CRP1 Protein: (dis)similarities between Arabidopsis thaliana and Zea mays

Roberto Ferrari††, Luca Tadini††, Fabio Moratti§§, Marie-Kristin Lehniger‡, Alex Costa†, Fabio Rossi¶, Monica Colombo∥‖, Simona Masiero†, Christian Schmitz-Linneweber‡ and Paolo Pesaresi∗

†† Dipartimento di Bioscienze, Università degli studi di Milano, Milano, Italy, ‡ Max-Planck-Institut für Molekulare Pflanzenphysiologie, Potsdam-Golm, Germany, §§ Molecular Genetics, Institute of Biology, Humboldt University of Berlin, Berlin, Germany, ‡‡ Dipartimento di Biotecnologie Mediche e Medicina Traslazionale, Università degli studi di Milano, Milano, Italy, † Centre Ricerca e Innovazione, Fondazione Edmund Mach, San Michele all’Adige, Italy, †† Dipartimento di Scienze Agrarie e Ambientali - Produzione, Territorio, Agroenergia, Università degli studi di Milano, Milano, Italy

Biogenesis of chloroplasts in higher plants is initiated from proplastids, and involves a series of processes by which a plastid able to perform photosynthesis, to synthesize amino acids, lipids, and phytohormones is formed. All plastid protein complexes are composed of subunits encoded by the nucleus and chloroplast genomes, which require a coordinated gene expression to produce the correct concentrations of organellar proteins and to maintain organelle function. To achieve this, hundreds of nucleus-encoded factors are imported into the chloroplast to control plastid gene expression. Among these factors, members of the Pentatricopeptide Repeat (PPR) containing protein family have emerged as key regulators of the organellar post–transcriptional processing. PPR proteins represent a large family in plants, and the extent to which PPR functions are conserved between dicots and monocots deserves evaluation, in light of differences in photosynthetic metabolism (C3 vs. C4) and localization of chloroplast biogenesis (mesophyll vs. bundle sheath cells). In this work we investigated the role played in the process of chloroplast biogenesis by At5g42310, a member of the Arabidopsis PPR family which we here refer to as AtCRP1 (Chloroplast RNA Processing 1), providing a comparison with the orthologous ZmCRP1 protein from Zea mays. Loss-of-function atcrp1 mutants are characterized by yellow-albinotic cotyledons and leaves owing to defects in the accumulation of subunits of the thylakoid protein complexes. As in the case of ZmCRP1, AtCRP1 associates with the 5′ UTRs of both psaC and, albeit very weakly, petA transcripts, indicating that the role of CRP1 as regulator of chloroplast protein synthesis has been conserved between maize and Arabidopsis. AtCRP1 also interacts with the petB-petD intergenic region and is required for the generation of petB and petD monocistronic RNAs. A similar role has been also attributed to ZmCRP1, although the direct interaction of ZmCRP1 with the petB-petD intergenic region has never been reported, which could indicate that AtCRP1 and ZmCRP1 differ, in part, in their plastid RNA targets.

Keywords: PPR, anterograde signaling, chloroplast, biogenesis, RNA metabolism
INTRODUCTION

In land-plants, nuclear-encoded pentatricopeptide repeat (PPR) containing proteins constitute a large family, which regulates organelle gene expression at the RNA level (Lurin et al., 2004; O’Toole et al., 2008; Barkan and Small, 2014). They are, indeed, a major constituent of the genome-coordinating anterograde signaling pathway that evolved to adapt the expression of the organellar genomes in response to endogenous and environmental stimuli that are perceived by the nucleus (Woodson and Chory, 2008).

A typical PPR motif is characterized by a degenerate 35-amino acid repeat that folds into two antiparallel alpha helices (Small and Peeters, 2000). PPR proteins contain a tandem array of 2–30 PPR motifs, which stack together to form a superhelix with a central groove that allows the protein to bind RNA (Lurin et al., 2004; Rivals et al., 2006). According to the characteristics of their repeats, PPR proteins are generally classified into P and PLS sub-families. The P-type proteins are implicated in the determination and stabilization of 5′ and/or 3′ RNA termini, RNA splicing and translation of specific RNAs in chloroplasts and mitochondria, while PLS-type proteins are generally involved in RNA editing (Barkan and Small, 2014). Higher plants harbor several hundreds of PPR proteins, which generally have distinct, non-redundant functions in organelle biogenesis, plant growth and development and adaptation to environmental cues (Barkan and Small, 2014; Manna, 2015), as revealed by the high number of ppr mutants with distinct phenotypes. This is due to their ability to recognize primary RNA sequences, with each protein having different target sites, thus implying that the elucidation of the primary role of each PPR protein is greatly facilitated by the identification of its RNA targets.

The detection of few native PPR-RNA interactions through RNA immunoprecipitation on microarray (RIP-Chip) analyses and in vitro binding assays using PPR recombinant proteins, together with PPR crystal structures indicate that PPR proteins bind their cognate RNA targets in a sequence specific manner (Meierhoff et al., 2003; Schmitz-Linneweber et al., 2005, 2006; Williams-Carrier et al., 2008; Yin et al., 2013; Okuda et al., 2014; Shen et al., 2016). The code describing how PPR proteins recognize specific nucleotides of their RNA targets relies primarily on two amino acids that are within a single PPR motif, specifically the fifth residue in the first helix and the last residue on the loop interconnecting adjacent motifs (Barkan et al., 2012; Yin et al., 2013; Cheng et al., 2016). However, the current understanding of the code does not allow accurate large-scale computational predictions of PPR targets (Takenaka et al., 2013; Kindgren et al., 2015; Hall, 2016; Harrison et al., 2016). Predictive power is constrained by the fact that the code is degenerate and by the low accuracy of current methods used for the identification of PPR domains, which in turn leads to mismatches in the amino acid/nucleotide alignments. However, a more robust annotation of PPR domains has recently been conducted and made available at the PlantPPR database1 (Cheng et al., 2016). Furthermore, more PPR-RNA interactions as well as crystal structures of PPR-RNA complexes need to be characterized in different species in order to improve the understanding of the code. This would also help to determine if the amino acid sequences of the PPR domains coevolved with the nucleotide sequences of their RNA targets and ultimately to determine whether there is functional conservation of PPR proteins among land plants.

The function of PPR proteins, and more generally the function of the nuclear gene complement involved in organelar RNA metabolism, have been primarily studied in maize, since the large seed reserves of maize support rapid heterotrophic growth of non-photosynthetic mutants and provide ready access to non-photosynthetic tissues for molecular biology and biochemical studies (Belcher et al., 2015). However, the degree of functional conservation of PPR proteins between maize and other species, including Arabidopsis thaliana, has yet to be investigated. The question is of particular interest since the elaboration of the thylakoid membrane system and the biogenesis of the multi-subunit photosynthetic complexes appear to have major differences between monocotyledonous and dicotyledonous plants (Pogson et al., 2015). Indeed in maize, and more generally in monocots, the process of chloroplast development from the proplastid to functional chloroplasts can be observed as a gradient along the leaf blade, whereas in dicots, such as Arabidopsis thaliana, the development of chloroplasts differs between developmental stages, plant organs – i.e., chloroplast development is different in cotyledons and leaves – and plant tissues (Pogson and Albrecht, 2011; Jarvis and Lopez-Juez, 2013).

The majority of PPR proteins are conserved at sequence level between dicots (Arabidopsis) and monocots (rice) (O’Toole et al., 2008). Orthologous pairs can readily be identified and in a number of cases, primary sequence conservation can be traced back to the roots of all embryophytes (O’Toole et al., 2008). As a matter of fact, functional differences between orthologous PPR proteins of maize and Arabidopsis have been observed. For example, the molecular phenotypes resulting from loss of the orthologous PPR proteins ATP4 (maize) and SVR7 (Arabidopsis) differ substantially (Liu et al., 2010; Zoschke et al., 2012, 2013a,b), as do the molecular defects in maize and Arabidopsis mutants lacking the PGR3 protein (Yamazaki et al., 2004; Cai et al., 2011; Belcher et al., 2015). Thus, the extent to which lessons on PPR proteins learnt from maize can be extrapolated to dicots, such as Arabidopsis, and more broadly to other organisms, needs further investigation.

