Effect of the Secondary Structure in the *Euglena gracilis* Chloroplast Ribulose-bisphosphate Carboxylase/Oxygenase Messenger RNA on Translational Initiation*

(Received for publication, September 22, 1993, and in revised form, November 19, 1993)

Ja Seok Koo and Linda L. Spremulli†
From the Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290

The results reported in the previous paper indicate that the translational start site of the *Euglena gracilis* chloroplast mRNA for the large subunit of ribulose-bisphosphate carboxylase/oxygenase (*rbcL*) is not defined by primary sequence elements (Koo, J. S., and Spremulli, L. L. (1994) *J. Biol. Chem.* 269, 7494–7500). In the work presented here, the effects of secondary structure in the 5′-untranslated leader of the *rbcL* mRNA have been examined. Only weak secondary structure can be detected in the 5′-untranslated leader of the *rbcL* message by enzymatic and computer analysis. Further reduction of the weak secondary structure of this message by site-directed mutagenesis does not significantly affect the ability of this message to participate in initiation complex formation. The secondary structure near the translational start site was increased by the introduction of an inverted repeat sequence and by site-directed mutagenesis. Messages with increased secondary structure are much less active in initiation complex formation if the structural element introduced is within ~10 nucleotides of the start codon. These results suggest that the translational start site in this chloroplast mRNA is specified by the presence of an AUG codon in an unstructured or weakly structured region of the mRNA. No specific sequences around the start codon, either upstream or immediately downstream, were found to have important information directing the chloroplast ribosome to the start site of this mRNA.

In the accompanying paper (48), the role of the length and the primary sequence of the 5′-untranslated leader in the mRNA encoding the large subunit of ribulose-bisphosphate carboxylase/oxygenase (*rbcL*) in *Euglena gracilis* in the activity of the mRNA in initiation was studied. The results obtained indicate that the full 55-nucleotide leader is required for maximal efficiency in initiation complex formation. This message does not have a Shine-Dalgarno sequence 5′ to the start codon, and no clear primary sequence elements appear to be present that specify a particular AUG codon as the start site. The previous results lead to the suggestion that weakly structured regions of the mRNA having an AUG codon allow 30 S ribosomal subunits to have access to the start region and subsequently lead to the formation of an initiation complex at this position.

The secondary structure of a mRNA is believed to be one of the most important elements regulating the efficiency of initiation in prokaryotes (1–5). Messenger RNAs having relatively unstructured translational initiation regions appear to be expressed efficiently, while those having stable secondary structures in or around the ribosome-binding site are generally less efficient (1–3, 6–8). In the prokaryotic system, a number of factors play a role in determining the specificity and efficiency of a translational start site (9–13). These include the start codon, the Shine-Dalgarno sequence (14), and the exact sequence of the ribosome-binding site (15–19). Although primary sequence elements appear to be the major determinants for specifying initiation in prokaryotes, these elements must be accessible to the ribosome (6, 7) in order to function efficiently. Hence, secondary structure in the mRNA is also crucial. For example, the destruction of stem-loop structures containing the initiation codon or the Shine-Dalgarno sequence has been shown to improve the efficiency of translation of mRNAs (4, 21–23). A careful analysis of translational efficiency as a function of secondary structure has been carried out with the coat protein cistron of bacteriophage MS2 as a model system (4). The results of this analysis clearly show an inverse correlation between translational efficiency and the stability of the secondary structure of the mRNA in the initiation region.

The chloroplast translational system resembles the prokaryotic system in several general ways (24); however, almost half of the messages in chloroplasts do not have a Shine-Dalgarno sequence in the ribosome-binding site. In this report, the effects of RNA structure in the initiation region on the efficiency of initiation complex formation with one of these mRNAs has been examined.

