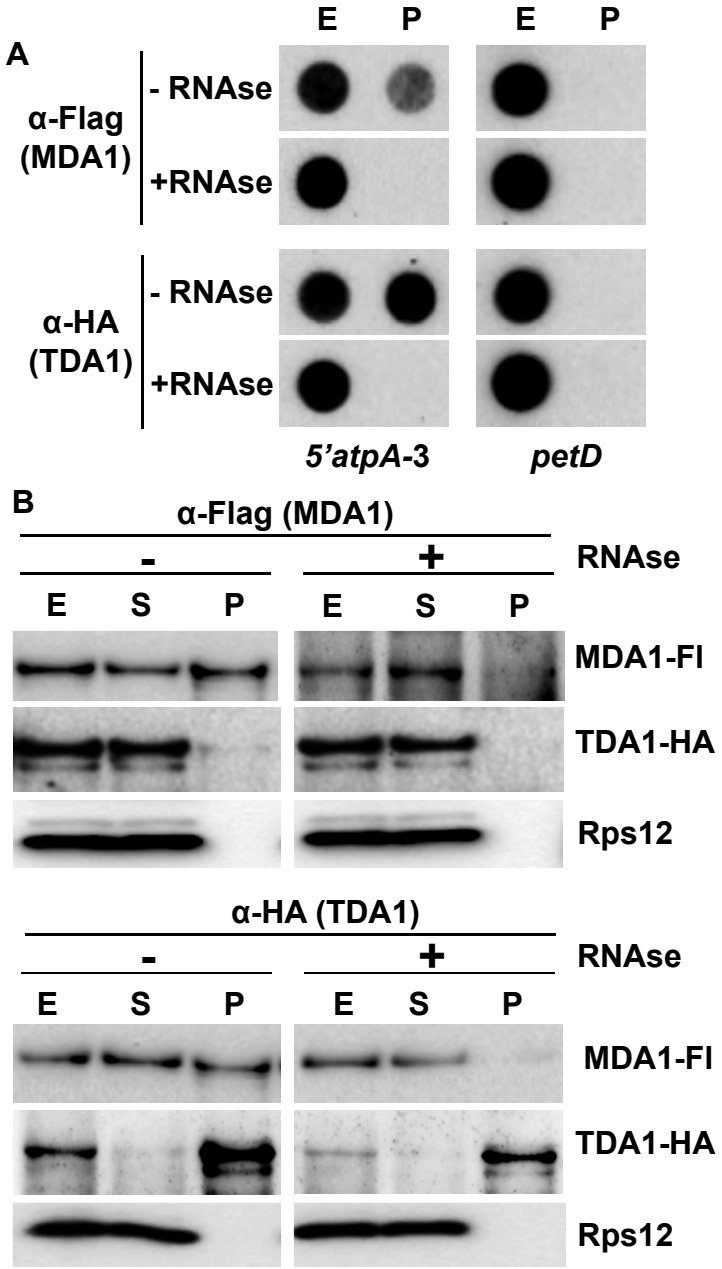


Figure 3. The *atpA* 5'UTR mediates the *in-vivo* interaction of MDA1 and TDA1.
A) RNA immunoprecipitation with anti-Flag and anti-HA magnetic beads identifying RNAs associated with MDA1-Fl and TDA1-HA in untreated and RNAse-treated *mt*MT samples. Dot blot of input (E) and beads pellet (P) RNA fractions for each sample analysed with 5'*atpA-3* (shown in Figure 3D) and *petD* (control) probes.
B) Immuno-detection of MDA1-Fl and TDA1-HA in the input (E), supernatant (S) and pellet (P) samples from the same immuno-precipitations as in (A). The antibody against the ribosomal protein Rps12 was used as control.

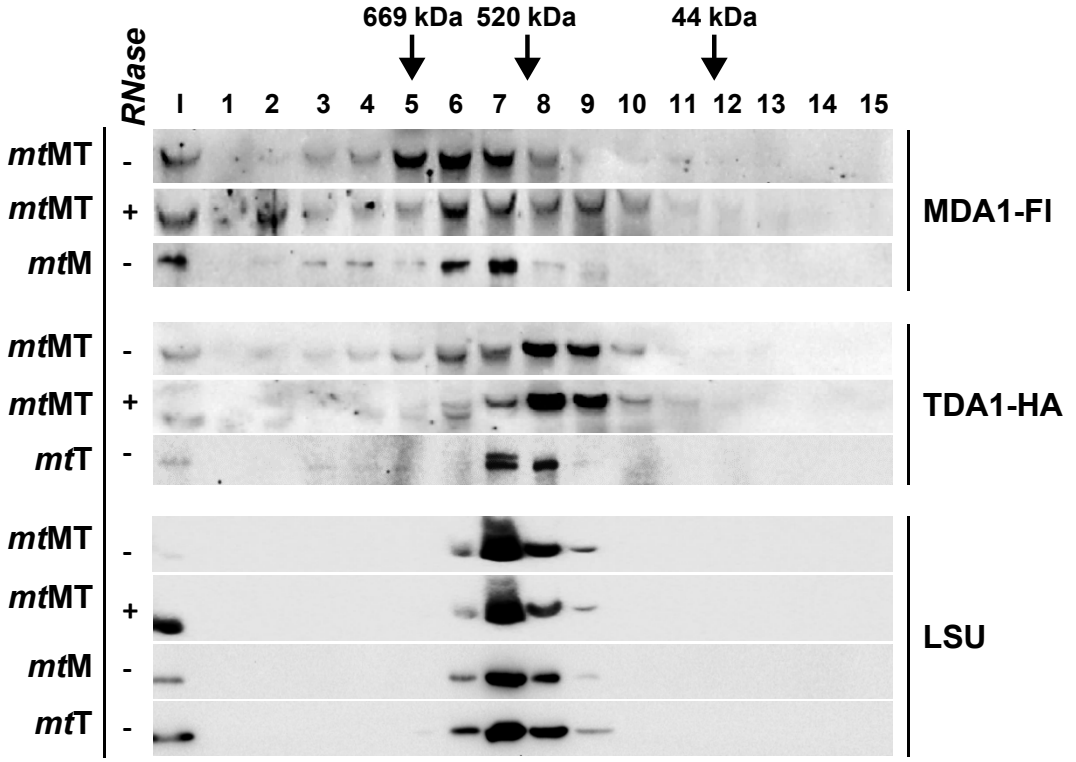


Figure 4. Interaction of MDA1, TDA1 and *atpA* mRNA probed by size-exclusion chromatography.
Soluble cell extract from strains *mt*MT, either untreated or RNase-treated, *mt*M or *mt*T, respectively expressing both MDA1-Fl and MDA1-HA, MDA1-Fl alone or MDA1-HA alone were fractionated on a Superose 6 column. After gel electrophoresis, fractions were immunodecorated with anti Flag or HA antibodies. Molecular masses of the complexes found in each fraction were estimated by comparison with protein standards of the MHW gel filtration calibration kit and with the elution profile of the RuBisCO holoenzyme (represented by the large subunit, LSU).

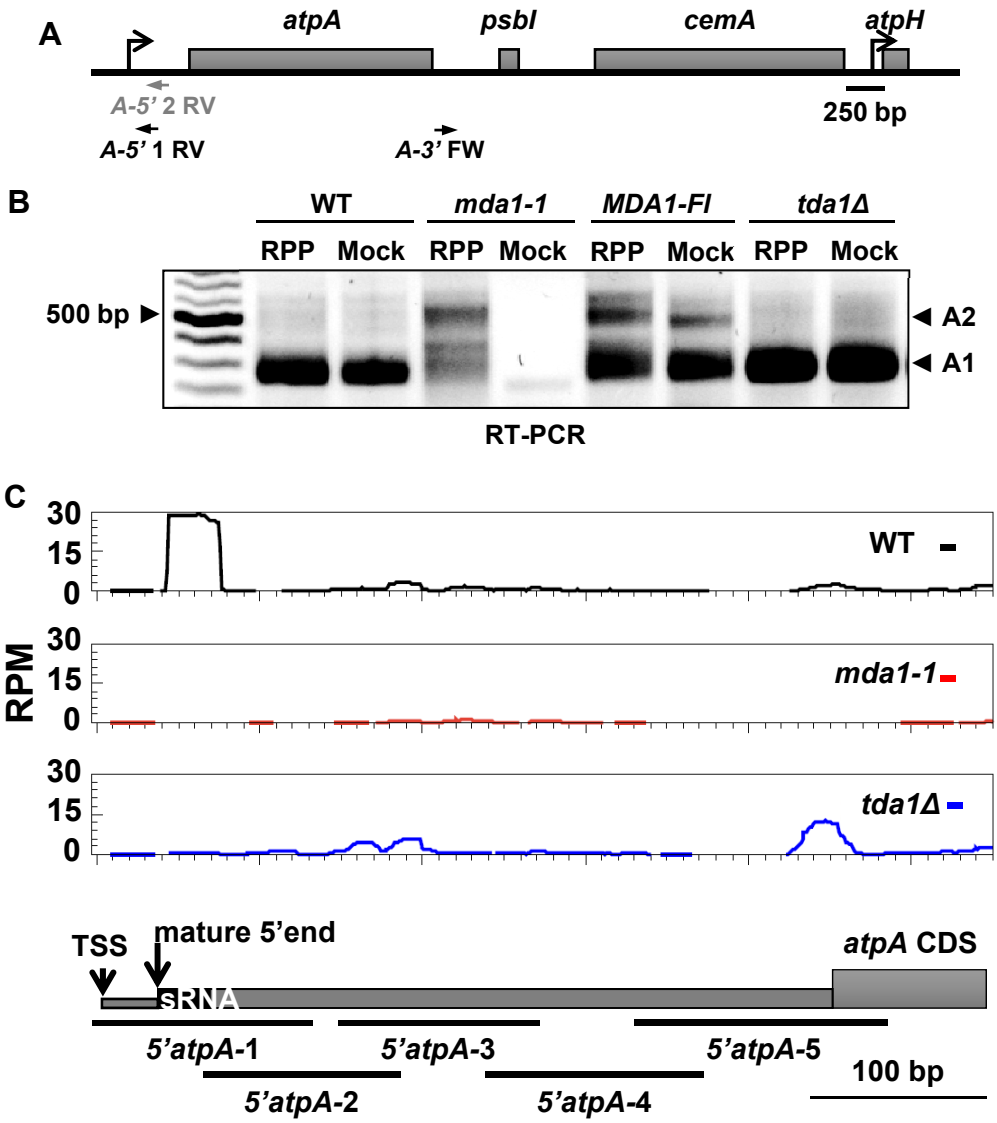


Figure 5. Mapping of *atpA* mRNA 5'ends and sRNA accumulation in presence or absence of MDA1 and TDA1.

A) Schematic representation of the *atpA-psbI-cemA-atpH* gene cluster. Coding sequences are represented by grey boxes, promoters by bent arrows. Positions of the primers used for retro-transcription (*A-5'* 2 RV; grey) and cRT-PCR (*A-5'* 1 RV and *A-3'* FW; black) are shown.

B) Agarose gel electrophoresis of the cRT-PCR amplicons from RPP- and mock-treated samples from the WT, *mda1-1*, *MDA1-FI* and *tda1Δ* strains.

C) Distribution of sRNAs in mutant strains *mda1-1* (red) and *tda1*Δ (blue) along the *atpA* 5'UTR compared to that in the wild-type (black). Coverage is normalized as Reads Per Million (RPM) and averaged over two biological replicates. The lower scheme depicts the *atpA* 5'UTR region (same convention than in Fig. 2D).