In this context, we investigated here the function of and identified the RNA targets of the PPR protein At5g42310 from Arabidopsis thaliana, that shares high similarity with the well-characterized CRP1 (Chloroplast RNA Processing 1) protein from maize (ZmCRP1), and which we here refer to as AtCRP1. Our findings indicate that AtCRP1, like the orthologous ZmCRP1 (Barkan et al., 1994; Fisk et al., 1999; Schmitz-Linneweber et al., 2005), is essential for plant autotrophy since it plays a direct role in the accumulation of the cytochrome b6/f (Cyt b6/f) complex and of the PsAC subunit.

1http://www.plantppr.com
of photosystem I (PSI). Furthermore AtCRP1, similarly to ZmCRP1, is required for the accumulation of petB and petD monocistronic RNAs, indicating that the functional roles of CRP1 proteins are highly conserved between monocots and dicots.

**MATERIALS AND METHODS**

**Plant Material and Growth Conditions**

Arabidopsis thaliana atcrp1-1 (SALK_035048) (Alonso et al., 2003) and atcrp1-2 (SAIL_916A02) (Sessions et al., 2002) T-DNA insertion lines were identified by searching the T-DNA Express database. For promoter analyses, the putative AtCRP1 promoter region (AtCRP1p, −1062 to −2 upstream the translation starting codon) was cloned into pBGWFS7 destination vector and introduced into Arabidopsis wild type background, ecotype Columbia-0 (Col-0), by Agrobacterium tumefaciens-mediated transformation. AtCRP1-GFP transgenic lines were obtained by transformation of AtCRP1/atcrp1-1 heterozygous plants with either the AtCRP1 coding sequence fused to GFP under the control of 35S-CaMV promoter, cloned into pB7FWG2 vector, or the genomic locus fused to GFP under the control of the native promoter, cloned into a modified pGreenII vector (Gregis et al., 2009). The GUN1 coding sequence, devoid of the stop codon, was cloned into pB7RWG2 vector, carrying an RFP reporter gene. pB7FWG2, pBGWFS7, and pB7RWG2 plasmids were obtained from Flanders Interuniversity Institute for Biotechnology of Gent (Karimi et al., 2002). Primers used for amplification of the DNA fragments cloned into the vectors, reported above, are listed in Supplementary Table S2. Arabidopsis Col-0 and mutant plants were grown on soil under controlled growth chamber conditions with a 16 h light/8 h dark cycle at 22°C/18°C. In the case of mesophyll protoplast preparation, Arabidopsis plants were also grown on soil in a growth chamber under the above reported conditions. Moreover, phenotypic characterization and molecular biology analyses were also conducted on plants grown on Murashige and Skoog (MS) medium (Duchefa)1, supplemented with or without 1% (w/v) sucrose. Tobacco plants, employed for transient gene expression, were cultivated for 5–6 weeks in a greenhouse under a 12 h light/12 h dark cycle at 22°C/18°C.

**Protoplast Transformation**

Mesophyll protoplasts of Arabidopsis thaliana (Col-0) were isolated and transiently transformed according to Yoo et al. (2007) and Costa et al. (2012). Briefly, well-expanded rosette leaves from 3-to-5 week-old plants were cut into strips of 0.5–1 mm with a fresh razor blade. Leaf tissue was digested using an enzyme solution containing 1.25% cellulase Onozuka R-10 (Duchefa) and 0.3% Macerozyme R-10 (Duchefa) for 3 h at 23°C in the dark. The protoplast suspension was filtered through a 50 μm nylon mesh washed three times with W5 solution (154 mM NaCl, 125 mM CaCl2, 5 mM KCl, 2 mM MES, pH 5.7 adjusted with KOH) and used for PEG-mediated transformation. For each protoplast transformation 10 μg of a MidiPrep purified DNA (QIAGEN) plasmid harboring the 35S-CaMV::AtCRP1-GFP cassette was used. Protoplasts were maintained for 16–24 h at 23°C in the dark, before performing epifluorescent microscopy.

**Transient Expression in Nicotiana benthamiana Leaves**

Tobacco leaf infiltration was performed using A. tumefaciens strain GV3101/pMP90 carrying the specified constructs (see results for details) together with the p19-enhanced expression system (Voinnet et al., 2003), according to the method described by Waadt and Kudla (2008). The final OD600 for A. tumefaciens strains harboring 35S-CaMV::AtCRP1-GFP and 35S-CaMV::GUN1-RFP was 0.2 and 0.3, respectively. After infiltration, plants were incubated for 3–5 days under the conditions described above.

**Confocal Microscopy Analysis**

Confocal Scanning Laser Microscopy analyses were performed using an inverted microscope, Leica DMIRE2, equipped with a Leica TCS SP2 laser scanning device (Leica). For the simultaneous detection of GFP and chlorophyll autofluorescence the cells were excited (Arabidopsis mesophyll protoplasts or tobacco leaf cells) with the 488 nm line of the Argon laser and the emissions were collected between 515/535 and 650/750 nm, respectively. For RFP detection the cells were excited at 561 nm from a He/Ne laser and the emission was collected between 575/625 nm. Image analyses were performed with Fiji4, an open-source platform for biological-image analysis (Schindelin et al., 2012).

**Nucleic Acid Analyses**

Arabidopsis DNA was isolated according to Ihnatowicz et al. (2004). Isolation of total RNA from homozygous atcrp1-1 plants at four-leaf rosette stage and RNA gel blot analyses were performed as described by Meurer et al. (2002), using 10 μg of total RNA for each sample. For the RNA slot blot hybridization experiments, one-fourth of the RNA purified from each immunoprecipitation pellet and one-tenth of the RNA purified from the corresponding supernatant were applied to a nylon membrane with a slot-blot manifold and hybridized to specific radiolabeled probes (see Supplementary Table S2). 32P-labeled DNA probes, complementary to chloroplast genes, were amplified using the primer pairs listed in Supplementary Table S2. Four micrograms of total RNA, treated with TURBO DNA-free (Ambion by Life Technologies), were employed for first-strand cDNA synthesis using GoScript Reverse Transcription System (Promega) according to the supplier’s instructions. Quantitative Real-Time PCR (qRT-PCR) was carried out on an CFX96 Real-Time system (Bio-Rad), using the primer pairs

---

1http://signal.salk.edu/cgi-bin/tdnaexpress
2http://www.duchefa.com
3http://www.fiji.sc/
reported in Supplementary Table S2. The SAND (Remans et al., 2008) and ubiquitin transcripts were used as internal references. Data from three biological and three technical replicates were analyzed with Bio-Rad CFX Manager software (V3.1).

Immunoblot Analyses
For immunoblot analyses, total proteins were prepared as described by Martinez-Garcia et al. (1999). Total proteins, corresponding to 5 mg of leaf fresh-weight (100% of WT and atrcr1-1 samples) and isolated from plants at four-leaf rosette stage, were fractionated by SDS–PAGE (12% acrylamide [w/v]; (Schagger and von Jagow, 1987). Proteins were then transferred to polyvinylidene difluoride (PVDF) membranes (Ihnatowicz et al., 2004) and replicate filters were immunodecorated with antibodies specific for PSI (PsaA, PsaC, and PsaD), PSII (D1, PsbO) Cyt b6/f (PetA, PetB, and PetC), ATPase (ATPase-β) subunits, PSI (Lhca1, Lhca2) and PSII (Lhcb2, Lhcb3) antenna proteins, all obtained from Agrisera3. The GFP antibody was purchased from Life Technologies4.

Chloroplast Stomatal Preparation and Protein Immunoprecipitation
Intact chloroplasts were isolated from 11 days old Arabidopsis plants, according to Kunst (1998), and Kupsch et al. (2012) with some modifications. Chloroplasts were directly resuspended in 300–400 µl of extraction buffer [2 mM DTT, 30 mM HEPES-KOH, pH 8.0, 60 mM KOAc, 10 mM MgOAc and proteinase inhibitor cocktail (Sigma–Aldrich-P9599)]. Two independent stromal preparations were carried out and one of them was performed in the presence of 2% sodium deoxycholate in order to solubilize the membrane-attached AtCRP1 protein fraction. Chloroplasts were then disrupted by pulling them through a syringe (0.55 mm × 40 mm) 30–40 times. The solution was centrifuged at 21,000 × g at 4°C to separate the stromal from the membrane fraction.