**EXPERIMENTAL PROCEDURES**

**Materials**—General chemicals and 5-bromouridine 5′-triphosphate were purchased from Sigma. RNase T1, RNase V1, calf intestinal phosphatase, and avian myeloblastosis virus reverse transcriptase were obtained from United States Biochemical Corp. RNase T2 was purchased from Life Technologies, Inc. Escherichia coli tRNA was from Boehringer Mannheim. ATP, dNTP, and ddNTP were purchased from Pharmacia LKB Biotechnology Inc. (32)P1ATP (3000 Ci/mmol) was obtained from Dupont NEN. Construction of Plasmids and Preparation of Corresponding mRNAs—In previous work, the complete 5′-untranslated leader and exon 1 of the *E. gracilis rbcL* gene were fused in frame with a portion of the neomycin resistance gene, providing a construct designated pRbcN (25). Wild-type plasmid pRbcN was prepared as described previously (25). The plasmid template pRbcN X51R (see Fig. 1) was prepared by inserting a 48-base pair oligonucleotide in the inverse orientation at the XbaI site of the vector pRbcN X5 using standard methods. The starting plasmid for this construction has a unique XbaI recognition sequence between positions ~55 and ~50, but is otherwise identical to pRbcN (26). Other mutants were created by oligonucleotide-directed mutagenesis basically as described by Kunkel (27) and McClary et al. (28). Oligonucleotides for mutagenic reactions (Table I) were prepared in the Department of Pathology, Microbiology, and Immunology and at the Lineberger Cancer Center of the University of North Carolina at Chapel Hill.

Messenger RNAs were synthesized from corresponding plasmids (i.e.,

---

*This work was supported in part by National Institutes of Health Grant GM24963. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.
† To whom correspondence should be addressed.

---

*This is an Open Access article under the CC BY license.
UTP with 5-BrUTP during the in vitro transcription reaction. The incubated for 10 min at room temperature before use.

The 3'-terminal region of mRbcN from plasmid pRbcN was prepared by the substitution of mG (30) in place of uridine was prepared by the substitution of mG (30) in place of uridine was prepared by the substitution of 

Digestion with RNase V1 (0.06 unit) was carried out by heating the mixture (10 pl) to 90 °C for 3 min and followed by slow cooling to room temperature over a 30-min period. Alternatively, the heated mixture was placed directly on ice. Following this step, 8 pl of buffer (250 mM Tris-HCl, pH 8.3, 40 mM MgCl₂, 250 mM NaCl, and 5 mM dithiothreitol) and 1 pl of 5 mM dNTP were added. This reaction was divided into four aliquots of 4.5 μl each, and 4 μl of the dNTP for chain termination were added to the appropriate aliquot, giving final concentrations of 0.21 mM ddATP, 0.17 mM ddTTP, 0.11 mM ddGTP, and 0.15 mM ddCTP. The RNA sequencing reaction was started by the addition of 2 units of avian myeloblastosis virus reverse transcriptase in 1 μl and was incubated for 30 min at 42 °C. The reactions were analyzed on 6% polyacrylamide gels containing 7 M urea (33).

Computer Analysis of Secondary Structures of mRNA-Computer predictions of mRNA secondary structures were done using the MFold program (version 7.0) of the University of Wisconsin Genetics Computer Group software running on a VAX computer system. This program uses the algorithm of Zuker (34) for structural predictions. The data obtained from the primer extension reactions were used as parameters to impose restrictions on the folding patterns when needed.

For mRNA sequencing, 1 pmol of mRbcN and 1 pmol of [32P]-labeled primer were annealed in water (10 μl) by heating for 3 min at 90 °C, followed by slow cooling to room temperature over a 30-min period. Alternatively, the heated mixture was placed directly on ice. Following this step, 8 pl of buffer (250 mM Tris-HCl, pH 8.3, 40 mM MgCl₂, 250 mM NaCl, and 5 mM dithiothreitol) and 1 pl of 5 mM dNTP were added. This mixture was divided into four aliquots of 4.5 μl each, and 4 μl of the dNTP for chain termination were added to the appropriate aliquot, giving final concentrations of 0.21 mM ddATP, 0.17 mM ddTTP, 0.11 mM ddGTP, and 0.15 mM ddCTP. The RNA sequencing reaction was started by the addition of 2 units of avian myeloblastosis virus reverse transcriptase in 1 μl and was incubated for 30 min at 42 °C. The reactions were analyzed on 6% polyacrylamide gels containing 7 M urea (33).

