Solving the Riddle of the Evolution of Shine-Dalgarno Based Translation in Chloroplasts

Iddo Weiner,[†,1,2] Noam Shahar,[†,2] Pini Marco,[2] Iftach Yacoby,*,[2] and Tamir Tuller*,[1,3]

[1]Department of Biomedical Engineering, The Iby and Aladar Fleischman Faculty of Engineering, Tel Aviv University, Tel Aviv, Israel
[2]School of Plant Sciences and Food Security, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv, Israel
[3]The Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

[†]These authors contributed equally to this work.
*Corresponding author: E-mail: iftachy@tauex.tau.ac.il; tamirtul@tauex.tau.ac.il.
Associate editor: Xuhua Xia

Abstract

Chloroplasts originated from an ancient cyanobacterium and still harbor a bacterial-like genome. However, the centrality of Shine–Dalgarno ribosome binding, which predominantly regulates proteobacterial translation initiation, is significantly decreased in chloroplasts. As plastid ribosomal RNA anti-Shine–Dalgarno elements are similar to their bacterial counterparts, these sites alone cannot explain this decline. By computational simulation we show that upstream point mutations modulate the local structure of ribosomal RNA in chloroplasts, creating significantly tighter structures around the anti-Shine–Dalgarno locus, which in-turn reduce the probability of ribosome binding. To validate our model, we expressed two reporter genes (mCherry, hydrogenase) harboring a Shine–Dalgarno motif in the *Chlamydomonas reinhardtii* chloroplast. Coexpressing them with a 16S ribosomal RNA, modified according to our model, significantly enhances mCherry and hydrogenase expression compared with coexpression with an endogenous 16S gene.

Key words: chloroplast evolution, Shine–Dalgarno, mRNA translation, *Chlamydomonas reinhardtii*.

Chloroplasts are the intracellular plant organelles in which photosynthesis takes place. The origin of these organelles is an early event of endosymbiosis in which an ancient cyanobacterium was engulfed by a eukaryotic host (Gray 1993; Reyes-Prieto et al. 2007). During the adaptation to this new symbiosis, the majority (~95%) of plastid genes have either been lost or horizontally transferred to the nucleus. However, modern chloroplasts maintain a small circular genome consisting of ribosomal RNAs (rRNAs), transfer RNAs, and open reading frames coding for ribosomal proteins, elongation factors, photosynthetic complexes and other housekeeping functions (Wakasugi et al. 2001). On average, roughly 100 coding sequences (CDSs) are translated within chloroplasts by a bacterial-like 70S ribosome and a matching set of gene expression machinery (Peled-zehavi and Danon 2007; Zoschke and Bock 2018).

Although translation initiation (TI) has been described as a major rate-limiting step for overall chloroplast gene expression (Eberhard et al. 2002; Marin-Navarro et al. 2007; Peled-zehavi and Danon 2007), its dynamics and similarity to bacterial TI remain unclear. The canonical model for explaining TI in prokaryotes is the Shine–Dalgarno (SD) mechanism in which the 30S ribosomal subunit is recruited to the mRNA through base-pairing between an mRNA motif (SD sequence) found upstream of the START site and an anti-SD (aSD)-conserved sequence found at the 3′-edge of the 16S rRNA (Shine and Dalgarno 1974, 1975). In this TI model, the proximity of the binding site to the START codon (approximately 10 bp upstream) as well as the complementarity between SD and aSD are key features necessary for proper initiation

(Jacob et al. 1987; Ma et al. 2002; Salis et al. 2009; Shaham and Tuller 2018) (fig. 1A).

