The RRM domain-containing protein Rbp3 interacts with ribosomes and the 3′ ends of mRNAs encoding photosynthesis proteins

Luisa Hemm[1], Elisabeth Lichtenberg[2], Stefan Tholen[3], Viktoria Reimann[1], Kenta Kakazu[4], Sotaro Machida[4], Moontaha Mahbub[5], Oliver Schilling[3], Annegret Wilde[2] Satoru Watanabe[4], Conrad W. Mullineaux[5] and Wolfgang R. Hess[1,*]

[1]Genetics and Experimental Bioinformatics, Faculty of Biology, Institute of Biology III, University of Freiburg, Freiburg, Germany;
[2]Molecular Genetics of Prokaryotes, Faculty of Biology, Institute of Biology III, University of Freiburg, Freiburg, Germany;
[3]Institute for Surgical Pathology, Medical Center – University of Freiburg, Faculty of Medicine, Freiburg, Germany;
[4]Department of Bioscience, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo, Japan
[5]School of Biological and Behavioural Sciences, Queen Mary University of London, London, United Kingdom

*corresponding author: Wolfgang R. Hess
Email: wolfgang.hess@biologie.uni-freiburg.de

Author Contributions: WRH designed the project and secured funding. ST and OS carried out MS -based proteomic analyses. SW, SM and KK provided the Δ*rbp2/3* mutant and the BLI data. VR performed the microarray analyses, MM and CWM provided the spectroscopic analyses. EL and AW performed the RNA-FISH and microscopical analyses and contributed to the data interpretation. All other experiments and analyses were performed by LH. LH and WRH wrote the manuscript with input from all authors.

Competing Interest Statement: The authors declare that they have no competing interests.

Classification: Biological Sciences, Microbiology

Keywords cyanobacteria, gene expression, mRNA transport, photosynthesis, RNA binding proteins

This PDF file includes:

Main Text

Figures 1 to 6

Abstract

RNA recognition motif (RRM) domain proteins are crucial RNA-binding proteins (RBPs) across all domains of life. In cyanobacteria, single RRM domain proteins are involved in mRNA targeting to the thylakoid membrane and acclimation to certain stress conditions, but many details of their physiological functions and molecular targets have remained unknown. The model cyanobacterium *Synechocystis* sp. PCC 6803 has a family of three genes encoding the RRM domain-containing proteins Rbp1, Rbp2 and Rbp3. Here, we verified the RNA-binding activity of Rbp3 *in vivo* and show that cells of a Δ*rbp3* deletion strain had a lower PSI:PSII ratio and pigment content and were significantly smaller than wild-type cells. To identify the set of interacting molecules, co-immunoprecipitation experiments were performed with a strain expressing a C-terminally FLAG-tagged Rbp3. Mass spectrometry of the elution fraction suggested physical proximity between Rbp3, ribosomes, and a very small number of other proteins. The most highly enriched transcript in the co-eluting RNA fraction was the *psaAB* mRNA. This was corroborated by fluorescent *in situ* hybridization (FISH) analyses showing decreased *psaA* mRNA signals and colocalization with Rbp3-GFP signals and ribosomes. Other enriched mRNAs encode thylakoid, plasma membrane and carboxysome proteins. The Rbp3-mRNA interactions occurred preferentially towards the end of coding regions or the 3′UTRs, although some were also mapped to other regions. Binding assays using Bio-layer Interferometry validated the Rbp3-*psaAB* mRNA interaction, indicating a preference for folded RNA segments near or overlapping the respective stop codons.

Significance statement

The mechanisms by which proteins are produced at specific sites and inserted into the intricate membrane systems of photosynthetic cyanobacteria are only partially understood. While RRM domain proteins are well-studied RNA-binding proteins in eukaryotes, their functions in bacteria remain underexplored. This study reveals that the RRM domain protein Rbp3 in the cyanobacterium *Synechocystis* sp. PCC 6803 binds mRNAs encoding photosynthetic proteins, plasma membrane proteins and carboxysome proteins and localizes near ribosomes. The bound RNA segments are typically near the ends of coding regions, or in 5′ untranslated regions. These findings

suggest that Rbp3 is involved in targeting mRNAs to specific intracellular locations by interacting with structural elements within these mRNA molecules.

Introduction

RNA-binding proteins in bacteria

RNA-binding proteins (RBPs) play crucial roles in the physiological functions, processing and stability of RNA molecules. Although prokaryotes lack the extensive processing of pre-mRNAs that is typical of eukaryotes, they have a variety of RBPs that act in different steps of RNA-related processes. Among the best-studied RBPs in bacteria are the Rho factors, which terminate transcription (1), Hfq and ProQ, which mediate sRNA/mRNA interactions and can alter RNA stability (2, 3), and cold shock protein A, an important transcription antitermination factor (4). Little is known about putative homologs of these well-known RBPs in cyanobacteria. A structural homolog of Hfq exists in most cyanobacteria including the model *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis* 6803), but the residues interacting with RNA in other bacteria are not conserved (5), and RNA-binding activity was not detected (6). However, systematic studies revealed several candidate RBPs in cyanobacteria (7, 8), but only a few RBPs were analyzed in greater detail.

RRM domain-containing proteins in cyanobacteria

The RNA recognition motif (RRM; also called ribonucleoprotein domain (RNP) or RNA-binding domain (RBD)) (9, 10) is the most common RNA-binding domain. It consists of a four-stranded antiparallel beta-sheet, stacked on two alpha helices (11). The domain was first described in proteins of heterogeneous nuclear ribonucleoprotein particles, where it is involved in the splicing of heterogeneous nuclear RNA (9). In addition to splicing, eukaryotic RRM proteins are now known to be involved in a variety of RNA-related processes. Eukaryotic RRM proteins are associated with mRNA stability (12), poly(A)-binding (13), mRNA transport (14), translation (15) and degradation (16). In prokaryotes, most proteins harboring an RRM domain are found in species in the phylum cyanobacteria (17). These bacterial RRM domain proteins typically have a single RRM domain, whereas eukaryotic proteins usually have multiple RRM domains.

RRM-domain proteins in cyanobacteria were discovered in *Synechococcus* sp. PCC 6301 (18), and homologs were soon found in many other cyanobacteria (19–22). Genes encoding these RBPs are strikingly conserved, with at least three members throughout the entire cyanobacterial phylum (*SI Appendix*, **Fig. S1**), suggesting a function very relevant to the physiology of cyanobacteria. Distinct types of these RBPs

have been recognized (23, 24), proteins with a C-terminal glycine-rich domain, frequently becoming expressed under cold stress (20, 25, 26), short proteins that consist of only the RRM domain and the longer Type II proteins with a relatively long C-terminal region of unknown function.

Despite the presence of the RRM domain, RNA binding was only verified for the RBPs of *Anabaena variabilis* using an *in vitro* assay (25), for RbpA1 of *Anabaena* (*Nostoc*) sp. PCC 7120 in an *in vitro* RNA-binding assay using homopolymeric RNA molecules (23) and for the *Synechocystis* 6803 recombinant Rbp1 and Rbp2 in an *in vitro* assay using RNA-immunoprecipitation high-throughput sequencing (RIP-Seq) (24). Mahbub et al. performed RT-PCR on RNA co-immunoprecipitated with Rbp2 and Rbp3 (27). However, actual RNA binding *in vivo* has not been demonstrated, nor has the population of interacting transcripts *in vivo* been determined for any member of this protein family.

Three RRM domain-containing proteins are present in *Synechocystis* 6803, Rbp1, Rbp2, and Rbp3 (encoded by the genes *sll0517, ssr1480* and *slr0193*, respectively; *SI Appendix*, **Fig. S1**). Rbp1 is important under cold stress (28), although Rbp2 and Rbp3 are also somewhat cold-inducible (29). Rbp3 is involved in the acclimation to salt stress and the cold-induced lipid peroxidation (30). Indeed, the levels of some mRNAs encoding proteins involved in fatty acid biosynthesis and salt stress differed between a Δ*rbp3* mutant and the wild type, although the molecular basis of these differences remained unclear (30). Rbp3 consists of the RRM domain in the N-terminal 85 amino acids and an uncharacterized C-terminal section of 66 residues. It could be shown, that the salt tolerance in Δ*rbp3* mutants was restored when the Rbp3 C-terminus was fused to the N-terminus of Rbp1 suggesting that the salt-related aspects were not connected to the RRM domain (24). Conversely, the glycine-rich C-terminus of Rbp1 was dispensable under cold stress conditions (24).

Using fluorescent *in situ* hybridization (FISH) to probe the subcellular location of mRNAs encoding core subunits of the photosystems, we observed that Rbp2 and Rbp3 are involved in the targeting of *psaA* and *psbA* mRNAs towards the thylakoid membrane (27). Here, we analyzed the function of Rbp3 at the molecular and physiological level. We demonstrate actual RNA binding *in vivo* and show that this protein is relevant for the accumulation of a subset of mRNAs and determined the interacting proteins and transcripts. The data suggest that Rbp3 binds to the 3′ regions of mRNAs encoding photosynthesis core subunits and interacts closely with

ribosomes. Our results further suggest specialized functions for the different members of this protein family, but point also at some functional complementarity. Finally, we present a working model for the function of Rbp3.

Results

Rbp3 mutant strains show differences in growth and pigment content compared to wild type

To phenotype the Δ*rbp3* deletion mutant, spot assays were performed in triplicate with stepwise 10-fold dilutions (**Fig. 1A, B**). Growth of the *Synechocystis* 6803 wild type on agar medium was compared to that of Δ*rbp3* and a complementation strain in which a plasmid pVZ322-localized *rbp3* copy was introduced into Δ*rbp3* and expressed from its native promoter. Growth of the Δ*rbp3* strain appeared to slightly impaired compared to the other two strains. However, closer inspection revealed similar numbers of colonies in the three strains, while the colony sizes in Δ*rbp3* were significantly decreased compared to the wild type and complementation strain. After 8 days of growth under standard conditions, wild-type colonies had a median colony size of ~0.6 mm, while that of the Δ*rbp3* mutant was ~0.36 mm. This effect could be rescued in the complementation strain Δ*rbp3*+Rbp3 (**Fig. 1A**), although not completely (**Fig. 1B**). We conclude that the size, but not the number of colonies, was different in the absence of Rbp3 indicating that lack of Rbp3 slows growth, but does not induce cell death.

Whole-cell absorption spectra for the *Synechocystis* 6803 wild type and Δ*rbp3* indicated a lower cellular content of chlorophyll *a* and phycocyanin in Δ*rbp3* (**Fig. 1C**), which also showed a substantially lower PSI:PSII ratio in low-temperature fluorescence spectra (**Fig. 2D, E**). Both effects were complemented in the Δ*rbp3*+Rbp3 strain (**Fig. 1C; Fig. 2D, E**). Taken together, these measurements indicate that loss of *rbp*3 results in a lower cellular content of PSI.

From these experiments, we conclude that the loss of Rbp3 affects the growth of cells as well as their photosynthetic properties, as indicated by the lower pigment content. This phenotype could be rescued by introduction of a gene encoding Rbp3.

Verification of *in-vivo* RNA binding of Rbp3

Previous analyses suggested that Rbp3 is involved in the targeting of *psaA* and *psbA2* mRNAs towards the thylakoid membrane (27), but its actual RNA binding has not been directly demonstrated. To test the *in vivo* RNA binding of Rbp3, cultures expressing triple FLAG-tagged Rbp3 (strain Rbp3-3xFLAG) were used in an assay using the enzyme polynucleotide kinase (PNK) (31). The principle is outlined in **Fig. 2A.** After size-fractionating the complex and transfer to a nitrocellulose membrane, the labeled RNA was visualized by autoradiography and the protein by incubating the membrane with anti-FLAG antisera (**Fig. 2B**). For Rbp3, several bands were present in the protein panel, both for extracts from the crosslinked cultures and for non-crosslinked cultures used as negative controls. The most prominent band at ~20 kDa corresponded to the molecular mass of the Rbp3-3xFLAG monomer (19.6 kDa). Below this, three other bands were observed, which were likely N-terminally truncated Rbp3-3xFLAG fragments. Two bands >20 kDa, one at ~35 kDa and one at ~70 kDa, were common to both Rbp3 protein lanes, that is, with and without UV treatment. For the crosslinked sample, two bands were specific: a somewhat diffuse band at ~30 kDa and another at ~55 kDa. In the autoradiogram, a single prominent band was observed at ~30 kDa (**Fig. 2B**). Its position and somewhat diffuse appearance indicated that it belonged to the Rbp3-RNA ribonucleoprotein complex, which was also visible in the protein lane at the same molecular mass.

The sfGFP-3xFLAG control exhibited a prominent band just above the 35 kDa marker and several faint bands between ~70 kDa and ~100 kDa in the crosslinked protein sample, absent in the non-crosslinked sample, likely resulting from non-specific UV effects. A weaker band below the primary one indicated degraded sfGFP, implying UV treatment reduced protein stability. In the autoradiogram, only a very faint band was visible, possibly crosslinked sfGFP mRNA or RNA that had been near sfGFP during the crosslinking step. In comparison, the cross-linked Rbp3-3xFLAG sample showed a distinct band in the autoradiogram, matching a band in the western blot. This suggested that Rbp3 bound unknown RNA molecules *in vivo*, which could be isotope-labeled after protein purification and were protected from RNase digestion. The ~10 kDa mass difference between RNA-free Rbp3 (~20 kDa) and the RNA-Rbp3 complex indicated protected RNA fragments of approximately 60 nt. The diffused appearance of this band suggested a heterogeneous RNA population in length, composition, or both. The fact that Rbp3 was detected both with and without bound RNA suggested the existence of two pools of this protein.

