THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 269, No. 10, Issue of March 11, pp. 7494–7500, 1994
Printed in U.S.A.

Analysis of the Translational Initiation Region on the *Euglena gracilis* Chloroplast Ribulose-bisphosphate Carboxylase/Oxygenase (*rbcL*) Messenger RNA*

(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 chloroplast mRNAs from *Euglena gracilis* fall into two classes. One group of mRNAs from this organelle contains a Shine-Dalgarno sequence 5′ to the start codon, while the other group of mRNAs does not have a conserved sequence signal in the 5′-untranslated region. To investigate the start signals for *E. gracilis* chloroplast mRNAs that do not carry a Shine-Dalgarno sequence, 30 S initiation complex formation has been studied using a series of transcripts carrying the wild-type translational start site of ribulose-bisphosphate carboxylase/oxygenase (*rbcL*) or mutated derivatives of this site. Mutation of the start codon of the *rbcL* gene indicates that the chloroplast 30 S subunit is recognizing only the correct AUG codon. The analysis of the messages from a series of deletion mutants shows that a minimum of Δ20 residues 5′ to the AUG codon is required for activity. Maximal activity requires the full 55-base leader sequence. Surprisingly, a transcript carrying the inverse complement of 48 bases in the leader is Δ60% as active as the wild-type message in promoting initiation complex formation. Introduction of a Shine-Dalgarno sequence in the 5′-leader increases the activity of the mRNA only Δ1.4–2-fold. The presence of an oligodeoxynucleotide containing a strong Shine-Dalgarno sequence does not significantly inhibit the formation of initiation complexes at the *rbcL* start site. Similar results are obtained when initiation complexes are formed with initiation factors from either *E. gracilis* chloroplasts or *Escherichia coli*.

During the past several years, progress has been made in understanding the mechanism of protein biosynthesis in chloroplasts. Chloroplast ribosomes are 70 S particles that consist of 30 S and 50 S subunits. Two chloroplast translational initiation factors have been purified from *Euglena gracilis*. Chloroplast initiation factor 2 (IF-2_{chl})[1] promotes the binding of fMet-tRNA to 30 S ribosomal subunits (1). This factor is structurally complex (1, 2) and does not have significant activity on bacterial ribosomes (3). Chloroplast initiation factor 3 (IF-3_{chl}) promotes the dissociation of ribosomes and facilitates initiation complex formation (4). IF-3_{chl} is active on *Escherichia coli* ribosomes, but has many physical properties that distinguish it from the corresponding prokaryotic factor (5).

Little detailed information is available on the precise nucleotide sequences in chloroplast mRNAs that specify the start site for translation. In prokaryotes, a polypurine sequence (such as GGAGG) is centered ~10 nucleotides upstream of the start codon. This sequence (the Shine-Dalgarno sequence) hydrogen-bonds to a polypyrimidine sequence (CCUCC) near the 3′-end of the small subunit rRNA (16 S rRNA) during initiation complex formation. This rRNA-mRNA interaction facilitates the correct selection of the start codon by the 30 S subunit (6–8). In contrast, in the eukaryotic cytoplasmic system, 40 S ribosomal subunits carrying the initiator tRNA interact with the cap structure on the mRNA and, in the presence of appropriate initiation factors, scan the mRNA until it reaches the first AUG codon from the 5′-end, where initiation generally occurs (9). In higher plants, statistical analysis of putative start sites on chloroplast mRNAs indicates that some of these messages contain Shine-Dalgarno sequences 5′ to the start codon (10). Other mRNAs do not appear to have an equivalent sequence. The signals used to designate a certain AUG codon as a start codon in this latter class of mRNAs remain to be clarified.

The analysis of putative translational start sites based on DNA sequence information from the chloroplasts of *E. gracilis* has been complicated by the presence of Group III introns, which are difficult to predict based on sequence information alone (11, 12). However, it appears that less than half of the chloroplast mRNAs in this organism have traditional Shine-Dalgarno sequences (13). Previously, we proposed that the translational machinery in the chloroplasts of *E. gracilis* has two potential mechanisms available for the recognition of translational start sites (13). In one class of sites, a sequence equivalent to the Shine-Dalgarno sequence is present and can hydrogen-bond to the CUCCC sequence found at the 3′-end of the small subunit rRNA in *E. gracilis* chloroplasts (14). In the second class of mRNAs, no Shine-Dalgarno sequence is present. In these mRNAs, the A/U content of the region just 5′ to the AUG start codon is very high (generally >90%). This portion of the mRNA may therefore have very little stable secondary structure. In this class of mRNAs, the AUG start codon may be selected because it resides in a region of the mRNA that has little secondary structure, making it readily available for interaction with the 30 S ribosomal subunit (13).

The investigation into the nucleotide sequence signals in *E. gracilis* chloroplast mRNAs has been facilitated by the development of an *in vitro* system that measures the formation of initiation complexes (13). In this paper, we have examined the role of primary sequence information in initiation complex formation with the chloroplast mRNA encoding the large subunit of ribulose-bisphosphate carboxylase/oxygenase. Although many chloroplast mRNAs are polycistronic, the message encoding the large subunit is monocistronic and has a 55-base 5′-untranslated leader. This mRNA does not contain a Shine-Dalgarno sequence and thus provides a model system in which to delineate the translational initiation signals recognized by ribosomes in one class of chloroplast mRNAs. In the accompanying paper (44), the effects of secondary structural

* 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 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed.

