

Article

Light-Dependent Translation Change of *Arabidopsis psbA* Correlates with RNA Structure Alterations at the Translation Initiation Region

Piotr Gawroński ^{1,†}, Christel Enroth ^{2,†}, Peter Kindgren ³, Sebastian Marquardt ³, Stanisław Karpinski ¹, Dario Leister ⁴, Poul Erik Jensen ⁵, Jeppe Vinther ² and Lars B. Scharff ^{3,*}

¹ Department of Plant Genetics, Breeding and Biotechnology, Institute of Biology, Warsaw University of Life Sciences, Nowoursynowska 159, 02-776 Warsaw, Poland; piotr_gawronski@sggw.pl (P.G.); stanislaw_karpinski@sggw.pl (S.G.)

² Department of Biology, Section for Computational and RNA Biology, University of Copenhagen, Ole Maaløes Vej 5, 2200 København N, Denmark; kroysler@hotmail.com (C.E.); jvinther@bio.ku.dk (J.V.)

³ Copenhagen Plant Science Centre, Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Denmark; peter.kindgren@slu.se (P.K.); sebastian.marquardt@plen.ku.dk (S.M.)

⁴ Plant Molecular Biology, Department Biology I, Ludwig-Maximilians-University Munich, Großhadernerstr. 2-4, 82152 Planegg-Martinsried, Germany; leister@lmu.de

⁵ Department of Food Science, University of Copenhagen, Rolighedsvej 26, 1958 Frederiksberg C, Denmark; peje@food.ku.dk

* Correspondence: scharff@plen.ku.dk

† These authors contributed equally to this work.

Citation: Gawroński, P.; Enroth, C.; Kindgren, P.; Marquardt, S.; Karpinski, S.; Leister, D.; Jensen, P.E.; Vinther, J.; Scharff, L.B. Light-dependent translation change of *Arabidopsis psbA* correlates with RNA structure alterations at the translation initiation region. *Cells* **2021**, *10*, 322. <https://doi.org/10.3390/cells10020322>

Academic Editor: Philippe Giegé

Received: 5 January 2021

Accepted: 29 January 2021

Published: 4 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Abstract: mRNA secondary structure influences translation. Proteins that modulate the mRNA secondary structure around the translation initiation region may regulate translation in plastids. To test this hypothesis, we exposed *Arabidopsis thaliana* to high light, which induces translation of *psbA* mRNA encoding the D1 subunit of photosystem II. We assayed translation by ribosome profiling and applied two complementary methods to analyze in vivo RNA secondary structure: DMS-MaPseq and SHAPE-seq. We detected increased accessibility of the translation initiation region of *psbA* after high light treatment, likely contributing to the observed increase in translation by facilitating translation initiation. Furthermore, we identified the footprint of a putative regulatory protein in the 5' UTR of *psbA* at a position where occlusion of the nucleotide sequence would cause the structure of the translation initiation region to open up, thereby facilitating ribosome access. Moreover, we show that other plastid genes with weak Shine-Dalgarno sequences (SD) are likely to exhibit *psbA*-like regulation, while those with strong SDs do not. This supports the idea that changes in mRNA secondary structure might represent a general mechanism for translational regulation of *psbA* and other plastid genes.

Keywords: chloroplast translation; regulation; mRNA secondary structure; RNA structure probing; high light; gene expression; plastid; *Arabidopsis thaliana*

1. Introduction

The secondary structure of mRNA is important for many translation related processes in bacteria and bacteria-derived eukaryotic organelles. This includes the efficiency of translation initiation [1–5], the recognition of start codons [6–8], and ribosome pausing [9–11]. In addition, changes in mRNA secondary structure can regulate translation initiation. Some of the mechanisms involved, such as riboswitches [12] and RNA thermometers [13,14], are independent of proteins, whereas others depend on the

binding of small RNAs or proteins to either activate or repress translation by modifying mRNA secondary structure [15,16].

It has been proposed that changes in mRNA secondary structure regulate translation in plastids, i.e., plant organelles derived from cyanobacteria [17–20]. This is not surprising, as bacterial-type 70S ribosomes synthesize proteins in plastids, and the process shows many similarities to translation in bacteria. Indeed, translational regulation is a major determinant of gene expression in plastids [21–24]. The intrinsic mRNA features that determine the efficiency of start codon recognition in plastids of higher plants, and hence, the efficiency of translation initiation, are well characterized: (a) Shine-Dalgarno sequences hybridize to the anti-Shine-Dalgarno sequence at the tail of the 16S rRNA and thereby position the start codon so that it can bind to the initiator tRNA; (b) local minima of mRNA secondary structure around the start codon make it accessible for the ribosome, whereas other AUGs are masked by folded RNA [6,8,25–28].

Compared to the intrinsic mRNA features determining the efficiency of translation initiation, we understand much less about the molecular mechanisms regulating translation in plastids. One hypothesis is based on *in vitro* findings that some RNA-binding proteins can alter the structure of the translation initiation region of their target mRNAs in a way which activates translation initiation [17–20]. In the absence of such a protein, the Shine-Dalgarno sequence and/or the start codon are occluded by mRNA secondary structure; therefore, translation efficiency is low. The binding of the regulatory protein shifts the structural equilibrium to an RNA conformation that makes these *cis*-elements, which are essential for translation initiation, accessible to the ribosome, and thereby activates or upregulates translation.

Here, we tested this hypothesis by exposing *Arabidopsis thaliana* plants to high light. In higher plants, this condition is known to induce the translation of the plastid-encoded *psbA* mRNA (encoding the D1 subunit of photosystem II) on the level of translation initiation [29,30]. This increase of *psbA* translation counteracts the increase in D1 turnover due to photodamage [31,32]. *psbA* translation is activated by specific RNA binding proteins [33–38]. Changes in the *psbA* mRNA *in vivo* secondary structure and translation efficiency were analyzed; our findings support the hypothesis that *psbA* is regulated by an RNA-binding protein that increases the accessibility of the Shine-Dalgarno sequence under high light conditions. Moreover, our analysis of the relationship between mRNA secondary structure and translation suggests that this mechanism is generally used to regulate translation of plastid-encoded genes that, like *psbA*, possess a weak Shine-Dalgarno sequence.

2. Materials and Methods

2.1. Plant Material

For DMS probing, *Arabidopsis thaliana* wild-type (ecotype Col-0) plants were grown in Jiffy pots (Jiffy Products, Zwijndrecht, Netherlands) for 17–18 days at 22 °C and 150 $\mu\text{E m}^{-2} \text{s}^{-1}$ in long-day conditions (16 h day/8 h night). Then, they were either kept for 1 h in dim light ($\sim 10 \mu\text{E m}^{-2} \text{s}^{-1}$, low light control) or shifted for the same time to 1000 $\mu\text{E m}^{-2} \text{s}^{-1}$ white light (high light treatment), supplied by an SL 3500-W-D LED lamp (Photon Systems Instruments, Drasov, Czech Republic). The young plants were DMS probed at noon. Plants treated in the same way were used for polysome analysis (Supplemental Figure S1).

For NAI-N₃ probing, *A. thaliana* plants were grown for seven weeks at 20 °C in short-day (8 h day/16 h night), low light conditions (140–160 $\mu\text{E m}^{-2} \text{s}^{-1}$). The low light sample was harvested at noon, while the high light sample was transferred at noon to the following conditions: 4 h high light [1200 $\mu\text{E m}^{-2} \text{s}^{-1}$]; 16 h dark; 4 h high light [1200 $\mu\text{E m}^{-2} \text{s}^{-1}$]; and then the leaf material was harvested. The temperature of the growing chamber was set to 20 °C, but owing to the heat emitted by the lamps, the leaves were exposed to temperatures of up to 30 °C. Young leaves with a maximum length of 20 mm were

harvested into liquid nitrogen (rosette diameter at this growth stage was 68 ± 3 mm). Plants treated the same way were used for ribosome profiling and RNA-seq (Figures 1G, 3H, 4 and Supplemental Figure S11A–D) as well as polysome analysis (Supplemental Figure S7).

2.2. Determining Photosynthetic Parameters

Chlorophyll *a* fluorescence parameters were measured in triplicate using a MAXI IMAGING-PAM M-series instrument (Walz, Effeltrich, Germany). Plants were dark-acclimated for 30 min. For F_0 and F_m determination, plants were exposed to a saturating pulse, followed by 5 min of blue (450 nm) actinic light ($81 \mu\text{E m}^{-2} \text{s}^{-1}$). In an actinic light phase, saturating light pulses were applied at 20-s intervals. Results were calculated for the last saturating pulse during the actinic light period. Maximum quantum yield of photosystem II (F_v/F_m) and electron transport rate (ETR) parameters were calculated as described previously [39].

2.3. Polysome Analysis

Polysome analysis using sucrose gradients for separation of free mRNA and polysome complexes was done as described previously [40]. The *psbA* and *rbcL* probes were amplified from total plant DNA using gene-specific primers (see Supplemental Table S4), radioactively labelled with $\alpha^{32}\text{P}$ [CTP] using the Megaprime DNA Labeling System (GE Healthcare Life Sciences, Chicago, IL, USA), and hybridized at 65°C .

2.4. Ribosome Profiling (Ribo-seq)

Ribosome profiling was done as described before [11,41,42]. Three biological replicates (each consisting of material from at least three plants) for each treatment were analyzed as follows: 400 mg of deep-frozen, ground leaf material was thawed on ice in 5 mL of extraction buffer (200 mM Tris/HCl pH 8.0, 200 mM KCl, 35 mM MgCl_2 , 0.2 M sucrose, 1% Triton X-100, 2% polyoxyethylen-10-tridecyl-ether, 5 mM dithiothreitol, 100 $\mu\text{g}/\text{mL}$ chloramphenicol, 50 $\mu\text{g}/\text{mL}$ cycloheximide). The extract was centrifuged for 5 min at 13,200 g and 4°C . 600 μL of the supernatant was removed for analysis by RNA-seq (subsection 2.5.), and the remaining supernatant was centrifuged for 10 min at 15,000 g and 4°C . CaCl_2 was added to the resulting supernatant to a concentration of 5 mM, followed by 750 units of micrococcal nuclease (Thermo Fisher Scientific, Waltham, MA, USA), and the mixture was incubated for 1 h at room temperature. The digested extract was loaded on a 2-mL sucrose cushion (40 mM Tris/acetate pH 8.0, 100 mM KCl, 15 mM MgCl_2 , 1 M sucrose, 5 mM dithiothreitol, 100 $\mu\text{g}/\text{mL}$ chloramphenicol, 50 $\mu\text{g}/\text{mL}$ cycloheximide) and centrifuged for 3 h at 55,000 g and 4°C in a Type 70 Ti rotor (Beckman Coulter, Brea, CA, USA). The pellet was dissolved in 1% SDS, 10 mM Tris/HCl pH 8.0, and 1 mM EDTA. RNA was purified using the PureLink miRNA Isolation Kit (Thermo Fisher Scientific). The 16 to 42-nt fraction was isolated by electrophoresis and treated with T4 polynucleotide kinase before library preparation using the TruSeq Small RNA Library Preparation Kit (Illumina, San Diego, CA, USA). Sequencing was performed on the HiSeq 4000 platform (Illumina).

2.5. RNA-seq

For each treatment, three biological replicates were analyzed. RNA was purified from 600 μL of leaf extract (subsection 2.4.) using the RNAsasy Plant Mini Kit (Qiagen, Hilden, Germany). The RNA was treated with the Ribo-Zero™ rRNA Removal Kit (Plant Leaf) (Illumina), libraries were prepared using the TruSeq RNA Library Prep Kit v2 (Illumina) and sequenced on the HiSeq 4000 platform (Illumina).