Secondary Structure Analysis of Initiation Region of mRbcN—The 5'-untranslated leader of mRbcN is ~90% A and U residues (Fig. 1). Such an A/U-rich region would be expected to have little, or at most only weak, secondary structure. To examine this possibility, RNA structural analysis was carried out on mRbcN using biochemical probing and computer predictions. Primer extension analysis was performed on intact mRNA (Fig. 2A, Ct lane) and on RNA that had been digested with RNases T1, T2, and V1. Each RNase digestion was optimiz
mRNA translation. (A) The predicted secondary structures are shown for the translational initiation region of mRbcN. The first and second structures are derived from an analysis of the primer extension data, while the third structure is predicted by the MFold program (34).

Role of Reduced Secondary Structures in 5'-Untranslated
Leader of \textit{mRbcN}—The possible role, if any, of the stem-loop formations present in the RNA at equilibrium. This mutated mRNA has changes both in the primary sequence adjacent to the start initiation codon and in the potential secondary structure near the start site. It should therefore also be useful for the further clarification of the importance of the primary sequences upstream of the AUG codon.

To confirm that this mRNA had the predicted secondary structure near the translational initiation region, it was subjected to primer extension analysis. This analysis (data not shown) suggested that the 5'-untranslated leader was largely single-stranded and that no unexpected structure had been introduced (Fig. 3A). The activity of mRbcN As in initiation complex formation was tested as a function of the concentration of the mRNA and was compared to the activity of the wild-type mRNA. As shown in Fig. 3B, this mRNA has essentially the same activity as mRbcN. These results have two implications. First, a defined primary sequence between positions -1 and -20 is not required for the activity of mRbcN in initiation. This result is in agreement with the data presented in the previous paper (48), which argue against a major role for a defined primary sequence in specifying the correct translational start site for this chloroplast mRNA. Second, the data argue that the putative weak stem-loop structure just upstream of the start codon does not play a positive role in initiation and that the 5'-untranslated leader of the wild-type mRNA is already sufficiently unstructured to allow maximal efficiency in initiation.

The second construct prepared in this series replaced the stem-loop structure near the 5'-end of the mRbcN with a series of U residues (mRbcN Us; Figs. 1 and 3A). A similar uridine-rich sequence (U5AU4) is located just upstream of the start codon in mRbcN. Uridine-rich sequences have been suggested to enhance translational efficiency with \textit{E. coli} mRNAs (16), and ribosomal protein S1 appears to have a high affinity for uridine-rich regions (17). It should be noted that the stem-loop structure being replaced here (\(\Delta G = -3.1\) kcal/mol) is present in at least a population of mRbcN conformers that can be detected by primer extension analysis (Fig. 2). This stem-loop structure may be partially stabilized by its GUAA tetraloop. RNA loops with the sequence GNRA (where N = A, C, G, or U and R = A or G) appear to be significantly more stable than other loop sequences (41–43).

Primer extension analysis of mRbcN Us was carried out and indicated that the stem-loop predicted to be present at the 5'-end of mRbcN has indeed been eliminated (data not shown). The activity of mRbcN Us in initiation complex formation was determined as a function of the mRNA concentration and was compared to the activity of the wild-type mRNA (Fig. 3B). The mRNA was \(-70\%\) as efficient as mRbcN, suggesting that it does not differ significantly from the wild-type RNA in its ability to direct the initiation of translation. This result indicates that the weak hairpin structure near the 5'-end of mRbcN is not necessary for the activity of this mRNA in translation. In addition, the data indicate that the specific primary sequence from positions -36 to -50 is not critical for ribosome recognition. This observation is in agreement with the results obtained with several of the mutants described in the previous paper (48).

Effect of Enhancing Stability of Secondary Structure of \textit{mRbcN} on Its Activity in Initiation Complex Formation—The results outlined above and in the accompanying paper (48) lead to the suggestion that the 5'-untranslated leader of mRbcN does not provide any unique sequence information that is directly responsible for specifying a specific AUG as the start codon. Rather, it appears that the low degree of secondary structure in this region of the mRNA might lead to the selection of a particular AUG codon for initiation.