The isolated stromal fraction was diluted with one volume of coimmunoprecipitation (CoIP) buffer (150 mM NaCl, 20 mM Tris-HCl pH 7.5, 2 mM MgCl₂, 0.5% Nonidet P-40 and 0.5 μg/mL Aprotinin). Five microliters of mouse anti-GFP antibody (Roche, No. 11814460001) were added to the stromal fraction and incubated for 1 h at 4°C and 13 rpm on an overhead shaker. Thereafter the coimmunoprecipitation was performed as described by Kupsch et al. (2012). Successful precipitation of AtCRP1-GFP was confirmed by immunoblot analyses, using the same GFP antibody.

RNA Extraction and Labeling for RIP-Chip Assay
RNA immunoprecipitation-chip analyses were performed using a tiling microarray covering the complete Arabidopsis chloroplast genome (Kupsch et al., 2012). The coimmunoprecipitated RNA was isolated from pellet and supernatant fractions either by phenol-chloroform extraction or using the Direct-zol™ RNA MiniPrep kit (Zymo Research). For the phenol-chloroform extraction, RNA samples were incubated in 1% SDS and 5 mM EDTA at room temperature for 5 min to dissociate RNA-protein complexes. RNA was phenol-chloroform extracted, ethanol precipitated with the addition of Glycoblue™ Coprecipitant (Thermo Fisher Scientific), washed twice with 75% ethanol, air-dried and resuspended in 20 µl RNase-free water. For the replicate, RNA was extracted using the Direct-zol™ RNA MiniPrep kit (Zymo Research) according to the manufacturer’s instructions. Before the extraction 2 µg yeast RNA was added to the coimmunoprecipitated RNA pellet. The entire RNA of the pellet fraction and 2 µg RNA of the supernatant fraction were used for labeling. The pellet and supernatant RNA were labeled with 0.5 µl Cy5 and 1 µl Cy3 dye, respectively (aRNA labeling kit, Kreatech Diagnostics). Labeling reaction, microarray hybridization, scanning, and evaluation were performed as described in Kupsch et al. (2012). Only PCR products for which more than half of all replicate spots (24 per PCR product spanning two experiments) passed our quality assessment (Kupsch et al., 2012) and were used in this analysis (Supplementary Table S1).

In silico Prediction of AtCRP1 Binding Sites
The putative AtCRP1 binding motif, i.e., the nucleotide preference for each of the amino acid pairs at the fifth and last position of PPR domains, was predicted in silico using the reported weighting schemes (Barkan et al., 2012; Barkan and Small, 2014; Harrison et al., 2016). The software FIMO5, which analyizes sequence databases for occurrences of known motifs (Grant et al., 2011), was employed to identify the potential binding sites of AtCRP1 within the regions enriched in our RIP-Chip experiment. Furthermore, the same regions were searched for the presence of sRNA native footprints, by consulting the JBrowse sRNA database6. Numbers that delimit the native footprints refer to the chloroplast genome of Arabidopsis thaliana (NC_000932.1).

β-Glucuronidase (GUS) Assay
For GUS histochemical detection, plant material was fixed in 90% acetone at −20°C for 1 h. Samples were then washed three times with NaPi buffer (NaH₂PO₄ 50 mM, Na₂HPO₄ 50 mM; pH 7.0) and stained overnight at 37°C with X-gluc solution [1 mM 5-bromo-4-chloro-3-indolyl-β-D-glucuronide, 2 mM K₃/K₄Fe(CN)₆, 0.1% Triton (v/v), 10 mM EDTA, 50 mM NaPi pH 7.0]. 70% EtOH (v/v) was used as washing solution. Stained samples were then stored at 4°C and observed using a Zeiss Axioshot D1 microscope equipped with differential interference contrast (DIC) optics. Images were recorded with an AxioCam MRC5 camera (Zeiss) using the Axiovision program (v.4.1).

3http://www.agrisera.com/en/artiklar/plantagal-cell-biology/index.html
4http://www.thermofisher.com
5http://meme-suite.org/tools/fimo
6https://www.molgen.hu-berlin.de/projects-jbrowse-athlonia.php
RESULTS

AtCRP1 Is a PPR Protein Imported into the Chloroplast

The Maize Genetics and Genomics Database (Lawrence et al., 2004) was used to identify the At5g42310 gene as the Arabidopsis ortholog of ZmCRP1 (see also Belcher et al., 2015). At5g42310 encodes a polypeptide of 709 amino acids with a calculated molecular mass of 80 kDa. Intron number (three) and position are conserved between the two genes, and BLASTP query of public Arabidopsis sequence database with ZmCRP1 amino acid sequence detected At5g42310 protein as the top hit with 55% sequence identity and 72% sequence similarity (Figure 1).

AtCRP1 is annotated as a PPR protein and shares with ZmCRP1 15 PPR tandem repeats, which were predicted by using the PlantPPR database (Cheng et al., 2016). All PPR motifs are of 35 aa, with the exception of P0 which consists of 30 aa and P2 of 37 aa in Arabidopsis and 38 aa in maize. The fifth and the last residue of each PPR domain form the amino acid pairs that specify the RNA target molecules (Cheng et al., 2016), and are labeled with gray and black stars in Figure 1. The ChloroP server (Emanuelsson et al., 1999)\(^9\) predicted the presence of a cTP of 54 residues (see amino acid residues in italics in Figure 1), indicating that AtCRP1, like ZmCRP1, could be imported into the chloroplast. To corroborate the in silico prediction, the AtCRP1-GFP fusion protein was expressed in transiently transformed Arabidopsis protoplasts (Figure 2). In agreement with the ChloroP prediction, the chimeric protein (GFP fluorescence) accumulated within the chloroplast in distinct fluorescent foci (CHL, autofluorescence of chloroplast chlorophylls, Figure 2A), resembling the nucleoid complexes. Indeed, AtCRP1-GFP chimera co-localized perfectly with the GUN1-RFP fusion protein, used as a nucleoid marker in this assay (RFP fluorescence, Figure 2B), (Koussevitzky et al., 2007; Colombo et al., 2016; Tadini et al., 2016), in tobacco leaf cells. To further localize AtCRP1, chloroplasts were fractionated to separate the stroma and thylakoid compartments. Immunoblot analysis, using a GFP specific antibody, allowed detection of AtCRP1-GFP specific signal in total chloroplasts, as well as in thylakoids and in the stromal fraction, indicating that the nucleoid AtCRP1 protein is both associated to membranes and soluble in the stroma (Figure 2C). These findings are in agreement with the identification of AtCRP1 as part of Megadalton complexes in the chloroplast stroma (Olinares et al., 2010), as well as in the grana of thylakoid membranes (Tomiziolli et al., 2014).

AtCRP1 Is Essential for Plant Autotrophy

To investigate the role that AtCRP1 plays in Arabidopsis, two lines carrying T-DNA insertions into the coding sequence of At5g42310, renamed atcrp1-1 (Salk_035048) and atcrp1-2
Both T-DNA insertions completely suppressed the accumulation of the corresponding transcripts in homozygous mutant seedlings (Figure 3B), which were characterized by a paler pigmentation of cotyledons, visible even at the fully mature embryo stage (Figure 4A), and leaves (Figures 4B,C), and found to be seedling lethal under autotrophic growth conditions on soil and MS medium without sucrose, but able to develop yellow-albino rosette leaves and sterile inflorescence when sucrose was provided in the medium (Figure 4C). The mutant phenotype could be rescued by Agrobacterium tumefaciens-mediated transformation of heterozygous plants with either the appropriate coding sequence fused to the 35S promoter of cauliflower mosaic virus (35S-CaMV::AtCRP1-GFP), or the genomic sequence including a 1-Kbp fragment of the promoter region (AtCRP1p::AtCRP1-GFP), corroborating a direct correspondence between genotype and phenotype, and indicating that the AtCRP1-GFP chimera was fully functional, in both cases (Figure 4D). Interestingly, complemented plants carrying the AtCRP1-GFP construct under the control of the native promoter showed a fivefold increase in AtCRP1 gene expression (Figure 4E), most probably as consequence of the T-DNA insertion in a highly expressed