2.6. Determination of 3'-Ends of Plastid Transcripts

Determination of 3'-ends was done with the protocol described previously [43]. Briefly, a DNA linker (NEB, Ipswich, MA, USA) was ligated to free 3'-ends of 1 µg of denatured total RNA. After ligation, the RNA was fragmented in an alkaline solution (100 mM NaCO₃, 2 mM EDTA) for 30 min. The RNA was subsequently precipitated, dissolved, and loaded on a 15% TBE-Urea gel. Fragments in the size range of 50–150 bp were cut out and precipitated overnight. The fragmented RNA was then used to synthesize cDNA with Superscript III (Thermo Fisher Scientific) and a primer that annealed to the ligated linker. The cDNA was loaded on a 10% TBE-Urea gel, and products in the size range of 85–160 bp were cut out and precipitated. Products were then circularized with CircLigase (Epicentre, Madison, WI, USA) according to the manufacturer's instructions and used as a template for PCR amplification. The reaction included a primer with a barcode for sequencing purposes. Amplified PCR products were loaded on an 8% TBE gel and products around 150 bp were cut out. The sequencing library was run on a Bioanalyzer (Agilent, Santa Clara, CA, USA) to confirm that no contaminations from the library construction were present. Sequencing was done on a MiSeq platform (Illumina).

2.7. Processing of Ribo-seq and RNA-seq Reads

Arabidopsis thaliana (Col-0) genomic, transcriptomic, and noncoding RNA sequences, and the GFF3 annotation file were downloaded from Ensembl Plants (<http://plants.ensembl.org>, release 41). This annotation file lacks the annotation of the plastid transcripts. We added our own annotation using a manually curated data set. The 5' ends are based on primer extension data from the RNA secondary structure probing with NAI-N₃ (subsection 2.13. and 2.15.). The 3' ends are based on the 3'-end mapping set (subsection 2.6.). When there were multiple transcripts for one gene, the longest transcript detected was chosen for the annotation file. The sequences of coding regions were corrected for editing as detected by RNA-seq. Start codons and missing exons were corrected using GeSeq [44] plus corrections based on the ribosome profiling data. *rps16* was not spliced and was therefore characterized as a pseudogene as described previously [45]. From the downloaded transcriptome, plastid sequences were replaced with the new set.

Adapter sequences were removed using TrimGalore! (version 0.4.5; http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Alignments were performed with STAR (version 2.6.0a;) [46] with following settings: `–outFilterMismatchNmax 2–outMultimapperOrder Random–outSAMmultNmax 1–alignIntronMax 1–alignIntronMin 2` allowing for two mismatches, ungapped alignment on the transcriptome, and random assignment of reads that mapped to more than one location. Reads that mapped to noncoding RNAs were removed from the analysis. Unaligned reads were used as an input in an alignment to the transcriptome. Reads whose alignment length was between 28 nt and 40 nt and which mapped in a “sense” direction were used for further analysis. To assign each footprint to the P-site of the ribosome, we used the 5'-end of a mapped footprint and 23-nt offset as described previously [11]. Only reads whose P-site overlapped with coding regions (CDSs) were used for further analysis. When a P-site overlapped with more than one CDS (e.g., in partially overlapping *psbD-psbC*), the read was assigned randomly to one of the CDSs. RNA-seq reads were mapped to the transcriptome in a similar way, but reads with more than 5% mismatches were removed. Reads that mapped in both directions (unstranded library) and those that overlapped by at least 1 nt with CDSs were used for further analysis. Similarly to Ribo-seq, random assignment was used when a read overlapped with more than one CDS (e.g., *psbD-psbC*). Based on counts of reads mapped to the CDSs, RPKM (reads per kilobase per million mapped reads) values were calculated using normalization to the total number of mapped reads for each sample and the length of the CDS. For the analysis of footprints of putative regulatory proteins, reads with an aligned length between 18 nt and 40 nt were used.

2.8. Calculation of Translation Efficiency and Analysis of Differential Gene Expression

Translation efficiency (TE) was calculated using RiboDiff (version 0.2.1) [47] and counts of reads were mapped to the CDSs. Genes with $p < 0.01$ were considered to be significantly changed. For the differential gene expression analysis, RNA-seq reads were pseudo-aligned to the transcriptome using Salmon (version 0.9.1) [48] with default parameters. Transcript-level abundances were imported into R using tximport [49] and analyzed using the DESeq2 package [50].

2.9. Gel-Blot Analysis of Small RNAs

RNA was extracted from leaf material harvested in low light and high light (same material as used for ribosome profiling, RNA-seq, and RNA secondary structure probing with NAI-N₃) by adding 666 μ L of extraction buffer (subsection 2.4.) to frozen, ground material. The RNA was purified from the extract using a phenol/chloroform/isoamyl alcohol step and isopropanol precipitation. The gel blot was done as described before [51]: 10 μ g total RNA was separated on 15% polyacrylamide TBE urea gels (Biorad, Hercules, CA, USA) and transferred in a wet-blot setup with 0.5 TBE buffer to a Hybond-N membrane (GE-Healthcare Life Sciences). The RNA was cross-linked to the membrane with 0.16 M *N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride in 0.13 M 1-methylimidazole (pH 8.0) at 60 °C for 1 h. The probe (see Supplemental Table S4) was labelled at the 5' end with γ -³²P-ATP and hybridized at 60 °C using standard protocols.

2.10. RNA Structure Probing with DMS

Three biological replicates (each consisting of at least three plants) for each treatment were used. Young, 17–18 days old plants were collected into 10 mL of DMS reaction buffer (100 mM KCl, 40 mM HEPES pH 7.5, 0.5 mM MgCl₂). Dimethyl sulfate (DMS) was added to a concentration of 5% (*w/v*) and the reaction was performed at 24–25 °C (DMS+ samples). In parallel, negative control (DMS-) samples were prepared by adding water in place of DMS. The young plants were incubated for 6 min at either $\sim 10 \mu\text{E m}^{-2} \text{ s}^{-1}$ (low light control) or $1000 \mu\text{E m}^{-2} \text{ s}^{-1}$ (high light treatment) while the solution was held horizontally and hand mixed. The high light treatment caused a 1 °C temperature increase in the reaction buffer. The reaction was stopped by adding 20 mL of ice-cold 30% β -mercaptoethanol and incubating for 1 min on ice. Afterwards, the liquid was removed, and the plants were washed twice with distilled water and frozen in liquid nitrogen. RNA was extracted using the Spectrum Plant Total RNA Kit (Merck, Darmstadt, Germany). DNA was removed using the Turbo DNA-free kit (Thermo Fisher Scientific). cDNA was produced using 1- μ g aliquots of RNA as template, 0.5 μ M target-specific primer (see Supplemental Table S4), 100 units TGIRT-III (InGex, St. Louis, MO, USA) reverse transcriptase in TGIRT buffer (50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM MgCl₂), 1 mM dNTPs, 5 mM dithiothreitol, and 4 units of Murine RNase Inhibitor (NEB)). The mixture was incubated for 2 h at 57 °C. RNA was removed by adding five units of RNase H (NEB) and incubating for 20 min at 37 °C. RNase H was inactivated by 20-min incubation at 65 °C. cDNA was purified using 1.8X strength Ampure XP beads (Beckman Coulter). The region of interest was amplified with specific primers (see Supplemental Table S4) and the Q5 DNA polymerase (NEB), indexed by PCR using primers containing Illumina indexes (see Supplemental Table S4), and sequenced on a MiSeq (Illumina) sequencer (2 \times 300 bp).

For in vitro DMS probing, 5 μ g (20 μ L) of DNase treated RNA in water was heat denatured for 2 min at 95 °C and quickly transferred to ice. 80 μ L of DMS reaction buffer (100 mM KCl, 40 mM HEPES pH 7.5, 0.5 mM MgCl₂) and 100 U of Murine RNase Inhibitor (NEB) were added, followed by incubation with mixing at 25 °C for 5 min. Next, DMS (or water for DMS- samples) was added to the final concentration of 5% (*w/v*) and samples were incubated for 5 min at 25 °C with gentle mixing. The reaction was terminated by adding 200 μ L of ice cold 30% β -mercaptoethanol and incubating for 1 min on ice. RNA

was recovered by ethanol precipitation. cDNA synthesis, library preparation, and sequencing was done as described above.

2.11. DMS-MapSeq Analysis

Adapter sequences from reads were removed using TrimGalore! (version 0.4.5) with the following settings: `—fastqc—quality 35—length 75` (and `—max_length 200` for reverse reads). Reads were mapped to the *psbA*, *rbcL*, and 16S rRNA using bowtie2 (version 2.3.4.1) [52] separately for forward and reverse reads with following settings: `—local—very-sensitive-local -p 12 -U`. Mutation frequencies for *psbA*, *rbcL*, and 16S rRNA regions located between the primers used for amplification were calculated using the pileup function from the Rsamtools package [53]. For further analysis, substitutions and deletions at nucleotides with coverage higher than 1500 reads and not bound by primers were used. Raw DMS reactivities from DMS- were subtracted from DMS+ samples and all negative values set to 0. Next, DMS reactivities were normalized separately for G/U and A/C by dividing the reactivities by the mean reactivity of the most highly reactive nucleotides (90th–99th percentile) of each transcript followed by 99% winsorization to remove extremely high values, as described earlier [27]. For structure prediction, RNA sequences were folded by the Fold program from RNAstructure (version 6.2) [54] with normalized DMS reactivities for all nucleotides used as soft constraints. Fold program parameters were as follows: `—md 500—t 298.15`. For the *psbA* high light samples, the protein binding site was forced to be single-stranded. Structures were visualized using VARNA [55].

2.12. Structure Analysis of 16S rRNA

The crystal structure of the chloroplast 70S ribosome [56] was downloaded from PDB (<https://www.rcsb.org/>, entry 5 × 8P). Surface residues (i.e., solvent accessible) were calculated in PyMOL using the FindSurfaceResidues module. Residues with an area > 2.5 Å² were considered as solvent accessible.

2.13. RNA Secondary Structure Probing with NAI-N₃

Ten samples in all were prepared, eight of which were derived from plant material grown under low light control conditions and two from high light material. All but one of the samples were structure-probed using the SHAPE reagent NAI-N₃; the exception was exposed to a mock treatment using DMSO as a probing control [57]. All samples, except the DMSO control and one low light sample, were selected for probing of induced termination [58]. One low light sample was subjected to probing of in vitro-folded RNA. The others were probed using homogenized, flash-frozen leaf tissue, imitating in vivo conditions. The sample probed in vitro, as well as two low light and the two high light samples, were depleted of rRNA, whereas all others were comprised of total RNA. For low light conditions, there were two biological replicates each for the total RNA and the rRNA-depleted RNA. For the total RNA, an additional technical replicate was generated by splitting the sample after DNase treatment (subsection 2.15.). For the high light conditions, two biological replicates were analyzed.