An initial investigation into the role of secondary structure in the initiation of mRbcN was carried out by substituting 5-bromouridine for uridine in the RNA. This change was easily accomplished by replacing UTP with 5-BrUTP during the \textit{in vitro} synthesis of mRbcN (producing mRbcN UBr). RNAs having 5-bromouridine in place of uridine have stronger secondary
Weak Secondary Structure Controls Chloroplast Translation

structures due to an increase in base stacking interactions (30). When mRbcN UBr was tested for the ability to participate in 30S initiation complex formation, it was only ~20–30% as active as mRbcN in this assay (data not shown). This result suggests that enhanced secondary structure in the message may reduce ribosome access to the start codon and inhibit initiation.

Effect of Inverted Repeat Sequences Creating Secondary Structures in 5′-Untranslated Leader on Initiation—A significant increase in the secondary structure of the 5′-untranslated leader region of mRbcN was accomplished by creating a message that had a duplication of 48 bases in the leader region present as a 48-base pair inverted repeat (mRbcN X5IR; Fig. 1). Primer extension analysis was carried out to determine the secondary structures of this mRNA (Fig. 4A). In this analysis, the mRNA was mapped for intrinsic reverse transcriptase stops. Strong reverse transcriptase stop signals were observed throughout the region between positions -13 and -45. These stops reflect the stem-loop that is predicted to form due to the inverted repeat in this leader (Fig. 4B). The observation that there are numerous stops along the 3′-edge of this stem suggests that reverse transcriptase can partially melt this structure or that this stem is partially opened due to thermal energy. The calculated total free energy value ($\Delta G$) for the stem shown here suggests that it is a strong stem. However, the bottom of the stem is largely composed of A:U base pairs, which may be partially opened at the temperature used for the initiation complex assays (37°C).

The mRNA containing the inverted repeat structure was tested for the ability to participate in initiation complex formation (Fig. 5) and had ~20% of the activity observed with the wild-type mRNA. The significant reduction in activity observed with this mRNA is most likely due to the presence of the secondary structure created by the inverted repeat present. This structure would interfere with the binding of 30 S subunits to the start site on the mRNA.

Inhibition of Translational Initiation by Strong Stem-Loop Structures in 5′-Untranslated Leader—The experiments described above have consistently indicated that the information for the selection of the start site on mRbcN by the chloroplast 30 S subunit resides in the unstructured single-stranded nature of the leader region of this mRNA. If this idea is correct, it is logical to predict that the introduction of stronger secondary structures would inhibit initiation complex formation. To test this prediction, three additional mutants were constructed that had strong stem-loop structures at different positions within the leader region (Fig. 1). The first mutant mRNA has a strong stem-loop structure composed of G and C residues immediately upstream of the AUG start codon (mRbcN GC2). The second has the strong stem-loop located 10 residues 5′ to the start AUG codon (mRbcN GC10), while the third construct has the

![Fig. 4. Secondary structure of message containing inverted repeat in 5′-untranslated region.](image)

![Fig. 5. Effect of dyad symmetry sequences in 5′-untranslated leader of mRbcN on initiation.](image)
Weak Secondary Structure Controls Chloroplast Translation

The presence of these stem-loop structures was confirmed by primer extension analysis (Fig. 6A), and the predicted secondary structures of these mRNAs are shown in Fig. 6B. Mapping mRbcN GC2 gave strong reverse transcriptase stop signals at positions -1 to -7. These stop signals presumably arise from the strong stem introduced at this position. Cleavage of mRbcN GC2 with RNase T2 resulted in strong signals from positions -10 to -14. Since RNase T2 preferentially cuts residues in single-stranded loops, these results agree with the introduction of the predicted stem-loop structure. RNase V1 provided evidence for the 5'-side of the stem-loop by giving a signal from the C residue at position -17.

The hairpin structure in mRbcN GC2 is immediately upstream of the AUG codon and may actually encompass the start codon. As expected, this message is practically inactive in initiation complex formation (Fig. 7). Presumably, the strong secondary structure present prevents 30 S ribosomal subunits from gaining access to the start codon.