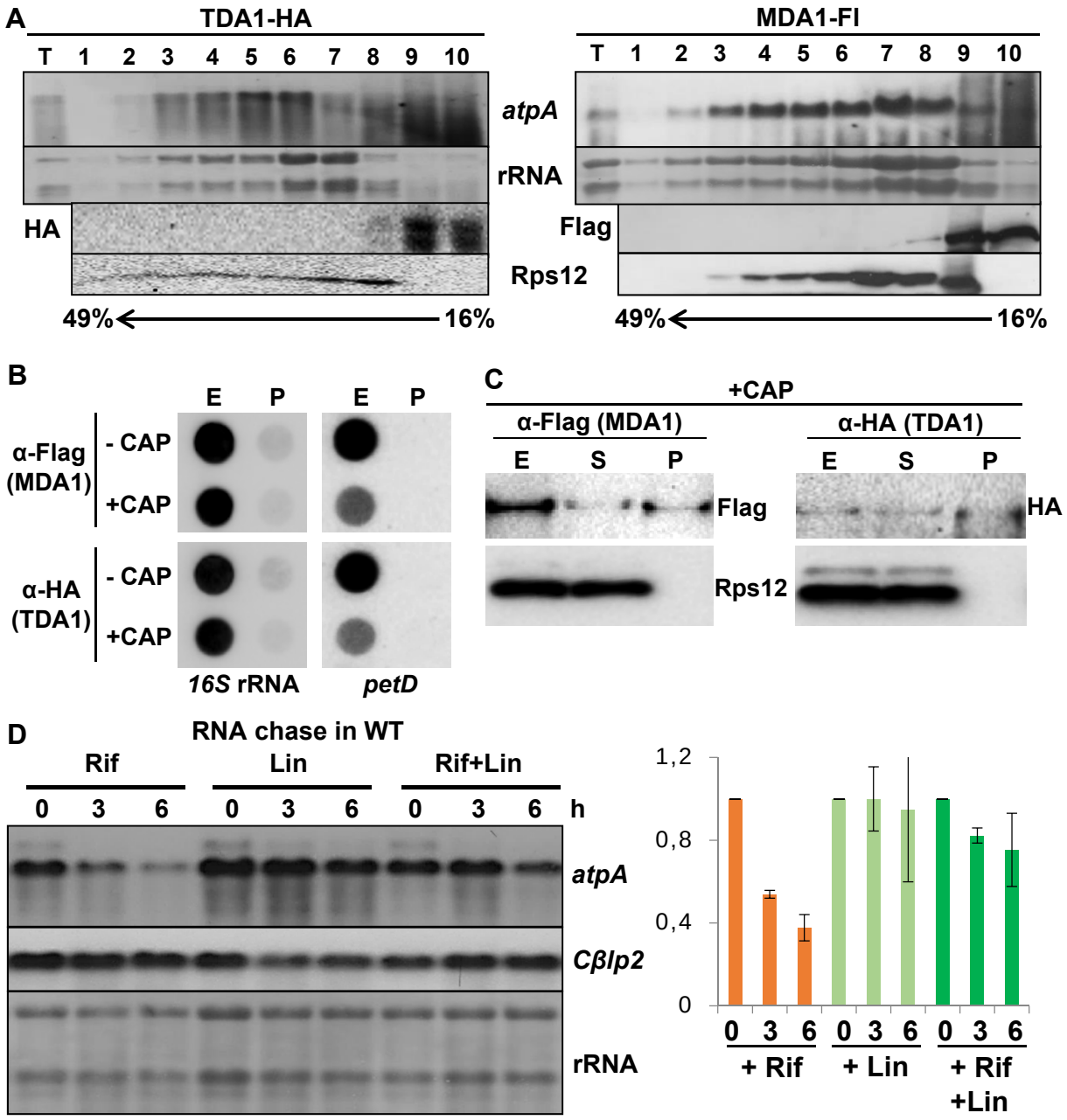


Figure 6. MDA1 and TDA1 are not associated with the *atpA* mRNAs engaged in translation, which are destabilized.
A) Solubilized whole-cell extracts (T) from *TDA1-HA* and *MDA1-Fl* complemented strains collected in presence of CAP (100 µg ml^{-1}) were loaded on sucrose gradients. After ultracentrifugation, ten fractions were collected and used for both RNA and protein extraction. The distribution of *atpA* along the gradient was analysed by RNA blotting (ribosomal RNAs were detected by methylene blue staining of the filters). The distributions of TDA1, MDA1 and of ribosomal protein Rps12 were analysed by immunoblotting. Sucrose concentrations are indicated at the bottom.
B) RNA co-immuno-precipitated with TDA1-HA and MDA1-Fl in untreated and CAP-pre-treated *mt*MT samples analysed by dot-blot with *16S* rRNA and *petD* (control) probes.
C) MDA1, TDA1 and ribosomal protein Rps12 immunodetection in the input (E), supernatant (S) and pellet (P) fractions from the CAP-treated samples shown in B.
D) Wild-type cells were treated with rifampicin (Rif, 350 µg ml^{-1}), lincomycin (Lin, 500 µg ml^{-1}) or both (Rif+Lin). RNA was extracted just before or 3 and 6 hr after inhibitors addition. Accumulation of *atpA* and *Cβlp2* (nuclear control) were analysed by RNA blot, ribosomal RNAs were detected by methylene blue staining of the filters. For each condition, the percentage of *atpA* mRNA amounts relative to time 0 is shown on the right (n=3).

SUPPLEMENTAL DATA:

Suppl. Figure S1: Annotated sequence of the MDA1 protein.
A) Amino acid sequence of MDA1. The predicted chloroplast targeting peptide is written in red, OPR repeats are alternatively boxed or grey-highlighted. Intron locations (downward triangles) are also shown, while the site of insertion of the 3xFlag tag is highlighted in blue and boxed. The two lines below the sequence respectively depict the predicted secondary structure of the protein (H: α-helical region written in red and highlighted in yellow; E: extended strand; T: turn confirmation, C: the rest) and the predicted disorder (O: ordered; D: disordered). Prediction were done with the Scratch protein predictor suite (http://scratch.proteomics.ics.uci.edu/).
B) Alignment of the OPR repeats from *C. reinhardtii* MDA1 protein. Residues matching the OPR consensus are highlighted in grey.

Supplemental Figure S2: Conservation of MDA1 proteins among Chlamydomonadales algae.
MDA1 orthologues were retrieved from the GenBank database by BLAST searches, aligned with the MUSCLE software using default option followed by manual editing to improve the alignment. OPR repeats shared by the different proteins are shown on top of sequences. Additional species-specific OPR repeats are highlighted in grey. Residues conserved in more than half of the sequences are written in red, while conservative substitutions are written in blue. The residue after which the 3xtag is inserted is written in green, boxed and highlighted in blue.
Abbreviations of species names are as follows: *Csp: Chlamydomonas sphaeroides; Cas: Chlamydomonas asymetrica; Gpe: Gonium pectorale; Tso: Tetrabaena socialis; Vca: Volvox carteri; Esp: Eudorina sp.; Cde: Chlamydomonas debaryana; Cre: Chlamydomonas reinhardtii; Yun: Yamagishiella unicocca*
Accession numbers: *Cas*: BDDA01000178.1; *Csp*: BDDC01000013.1; *Gpe*: LSYV01000003.1; *Tso*: PGGS01000476.1; *Vca*: ACJH01001813.1; *Esp*: BDSJ01000002.1;*Cde*: BDDB01000001.1; *Yun*: BDSL01000019.1.

Suppl. Figure S3: Complementation of *mda1-1* with BAC 16F5.
Supplemental Figure S3. Complementation of *mda1-1* with BAC 16F5.
A) Schematic representation of the genomic region encompassed by the BAC clone 16F5.
B) Immunoblot analysis showing the accumulation of ATPase subunit α in the wild type, *tda1Δ*, *mda1-1* and in five phototrophic clones recovered after transformation of *mda1-1* with the BAC clone 16F5.
C) Multiplex PCR for the genotyping of the strains used in (B), except *tda1Δ*. Amplification with primers MDA1 1+4 generates a wild-type allele-specific product of 333 bp; amplification with primers MDA1 3+2 generates a mutated allele-specific product of 181 bp; amplification with primers MDA1 1+2 generates a non-allele-specific product of 479 bp, used as PCR internal control. N.C. represents the negative control (PCR reaction with no template). Primers and amplicons are schematized in the lower panel. The deletion present in *mda1-1* is shown as a box with obliquus lines.

Supplemental Figure S4: MDA1 binds the *atpA* 5'UTR in absence of TDA1.
A) Detection of MDA1-Fl in the input (E), supernatant (S) and pellet (P) samples in samples from the *mt*M strain expressing the tagged version of MDA1 in the absence of TDA1, immuno-

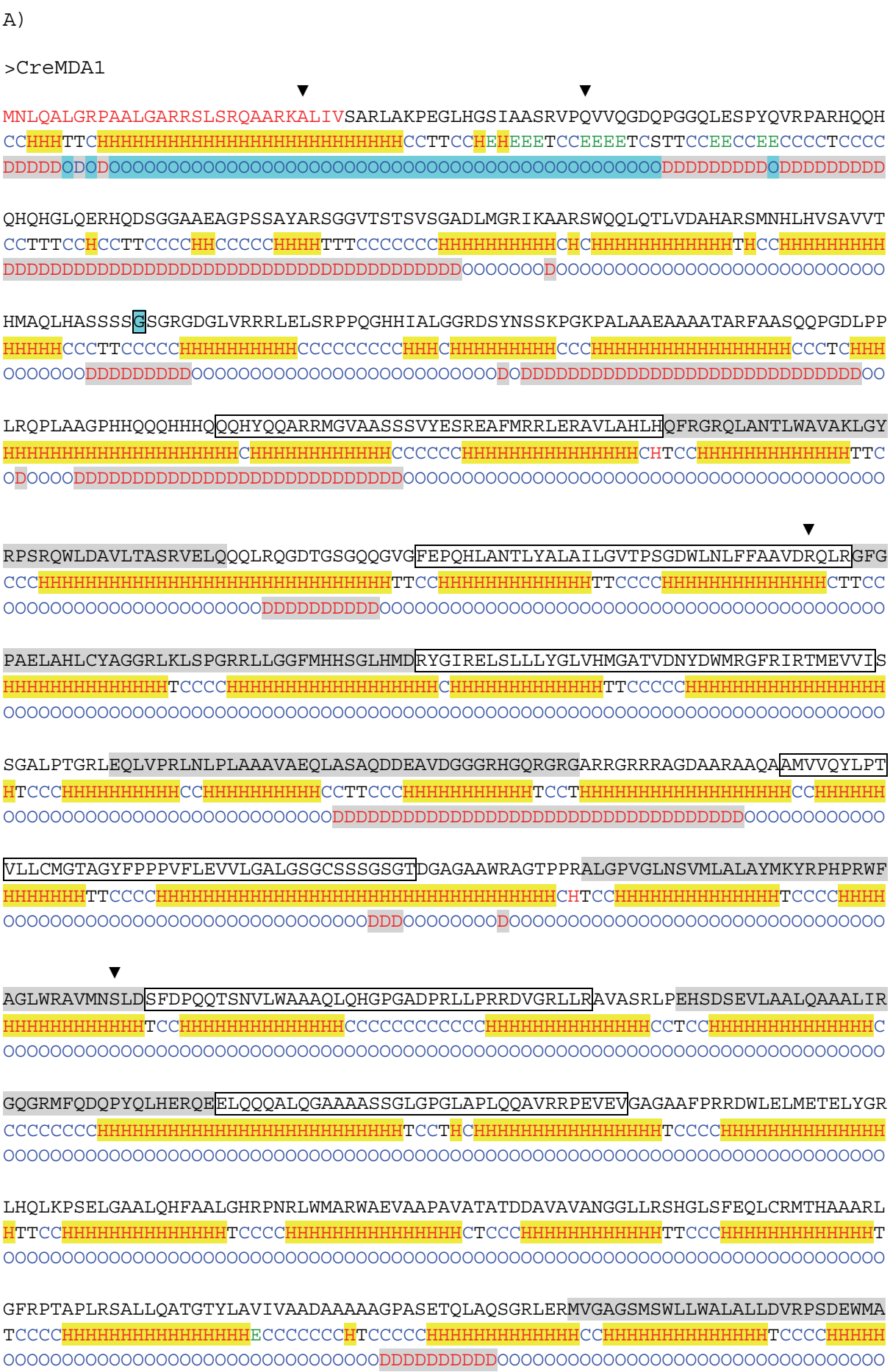
precipitated with anti-Flag magnetic beads. An antibody against ATP synthase subunit β was used as control
B) RNA immuno-precipitated in the same samples than in A), assessed by dot-blot, using the probes indicated at the bottom of the panel. The *petD* probe is used as negative control.