Although SD-based TI is prevalent in a wide range of prokaryotes (e.g., proteobacteria, firmicutes, archaea, etc.), the extent of its role in chloroplasts has been a source of ongoing research; since the availability of plastid sequences first began rising, the search for differences between chloroplast and bacterial TI features has attracted attention (Marin-Navarro et al. 2007; Peled-zehavi and Danon 2007; Zoschke and Bock 2018). Although such differences often tend to be attributed to unique evolutionary adaptations taken on by chloroplasts postendosymbiosis, computational works performed on several organisms have shown that the SD motif presence on mRNAs is decreased both in chloroplasts (Scharff et al. 2011) and in cyanobacteria (Ma et al. 2002; Starmer et al. 2006; Nakagawa et al. 2010; Scharff et al. 2011; Voigt et al. 2014). These data suggest that the decrease in SD centrality occurred prior to endosymbiosis. To present this on large-scale, we calculated the 5′-UTR position-wise conservation for a representative from each sequenced chloroplast-containing genera, and compared the results with the equivalent cyanobacteria group of representatives (see supplementary data sets, Supplementary Material online, for database details). As an outgroup, we used a similarly built database of proteobacteria—which are a large and diverse group, including key model organisms (e.g. *Escherichia coli*, *Helicobacter pylori*, *Pseudomonas aeruginosa*, etc.) from which the SD TI model was originally deduced. From this large scale analysis, it is evident that proteobacteria show a clear SD signal centered roughly nine bases upstream from the START site (fig. 1B).
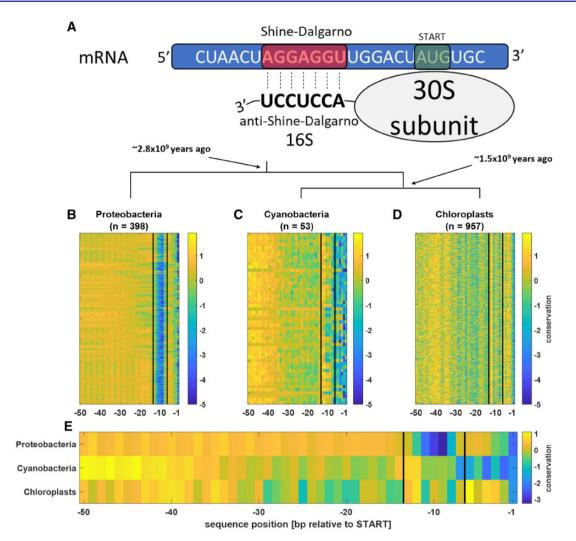
FIG. 1. There is no global SD motif conservation signal on chloroplast or cyanobacterial mRNAs. (*A*) Schematic illustration of the SD mechanism. A conserved element near the 3′-edge of the 16S rRNA spontaneously binds an mRNA motif located slightly upstream of the START site by base-pairing. This interaction properly positions the small ribosomal subunit on the start codon to initiate translation. (*B*–*D*) Sequence conservation (measured as Shannon Entropy Z-score, see supplementary methods, Supplementary Material online, for details and PSSM analysis) along the 5′-UTR. In each group the rows represent a single organism, selected to represent its genus. Lower values report on higher conservation in the specific position. The space between the vertical black lines represents the canonical location of the SD motif on the mRNA. The given number of rows (*n*) corresponds to the number of unique genera retrieved for each group. The dendrogram depicts the phylogenetic relations between the groups. Node dating is based on phylogenetic reconstruction evaluations (Reyes-Prieto et al. 2007; Flannery and Walter 2012; Schirrmeister et al. 2013). (*E*) The mean value for each position across all rows in (*B*)–(*D*).

On the other hand, although chloroplasts and cyanobacteria exhibit different conservation patterns (fig. 1C and D and supplementary fig. S1, Supplementary Material online), they both show no clear conservation signal in the SD area (fig. 1E). To validate that this observation is not an artifact caused by plastid annotation errors (Gallaher et al. 2018), we conducted a robustness test in which we iteratively omitted the annotated 5′-UTRs which contained the lowest SD scores. These omissions had negligible effects on the conservation maps (supplementary fig. S2, Supplementary Material online).

As positioning plays a key role in the SD mechanism, this global perspective suggests that the SD mechanism in the ancient cyanobacterium engulfed during the formation of chloroplasts has already diverged significantly from its canonical role in proteobacteria. Using a different approach, we applied the DAMBE (Xia 2013, 2017; Prabhakaran et al. 2015; Abolbaghaei et al. 2017) 5′-UTR analysis tool to identify putative SD locations in gene resolution (supplementary fig. S3, Supplementary Material online). These data show that the typical SD positioning prevalent in proteobacteria is eroded in both cyanobacteria and chloroplasts, thus supporting the same conclusion. These observations indicate that symbiosis with the nucleus (which applies to chloroplasts alone) could not be a main driver toward the reduced role of SD in chloroplasts. Alternatively, such drivers might include adaptation to oxygenic photosynthesis or slower growth rates, as these features differentiate proteobacteria from cyanobacteria and chloroplasts alike (Karlin et al. 2001; Ma et al. 2002).