The second construct places the same strong stem-loop structure 10 nucleotides 5' to the AUG codon (mRbcN GC10). The secondary structure of this mRNA was probed by primer extension analysis (Fig. 6A), and the results of this analysis combined with the analysis by the MFold program are shown in Fig. 6B. Reverse transcription of mRbcN GC10 gave a strong stop signal at position -10. This stop appears to dominate the analysis of the mRNA regardless of cleavage with RNase. This stop signal presumably arises from the base of the G/C-rich stem-loop at position -10. The activity of mRbcN GC10 in initiation complex formation is significantly reduced (Fig. 7), and mRbcN GC10 has only ~20% of the activity seen with the wild-type mRbcN message. This reduced activity is presumably due to steric hindrance created by the stem-loop structure on the mRNA, which would interfere with the access of the 30 S subunit to the start codon. The stem-loop in this mRNA would encompass a significant portion of the normal ribosome-binding site, which usually spans positions -20 to +15.

Finally, a mutant of mRbcN was prepared in which the hairpin structure was moved between positions -33 and -52, which should lie outside of the ribosome-binding site (mRbcN GC33). Primer extension analysis of mRbcN GC33 (Fig. 6A) gave a strong stop signal for reverse transcription at position -33 presumably due to the stem-loop that had been introduced into the RNA. RNase T2 probing gave signals from residues -41 and -42, which are expected to be part of the loop region. Finally, RNase V1 gave a signal at position -50 on the 5'-side of the predicted stem. The ability of mRbcN GC33 to participate in initiation complex formation was examined (Fig. 7). Interestingly, the mRbcN GC33 mRNA has essentially the same activity as the wild-type mRbcN. This observation indicates that a
Weak Secondary Structure Controls Chloroplast Translation

The remainder of the mRNA can be weakly structured (wild-type RNA), unstructured (Fig. 3), or in a strong secondary structure (Fig. 7).

**Effects of Downstream Sequences on Initiation Complex Formation with mRbcN**—In prokaryotic systems, the presence of an additional sequence element located immediately downstream of the AUG codon has been suggested to enhance the efficiency of translational initiation (44, 45). This downstream sequence is believed to hydrogen-bond to the 16S rRNA present in the 30 S ribosomal subunit. Computer analysis of the nucleotide sequence of mRbcN just downstream of the start codon indicates the presence of possible base pairing of residues 459–465 of the 16S rRNA and nucleotides +7 to +13 (CCUC-CAA) of mRbcN. In addition, residues +9 to +15 (UCAAACU) of mRbcN can potentially hydrogen-bond to the sequence from positions 12 to 18 or from positions 971 to 977 of the small subunit ribosomal RNA. A mutant in which this hypothetical hydrogen bonding was disrupted (mRbcN DB; Fig. 1) was prepared. The activity of mRbcN DB in initiation complex formation is essentially the same as that observed with mRbcN (data not shown). This observation indicates that no specific primary sequence information resides within the ribosome-binding site downstream of the start codon.

**DISCUSSION**

The results presented here and in the previous paper (48) suggest that there is no direct primary sequence information used to specify the translational start site on the rbcL mRNA. The translational start site of this mRNA has the AUG start codon positioned in an unstructured or very weakly structured region of the message, making the AUG codon accessible to the 30 S subunit. Introduction of strong secondary structure elements close to the AUG codon reduces its activity in initiation complex formation significantly. These results argue that the only major determinant defining a particular AUG as the start codon in this message is its presence in an accessible region of the RNA. The rbcL mRNA is representative of one class of *E. gracilis* chloroplast mRNAs that lack Shine-Dalgarno sequences and that appear to have little structure in the translational start site. The results observed here can thus probably be extrapolated to the initiation regions of about half of the mRNAs in the chloroplasts of *E. gracilis*.