A 2 M NAI-N₃ stock solution was prepared by mixing dropwise 0.15 g of 2-azidomethyl nicotinic acid dissolved in 210 µL DMSO with 0.14 g of carbonyldiimidazole in 210 µL DMSO, and letting the two react for 1 h. Probing was done by adding to 100 mg of deep frozen, ground leaf material 540 µL extraction buffer (0.92 M HEPES/KOH pH 8.0, 5 mM MgCl₂, 0.5 mg/mL heparin, 1% Triton X-100, 2% polyoxyethylen-10-tridecyl-ether) premixed with 60 µL of 1 M NAI-N₃ in DMSO (giving a final concentration of 100 mM). The sample was incubated for 2 min at room temperature. The reaction was stopped by addition of β-mercaptoethanol to a final concentration of 1.4 M. Cell debris was removed by centrifugation for 5 min at 13,200 g and 4 °C. RNA was isolated using phenol/chloroform/isoamylalcohol (25:24:1) extraction and isopropanol precipitation. DNA was removed using the RNase-Free DNase Set (Qiagen) according to the

manufacturer's protocol. For some of the samples (see above), rRNA was depleted using the Ribo-Zero Bacterial rRNA Removal Kit (Illumina). To evaluate the efficiency of subsequent selection of probed RNA, before cDNA synthesis 1% and 2% of *E. coli fhlA220* mRNA was spiked into total RNA and rRNA-depleted samples, respectively. The cDNA synthesis was carried out with modifications as described [59]. Specifically, 1 μ L 50 μ M random primer (RT_15xN, see Supplemental Table S4) was annealed to 8 μ L of total RNA or rRNA-depleted RNA by incubation at 65 °C for 5 min and then transferred to ice. A 28- μ L aliquot of a mastermix consisting of transcription buffer (250 mM HEPES pH 8.3, 375 mM KCl, 15 mM MgCl₂), 7.5 μ L 2.5 mM dNTPs, 7.5 μ L sorbitol (3.3 M)/trehalose (0.6 M), 500 units PrimeScript Reverse Transcriptase (Takara, Kyoto, Japan), and 3 μ L water was added to each sample. Samples were then incubated at 25 °C for 30 s, 42 °C for 30 min, 50 °C for 10 min, 56 °C for 10 min, and 60 °C for 10 min, and subsequently purified using AMPure XP RNA beads (Beckman Coulter) according to the manufacturer's protocol. The samples were biotinylated as described earlier [57]. Full-length cDNA was selected using 100 μ L MPG Streptavidin beads (PureBiotech, Middlesex, NJ, USA) per sample as described [59] with minor alterations. The beads were blocked with 1.5 μ L of a 20 μ g/ μ L *E. coli* tRNA mix for 60 min at room temperature, separated from the supernatant on a magnetic stand, and washed twice with 50 μ L of wash buffer 1 (4.5 M NaCl, 50 mM EDTA pH 8.0), followed by resuspension in 80 μ L wash buffer 1. The beads were then mixed with 40 μ L of cDNA/RNA sample and incubated at room temperature for 30 min with vortexing every 10 min. After 5 min on a magnetic stand, the supernatant was removed, and the beads were washed a total of six times with 150 μ L of the following wash buffers: 1 \times wash buffer 1, 1 \times wash buffer 2 (300 mM NaCl, 1 mM EDTA pH 8.0), 2 \times wash buffer 3 (20 mM Tris-HCl pH 8.5, 1 mM EDTA pH 8.0, 500 mM NaOAc pH 6.1, 0.4% SDS), 2 \times wash buffer 4 (10 mM Tris-HCl pH 8.5, 1 mM EDTA pH 8.0, 500 mM NaOAc pH 6.1). To release the cDNA from the beads, 60 μ L of 50 mM NaOH was added and the samples were incubated for 10 min at room temperature. The eluate was removed after separation on a magnetic stand and mixed with 12 μ L of 1 M Tris-HCl (pH 7), followed by ethanol precipitation. Libraries were prepared as previously described [58] with minor modifications. A mixture consisting of 1 μ L 10 \times Circligase reaction buffer (Epicentre), 0.5 μ L 1 mM ATP, 0.5 μ L 50 mM MnCl₂, 2 μ L 50% PEG 6000, 2 μ L 5 M betaine, 0.5 μ L 100 μ M Ligation_adapter oligonucleotide (see Supplemental Table S4), and 50 units of Circligase (Epicentre) was added to 3 μ L of cDNA and incubated at 60 °C for 2 h and 68 °C for 1 h, followed by enzyme inactivation at 80 °C for 10 min. The ligated cDNA was purified by ethanol precipitation and resuspended in 20 μ L water, of which 5 μ L was used for PCR with 45 μ L PCR reaction mix (3 μ L of PCR_forward primer, 2.5 μ L of indexed reverse primer (see Supplemental Table S4), 10 μ L of Phusion 5 \times HF buffer, 4 μ L of 2.5 mM dNTPs, 24.5 volume water, and 2 units Phusion Polymerase (NEB)). The PCR was conducted with the following cycles: 1 \times (98 °C for 3 min), 5 \times (98 °C for 80 s; 64 °C for 15 s; 72 °C for 1 min), 16 \times (98 °C for 80 s; 72 °C for 45 s), 1 \times (72 °C for 5 min) and purified using AMPure XP beads, eluting the PCR product in 30 μ L water. The molar distribution of the individual samples was analyzed using a Bioanalyzer High sensitivity chip (Agilent) and used to pool samples equally followed by size selection (200–600 bp range) on an E-gel 2% SizeSelect gel (Invitrogen). The size-selected library was precipitated and resuspended in 20 μ L of water followed by AMPure XP bead (ratio 1:1.8) purification. The library was sequenced on the Illumina NextSeq system with the 75 bp single-end protocol.

2.14. In Vitro RNA Secondary Structure Probing with NAI-N₃

DNA-depleted RNA was folded in vitro and SHAPE probed as described [57], with modifications. Specifically, 10 μ g RNA in water was heat-denatured for 2 min at 95 °C and transferred to ice. SHAPE reaction buffer (100 mM HEPES pH 8.0, 6 mM MgCl₂, 100 mM NaCl) and 400 units of RiboLock RNase inhibitor (Thermo Fisher Scientific) were then added, followed by incubation for 5 min at 37 °C. Subsequently, NAI-N₃ was added to a final concentration of 100 mM, followed by gentle mixing and incubation at 37 °C for

10 min. The reaction was terminated with β -mercaptoethanol (1.4 M final concentration) and the RNA was recovered by ethanol precipitation. Reverse transcription, biotinylation, selection of probed sequences, library preparation, and sequencing were done as described above.

2.15. SHAPE Data Analysis

Data analyses were conducted either on a Debian Linux server as command line functions or in RStudio (V. 1.1.456). Adapter sequences, short reads, and low quality 3'-ends were removed from the reads using cutadapt v. 1.15 [60] with the options cutadapt -a GATCGGAAGAGCACACGTCT—quality-cutoff 17—minimum-length 40. The random barcodes incorporated into the 3' adapter were removed and saved for later analysis using preprocessing.sh [61] with the options -b NNNNNNN and -t 15 for barcode and trimming length respectively. Sequenced reads were mapped using Bowtie2 v2.2.3 [52] with the options—norc -N 1 -D 20 -R 3 -L 15. The reads were mapped to a fasta file containing manually annotated Arabidopsis transcripts (subsection 2.7.). Using the barcode and sequence information, the counts from observed unique barcodes were summarized with summarize_unique_barcodes.sh with the options -t -k to trim untemplated nucleotides and to produce a k2n file, respectively [61]. To account for bias during library preparation the estimated unique barcodes were calculated with the R package “RNAprobR” (v. 1.2.o) function readsamples() with the euc = “HRF-Seq” option [61] using Rstudio Version 1.1.456. Finally, the count data from the estimated unique counts were compiled with the original fasta file to create positional information using the RNAprobR function comp(). The compiled data were subsequently normalized by a 90% winsorization, whereby all values in a sliding 51-nt window were set to the 98th percentile. Comparison of the samples revealed that two samples were extreme outliers (Supplemental Figure S8) and they were excluded from further analysis. Both were low light samples, one of which contained total RNA (LL4), the other one had been depleted of rRNA (LL5). The remaining samples included in the following analysis were: the low light samples LL1 (total RNA), LL2 (total RNA, technical replicate of LL1), and LL3 (rRNA depleted) plus the DMSO control, the not selected control and one in vitro-folded sample; for high light conditions, HL1 and HL2 (both rRNA depleted). For structural analysis only genes with on average more than 10 reverse transcription stops per nt were used (Supplemental Figure S8). Positions in these genes that were missing swinsor values in at least one of the LL or HL samples were not analyzed (e.g., positions 9 and 7 in *rbcL*; Figure 3). Normalized swinsor values for selected motifs (i.e., start, SD, as-SD, and as-start) were calculated by dividing, for each nucleotide, the swinsor value by the average swinsor for that nucleotide in all LL samples. The SDs were identified by in silico hybridization of the anti-SD CCUCCU of the 16S rRNA to nucleotides -22 to -2 of each 5' UTR at 20 °C using Free2bind [62]. The same program was also used to determine the strength of the interaction between SDs and anti-SD, and the SDs were classified into strong and weak categories.

2.16. Receiver Operating Characteristic (ROC) Curve

Using the roc() function from the pROCpackage (v. 1.9.1) in R, a receiver operating characteristic (ROC) curve was generated using the dot-bracket structure from the Arabidopsis 18S rRNA obtained from the “The comparative RNA web” (CRW) site [63] as predictor and the swinsor normalized termination counts as response. From the generated ROC curve, the area under the curve was calculated. To assess the quality of DMS data, we performed receiver operating characteristic (ROC) using pROC package [64] based on crystal structure of chloroplast ribosome [56], as described earlier [27].

2.17. Data Analyses

All data analyses were performed in R [65] and plotted using ggplot2 [66].

3. Results

We tested whether changes in mRNA secondary structure would influence translation in chloroplasts using a well-known example for translation regulation as a starting point: the high light induced upregulation of *psbA* translation, the mRNA coding for the D1 subunit of photosystem II [30,31]. First, we validated that in young *Arabidopsis thaliana* plants (17–18 days old) *psbA* translation was induced by exposure to high light for one hour. For this, we extracted polysomes, size-fractionated them in sucrose gradients, and analyzed the distribution of *psbA* mRNA by RNA gel blot analysis. As expected, we observed a prominent shift of *psbA* mRNA into denser fractions (relative to low light controls), which indicates increased loading of ribosomes and higher translation initiation rates in high light (Supplemental Figure S1). Polysome analysis is independent of the mRNA level, because it fractionates the mRNA according to its ribosome loading.

3.1. mRNA Secondary Structure Changes in the *psbA* Translation Initiation Region

Next, we focused our analysis on the translation initiation region of the *psbA* mRNA and analyzed its in vivo secondary structure using dimethyl sulfate (DMS) probing. DMS was described to methylate only N1 of adenosines and N3 of cytidines of single-stranded and accessible RNA [67]. However, recently, it was demonstrated that under alkaline conditions DMS can probe also guanosines and uridines [68]. As the chloroplast stroma is slightly alkaline, all four nucleotides of chloroplast RNA can be probed, although the probing of adenosines and cytidines is more reliable [27]. High DMS probing at a nucleotide indicates a single-stranded confirmation. Low probing can be caused by double stranded regions, protein binding or compact RNA secondary structure preventing DMS access [67]. DMS efficiently enters cells [69], including those of *Arabidopsis* plants [70], and is therefore suited for in vivo structural probing. DMS-reactivity of probed nucleotides can be quantified by mutational profiling (MaP) using a thermostable group II reverse transcriptase (TGIRT), which during reverse transcription incorporates mutations in the cDNA at the reacted positions [71]. The young plants were exposed to either low light or high light for one hour and then, in the same light regime, incubated in a DMS solution for six minutes (subsection 2.10.). The probing did not cause browning of the leaves as previously observed [72], and the quality of the extracted RNA was not affected by the treatment (Supplemental Figure S2). Using gene-specific primers, we analyzed the translation initiation region of *psbA* and, as a control, helix 33 of the plastid 16S rRNA (see Supplemental Table S1 for the coverage). In parallel, we probed purified RNA that had been refolded in vitro (Figures 1 and Supplemental Figures S3–S5). In addition, we analyzed young plants treated with water instead of DMS under low light and high light conditions and found a very low background level of mutations using our protocol (Supplemental Figures S3A, S4 and S5B). As expected for DMS, adenosines and cytidines are statistically significantly more probed than guanosines and uridines (Supplemental Figure S3A). Furthermore, the observed DMS probing for helix 33 of the 16S rRNA corresponds nicely with the rRNA structure previously described for plastid ribosomes (Ahmed et al., 2017) (Supplemental Figure S5) and is similar for low and high light conditions (Supplemental Figures S3C and S5B). The structure signals for guanosines and uridines in vivo were, as expected [27,68], weaker than those for adenosines and cytidines, but still informative compared to the in vitro, protein-free control (Supplemental Figure S5A). In addition, the reproducibility of the probing of adenosines and cytidines was better than for guanosines and uridines (Supplemental Figure S3B).

The DMS-MaPseq results for the *psbA* translation initiation region were highly reproducible (Supplemental Figure S3B–D). Obvious differences were detected around the Shine-Dalgarno sequence and the start codon (Figure 1A,E,F). In high light, both

elements had higher DMS probing than in the low light samples. This was true when the probing of all four nucleotides was considered, as well as when only the more reliable data at adenosines were considered (Figure 1E,F). The increased DMS probing indicates that these RNA regions are more single-stranded and accessible under high light conditions, which is in agreement with the observed increase of *psbA* translation (Figure 1B, Supplemental Figure S1). Interestingly, an upstream sequence, complementary to the Shine-Dalgarno sequence, also displayed increased DMS probing under high light conditions, suggesting that this sequence might interact with the Shine-Dalgarno sequence under low but not under high light conditions, and thus could control translational activation (Figure 1A,D). This would be in agreement with the hypothesis that translation efficiency is low when the Shine-Dalgarno sequence and/or the start codon are occluded in a double-stranded region. The opening of the structure would make these elements more accessible, which should boost translation efficiency.