(Sail_916A02), were obtained from the T-DNA Express Arabidopsis mutant collection (Figure 3A; see also Materials and Methods).
FIGURE 4 | Effects of loss of AtCRP1 on plant development. (A) Images of isolated fully mature embryos (bent cotyledon stage) from WT (Col-0) and atcrp1-1 seeds. The lack of AtCRP1 protein did not alter embryo development, although mutant embryos were slightly larger and paler than Col-0. (B) atcrp1-1 seeds were able to germinate on soil, giving rise to yellow seedlings that accounted for about one-quarter of all seedlings, indicative of a monogenic recessive trait. Mutant seedlings did not survive past the cotyledon stage. (C) Mutant seedlings showed albino cotyledons when grown on MS medium without sucrose and arrested at the cotyledon stage as in (B). However, when atcrp1-1 seedlings were grown on MS medium supplemented with 1% sucrose, they showed yellow-albino cotyledons at 8 days after sowing and were able to develop up to 8–10 true leaves after 35 days. (D) The atcrp1-1 seedling lethal phenotype could be fully rescued by Agrobacterium tumefaciens-mediated transformation of AtCRP1/atcrp1-1 heterozygous plants with either the AtCRP1 coding sequence fused to GFP under the control of 35S-CaMV promoter (35S::AtCRP1-GFP#12), or the genomic sequence fused to GFP under the control of native promoter (AtCRP1p::AtCRP1-GFP#5). (E) Real-time PCR to monitor the expression of AtCRP1 gene in WT and complemented plants. Gene expression was normalized with respect to the level of AtCRP1 transcripts in Col-0, and SAND and ubiquitin were used as internal references. The bars indicate standard deviations. (F) Col-0 and 35S::AtCRP1-GFP#16 transgenic line with about 15-folds more AtCRP1 transcripts than WT. In this case the transgenic line shows WT-like rosette, but it is characterized by shorter and paler stems, bleached cauline leaves, together with sterile flowers (Figures 4E,F).

euchromatin region of the nuclear genome. Furthermore, a complete rescue of mutant plant phenotype could only be observed in 35S::AtCRP1-GFP transgenic lines with a limited accumulation of AtCRP1 transcripts (Figures 4D,E). Higher AtCRP1 expression levels (around 15-folds in comparison to WT) led to transgenic plants with WT-like rosette but shorter and paler stems, bleached cauline leaves, together with sterile flowers (Figures 4E,F).

Temporal and spatial expression patterns of AtCRP1, monitored by fusing the promoter region of the gene upstream of the GUS reporter gene (see also Materials and Methods), support further the key role played by AtCRP1 during early stages of seedling and leaf development (Figure 5). The GUS staining could, indeed, be detected in young cotyledons and in the upper portion of the hypocotyl (Figure 5A). Furthermore, intense GUS signals were observable in young developing leaves.
in cotyledons and leaves using quantitative Real-Time PCR were also obtained by monitoring the expression of AtCRP1 with respect to the level of transcripts. Gene expression was normalized and used to calculate median enrichment ratios of the red and green fluorescence signals for each PCR product among the 24 replicate spots on two arrays. Ratios of the red and green fluorescence signals for each spot on the microarray. Two biological replicate experiments were performed with stroma from AtCRP1-GFP expressing plants and two with WT stroma. Data from the four assays were normalized and used to calculate median enrichment ratios of the red and green fluorescence signals for each PCR product among the 24 replicate spots on two arrays (Supplementary Table S1). To identify enrichment of RNA species specifically in the AtCRP1-GFP immunoprecipitation,(qRT-PCR). In general, a high level of expression of AtCRP1 was observed in green developing tissues, such as young cotyledons and leaves, whereas the expression decreased in older tissues (Figure 5E).

**atcrp1 Mutant Chloroplasts Fail to Accumulate Cytochrome b6/f Protein Complex and the PsaC Subunit of PSI**

The albino pigmentation of atcrp1 seedlings, together with their inability to grow under autotrophic conditions, indicated a defect in the thylakoid-associated photosynthetic apparatus. To verify this assumption, immunoblot analyses with antibodies specific for single subunits of the four major thylakoid protein complexes were performed on total leaf proteins. Leaf samples were harvested from atcrp1 plants at the four-leaf rosette stage and grown on MS-medium supplemented with 1% sucrose (Figure 6; see also Materials and Methods). Under standard light conditions (50 µmol photons m−2 s−1), subunits of Photosystem I (PsaA, PsaC, and PsaD), Photosystem II (D1, PsbO), Light harvesting complexes (Lhca1, Lhca2, Lhcb2, and Lhcb3) and ATPase (ATPase-β) accumulated to levels lower than 10% with respect to wild type plants. Furthermore, subunits of the Cyt b6/f (PetA, PetB, and PetC) and PSI (PsaC) were below the limits of immunoblot detection.

In summary, these results indicate a general reduction of thylakoid protein complex subunits in atcrp1 leaves, with a particularly severe effect on the accumulation of the Cyt b6/f complex and PsaC.

**AtCRP1 Is Associated In vivo with psaC and petB-petD Transcripts**

ZmCRP1 has been previously demonstrated to associate with the psaC and petA mRNAs in vivo by RIP-Chip analyses (Schmitz-Linneweber et al., 2005). To investigate whether AtCRP1 shares with ZmCRP1 the RNA targets, the same RIP-Chip approach employed in maize was used here. Stroma from plants expressing AtCRP1-GFP, under the control of the native promoter (AtCRP1p::AtCRP1-GFP), was isolated and the fusion protein was immunoprecipitated using an anti-GFP antibody. RNA was purified from the immunoprecipitation pellets and supernatants and was labeled with Cy5 (red) and Cy3 (green) fluorescent dyes, respectively. The two RNA fractions from AtCRP1-GFP immunoprecipitations (IPs) and from mock IPs were competitively hybridized to a chloroplast genome tiling microarray (Kupsch et al., 2012). Enrichment of RNA is reflected in the ratio of red to green fluorescence for each spot on the microarray. Two biological replicate experiments were performed with stroma from AtCRP1-GFP expressing plants and two with WT stroma. Data from the four assays were normalized and used to calculate median enrichment ratios of the red and green fluorescence signals for each PCR product among the 24 replicate spots on two arrays (Supplementary Table S1). To identify enrichment of RNA species specifically in the AtCRP1-GFP immunoprecipitation,
Ferrari et al. | AtCRP1-Mediated Chloroplast Biogenesis

we plotted the difference in median enrichment ratio for each DNA fragment between the AtCRP1-GFP mock experiment and the position of the product on the plastid chromosome (Figures 7A,B).

Four prominent peaks of differential enrichment were observed. One of them corresponds to the 5′UTR of psaC transcript, a target already recognized as a ligand of ZmCRP1 in RIP-Chip assays (Schmitz-Linneweber et al., 2005). A second RNA target is represented by the petB-petD intergenic region. This RNA was not identified to interact with ZmCRP1 by RIP-Chip analysis, however, ZmCRP1 is known to aid in maturation of this particular intergenic region (Barkan et al., 1994). Interestingly, the observed enrichment of rps15 transcripts might uncover a further, novel target of AtCRP1, whereas the enrichment of psbM/trnD transcripts is often observed in RIP-Chip experiments, thus this peak was considered an artifact.

To corroborate the RIP-Chip data, the AtCRP1-associated RNAs were analyzed by slot blots (Figure 7C). RNA purified from immunoprecipitation pellets and supernatants were probed with PCR fragments that detected the most highly enriched sequences in the RIP-Chip assay. The data confirmed that the psaC and petB-petD transcripts are highly enriched in the AtCRP1-GFP immunoprecipitates, but not the rps15 RNA. ZmCRP1 was also reported to be associated with RNAs of the petA region (Schmitz-Linneweber et al., 2005; Williams-Carrier et al., 2008), however, no enrichment of petA transcripts could be observed in the AtCRP1-GFP RIP-Chip assay (Figure 7A) and a low enrichment was detected in the slot blot assay (Figure 7C), possibly indicating that the interaction of AtCRP1 with petA transcripts is not very stable. In general, our analysis cannot exclude the possibility that CRP1 binds to additional target RNAs, for example when interactions take place at chloroplast membranes. Since we are not using cross-linked material, weak RNA-protein interactions might be lost during our assay.