The majority of Rbp3 RNA interactors are transcripts encoding membrane-associated proteins

The results of the PNK assay suggested that Rbp3 was interacting *in vivo* with an RNA population heterogenous in length, composition, or both (**Fig. 2B**). To identify these transcripts, CLIP-Seq analyses were performed. Cultures expressing Rbp3-3xFLAG or sfGFP-3xFLAG were crosslinked using 1.25 J cm^{-2} UV-C. After enrichment of the samples with anti-FLAG magnetic beads (*SI Appendix,* **Fig. S2**), the protein was proteinase K digested and the RNA purified. On average, we obtained 1130 ± 269 ng of RNA from 200 mL of UV-treated Rbp3-3xFLAG cultures and 583 ± 107 ng from sfGFP-3xFLAG cultures. For control, we prepared total RNA from 25 mL aliquots of the same cultures after UV treatment. All samples were prepared in biological triplicates. After reverse transcription and sequencing, the RNA sequencing data were mapped to the genome of *Synechocystis* 6803, followed by a DESeq2 (32) analysis of transcripts enriched in the Rbp3-3xFLAG pulldown compared to the sfGFP-3xFLAG pulldown. A total of 229 transcripts were significantly enriched ($\log_2 FC > 1$, $p \leq 0.05$), including many mRNAs encoding membrane-associated proteins (**Fig. 2C**). Of these enriched transcripts, 41 were mRNAs encoding proteins associated with the thylakoid membrane, 21 encoded proteins associated with the plasma membrane, and seven proteins involved in carboxysome formation and carbon fixation. Additionally, 16 transcripts encode cytoplasmic proteins, for all other no location was annotated by a GO term. The transcripts with the highest enrichment were *psaA* and *psaB*, organized in a dicistronic operon and encoding the core proteins of photosystem (PS) I. Other highly enriched transcripts were *epsB* (*sll0923*) encoding a cell membrane protein, the tyrosine kinase Wzc related to the production of exopolysaccharides (33), *atpH (ssl2615)* encoding the small, membrane-intrinsic ATP synthase subunit c, *apcE* encoding the core-membrane linker polypeptide ApcE (or L_{CM}), which anchors the giant phycobilisome light-harvesting structure to the membrane (34), *cpcL/cpcG2* encoding an alternative linker protein that facilitates energy transfer from a small phycobilisome type directly to PSI (35), and *psbA* encoding the PSII core protein D1. Additionally, the transcript for the cysteine-rich small protein IsiE, which is encoded in the *isiA-isiB* intergenic region (36) was highly enriched. For the complete list of 229 transcripts enriched in the co-IP with Rbp3 after UV cross-linking, see ***SI Appendix***, **Table S1**.

Spectroscopic measurements indicated that the Δ*rbp3* mutant had a lowered PSI:PSII ratio (**Fig. 2D, E**). In order to test whether deletion of *rbp3* had an effect on the *psaA* mRNA content per cell we used *in situ* FISH analysis. The results showed a significantly reduced *psaA* mRNA content per cell in Δ*rbp3* compared to the wild type and the complementation control (Δ*rbp3*+Rbp3-sfGFP) in which Rbp3 was fused to GFP and reintroduced into the Δ*rbp3* mutant (**Fig. 2F, G**).

Based on these analyses, we conclude that Rbp3 binds several different mRNAs, the most abundant of which are the *psaA* and *psaB* mRNAs. The interaction of Rbp3 with these mRNAs is relevant for their accumulation level per cell and this should influence the number of PSI units.

Specific transcript regions are enriched in the Rbp3 interactome

The enrichment of certain mRNAs in the analysis shown in **Fig. 2C** gives no information whether certain regions in these transcripts were more highly enriched than others. However, coverage plots of selected mRNAs indicated a highly uneven distribution of reads retained from crosslinking to Rbp3. For the dicistronic *psaAB* transcript, a peak was observed close to the 3′ region and overlapping the stop codon, whereas for *desC* the peak was in the 5'UTR (**Fig. 3**). Enrichment in regions close to the mRNA 3′ end was also observed for *psbA2* (***SI Appendix,* Fig. S3A**), while we observed for *rne* encoding RNase E, a clear peak in the middle of its 583 nt long 5′UTR, beginning exactly at the previously defined U-rich site (37) (***SI Appendix,* Fig. S3B**).

Since no sequence similarity was detected between the peaks, we calculated the secondary structures of selected peak regions and the surrounding segments using RNAfold (38). For the transcript segments enriched for *psaAB* and *psbA2*, stem-loop structures were predicted that included the respective stop codon and short stretches of unpaired nucleotides. These structures are located upstream of the Rho-independent terminators (**Fig. 3B, *SI Appendix,* Fig. S3A**). For the element at the end of the *psaB* mRNA, we tested these interactions by Bio-layer Interferometry (BLI) using recombinant Rbp3 and biotin-labeled 50 nt ribooligonucleotides representing the most-enriched region of the *psaB* mRNA. For comparison, we tested the 50 nt sequences immediately before or after the peak segment (**Fig. 3C)**. The binding height towards the sequence of the coverage peak containing the STOP codon and the predicted Rbp3 binding site was higher than the binding height towards

neighboring sequences. To gain further insight in the likely interacting RNA motifs, we used RRMScorer, a recently developed algorithm that predicts the most probable interacting nucleotide segments (11). RRMScorer identified a 5′-UUNGN-3′ motif (*psaAB*: UUAGG; *psbA2*: UUGGU; *desC*: UUCGA) in the three shown enriched transcript segments with confidence scores between -0.248 and -0.374, placing them clearly in the likely-binder region (11). These motifs include or are directly adjacent to unpaired nucleotide(s). Previously characterized RRM-domain containing RBPs bind single-stranded as well as structured RNA segments (11). Therefore, these results are in line with previous reports on other RRM-containing RBPs. For transcripts encoding carboxysome shell proteins or RuBisCO subunits, enrichment was observed along the entire length of the transcript (***SI Appendix,* Fig. S3D-G**).

Rbp3 co-immunoprecipitates with the ribosome and LabA/NYN-domain proteins

The appearance of differently-sized signals for Rbp3 binding RNA and not binding RNA in our PNK assay (**Fig. 2B**) and the previous observation of involvement in thylakoid membrane targeting (27) posed the question if one or the other Rbp3 fraction would interact with other cellular structures. To identify possibly interacting proteins, co-immunoprecipitation (co-IP) analysis of Rbp3-3xFLAG was performed followed by mass spectrometry-based identification of the enriched proteins. A parallel analysis using sfGFP-3xFLAG as bait protein was used as control. Both analyses were done in triplicates and in total 686 proteins were detected. From these, 64 were significantly enriched ($\log_2$FC > 1, p = 0.05) for co-immunoprecipitating with Rbp3 (***SI Appendix,* Table S2, Fig. 4**). Forty-one of these 62 proteins belong to the ribosome and were enriched with a $\log_2$FC > 2. Four ribosomal proteins could not or only be detected in one or two replicates in the sfGFP control (**Fig. 4B**). Two proteins, Slr1870 and Slr1859 were only detected in the Rbp3 pulldown. Slr1870 is an uncharacterized protein that harbors a LabA/NYN domain in the N-terminal and central two thirds, and an OST-HTH domain in the C-terminal third. Interestingly, with Slr0650 there was another protein with a similar structure as Slr1870 highly enriched ($\log_2$FC of +3.97). The enriched protein Slr1859 was characterized as an anti-sigma factor antagonist (39). Its gene *slr1859* belongs to a cluster that previously was characterized as involved in the regulation of carbon metabolism, particularly glucose metabolism (40). Besides the enrichment of ribosomal proteins, proteins involved in RNA degradation like RNase J

(Rnj) or the polyribonucleotide nucleotidyltransferase (Pnp) were enriched. Another enriched protein was the flavodiiron protein Flv2, which is involved in alternative electron transport and photoprotection (41). The uncharacterized protein Sll0518 that harbors two oligonucleotide/oligosaccharide-binding (OB-fold) domains and is encoded directly downstream of Rbp1, had a $\log_2$FC of +3.69. In addition, seven more such potentially relevant other proteins were enriched, Slr0006, a TsaC fold protein involved in the N6-threonylcarbamoyladenosine modification of tRNA (42), the PilL-C protein Slr0322 (43), the thylakoid-associated ssDNA-binding protein Slr1034, as well as Slr0635, Slr1223 and Slr1854.

The ribosome associated protein Rps1A was significantly enriched whereas Rps1B showed a not significant enrichment due to high enrichment of this protein in the second replicate of the sfGFP pulldown. Both Rps1A and Rps1B were suggested to be involved in the Shine–Dalgarno-independent initiation of translation (44, 45).

To independently verify the association of ribosomes and Rbp3, RNA-FISH analysis of 16S rRNA was performed in cells expressing Rbp3-sfGFP. The Pearson's correlation coefficient in the overlay between the sfGFP and the TAMRA (16S rRNA) signal was close to 1 (median: 0.935), strongly supporting the close interaction between ribosomes and Rbp3 (**Fig. 4C**). FISH labelling of the *psaA* mRNA in the same strain and microscopic analysis revealed a significant overlap with Rbp3-sfGFP fluorescence (**Fig. 4D**). The Pearson's correlation coefficient in the overlay between the sfGFP signal and the *psaA* mRNA was positive (median: 0.352) supporting their interaction. That the correlation was not higher is consistent with the detection of two Rbp3 forms in the PNK assay, one binding RNA and the other non-binding (**Fig. 2B**), as well as the fact that our UV cross-linking experiment revealed interaction also with other transcripts than the *psaAB* mRNA (**Fig. 2C**).

Transcriptomic analyses of the Δ*rbp2*Δ*rbp3* double mutant

To study the effect of a deletion of Rbp3 and to prevent complementation by Rbp2 we analyzed the transcriptome of the Δ*rbp2*Δ*rbp3* double mutant that was used previously in the investigation of mRNA thylakoid membrane targeting (27). The Rbps in cyanobacteria, especially Rbp1, are considered to be important factors in cold stress acclimation (17, 28). However, Rbp2 and Rbp3 are also somewhat cold-inducible (29, 30). Therefore, we analyzed the transcript levels 1 h, 2 h and 4 h after transferring the

cultures from 30°C to 20°C. In addition, since Rbp3 affected photosynthetic properties (**Fig. 1C** and **2D, E**), we performed an analysis 0.5 h, 1 h and 2 h following a shift to high light (250 µmol photons m^{-2} s^{-1}). Subsequently, we performed microarray analyses of selected samples to identify the full set of transcriptomic differences.

Transcriptomic differences in the comparison of Δ*rbp2*Δ*rbp3* versus wild type under non-stress conditions and after shift to 20°C

Effects on *psaAB* mRNA accumulation were already observed 1 h after cold stress induction and 0.5 h after shift to high light (**Fig. 5A, B**). Therefore, we focused on the comparison of transcriptomic data from these time points to the respective non-stress control. Under non-stress conditions, comparing *Δrbp2Δrbp3* to wild type, the levels of 8 transcripts were significantly increased, and of 37 transcripts decreased (**Fig. 5C, *SI Appendix,* Table S3**). With a $\log_2$FC of -0.3, the *psaA* and *psaB* mRNA levels showed only a small decrease under these steady-state conditions **(Fig. 5C**). The most decreased transcript levels ($\log_2$FC < 2.0), besides *rbp2* and *rbp3*, were observed for *sll0788*, which encodes a DUF305 membrane protein and the plasmid pSYSX-located genes *slr6040* and *slr6039* (*hik31a* operon*),* along with a slightly decreased level for *slr6041* (**Fig. 5C**). The two *hik31* operons are involved in the control of heterotrophic growth, of acclimation to dark and low O_2 conditions, and in the responses to several stress conditions (46). Moreover, the *hik31* operon is involved in copper resistance of *Synechocystis* 6803 (47). The eight upregulated transcripts included the complete transcriptional unit *pilA1*/*pilA2*/*sll1696*. In addition, *ndhC* encoding the NDH-1 complex core subunit NdhC was upregulated (which was also enriched in the Rbp3 CLIP-Seq analysis, **Fig. 5C, *SI Appendix,* Table S1**). Interestingly, the *rbp1* transcript was upregulated when *rbp2* and *rbp3* were lacking, indicating a complementary function of the remaining RRM-domain-containing protein. Furthermore, *slr0552* and *slr0601,* encoding proteins of unknown function and *slr0460* encoding an IS701-like transposase were upregulated.

After 1 h of cold stress, 61 transcripts showed a significantly increased level and 64 a significantly decreased level in the double mutant compared to wild type (**Fig. 5D**). Notably, *psaA* ($\log_2$FC = -1.64) and *psaB* ($\log_2$FC = -2.17), the most enriched mRNAs in our CLIP-Seq analysis (**Fig. 2C**), showed the most decreased transcript levels. The levels of two other transcripts encoding PSI proteins were decreased in Δ*rbp2*Δ*rbp3*

as well, *psaJ* (*sml0008*) with a $\log_2$FC of -1.43 and *psaF* (*sll0819*) with a $\log_2$FC of -1.25. Other highly downregulated genes ($\log_2$FC < -2) were again the *hik31a* operon, two mRNAs encoding proteins of unknown function (*sll0788* and *sll8018*), as well as *sll1621* encoding peroxiredoxin. Several downregulated transcripts were identified as Rbp3 interaction partners in CLIP-Seq, such as ATP synthase genes (*atpA*, *atpH, atpG* and *atpD*), mRNAs encoding components of the phycobilisome (*cpcD*), the cytochrome b_6f complex (*petA* and *petD*), and proteins involved in carbon fixation (*rbcS* and *ccmM*) (***SI Appendix,* Table S3**).

Five transcripts showed highly increased levels ($\log_2$FC > 1.5, p = 0.05) in Δ*rbp2*Δ*rbp3* compared to wild type (**Fig. 5D**). Two of these transcripts (*pilA2* and *sll1696*) already showed higher levels under normal conditions. The other three transcripts included the sRNA *csiR1* and *slr1214* encoding the PatA-type response regulator NixC (43), also called LsiR (48). Gene *slr1214* is located downstream of the highly abundant *csiR1* sRNA and does not have an own promoter. The end of the *csiR1* transcript was significantly enriched in the Rbp3-3xFLAG CLIP-Seq pulldown, indicating a regulation at the RNA level (***SI Appendix,* Fig. S3G**). The full transcriptome dataset is visualized in **Dataset 1 Cold**.

Transcriptomic differences in the comparison of Δ*rbp2*Δ*rbp3* versus wild type after shift to high light

In cells grown under standard light conditions, the steady-state levels of 58 transcripts were significantly increased and of 58 transcripts, besides *rbp2* and *rbp3,* significantly decreased ($\log_2$FC ≥ $|1|$, p = 0.05) in the double mutant compared to wild type (**Fig. 5E, *SI Appendix,* Table S4**). Notably, the *hik31a* operon, *sll0788, sll8018*, and *sbtA* and *sbtB*, encoding the bicarbonate transporter SbtA and its regulator SbtB (49), were strongly decreased.

Among the 58 transcripts with increased levels were three non-coding RNAs: *ncl0430*, a small asRNA transcribed from the complementary strand in the *psaA-psaB* intergenic spacer, *ncl0680,* constituting the 5′UTR of the *kaiC2B2* operon and *ncr0860* constituting the 5′UTR of the *hik34* gene (*slr1285).* Moreover, four transcripts encoding proteins of unknown function (*sll1696*, *slr0959*, *slr1915* and *slr1674*). Among the most increased transcripts were *slr0853* encoding the peptide-alanine-alpha-N-acetyltransferase RimI, *rbp1* and *ssr2227*, which encodes a putative transposase.