Previous experimental approaches aiming to clarify the functional role of the SD mechanism in chloroplasts yielded a mixture of results. Although mutating a 5′-UTR SD motif is expected to reduce the gene's translation efficiency, this

outcome was only observed in several genes (Hirose et al. 1998; Nickelsen et al. 1999; Hirose and Sugiura 2004), whereas in others no significant effects were found (Sakamoto et al. 1994; Fargo et al. 1998; Nickelsen et al. 1999; Hirose and Sugiura 2004). In a recent report, researchers mutated the 16S aSD sequence and applied ribosomal profiling to quantify changes in translation rates across the tobacco chloroplast genome (Scharff et al. 2017). Interestingly, it was found that SD binding is required for translation of only a subset of plastid genes, thus giving a broader context to previous reports.

However, it is well established that unlike in model bacteria, translation constitutes a major regulatory phase in chloroplast gene expression (Kim et al. 1993; Eberhard et al. 2002; Zoschke and Bock 2018). This notion corresponds well with the decline in centrality of the spontaneous and energy-independent (Peled-zehavi and Danon 2007) SD TI mechanism (fig. 1, and references above), which is more suitable for short-lived mRNAs (Dierstein 1984; Brawerman 1987; Selinger et al. 2003) and regimes with strong transcriptional regulation. Thus, alterations in the plastid 16S rRNA 3′-edge that would add regulatory steps to the spontaneous SD mechanism and acclimate it to the chloroplast gene expression environment could be expected. As the 16S aSD sequence itself is highly conserved between chloroplasts and model bacteria (Marin-Navarro et al. 2007; Peled-zehavi and Danon 2007; Zoschke and Bock 2018), we hypothesized that structural alteration of the rRNA could be taking place instead.

To test this hypothesis, we examined the 3′-edges of all 16S rRNAs in our database. Although the aSD is clearly conserved in chloroplasts (and cyanobacteria), they contain a typical pattern of point mutations, clustered in two loci upstream of the aSD sequence (fig. 2A, positions 4–5 and 10–11). Although these base substitutions do not directly affect the aSD sequence, they could change the local rRNA secondary structure, and thus indirectly complicate the SD:aSD binding event. To examine this theory, we simulated the local folding of the 16S rRNA 3′-edge in all organisms in our database. In our simulations, we used the ViennaRNA package (Lorenz et al. 2011), which is based on empirical parameters (Mathews et al. 1999, 2004) and was shown to be highly accurate for predicting local RNA structures (Zhao et al. 2018). Yet, the structure of long complex RNA polymers is challenging to predict (Gutell 2015; Bevilacqua et al. 2016). Thus, to define the relevant region for simulation, we examined the 30S ribosomal subunit PDB structures from *E. coli* (Svidritskiy et al. 2018), and the *Spinacia oleracea* chloroplast (Boerema et al. 2018). We chose the region from nucleotide -28 relative to the aSD sequence until the 3′-edge of the 16S rRNA, because unlike upstream regions which bind ribosomal proteins and distant rRNA parts, the RNA only interacts with itself in this area (fig. 2B and supplementary fig. S4, Supplementary Material online). Interestingly, our folding simulations of this region show that the proteobacteria rRNA structures differ significantly from those of chloroplasts and cyanobacteria in several aspects (supplementary fig. S5, Supplementary Material online). To broadly quantify these effects across our database, for each organism we computed

the rRNA folding structure that minimizes free energy and counted the number of aSD nucleotides found in base-pairing interactions in each of these structures. Our simulation results clearly show that a significantly larger portion of aSD nucleotides is paired in chloroplasts and cyanobacteria, compared with proteobacteria (fig. 2C). As opposed to the one deterministic 16S end point provided by genome annotation, mature 16S rRNA molecules were shown to have a variety of end points (Wei et al. 2017; Silke et al. 2018). To test the effect of this phenomenon on our observations, we randomly added subsequent genomic DNA bases to the annotated 16S sequences and repeated the analysis described above; we found that introducing such diversity into the database has negligible effect on the above conclusion (supplementary fig. S6, Supplementary Material online).