To support further the RIP-Chip findings, AtCRP1 target RNAs were interrogated for the presence of native footprints at the JBrowse database. The JBrowse database provides annotations of Arabidopsis thaliana organellar short RNA (sRNA), thought to be generated from protein-mediated temporary protection of target RNAs against exonucleolytic degradation (Ruwe et al., 2016; see also Figure 8). sRNAs were found within the 5′UTR of psaC (corresponding to the 117633–117597 region of chloroplast genome) and the petB-petD intergenic region (region 76318–76358), and an sRNA was also annotated in the 5′UTR of petA (region 61615–61643). Furthermore, AtCRP1 predicted RNA binding motifs were shown to co-map with the native footprints, when the corresponding sequences were searched for the occurrence of the consensus binding motif with the FIMO program in the MEME suite (Figure 8B; Takenaka et al., 2013). A short RNA has been also mapped upstream of rps15, but this region was not enriched in the RIP-Chip assay and the match with the predicted binding site of AtCRP1 is weaker than for the psaC, petB-petD, and petA sRNAs.

---

11https://www.molgen.hu-berlin.de/projects-jbrowse-athaliana.php
12http://meme-suite.org/tools/fimo
FIGURE 7 | AtCRP1 RIP-Chip data plotted according to gene order within the plastid genome. (A) Differential enrichment ratios obtained by RNA immunoprecipitation (RIP)-Chip analysis. The enrichment ratios \( \frac{F_{635}}{F_{532}} \) obtained from an assay of AtCRP1p::AtCRP1-GFP chloroplast stroma extract were normalized with respect to a control assay that used WT (Col-0) chloroplast stroma extract (both assays were performed in duplicate). The median-normalized values for replicate spots were plotted according to gene order within the plastid genome. Fragments for which fewer than 13 spots per experiment (AtCRP1p::AtCRP1-GFP/WT) passed our manual quality control and/or yielded an F532 signal below background were excluded and appear as gap in the curve. The enrichment of psaC5′UTR is in agreement with previous findings obtained by RIP-Chip analysis on CRP1 from maize (Schmitz-Linneweber et al., 2005).

(B) Immunoblot analysis of protein fractions obtained from immunoprecipitation experiments using the anti-GFP mouse antibody and stroma material from Col-0 and AtCRP1-GFP plants. Equal volumes of supernatant and pellet preparations were loaded onto the gel. Note that the pellet from AtCRP1-GFP immunoprecipitation gave a stronger signal than the corresponding supernatant, implying quantitative precipitation of AtCRP1-GFP. The fact that no signal was obtained with Col-0 extracts demonstrates the specificity of the antibody. The RbcL migration region of the Ponceau S stained nylon membrane, after transfer from SDS-PAGE, was used to verify equal loading.

(C) Verification of AtCRP1 RNA targets. Coimmunoprecipitations and RNA extractions from AtCRP1-GFP and Col-0 samples were performed as for RIP-Chip assays. The RNAs were then analyzed by slot-blot hybridization with the indicated probes (see also Materials and Methods and Supplementary Table S2). The ATPase-α probe hybridization was included as a control. SUP, supernatant.
In summary, the RIP-Chip and slot blot data together with the colocalization of native footprints and AtCRP1 RNA binding motifs indicate that AtCRP1 likely binds directly to the 5′UTR of psaC and the petB-petD intergenic region and possibly to the 5′UTR of petA. On the contrary, the absence of an AtCRP1-specific footprint within the rps15 RNA, together with the failure of slot blot enrichment, makes any AtCRP1−rps15 interaction unlikely.

**AtCRP1 Is Required for the Correct Processing of *psbB-psbT-psbH-petB-petD* Transcripts**

To assess whether the lack of Cyt *b/f* complex and PsA subunit, together with the marked reduction of all protein complex subunits observed in *atcrp1-1* thylakoids, was caused by deficiencies in transcript accumulation and AtCRP1-dependent transcript processing, we probed the identified AtCRP1 RNA targets and other plastid transcripts by gel blot hybridization (Figure 9).

We investigated the transcripts encoding the subunits CP47 (*psbB*), T (*psbT*), and H (*psbH*) of photosystem II (PSII), subunits A (*psaA*) and C (*psaC*) of PSI, Cyt *f* (*petA*), Cyt *b* (*petB*) and subunit IV (*petD*) of cytochrome *b/f* and the alpha subunit of ATPase (ATPase-α). All these transcripts accumulated in *atcrp1-1* plastids to levels lower than WT, indicating that global plastid gene expression is affected by the *atcrp1-1* mutation, and explaining the marked reduction of thylakoid protein accumulation observed in *atcrp1-1* leaves.

Furthermore, the plastid polycistronic transcription unit *psbB-psbT-psbH-petB-petD* showed some striking alteration of transcript pattern in *atcrp1* samples (Figure 9). In particular, the monocistronic *petB* (band #4, 0.8 Kb), the dicistronic *psbH-petB* (band #3; 1.2 Kb) and the unspliced *petB* (band #2, 1.6 Kb) transcripts were barely detectable in the mutant, whereas the *petB*-unspliced *petD*-spliced dicistronic transcript (band #1, 2.2 Kb), detected with probes D, E, F, and H, accumulated to even higher levels in *atcrp1* plastids, presumably due to the failure of AtCRP1-dependent processing between the *petB* and *petD* coding regions, as also shown in zmcrl mutant plants (Barkan et al.,...
FIGURE 9 | Transcript patterns of chloroplast genes in Col-0 and atcrp1-1 mutant leaves. (A) The structure of the psbB gene cluster and probes A to H used in RNA gel blots analysis in (B) are shown. Furthermore, processed and spliced transcripts that accumulate differentially between Col-0 and mutant chloroplasts are drawn to scale and numbered from 1 to 4. Upward arrow indicates transcripts that accumulate to higher levels in atcrp1-1 than Col-0 chloroplasts, whilst the downward arrow is used for transcripts less abundant or absent in mutant samples. The putative binding site of AtCRP1 within the petB-petD intergenic region is also indicated. (B) RNA gel blot analysis of the psbB gene cluster were performed using probes indicated as A to H, whilst petA, ATPase-A, psaC, and psaA specific probes are described in section "Materials and Methods." The identity of labeled transcripts (1–4), shown in (A) together with their size, was established based on the hybridization pattern, transcript size and on data reported in Meierhoff et al. (2003) and Stoppel et al. (2011). Asterisks indicate the mature transcript forms. A portion of the ethidium bromide stained Agarose gels, containing the cytosolic 25S rRNA, is included, as loading control, below each filter. One out of three Northern-blots for each transcript-specific probe is shown.

1994; Fisk et al., 1999). In contrast with maize, monocistronic and spliced petD transcripts of ~600 nucleotides do not accumulate to significant levels in Arabidopsis, and thus its absence was not observed in atcrp1 plastids (Barkan et al., 1994; Barkan, 2011).

Moreover, the lack of the PsuC and PetA subunits could be the consequence of the simultaneous decrease of transcript accumulation and a possible defect in AtCRP1-dependent activation of psaC and petA translation, as shown in Zea mays (Barkan et al., 1994; Schmitz-Linneweber et al., 2005). However, the specific regulatory role of AtCRP1 in plastid protein translation is difficult to verify, owing to the general and pleiotropic decrease of mature plastid rRNA in atcrp1-1 leaves, in spite of WT-like accumulation of rrn23 and rrn4.5 precursor forms (Figure 10). This rRNA accumulation pattern...
FIGURE 10 | Plastid rRNA accumulation in Colo-0 and atcrp1-1 mutant leaves. (A) Schematic representation of the chloroplast rrn operon. Probes used in Northern blot analysis are indicated as black bars under each rRNA gene (A–D). (B) RNA gel blot analysis of plastid rRNAs were performed using the probes A-to-D described above. For loading control, a methylene blue stained filter is shown. One out of three Northern-blots for each transcript-specific probe is shown.

is very similar to the ones of mutants with impaired chloroplast translation and has been interpreted as a secondary consequence of reduced plastid protein synthesis (Tiller et al., 2012; Tadini et al., 2016).

**DISCUSSION**

In this study we have investigated the role of AtCRP1 in the biogenesis of dicotyledonous-C3 chloroplasts and compared its function to the already characterized monocotyledonous-C4 chloroplast counterpart, ZmCRP1. Both proteins are essential for chloroplast biogenesis and photosynthetic activity, since they are required for the processing and translation of specific plastid transcripts encoding subunits of the thylakoid protein complexes. Our results indicate that AtCRP1 and ZmCRP1 have very similar RNA targets and the main functional divergences are most likely due to the distinct localization of the two proteins inside the chloroplast and the partially different affinity for the RNA targets (see Table 1).