After 30 minutes of high light exposure (250 µmol photons m^{-2} s^{-1}), the levels of 22 transcripts were significantly increased ($log_2FC > 1$) and the levels of 40 transcripts decreased ($log_2FC < -1$) (**Fig. 5F**). Again, the *hik31a* operon, as well as the transcripts of *sll8018*, *sll1621* and *sll0788* showed the strongest decreases. Additionally, transcript levels of the cell division gene *sepF* (50) were notably lower in Δ*rbp2*Δ*rbp3*. The transcript levels of some photosynthetic proteins showed a strong decrease ($log_2FC < -1.5$), namely *cpcD*, *psaB* and *cpcC1*. These three transcripts were also significantly enriched in the Rbp3 CLIP-Seq dataset, as were other transcripts showing a decreased level upon light stress (***SI Appendix,* Table S1** and **S4**).

Of the 22 transcripts with increased levels, only four exhibited a $log_2FC > 1.5$ and only *rbp1* had a $log_2FC > 2$, underscoring the importance of the only remaining RRM-domain protein in Δ*rbp2*Δ*rbp3*. The other three transcripts with highly increased levels were *sll1696* in the *pilA1A2* operon, the transcript for the phycobilisome degrading protein NblD (51) and the non-coding RNA *ncr0380*. The full transcriptome dataset is visualized in **Dataset 2 HL**.

In summary, in cells deficient in both the *rbp3* and *rbp2* genes, the transcript levels of *psaA* and *psaB* were significantly lower under cold stress and high light stress compared to wild type (**Fig. 5**). The most pronounced reduction in *psaAB* transcript levels occurred 1 h after shifting from 30°C to 20°C (**Fig. 5A**). This decrease was detected by microarray analysis **(Fig. 5C, D, E** and **F**) and northern blot (**Fig. 5A** and **B**). Additionally, Rbp1 transcript levels surged in all double mutant cultures, except for the 1 h cold stress cultures, where Rbp1 expression was already elevated in the wild type. Transcript levels of several stress response proteins were also altered following *rbp2* and *rbp3* deletion.

Discussion

All sequenced cyanobacterial genomes harbor a gene family encoding RRM domain-containing RBPs, typically consisting of three members, but sometimes more, such as eight in the N_2-fixing, multicellular *A. variabilis* (25). RNA binding was previously verified for these eight Rbps using an *in vitro* assay (25). More recently, *in vitro* RIP-Seq analyses of *Synechocystis* 6803 recombinant Rbp1 and Rbp2 after total RNA incubation identified transcripts mainly related to iron uptake (24).

Here, we show that Rbp3, one of the three different RRM-type RBPs in *Synechocystis* 6803, binds RNA *in vivo* (**Fig. 2B**) and used CLIP-Seq to determine the interacting RNA population (**Fig. 2C**). The results are consistent with our previous report that Rbp3 is involved in targeting mRNAs encoding the core subunits of both photosystems to the thylakoid membrane, and demonstrate that Rbp3 physically interacts with them. Our CLIP-Seq results also identified several other mRNAs interacting with Rbp3. Most prominent were mRNAs encoding linker proteins for the phycobilisome such as *apcE, cpcC1* and *cpcL*, as well as a protein of the ATP synthase complex that resides in the same membrane system. It is not unlikely that thylakoid membrane interactions of these proteins follow similar principles. This is in line with recent observations in an *Anabaena* sp. PCC 7120 deletion mutant lacking RbpG, the ortholog of *Synechocystis* 6803 Rbp3, which showed perturbed thylakoid membrane organization and function (52).

However, some mRNAs enriched by CLIP-Seq, such as *epsB*, encode proteins that are more likely to interact with the plasma membrane and even non-membrane proteins, most notably carboxysome shell proteins and RuBisCO (*rbcLXS* operon, see ***SI Appendix,* Fig. S3D**). These findings suggest a broader role for Rbp3, while supporting its previously proposed function, along with Rbp2, in localizing specific photosynthetic mRNAs to the thylakoid surface (27).

As interacting sites within the bound transcripts, we found mainly predicted folded secondary structure elements in which the RRMscorer algorithm identified U-rich pentanucleotide motifs (UUNGN) with at least one unpaired nucleotide or U-G base pair. Eukaryotic RRM proteins are highly flexible, typically binding single-stranded RNA, and occasionally single-stranded DNA (53, 54), although structured RNA elements are also documented (11). It is interesting to note that sequences within mRNA 3′UTRs have been recognized to contribute to the correct targeting of the encoded proteins to intracellular membranes (55, 56). The term memRBPs has been proposed for the "membrane-associated RNA-binding proteins" involved (57), and well-characterized memRBPs, like the fungal Rrm4, contain RRM domains (58). The analogy between Rbp3, which interacts with the 3′UTRs of mRNAs encoding key proteins of the photosynthetic machinery, and eukaryotic RRM proteins, which are involved in transport by targeting stem-loop elements in mRNA 3′UTRs, is striking. Thus, Rbp3 can be considered as a bacterial memRBP.

To perform their transport functions, memRBPs must interact with other cellular structures; for instance, Rrm4 shuttles along microtubules (58, 59). Although several elements of these mRNA transport systems have been elucidated, they are far from fully understood (60).

Rbp3 likely does not function independently of other proteins. Mechanisms must exist to sort and recognize the targeted membrane or cellular compartment and to prevent premature translation. Therefore, we enriched proteins by co-IP to identify those interacting directly with Rbp3 or forming larger complexes nearby. A clear result of this approach was the enrichment of the entire set of ribosomal proteins (**Fig. 4**), indicating that complete ribosomes, not just subunits, were near Rbp3. However, recent fractionation of ribonucleoprotein complexes by Grad-Seq did not indicate co-fractionation of Rbp3 and ribosomes (7, 8). Thus, while a direct association of Rbp3 with ribosomes cannot be definitively excluded, there is no direct evidence supporting it, leading to the model shown in **Fig. 6A** in which Rbp3 is linked to ribosomes via a bound mRNA molecule.

For mRNAs encoding PS core subunits, Rbp3 interacts with regions close to, or overlapping the stop codons. These mRNAs are highly translated and associated with polysomes, explaining the enrichment of ribosomal subunits in the Rbp3 co-IP. When the first ribosomes reach the stop codon region, their helicase activity likely unwinds the secondary structures where Rbp3 resides. As a result, it may be released and initiate interaction with another mRNA. Nascent polypeptides probably begin inserting into the membrane, keeping the mRNA-ribosome complexes associated with thylakoid surfaces, even after Rbp3 is released from the 3′ region. However, somehow membrane contact must be established and maintained. A conserved PDPRWA sequence motif within the C-terminal region of Rbp3 orthologs (24) is located within a large intrinsically disordered region of the protein (**Fig. 6B**). Prediction of putative lipid binding sites using the DisoLipPred algorithm, which was also trained on bacterial datasets (61), yielded a high score, suggesting its possible involvement in the interaction of Rbp3 with the membrane surface.

An unexpected finding in our co-IP experiments was the enrichment of Slr1870 and Slr0650, two structurally similar proteins with a LabA/NYN domain and an additional OST-HTH domain in Slr1870. A third such protein, Slr0755, was not enriched. The OST-HTH domain was predicted as RNA-binding and some eukaryotic proteins with that domain nucleate or organize ribonucleoprotein complexes (62). The relationship

between Rbp3 and these proteins is unclear, but a similar structure is found in PIN-related RNases, which may be involved in tRNA or rRNA maturation. These proteins belong to the low-amplitude and bright (LabA) superfamily, including the *Synechococcus elongatus* LabA, which is linked to circadian timing (63). A LabA paralog, LalA, was described in *S. elongatus*, which is also involved in circadian gene expression, as well as in cell growth regulation (64). Recently, the *S. elongatus* Rbp2 homolog was reported to interact with the circadian clock protein KaiC, and disruption of its gene resulted in altered gene expression rhythms and KaiC localization (65). This suggests a connection between the cyanobacterial Rbps and the circadian clock system; however, Slr0755, the ortholog of *S. elongatus* LabA was not enriched in our co-IP, and the KaiC-interacting protein was Rpb2 and not Rbp3 addressed here.

Another interesting interacting protein was Slr0006, which binds tightly to membranes (66) and colocalizes with 30S ribosomal proteins and rRNA (67). Slr0006 has a TsaC fold, indicating its role in the N6-threonylcarbamoyladenosine modification in most ANN-decoding tRNAs (42), aligning with the enrichment of ribosomal proteins and other RNA maturation factors.

There is evidence of cross-talk between the three Rbps in *Synechocystis* 6803, as indicated by the upregulation of Rbp1 in the Δ*rbp2*Δ*rbp3* double mutant (**Fig. 5C, E, F**) and the fact that it is possible to generate double deletion mutants between any combination of these *rbp* genes, but not triple deletions. How this cross-talk is mediated is currently unknown, but we noticed the co-IP of possible regulatory proteins with Rbp3 such as Sll0518 (**Fig. 4A**). The gene *sll0518* is located directly downstream of *sll0517* encoding Rbp1, making it a candidate for possible cross-talk between different Rbps, especially during cold stress acclimation.

Similar targeting mechanisms probably exist for plasma membrane proteins (68) and, as our data indicate, potentially for cellular microcompartments, like carboxysomes. While some components must be distinct, others might be shared. Clarification of the functions of the here identified non-ribosomal proteins and of the transcript segments that distinguish them from other transcripts is a promising topic for future research.

Materials and Methods

Strains and culturing conditions

Synechocystis 6803 GT-I (69) wild type and mutant strains were cultivated in copper-free BG11 supplemented with 20 mM TES buffer pH 7.5, under continuous white light of 50 µmol photons m^{-2} s^{-1} at 30°C (standard conditions). Mutant strains containing pVZ322 were cultivated in the presence of 50 µg/mL kanamycin and 2.5 µg/mL gentamycin, while Δ*rbp3* was kept at 50 µg/mL kanamycin. The Δ*rbp2*Δ*rbp3* double mutant strain was kept at 10 µg/mL chloramphenicol and 50 µg/mL kanamycin in the pre-cultures. Antibiotics were omitted during experimental assays. To test the acclimation responses of the Δ*rbp2*Δ*rbp3* double mutant and wild type to downshifts in temperature or to upshifts in light intensity, cultures were transferred at an OD_{750} of ~ 1.0 from 30°C to 20°C, or at an OD_{750} of ~0.6 from a light intensity of 50 to 250 µmol photons m^{-2} s^{-1}. Kanamycin was omitted during co-IP experiments. For mRNA-FISH and Rbp3-GFP localization microscopy, *Synechocystis* 6803 PCC-M (70) was used due its larger cell size, facilitating high-resolution microscopy and protein localization.

Construction of mutant and tagged strains

Synechocystis 6803 GT-I (69) served as the wild type and background for constructing Δ*rbp3* and Δ*rbp2*Δ*rbp3* mutants (27). Strains harboring the pVZ322 plasmid with the *rbp3-3xFLAG* gene under control of the *petE* promoter (P_{petE}) were previously described (27). To construct the sfGFP-3xFLAG expression plasmid, the *sfGFP* sequence with overlaps to P_{petE} was PCR-amplified from the pXG10-SF plasmid (71) (primer and plasmid details in ***SI Appendix,*** **Table S5**). For subcloning, pJet1.2-P*petE-slr0193* (27) was used, replacing the *slr0193* sequence with *sfGFP* using AQUA cloning. The constructed vector was subjected to inverse PCR to introduce the 3xFLAG tag to the C-terminus of sfGFP followed by *Dpn*I digest (Thermo Fisher Scientific), 5′ phosphorylation (T4 PNK, Thermo Fisher Scientific), self-ligation (T4 DNA ligase, Thermo Fisher Scientific) and heat-shock transformation into *E. coli* DH5α. The P_{petE}-*sfGFP*-3xFLAG fragment was amplified and inserted by AQUA cloning into the self-replicating, multi-host vector pVZ322 linearized by restriction with *Hind*III and *Xho*I. To obtain the complementation strain with the pVZ322::*rbp3* plasmid, the *rbp3* gene was amplified from genomic DNA including the native promoter and 3′UTR using specific primers (***SI Appendix,*** **Table S5**). Cloning into pVZ322 and

transformation of *Synechocystis* 6803 GT-I was done as described for the strain expressing the triple FLAG-tagged Rbp3 (27).

To generate *rbp3* knockouts in the PCC-M strain, total gDNA of Δ*rbp3* GT-I cells containing the *rbp3* knockout cassette was used to transform PCC-M wild-type cells. Transformants were selected on kanamycin (50 µg/mL) containing BG11 plates. Complete segregation was confirmed by colony PCR (primers FP and RP, ***SI Appendix,* Table S5**).

To generate PCC-M Δ*rbp3* + P_{rbp3}-Rbp3-sfGFP strains, the *rbp3* knockout strain was transformed with the pVZ322 plasmid containing the *rbp3-sfGFP* gene under the control of the native *rbp3* promoter (P_{rbp3}). To obtain the sfGFP-tagged version of Rbp3, the 3xFLAG fragment in the *rbp3-3xFLAG* construct was replaced by the *sfGFP* gene by AQUA cloning. Successful exchange of the 3xFLAG segment by the *sfGFP* gene was confirmed by PCR and by Sanger sequencing (GATC, Eurofins). Deletion of *rbp3* in PCC-M background led to smaller cells and a reduced chlorophyll content similar to the *rbp3* mutant in the GT-I strain, a phenotype which could be complemented by Rbp3-GFP indicating that the GFP fusion did not interfere with the functionality (***SI Appendix,* Fig. S4**).

Protein induction, cell harvest and lysis

The expression of *rbp3-3xFLAG* as well as *sfGFP-3xFLAG* was induced in corresponding culture volumes at an OD_{750} of ~ 0.6 by adding Cu_2SO_4 to a final concentration of 0.31 µM. After another 24 h of cultivation, the cells were harvested by centrifugation (4,000 *x* g, 4°C, 20 min). Cell pellet is resuspended in the respective lysis buffer. For cell lysis, mechanical disruption using a prechilled Precellys homogenizer (Bertin Technologies) was used. To remove intact cells and glass beads, the culture was centrifuged (800 *x* g, 5 min, 4°C), and the supernatant was collected for further processing. To solubilize membrane proteins, the cell lysates were incubated for 1 h in the presence of 2% n-dodecyl β-D-maltoside in the dark at 4°C. Afterwards, cell debris was removed by centrifugation (21,000 *x* g, 4°C 1 h). Supernatant was then used respectively.