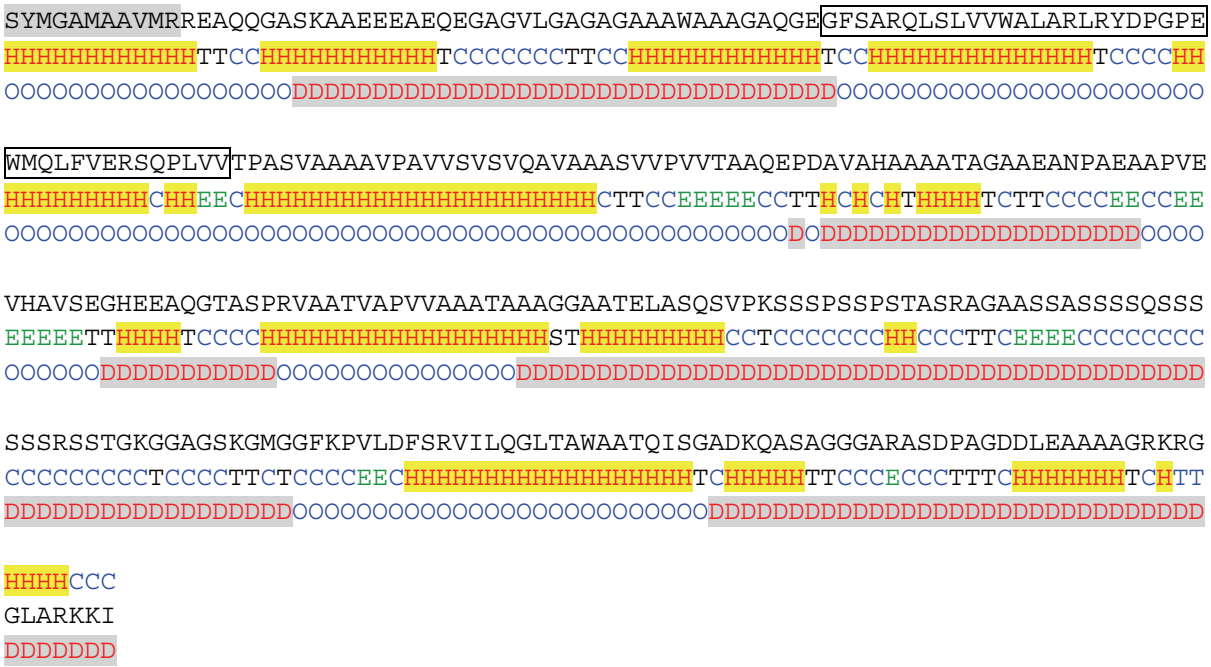
Supplemental Figure S5: Mapping of di- and tri-cistronic *atpA* mRNA 5'ends in presence or absence of MDA1.
A) Schematic representation of the *atpA-psbI-cemA-atpH* gene cluster. Coding sequences are represented by grey boxes, promoters by bent arrows. Positions of the primers used for retro-transcription (*A-5'* 2 RV; grey) and cRT-PCR (black) are shown.
B) cRT-PCR amplicons around the 5'*atpA*/3'*psbI* junction (primer *psbI*-3' FW) from RPP- and mock-treated samples from the WT, *mda1-1*, and *MDA1-Fl* strains.
C) cRT-PCR amplicons around the 5'*atpA*/3'*cemA* junction (primer *cemA*-3' FW) from RPP- and mock-treated samples from the WT, *mda1-1*, and *MDA1-Fl* strains.

Supplemental Figure S6: Stability of TDA1-Fl and MDA1-Fl.
Stability of TDA1-Fl (A) and MDA1-Fl (B), assessed by immuno-chase in the *TDA1-Fl* (A) and *MDA1-Fl* and *mt*M strains (B), at the indicated time point after addition of cycloheximide, an inhibitor of cytosolic translation.

Supplemental Figure S7: Mapping of the *mda1-1* mutation.
Proportion of S1D2 polymorphisms in the sequencing of pooled ac- (blue) and ac+ (red) progeny from the crossing *mda1-1* x S1D2 (see Materials and Methods), plotted along chromosome 16. The value 0 corresponds to a 100% S1D2 polymorphism. The arrow indicates the position of the 7 bp deletion identified in *mda1-1*.

A)



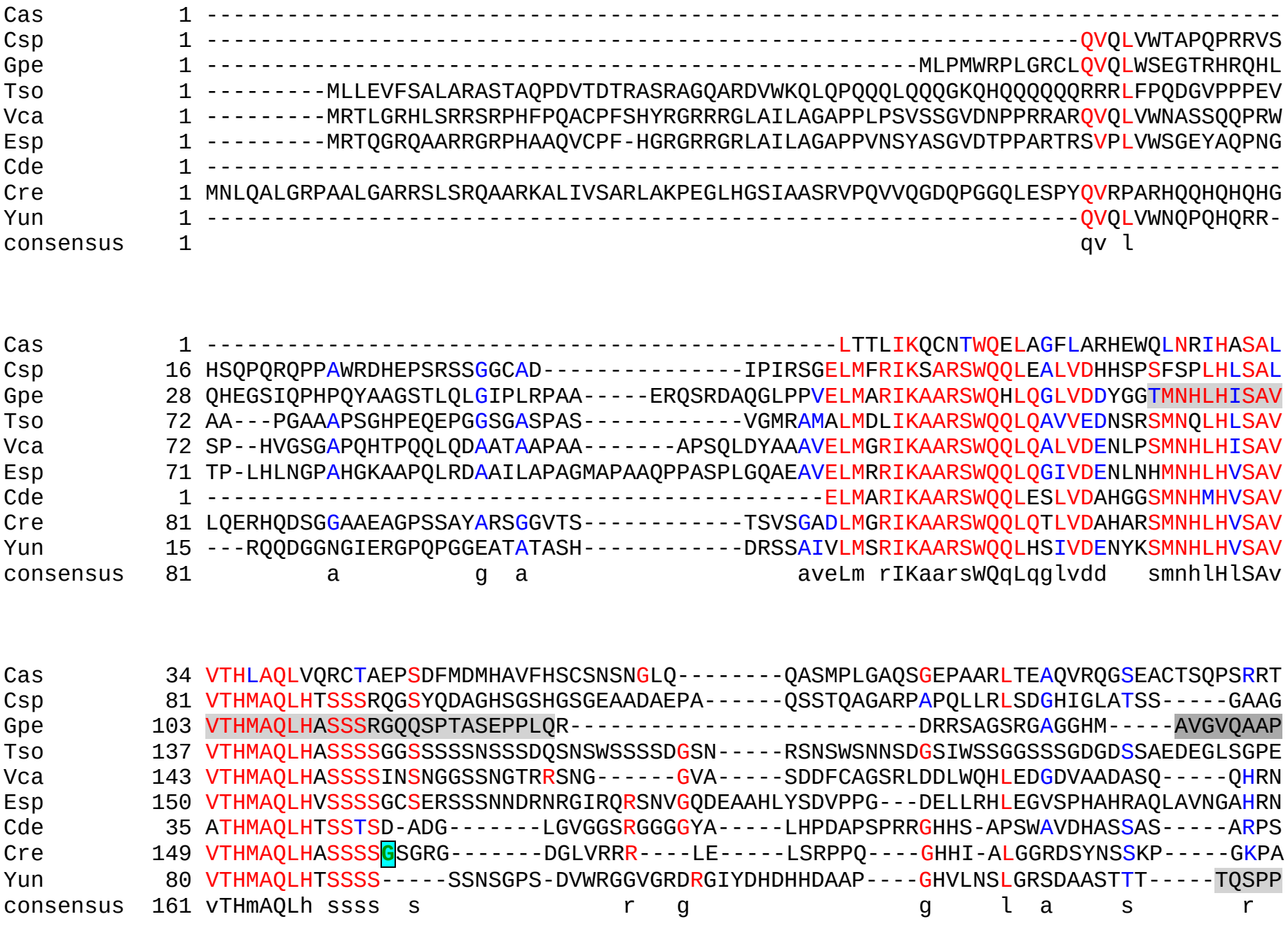


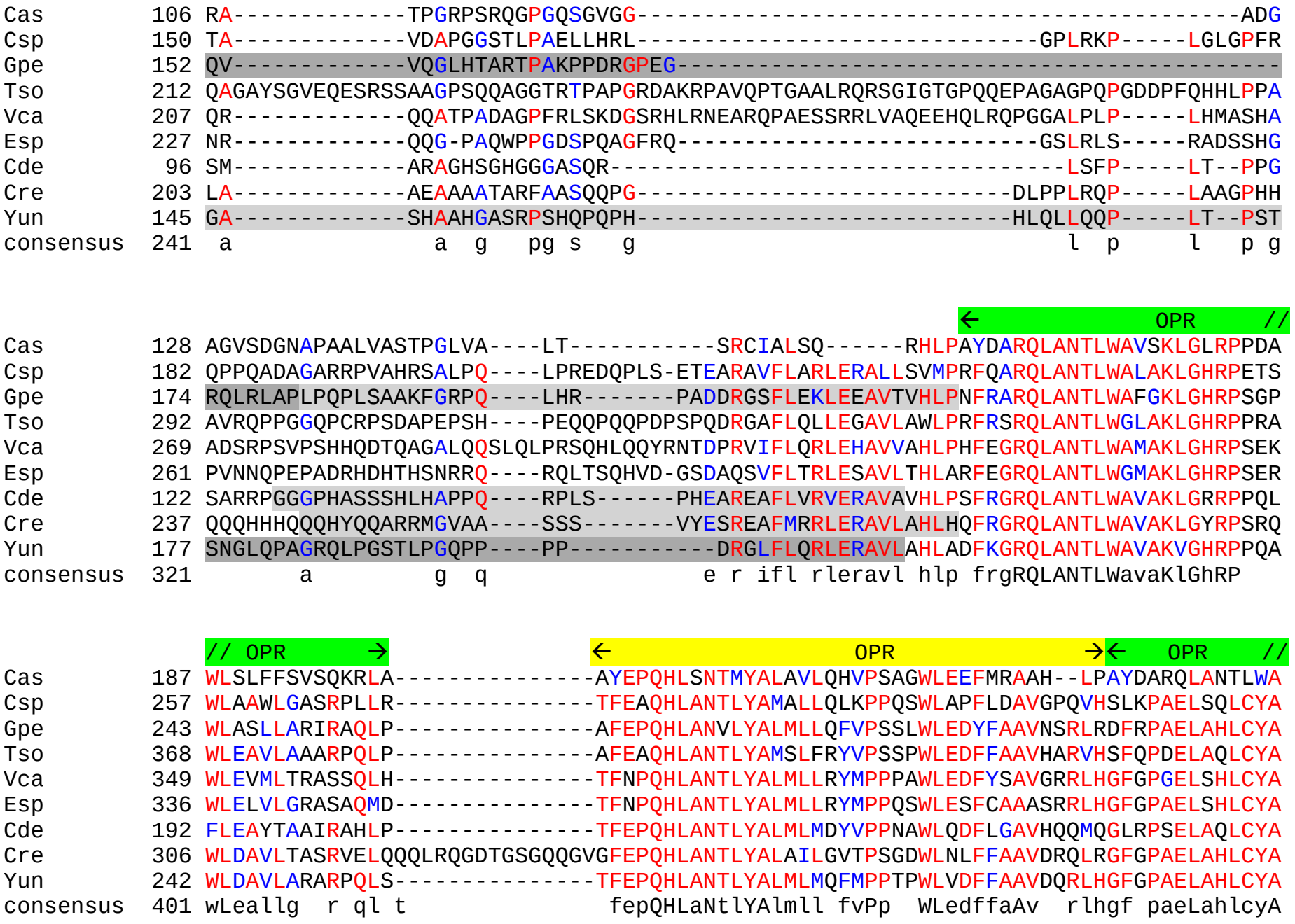
▼: position of introns; OPR repeats; G: tag insertion point