As ribosome binding in the SD mechanism requires base-pairing between the mRNA SD motif and the rRNA aSD sequence, the occupation of the latter by other bonds reduces its affinity to the mRNA and lowers the probability of spontaneous binding; thus explaining the less canonical role that SD plays in chloroplast TI. According to our simulation results, examining the most common structures received from each group (fig. 2D) reveals the loss of the well-described bacterial hairpin tetraloop (Woese et al. 1980; Noller and Woese 1981; Cannone et al. 2002) in cyanobacteria and chloroplasts. However, by using a Boltzman equilibrium function (Ding and Lawrence 2003) to compute the probability of receiving suboptimal structures (i.e., RNA structures with higher free energy), we were able to observe that a tetraloop-containing structure occurs in these groups in lower but nonnegligible probability (supplementary fig. S7, Supplementary Material online)—suggesting that structural elasticity and transition between transient states may occur in this region.

The main drivers of this conformational change (fig. 2D) are four point mutations clustered in two loci (fig. 2A, positions 4–5, 10–11). Other point mutations and the slightly longer proteobacterial tails (fig. 2A) have negligible effects on this observation (fig. 2D and supplementary fig. S6, Supplementary Material online). Interestingly, we also observed a significant positive correlation between the openness of the aSD element (i.e., the number of exposed aSD nucleotides) and the conservation of the mRNA SD motif within the proteobacterial group (supplementary fig. S8, Supplementary Material online). This observation indicates that the coevolutionary dependence between these two traits occurs in proteobacteria as well.

According to this model, simply changing these two clusters of altered nucleotides into their proteobacterial form will modify the local folding at the 3′-edge of the 16S rRNA and expose the aSD sequence to facilitate easier binding to SD motifs. To test this theory, we designed a set of plasmids for chloroplast expression in the model green alga *Chlamydomonas reinhardtii*; plasmids A and B both code for a codon-optimized reporter gene (we used two different reporters: hydrogenase and mCherry) with a clear SD motif properly situated upstream from the START codon. Plasmid A also codes the natural *C. reinhardtii* 16S rRNA gene, whereas
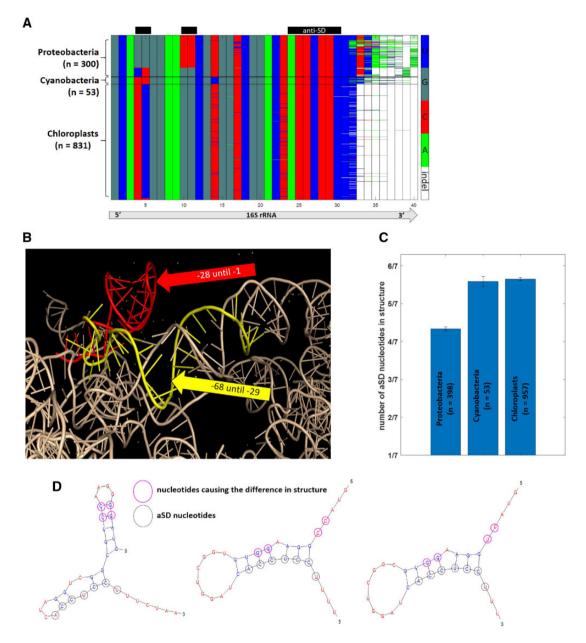
FIG. 2. Local folding at the 16S 3′-rRNA edge affects the aSD structure. (*A*) Multiple sequence alignment of the 16S rRNA 3′-edge database, composed of a single representative species from each genus. The black boxes above the matrix highlight important regions: The aSD box represents the conserved region where this sequence is found, whereas empty boxes represent columns with high variability between the different groups. Several outlying species which introduce indels in the body of the alignment were excluded for visualization purposes. The full alignment can be seen in supplementary figure S9, Supplementary Material online. The *X*-axis coordinates are given according to the full alignment, and not according to individual sequences (e.g., the majority of sequences are shorter than 40 and contain several indels at their 3′-edge). (*B*) Examination of the 3′-edge of the 16S rRNA in the 30S ribosomal subunit from the chloroplast of *Spinacia oleracea*. The reference point is the aSD sequence (e.g., −1 means one nucleotide upstream from the aSD sequence). The 3′-edge of the molecule (red) only interacts with itself, whereas the region immediately upstream (yellow) interacts with distant parts of the rRNA. (*C*) After simulating the 16S 3′-edge secondary structure of each organism, the number of aSD nucleotides found in base-pairing interactions was computed. The bars show the mean ± SE of this value across all organisms in each of the groups. (*D*) Typical (most abundant) rRNA secondary structures, with their corresponding consensus sequences. aSD nucleotides are marked in dark green, whereas the nucleotides that cause the structure discrepancy between groups are circled in pink.