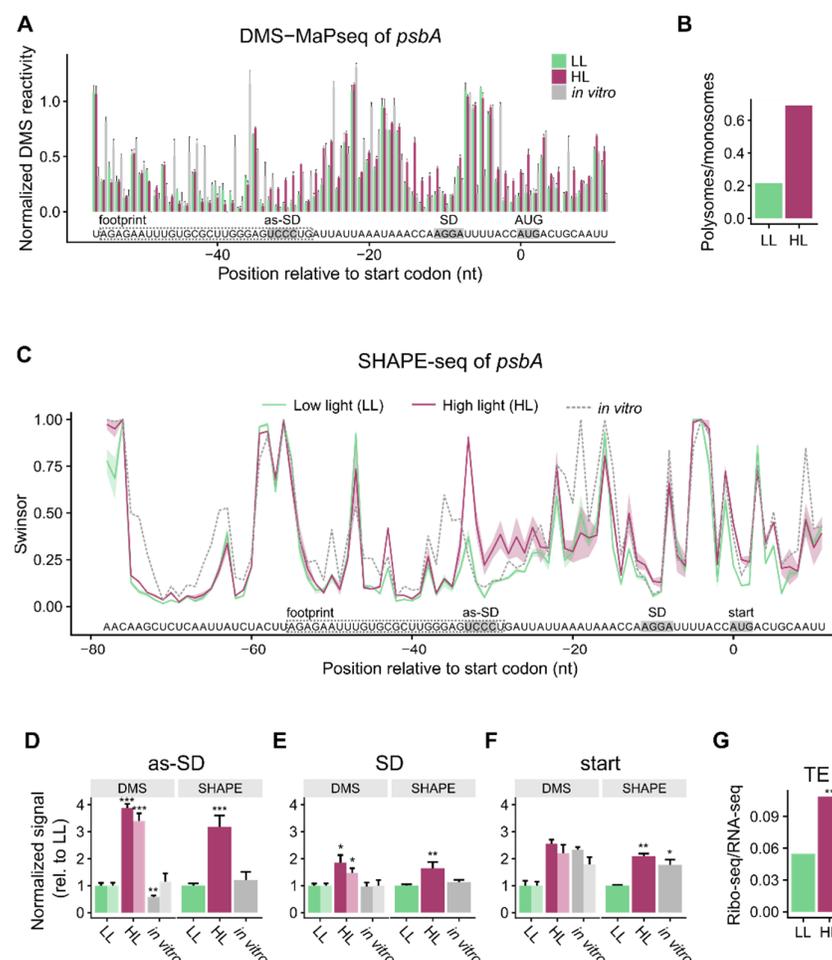


Figure 1. mRNA secondary structure changes in the translation initiation region of *psbA*: Increased accessibility of the Shine-Dalgarno sequence and the start codon correlates with increased translation efficiency. (A) DMS-based mapping of *psbA* mRNA secondary structure of young, 17–18-days-old plants as determined by MaPseq. Normalized DMS reactivities are given to account for the differences between adenosines/cytidines vs. guanosines/uridines reactivities (Supplemental Figures S3A and S4A). Furthermore, the values of a control without added DMS (Supplemental Figure S4A) were subtracted and outliers were removed by winsorization (only the 90th percentile is retained). The information obtained at adenosines/cytidines is more reliable than at guanosines/uridines (Supplemental Figures S3B and S5A) [27]. High normalized DMS reactivity indicates single-stranded nucleotides. The data for the low light (LL) control are shown in light green, the high light (HL) samples in dark red, and the mRNA that was allowed to fold *in vitro* in

gray. The error bars indicate the mean standard error. The start codon (start), Shine-Dalgarno sequence (SD), and a sequence that can bind the SD (as-SD) are marked. The position of the footprint of a putative regulatory protein is given as a dashed box (Supplemental Figure S11A). A comparison of the DMS-probed RNA with a water-treated control is shown in Supplemental Figure S4A. (B) Polysome analysis of *psbA* translation in 17–18-days-old plants (Supplemental Figure S1). The ratio of the *psbA* mRNA amount in the five most dense fractions (polysomes) to the amount in the five least dense fractions (monosomes) is given. (C) SHAPE analysis (NAI-N₃ probing) of young leaves of seven-week-old plants. SHAPE signals indicate the extent to which each nucleotide is unpaired. Swinsor values are the termination counts, i.e., how often reverse transcription was stopped at each nucleotide by a bound NAI-N₃ probe, normalized by winsorization (only the 90th percentile is retained, outliers are discarded). High swinsor values indicate unpaired nucleotides, low swinsor values base-paired nucleotides. The shaded areas around the lines indicate the mean standard error. (D) Average of mRNA secondary structure at the sequence binding the Shine-Dalgarno sequence (as-SD), as revealed by DMS and SHAPE values normalized to the low light values. The DMS columns in darker color represent the more reliable reactivities at adenosines/cytidines, the lighter the reactivities at all four nucleotides. Asterisks here (and in (E,F)) indicate statistically significant changes compared to LL (*P*-values calculated with the Wilcoxon rank sum test; * = *p* < 0.05, ** = *p* < 0.01, and *** = *p* < 0.001), error bars indicate mean standard error. (E) Average structure at the Shine-Dalgarno sequence (SD). (F) Average structure at the start codon (start). (G) Change in translation efficiency (ratio footprints/transcript levels) of *psbA* mRNA in young leaves of seven-week-old plants (Figures 4, S12, and S13). Asterisks indicate statistically significant changes (calculated with RiboDiff; ** = *p* < 0.01).

To further validate the observed structural changes in the *psbA* mRNA, we used a complementary method, selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) [73,74], to probe the RNA structure. Furthermore, to test the robustness of the response, we applied SHAPE to analyze *psbA* secondary structure changes in response to high light acclimation in mature plants. Arabidopsis plants grown in short-day, low light conditions were acclimated to high light by exposing seven-week-old plants to four hours high light, 16 h dark, and again four hours high light. To analyze translation, we used young leaves, which were found to be more capable of acclimating to the high light conditions than mature, fully expanded leaves (Supplemental Figure S6). In agreement with the results obtained for the long day-grown young plants used for DMS probing (Supplemental Figure S1), *psbA* translation in seven-week-old plants was increased after one hour in high light (Supplemental Figure S7). This increase was still present after one day in the above-described acclimation conditions (Supplemental Figure S7). For RNA secondary structure probing, we analyzed RNA after one day high light acclimation using the SHAPE reagent NAI-N₃ [73,74]. Like other SHAPE reagents, NAI-N₃ reacts with the 2'-hydroxyl groups in the RNA backbone when the RNA adopts specific conformations that are characteristic for flexible single-stranded RNA, but it does so much less efficiently if their flexibility is constrained by base pairing [75,76]. Hence, NAI-N₃ effectively probes for the presence of single-stranded nucleotides. The SHAPE reactivity profile can be read out by mapping termination sites of reverse transcription caused by the introduction of SHAPE adducts, and NAI-N₃ can be used for intracellular probing experiments [73]. We added NAI-N₃ to flash-frozen leaf samples and probed the RNA during the thawing of the high light and low light samples. In addition, we performed in vitro probing on purified RNA, which had been refolded in vitro. We performed SHAPE selection on the samples, as previously described [58], and the counts obtained were normalized for coverage using Smooth Winsorization [61] to give SHAPE reactivities between 0 and 1.

First, we investigated the correlation between the replicates in our samples using PCA analysis (Supplemental Figure S8A). As expected, the quality of the probing signal was dependent on the sequence coverage; therefore, we limited our analysis to RNAs having on average more than 10 termination counts per nucleotide (Supplemental Figure S8B). In the PCA plot, the in vitro probing data are clearly separated from the in vivo samples, and the two high light samples cluster together. Among the five low light

samples, three samples clustered together, whereas the remaining two deviated both from the other three and from each other; therefore, we excluded these two samples from our further analysis. Next, we checked the structural signal in the dataset by comparing the SHAPE probing data for the Arabidopsis 18S rRNA with the known secondary structure of this RNA. For all samples, except those having low coverage of the 18S rRNA owing to prior rRNA depletion, we observed a signal for RNA structure (Supplemental Figure S9).

For the translation initiation region of *psbA*, we observed good reproducibility of SHAPE reactivities among replicates (Figure 1C). The Shine-Dalgarno sequence, start codon, and the sequence that can potentially bind the Shine-Dalgarno sequence (as-SD) showed higher SHAPE reactivity in high light samples than in low light controls (Figure 1C–F). The SHAPE reactivities correlate with the DMS-MaP signal observed in the translation initiation region, especially with the more reliable DMS probing of adenosines and cytidines in the region from the as-SD to the start codon (Figure S10). Thus, using two different chemical probes, we find that the *psbA* translation initiation region becomes more accessible under high light conditions (Figure 1A,C,E,F, Supplemental Figure S10). The effect correlates well with increased *psbA* translation (Figure 1B,G) and is observed after both short-term high light stress in young plants and long-term high light acclimation of young leaves of seven-week-old plants.

3.2. Change of mRNA Secondary Structure of *psbA* Translation Initiation Region Likely Caused by Protein Binding

One potential means of altering mRNA secondary structure is the binding of RNA-binding proteins. Analyzing the reads from MNase-digested RNA (subsection 2.7.), we detected a footprint of a putative regulatory protein in the 5' UTR of *psbA* (Supplemental Figure S11A). We confirmed the footprint by northern blot analysis of small RNAs isolated without prior RNase treatment using a probe specific for the footprint sequence (Supplemental Figure S11D) as it was done previously [51,77]. A central part of this footprint has previously been described as a site where HCF173 binds alone or with other unknown proteins [36]. HCF173 is a protein that activates *psbA* translation [34,35]. The detected footprint is located upstream of the Shine-Dalgarno sequence. We assessed the possible influence of a bound protein on the *psbA* mRNA secondary structure by predicting the structure using DMS reactivities and the position of the bound protein as constrains. The prediction revealed that the footprint contains sequences that can bind to the Shine-Dalgarno sequence and the start codon. In low light, these *cis*-elements are part of a double-stranded structure (Figure 2A). In high light, the Shine-Dalgarno sequence and the start codon are in a largely single-stranded structure and therefore more accessible (Figure 2B). The abundance of the footprint increased in high light compared to low light (Supplemental Figure S11A,B,D). However, we also observed increased *psbA* mRNA levels under high light conditions (Supplemental Figure S11C), and this could potentially explain the increased accumulation of small RNAs stemming from this region.

As an alternative approach to distinguish between double-stranded RNA and a bound protein, we analyzed the DMS reactivity at the nucleotides of the footprint and around the Shine-Dalgarno sequence that are predicted to pair in low light (Figure 2A). DMS probing is sensitive to protein binding [78,79], therefore DMS reactivity is low both for a nucleotide bound to a protein and a nucleotide involved in base pairing. High DMS reactivity indicates a single-stranded, not protein-bound nucleotide. The half of the stem loop to which the protein binds had low DMS reactivities both in low light and high light (Figure 2C). In contrast, the DMS reactivities of the other half of the stem loop, at the sequence around the Shine-Dalgarno sequence and the start codon, increased in high light (Figure 2C). This suggests that these nucleotides pair in low light to nucleotides of the protein-binding site. In high light, a protein prevents the formation of the double-stranded structure and thereby increases the accessibility of the *cis*-elements required for translation initiation and *psbA* translation. The analysis of DMS reactivities only at adenosines and cytidines showed the same trend as the one for the DMS reactivities at all

four nucleotides (Supplemental Figure S11E). Interestingly, whereas the average DMS reactivities of paired nucleotides at the footprint were similar in low light and high light (Figure 2C), the DMS reactivities of the nucleotides that can bind the Shine-Dalgarno sequence increased in high light (Figure 1A,D). This sequence is located at the 3' end of the footprint (Figure 1A) but is still part of the footprint, which is protected from nuclease attack. A possible explanation is that the protein binding at the footprint interacts only with some of the nucleotides and therefore does not influence the DMS reactivities of the other nucleotides (compare also Figure S10). The SHAPE reactivities differ from the DMS reactivities (Supplemental Figure S11F), which can be caused by differences between SHAPE reagents and DMS in the sensitivity to protein binding. DMS reactivity is low in case of bound proteins [78,79], whereas bound proteins are not always detected as nucleotides with low SHAPE reactivity [73,80,81] (compare also Figure S10).