B)

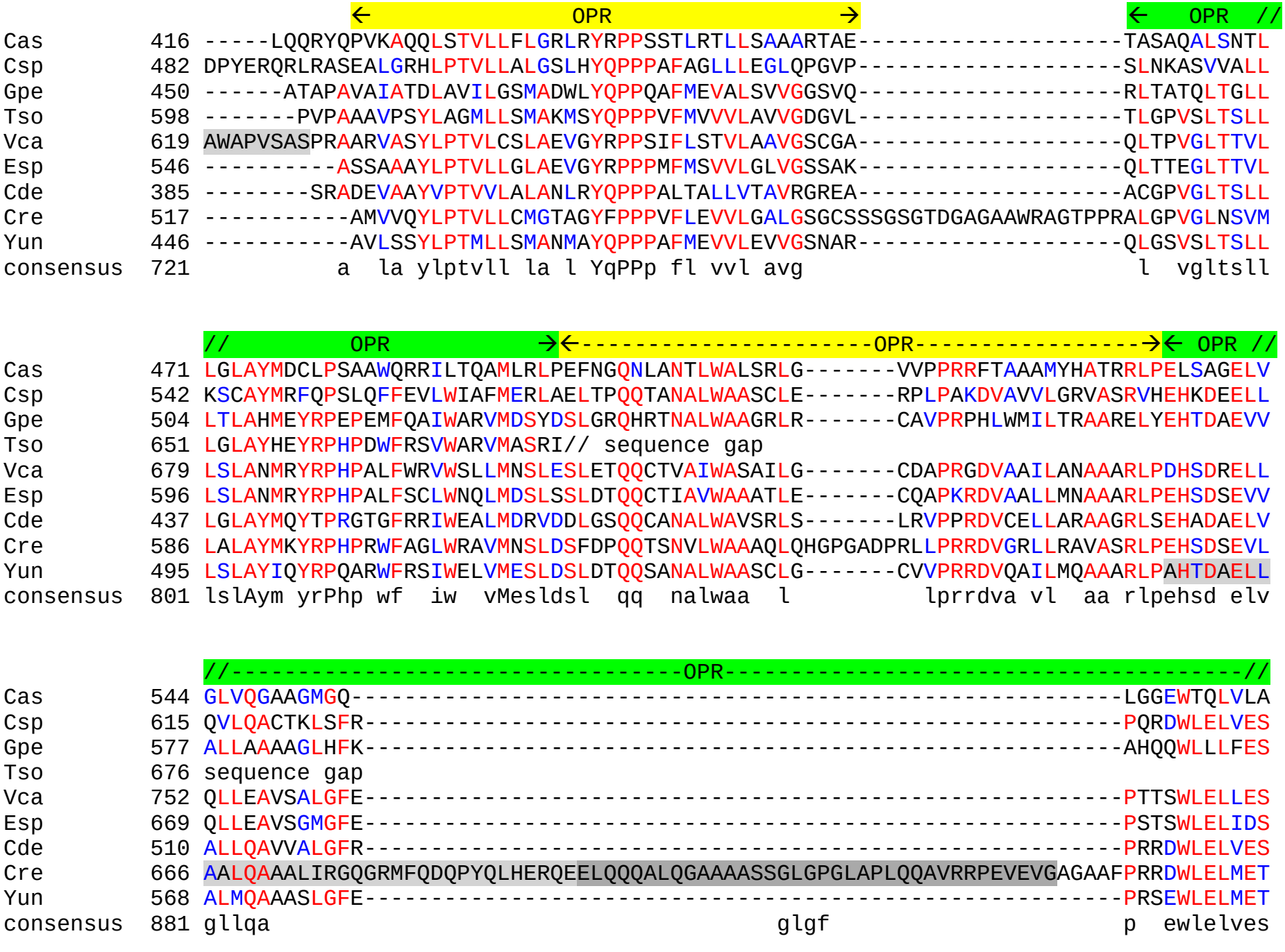
```
QQHYQQARRMGVAASSSVYeS-REAFMRRLERAVLAHLH
QFRGRQLANTLWAVAKLGYrP-SRQWLDAVLTASRVELQ
GFEPQHLANTLYALAILGVtP-SGDWLNLFFAAVDRQLR
GFGPAELAHLCYAGGRLKLsP-GRRLLGGFMHHSGLHMD
RYGIRELSLLLYGLVHMGAtvDNYDWMRGFRIRTMEVVI
RLNLPLAAAVAEQLASAQ----DDEAVDGGGRHGQRGRG
AMVVQYLPTVLLCMGTAGYfP-PPVFLEVVLGALGSGCS
ALGPVGLNSVMLALAYMKYrP-HPRWFAGLWRAVMNSLD
SFDPQQTSNVLWAAAQLQHgP-GADPRLLPRRDVGRLLR
EHSDSEVLAALQAAALIRGQ--GRMFQDQPYQLHERQEE
MVGAGSMSWLLWALALLDVrP-SDEWMASYMGAMAAVMR
GFSARQLSLVVWALARLRYdP-GPEWMQLFVERSQPLVV
AAVPAVVSVSVQAVAAASVvP-VVTAAQEPDAVAHAAAA
```

Supplemental Figure S1. Annotated sequence of the MDA1 protein.
A) Amino acid sequence of MDA1. The predicted chloroplast targeting peptide is written in red, OPR repeats are alternatively boxed or grey-highlighted. Intron locations (downward triangles) are also shown, while the site of insertion of the 3xFlag tag is highlighted in blue and boxed. The two lines below the sequence respectively depict the predicted secondary structure of the protein (H: α-helical region written in red and highlighted in yellow; E: extended strand; T: turn confirmation, C: the rest) and the predicted disorder (O: ordered; D: disordered). Prediction were done with the Scratch protein predictor suite (http://scratch.proteomics.ics.uci.edu/).
B) Alignment of the OPR repeats from *C. reinhardtii* MDA1 protein. Residues matching the OPR consensus are highlighted in grey.