Production of recombinant Rbp3

To obtain the intact recombinant Rbp3 protein, it was first expressed as a GST fusion protein (GST-Rbp3) and then the GST tag was removed. The *rbp3* gene (*slr0193*) was amplified and ligated into vector pGEX-6P-1 using the In-Fusion Cloning kit (TaKaRa, Shiga, Japan) and the PCR primers pGEX-6P-1_inv-f/-r and Rbp3_iVEC-f/-r (***SI Appendix,* Table S5**). After verification of the sequences, the plasmids were introduced into *E. coli* Rosetta competent cells (TaKaRa) and used for the expression and purification of recombinant GST-Rbp3 protein as follows. The resulting *E. coli* strain was grown in 2 L of LB medium at 37°C. When the culture reached an OD_{600} of 0.6, IPTG was added to the medium at a final concentration of 1 mM and the growth temperature was shifted to 15°C. After 24 h, cells were harvested by centrifugation, washed with purification buffer (10% w/v glycerol, 100 mM NaCl, 20 mM HEPES-NaOH), and stored at −80°C until use. For protein purification, frozen cells were suspended in 200 mL of purification buffer. The cells were disrupted by sonication and centrifugated at 10,460 × *g* for 30 min at 4°C, and protein affinity purification was performed using an ÄKTA start chromatography system with a GSTrap column (Cytiva, MA, USA) according to the manufacturer's instructions. The protein was eluted with 50 mM reduced glutathione and dialyzed against the dialysis buffer 1 (100 mM HEPES-KOH (pH 7.6), 100 mM NaCl, 1 mM 1,4-dithiothreitol (DTT), dialysis buffer 2 (50% w/v dialysis buffer 1, 50% w/v glycerol). Following dialysis, GST-Rbp3 was cleaved by PreScission Protease (Cytiva) to remove the GST-tag and dialyzed again. The purity of the protein was confirmed by SDS–PAGE and used for the BLI analysis.

SDS-PAGE and Western blotting

Protein separation was done on an SDS-PAGE and protein detection with a subsequent western blot as described before (72). In short, after co-IP beads in elution buffer were mixed in a 5:1 ration with 5x protein loading dye (25 mM Tris-HCl pH 6.8, 25% glycerol, 10% SDS, 50 mM DTT, 0.05% bromophenol blue) and boiled for 5 min at 95°C. Equal volumes of supernatant were separated on a 15% SDS-PAA gel (73). After separation the proteins were transferred to a nitrocellulose membrane (Hybond™-ECL, Cytiva) at 1.5 A cm^{-2} for 1 h. After transfer, the membrane was blocked in 5% skimmed milk powder in TBS-T (50 mM Tris-HCl, 150 mM NaCl, 0.05% Tween20, pH 7.5) overnight at 4°C or for 1 h at room temperature. After blocking, the membrane was washed three times with TBS-T before incubating with mouse monoclonal anti-FLAG coupled to HRP (#A8592, Sigma) at a 1:5,000 titer in

TBS-T. After incubation for 1 h at room temperature, signals were visualize using Western-Blot ECL spray (Advanta) and the FUSION SL Transilluminator (Vilber Lourmat).

RNA preparation, Northern blotting and hybridization

Culture for RNA extraction was harvested by vacuum filtration on a hydrophilic polyether sulfone filter (Pall Supor-800) and snap frozen in liquid nitrogen. For RNA extraction, 1 mL PGTX (74) was given to the filter followed by an incubation of 15 min at 65°C. RNA was washed and precipitated as described previously (75).
The RNA concentration was determined by a Nanodrop spectrophotometer (peqlab). Equal amounts of RNA were diluted in water and 1:1 mixed with 2x RNA Gel Loading Dye (Thermo). Afterwards the samples were denatured at 72°C for 5 min.
RNA was separated on a 1.5% formaldehyde-agarose gel and RNA transfer to a positively charged nylon membrane (Amersham Hybond-N+, GE Healthcare) occurred overnight using a capillary blot system. Synthesis of probes and detection of radioactive signal has been described before (72).

Microarray analysis

The here used Agilent microarrays have been designed for the direct detection of transcript after chemical labeling, without cDNA synthesis (76). The microarrays contain oligonucleotide probes for all annotated mRNAs and the majority of non-coding RNAs. Total RNA (5 µg) extracted from the wild type and deletion mutants grown at standard temperature (30°C) and after 1 h of cold stress (20°C) or after 0.5 h at high light was directly labeled using the ULS labeling kit for Agilent gene expression arrays with Cy3 according to the manufacturer's protocol for Agilent one color microarrays (Kreatech Diagnostics, B.V., Netherlands). Array analysis was performed using two biological replicates and an internal technical replicate. Raw data were normexp background corrected (77) and quantile normalized using the limma R package (78). The transcriptome differences were calculated between wild type and mutant for the two different temperatures and the two different light conditions. Significance criteria for differential expression were $\log_2 FC \geq |1|$, adj. p -value ≤ 0.05. P-values were adjusted for multiple testing with the Benjamini-Hochberg method. The array data has been deposited in the GEO database under the accession number GSE269515.

Protein preparation, proteomic sample preparation and analyses by MS

After harvesting, the cell pellet was resuspended in 1.5 mL FLAG-MS buffer (50 mM HEPES/NaOH pH 7, 5 mM $MgCl_2$, 25 mM $CaCl_2$, 150 mM NaCl) containing protease inhibitor (c0mplete, Roche). The cells were lysed as described above. Co-IP was performed by incubating 50 µL packed volume of Anti-FLAG M2 magnetic beads (Sigma) with the cleared cell lysate (1 h, 4°C, rotating). Afterwards the supernatant was removed using the DynaMag™-2 Magnet rack (Thermo Fisher Scientific) and the magnetic beads were washed 6x with FLAG-MS buffer.

Proteins were further processed according to the manufacturer's protocol for S-Trap micro filters (Protifi), and using the provided buffers. Briefly, 60 µL of the 1x lysis buffer (5% SDS, 50 mM TEAB pH 8.5) was added to the beads and incubated at 95°C for 10 min to elute the proteins. In the following process 46 µL of eluate were used. Proteins were reduced by incubation with reduction buffer (end concentration 5 mM Tris(2-carboxyethyl)phosphine (TCEP)) for 15 min at 55°C and alkylated by incubation with alkylation buffer (end concentration 20 mM methyl methanethiosulfonate (MMTS)) for 10 min at room temperature in the dark. Afterwards, a final concentration of 1.2% phosphoric acid and then six volumes of binding buffer (90% methanol; 100 mM TEAB; pH 7.55) were added. After gentle mixing, the protein solution was loaded to an S-Trap filter and spun 3 times at 10,000 *x* g for 30 sec. The filter was washed three times using 150 µL of binding buffer. Sequencing-grade trypsin (Promega, 1:25 enzyme:protein ratio) diluted in 20 µL digestion buffer (50 mM TEAB) were added to the filter and digested at 47°C for 2 h. To elute peptides, three step-wise buffers were applied: a) 40 µL 50 mM TEAB, b) 40 µl 0.2% formic acid in H_2O, and c) 50% acetonitrile and 0.2% formic acid in H_2O. Each step was eluted by centrifugation (10,000 *x* g for 1 min). The peptide solutions were combined and dried in a DNA SpeedVac DNA120 (Savant). For LC-MS/MS measurements 800 ng of peptides were analyzed on an LTQ-Orbitrap Elite mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-nLCTM 1000 UHPLC system (Thermo Fisher Scientific). The column setup consisted of a Acclaim™ PepMap™ 100 C18 column (Thermo Fisher Scientific, Cat. No. 164946) and a 200 cm µPac GEN1 analytical column (PharmaFluidics, 55250315018210) coupled to a Nanospray FlexTM ion source (Thermo Scientific, ES071) and a fused silica emitter (MS Wil, TIP1002005-5).

For peptide separation, a linear gradient of buffer B (0.1% formic acid in 80% acetonitrile, Fluka) was applied, increasing from 8 to 55% over 80 min and from 55 to 100% over the next 40 min. Peptides were analyzed using data-dependent acquisition (DDA). In the positive ion mode, a full scan ranging from m/z 400 to 1600 with a resolution of 60,000 and a maximum ion accumulation time of 40 ms was performed. The top 15 precursors of the highest abundance in the full scan were selected and fragmented by collision-induced dissociation (CID) function and analyzed in MS/MS with a normalized collision energy of 35%, 10 ms activation time, and a minimum TIC threshold of 10,000 for MS/MS selection. The following dynamic exclusion settings included: repeat counts 1; repeat duration 15 s; exclusion duration 35 s.

Raw data were analyzed with MaxQuant (v 1.6.17.0) with the built-in Andromeda peptide search engine (79). The false discovery rate (FDR) at both the protein and peptide level was set to 1%. Two missed cleavage sites were allowed, no variable modifications were allowed, and carbamidomethylation of cysteines was set as fixed modification. The match between runs option was selected. For label free quantification the MaxLFQ algorithm was applied using the standard settings. Only unique peptides were used for quantification. The database "Synechocystis sp. (strain PCC 6803 / Kazusa)" was downloaded from https://www.uniprot.org/proteomes/UP000001425 on Jun 9th, 2022.

Data were normalized on peptide level by equalizing the medians using the MSstats package (v. 3.20.3) in R (v. 4.0.4) (80). Subsequently, protein LFQ intensities were $\log_2$ transformed. P-values were adjusted using the Benjamini-Hochberg procedure.

The proteome raw data acquired by MS were deposited at the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (81) under the identifier PXD045584. The intensities were compared using LFQ (label-free quantification) values (82) with Perseus (version 1.6.1.3 (83)). In summary, contaminants, reverse sequences and proteins only identified by site were removed from the matrix and LFQ intensities were $\log_2$-transformed. Before t-test and visualization using a volcano plot, the missing values were replaced by imputation with the normal distribution for each column separately (default settings). For hierarchical clustering (default parameters), only proteins with three valid values in at least one declared group (Rbp3_3×FLAG and sfGFP_3xFLAG) were considered.

Crosslinked immunoprecipitation and sequencing (CLIP-Seq)

CLIP-Seq was carried out as described (84). Rbp3-3×FLAG and sfGFP-3×FLAG expression was induced in 250 mL *Synechocystis* cultures with 0.31 µM Cu_2SO_4 at an OD_{750} of 0.6-0.7. After 24 h, the cultures were transferred to 21 x14.5 x 5.5 cm^3 plastic trays and crosslinked in a UV Stratalinker 2400 (Stratagene) three times at 0.45 J/cm^{-2}. 25 mL culture aliquots were collected to prepare total RNA for RNA-seq using the hot-phenol extraction method (75). The remaining culture was collected as described above. Cell pellets were resuspended in 1 mL lysis buffer (20 mM Tris/HCl pH 8.0, 1 mM $MgCl_2$, 150 mM KCl, 1 mM DTT) with 1 U of RNase inhibitor per sample (RiboLock, Thermo Fisher Scientific) and lysed as described above. Supernatants were co-immunoprecipitated by incubation with 50 µL packed bead volume of Anti-FLAG M2 magnetic beads (Sigma) for 1 h at 4°C.The DynaMag™-2 magnetic rack (Thermo Fisher Scientific) was used for flow-through removal and washing. Beads were washed twice with FLAG buffer, then twice with high salt FLAG buffer (50 mM HEPES/NaOH pH 7, 5 mM $MgCl_2$, 25 mM $CaCl_2$, 1 M NaCl, 10% glycerol, 0.1% Tween20), and then washed two more times with FLAG buffer. Beads were subjected to proteinase K digestion (1% SDS, 10 mM Tris-HCl pH 7.5, 100 µg/mL proteinase K) for 30 min and 650 rpm at 50°C. Then, 200 µL PGTX (74) was added and the samples were boiled for 15 min at 65°C to extract protein-bound RNA. After adding 140 µL chloroform/isoamyl alcohol (24:1), the samples were incubated for 10 min at room temperature with agitation. For phase separation, the samples were centrifugated in a swing-out rotor at 3,260 x g for 3 min at room temperature. The upper aqueous phase was transferred to an RNase-free tube and mixed in a 1:1 ratio with 100% EtOH and processed via the clean and concentrator kit (Zymo Research). RNA was eluted in 15 µL RNase-free ddH_2O. Construction of cDNA libraries and sequence analysis were performed by a commercial provider (Vertis Biotechnologie AG Freising, Germany). The RNA samples were fragmented using ultrasound (4 pulses of 30 s each at 4°C). An oligonucleotide adapter was ligated to the RNA 3' ends. M-MLV reverse transcriptase and a 3′ primer complementary to the adapter sequence were used for first-strand cDNA synthesis. The obtained cDNA was purified and the 5' Illumina TruSeq sequencing adapter ligated to the cDNA 3' ends. The cDNA was amplified by 13 cycles of PCR (15 to 18 for the negative control), to about 10-20 ng/µL using a high-fidelity DNA polymerase and introducing TruSeq barcode sequences as part of the 5' and 3' TruSeq sequencing adapters. The cDNA was purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics) and analyzed by capillary

electrophoresis. For Illumina NextSeq sequencing, the samples were pooled in approximately equimolar amounts. The cDNA pool was 200 – 600 bp size-fractionated via a preparative agarose gel and sequenced on an Illumina NextSeq 500 system with a read length of 75 bp.

Raw reads of the RNA-seq libraries were analyzed using the galaxy platform. The galaxy workflow can be accessed and reproduced at the following link: https://usegalaxy.eu/u/luisa_hemm/w/poi-3xflag-clip. In short, read quality was check with FastQC in the beginning. After the quality check, reads were mapped to the *Synechocystis* chromosome (GenBank accession NC_000911.1) and the 4 large plasmids pSYSA, pSYSG, pSYSM and pSYSX (NC_005230.1, NC_005231.1, NC_005229.1, NC_005232.1) using BWA-MEM (85) with default parameters. Resulting data were filtered using the BAM filter (86) to remove unmapped reads as well as reads smaller than 20 nucleotides. The annotation gff-file was based on the reference GenBank annotations complemented by additional genes based on transcriptomic analyses (87). htseq-count (88) (mode union, nonunique all, minaqual 0) was used to count the reads per transcript. Only transcripts with ≥ 10 reads in at least three samples were subjected to differential expression analysis with DESeq2 (32). To visualize coverage of the reads, bam files were transformed into wiggle files using RSeQC (89) and summarizing the reads from the triplicate analyses for each position.