plasmid B codes for a slightly modified version of this 16S gene in which the 3′-edge was mutated to match the canonical proteobacterial sequence (fig. 3A). We transformed each of these plasmids into the *C. reinhardtii* chloroplast, isolated a strain from each transformation group, validated that its chloroplast was properly engineered, and drove it to homoplasmicity (supplementary fig. S10, see supplementary methods, Supplementary Material online). Subsequently, we confirmed the presence of the reporter gene's protein product by immunoblotting (supplementary fig. S11, Supplementary Material online). To quantify the differences in translation efficiency between the two groups of
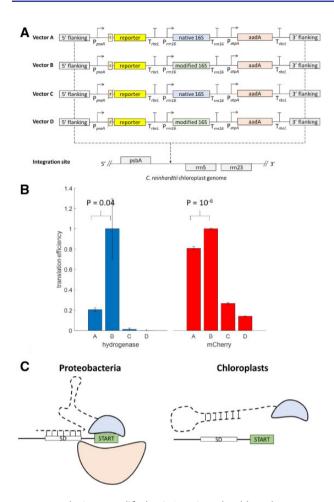
FIG. 3. Introducing a modified 16S rRNA into the chloroplast genome enhances the expression of a reporter gene with an SD motif. (A) Schematic diagram of the chloroplast expression vectors. Vectors A and B were completely identical, except for point mutations at the 3′-edge of the 16S gene in vector B, which altered its secondary structure. Vectors C and D are identical to A and B with the only difference being the deletion of the SD motif. All plasmids were integrated into an intergenic region between the psbA and rrn5 genes in the *Chlamydomonas reinhardtii* plastome. (B) Translation efficiency (protein abundance/mRNA abundance) for each reporter gene set. Data shown is mean $\pm$ SE. P-values were computed using a standard t-test. The additional P-values are: hydrogenase: A–C: 10^{-3}, A–D: 10^{-3}, B–C: 0.04, B–D: 0.04, C–D: 0.2; mCherry: A–C: 10^{-5}, A–D: 10^{-6}, B–C: 10^{-6}, B–D: 10^{-7}, C–D: 10^{-4}. (C) Schematic illustration of the model representing the altered role of SD in chloroplast TI; although bacterial 16S rRNAs are open and ready to base-pair with the mRNA, mutations in the chloroplast 16S 3′-edges tighten the structure of the aSD sequence and hamper spontaneous interactions with the mRNA.

transformants, we first measured overall expression for each reporter set; mCherry fluorescence was measured by flow cytometry, whereas hydrogenase activity was measured using the MV-enzyme quantification assay (Weiner, Atar, et al. 2018, see supplementary methods, Supplementary Material online). In both cases, we observed that our reporter gene was significantly more highly expressed in the clones transformed with plasmid B (supplementary fig. S12A and B, Supplementary Material online). In the next step, we used droplet digital PCR (see supplementary methods,

Supplementary Material online) to perform absolute quantification of transcript abundance in all clones (supplementary fig. S12C and D, Supplementary Material online) and computed the translation efficiency by dividing protein abundance to transcript abundance; these results clearly show that translation efficiency is significantly higher in group B (fig. 3B).

Subsequently, we created two additional plasmids (C and D) for each reporter gene which were identical to plasmids A and B, with the only difference being the deletion of the SD motif (achieved by mutating A↔T and G↔C) upstream from the reporter CDSs. These vectors, in which a key component of the SD mechanism has been knocked out, were built to serve as a negative control which further proves that our reporter gene is controlled by the SD mechanism (fig. 3A). They were put through the same process as plasmids A and B, and the overall reporter gene expression was measured as described above (supplementary figs. S10–12, Supplementary Material online). We observed that translation efficiency in these clones was significantly reduced by the deletion of the SD motif (fig. 3B), as expected. It is important to note that the mCherry clones C and D showed some leakiness due to an internal synthetic SD element near the 5′-edge of the CDS, inflicted by its codon composition (supplementary fig. S13, Supplementary Material online).