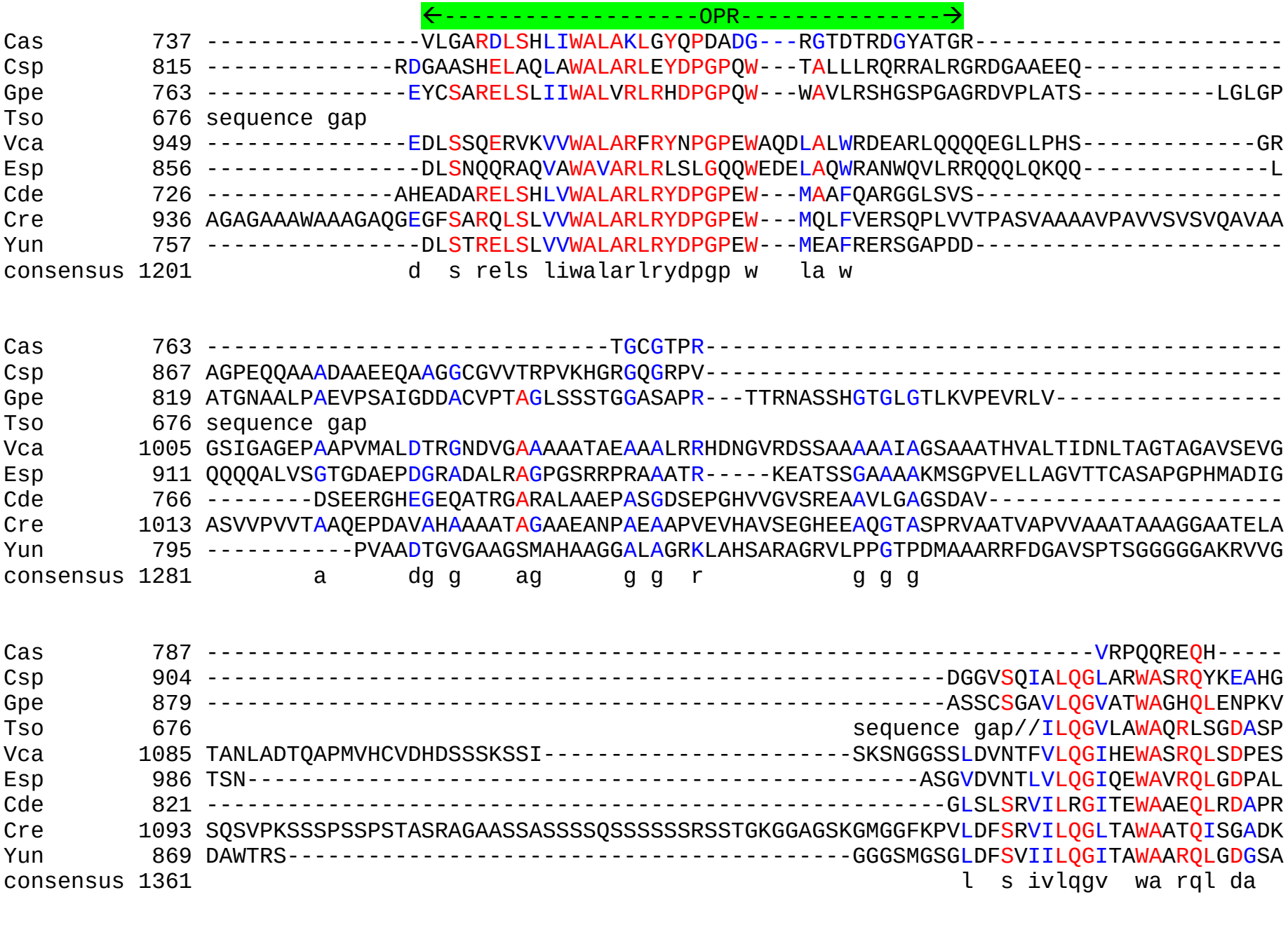

Cas 1 ----------QVQLVWTAPQPRRVS
Csp 1 ----------MLPMWRPLGRCLQVQLWSEGTRHRQHL
Gpe 1 ---------MRTLGRHLSRRSRPHFPQACPFSHYRGRRRGLAILAGAPPLPSVSSGVDNPPRRARQVQLVWNASSQQPRW
Tso 1 ---------MRTQGRQAARRGRPHAAQVCPF-HGRGRRGRLAILAGAPPVNSYASGVDTPPARTRSVPLVWSGEYAQPNG
Vca 1 ----------QVQLVWNQPQHQRR-
consensus 1 qv l
Cas 1 ----------LTTLIKQCNTWQELAGFLARHEWQLNRIHASAL
Csp 16 HSQPQRQPPAWRDHEPSRSSGGCAD----------IPIRSGELMFRIKSARSWQQLEALVDHHSPSFSPLHLSAL
Gpe 28 QHEGSIQPHPQYAAGSTLQLGIPLRPAA-----ERQSRDAQGLPPVELMARIKAARSWQHLQGLVDDYGGTMNHLHISAV
Tso 72 AA---PGAAAPSGHPEQEPGGSGASPAS----------VGMRAMALMDLIKAARSWQQLQAVVEDNSRSMNQLHLSAV
Vca 72 SP--HVGSGAPQHTPQQLQDAATAAPAA-------APSQLDYAAAVELMGRIKAARSWQQLQALVDENLPSMNHLHISAV
Esp 71 TP-LHLNGPAHGKAAPQLRDAAILAPAGMAPAAQPPASPLGQAEAVELMRRIKAARSWQQLQGIVDENLNHMNHLHVSAV
Cde 1 ----------ELMARIKAARSWQQLESLVDAHGGSMNHMHVSAV
Cre 81 LQERHQDSGGAAEAGPSSAYARSGGVTS----------TSVSGADLMGRIKAARSWQQLQTLVDAHARSMNHLHVSAV
Yun 15 ---RQQDGGNGIERGPQPGGEATATASH----------DRSSAIVLMSRIKAARSWQQLHSIVDENYKSMNHLHVSAV
consensus 81 a g a aveLm rIKaarsWQqLqglvdd smnhlHlSAv
Cas 34 VTHLAQLVQRCTAEPSDFMDMHAVFHSCSNSNGLQ-------QASMPLGAQSGEPAARLTEAQVRQGSEACTSQPSRRT
Csp 81 VTHMAQLHTSSSRQGSYQDAGHSGSHGSGEAADAEPA------QSSTQAGARPAPQLLRLSDGHIGLATSS-----GAAG
Gpe 103 VTHMAQLHASSSRGQQSPTASEPPLQR----------DRRSAGSRGAGGHM-----AVGVQAAP
Tso 137 VTHMAQLHASSSSGGSSSSSNSSSDQSNSWSSSSDGSN-----RSNSWSNNSDGSIWSSGGSSSGDGDSSAEDEGLSGPE
Vca 143 VTHMAQLHASSSSINSNGGSSNGTRRSNG------GVA-----SDDFCAGSRLDDLWQHLEDGDVAADASQ-----QHRN
Esp 150 VTHMAQLHVSSSSGCSERSSSNNDRNRGIRQRSNVGQDEAAHLYSDVPPG---DELLRHLEGVSPHAHRAQLAVNGAHRN
Cde 35 ATHMAQLHTSSTSD-ADG-------LGVGGSRGGGGYA-----LHPDAPSPRRGHHS-APSWAVDHASSAS-----ARPS
Cre 149 VTHMAQLHASSSSGSGRG-------DGLVRRR----LE-----LSRPPQ----GHHI-ALGGRDSYNSSKP-----GKPA
Yun 80 VTHMAQLHTSSSS-----SSNSGPS-DVWRGGVGRDRGIYDHDHHDAAP----GHVLNSLGRSDAASTTT-----TQSPP
consensus 161 vTHmAQLh ssss s r g g l a s r

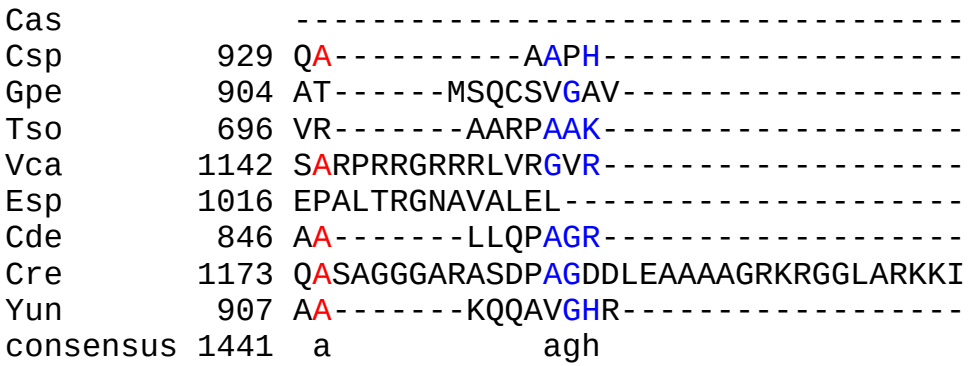

Cas 106 RA----------TPGRPSRQGPGQSGVGG----------ADG
Csp 150 TA----------VDAPGGSTLPAELLHRL----------GPLRKP-----LGLGPFR
Gpe 152 QV----------VQGLHTARTPAKPPDRGPEG----------
Tso 212 QAGAYSGVEQESRSSAAGPSQQAGGTRTPAPGRDAKRPAVQPTGAALRQRSGIGTGPQQEPAGAGPQPGDDPFQHHLPPA
Vca 207 QR----------QQATPADAGPFRLSKDGSRHLRNEARQPAESSRRLVAQEEHQLRQPGGALPLP-----LHMASHA
Esp 227 NR----------QQG-PAQWPPGDSPQAGFRQ----------GSLRLS-----RADSSHG
Cde 96 SM----------ARAGHSGHGGGASQR----------LSFP-----LT--PPG
Cre 203 LA----------AEAAAATARFAASQQPG----------DLPPLRQP-----LAAGPHH
Yun 145 GA----------SHAAHGASRPSHQPQPH----------HLQLLQQP-----LT--PST
consensus 241 a a g pg s g l p l p g
← OPR //
Cas 128 AGVSDGNAPAALVASTPGLVA----LT----------SRCIALSQ------RHLPAYDARQLANTLWAVSKLGLRPPDA
Csp 182 QPPQADAGARRPVAHRSALPQ----LPREDQPLS-ETEARAVFLARLERALLSVMPRFQARQLANTLWALAKLGHRPETS
Gpe 174 RQLRLAPLPQPLSAAKFGRPQ----LHR-------PADDRGSFLEKLEEAVTVHLPNFRARQLANTLWAFGKLGHRPSGP
Tso 292 AVRQPPGGQPCRPSDAPEPSH----PEQQPQQPDPSPQDRGAFLQLLEGAVLAWLPRFRSRQLANTLWGLAKLGHRPPRA
Vca 269 ADSRPSVPSHHQDTQAGALQQSLQLPRSQHLQQYRNTDPRVIFLQRLEHAVVAHLPHFEGRQLANTLWAMAKLGHRPSEK
Esp 261 PVNNQPEPADRHDHTHSNRRQ----RQLTSQHVD-GSDAQSVFLTRLESAVLTHLARFEGRQLANTLWGMAKLGHRPSER
Cde 122 SARRPGGGPHASSSHLHAPPQ----RPLS------PHEAREAFLVRVERAVAVHLPSFRGRQLANTLWAVAKLGRRPPQL
Cre 237 QQQHHHQQQHYQQARRMGVAA----SSS-------VYESREAFMRRLERAVLAHLHQFRGRQLANTLWAVAKLGYRPSRQ
Yun 177 SNGLQPAGRQLPGSTLPGQPP----PP----------DRGLFLQRLERAVLAHLADFKGRQLANTLWAVAKVGHRPPQA
consensus 321 a g q e r ifl rleravl hlp frgRQLANTLWavaKlGhRP
// OPR → ← OPR →← OPR //
Cas 187 WLSLFFSVSQKRLA----------AYEPQHLSNTMYALAVLQHVPSAGWLEEFMRAAH--LPAYDARQLANTLWA
Csp 257 WLAAWLGASRPLLR----------TFEAQHLANTLYAMALLQLKPPQSWLAPFLDAVGPQVHSLKPAELSQLCYA
Gpe 243 WLASLLARIRAQLP----------AFEPQHLANVLYALMLLQFVPSSLWLEDYFAAVNSRLRDFRPAELAHLCYA
Tso 368 WLEAVLAAARPQLP----------AFEAQHLANTLYAMSLFRYVPSSPWLEDFFAAVHARVHSFQPDELAQLCYA
Vca 349 WLEVMLTRASSQLH----------TFNPQHLANTLYALMLLRYMPPPAWLEDFYSAVGRRLHGFGPGELSHLCYA
Esp 336 WLELVLGRASAQMD----------TFNPQHLANTLYALMLLRYMPPQSWLESFCAAASRRLHGFGPAELSHLCYA
Cde 192 FLEAYTAAIRAHLP----------TFEPQHLANTLYALMLMDYVPPNAWLQDFLGAVHQQMQGLRPSELAQLCYA
Cre 306 WLDAVLTASRVELQQQLRQGDTGSGQQGVGFEPQHLANTLYALAILGVTPSGDWLNLFFAAVDRQLRGFGPAELAHLCYA
Yun 242 WLDAVLARARPQLS----------TFEPQHLANTLYALMLMQFMPPTPWLVDFFAAVDQRLHGFGPAELAHLCYA
consensus 401 wLeallg r ql t fepQHLaNtlYAlmll fvPp WLedffaAv rlhgf paeLahlcyA

```
                 //      OPR           →←                    OPR                 →
Cas        250 VSKLGLRPPDAWLSLFFSVSQKRL-AAYEQHLSNTMYALAVL---QHV-PSAGWLEEFMRAAHNAAPGMSPQELSNLVYA
Csp        322 AGRLGLTLPQQLVDQLTQHAGVHASRYGLRELALLLHGVTRV---GAE-LRPTWMRTIRIRISTLVL------------Q
Gpe        308 AGKLRVAPHSSLLGGIFRHTSINFRQYGMRELALLLYGLVHM---GGE-LNPDWVRQYRVRVCALLV------------E
Tso        433 AGRLGLQPRGEQLAAIRLHATLNMHRYGMSELVLLLHGIVRMGDPGPPRHQQVWMRAYRKRVLSLLI------------E
Vca        414 AGKLGLQPGRELLGGVLHHSLMHMQAYGVRELALLAYGVVHM---GAR-LHADWLRVYRKRVTGLVI------------E
Esp        401 AGKLGLRPDRGLLGGVLHHSAMHIEAYGMRELALLLYGVVNM---GAV-LPSDWLRVYRKRVCVLVL------------Q
Cde        257 AGKMRLQPSRALLGHAMMHSAVHMDRYGIRELALLLHGLVHM---GAS-LSHSWLRQYRIRLAAVVI------------E
Cre        386 GGRLKLSPGRRLLGGFMHHSGLHMDRYGIRELSLLLYGLVHM---GATVDNYDWMRGFRIRTMEVVI------------S
Yun        307 AGKLRLQPGRKVLGGVLTHSMVHMESYGIRELTLLLYGVVHM---GRS-HHQAWMRRYRMRVTSLLI------------K
consensus  481 agklgl p r llggll hs vhm  YglreLalllyglvhm   ga  l   Wlr yri            rv  lvi

Cas        325 LGKLSVQPTRSTCELLLRLTLHMLSER---RVNPQ-EASIVAHGFAGMGVR---PPPQWLECFCQSFADLVRRTDTQLSS
Csp        386 SGALSAK------------ALRMLGPS----LNRC-VVPLAIEQLAMLRRRWGADVEAANAAATRAAAEEAEGLPGRPSA
Gpe        372 SEVVPPP------------TLQRWGPH----LHPNALGPLVKARL-REVPQ---DSSRGDVDAAQSSSGEPSPLGPLAGS
Tso        501 WGLVRRP------------LLASVAPR----LNKQ-LRQLAAEHLGAGKRQ-----PQLRETQQQPQQQPQQQQQRQQSS
Vca        478 SGALPAR------------FLERLAEQLGQQQQLD-LAELARAQL-KMRQQ-----QLSQQQYPPGQHQQYHIHHTTSEQ
Esp        465 TGTLPVQ------------LLEKLAAA-ERQVHPE-LAEAARAQL-ALRLS------EQQPQHQPRLLRPGRVSAPTAGK
Cde        321 SGAVPPA------------LLQRLLPH----LNAG-LAAAAEAQL-----------------ARQSAAPGGSTEKRRRRG
Cre        451 SGALPTG------------RLEQLVPR----LNLP-LAAVAEQL----------------ASAQDDEAVDGGGRHGQRGR
Yun        371 AEALSPR------------VLERLMPR----LNKP-VAELANAHL-RGRQQ-----QQQQLRQHQTLEGAEQQQHFQDEW
consensus  561 sgalp               L rlgpr    ln    la la aql   r               q

Cas        398 SRGTGVD-----------------------------NMAERDQDKIG---------------------------------
Csp        449 RDESGPD-----------------------------GSGQAKQEQDR-----------------HRWLQLQDRELRRRQ
Gpe        432 SEASGER-----------------------------LAQETTRAGGG---------------------------------
Tso        559 QAGEGTD--------GATDAGGGVREAGAAAPVAGGGGGGVRPFSPA---------------------------------
Vca        539 HHEQQHQQELGLMEHGSSHVWPSVTRGDGLSGGTGPGAASTSSTNPESLLPPSFSSSPVDNSIDSIMATTISVATVAPAA
Esp        524 PAAEGHA-------------------------LNHGGGDSAVQNPE---------------------------------
Cde        367 RSRQGAE-----------------------------GSSQAADDDGG---------------------------------
Cre        499 GARRGRR-----------------------------RAGDAARAAQA---------------------------------
Yun        428 QQRTLAD-----------------------------GSSAAPRTSEA---------------------------------
consensus  641     g d                             gaa       g
```