PNK assay

Procedures followed previously described methods (7, 31). Rbp3-3×FLAG and sfGFP-3×FLAG expression was induced in 250 mL *Synechocystis* cultures with 0.31 µM Cu_2SO_4 at an OD_{750} of 0.6 to 0.7. After 24 h, cultures were transferred to 21 x14.5 x 5.5 cm^3 plastic trays on ice and crosslinked three times at 0.75 J/cm^{-2} using a UV Stratalinker 2400 (Stratagene) to induce covalent bonds between protein and RNA. Negative controls were not crosslinked. Cells were collected by centrifugation and the pellets were resuspended in 800 µL of NT-P buffer (50 mM NaH_2PO_4, 300 mM NaCl, 0.05% Tween, pH 8.0) containing c0mplete protease inhibitor (Roche). Cell lysis supernatants were co-immunoprecipitated by incubation with 80 µL packed bead volume of Anti-FLAG M2 magnetic beads (Sigma) for 1 h at 4°C with slight rotation. The DynaMag™-2 magnetic rack (Thermo Fisher Scientific) was used for flow-through

removal and washing using 2 mL per wash step. Beads were washed twice with high salt NT-P buffer (50 mM NaH_2PO_4, 1 M NaCl, 0.05% Tween, pH 8.0), once with NT-P buffer and once with benzonase buffer (50 mM Tris-HCl, 1 mM $MgCl_2$, pH 8). Beads were resuspended in 100 µL benzonase buffer containing 25 U benzonase (Sigma). After incubation for 10 min at 37°C and 900 rpm, beads were washed once with high-salt NT-P buffer and twice with FastAP buffer (10 mM Tris-HCl, 5 mM $MgCl_2$, 100 mM KCl, pH 8). Beads were then resuspended in 150 µL FastAP buffer and incubated with 2 U Fast-AP (Thermo) for 30 min at 37°C and 900 rpm. Beads were washed once with high-salt NT-P buffer and twice with PNK buffer (50 mM Tris-HCl, 10 mM MgCl2, 0.1 mM spermidine, pH 8) and then resuspended in 100 µL PNK buffer containing 10 µCi of [γ-^{32}P]-ATP and 1 µL PNK (Thermo). After incubation at 37°C for 30 min at 900 rpm beads were washed once with high salt NT-P buffer and twice with NT-P buffer. For elution, beads were resuspended in 35 µL TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7.5) with 0.5 µg/µL FLAG peptide (Sigma) and incubated at room temperature for 15 min at 900 rpm. To the supernatant 5 µL protein loading dye (25 mM Tris-HCl pH 6.8, 25% glycerol, 10% SDS, 50 mM DTT, 0.05% bromophenol blue) was added and samples were boiled for 5 min at 95°C. Proteins were separated by SDS-PAGE and western blotting was done as described above. After transfer to the membrane, the membrane was exposed for several hours to a phosphoimaging screen (Kodak) and visualized using a Typhoon FLA 9500 (GE Healthcare), followed by western blot incubation.

Spectrometric measurements

Whole-cell absorption spectra were obtained using an SLM Aminco DW-2000 spectrophotometer. Cyanobacteria were grown to mid-log phase in BG11 medium supplemented with 20 mM TES buffer pH 7.5 under continuous white light at 15-20 µmol photons m^{-2} s^{-1} at 30°C. Spectra were recorded from 400 nm to 750 nm and normalised at 750 nm. Data presented are the representative of similar results from 3 independent cultures. Fluorescence emission spectra were recorded at 77 K with a Perkin Elmer LS55 fluorescence spectrometer with a liquid nitrogen housing. Cultures were grown as for the absorption spectra, concentrated to 5 µM chlorophyll in BG11 medium and loaded into silica capillary tubes. The tubes were dark-adapted for 10 min and then snap frozen in liquid nitrogen. Fluorescence emission spectra were recorded

from 620-750 nm with slit-width 3-5 nm and excitation at 435 nm for chlorophyll *a* or at 600 nm for phycocyanin. Spectra were normalized to the PSI peak (~735 nm) or to the phycocyanin/allophycocyanin peak (~660 nm) and are representative of 3 independent and similar results.

Bio-Layer Interferometry

Recombinant Rbp3 protein interactions with RNA oligonucleotide ligands were examined using the BLI system, BLItz (Fortebio). Ligand RNA oligonucleotides were synthesized with a biotinylated 5' end (Eurofins) and used for BLI analysis. The binding buffer (4% glycerol, 10 mM Tris-HCl pH 7.5, 1 mM $MgCl_2$, 0.5 mM EDTA, 50 mM NaCl, 0.5 mM DTT, 0.05 mg/mL poly(dI-dC)·poly(dI-dC), 0.05% [w/v] BSA, 0.01% Tween 20) was used to dilute both ligand RNA and analyte Rbp3 protein. Experiments were conducted at room temperature. 0.5 pmol of the biotinylated ligand RNA oligonucleotides were captured on a streptavidin biosensor (Sartorius) that had been equilibrated with binding buffer. Rbp3 was bound to the ligand RNA-reacted biosensor proteins at concentrations of 0.5, 1, 2 or 3 pmol/µL, and the association was analyzed. The same binding buffer was used for the dissociation step.

mRNA fluorescent *in situ* hybridization

mRNA-FISH was performed as described (90), with adjustments for *Synechocystis* 6803 according to reference (27). Briefly, the target RNA was hybridized with 44–48 oligonucleotides, each 20 nt long, fused to a TAMRA fluorophore at the 3' end. Oligonucleotide probes for mRNA-FISH were sourced from LGC Biosearch Technologies (***SI Appendix,* Table S6)**. For further procedural details, including cell fixation, permeabilization, hybridization and buffer compositions, see reference (27).
For imaging hybridized *Synechocystis* 6803 PCC-M cells, fixed samples were mounted on a high precision cover glass covered by a 1.5% agarose pad to minimize cell movement. Imaging was conducted using an LSM 880 confocal microscope equipped with an Airyscan detector (Carl Zeiss) and a Plan-Apochromat 63×/1.4NA oil objective lens. Images were recorded in 16-bit format, at 488 × 488 pixels resolution, 8x line averaging and a zoom value of 10. The 561 nm, 488 nm, and 633 nm lasers were used to excite the TAMRA fluorophore, Rbp3-sfGFP fusion protein, and thylakoid membrane photosynthetic pigments, respectively. To prevent signal bleed-through, a secondary

beam splitter and filter set were installed before the Airyscan detector. A short pass 615 nm beam splitter specified TAMRA emission, and a band pass filter isolating emission from 495–550 nm was used for sfGFP emission. Image reconstruction was performed using the Airyscan processing software with automatic strength.

Prior to in-depth image analysis, autofluorescence from RNA-FISH and Rbp3-sfGFP was removed by background subtraction as described (27). The ImageJ plugin EzColocalization was used to determine the cell area, calculate mean fluorescence intensity and perform colocalization analysis (91, 92). Binary images of individual cell outlines were created from the thylakoid membrane channel using the auto threshold command. The adjustable watershed with a tolerance of 5 was applied to distinguish between adjacent cells. Cells not fully captured in the images were excluded from further analysis. Colocalization of hybridized target mRNA and Rbp3-sfGFP fluorescence was assessed by calculating the Pearson correlation coefficient using Costes thresholding (93).

Data availability

The CLIP-seq data have been deposited in the SRA database under the accession number PRJNA1104562. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (94) via the PRIDE (81) partner repository with the dataset identifier PXD045584. Furthermore, all mass spectrometry proteomics datasets used and/or analyzed during this study are available online at the MassIVE repository (http://massive.ucsd.edu/; dataset identifier: MSV000092937). The microarray data have been deposited in the GEO database under the accession number GSE269515.

Funding

We appreciate the support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) to LH, EL, OS, AW and WRH through the graduate school MeInBio - 322977937/GRK2344. OS acknowledges support by DFG (SCHI 871/11-1). The Proteomic Platform – Core Facility was supported by the Medical Faculty of the University of Freiburg to O.S. (2021/A3-Sch). CWM acknowledges support from the Biotechnology and Biological Research Council (BB/W001012/1).

References

1. D. L. Kaplan, M. O'Donnell, Rho factor: transcription termination in four steps. *Curr. Biol. CB* **13**, R714-716 (2003).
2. H. Aiba, Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr. Opin. Microbiol.* **10**, 134–139 (2007).
3. S. G. Chaulk, *et al.*, ProQ is an RNA chaperone that controls ProP levels in *Escherichia coli*. *Biochemistry* **50**, 3095–3106 (2011).
4. W. Bae, B. Xia, M. Inouye, K. Severinov, *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7784–7789 (2000).
5. A. Bøggild, M. Overgaard, P. Valentin-Hansen, D. E. Brodersen, Cyanobacteria contain a structural homologue of the Hfq protein with altered RNA-binding properties. *FEBS J.* **276**, 3904–3915 (2009).
6. N. Schuergers, *et al.*, Binding of the RNA chaperone Hfq to the type IV pilus base is crucial for its function in *Synechocystis* sp. PCC 6803. *Mol. Microbiol.* **92**, 840–852 (2014).
7. M. Brenes-Álvarez, *et al.*, RNA-binding proteins identified by R-DeeP/TripepSVM are involved in heterocyst differentiation. [Preprint] (2024). Available at: https://www.biorxiv.org/content/10.1101/2024.04.03.587981v1 [Accessed 25 April 2024].
8. M. Riediger, *et al.*, Analysis of a photosynthetic cyanobacterium rich in internal membrane systems via gradient profiling by sequencing (Grad-seq). *Plant Cell* **33**, 248–269 (2021).
9. G. Dreyfuss, M. S. Swanson, S. Piñol-Roma, Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. *Trends Biochem. Sci.* **13**, 86–91 (1988).
10. C. Maris, C. Dominguez, F. H.-T. Allain, The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* **272**, 2118–2131 (2005).
11. J. Roca-Martínez, H. Dhondge, M. Sattler, W. F. Vranken, Deciphering the RRM-RNA recognition code: A computational analysis. *PLOS Comput. Biol.* **19**, e1010859 (2023).
12. A. Zhang, W.-F. Liu, Y.-B. Yan, Role of the RRM domain in the activity, structure and stability of poly(A)-specific ribonuclease. *Arch. Biochem. Biophys.* **461**, 255–262 (2007).
13. A. B. Sachs, R. W. Davis, R. D. Kornberg, A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. *Mol. Cell. Biol.* **7**, 3268–3276 (1987).
14. L. Olgeiser, *et al.*, The key protein of endosomal mRNP transport Rrm4 binds translational landmark sites of cargo mRNAs. *EMBO Rep.* **20**, e46588 (2019).
15. M. Grskovic, M. W. Hentze, F. Gebauer, A co-repressor assembly nucleated by Sex-lethal in the 3′UTR mediates translational control of Drosophila msl-2 mRNA. *EMBO J.* **22**, 5571–5581 (2003).
16. G. Caponigro, R. Parker, Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. *Genes Dev.* **9**, 2421–2432 (1995).
17. K. Maruyama, N. Sato, N. Ohta, Conservation of structure and cold-regulation of RNA-binding proteins in cyanobacteria: Probable convergent evolution with eukaryotic glycine-rich RNA-binding proteins. *Nucleic Acids Res.* **27**, 2029–2036 (1999).

18. M. Sugita, M. Sugiura, The existence of eukaryotic ribonucleoprotein consensus sequence-type RNA-binding proteins in a prokaryote, *Synechococcus* 6301. *Nucleic Acids Res.* **22**, 25–31 (1994).
19. M. E. Mulligan, D. M. Jackman, S. T. Murphy, Heterocyst-forming filamentous cyanobacteria encode proteins that resemble eukaryotic RNA-binding proteins of the RNP family. *J. Mol. Biol.* **235**, 1162–1170 (1994).
20. N. Sato, A cold-regulated cyanobacterial gene cluster encodes RNA-binding protein and ribosomal protein S21. *Plant Mol. Biol.* **24**, 819–823 (1994).
21. M. E. Mulligan, T. J. Belbin, Characterization of RNA-binding protein genes in cyanobacteria. *Nucleic Acids Symp. Ser.* 140–2 (1995).
22. T. Kaneko, *et al.*, Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3**, 109–136 (1996).
23. T. Hamano, *et al.*, Characterization of RNA-binding properties of three types of RNA-binding proteins in *Anabaena* sp. PPC 7120. *Cell. Mol. Biol. Noisy--Gd. Fr.* **50**, 613–624 (2004).
24. Y. Zhang, D. Wu, Y. Wang, X. Xu, Two types of C-terminal regions of RNA-binding proteins play distinct roles in stress tolerance of Synechocystis sp. PCC 6803. *FEMS Microbiol. Lett.* **369**, fnac021 (2022).
25. N. Sato, A family of cold-regulated RNA-binding protein genes in the cyanobacterium *Anabaena variabilis* M3. *Nucleic Acids Res.* **23**, 2161–2167 (1995).
26. N. Sato, K. Maruyama, Differential regulation by low temperature of the gene for an RNA-binding protein, rbpA3, in the cyanobacterium *Anabaena variabilis* strain M3. *Plant Cell Physiol.* **38**, 81–86 (1997).
27. M. Mahbub, *et al.*, mRNA localization, reaction centre biogenesis and thylakoid membrane targeting in cyanobacteria. *Nat. Plants* **6**, 1179–1191 (2020).
28. X. Tan, *et al.*, Role of Rbp1 in the acquired chill-light tolerance of cyanobacteria. *J. Bacteriol.* **193**, 2675–2683 (2011).
29. Q. Tang, X. Tan, X. Xu, Effects of a type-II RNA-binding protein on fatty acid composition in *Synechocystis* sp. PCC 6803. *Chin. Sci. Bull.* **55**, 2416–2421 (2010).
30. Y. Wang, X. Xu, Effects of Rbp3 on lipid peroxidation and salt tolerance in *Synechocystis* sp. PCC 6803. *FEBS Lett.* **587**, 1446–1451 (2013).
31. T. S. Stenum, *et al.*, RNA interactome capture in *Escherichia coli* globally identifies RNA-binding proteins. *Nucleic Acids Res.* **51**, 4572–4587 (2023).
32. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
33. S. B. Pereira, *et al.*, The role of the tyrosine kinase Wzc (Sll0923) and the phosphatase Wzb (Slr0328) in the production of extracellular polymeric substances (EPS) by *Synechocystis* PCC 6803. *MicrobiologyOpen* **8**, e00753 (2019).
34. V. Capuano, A. S. Braux, N. Tandeau de Marsac, J. Houmard, The “anchor polypeptide” of cyanobacterial phycobilisomes. Molecular characterization of the *Synechococcus* sp. PCC 6301 apcE gene. *J. Biol. Chem.* **266**, 7239–7247 (1991).
35. L. Zheng, *et al.*, Cryo-EM and femtosecond spectroscopic studies provide mechanistic insight into the energy transfer in CpcL-phycobilisomes. *Nat. Commun.* **14**, 3961 (2023).