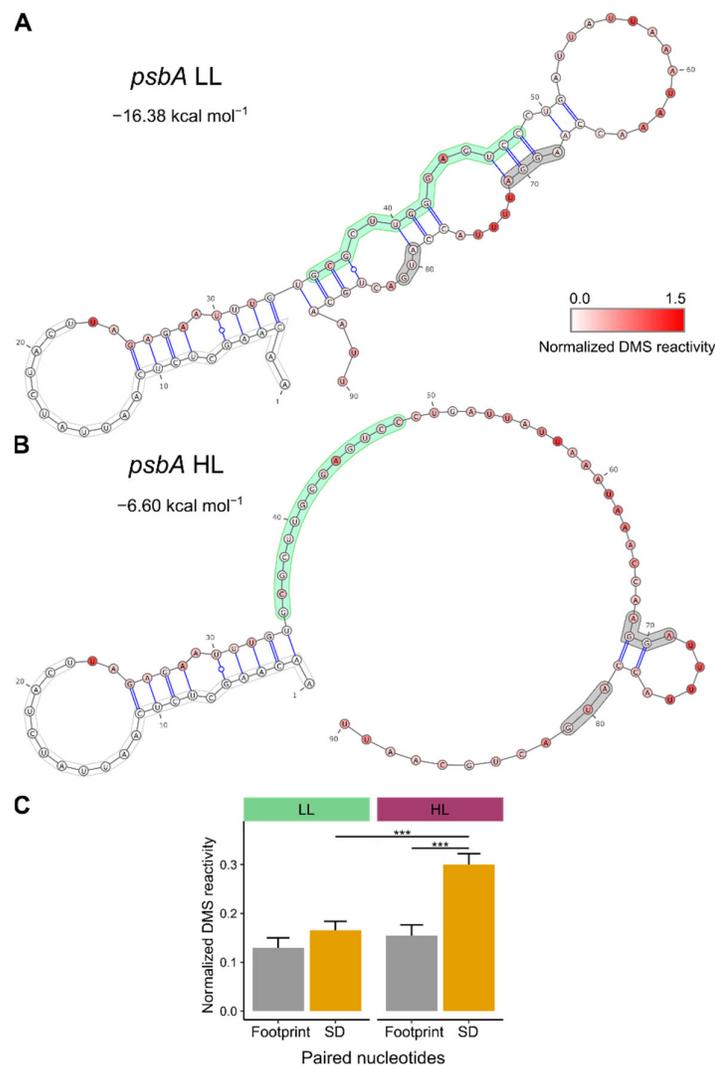


Figure 2. Footprint of a putative regulatory protein bound to the 5' UTR of *psbA*: **(A)** Predicted mRNA secondary structure of the *psbA* translation initiation region in low light (LL) using normalized DMS reactivities (Figure 1A) as constrains. The white box marks the position of the primer used to amplify the cDNA. For this region no DMS reactivities could be obtained. The green box marks the footprint of a RNA binding protein [36] (Supplemental Figure S11A–D), the grey boxes indicate the Shine-Dalgarno sequence (AGGA) and the start codon (AUG). For each nucleotide, the normalized DMS reactivity is shown in a color code. The kcal mol^{-1} value for the

strength of the RNA structure is given. **(B)** Predicted mRNA secondary structure in high light (HL) using normalized DMS reactivities (Figure 1A) and the protein binding site (forced to be single-stranded) as constrains. For the structure predictions for in vitro-folded RNA see Supplemental Figure S11G. **(C)** Normalized DMS reactivities of the nucleotides predicted to form base pairs in low light **(A)** between the region of the footprint (between nucleotides 35–48) and the region including the start codon and the Shine-Dalgarno sequence (SD) (between nucleotides 69–86). The average normalized DMS reactivities are shown separately for both regions. Nucleotides in these regions predicted not to be paired are excluded. There is no significant difference between low light and high light for the low DMS reactivities at the footprint side, which can indicate both double-stranded RNA and a bound protein. In contrast, the DMS reactivities at the SD side significantly increase in high light indicating a shift to single-stranded RNA. This suggests that in low light a stem loop structure is formed **(A)**, whereas in high light a protein is bound to the *psbA* translation initiation region making the SD and the start codon accessible **(B)**. Asterisks indicate statistically significant changes (p -values calculated with the Wilcoxon rank sum test; *** = $p < 0.001$), error bars indicate mean standard error. For the separately analyzed DMS reactivities at adenosines and cytidines as well as SHAPE reactivities see Supplemental Figure S11E,F.

3.3. mRNA Secondary Structure of the Translation Initiation Regions of *rbcL*

As an additional example, we examined the translation initiation region of *rbcL*, which encodes the large subunit of RuBisCO. In contrast to *Nicotiana tabacum* [30], in Arabidopsis, *rbcL* translation is increased after a shift to high light in young plants (Supplemental Figures S1 and 3G) and in young leaves of seven-week-old plants (Figures 3H and 4). However, using DMS-probing of high light-treated young plants, we observed a slight decrease of DMS reactivity at the Shine-Dalgarno sequence and no structural change at the start codon of *rbcL* (Figure 3A,C,E). Furthermore, in high light-treated young leaves, the Shine-Dalgarno sequence and the start codon show a reduction in SHAPE reactivity, indicating that the translation initiation region of *rbcL* is more compactly folded and less accessible in high light conditions (Figure 3B,C,E). In this case, our data do not support that translation initiation is regulated by the accessibility of the Shine-Dalgarno sequence and start codon. Moreover, we also did not observe significant changes in the accessibility of the sequences that have the potential to interact with the start codon and the Shine-Dalgarno sequence (Figure 3D,F).

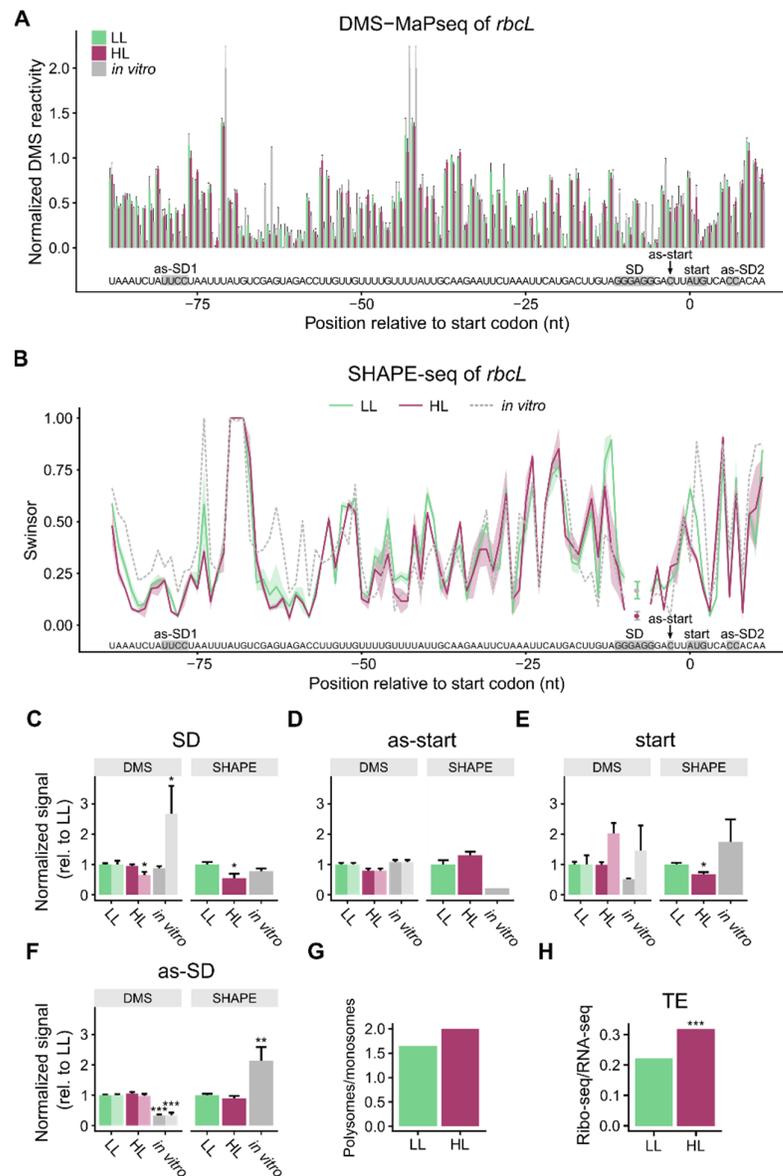


Figure 3. mRNA secondary structure changes in the translation initiation region of *rbcL*: (A) DMS probing of RNA from young, 17–18-days-old plants exposed to low light (LL, light green), high light (HL, dark red), and of *in vitro*-folded RNA (gray). Normalized DMS reactivities are shown (compare Figure 1). The error bars indicate the mean standard error. The start codon (start), Shine-Dalgarno sequence (SD), and sequences that can bind the start codon (as-start) and the SD (as-SD1 and as-SD2) are marked. One portion of the SD can bind to a sequence in the coding region (as-SD2), the other one to a sequence upstream in the 5' UTR (as-SD1). A comparison of the DMS-probed RNA with a water-treated control is shown in Supplemental Figure S4B. (B) SHAPE analysis (NAI- N_3 probing) of RNA from young leaf tissue obtained from seven-week-old plants presented as swinsor normalized termination counts. The shaded areas around the lines indicate the mean standard error. Positions -9 and -7 were not analyzed, because for at least one of the LL or HL samples the swinsor value was missing. (C) Average incidence of structure at the Shine-Dalgarno sequence (SD) as shown by DMS and SHAPE values normalized to the low light values. The darker colored columns represent the more reliable reactivities at adenosines/cytidines, the columns in lighter colors are the reactivities at all four nucleotides. Asterisks here (and in (D) to (F)) indicate statistically significant changes compared to LL (p -values calculated with the Wilcoxon rank sum test; * = $p < 0.05$, ** = $p < 0.01$, and *** = $p < 0.001$), error bars indicate mean standard error. (D) Average mRNA secondary structure at the nucleotide binding the start codon (as-start). (E) Average structure at the start codon (start). (F) Average structure at the sequences

binding the SD (as-SD1 and as-SD2). (G) Change in translation in 17-18-days-old plants (polysome fractions/monosome fractions, Figure S1). (H) Change in translation efficiency (ratio footprints/transcript levels, Figure 4) in young leaves of seven-week-old plants. Asterisks indicate statistically significant changes compared to LL (calculated with RiboDiff; *** = $p < 0.001$).

3.4. mRNA Secondary Structure and Translation Efficiency

As shown above, structural changes of mRNA seem to be important for the high light-induced translational activation of the *psbA* mRNA, but not the *rbcL* mRNA (Figures 1 and 3). We therefore wanted to see if there was a general correlation between structural changes and translation efficiency, or if this was a phenomenon unique to the *psbA* mRNA. Our SHAPE probing experiment of young leaves of seven-week-old plants had sufficient sequencing coverage to allow the analysis of 16 genes, including *psbA* and *rbcL*. Using the same plant material, the translation of these genes in the same plants was analyzed by ribosome profiling. This method is based on the sequencing of nuclease-protected mRNA footprints of ribosomes, which provide, when quantified per reading frame, a proxy for the synthesis rate of the corresponding protein [82]. The reproducibility between replicates was good (Supplemental Figure S12). The translation efficiency was calculated by dividing the amount of ribosome footprints for each reading frame by the transcript levels determined by RNA-seq (Supplemental Figure S13). For several genes, a statistically significant reduction in translation efficiency was noted in high light (Figure 4). An exception was *psbA* whose translation efficiency was increased (Figure 4), which indicates increased translation initiation [29,30] and is in accordance with the results of our polysome analysis (Supplemental Figure S7). Furthermore, we analyzed ribosome pausing on selected genes to determine if the increased amounts of detected ribosome footprints are derived from increased translation or increased pausing. The positions of ribosome pause sites were very similar in low and high light (Supplemental Figure S14). In the case of *rbcL* and *psaA* (encoding the PsaA subunit of photosystem I), the extent of ribosome pausing was also very similar (Supplemental Figure S14). In the case of *psbA*, the magnitude of changes in the extent of ribosome pausing is much smaller than the observed change of translation efficiency (Figure 1G, Supplemental Figure S14). This indicates that the changes of translation efficiency are not caused by altered ribosome pausing.