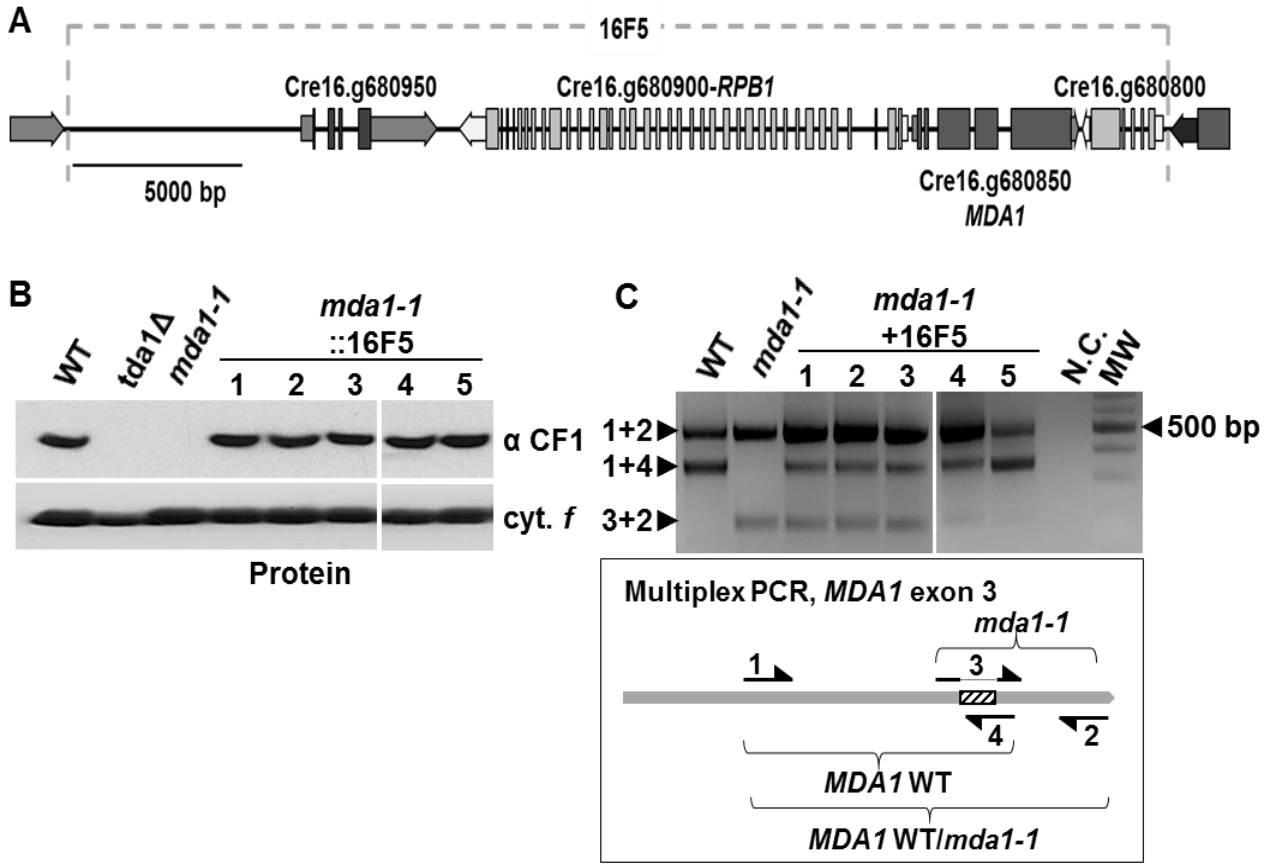

← OPR →
← OPR //
Cas 416 -----LQQRYQPVKAQQLSTVLLFLGRLRYRPPSSTLRTLLSAAARTAE----------TASAQALSNTL
Csp 482 DPYERQRLRASEALGRHLPTVLLALGSLHYQPPPAFAGLLLEGLQPGVP----------SLNKASVVALL
Gpe 450 ------ATAPAVAIATDLAVILGSMADWLYQPPQAFMEVALSVVGGSVQ----------RLTATQLTGLL
Tso 598 -------PVPAAAVPSYLAGMLLSMAKMSYQPPPVFMVVVLAVVGDGVL----------TLGPVSLTSLL
Vca 619 AWAPVSASPRAARVASYLPTVLCSLAEVGYRPPSIFLSTVLAAVGSCGA----------QLTPVGLTTVL
Esp 546 ----------ASSAAAYLPTVLLGLAEVGYRPPPMFMSVVLGLVGSSAK----------QLTTEGLTTVL
Cde 385 --------SRADEVAAYVPTVVLALANLRYQPPPALTALLVTAVRGREA----------ACGPVGLTSLL
Cre 517 ----------AMVVQYLPTVLLCMGTAGYFPPPVFLEVVLGALGSGCSSSGSGTDGAGAAWRAGTPPRALGPVGLNSVM
Yun 446 ----------AVLSSYLPTMLLSMANMAYQPPPAFMEVVLEVVGSNAR----------QLGSVSLTSLL
consensus 721 a la ylptvll la l YqPPp fl vvl avg l vgltsll
// OPR →←----------OPR----------→← OPR //
Cas 471 LGLAYMDCLPSAAWQRRILTQAMLRLPEFNGQNLANTLWALSRLG------VVPPRRFTAAAMYHATRRLPELSAGELV
Csp 542 KSCAYMRFQPSLQFFEVLWIAFMERLAELTPQQTANALWAASCLE------RPLPAKDVAVVLGRVASRVHEHKDEELL
Gpe 504 LTLAHMEYRPEPEMFQAIWARVMDSYDSLGRQHRTNALWAAGRLR------CAVPRPHLWMILTRAARELYEHTDAEVV
Tso 651 LGLAYHEYRPHPDWFRSVWARVMASRI// sequence gap
Vca 679 LSLANMRYRPHPALFWRVWSLLMNSLESLETQQCTVAIWASAILG------CDAPRGDVAAILANAAARLPDHSDRELL
Esp 596 LSLANMRYRPHPALFSCLWNQLMDSLSSLDTQQCTIAVWAAATLE------CQAPKRDVAALLMNAAARLPEHSDSEVV
Cde 437 LGLAYMQYTPRGTGFRRIWEALMDRVDDLGSQQCANALWAVSRLS------LRVPPRDVCELLARAAGRLSEHADAELV
Cre 586 LALAYMKYRPHPRWFAGLWRAVMNSLDSFDPQQTSNVLWAAAQLQHGPGADPRLLPRRDVGRLLRAVASRLPEHSDSEVL
Yun 495 LSLAYIQYRPQARWFRSIWELVMESLDSLDTQQSANALWAASCLG------CVVPRRDVQAILMQAAARLPAHTDAELL
consensus 801 lslAym yrPhp wf iw vMesldsl qq nalwaa l lprrdva vl aa rlpehsd elv
//----------OPR----------//
Cas 544 GLVQGAAGMGQ----------LGGEWTQLVLA
Csp 615 QVLQACTKLSFR----------PQRDWLELVES
Gpe 577 ALLAAAAGLHFK----------AHQQWLLLFES
Tso 676 sequence gap
Vca 752 QLLEAVSALGFE----------PTTSWLELLES
Esp 669 QLLEAVSGMGFE----------PSTSWLELIDS
Cde 510 ALLQAVVALGFR----------PRRDWLELVES
Cre 666 AALQAAALIRGQGRMFQDQPYQLHERQEELQQQALQGAAAASSGLGPGLAPLQQAVRRPEVEVGAGAAFPRRDWLELMET
Yun 568 ALMQAAASLGFE----------PRSEWLELMET
consensus 881 gllqa glgf p ewlelves

```
               //--->←--------------OPR----------------→                  ←------OPR-----→
Cas        566 QMFPQLEQLSAGSLATVLAALARSRQRPQAPWLARALRHMQPQLG------------------TATGTALTGAVWALAVM
Csp        638 VLYDRLPDMRPCDVLAVLCHLGGVRHRPPVPWLRRWAEVAVPSLR------------------SADLQQLAALSRACALL
Gpe        600 HLFKRLPSFKPFELVSVLQSLVDLGHKPDRIWLLRWAEALKPGVA------------------ELGLTHISSAAYAVARL
Tso        676 sequence gap
Vca        775 DLYNRLPRMEPPQVAALLLPLAGLGHKPHRLWMARWAEAMAAGLP------------------ELGLRHMAAAAYGAARL
Esp        692 DLYNRLPRMRPDELAQLLLPLANLGHKPHRLWMARWSEAILPGLR------------------ALSLRHLATAACGAARL
Cde        533 YLYDRLPRMQPYELNSVLVCMASLGHQPPRVWMARWSEAALPKLS------------------HLGLQQLSALVHAATKL
Cre        746 ELYGRLHQLKPSELGAALQHFAALGHRPNRLWMARWAEVAAPAVATATDDAVAVANGGLLRSHGLSFEQLCRMTHAAARL
Yun        591 HLYNRLPHMRPFEAAAALHRMADLGHKPHRVWLARWSETLLAGVA------------------ELALQDVAVAGYAAARL
consensus  961  ly rlp mrp ela vl  la lghrp riwlarw e m pgl                    l l  l  a  aaarl

               --------OPR--------→
Cas        628 AYQPSVAWRQAFLSAVDAHISAT---LA-TAPHTPSPAVTRRTTQLKQHPATA------QAGGGRSP-----AKRQNEGV
Csp        700 GARPPARMREALLQVSGERLAAI---AAAQHGDRFKPDGERQPAVPAEAERNA------GAEPPRTKSLFVLEREASPSR
Gpe        662 GFRPPAGLRAELLRVSAQRMAAA---PAASREGVELRSTSGQQSSEQMNAAAR------GTSPSRTK-------------
Tso        676 sequence gap
Vca        837 GFRPPSRLCGELLAATAAELGST------ATAAVEATAAQQQQQQDGGKQQQL------SRLGGRQGDEPKLPHRARSAW
Esp        754 GFRPPPQPRSELLSAAAAGLAAA---LSGGAAAATSAPPPPQLRE----DQPA------GYETGQVG-------RVRPAW
Cde        595 GFRPPTRLRRELVRAAGERLAAVAAFGAEGVEEEAAAAAGVEVREPLELPAPRFGAREGALSPGQKPQQQPLIQGRAAQP
Cre        826 GFRPTAPLRSALLQATGTYLAVI---VAADAAAAAGPASETQLAQSGRLER-----------------------------
Yun        653 GYRPPASLCSQLVSVAAERMVIVEDQTAASADARTAAAGEPQQQE-QEVEVPL------GLQPARLC-------------
consensus 1041 gfrpp  lr  ll a g hlaa     aa   g    a   q                 g   gr

                                 ←-------------OPR---------------→
Cas        693 HRPA-------PPPAHGLLLRPCDYAVVLWSLAALDVVPPLEWTSRFLAAVQSSLP------------------------
Csp        771 RPRG-------PPKLLGGEWSLATLRDLLVSTAVLDIRPDGAWMASYMSCLAVAL-------------------------
Gpe        720 -LRK-------GVDNGGEFQRIRVLSRLLWSLAMLDIRPDSEWLAAYLGCLGASL-------------------------
Tso        676 sequence gap
Vca        905 ERRH-------GDRAHSSRHLPETLSRLLWALAVLDIRPDAGWLSSYMECLGFMR-------------------------
Esp        814 TRRL-------GDRSS--GQLPETLTRLLWALAVLDIRPDGAWVSCYLACLRTVVG------------------------
Cde        675 LRRQLRSAASSGDVVTPGELQPQTLSRLLWSLAMLDVSPDGAWMAGYMAALRVSMA------------------------
Cre        879 ------------------MVGAGSMSWLLWALALLDVRPSDEWMASYMGAMAAVMRREAQQGASKAAEEEAEQEGAGVLG
Yun        713 RLRN-------RDGVAGARCRCETLSRLLWALAMLDVRPGGDWMSSYMACLGSSLE------------------------
consensus 1121   r             g      tlsrllw lavldvrpdg wm  ymacl
```