36. M. Riediger, M. A. Hernández-Prieto, K. Song, W. R. Hess, M. E. Futschik, Genome-wide identification and characterization of Fur-binding sites in the cyanobacteria Synechocystis sp. PCC 6803 and PCC 6714. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* **28**, dsab023 (2021).
37. S. Watanabe, *et al.*, Regulation of RNase E during the UV stress response in the cyanobacterium *Synechocystis* sp. PCC 6803. *mLife* **2**, 43–57 (2023).
38. D. H. Mathews, *et al.*, Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci.* **101**, 7287–7292 (2004).
39. L. Shi, K. M. Bischoff, P. J. Kennelly, The icfG gene cluster of *Synechocystis* sp. strain PCC 6803 encodes an Rsb/Spo-like protein kinase, protein phosphatase, and two phosphoproteins. *J. Bacteriol.* **181**, 4761–4767 (1999).
40. L. Beuf, S. Bédu, M. C. Durand, F. Joset, A protein involved in co-ordinated regulation of inorganic carbon and glucose metabolism in the facultative photoautotrophic cyanobacterium *Synechocystis* PCC6803. *Plant Mol. Biol.* **25**, 855–864 (1994).
41. A. Santana-Sanchez, *et al.*, Flavodiiron proteins 1-to-4 function in versatile combinations in O2 photoreduction in cyanobacteria. *eLife* **8**, e45766 (2019).
42. L. L. Carvalho, T. A. Salminen, K. M. Dahlström, Slr0006-like proteins: A TsaC/TsaC2/YciO subfamily exclusive to cyanobacteria. *Mol. Phylogenet. Evol.* **109**, 1–10 (2017).
43. R. Narikawa, *et al.*, Novel photosensory two-component system (PixA-NixB-NixC) involved in the regulation of positive and negative phototaxis of cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol.* **52**, 2214–2224 (2011).
44. M. Mutsuda, M. Sugiura, Translation initiation of cyanobacterial rbcS mRNAs requires the 38-kDa ribosomal protein S1 but not the Shine-Dalgarno sequence: development of a cyanobacterial in vitro translation system. *J. Biol. Chem.* **281**, 38314–38321 (2006).
45. S. Nakagawa, Y. Niimura, K. Miura, T. Gojobori, Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 6382–6387 (2010).
46. S. Nagarajan, D. M. Sherman, I. Shaw, L. A. Sherman, Functions of the duplicated hik31 operons in central metabolism and responses to light, dark, and carbon sources in *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* **194**, 448–459 (2012).
47. J. Giner-Lamia, L. López-Maury, J. C. Reyes, F. J. Florencio, The CopRS two-component system is responsible for resistance to copper in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol.* **159**, 1806–1818 (2012).
48. J.-Y. Song, *et al.*, Near-UV cyanobacteriochrome signaling system elicits negative phototaxis in the cyanobacterium *Synechocystis* sp. PCC 6803. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 10780–10785 (2011).
49. X.-Y. Liu, *et al.*, Structures of cyanobacterial bicarbonate transporter SbtA and its complex with PII-like SbtB. *Cell Discov.* **7**, 63 (2021).
50. M. Marbouty, C. Saguez, C. Cassier-Chauvat, F. Chauvat, Characterization of the FtsZ-interacting septal proteins SepF and Ftn6 in the spherical-celled cyanobacterium *Synechocystis* strain PCC 6803. *J. Bacteriol.* **191**, 6178–6185 (2009).
51. V. Krauspe, *et al.*, Discovery of a small protein factor involved in the coordinated degradation of phycobilisomes in cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2012277118 (2021).

52. K. Wang, M. Mahbub, G. Mastroianni, C. W. Mullineaux, mRNA localization and thylakoid protein biogenesis in the filamentous heterocyst-forming cyanobacterium Anabaena sp. PCC 7120. [Preprint] (2024). Available at: https://www.biorxiv.org/content/10.1101/2024.04.17.589878v1 [Accessed 2 May 2024].
53. J. Ding, *et al.*, Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev.* **13**, 1102–1115 (1999).
54. B. Soubise, Y. Jiang, N. Douet-Guilbert, M.-B. Troadec, RBM22, a key player of pre-mRNA splicing and gene expression regulation, is altered in cancer. *Cancers* **14**, 643 (2022).
55. C. Mayr, Regulation by 3′-untranslated regions. *Annu. Rev. Genet.* **51**, 171–194 (2017).
56. W. Ma, C. Mayr, A membraneless organelle associated with the endoplasmic reticulum enables 3'UTR-mediated protein-protein interactions. *Cell* **175**, 1492-1506.e19 (2018).
57. J. Béthune, R.-P. Jansen, M. Feldbrügge, K. Zarnack, Membrane-associated RNA-binding proteins orchestrate organelle-coupled translation. *Trends Cell Biol.* **29**, 178–188 (2019).
58. J. König, *et al.*, The fungal RNA-binding protein Rrm4 mediates long-distance transport of ubi1 and rho3 mRNAs. *EMBO J.* **28**, 1855–1866 (2009).
59. P. Becht, J. König, M. Feldbrügge, The RNA-binding protein Rrm4 is essential for polarity in *Ustilago maydis* and shuttles along microtubules. *J. Cell Sci.* **119**, 4964–4973 (2006).
60. S.-K. Devan, *et al.*, A MademoiseLLE domain binding platform links the key RNA transporter to endosomes. *PLOS Genet.* **18**, e1010269 (2022).
61. A. Katuwawala, B. Zhao, L. Kurgan, DisoLipPred: accurate prediction of disordered lipid-binding residues in protein sequences with deep recurrent networks and transfer learning. *Bioinforma. Oxf. Engl.* **38**, 115–124 (2021).
62. V. Anantharaman, D. Zhang, L. Aravind, OST-HTH: a novel predicted RNA-binding domain. *Biol. Direct* **5**, 13 (2010).
63. Y. Taniguchi, *et al.*, labA: a novel gene required for negative feedback regulation of the cyanobacterial circadian clock protein KaiC. *Genes Dev.* **21**, 60–70 (2007).
64. Y. Taniguchi, T. Nishikawa, T. Kondo, T. Oyama, Overexpression of lalA, a paralog of labA, is capable of affecting both circadian gene expression and cell growth in the cyanobacterium *Synechococcus elongatus* PCC 7942. *FEBS Lett.* **586**, 753–759 (2012).
65. B. M. McKnight, *et al.*, Roles for the *Synechococcus elongatus* RNA-binding protein Rbp2 in regulating the circadian clock. *J. Biol. Rhythms* **38**, 447–460 (2023).
66. D. Carmel, P. Mulo, N. Battchikova, E.-M. Aro, Membrane attachment of Slr0006 in *Synechocystis* sp. PCC 6803 is determined by divalent ions. *Photosynth. Res.* **108**, 241–245 (2011).
67. D. Carmel, *et al.*, Structural model, physiology and regulation of Slr0006 in *Synechocystis* PCC 6803. *Arch. Microbiol.* **195**, 727–736 (2013).
68. M. Mahbub, C. W. Mullineaux, Locations of membrane protein production in a cyanobacterium. *J. Bacteriol.* **205**, e0020923 (2023).
69. Y. Kanesaki, *et al.*, Identification of substrain-specific mutations by massively parallel whole-genome resequencing of *Synechocystis* sp. PCC 6803. *DNA Res.* **19**, 67–79 (2012).

70. D. Trautmann, B. Voß, A. Wilde, S. Al-Babili, W. R. Hess, Microevolution in cyanobacteria: Re-sequencing a motile substrain of *Synechocystis* sp. PCC 6803. *DNA Res.* **19**, 435–448 (2012).
71. C. P. Corcoran, *et al.*, Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol. Microbiol.* **84**, 428–445 (2012).
72. A. Kaltenbrunner, *et al.*, Regulation of pSYSA defense plasmid copy number in *Synechocystis* through RNase E and a highly transcribed asRNA. *Front. Microbiol.* **14**, 1112307 (2023).
73. U. K. Laemmli, SDS-page Laemmli method. *Nature* **227**, 680–685 (1970).
74. F. L. Pinto, A. Thapper, W. Sontheim, P. Lindblad, Analysis of current and alternative phenol based RNA extraction methodologies for cyanobacteria. *BMC Mol. Biol.* **10**, 79 (2009).
75. S. Hein, I. Scholz, B. Voß, W. R. Hess, Adaptation and modification of three CRISPR loci in two closely related cyanobacteria. *RNA Biol.* **10**, 852–864 (2013).
76. B. Voß, W. R. Hess, The identification of bacterial non-coding RNAs through complementary approaches. *Handb. RNA Biochem. Second Complet. Revis. Enlarg. Ed.* 787–800 (2014).
77. M. E. Ritchie, *et al.*, A comparison of background correction methods for two-colour microarrays. *Bioinforma. Oxf. Engl.* **23**, 2700–2707 (2007).
78. M. E. Ritchie, *et al.*, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
79. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
80. M. Choi, *et al.*, MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* **30**, 2524–2526 (2014).
81. Y. Perez-Riverol, *et al.*, The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).
82. J. Cox, *et al.*, Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics MCP* **13**, 2513–2526 (2014).
83. S. Tyanova, *et al.*, The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
84. L. Hemm, *et al.*, Interactors and effects of overexpressing YlxR/RpnM, a conserved RNA binding protein in cyanobacteria. [Preprint] (2024). Available at: https://www.biorxiv.org/content/10.1101/2024.01.06.574455v2 [Accessed 25 April 2024].
85. H. Li, R. Durbin, Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
86. M. R. Breese, Y. Liu, NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets. *Bioinforma. Oxf. Engl.* **29**, 494–496 (2013).
87. M. Kopf, *et al.*, Comparative analysis of the primary transcriptome of *Synechocystis* sp. PCC 6803. *DNA Res.* **21**, 527–539 (2014).
88. G. H. Putri, S. Anders, P. T. Pyl, J. E. Pimanda, F. Zanini, Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics* **38**, 2943–2945 (2022).

89. L. Wang, S. Wang, W. Li, RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012).
90. S. O. Skinner, L. A. Sepúlveda, H. Xu, I. Golding, Measuring mRNA copy number in individual *Escherichia coli* cells using single-molecule fluorescent *in situ* hybridization. *Nat. Protoc.* **8**, 1100–1113 (2013).
91. J. Schindelin, *et al.*, Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
92. W. Stauffer, H. Sheng, H. N. Lim, EzColocalization: An ImageJ plugin for visualizing and measuring colocalization in cells and organisms. *Sci. Rep.* **8**, 15764 (2018).
93. K. W. Dunn, M. M. Kamocka, J. H. McDonald, A practical guide to evaluating colocalization in biological microscopy. *Am. J. Physiol. - Cell Physiol.* **300**, C723–C742 (2011).
94. E. W. Deutsch, *et al.*, The ProteomeXchange consortium at 10 years: 2023 update. *Nucleic Acids Res.* **51**, D1539–D1548 (2023).
95. K. Darty, A. Denise, Y. Ponty, VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* **25**, 1974–1975 (2009).
96. J. Jumper, *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
97. M. Mirdita, *et al.*, ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
98. G. Erdős, Z. Dosztányi, AIUPred: combining energy estimation with deep learning for the enhanced prediction of protein disorder. *Nucleic Acids Res.* gkae385 (2024). https://doi.org/10.1093/nar/gkae385.

Figures

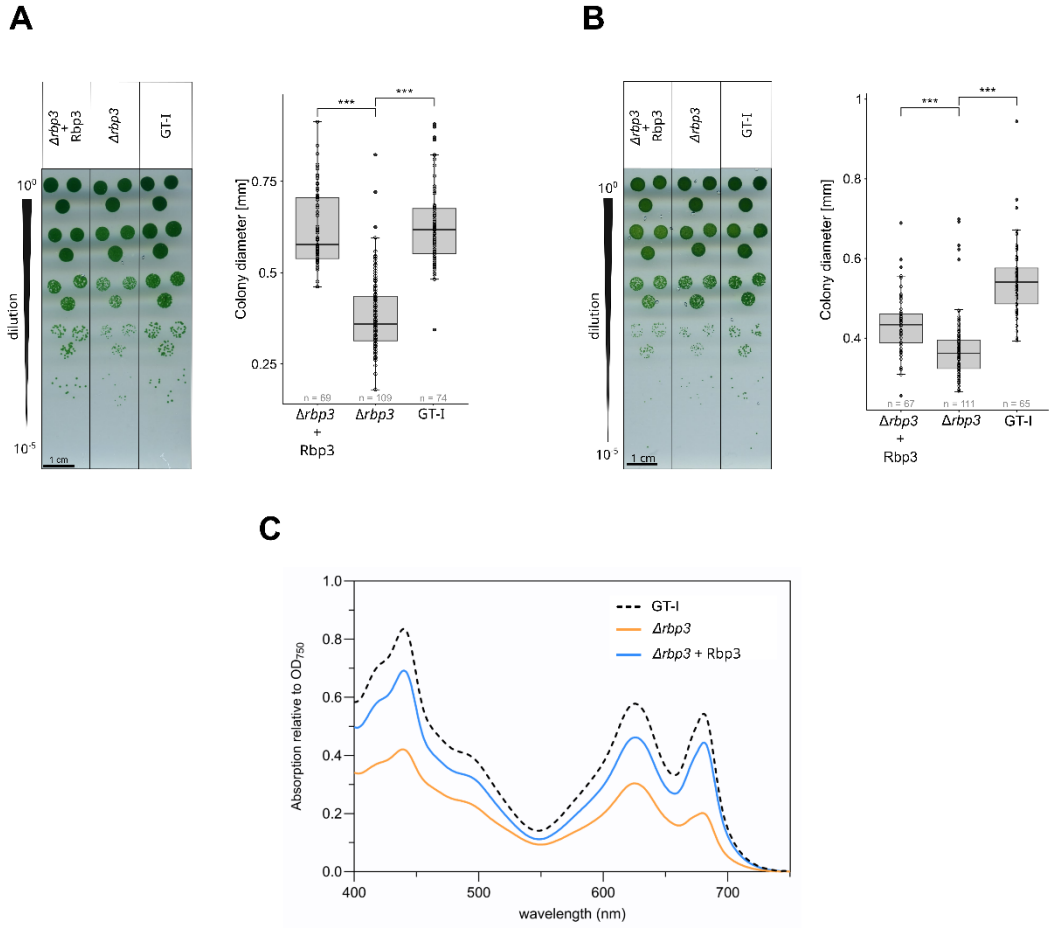

Fig. 1. Phenotypical analysis of Rbp3 mutants. (A) Spot assays of the indicated *Synechocystis* 6803 strains after 8 days of incubation under standard growth conditions (40 µmol photons m^{-2} s^{-1}, continuous light 30°C). **(B)** Spot assays under high light (300 µmol photons m^{-2} s^{-1}) and 5 days of incubation. In both conditions, the growth of wild type *Synechocystis* GT-I strain was compared to the Δ*rbp3* deletion mutant and the complementation strain Δ*rbp3*+Rbp3 that expressed *rbp3* from an introduced pVZ322 plasmid under control of its native promoter. Culture aliquots were inoculated in triplicates and diluted ten-fold by every step as indicated. Colony sizes were measured and evaluated (right panels). Significances were calculated by the Students t-test, for the numeric values see ***SI Appendix,* Table S7**. **(C)** Room-temperature absorption spectra of cell suspensions, normalized at 750 nm. The Δ*rbp3* mutant has lower pigment content compared to the wild type (WT) and the *rbp3* complementation mutant.