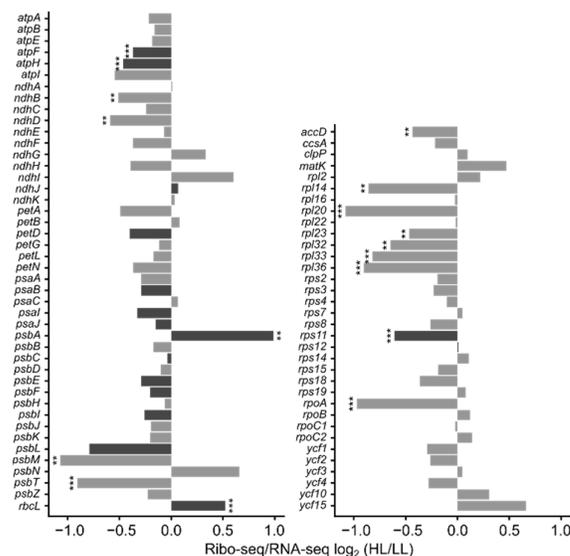


Figure 4. Changes in translation efficiency in response to high light treatment: The translation efficiencies (Ribo-seq/RNA-seq) for all plastid-encoded genes are shown as ratios of the high light (HL) to the low light (LL) scores, expressed as log₂ values. Young leaves of seven-week-old plants were analyzed (as in Figures 1C and 3B). The left panel lists the genes coding for subunits of the

photosynthetic complexes, the right panel shows the data for all other genes. Translation efficiency was determined from normalized read counts for the ribosomal footprints divided by those for the transcripts of each coding region (Figure S13). Asterisks indicate statistically significant changes (calculated with RiboDiff; ** = $p < 0.01$ and *** = $p < 0.001$). Genes with bars in darker gray had sufficient coverage to permit the analysis of mRNA secondary structure (Figures 1, 3 and 5).

If a large proportion of mRNAs is regulated through RNA structural changes similar to what we observed for *psbA* upon exposure to high light, a correlation would be expected between the changes in translation efficiency and the structural alterations at the start codon and/or Shine-Dalgarno sequence (SD). However, this is not the case for either the start codons (Figures 5A-1 and Supplemental Figure S15A) or the SDs (Figures 5A-4 and Supplemental Figure S15D). *psbA* and *rbcL* show a higher translation efficiency in light; yet, a clear correlation with mRNA structure changes can only be found for *psbA* (Figures 1 and 3). Interestingly, these two genes differ strongly regarding the strength of their SD: *rbcL* possesses a strong SD (hybridization to the anti-SD of the 16S rRNA -12.98 kcal mol⁻¹), whereas the SD of *psbA* is much weaker (-5.50 kcal mol⁻¹) (Supplemental Table S2). Regarding the strength of their SD, the 16 genes analyzed can be separated into a group with strongly interacting SDs (hybridization to the anti-SD of the 16S rRNA < -9 kcal mol⁻¹) and a group with weak or no SDs (> -6 kcal mol⁻¹) (Supplemental Table S2). In our set of 16 genes, there were only two genes without an SD, *rps11* and *rps12* (coding for the ribosomal proteins uS11c and uS12c, respectively). Therefore, SD-independent translation could not be investigated specifically, and these two genes were included in the group with weak or no SDs as appropriate. Using these two groups of genes for an analysis of the start codons, we still did not observe a significant correlation between the changes in SHAPE reactivities and the change in translation efficiency (Figures 5A-2,A-3 and Supplemental Figure S15B,C). In contrast, there was a clear difference between the groups regarding the structure at the SD. Genes with weak SDs showed a statistically significant correlation between the change in translation efficiency and the change in SHAPE reactivities in the SDs (Figures 5A-6,B and Supplemental Figure S15F). No such correlation was observed for genes with strong SDs (Figures 5A-5,C and Supplemental Figure S15E; the analysis of additional regions is included in Supplemental Figures S15G–L and S16).

Thus, our data (Figures 1, 2 and 5B) suggest that the structural accessibility of the SD region is central for the light-dependent translational regulation of mRNAs with weak SDs (such as *psbA*), whereas other mechanisms are likely to be more important for mRNAs with strong SDs (Figures 3 and 5C). In the case of the *psbA* mRNA, translational regulation seems to depend on the recruitment of specific proteins to the 5' UTR region and subsequent remodeling of the RNA structure.

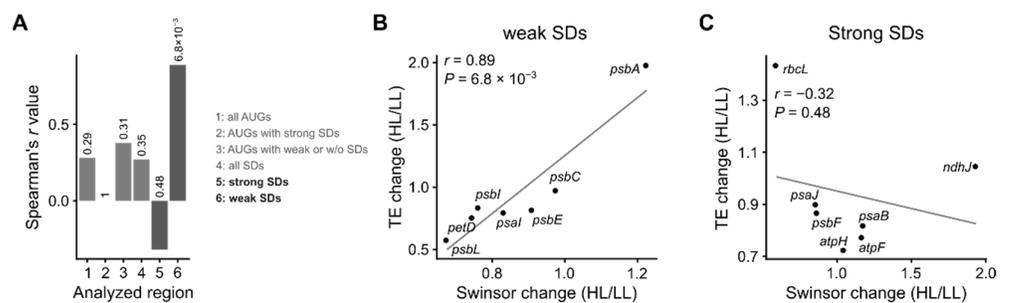


Figure 5. Correlations between changes in mRNA secondary structure and translation efficiency: The changes in mRNA secondary structure are calculated from the swinsor-normalized termination count values derived from NAI-N₃ probing by dividing the values from the high light exposed plants by those from the low light control plants (young leaves of seven-week-old plants). An increase of the swinsor value indicates a decrease in base pairing, i.e., less RNA secondary structure. Average changes for the indicated segments of each gene are given. The change in

translation efficiency is calculated by dividing the normalized read counts for the ribosomal footprints by those for the transcripts of each coding region, and then dividing the resulting values from the high light treatment by those from the low light control. Only genes with sufficient coverage of the mRNA secondary structure (on average at least 10 reverse transcription stops per nucleotide) are included. Spearman's r and p values are given. (A) Overview including all analyzed correlations. Columns 1–6 show Spearman's r for the correlation between the change in translation efficiency and the change in SHAPE reactivities for different gene regions. The corresponding P values are given above the respective column. (1) start codons (AUG); (2) start codons of genes with strong Shine-Dalgarno sequences (SDs) (hybridization to the anti-SD of the 16S rRNA <-9 kcal mol $^{-1}$); (3) start codons of genes with weak or no SD (>-6 kcal mol $^{-1}$); (4) SDs; (5) SDs of genes with strong SD (<-9 kcal mol $^{-1}$); and (6) SDs of genes with weak SD (>-6 kcal mol $^{-1}$). The plots for all these analyses can be found in Supplemental Figure S15, where also an analysis of additional regions is included (see also Supplemental Figure S16). The plots for the highlighted correlations (5) and (6) are shown in (B) (change of structure at the SDs of genes with strong SD) and (C) (SDs of genes with weak SD).

4. Discussion

The molecular mechanisms of translation regulation in plastids of higher plants have been elusive. In vitro data showed that binding of putative regulatory proteins influences the mRNA secondary structure of the region encompassing the start codon and/or the Shine-Dalgarno sequences (SD) [19,20]. It was postulated that such a mechanism might act to regulate translation in vivo. We tested this hypothesis by analyzing the secondary structure and translation efficiency of plastid mRNAs from plants exposed to low and high levels of light.

In high light, *psbA* was the plastid mRNA with the strongest increase in translation efficiency (Figure 4). This was expected, because the turnover of its protein product, D1, increases under these conditions [32]. Regulation of *psbA* translation differs between dark/light shifts and the response to increasing D1 turnover (PSII repair) [24,29]. At least in higher plants, the regulation in response to dark/light shifts happens on the level of translation elongation [29], whereas under conditions of high D1 turnover, *psbA* translation is induced on the level of translation initiation [29,30], as indicated also by polysome analysis (Supplemental Figures S1 and S7) and ribosome profiling (Figure 4). The *cis* elements required for initiation of *psbA* translation are not strongly conserved in higher plants: The *psbA* mRNA in *Arabidopsis* has a weak SD, whereas in some other species, e.g., *Nicotiana tabacum* and *Zea mays*, *psbA* completely lacks a SD (Supplemental Table S2) [8]. In contrast, the *trans* factors regulating *psbA* translation are probably conserved in higher plants. It has been reported that three proteins activate *psbA* translation, i.e., HCF173 [34], HCF244 [35], and LPE1 [83], whereas AtPDI6 is described as a negative regulator [84]. Furthermore, also the chlorophyll-binding proteins OHP1 and OHP2 are important for translation activation of *psbA* [37]. However, conflicting results indicate that LPE1 binds to *psbJ* and *psbN*, not *psbA* [85], and LPE1 was not found to be bound to *psbA* mRNA [36]. It has been proposed that HCF173 is one of the proteins contributing to the footprint detected in the *psbA* 5' UTR (Figures 2 and Supplemental Figure S11A,D) [36]. The footprint of HCF173 on the *psbA* mRNA is conserved over a wide range of higher plant species [36]. D1 is inserted cotranslationally into thylakoid membranes. HCF173 and HCF244 are bound to the thylakoids [35]. HCF244 is possibly recruited there via an interaction with OHP1 and OHP2 [86,87]. If HCF173, HCF244, LPE1, AtPDI6, OHP1 and/or OHP2 are involved in the regulation of *psbA* translation, it could be assumed that these proteins themselves and/or their expression are subject to light-dependent regulation. However, we did not observe any alterations in the transcript levels and translation efficiency of their genes during high light acclimation (Supplemental Table S3), indicating that light-dependent regulation of these proteins, if it occurs, must take place post-translationally or via protein-protein interactions. A complex of HCF244, OHP1, and OHP2 could link the amount of free D1 protein with *psbA* translation. When this complex binds D1, it cannot activate HCF173, whereas without D1

it can activate HCF173, which, in turn, activates *psbA* translation [38]. This would link the regulation of *psbA* translation directly to D1 photodamage and the D1 repair cycle. How the described proteins might activate *psbA* translation, either alone or as a complex, remains unknown.

Using DMS and a SHAPE reagent, NAI-N₃, we demonstrated that the degree of secondary structure of the Shine-Dalgarno sequence and the start codon in the *psbA* mRNA is reduced in vivo under high light conditions, and that this correlates with increased translational efficiency (Figure 1). This correlation is compatible with the hypothesis that translation is activated by making the SD and/or start codon more accessible. Furthermore, we and others found evidence for a possible binding site for a regulatory protein in a position where binding could result in structural changes of the translation initiation region as predicted by the hypothesis (Figure 2) [36]. These findings support the hypothesis that the regulation of *psbA* translation involves the modulation of mRNA secondary structure by protein binding.

There are indications that such a mechanism is used by other genes: in the case of genes with weak SDs, the change in mRNA secondary structure at the SD correlates with the change in translation efficiency (Figure 5B). Interestingly, the correlation is specific for genes with weak SDs. It is possible that strong SDs are more likely to hybridize to the anti-SD of the 16S rRNA, and therefore, are less amenable to regulation by alternative mRNA secondary structures. Accordingly, *rbcL* is an example of a gene with a strong SD (Supplemental Table S2) and here the increased translation efficiency under high light conditions cannot be explained by changes in the structure of the SD and the start codon (Figure 3). How translation of *rbcL* itself is regulated remains unknown. It is possible that distinct mechanisms for regulation of translation initiation exist, as plastids use two distinct mechanisms for start codon recognition [6,8].