**CRP1 Proteins Are Part of Chloroplast Nucleoids**

We detected AtCRP1 in the stroma and associated with thylakoid membranes (see Figure 2; Table 1), whereas ZmCRP1 is a stromal protein with no detectable association with chloroplast membranes (Fisk et al., 1999). The dual localization of AtCRP1 within the chloroplast is supported by proteomic studies that detected AtCRP1 in the grana-fraction of Arabidopsis thylakoids (Tomizioi et al., 2014) and in the stroma proteome, as part of Megadalton complexes (Olinares et al., 2010). In particular, AtCRP1 appeared to be highly enriched in fractions that contained ribosomal proteins, translation factors, RNA helicases and other PPR proteins, suggesting a major role of AtCRP1 in chloroplast gene expression. These data, together with the co-localization with GUN1 protein (see Figure 2), indicate that AtCRP1 is integral to chloroplast nucleoids (Koussevitzky et al., 2007; Colombo et al., 2016; Tadini et al., 2016), i.e., the DNA-containing structures without defined boundaries that harbor the plastid gene expression machinery (Pfalz and Pfannschmidt, 2013; Melonek et al., 2016). Similarly, ZmCRP1 was found to be highly enriched in the nucleoid fractions of maize plastids, together with proteins involved in DNA replication, organization and repair as well as transcription, mRNA processing, splicing and editing (Majeran et al., 2012), further supporting the involvement of CRP1 proteins in plastid gene expression.

**CRP1 Proteins Are Required for the Biogenesis of the Photosynthetic Apparatus**

The yellow-albinotic and seedling lethal phenotype exhibited by atcrp1 is very similar to the chlorophyll deficient and lethal phenotype of zmcrp1 plants (Barkan et al., 1994; Fisk et al., 2016).
TABLE 1 | Overview of the phenotypes of Arabidopsis and maize crp1 mutants and comparison of their molecular roles in chloroplast biogenesis.

|                      | atcrp1a            | zmcrp1b            |
|----------------------|--------------------|--------------------|
| Plant phenotype      | Seedling lethal with yellow-albino cotyledons and leaves. Plants are able to develop mature leaves and sterile flowers when grown on MS medium supplemented with sucrose | Seedling lethal with pale-green cotyledon and leaves. Plants are able to develop mature non-photosynthetic leaves thanks to the large reserves of maize seeds |
| CRP1 protein localization | AtCRP1 is a component of plastid nucleoids and it is found associated to thylakoid membranes and in the stroma | ZmCRP1 has been reported to be highly enriched in plastid nucleoids and to localize exclusively in the chloroplast stroma |
| Thylakoid protein content |                      |                    |
| PSI                  | (PsaA)             |                    |
| PSII                | /                  |                    |
| Cyt b/f             | –                  | –                  |
| ATPase              | –                  | =                  |
| RNA targets         | RIP-Chip           | Slot-Blot          |
| psaC                | psaC               | psaC               |
| petB-petD           | petB-petD          | petB-petD          |
| /                   | petA               | petA               |
| rps15               | /                  | /                  |
| Metabolism of chloroplast RNAs |                      |                    |
| Accumulation        |                      |                    |
| Processing defects  |                      |                    |
| psaC                | –                  | No                 |
| petB-petD           | /                  | Yes                |
| petA                | –                  | No                 |

aData are obtained from the present manuscript bData are obtained from Barkan et al. (1994, 2012), Fisk et al. (1999), Schmitz-Linneweber et al. (2005). –, marked reduction; /, complete absence; =, no changes; +, increase; n.r., not reported.

1999). Arabidopsis mutants die at the two-cotyledon stage after germination on soil, but can overcome seedling lethality on sucrose-containing media, where they develop mature leaves and sterile flowers (see Figure 4; Table 1). Similarly, non-photosynthetic zmcrp1 plants die at about 3 weeks after germination when seed reserves are exhausted. Furthermore, the atcrp1 phenotype appears to be typical of Arabidopsis mutants lacking components of the photosynthetic apparatus and not of the gene expression machinery or of the protein import apparatus, since the latter usually result in the premature arrest at the globular-to-heart stage of embryo development, when chloroplast biogenesis begins (Ruppel and Hangarter, 2007; Romani et al., 2012; Beeler et al., 2014). Nevertheless, the pale-green pigmentation of the mutant embryo at bent-cotyledon stage (see Figure 4) and the β-glucuronidase (GUS) activity observed in young developing cotyledons and rosette leaves, but not in older tissues (see Figure 5), indicate that AtCRP1 gene expression and protein accumulation is required during the very early stages of the photosynthetic apparatus assembly. Immunoblot data indicate, indeed, that AtCRP1, like ZmCRP1, might act as a nuclear-encoded anterograde regulatory component responsible for coordination of the accumulation of Cyt b/f and PSI protein complexes (see Figure 6). Besides their role in linear electron transport (LET), Cyt b/f and PSI indeed play a key role in Cyclic Electron Transport (CET), which has been reported to be enhanced in Arabidopsis green seeds and to be required for optimal seed vigor and seed germination rate (Allorent et al., 2015).

In contrast to zmcrp1 plants (Barkan et al., 1994), the absence of AtCRP1 destabilized the entire photosynthetic apparatus, as shown by the marked reduction of PSII core, ATPase and LHC protein levels. The general down-regulation of thylakoid complexes owing to defects in the intersystem electron transport chain appears to be a common feature of Arabidopsis photosynthetic mutants and provides clear evidence of a different adaptive response between monocot and dicot plants (Meurer et al., 1996; Varotto et al., 2000, 2002; Maiwald et al., 2003; Weigel et al., 2003; Ihnatowicz et al., 2004, 2007; Belcher et al., 2015). Furthermore, the atcrp1-1 phenotype, both in terms of plastid transcript and plastid protein accumulation, appears to be much more drastic than the one of other ppr mutants required for the processing and expression of psbB-psbT-psbH-petB-petD operon, such as hcf152 (Meierhoff et al., 2003), suggesting that the absence of AtCRP1 protein might affect the activity of other factors essential for plastid gene expression. As a matter of fact, rRNA abundance is markedly reduced in atcrp1-1 plastids, indicating a general reduction of protein synthesis, as consequence of pleiotropic effects.

RNA Targets: Commonalities and Divergences between AtCRP1 and ZmCRP1 Proteins

RNA immunoprecipitation-Chip and slot blot data suggest a physical interaction between AtCRP1 and the transcripts of
psaC, petB-petD and possibly petA, even though it is not known whether these interactions are direct or mediated by other factors (see Figure 7). However, all of these RNAs harbor a region where a native footprint is annotated, raising the tempting hypothesis that AtCRP1 is in fact the RNA-binding factor responsible for that footprint (see Figure 8). Furthermore, when these enriched fragments were searched for occurrences of the predicted binding motif of AtCRP1, each of them proved to contain a hit inside the footprint region, strongly suggesting that AtCRP1 could be the factor leaving those footprints. Nevertheless, the observation that the footprints identified in Arabidopsis psaC, petA, and petB-petD transcripts are larger than the 14 nucleotide size of the predicted AtCRP1 footprint (37, 29, and 41 nucleotides in psaC, petA, and petB-petD, respectively) supports the view that the binding of AtCRP1 to its targets in vivo could be stabilized by other protein partners. For instance, the peptide chain release factor B3 (PrfB3) has been also shown to be required for Arabidopsis autotrophic growth and for the stability of 3′ processed petB transcripts to adjust cytochrome b6 levels (Stoppel et al., 2011), thus possibly being an AtCRP1 specific protein partner. Similarly, PPR proteins involved in RNA stabilization and editing have been shown to interact with RNA Recognition Motif (RRM) proteins and other factors, indicating that larger protein complexes assembled around a PPR protein are likely to occur (Kupsch et al., 2012; Takenaka et al., 2014; Shi et al., 2015).