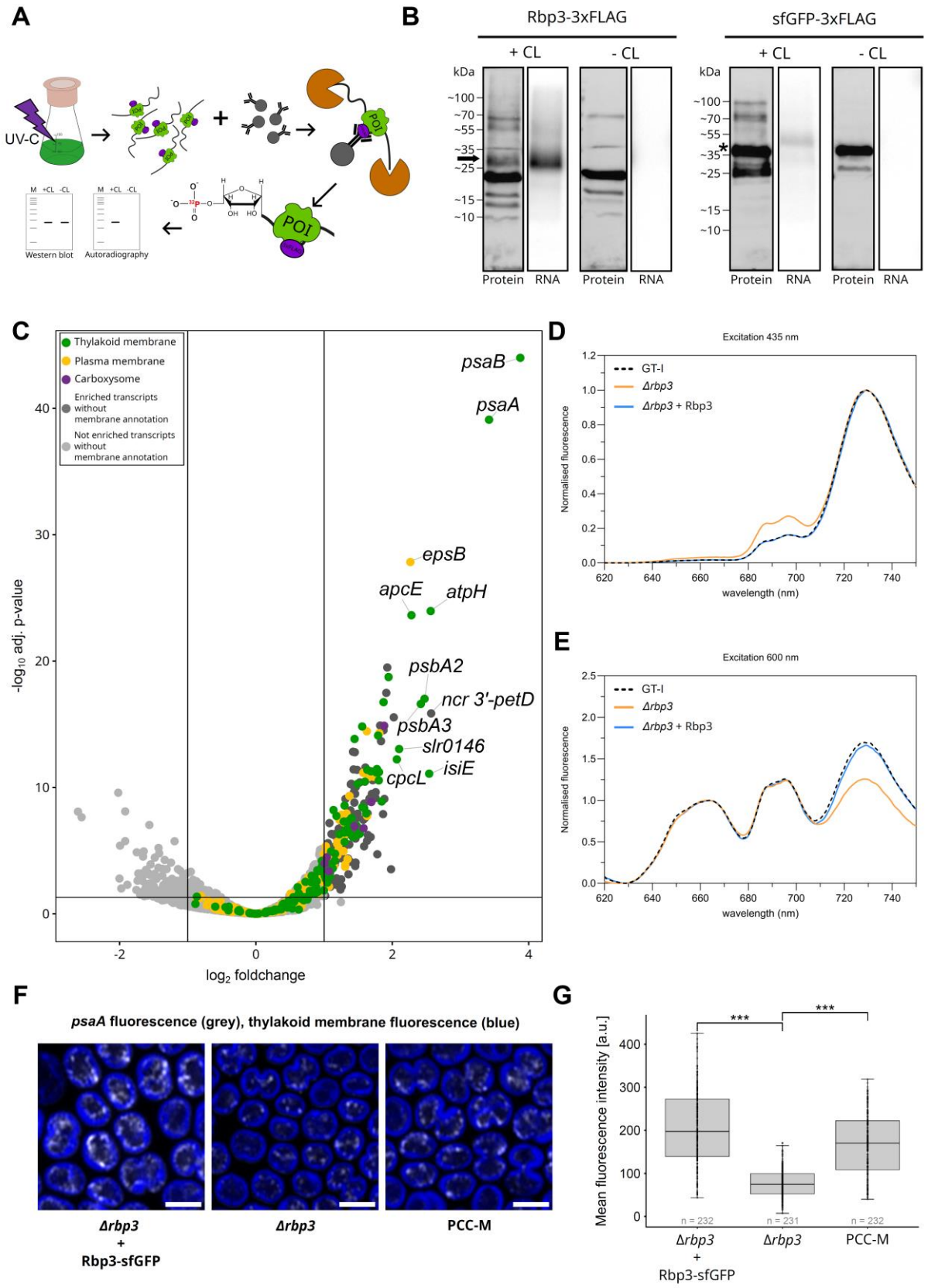

Fig. 2. Verification of RNA-binding of Rbp3. (A) Schematic overview of the PNK assay based on Stenum et al. (31). The protein of interest (POI) and RNA are crosslinked using UV-C light. After cell lysis, co-IP is performed using anti-FLAG M2

magnetic beads. Non-covalently bound nucleic acids are removed by benzonase treatment. The nuclease-protected, bound RNA is labeled with ^{32}P using PNK. Purified proteins are separated by SDS-PAGE (73), transferred to a nitrocellulose membrane and visualized by autoradiography for isotope-labeled RNA and by western blotting with anti-FLAG antisera for proteins. In controls, the crosslinking step is omitted. **(B)** Rbp3 and GFP (negative control) PNK assays. Protein expression was induced with 0.31 µM Cu_2SO_4, 24 h before crosslinking, and cultures were crosslinked using 2.25 J cm^{-2}. After separating complexes on a 15% SDS-PAA gel and western blotting, FLAG-tagged proteins were detected using an anti-FLAG M2 HRP coupled antibody (1:5,000) (Sigma). Several bands appeared in the protein lanes (left subpanels). The most prominent band, labeled by an arrow, indicates the position of Rbp3-3xFLAG at ~20 kDa, which belongs to the Rbp3-3xFLAG monomer in UV treated and untreated samples. For the sfGFP-3xFLAG the most prominent band in the UV treated and untreated sample runs at ~40 kDa, which refers to the monomeric sfGFP-3xFLAG. Bands beneath the signal of the respective monomer are degraded protein forms as these bands were increased in intensity after UV treatment. Bands above the respective monomers belong to potential multimers, stabilized by UV treatment since more of those bands appear in the UV treated samples. The band at ~40 kDa in the Rbp3-3xFLAG samples belongs to the sfGFP-3xFLAG. In the UV treated sample of Rbp3-3xFLAG two diffuse protein signals are present of which the one at ~30 kDa is overlapping with the autoradiography signal (left panel, + CL, RNA lane). The band at ~70 kDa might belong to a multimer of Rbp3. The band at ~35 kDa labeled with an asterisk probably belongs to sfGFP that spilled into the other lanes during loading. **(C)** Volcano plot showing transcripts enriched in the Rbp3_3×FLAG CLIP RNA samples in comparison to the sfGFP_3×FLAG CLIP RNA samples. Green dots indicating enriched transcripts that code for proteins involved in photosynthesis, yellow dots show enriched transcripts with the GO-term membrane, purple dots are transcripts that code for proteins associated with the carboxysome, dark grey dots show transcripts that are enriched but do not belong to any of the afore mentioned groups and light grey dots represent transcripts that were not enriched. The transcript for Rbp3 was removed from this analysis. Functionalities or gene names are given for transcripts enriched with a log_2 fold change ≥2. (**D**) Fluorescence emission spectra of *Synechocystis* 6803 PCC-M strains at 77K with chlorophyll excitation at 435 nm, normalized to PSI emission at 725 nm. The Δ*rbp3* mutant showed an increased PSII:PSI ratio compared to the other

two strains. **(E)** Fluorescence emission spectra of PCC-M strains at 77K with phycocyanin excitation at 600 nm, normalized to phycobilin emission at 660 nm. The spectra are representative of similar results from three independent cultures. **(F)** Confocal images showing the localization of *psaA* mRNA (TAMRA, grey) relative to the thylakoid membrane (chlorophyll, blue) in PCC-M (WT), PCC-M::*Δrbp3* and PCC-M::*Δrbp3* + Rbp3-sfGFP strains. **(G)** Quantification of *psaA* mRNA fluorescence intensity comparing PCC-M (WT), PCC-M::*Δrbp3* and PCC-M::*Δrbp3* + Rbp3-sfGFP strains. *n* = number of cells. *** significant difference ($p < 0.001$); scale bar = 2 µm. Significances were calculated by the students t-test assuming equal variances, for the numeric values see ***SI Appendix,* Table S8**.

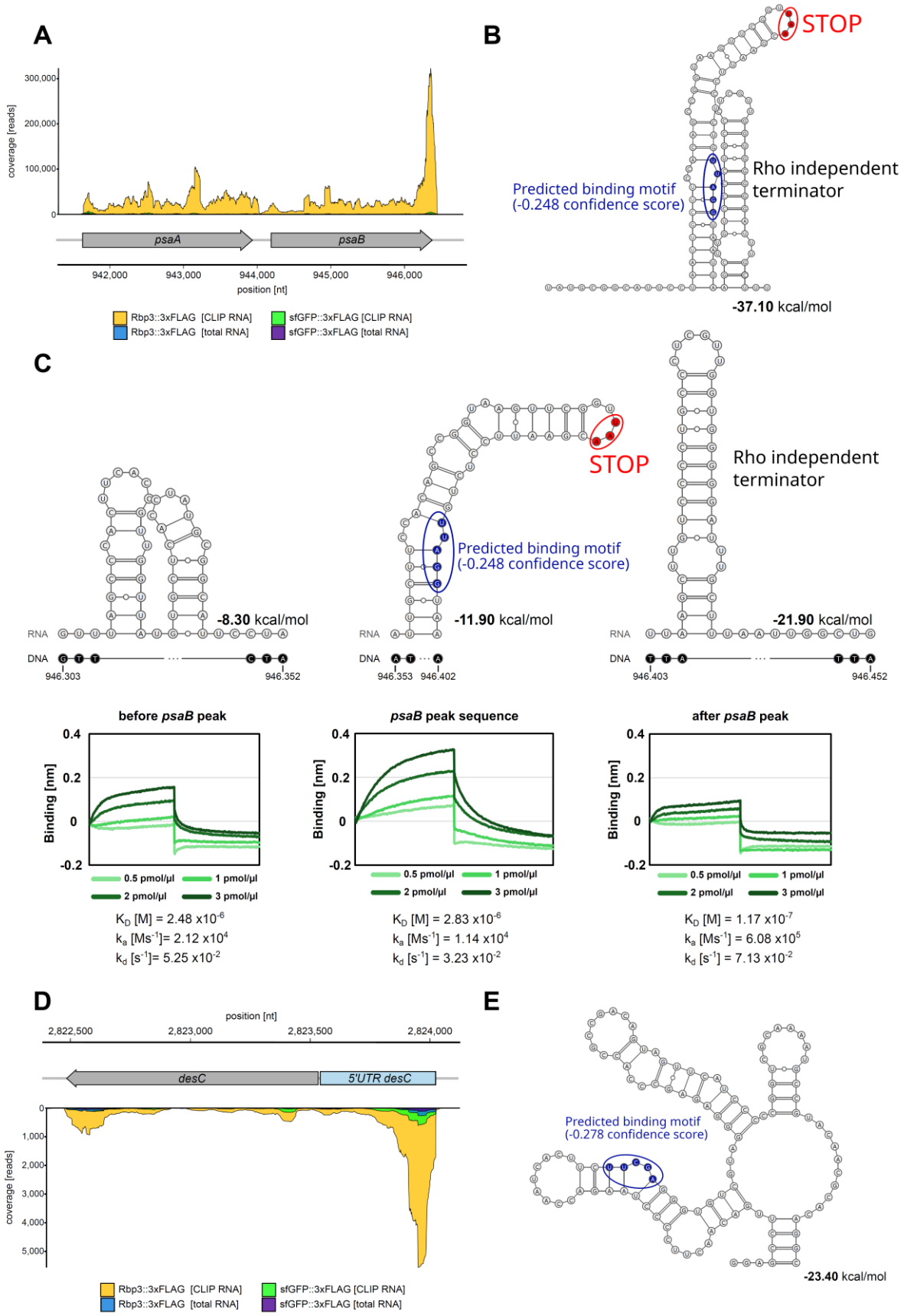

Fig. 3. Coverage of enriched mRNA reveals peaks in distinct areas of the transcripts. (A) Read coverage along the *psaAB* locus in the Rbp3-3xFLAG CLIP

RNA compared to the sfGFP-3xFLAG CLIP RNA. **(B)** Secondary structure prediction around the *psaB* read coverage peak. (**C**) Bio-Layer Interferometry of Rbp3 interactions with *psaB* coverage peak and neighboring sequences. Recombinant Rbp3 was used in an *in vitro* assay to determine binding affinities (lower panel) towards selected 50 nt RNA oligonucleotide sequences of the *psaB* coverage peak (upper panel, center), the 50 preceding nt of coding sequence (left) or the following 50 nt encompassing the Rho-independent terminator (right). (**D, E**) The *desC* (*sll0541*) gene encoding delta 9 acyl-lipid desaturase. In panels A and E the cDNA read coverages are plotted from four different types of samples. These samples were total RNA from cultures expressing the Rbp3-3xFLAG or sfGFP-3xFLAG and RNA obtained after crosslinking to Rbp3-3xFLAG or sfGFP-3xFLAG (negative control) and are color-coded as indicated. All cDNA libraries were sequenced in biological triplicates.

In panels B and C, the STOP codons are marked in red, and circled in blue the putative RRM binding sites predicted by the RRMScorer algorithm (11). The secondary structures were calculated by RNAfold (38) and displayed using VARNA (95). Further mRNA coverage plots are depicted in ***SI Appendix,* Fig. S3**.