It is important to note that the comparison of the structural changes in the translation initiation regions of *psbA* (Figure 1) and *rbcL* (Figure 3) indicated that the structure alterations were not a consequence of increased translation itself, e.g., by increased binding of the ribosome (including tRNA-fMet(CAU)) at the start codon. Both genes were upregulated at the level of translation, but the degree of secondary structure did not change in the same direction. Therefore, these structural changes could not have simply been caused by increased temperatures during high light treatment. Heat would normally be expected to decrease pairing; however, in contrast to *psbA*, the SD and start codon of *rbcL* were not paired to a lower extent in high light. Furthermore, the higher accessibility of the SD of *psbA* in high light was observed both with the DMS-MaPseq and the SHAPE-seq approach (Figure 1, S10), although the temperature difference caused by high light treatment in the DMS-MaPseq experiment was very mild (1 °C, subsection 2.10.) compared to the SHAPE-seq experiment (10 °C, subsection 2.1.).

The results for *psbA* and other plastid genes with weak SDs are in agreement with reports for *E. coli* that translation efficiency is determined by the extent of RNA secondary structure at the SD [4,5]. In bacteria, several mechanisms are described to regulate translation initiation by altering the accessibility of SDs, including RNA thermometers [13,14], binding of small RNAs and proteins [15,16], and riboswitches [12]. Synthetic riboswitches are also functional in plastids [88]. Our results (Figures 1 and 5B) suggest that in plastids, a similar mechanism, based on the manipulation of mRNA secondary structure by RNA-binding proteins, is used for the regulation of translation of *psbA* and other genes with weak SDs.

Supplementary Materials: The following are available online at www.mdpi.com/2073-4409/10/2/322/s1, Supplemental Figure S1: Polysome analysis of *psbA* and *rbcL* translation in young plants, Supplemental Figure S2: RNA quality assessment after DMS treatment, Supplemental Figure S3: DMS-MaPseq quality control including reproducibility, Supplemental Figure S4: DMS-MaPseq compared to water control, Supplemental Figure S5: DMS data reproduce known elements of the secondary structure of 16S rRNA, Supplemental Figure S6: Determination of photosynthetic parameters, Supplemental Figure S7: Polysome analysis of *psbA* translation in mature plants,

Supplemental Figure S8: Reproducibility of the mRNA secondary structure probing (NAI-N₃) data, Supplemental Figure S9: Structural signal in SHAPE (NAI-N₃) data for the 18S rRNA structure, Supplemental Figure S10: Correlation between DMS and NAI-N₃ probing of the mRNA secondary structure of the *psbA* translation initiation region, Supplemental Figure S11: Footprint of a putative regulatory protein bound to the 5' UTR of *psbA* and secondary structure of the *psbA* translation initiation region, Supplemental Figure S12: Reproducibility of translational and transcriptomic data, Supplemental Figure S13: Changes in translation and mRNA levels of plastid-encoded genes, Supplemental Figure S14: Ribosome pausing on *psbA*, *rbcL*, and *psaA*, Supplemental Figure S15: Correlations between changes in mRNA secondary structure and translation efficiency, Supplemental Figure S16: mRNA secondary structure of the coding regions of *psbA*, *rbcL*, and *psaA*, Supplemental Table S1: Number of mapped reads from the DMS-MaPseq analysis, Supplemental Table S2: Strength of binding of Shine-Dalgarno sequences to anti-Shine-Dalgarno sequences, Supplemental Table S3: Fold change of mRNA levels and translation efficiency of nuclear-encoded genes encoding factors possibly regulating plastid translation, Supplemental Table S4: List of used oligonucleotides used and their sequences.

Author Contributions: Conceptualization, P.G., C.E., D.L., P.E.J., J.V., and L.B.S.; formal analysis, P.G., C.E., and J.V.; investigation, P.G., C.E., P.K., and L.B.S.; writing—original draft preparation, P.G., C.E., and L.B.S.; writing—review and editing, all authors; visualization, P.G. and C.E.; supervision, S.M., S.K., P.E.J., and J.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Polish National Science Centre (Narodowe Centrum Nauki; grants SONATA12, UMO-2016/23/D/NZ3/02491 and MAESTRO6 2014/14/A/NZ1/00218), the H2020 European Research Council (ERC) (grant StG2017-757411), the H2020 Marie Skłodowska-Curie Actions (grant MSCA-IF 703085), a Hallas-Møller Investigator award by the Novo Nordisk Foundation (Novo Nordisk Fonden; grant NNF15OC0014202), a Copenhagen Plant Science Centre Young Investigator Starting Grant, the German Science Foundation (Deutsche Forschungsgemeinschaft; grant TR175), the Independent Research Fund Denmark (Danmarks Frie Forskningsfond; grant 7014-00322B) and the VILLUM Foundation (Villum Fonden; grant 13363).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All sequences were deposited in the Sequence Read Archive under BioProject number PRJNA699408.

Acknowledgments: We thank Paul Hardy and Stefanie Zintl for commenting on the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

- de Smit, M.H.; van Duin, J. Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 7668–7672.
- Kudla, G.; Murray, A.W.; Tollervey, D.; Plotkin, J.B. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **2009**, *324*, 255–258.
- Goodman, D.B.; Church, G.M.; Kosuri, S. Causes and effects of N-terminal codon bias in bacterial genes. *Science* **2013**, *342*, 475–479.
- Mustoe, A.M.; Busan, S.; Rice, G.M.; Hajdin, C.E.; Peterson, B.K.; Ruda, V.M.; Kubica, N.; Nutiu, R.; Baryza, J.L.; Weeks, K.M. Pervasive Regulatory Functions of mRNA Structure Revealed by High-Resolution SHAPE Probing. *Cell* **2018**, *173*, 181–195.e18.
- Bhattacharyya, S.; Jacobs, W.M.; Adkar, B.V.; Yan, J.; Zhang, W.; Shakhnovich, E.I. Accessibility of the Shine-Dalgarno Sequence Dictates N-Terminal Codon Bias in *E. coli*. *Mol. Cell* **2018**, *70*, 894–905.e5.
- Scharff, L.B.; Childs, L.; Walther, D.; Bock, R. Local absence of secondary structure permits translation of mRNAs that lack ribosome-binding sites. *PLoS Genet.* **2011**, *7*, e1002155.
- Nakagawa, S.; Niimura, Y.; Gojobori, T. Comparative genomic analysis of translation initiation mechanisms for genes lacking the Shine-Dalgarno sequence in prokaryotes. *Nucleic Acids Res.* **2017**, *45*, 3922–3931.
- Scharff, L.B.; Ehrnthaler, M.; Janowski, M.; Childs, L.H.; Hasse, C.; Gremmels, J.; Ruf, S.; Zoschke, R.; Bock, R. Shine-Dalgarno Sequences Play an Essential Role in the Translation of Plastid mRNAs in Tobacco. *Plant Cell* **2017**, *29*, 3085–3101.
- Wen, J.-D.; Lancaster, L.; Hodges, C.; Zeri, A.-C.; Yoshimura, S.H.; Noller, H.F.; Bustamante, C.; Tinoco, I. Following translation by single ribosomes one codon at a time. *Nature* **2008**, *452*, 598–603.
- Tuller, T.; Veksler-Lublinsky, I.; Gazit, N.; Kupiec, M.; Rupp, E.; Ziv-Ukelson, M. Composite effects of gene determinants on the translation speed and density of ribosomes. *Genome Biol.* **2011**, *12*, R110.

11. Gawroński, P.; Jensen, P.E.; Karpiński, S.; Leister, D.; Scharff, L.B. Pausing of Chloroplast Ribosomes Is Induced by Multiple Features and Is Linked to the Assembly of Photosynthetic Complexes. *Plant Physiol.* **2018**, *176*, 2557–2569.
12. Breaker, R.R. Riboswitches and Translation Control. *Cold Spring Harb. Perspect. Biol.* **2018**, *10*, a032797.
13. Neupert, J.; Karcher, D.; Bock, R. Design of simple synthetic RNA thermometers for temperature-controlled gene expression in *Escherichia coli*. *Nucleic Acids Res.* **2008**, *36*, e124.
14. Krajewski, S.S.; Narberhaus, F. Temperature-driven differential gene expression by RNA thermosensors. *Biochim. Biophys. Acta* **2014**, *1839*, 978–988.
15. Laursen, B.; Sørensen, H. Initiation of protein synthesis in bacteria. *Microbiol.* **2005**, *236*, 747–771.
16. Duval, M.; Simonetti, A.; Caldelari, I.; Marzi, S. Multiple ways to regulate translation initiation in bacteria: Mechanisms, regulatory circuits, dynamics. *Biochimie* **2015**, *114*, 18–29.
17. Stampacchia, O.; Girard-Bascou, J.; Zanasco, J.L.; Zerges, W.; Bennoun, P.; Rochaix, J.-D. A nuclear-encoded function essential for translation of the chloroplast *psaB* mRNA in *Chlamydomonas*. *Plant Cell* **1997**, *9*, 773–782.
18. Klinkert, B.; Elles, I.; Nickelsen, J. Translation of chloroplast *psbD* mRNA in *Chlamydomonas* is controlled by a secondary RNA structure blocking the AUG start codon. *Nucleic Acids Res.* **2006**, *34*, 386–394.
19. Prikryl, J.; Rojas, M.; Schuster, G.; Barkan, A. Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 415–420.
20. Hammani, K.; Cook, W.B.; Barkan, A. RNA binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5651–5656.
21. Barkan, A. Expression of Plastid Genes: Organelle-Specific Elaborations on a Prokaryotic Scaffold. *Plant Physiol.* **2011**, *155*, 1520–1532.
22. Sugiura, M. Plastid mRNA translation. *Methods Mol. Biol.* **2014**, *1132*, 73–91.
23. Sun, Y.; Zerges, W. Translational regulation in chloroplasts for development and homeostasis. *Biochim. Biophys. Acta* **2015**, *1847*, 809–820.
24. Zoschke, R.; Bock, R. Chloroplast Translation: Structural and Functional Organization, Operational Control, and Regulation. *Plant Cell* **2018**, *30*, 745–770.
25. Hirose, T.; Sugiura, M. Functional Shine-Dalgarno-Like Sequences for Translational Initiation of Chloroplast mRNAs. *Plant Cell Physiol.* **2004**, *45*, 114–117.
26. Zhang, J.; Ruf, S.; Hasse, C.; Childs, L.; Scharff, L.B.; Bock, R. Identification of cis-elements conferring high levels of gene expression in non-green plastids. *Plant J.* **2012**, *72*, 115–128.
27. Gawroński, P.; Pałac, A.; Scharff, L.B. Secondary Structure of Chloroplast mRNAs In Vivo and In Vitro. *Plants* **2020**, *9*, 323.
28. Fu, H.-Y.; Ghandour, R.; Ruf, S.; Zoschke, R.; Bock, R.; Schöttler, M.A. The availability of neither D2 nor CP43 limits the biogenesis of photosystem II in tobacco. *Plant Physiol.* **2020**.
29. Chotewutmontri, P.; Barkan, A. Multilevel effects of light on ribosome dynamics in chloroplasts program genome-wide and *psbA*-specific changes in translation. *PLoS Genet.* **2018**, *14*, e1007555.
30. Schuster, M.; Gao, Y.; Schöttler, M.A.; Bock, R.; Zoschke, R. Limited Responsiveness of Chloroplast Gene Expression during Acclimation to High Light in Tobacco. *Plant Physiol.* **2020**, *182*, 424–435.
31. Mulo, P.; Sakurai, I.; Aro, E.-M. Strategies for *psbA* gene expression in cyanobacteria, green algae and higher plants: From transcription to PSII repair. *Biochim. Biophys. Acta* **2012**, *1817*, 247–257.
32. Li, L.; Aro, E.-M.; Millar, A.H. Mechanisms of Photodamage and Protein Turnover in Photoinhibition. *Trends Plant Sci.* **2018**, *23*, 667–676.
33. Danon, A.; Mayfield, S.P.Y. Light regulated translational activators: Identification of chloroplast gene specific mRNA binding proteins. *EMBO J.* **1991**, *10*, 3993–4001.
34. Schult, K.; Meierhoff, K.; Paradies, S.; Töller, T.; Wolff, P.; Westhoff, P. The nuclear-encoded factor HCF173 is involved in the initiation of translation of the *psbA* mRNA in *Arabidopsis thaliana*. *Plant Cell* **2007**, *19*, 1329–1346.
35. Link, S.; Meierhoff, K.; Westhoff, P. The Atypical Short-Chain Dehydrogenases HCF173 and HCF244 are Jointly Involved in Translational Initiation of the *psbA* mRNA of *Arabidopsis thaliana*. *Plant Physiol.* **2012**, *160*, 2202–2218.
36. McDermott, J.J.; Watkins, K.P.; Williams-Carrier, R.; Barkan, A. Ribonucleoprotein Capture by in Vivo Expression of a Designer Pentatricopeptide Repeat Protein in *Arabidopsis*. *Plant Cell* **2019**, *31*, 1723–1733.
37. Chotewutmontri, P.; Williams-Carrier, R.; Barkan, A. Exploring the Link between Photosystem II Assembly and Translation of the Chloroplast *psbA* mRNA. *Plants* **2020**, *9*, 152.
38. Chotewutmontri, P.; Barkan, A. Light-induced *psbA* translation in plants is triggered by photosystem II damage via an assembly-linked autoregulatory circuit. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 21775–21784.
39. Klughammer, C.; Schreiber, U. Complementary PS II quantum yields calculated from simple fluorescence parameters measured by PAM fluorometry and the Saturation Pulse method. *PAM Appl. Notes* **2008**, *1*, 27–35.
40. Barkan, A. Nuclear Mutants of Maize with Defects in Chloroplast Polysome Assembly Have Altered Chloroplast RNA Metabolism. *Plant Cell* **1993**, *5*, 389–402.
41. Oh, E.; Becker, A.H.; Sandikci, A.; Huber, D.; Chaba, R.; Gloge, F.; Nichols, R.J.; Typas, A.; Gross, C.A.; Kramer, G.; et al. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* **2011**, *147*, 1295–1308.
42. Zoschke, R.; Watkins, K.P.; Barkan, A. A rapid ribosome profiling method elucidates chloroplast ribosome behavior in vivo. *Plant Cell* **2013**, *25*, 2265–2275.