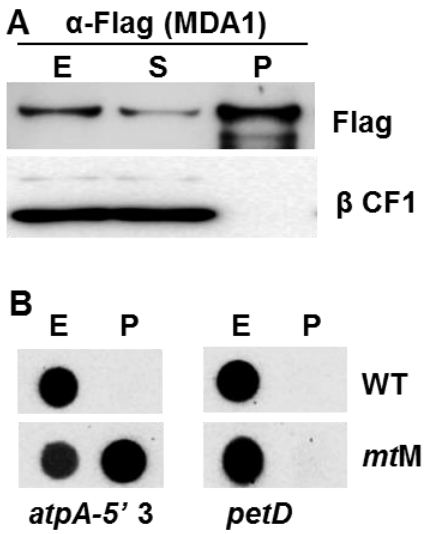

←--------------------OPR---------------→
Cas 737 ----------------VLGARDLSHLIWALAKLGYQPDADG---RGTDTRDGYATGR----------
Csp 815 --------------RDGAASHELAQLAWALARLEYDPGPQW---TALLLRQRRALRGRDGAAEEQ----------
Gpe 763 ----------------EYCSARELSLIIWALVRLRHDPGPQW---WAVLRSHGSPGAGRDVPLATS----------LGLGP
Tso 676 sequence gap
Vca 949 ----------------EDLSSQERVKVVWALARFRYNPGPEWAQDLALWRDEARLQQQQEGLLPHS----------GR
Esp 856 ----------------DLSNQQRAQVAWAVARLRLSLGQQWEDELAQWRANWQVLRRQQQLQKQQ----------L
Cde 726 --------------AHEADARELSHLVWALARLRYDPGPEW---MAAFQARGGLSVS----------
Cre 936 AGAGAAAWAAAGAQGEGFSARQLSLVVWALARLRYDPGPEW---MQLFVERSQPLVVTPASVAAAAVPAVVSVSVQAVAA
Yun 757 ----------------DLSTRELSLVVWALARLRYDPGPEW---MEAFRERSGAPDD----------
consensus 1201 d s rels liwalarlrydpgp w la w
Cas 763 ----------TGCGTPR----------
Csp 867 AGPEQQAAADAAEEQAAGGCGVVTRPVKHGRGQGRPV----------
Gpe 819 ATGNAALPAEVPSAIGDDACVPTAGLSSSTGGASAPR---TTRNASSHGTGLGTLKVPEVRLV----------
Tso 676 sequence gap
Vca 1005 GSIGAGEPAAPVMALDTRGNDVGAAAAATAEAAALRRHDNGVRDSSAAAAAIAGSAAATHVALTIDNLTAGTAGAVSEVG
Esp 911 QQQQALVSGTGDAEPDGRADALRAGPGSRRPRAAATR-----KEATSSGAAAAKMSGPVELLAGVTTCASAPGPHMADIG
Cde 766 --------DSEERGHEGEQATRGARALAAEPASGDSEPGHVVGVSREAAVLGAGSDAV----------
Cre 1013 ASVVPVVTAAQEPDAVAHAAAATAGAAEANPAEAAPVEVHAVSEGHEEAQGTASPRVAATVAPVVAAATAAAGGAATELA
Yun 795 -----------PVAADTGVGAAGSMAHAAGGALAGRKLAHSARAGRVLPPGTPDMAAARRFDGAVSPTSGGGGGAKRVVG
consensus 1281 a dg g ag g g r g g g
Cas 787 ----------VRPQQREQH-----
Csp 904 ----------DGGVSQIALQGLARWASRQYKEAHG
Gpe 879 ----------ASSCSGAVLQGVATWAGHQLENPKV
Tso 676 sequence gap//ILQGVLAWAQRLSGDASP
Vca 1085 TANLADTQAPMVHCVDHDSSSKSSI----------SKSNGGSSLDVNTFVLQGIHEWASRQLSDPES
Esp 986 TSN----------ASGVDVNTLVLQGIQEWAVRQLGDPAL
Cde 821 ----------GLSLSRVILRGITEWAAEQLRDAPR
Cre 1093 SQSVPKSSSPSSPSTASRAGAASSASSSSQSSSSSSRSSTGKGGAGSKGMGGFKPVLDFSRVILQGLTAWAATQISGADK
Yun 869 DAWTRS----------GGGSMGSGLDFSVIILQGITAWAARQLGDGSA
consensus 1361 l s ivlqgv wa rql da



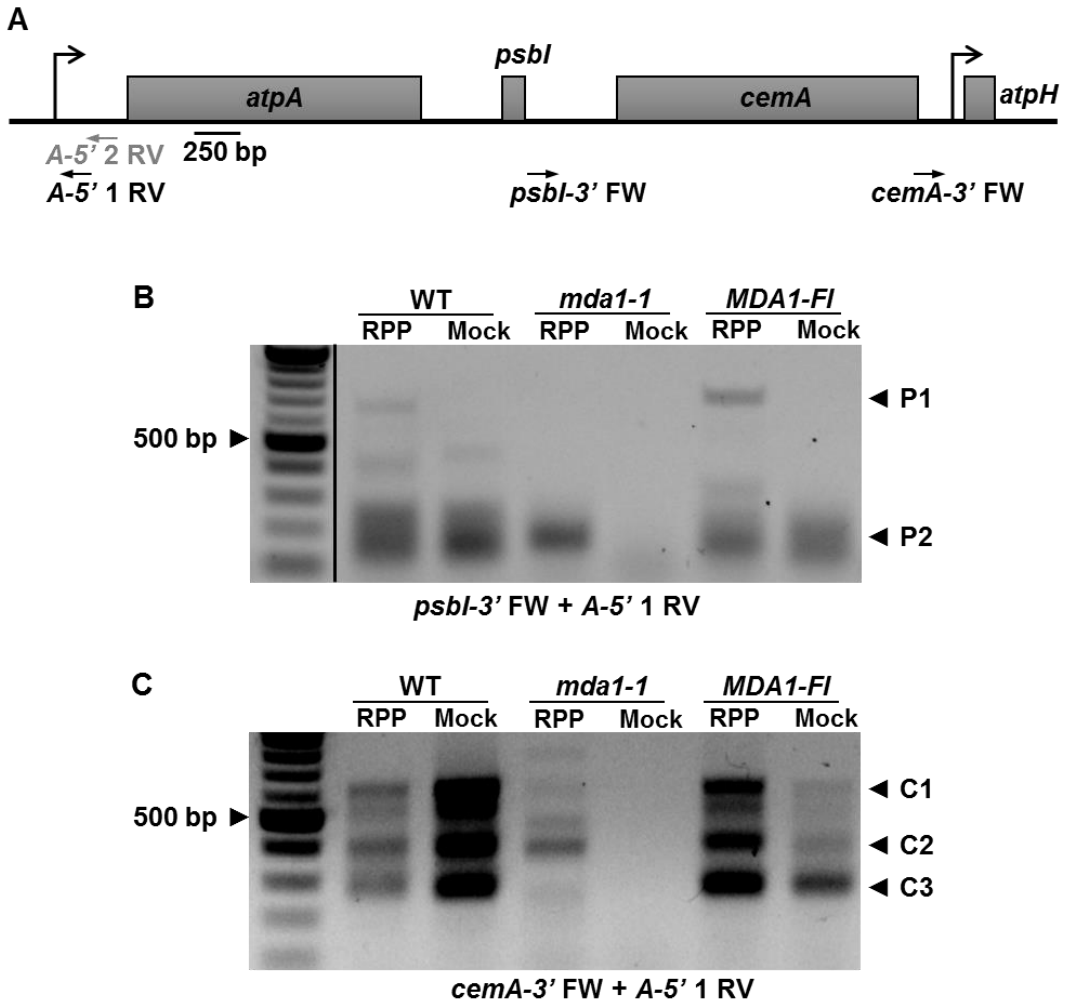
Supplemental Figure S2: Conservation of MDA1 among Chlamydomonadales.
MDA1 orthologues were retrieved from the GenBank database by BLAST searches, aligned with the MUSCLE software using default option followed by manual editing to improve the alignment. OPR repeats shared by the different proteins are shown on top of sequences. Additional species-specific OPR repeats are highlighted in grey. Residues conserved in more than half of the sequences are written in red, while conservative substitutions are written in blue. The residue after which the 3xtag is inserted is written in green, boxed and highlighted in blue.
Abbreviations of species names are as follows: *Csp: Chlamydomonas sphaeroides; Cas: Chlamydomonas asymetrica; Gpe: Gonium pectorale; Tso: Tetrabaena socialis; Vca: Volvox carteri; Esp: Eudorina sp.; Cde: Chlamydomonas debaryana; Cre: Chlamydomonas reinhardtii; Yun: Yamagishiella unicocca*
Accession numbers: *Cas*: BDDA01000178.1; *Csp*: BDDC01000013.1; *Gpe*: LSYV01000003.1; *Tso*: PGGS01000476.1; *Vca*: ACJH01001813.1; *Esp*: BDSJ01000002.1;*Cde*: BDDB01000001.1; *Yun*: BDSL01000019.1.