The interactions with the 5′UTR of psaC and petA have also been reported in the case of ZmCRP1 (Schmitz-Linneweber et al., 2005; Williams-Carrier et al., 2008), indicating that this feature of CRP1 function is conserved between Arabidopsis and maize. ZmCRP1 was also shown to bind directly to the 5′-UTR of petA transcripts by electrophoresis mobility shift assay (Williams-Carrier et al., 2008), favoring the possibility of a direct binding of CRP1 proteins to the corresponding RNA targets (see also Figure 8). Furthermore, ZmCRP1 has been proposed to directly control the translation of petA and psaC transcripts (Barkan et al., 1994), as shown through pulse labeling and polysome loading (in the case of petA), or deduced from the reduced association of psaC RNAs with ribosomes. Interestingly, the PsA subunit of PSI and the PetA subunit of Cyt b6/f could not be detected in atcrp1 thylakoids, despite the accumulation of the corresponding transcripts with no processing defects (see also Figure 9), suggesting that AtCRP1 plays a major role in translation regulation also in Arabidopsis. Unfortunately, the specific requirement of AtCRP1 in plastid protein translation cannot be verified by comparing Col-0 and atcrp1-1 leaves, due to the marked reduction of rRNA accumulation in atcrp1-1 plastids.

In addition to the defects in petA translation, the complete absence of Cyt b6/f protein complex observed in atcrp1 thylakoids can also be attributed to processing alterations of the psbB-psbH-psbT-psbH-psetD polycistronic transcription unit. The lack of the monocistronic petB, the dicistronic psbH-petB, and the unspliced petB transcripts, together with the direct binding of AtCRP1 to the petB-petD intergenic region, strongly support the role of AtCRP1 in the metabolism of petB and petD transcripts. PPR protein-derived RNA-footprints are considered to arise due to exonucleolytic activity (Ruwe et al., 2016). Since sRNAs corresponding to predicted binding sites of AtCRP1 are identified here, the most likely role for AtCRP1 is to block exonucleases from degrading the petB and petD transcripts. A similar defect in petB-petD maturation has been reported in zmcrp1 mutant plants (Barkan et al., 1994; Fisk et al., 1999; Schmitz-Linneweber et al., 2005), although no association was detected between ZmCRP1 and the petB-petD intergenic region (Schmitz-Linneweber et al., 2005), so it is still uncertain whether the role of ZmCRP1 is direct or indirect.

**CONCLUSION**

Taken together, the characterization of the functional role of AtCRP1 in chloroplast biogenesis has highlighted several features in common with the ZmCRP1. Both proteins appear to control, directly or indirectly, the expression of plastid genes encoding subunits of Cyt b6/f and PSI protein complexes. The coordination of the accumulation of these two protein complexes is fundamental to guarantee optimal photosynthesis in mature plants, but appears also to be important during seed germination, when cyclic electron transport is highly enhanced relative to LET.

Differences in RNA targets observed by immunoprecipitation and hybridization assays between AtCRP1 and ZmCRP1 might be explained by a broad affinity for RNA targets, but may also have technical reasons (GFP antibody for Arabidopsis versus direct anti-ZmCRP1 antibody in maize). Evidence in favor of conservation of PPR protein activity between different species has been reported for the PLS and P subfamilies (Choury et al., 2005; Bolle and Kempken, 2006; Choury and Araya, 2006); for instance, the maize MPPR6 protein can complement loss-of-function Arabidopsis mutants lacking the orthologous protein (Manavski et al., 2012). However, functional divergence has been also observed, as in the case of orthologous PPR proteins ATP4 (maize) and SVR7 (Arabidopsis) (Liu et al., 2010; Zoschke et al., 2012, 2013a,b). Further studies aimed to verify the degree of protein activity conservation between monocots and dicots are needed to extend our knowledge of PPR protein functions and the degree of protein function conservation. The parallel characterization of PPR orthologs, including the relationship between their protein structures and the corresponding target RNA species, may represent an underestimated and powerful strategy to precisely determine the PPR code, essential for a fast and accurate large scale prediction of PPR targets, and for the functional characterization of the PPR-mediated nucleus-to-chloroplast anterograde signaling pathway.

**AUTHOR CONTRIBUTIONS**

RF, LT, FM, FR, SM, M-KL, CS-L, and PP participated to the organization of the manuscript. RF, LT, FM, FR, SM, MC, AC,
and PP designed and carried out the experiments related to the molecular biology and biochemical characterization of atcrp1 mutants. RF, FM, LT, M-KL, and CS-L were involved in RIP-Chip and slot blot assays, as well as in the in silico identification of native footprints and prediction of AtCRP1 binding motif. PP wrote the manuscript.

FUNDING

This work was supported by ERA-NET Cofund FACE SURPLUS (BarPLUS grant id. 93) to PP. Work in the lab of CS-L was supported by DFG grant SCHM 1698/5-1.

REFERENCES

Allorent, G., Osorio, S., Vu, J. L., Falconet, D., Jouhet, J., Kuntz, M., et al. (2015). Adjustments of embryonic photosynthetic activity modulate seed fitness in Arabidopsis thaliana. New Phytol. 205, 707–719. doi: 10.1111/nph.13044

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., et al. (2003). Genome-wide insertion mutagenesis of Arabidopsis thaliana. Science 301, 653–657. doi: 10.1126/science.1086391

Barkan, A. (2011). Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. Plant Physiol. 155, 1520–1532. doi: 10.1104/pp.111.171231

Barkan, A., Rojas, M., Fujii, S., Yap, A., Chong, Y. S., Bond, C. S., et al. (2012). A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. PLoS Genet. 8:e1002910. doi: 10.1371/journal.pgen.1002910

Barkan, A., and Small, I. (2014). Pentatricopeptide repeat proteins in plants. Annu. Rev. Plant Biol. 65, 415–442. doi: 10.1146/annurev-arplant-050213-050159

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

Beeler, S., Liu, H. C., Stadler, M., Schreier, T., Eicke, S., Lue, W. L., et al. (2014). Plastidial NAD-dependent malate dehydrogenase is critical for embryo development and heterotrophic metabolism in Arabidopsis. Plant Physiol. 164, 1175–1190. doi: 10.1104/pp.113.233866

Belcher, S., Williams-Carrier, R., Stüffer, N., and Barkan, A. (2015). Large-scale genetic analysis of chloroplast biogenesis in maize. Biochem. Biophys. Acta 1847, 1004–1016. doi: 10.1016/j.bjba.2015.02.014

Bolle, N., and Kempken, F. (2006). Mono- and dicotyledonous plant-specific RNA editing sites are correctly edited in both in organellar systems. FEBS Lett. 580, 4443–4448. doi: 10.1016/j.febslet.2006.07.011

Cai, W., Okuda, K., Peng, L., and Shikanai, T. (2011). PROTON GRADIENT REGULATION 3 recognizes multiple targets with limited similarity and mediates translation and RNA stabilization in plastids. Plant J. 67, 318–327. doi: 10.1111/j.1365-313X.2011.04593.x

Cheng, S., Gutmann, B., Zhong, X., Ye, Y., Fisher, M. F., Bai, F., et al. (2016). Redefining the structural motif(s) that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. Plant J. 85, 532–547. doi: 10.1111/tpj.13121

Choury, D., and Araya, A. (2006). RNA editing site recognition in heterologous plant mitochondria. Curr. Genet. 50, 405–416. doi: 10.1007/s00299-006-0100-3

Choury, D., Farre, J. C., Jordana, X., and Araya, A. (2005). Gene expression studies in isolated mitochondria: Solanum tuberosum rps10 is recognized by cognate potato but not by the transcription, splicing and editing machinery of wheat mitochondria. Nucleic Acids Res. 33, 7058–7065. doi: 10.1093/nar/gki1017

Colombo, M., Tadini, L., Peracchis, C., Ferrari, R., and Pesaresi, P. (2016). GUN1, a jack-of-all-trades in chloroplast protein homeostasis and signaling. Front. Plant Sci. 7:1427. doi: 10.3389/fpls.2016.01427

ACKNOWLEDGMENT

We thank Hannes Ruwe and Gongwei Wang for helpful comments on the sRNA analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2017.00163/full

TABLE S1 | RNA immunoprecipitation (RIP-Chip) data.