Numerous observations from prokaryotic systems also argue that secondary structure present in the translational initiation region blocks access to the start site and reduces the efficiency of the message in translation (3, 22, 46). In contrast, the effect of RNA secondary structure is more complex in the eukaryotic cytoplasmic system. In the yeast *Saccharomyces cerevisiae*, a stem-loop structure with a predicted stability greater than −28 kcal/mol inhibited in vivo translation significantly, and a leader with a stem having a *GΔC* of −14 kcal/mol showed only about one-third the normal activity (35). Creation of a hairpin structure (Δ*G* = −30 kcal/mol) involving the AUG codon in the mammalian preproinsulin coding sequence did not reduce the yield of preproinsulin; however, a more stable stem-loop structure (Δ*G* = −50 kcal/mol) reduced the yield significantly. Downstream stem-loop structures may actually improve the recognition of an AUG initiation codon in eukaryotes by reducing the speed with which the 40 S subunit scans the mRNA and by providing more time for recognition of the AUG codon.

The data presented here and in the previous paper (48) argue strongly that primary sequence information does not play a direct role in specifying the start site of mRbcN. Rather, the basic information required to direct the chloroplast 30 S ribosomal subunit to the start site is an AUG codon present in a highly unstructured region of the RNA. The 30 S subunit may recognize the single-stranded region of the mRNA by using a single-strand RNA-binding protein such as ribosomal protein S1 (15). Recently, a chloroplast S1-like protein has been identified and characterized from spinach chloroplasts (20, 47). In *E. gracilis* ribosomes, an S1-like protein has been detected by Western blotting using *E. coli* anti-S1 antibody (data not shown). Interestingly, the chloroplast S1-like protein from spinach preferentially binds polya, while the *E. coli* S1 protein binds most strongly to polyU, although it will also bind polya and polyC (15, 47). The 30 S subunit interacting with the selected single-stranded region of the mRNA via an S1-like protein might then locate the nearby AUG codon. In the presence of Met-tRNA and initiation factors, the initiation complex will form at this position, and the interaction between the mRNA and the 30 S subunit will be stabilized by the codon-anticodon interaction between the mRNA and the initiator tRNA bound to the small subunit.

**REFERENCES**

1. de Smit, M. H., and van Duijn, J. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 7668–7672
2. Munson, L. M., Stormo, G. D., Niece, B. L., and Reznikoff, W. S. (1984) *J. Mol. Biol.* 177, 683–693
3. Wang, C.-C., Roney, W. B., Alston, R. L., and Spremulli, L. L. (1989) *J. Biol. Chem.* 264, 10,459–10,461
4. Gross, G., Miecke, C., Hollatz, I., Bloeker, H., and Frank, R. (1990) *J. Biol. Chem.* 265, 17627–17636
5. Hahn, V., Dorne, A.-M., Mache, R., Ebel, S., and Stiegler, P. (1988) *Plant Mol. Biol.* 19, 799–814
6. Hartz, D., Mcpheeters, D., and Gold, L. (1991) *Proc. Natl. Acad. Sci. U. S. A.* 88, 5695–5699
7. Zhang, J., and Deutscher, M. P. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 4474–4478
37. Knapp, G. (1989) Methods Enzymol. 180, 193–212
38. Vary, C. P., and Vournakis, J. (1984) Nucleic Acids Res. 12, 6763–6778
39. Lowman, H. B., and Draper, D. (1986) J. Biol. Chem. 261, 5396–5403
40. Friere, M. M., Kierzek, R., Jaeger, J., Sugimoto, N., Caruthers, M., Neilson, T., and Turner, D. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 9379–9377
41. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1992) Science 256, 217–219
42. Woese, C. R., Winker, S., and Gutell, R. R. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 8467–8471
43. Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Gayle, M., Guild, N., Stormo, G., D’Aubenton-Carafa, Y., Uhlenbeck, O. C., Tinoco, I., Jr., Brody, E. N., and Gold, L. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 1364–1368
44. Sprengart, M. L., Fatcher, H. P., and Fuchs, E. (1990) Nucleic Acids Res. 18, 1719–1723
45. Loechel, S., Inamine, J. M., and Hu, P. C. (1991) Nucleic Acids Res. 19, 6905–6911
46. Altuvia, S., Kornitzer, D., Teff, D., and Oppenheim, A. B. (1989) J. Mol. Biol. 210, 265–280
47. Franzetti, B., Carol, P., and Mache, R. (1992) J. Biol. Chem. 267, 19075–19081
48. Koo, J. S., and Spremulli, L. L. (1994) J. Biol. Chem. 269, 7494–7500