43. Marquardt, S.; Escalante-Chong, R.; Pho, N.; Wang, J.; Churchman, L.S.; Springer, M.; Buratowski, S. A chromatin-based mechanism for limiting divergent noncoding transcription. *Cell* **2014**, *157*, 1712–1723.
44. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq—Versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11.
45. Roy, S.; Ueda, M.; Kadowaki, K.; Tsutsumi, N. Different status of the gene for ribosomal protein S16 in the chloroplast genome during evolution of the genus *Arabidopsis* and closely related species. *Genes Genet. Syst.* **2010**, *85*, 319–326.
46. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21.
47. Zhong, Y.; Karaletsos, T.; Drewe, P.; Sreedharan, V.T.; Kuo, D.; Singh, K.; Wendel, H.-G.; Rättsch, G. RiboDiff: Detecting changes of mRNA translation efficiency from ribosome footprints. *Bioinformatics* **2017**, *33*, 139–141.
48. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **2017**, *14*, 417–419.
49. Sonesson, C.; Love, M.I.; Robinson, M.D. Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000Research* **2015**, *4*, 1521.
50. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550.
51. Loizeau, K.; Qu, Y.; Depp, S.; Fiechter, V.; Ruwe, H.; Lefebvre-Legendre, L.; Schmitz-Linneweber, C.; Goldschmidt-Clermont, M. Small RNAs reveal two target sites of the RNA-maturation factor Mbb1 in the chloroplast of *Chlamydomonas*. *Nucleic Acids Res.* **2014**, *42*, 3286–3297.
52. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359.
53. Morgan, M.; Pages, H.; Obenchain, V.; Hayden, N. *Rsamtools: Binary Alignment (BAM), FASTA, Variant Call (BCF), and Tabix File Import*; 2017.
54. Mathews, D.H.; Turner, D.H.; Watson, R.M. RNA Secondary Structure Prediction. *Curr. Protoc. Nucleic Acid Chem.* **2016**, *67*, 11.2.1–11.2.19.
55. Darty, K.; Denise, A.; Ponty, Y. VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* **2009**, *25*, 1974–1975.
56. Ahmed, T.; Shi, J.; Bhushan, S. Unique localization of the plastid-specific ribosomal proteins in the chloroplast ribosome small subunit provides mechanistic insights into the chloroplastic translation. *Nucleic Acids Res.* **2017**, *45*, 8581–8595.
57. Flynn, R.A.; Zhang, Q.C.; Spitale, R.C.; Lee, B.; Mumbach, M.R.; Chang, H.Y. Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat. Protoc.* **2016**, *11*, 273–290.
58. Poulsen, L.D.; Kielpinski, L.J.; Salama, S.R.; Krogh, A.; Vinther, J. SHAPE Selection (SHAPES) enrich for RNA structure signal in SHAPE sequencing-based probing data. *RNA* **2015**, *21*, 1042–1052.
59. Takahashi, H.; Lassmann, T.; Murata, M.; Carninci, P. 5′ end-centered expression profiling using cap-analysis gene expression and next-generation sequencing. *Nat. Protoc.* **2012**, *7*, 542–561.
60. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **2011**, *17*, 10.
61. Kielpinski, L.J.; Sidiropoulos, N.; Vinther, J. Reproducible Analysis of Sequencing-Based RNA Structure Probing Data with User-Friendly Tools. *Methods Enzymol.* **2015**, *558*, 153–180.
62. Starmer, J.; Stomp, A.; Vouk, M.; Bitzer, D. Predicting Shine-Dalgarno sequence locations exposes genome annotation errors. *PLoS Comput. Biol.* **2006**, *2*, e57.
63. Cannone, J.J.; Subramanian, S.; Schnare, M.N.; Collett, J.R.; D’Souza, L.M.; Du, Y.; Feng, B.; Lin, N.; Madabusi, L.V.; Müller, K.M.; et al. The comparative RNA web (CRW) site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinform.* **2002**, *3*, 2.
64. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **2011**, *12*, 77.
65. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation Statistical Computing: Vienna, Austria, 2018.
66. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*, 2nd ed.; Use, R., Ed.; Springer International Publishing: Cham, Germany, 2016; ISBN 978-3-319-24275-0.
67. Mitchell, D.; Assmann, S.M.; Bevilacqua, P.C. Probing RNA structure in vivo. *Curr. Opin. Struct. Biol.* **2019**, *59*, 151–158.
68. Mustoe, A.M.; Lama, N.N.; Irving, P.S.; Olson, S.W.; Weeks, K.M. RNA base-pairing complexity in living cells visualized by correlated chemical probing. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 24574–24582.
69. Wells, S.E.; Hughes, J.M.; Igel, A.H.; Ares, M. Use of dimethyl sulfate to probe RNA structure in vivo. *Methods Enzymol.* **2000**, *318*, 479–493.
70. Ding, Y.; Tang, Y.; Kwok, C.K.; Zhang, Y.; Bevilacqua, P.C.; Assmann, S.M. In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **2014**, *505*, 696–700.
71. Zubradt, M.; Gupta, P.; Persad, S.; Lambowitz, A.M.; Weissman, J.S.; Rouskin, S. DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods* **2017**, *14*, 75–82.
72. Wang, Z.; Wang, M.; Wang, T.; Zhang, Y.; Zhang, X. Genome-wide probing RNA structure with the modified DMS-MaPseq in *Arabidopsis*. *Methods* **2019**, *155*, 30–40.
73. Spitale, R.C.; Crisalli, P.; Flynn, R.A.; Torre, E.A.; Kool, E.T.; Chang, H.Y. RNA SHAPE analysis in living cells. *Nat. Chem. Biol.*

- 2013, 9, 18–20.
74. Spitale, R.C.; Flynn, R.A.; Torre, E.A.; Kool, E.T.; Chang, H.Y. RNA structural analysis by evolving SHAPE chemistry. *Wiley Interdiscip. Rev. RNA* **2014**, *5*, 867–881.
 75. Merino, E.J.; Wilkinson, K.A.; Coughlan, J.L.; Weeks, K.M. RNA structure analysis at single nucleotide resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE). *J. Am. Chem. Soc.* **2005**, *127*, 4223–4231.
 76. McGinnis, J.L.; Dunkle, J.A.; Cate, J.H.D.; Weeks, K.M. The mechanisms of RNA SHAPE chemistry. *J. Am. Chem. Soc.* **2012**, *134*, 6617–6624.
 77. Ruwe, H.; Wang, G.; Gusewski, S.; Schmitz-Linneweber, C. Systematic analysis of plant mitochondrial and chloroplast small RNAs suggests organelle-specific mRNA stabilization mechanisms. *Nucleic Acids Res.* **2016**, *44*, 7406–7417.
 78. Kwok, C.K.; Ding, Y.; Tang, Y.; Assmann, S.M.; Bevilacqua, P.C. Determination of in vivo RNA structure in low-abundance transcripts. *Nat. Commun.* **2013**, *4*, 2971.
 79. Talkish, J.; May, G.; Lin, Y.; Woolford, J.L.; McManus, C.J. Mod-seq: High-throughput sequencing for chemical probing of RNA structure. *RNA* **2014**, *20*, 713–720.
 80. Spitale, R.C.; Flynn, R.A.; Zhang, Q.C.; Crisalli, P.; Lee, B.; Jung, J.-W.; Kuchelmeister, H.Y.; Batista, P.J.; Torre, E.A.; Kool, E.T.; et al. Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* **2015**, *519*, 486–490.
 81. Kenyon, J.C.; Prestwood, L.J.; Lever, A.M.L. A novel combined RNA-protein interaction analysis distinguishes HIV-1 Gag protein binding sites from structural change in the viral RNA leader. *Sci. Rep.* **2015**, *5*, 14369.
 82. Ingolia, N.T.; Ghaemmaghami, S.; Newman, J.R.S.; Weissman, J.S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **2009**, *324*, 218–223.
 83. Jin, H.; Fu, M.; Duan, Z.; Duan, S.; Li, M.; Dong, X.; Liu, B.; Feng, D.; Wang, J.; Peng, L.; et al. LOW PHOTOSYNTHETIC EFFICIENCY 1 is required for light-regulated photosystem II biogenesis in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E6075–E6084.
 84. Wittenberg, G.; Levitan, A.; Klein, T.; Dangoor, I.; Keren, N.; Danon, A. Knockdown of the Arabidopsis thaliana chloroplast protein disulfide isomerase 6 results in reduced levels of photoinhibition and increased D1 synthesis in high light. *Plant J.* **2014**, *78*, 1003–1013.
 85. Williams-Carrier, R.; Brewster, C.; Belcher, S.E.; Rojas, M.; Chotewutmontri, P.; Ljungdahl, S.; Barkan, A. The Arabidopsis pentatricopeptide repeat protein LPE1 and its maize ortholog are required for translation of the chloroplast psbJ RNA. *Plant J.* **2019**, *99*, 56–66.
 86. Hey, D.; Grimm, B. ONE-HELIX PROTEIN2 (OHP2) Is Required for the Stability of OHP1 and Assembly Factor HCF244 and Is Functionally Linked to PSII Biogenesis. *Plant Physiol.* **2018**, *177*, 1453–1472.
 87. Myouga, F.; Takahashi, K.; Tanaka, R.; Nagata, N.; Kiss, A.Z.; Funk, C.; Nomura, Y.; Nakagami, H.; Jansson, S.; Shinozaki, K. Stable Accumulation of Photosystem II Requires ONE-HELIX PROTEIN1 (OHP1) of the Light Harvesting-Like Family. *Plant Physiol.* **2018**, *176*, 2277–2291.
 88. Verhounig, A.; Karcher, D.; Bock, R. Inducible gene expression from the plastid genome by a synthetic riboswitch. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 6204–6209.