TABLE S2 | Oligonucleotides employed for cloning, genotyping, northern blot, slot blot and qRT-PCR assays.
Kupsch, C., Ruwe, H., Gusewski, S., Tillich, M., Small, I., and Schmitz-Linneweber, C. (2012). Arabidopsis chloroplast RNA binding proteins CP31A and CP29A associate with large transcript pools and confer cold stress tolerance by influencing multiple chloroplast RNA processing steps. Plant Cell 24, 4266–4280. doi: 10.1105/tpc.112.103002

Lawrence, C. J., Dong, Q., Polacco, M. L., Seigfried, T. E., and Brendel, V. (2004). MaizeGDB, the community database for maize genetics and genomics. Nucleic Acids Res. 32, D393–D397. doi: 10.1093/nar/gkh011

Li, X., Yu, F., and Rodermel, S. (2010). An Arabidopsis pentatricopeptide repeat protein, SUPPRESSOR OF VARIATION7, is required for FtsH-mediated chloroplast biogenesis. Plant Physiol. 154, 1588–1601. doi: 10.1104/pp.110.164111

Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., et al. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16, 2089–2103. doi: 10.1105/tpc.104.022236

Maiwald, D., Dietzmann, A., Jahns, P., Pesaresi, P., Joliot, P., Joliot, A., et al. (2003). Knock-out of the genes coding for the Rieske protein and the ATP-synthase delta-subunit of Arabidopsis. Effects on photosynthesis, thylakoid protein composition, and nuclear chloroplast gene expression. Plant Physiol. 133, 191–202. doi: 10.1104/pp.103.021490

Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., et al. (2012). MCP201 clustering. Trends Plant Sci. 2012.11.003

Ruwe, H., Wang, G., Gusewski, S., and Schmitz-Linneweber, C. (2016). Systematic analysis of plant mitochondrial and chloroplast small RNAs suggests organelle-specific mRNA stabilization mechanisms. Nucleic Acids Res. 44, 7406–7417. doi: 10.1093/nar/gkw466

Schagger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. 166, 368–379. doi: 10.1006/abio.1997.0176

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682. doi: 10.1038/nmeth.1927

Schmitz-Linneweber, C., Williams-Carrier, R., and Barkan, A. (2005). RNA immunoprecipitation and microarray analysis show a chloroplast Pentatricopeptide repeat protein to be associated with the 5′ region of mRNAs whose translation it activates. Plant Cell 17, 2791–2804. doi: 10.1105/tpc.105.034454

Schmitz-Linneweber, C., Williams-Carrier, R. E., Williams-Voelker, P. M., Kroeger, T. S., Vichas, A., and Barkan, A. (2006). A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 pre-mRNA. Plant Cell 18, 2650–2663. doi: 10.1105/tpc.105.046110

Sessions, A., Burke, E., Presting, G., Aux, G., Mcelver, J., Patton, D., et al. (2002). A high-throughput Arabidopsis reverse genetics system. Plant Physiol. 141, 2985–2994. doi: 10.1104/pp.104.046380

Shen, C., Zhang, D., Guan, Z., Liu, Y., Yang, Z., Yang, Y., et al. (2016). Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. Nat. Commun. 7:11285. doi: 10.1038/ncomms11285

Shi, X., Hanssen, M. R., and Bentolila, S. (2015). Two RNA recognition motif-containing proteins are plant mitochondrial editing factors. Nucleic Acids Res. 43, 3814–3825. doi: 10.1093/nar/gkv245

Small, I. D., and Peeters, N. (2000). The PPR motif - a TPP-related motif prevalent in plant organellar proteins. Trends Biochem. Sci. 25, 46–47. doi: 10.1016/S0968-0004(99)01520-0

Stoppel, R., Lezhneva, L., Schwenkert, S., Torabi, S., Felder, K., Meierhoff, K., et al. (2011). Recruitment of a ribosomal release factor for light- and stress-dependent regulation of petB transcript stability in Arabidopsis chloroplasts. Plant Cell 23, 2680–2695. doi: 10.1105/tpc.111.085324

Tadini, L., Pesaresi, P., Kleine, T., Rossi, F., Guljamow, A., Sommer, F., et al. (2016). GUN1 controls accumulation of the plastid ribosomal protein S1 at the protein level and interacts with proteins involved in plastid protein homeostasis. Plant Physiol. 170, 1817–1830. doi: 10.1104/pp.15.02033

Takanaka, M., Verbitskiy, D., Zehrmann, A., Hartel, B., Bayer-Cassarz, E., Glass, F., et al. (2014). RNA editing in plant mitochondria-connecting RNA target sequences and acting proteins. Mitochondrion 19(1B), 191–197. doi: 10.1016/j.mito.2014.04.005

Takanaka, M., Zehrmann, A., Brennicke, A., and Graichen, K. (2013). Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. PLoS ONE 8:e65343. doi: 10.1371/journal.pone.0065343

Tillet, N., Weingartner, M., Thiele, W., Maximova, E., Schütter, M. A., and Bock, R. (2012). The plastid-specific ribosomal proteins of Arabidopsis thaliana can be divided into non-essential proteins and genuine ribosomal proteins. Plant J. 69, 302–316. doi: 10.1111/j.1365-313X.2011.04791.x
Tomizioli, M., Lazar, C., Brugiere, S., Burger, T., Salvi, D., Gatto, L., et al. (2014). Deciphering thylakoid sub-compartments using a mass spectrometry-based approach. *Mol. Cell. Proteom.* 13, 2147–2167. doi: 10.1074/mcp.M114.040923

Varotto, C., Pesaresi, P., Jahns, P., Lessnick, A., Tizzano, M., Schiavon, F., et al. (2002). Single and double knockouts of the genes for photosystem I subunits G, K, and H of *Arabidopsis*. Effects on photosystem I composition, photosynthetic electron flow, and state transitions. *Plant Physiol.* 129, 616–624. doi: 10.1104/pp.002089

Varotto, C., Pesaresi, P., Meurer, J., Oelmuller, R., Steiner-Lange, S., Salamini, F., et al. (2000). Disruption of the *Arabidopsis* photosystem I gene psaE1 affects photosynthesis and impairs growth. *Plant J.* 22, 115–124. doi: 10.1046/j.1365-313x.2000.00717.x

Waadt, R., and Kudla, J. (2008). In planta visualization of protein interactions using Bimolecular Fluorescence Complementation (BiFC). *CSH Protoc* 2008, dbrot4995.

Weigel, M., Varotto, C., Pesaresi, P., Finazzi, G., Rappaport, F., Salamini, F., et al. (2003). Plastocyanin is dispensable for photosynthetic electron flow in *Arabidopsis thaliana*. *J. Biol. Chem.* 278, 31286–31289. doi: 10.1074/jbc.M302876200

Woodson, J. D., and Chory, J. (2008). Coordination of gene expression between organelar and nuclear genomes. *Nat. Rev. Genet.* 9, 383–395. doi: 10.1038/nrg2348

Yamazaki, H., Tasaka, M., and Shikanai, T. (2004). PPR motifs of the nucleus-encoded factor, PGR3, function in the selective and distinct steps of chloroplast gene expression in *Arabidopsis*. *Plant J.* 38, 152–163. doi: 10.1111/j.1365-313X.2004.02035.x

Yin, P., Li, Q., Yan, C., Liu, Y., Liu, J., Yu, F., et al. (2013). Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* 504, 168–171. doi: 10.1038/nature12651

Yoo, S. D., Cho, Y. H., and Sheen, J. (2007). *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* 2, 1565–1572. doi: 10.1038/nprot.2007.199

Zoschke, R., Kroeger, T., Belcher, S., Schottler, M. A., Barkan, A., and Schmitz-Linneweber, C. (2012). The pentatricopeptide repeat-SMR protein ATP4 promotes translation of the chloroplast atpB/E mRNA. *Plant J.* 72, 547–558. doi: 10.1111/j.1365-313X.2012.05081.x

Zoschke, R., Qu, Y., Zubo, Y. O., Borner, T., and Schmitz-Linneweber, C. (2013a). Mutation of the pentatricopeptide repeat-SMR protein SVR7 impairs accumulation and translation of chloroplast ATP synthase subunits in *Arabidopsis thaliana*. *J. Plant Res.* 126, 403–414. doi: 10.1007/s10265-012-0327-1

Zoschke, R., Watkins, K. P., and Barkan, A. (2013b). A rapid ribosome profiling method elucidates chloroplast ribosome behavior in vivo. *Plant Cell* 25, 2265–2275. doi: 10.1105/tpc.113.111567

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Ferrari, Tadini, Monatti, Lehninger, Costa, Rossi, Colombo, Masiero, Schmitz-Linneweber and Pesaresi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.