As homologous recombination between the untouched endogenous 16S and the inserted synthetic 16S could potentially erase our modifications, we confirmed the presence of our inserted 16S in all clones following the gene expression analyses (supplementary fig. S14, Supplementary Material online). Alongside our computational analysis, these discrepancies in protein expression indicate that loose aSD structures are important for spontaneous aSD:SD base-pairing (fig. 3C).

Recently, it has been shown that SD-mediated ribosome binding still occurs in the tobacco chloroplast (Scharff et al. 2017), thus confirming that all machinery required for proper SD-mediated ribosome binding remained intact in plastids and still plays a role in controlling chloroplast TI (Scharff et al. 2017). This explains the necessity of the aSD element, and together with our model it might also explain the conservation of the upstream mutations (fig. 2A, positions 4–5, 10–11); according to this theory, these mutations serve as part of an additional regulatory element which lowers the spontaneity of the aSD:SD interaction and acclimates the SD mechanism to the chloroplast environment in which regulation at the translational level is predominant. However, unlike in proteobacteria where SD is the canonical TI model, Scharff et al. have shown that it is a dominant TI mechanism for only a specific subset of genes. Importantly, this work shows that genes in which SD is essential for TI also tend to have strong mRNA secondary structures in the vicinity of their START sites, an observation originally used to form the start codon accessibility hypothesis (Nakamoto 2006; Scharff et al. 2011, 2017). Together with our data, these observations could also suggest that mRNAs controlled by SD might require a certain 5′-UTR secondary structure to unfold the 16S edge and reveal the aSD sequence in order to facilitate

ribosome binding (supplementary fig. S15, Supplementary Material online).

A useful byproduct of this work is the establishment of a generalist method for enhancing heterologous expression in chloroplasts. Our baseline plasmid in this work, pLM21, is known to drive high chloroplast expression in *C. reinhardtii* (Tibiletti et al. 2016; Sawyer et al. 2017; Weiner, Shahar, et al. 2018); yet by adding a slightly modified copy of the 16S rRNA (fig. 3A), we were able to roughly double the amount of protein achieved (supplementary fig. S12A and B, Supplementary Material online). As our simulations show that the aSD sequence is concealed in nearly all chloroplasts (fig. 2), we expect that adding a modified 16S rRNA to any plastid expression vector would similarly enhance the translation of a target transcript harboring an SD motif in any chloroplast.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

Author Contributions

I.W., N.S., I.Y., and T.T. designed the research; I.W. and N.S. performed the experimental and computational procedures; P.M. analyzed the PDB ribosome structures; I.W., N.S., I.Y., and T.T. wrote the paper. All authors declare no conflict of interest. No conflicts, informed consent, human or animal rights are applicable

References

Abolbaghaei A, Silke JR, Xia X. 2017. How changes in anti-SD sequences would affect SD sequences in *Escherichia coli* and *Bacillus subtilis*. *G3* 7:1607–1615.

Bevilacqua PC, Laura ER, Zhao S, Assmann SM. 2016. Genome-wide analysis of RNA secondary structure. *Annu Rev Genet*. 50:235–266.

Boerema AP, Aibara S, Paul B, Tobiasson V, Kimanius D, Forsberg BO, Wallden K, Lindahl E, Amunts A. 2018. Structure of the chloroplast ribosome with chl-RRF and hibernation-promoting factor. *Nat Plants*. 4:212.

Brawerman G. 1987. Determinants of messenger RNA stability minireview. *Cell* 48(1):5.

Cannone JJ, Subramanian S, Schnare MN, Collett JR, Souza LMD, Du Y, Feng B, Lin N, Madabusi LV, Müller KM. 2002. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* 3(1):2.

Dierstein R. 1984. Synthesis of pigment-binding protein in toluene-treated *Rhodopseudomonas capsulata* and in cell-free systems. *Eur J Biochem*. 138(3):509–518.

Ding Y, Lawrence CE. 2003. A statistical sampling algorithm for RNA secondary structure prediction. *Nucleic Acids Res*. 31(24):7280–7301.

Eberhard S, Drapier D, Wollman F-A. 2002. Searching limiting steps in the expression of chloroplast- encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. *Plant J*. 31(2):149–160.

Fargo DC, Zhang M, Gillham NW, Boynton JE. 1998. Shine-Dalgarno-like sequences are not required for translation of chloroplast mRNAs in *Chlamydomonas reinhardtii* chloroplasts or in *Escherichia coli*. *Mol Gen Genet*. 257(3):271–282.