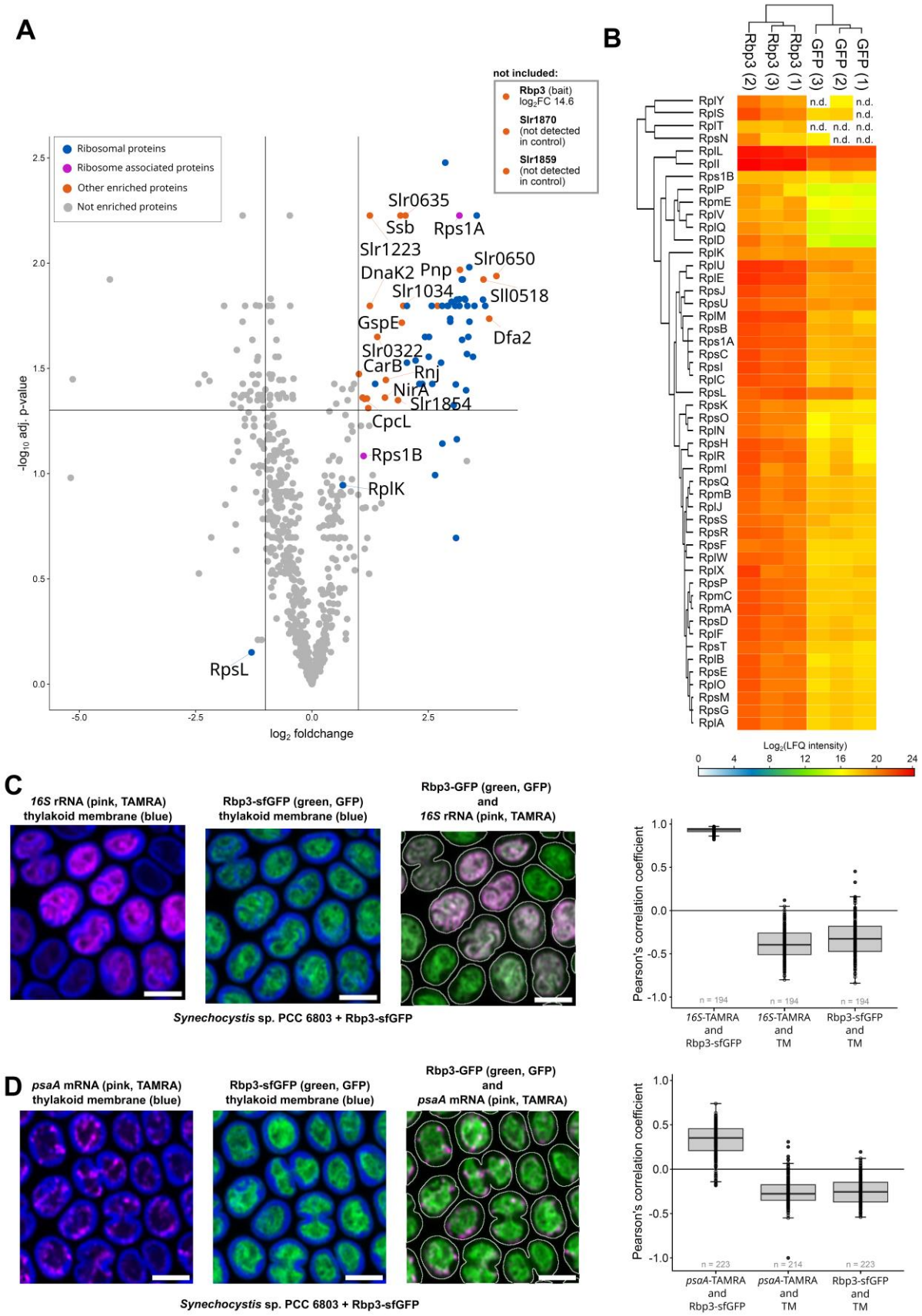

Fig. 4. **Protein interaction partner of Rbp3 (A)** Volcano plot showing the protein interaction partner of Rbp3-3xFLAG identified by mass spectrometry compared to sfGFP-3xFLAG. Further details of the enriched proteins can be found in ***SI Appendix,***

Table S2. For the complete list of detected proteins, see ***SI Appendix,*** **Table S9**. **(B)** Heatmap showing the $\log_2$ (LFQ intensity) of all ribosomal proteins that could be detected. Hierarchical clustering was done using the Perseus software (83). Rbp3-3xFLAG and sfGFP-3xFLAG expression was induced 24 h prior to harvesting with 0.31 µM Cu_2SO_4. For the analysis, the Rbp3-3xFLAG as well as the sfGFP-3xFLAG bait protein were not included in the plot. Details of statistical analysis in ***SI Appendix,*** **Table S10**. **(C)** Left panel: Localization of hybridized 16S rRNA (TAMRA, magenta) and Rbp3-sfGFP (sfGFP, green) relative to the thylakoid membrane (blue) or the cellular outline (white) in PCC-M::*Δrbp3* + Rbp3-sfGFP cells. Right panel: Colocalization quantification of fluorescent pair (left panel, third image) by calculation of the Pearson's correlation coefficient. *n* = number of cells. **(D)** Left panel: Localization of hybridized *psaA* mRNA (magenta) and Rbp3-sfGFP (sfGFP, green) relative to the thylakoid membrane (blue) or the cellular outline (white) in PCC-M::*Δrbp3* + Rbp3-sfGFP cells. Right panel: Colocalization quantification of fluorescent pair (left panel, third image) by calculation of the Pearson's correlation coefficient. *n* = number of cells. Scale bar = 2 µm in panels C and D. For the numeric values in the calculation of statistical significances in panels C and D, see ***SI Appendix,*** **Table S11**.

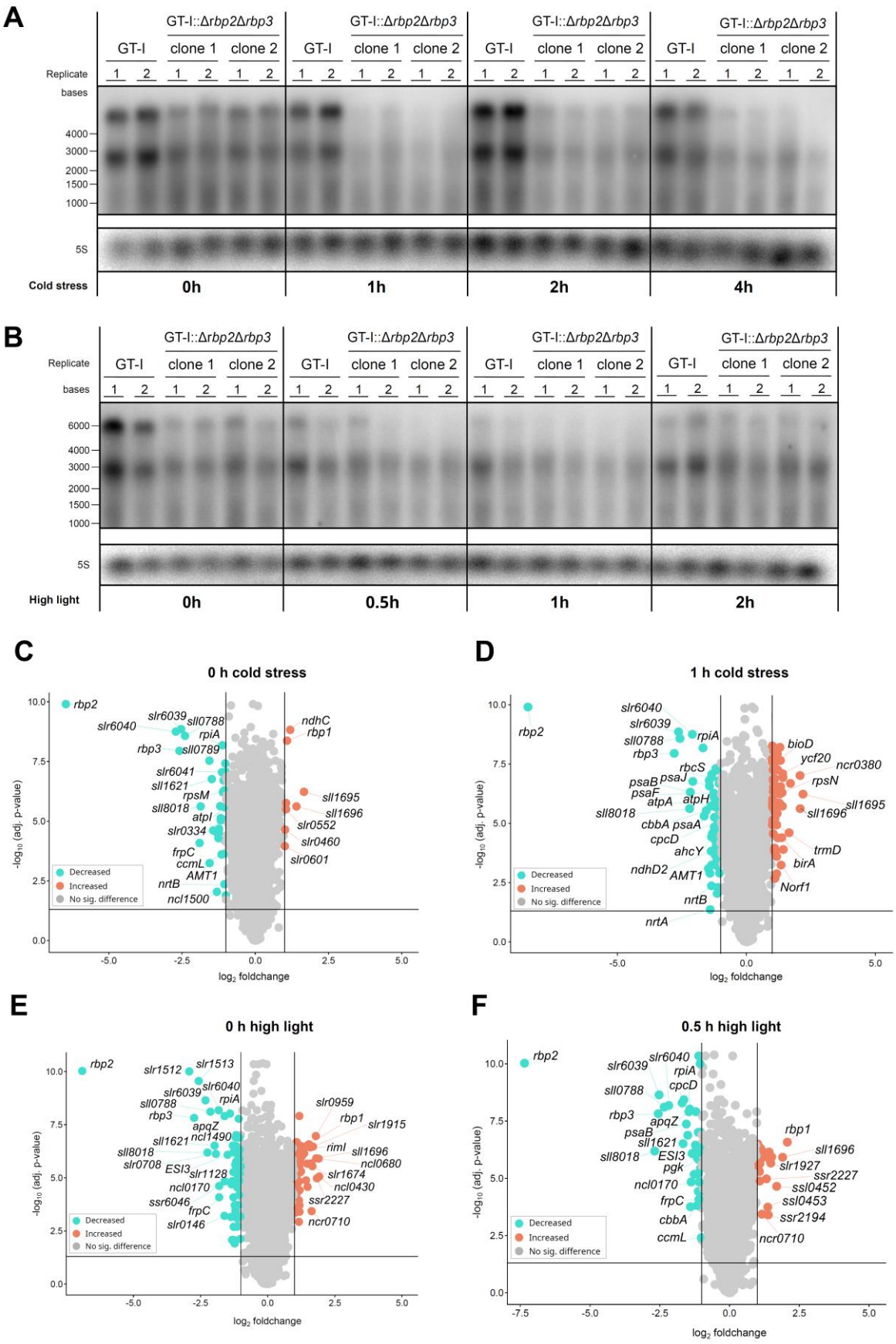

Fig. 5. **Transcriptomic analysis of Δ*rbp2*Δ*rbp3* double mutant. (A)** Time course of *psaAB* mRNA accumulation in *Synechocystis* 6803 wild type (WT) or the Δ*rbp2*Δ*rbp3*

mutant (C1, C2) after transferring the cultures from 30°C to 20°C for the indicated times. **(B)** Time course of *psaAB* mRNA accumulation after transferring the cultures from 50 to 250 µmol photons m^{-2} s^{-1}. In both panels A and B, 5 µg of total RNA were separated on 1.5% formaldehyde-agarose gels and the RiboRuler High Range RNA ladder (Thermo Fisher) served as molecular mass standard. The single-stranded RNA probe encompassed the region from 100 to 252 nt after the *psaA* start codon. Hybridization against the 5S rRNA served as loading control. (**C, D, E and F)** Volcano plots showing transcripts with significantly increased or decreased levels in the comparison of the double mutant against wild type, 0.5 h after shift from 40 to 250 µmol photons m^{-2} s^{-1}, or 1 h after shifting the cultures from 30°C to 20°C. The respective comparison at T = 0 h is also shown for both conditions. Transcriptomic changes were captured using a custom-made microarray. Salmon dots display increased transcript levels; turquoise dots display decreased transcript levels. Both experiments were performed with two replicates each. See **Tables S3** and **S4** and Datasets 1 and 2 for the complete set of transcriptomic differences.

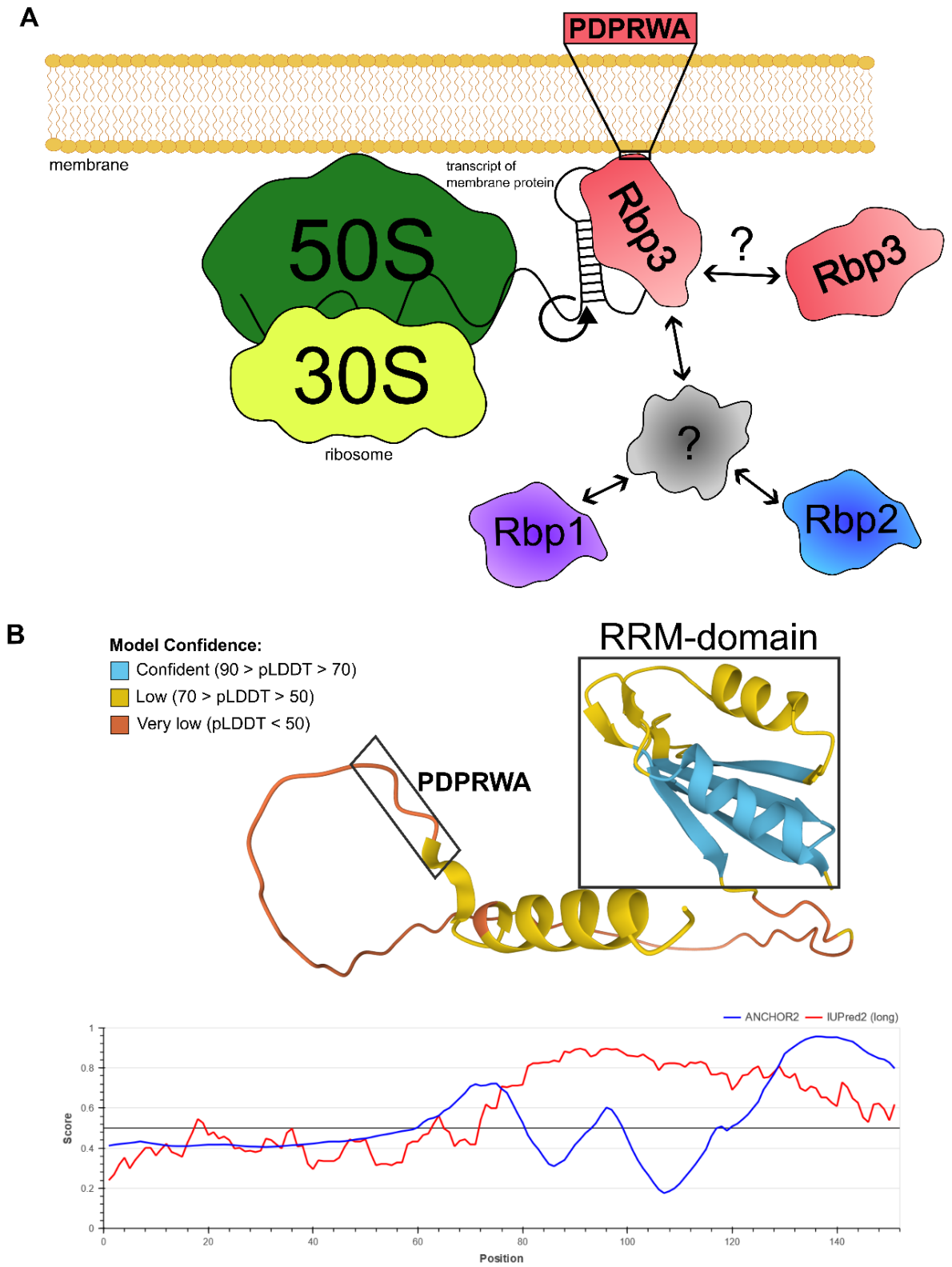

Fig. 6. Proposed model of transcript localization towards the membrane mediated by Rbp3. (A) Rbp3 binds distinct mRNA motifs such as a secondary structure overlapping the stop codon in case of the *psaAB* mRNA. The Rbp3 interaction with the membrane surface might involve a PDPRWA sequence motif conserved within

the C-terminal section of Rbp3 orthologs (24), identified by the DisoLipPred algorithm (61) as a lipid binding site. These mRNAs are translated near the membrane surface, explaining the high co-enrichment of ribosomal proteins with Rbp3. The helicase activity of the ribosome may unwind the RNA secondary structure, releasing Rbp3. Our data indicate that Rbp3 exists in both RNA-binding and non-binding states. The place where RNA binding is initiated and the factors involved in intracellular transport are unknown (question marks). The other RRM proteins, Rbp2 and Rbp1, likely also bind mRNAs with specific, potentially overlapping targets with Rbp3. Evidence suggests cross-talk among these Rbps, though its mechanisms are unclear. **(B)** Rbp3 structure predicted by Alphafold (96, 97). The RRM domain characterized by a four-stranded antiparallel beta-sheet and two alpha helices, is boxed. The structure ends with a C-terminal short beta fold followed by an alpha helix. Between the RRM domain and the C-terminal structures, a long intrinsically disordered region is predicted, at the end of which is the PDPRWA element. The reliability of the prediction is color-coded; the intrinsically disordered region is confirmed by the AIUPred algorithm (98).