Supplemental Figure S3. Complementation of *mda1-1* with BAC 16F5.
A) Schematic representation of the genomic region encompassed by the BAC clone 16F5.
B) Immunoblot analysis showing the accumulation of ATPase subunit α in the wild type, *tda1Δ*, *mda1-1* and in five phototrophic clones recovered after transformation of *mda1-1* with the BAC clone 16F5.
C) Multiplex PCR for the genotyping of the strains used in (B), except *tda1Δ*. Amplification with primers MDA1 1+4 generates a wild-type allele-specific product of 333 bp; amplification with primers MDA1 3+2 generates a mutated allele-specific product of 181 bp; amplification with primers MDA1 1+2 generates a non-allele-specific product of 479 bp, used as PCR internal control. N.C. represents the negative control (PCR reaction with no template). Primers and amplicons are schematized in the lower panel. The deletion present in *mda1-1* is shown as a box with obliquus lines.



Supplemental Figure S4: MDA1 binds the *atpA* 5'UTR in absence of TDA1.
A) Detection of MDA1-Fl in the input (E), supernatant (S) and pellet (P) samples in samples from the *mt*M strain expressing the tagged version of MDA1 in the absence of TDA1, immuno-precipitated with anti-Flag magnetic beads. An antibody against ATP synthase subunit β was used as control
B) RNA immuno-precipitated in the same samples than in A), assessed by dot-blot, using the probes indicated at the bottom of the panel. The *petD* probe is used as negative control.

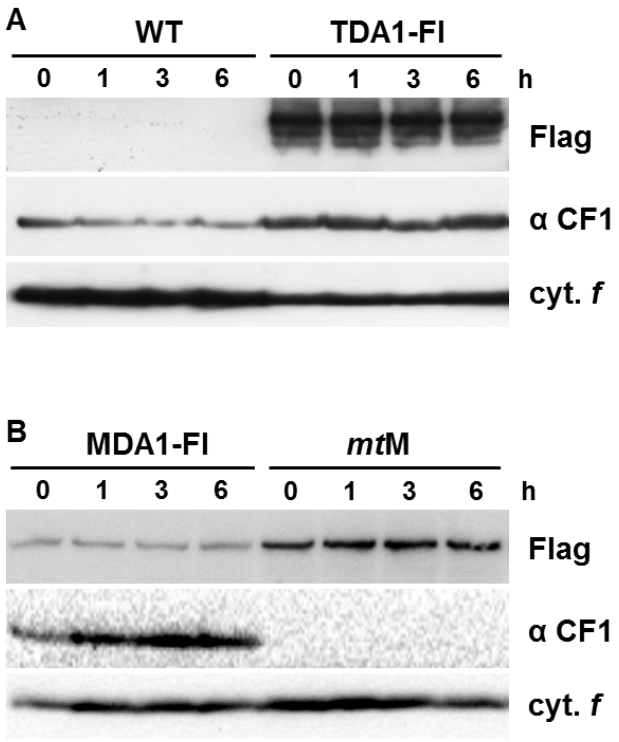


Supplemental Figure S5: Mapping of di- and tri-cistronic *atpA* mRNA 5'ends in presence or absence of MDA1.

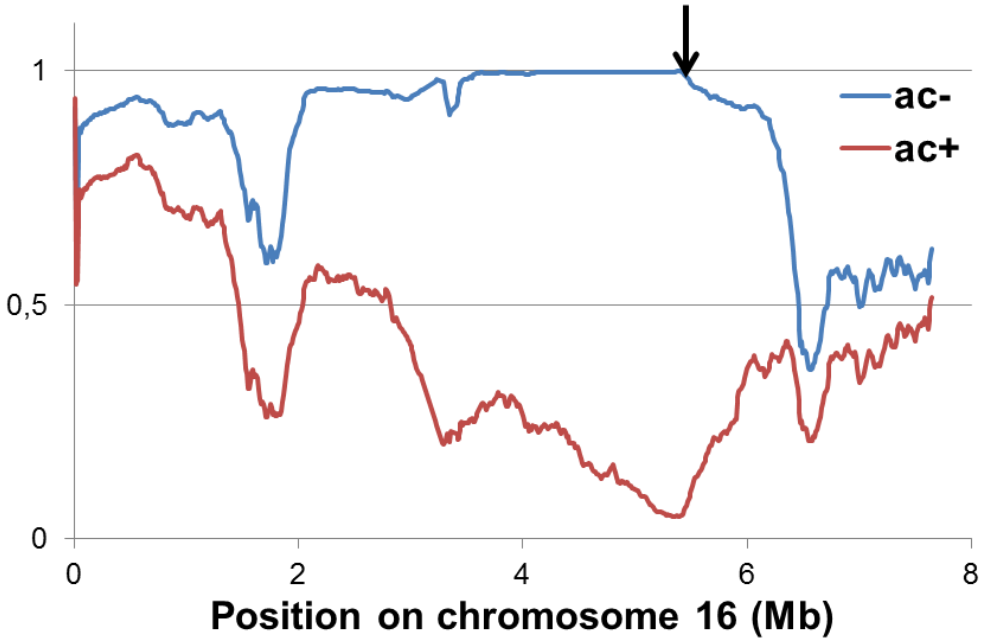
A) Schematic representation of the *atpA-psbI-cemA-atpH* gene cluster. Coding sequences are represented by grey boxes, promoters by bent arrows. Positions of the primers used for retro-transcription (*A-5'* 2 RV; grey) and cRT-PCR (black) are shown.

B) cRT-PCR amplicons around the 5'*atpA*/3'*psbI* junction (primer *psbI*-3' FW) from RPP- and mock-treated samples from the WT, *mda1-1*, and *MDA1-Fl* strains.

C) cRT-PCR amplicons around the 5'*atpA*/3'*cemA* junction (primer *cemA*-3' FW) from RPP- and mock-treated samples from the WT, *mda1-1*, and *MDA1-Fl* strains.



Supplemental Figure S6: Stability of TDA1-Fl and MDA1-Fl.
Stability of TDA1-Fl (A) and MDA1-Fl (B), assessed by immuno-chase in the *TDA1-Fl* (A) and *MDA1-Fl* and *mt*M strains (B), at the indicated time point after addition of cycloheximide, an inhibitor of cytosolic translation.



Supplemental Figure S7: Mapping of the *mda1-1* mutation.
Proportion of S1D2 polymorphisms in the sequencing of pooled ac- (blue) and ac+ (red) progeny from the crossing *mda1-1* x S1D2 (see Materials and Methods), plotted along chromosome 16. The value 0 corresponds to a 100% S1D2 polymorphism. The arrow indicates the position of the 7 bp deletion identified in *mda1-1*.

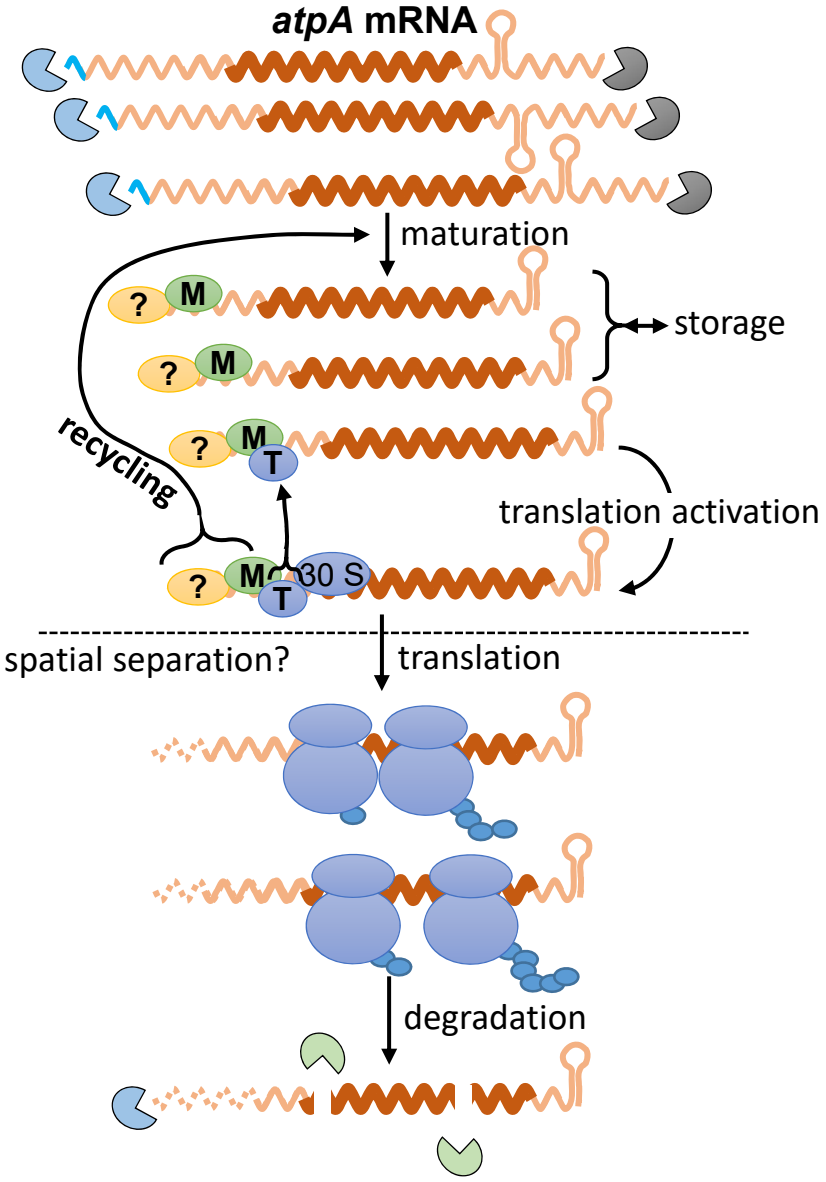


Figure 7: Working model for the post-transcriptional control of the *atpA* gene expression. *De novo* transcribed *atpA* mRNA molecules (represented in the monocistronic form for sake of simplicity), possessing a primary 5'end (in blue), are stabilized by the binding of MDA1 (M) to a region located in the middle of the 5' UTR. The binding of MDA1 is likely cooperative with that of another factor (question mark) that binds the 5'end of the transcript, protecting it from exonucleases and defining the maturation site. MDA1 can bind *atpA* in absence of TDA1 and its level determines the size of the stable *atpA* mRNA pool. TDA1 (T) interacts with the 5'UTR of a fraction of this pool in close proximity with MDA1. The binding of TDA1 activates the translation of the *atpA* mRNAs. When translation is initiated, MDA1 and TDA1 dissociate from the 5'UTR of ribosome-bound *atpA* mRNAs, leaving them unprotected. Upon translation initiation, the mRNAs might be spatially re-located inside the chloroplast. After or during translation, the mRNAs are degraded by nuclease and are not re-bound by MDA1, which is recycled to bind newly transcribed *atpA* mRNAs; TDA1 is probably also recycled.