Flannery DT, Walter MR. 2012. Archean tufted microbial mats and the Great Oxidation Event: new insights into an ancient problem. *Aust J Earth Sci*. 59(1):1–11.

Gallaher SD, Fitz-Gibbon ST, Strenkert D, Purvine SO, Pellegrini M, Merchant SS. 2018. High-throughput sequencing of the chloroplast and mitochondrion of *Chlamydomonas reinhardtii* to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates. *Plant J*. 93(3):545–565.

Gray MW. 1993. Origin and evolution of organelle genomes. *Curr Opin Genet Dev*. 3(6):884–890.

Gutell RR. 2015. rRNA—the evolution of that magic molecule. *RNA* 21(4):627–629.

Hirose T, Kusumegi T, Sugiura M. 1998. Translation of tobacco chloroplast rps14 mRNA depends on a Shine-Dalgarno-like sequence in the 5 P -untranslated region but not on internal RNA editing in the coding region. *FEBS Lett*. 430(3):257–260.

Hirose T, Sugiura M. 2004. Functional Shine-Dalgarno-like sequences for translational initiation of chloroplast mRNAs. *Plant Cell Physiol*. 45(1):114–117.

Jacob WF, Santer M, Dahlberg AE. 1987. A single base change in the Shine-Dalgarno region of 16S rRNA of *Escherichia coli* affects translation of many proteins. *Proc Natl Acad Sci U S A*. 84(14):4757–4761.

Karlin S, Mrázek J, Campbell A, Kaiser D. 2001. Characterizations of highly expressed genes of four fast-growing bacteria. *J Bacteriol*. 183(17):5025–5040.

Kim M, Christopher DA, Mullet JE. 1993. Direct evidence for selective modulation of psbA, rpoA, rbcL and 16S RNA stability during barley chloroplast development. *Plant Mol Biol*. 22(3):447–463.

Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. *Algorithms Mol Biol*. 6(1):26.

Ma J, Campbell A, Karlin S. 2002. Correlations between Shine-Dalgarno sequences and gene features and operon structures. *J Bacteriol*. 184(20):5733–5745.

Marin-Navarro J, Manuell A, Wu J, Mayfield SP. 2007. Chloroplast translation regulation. *Photosynth Res*. 94:359–374.

Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH. 2004. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci U S A*. 101(19):7287–7292.

Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol*. 288(5):911–940.

Nakagawa S, Niimura Y, Miura K, Gojobori T. 2010. Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proc Natl Acad Sci U S A*. 107(14):6382–6387.

Nakamoto T. 2006. A unified view of the initiation of protein synthesis. *Biochem Biophys Res Commun*. 341(3):675–678.

Nickelsen J, Fleischmann M, Boudreau E, Rahire M, Rochaix J. 1999. Identification of cis-acting RNA leader elements required for chloroplast psbD gene expression in *Chlamydomonas*. *Plant Cell* 11(5):957–970.

Noller HF, Woese CR. 1981. Secondary structure of 16S ribosomal RNA. *Science* 212(4493):403–411.

Peled-zehavi H, Danon A. 2007. Translation and translational regulation in chloroplasts. *Top Curr Genet*. 19:249–281.

Prabhakaran R, Chithambaram S, Xia X. 2015. *Escherichia coli* and Staphylococcus phages: effect of translation initiation efficiency on differential codon adaptation mediated by virulent and temperate lifestyles. *J Gen Virol*. 96(5):1169.

Reyes-Prieto A, Weber APM, Bhattacharya D. 2007. The origin and establishment of the plastid in algae and plants. *Annu Rev Genet*. 41:147–168.

Sakamoto W, Xuemei C, Karen K, David S. 1994. Function of the *Chlamydomonas reinhardtii* petD 5′ untranslated region in regulating the accumulation of subunit IV of the cytochrome b6/f complex. *Plant J*. 6(4):503–512.

Salis HM, Mirsky EA, Voigt CA. 2009. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol*. 27(10):946.

Sawyer A, Bai Y, Lu Y, Hemschemeier A, Happe T. 2017. Compartmentalisation of [FeFe]-hydrogenase maturation in *Chlamydomonas reinhardtii*. *Plant J*. 90(6):1134–1143.

Scharff LB, Childs L, Walther D, Bock R. 2011. Local absence of secondary structure permits translation of mRNAs that lack ribosome-binding sites. *PLoS Genet*. 7(6):e1002155.

Scharff LB, Ehrnthaler M, Janowski M, Childs LH, Hasse C, Gremmels J, Ruf S, Zoschke R, Bock R. 2017. Shine-Dalgarno sequences play an essential role in the translation of plastid mRNAs in tobacco. *Plant Cell* 29(12):3085–3101.

Schirrmeister BE, Vos JM, De Antonelli A, Bagheri HC. 2013. Evolution of multicellularity coincided with increased diversification of cyanobacteria and the Great Oxidation Event. *Proc Natl Acad Sci U S A*. 110(5):1791–1796.

Selinger DW, Saxena RM, Cheung KJ, Church GM, Rosenow C. 2003. Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res*. 13(2):216–223.

Shaham G, Tuller T. 2018. Genome scale analysis of *Escherichia coli* with a comprehensive prokaryotic sequence-based biophysical model of translation initiation and elongation. *DNA Res*. 25(2):195–205.

Shine J, Dalgarno L. 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci U S A*. 71(4):1342–1346.

Shine J, Dalgarno L. 1975. Determinant of cistron specificity in bacterial ribosomes. *Nature* 254(5495):34.

Silke JR, Wei Y, Xia X. 2018. RNA-Seq-based analysis reveals heterogeneity in mature 16S rRNA 3′ termini and extended anti-Shine-Dalgarno motifs in bacterial species. *G3* 8(12):3973–3979.

Starmer J, Stomp A, Vouk M, Bitzer D. 2006. Predicting Shine-Dalgarno sequence locations exposes genome annotation errors. *PLoS Comput Biol*. 2(5):e57.

Svidritskiy E, Demo G, Korostelev AA. 2018. Mechanism of premature translation termination on a sense codon. *J Biol Chem*. 293(32):12472–12479.

Tibiletti T, Auroy P, Peltier G, Caffarri S. 2016. *Chlamydomonas reinhardtii* PsbS protein is functional and accumulates rapidly and transiently under high light. *Plant Physiol*. 171(4):2717–2730.

Voigt K, Sharma CM, Mitschke J, Joke Lambrecht S, Voß B, Hess WR, Steglich C. 2014. Comparative transcriptomics of two environmentally relevant cyanobacteria reveals unexpected transcriptome diversity. *ISME J*. 8(10):2056.

Wakasugi T, Tsudzuki T, Sugiura M. 2001. The genomics of land plant chloroplasts: gene content and alteration of genomic information by RNA editing. *Photosynth Res*. 70(1):107–118.

Wei Y, Silke JR, Xia X. 2017. Elucidating the 16S rRNA 3′ boundaries and defining optimal SD/aSD pairing in *Escherichia coli* and *Bacillus subtilis* using RNA-Seq data. *Sci Rep*. 7(1):17639.

Weiner I, Atar S, Schweitzer S, Eilenberg H, Feldman Y, Avitan M, Blau M, Danon A, Tuller T, Yacoby I. 2018. Enhancing heterologous expression in *Chlamydomonas reinhardtii* by transcript sequence optimization. *Plant J*. 94(1):22–31.

Weiner I, Shahar N, Feldman Y, Landman S, Milrad Y, Ben-Zvi O, Avitan M, Dafni E, Schweitzer S, Eilenberg H, et al. 2018. Overcoming the expression barrier of the ferredoxin-hydrogenase chimera in *Chlamydomonas reinhardtii* supports a linear increment in photosynthetic hydrogen output. *Algal Res*. 33:310–315.

Woese CR, Magrum LJ, Gupta R, Siegel RB, Stahl DA, Kop J, Crawford N, Brosius J, Gutell R, Hogan JJ, et al. 1980. Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence. *Nucleic Acids Res*. 10:2275–2294.

Xia X. 2013. DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Mol Biol Evol*. 30(7):1720–1728.

Xia X. 2017. DAMBE6: new tools for microbial genomics, phylogenetics, and molecular evolution. *J Hered*. 108(4):431–437.

Zhao Y, Wang J, Zeng C, Xiao Y. 2018. Evaluation of RNA secondary structure prediction for both base-pairing and topology. *Biophys Rep*. 4(3):123–132.

Zoschke R, Bock R. 2018. Chloroplast translation: structural and functional organization, operational control and regulation. *Plant Cell* 30(4):745–770.