[1] The abbreviation used is: IF-2_{chl}, chloroplast initiation factor 2.

TABLE I
Mutagenic oligonucleotides

Plasmid	Oligonucleotide
pRbcN UAU	Oligo 1: TTTTATTATT AAAATNNNTC ACCTCAAACTG
pRbcN GAC	Oligo 1: TTTTATTATT AAAATNNNTC ACCTCAAACTG
pRbcN Δ2C	Oligo 2: TAATACGACT CACTATAGTT TTATTTTATT ATTAAAATAT GTCAC
pRbcN Δ28	Oligo 3: CGACTCACTA TAGAATATAA TTTTTATTTT ATTATTAAAA
pRbcN Δ40	Oligo 4: CTCACTATAG AGTGATTTTT CG
pRbcN Δ21,40	Oligo 5: AAATAACTGT AATTTTATTT TATTATT
pRbcN SD	Oligo 6: TCGAATATAA TTTTTATTTT GGGAGTAAAA TATGCACC
pRbcN SDΔ20	Oligo 7: GACTCACTAT AGTTTTATTT TGGG

elements on the recognition of this translational start site are explored.

EXPERIMENTAL PROCEDURES

Materials—General chemicals were purchased from Sigma. RNasin, *Hin*dIII, T4 DNA polymerase, and T7 RNA polymerase were purchased from Promega Biotec. The Sequenase™ version 2.0 kit, T4 polynucleotide kinase, calf intestine alkaline phosphatase, T7 DNA polymerase, and helper phage M13KO7 were obtained from U. S. Biochemical Corp. Deoxyribonucleoside triphosphates and ribonucleoside triphosphates were from Pharmacia LKB Biotechnology Inc. α-^{35}S-dATP (1200 Ci/mmol) and [5,6-^{3}H]UTP (43.3 Ci/mmol) were obtained from DuPont NEN. Type HA 0.45-μm nitrocellulose filter paper was purchased from Millipore Corp. The GeneClean™ DNA purification kit was from BIO 101, Inc. Yeast tRNA was from Boehringer Mannheim. Synthetic oligonucleotide primers for sequencing and mutagenic reactions (see Table I) were prepared in the Departments of Pathology and of Microbiology and Immunology and at the Lineberger Cancer Center of the University of North Carolina at Chapel Hill.

Plasmid Preparation and Transcription—The preparation of the plasmid pRbcN has been described previously (13). A derivative of this construct (pRbcN X2) has had two *Xba*I sites introduced upstream of the start site of the RbcN open reading frame (15). A derivative of pRbcN X2 was prepared in which a large portion of the 5′-untranslated leader sequence was present as the inverse complement. Digestion of pRbcN X2 with *Xba*I released a 48-base pair fragment from the 5′-untranslated leader region. The 48-base pair fragment and the remaining large fragment of the plasmid were separated by electrophoresis on a 2% agarose gel. The large fragment was extracted from the gel using the GeneClean™ method according to the manufacturer's instructions. The 48-base pair fragment was recovered from the gel by electrophoresis onto Whatman DE81-cellulose paper (16) and was ligated with the large fragment, which had been treated with calf intestine alkaline phosphatase (17). The ligation mixture was used to transform *E. coli* JM101 (18, 19). A derivative (pRbcN XF) in which the 48-base pair fragment had been inserted in the reverse orientation was identified by DNA sequencing.

In vitro mutagenesis of the AUG start codon of the *rbcL* gene was carried out using oligonucleotide 1 (Table I). During synthesis, this oligonucleotide was doped at the positions designated N with 33% of the wild-type nucleotide and 67% of an equimolar mixture of the other 3 nucleotides. Deletions of portions of the 5′-untranslated leader region of the *rbcL* gene were carried out using oligonucleotides 2–5 (Table I). The insertion of a Shine-Dalgarno sequence (GGGAG) into the 5′-untranslated leader region and a partial deletion derivative of this construct were prepared using oligonucleotides 6 and 7. Site-directed mutagenesis was carried out according to the method of Kunkel *et al.* (20, 21) and McClary *et al.* (22). Mutants of interest were identified by DNA sequencing.

Prior to transcription, plasmids were prepared as described (17), further purified by centrifugation in CsCl gradients, and linearized by digestion with *Hin*dIII. *In vitro* transcription was carried out as described previously (13).

Preparation of Formylmethionyl-tRNA, Ribosomal Subunits, and Initiation Factors—[^{35}S]Formylmethionyl-tRNA$_i^{Met}$ was prepared from yeast tRNA as described (23). A high salt wash of *E. coli* ribosomes was prepared as described (24); concentrated by ammonium sulfate precipitation; and dialyzed against 50 mM Tris-HCl, pH 7.6, 50 mM NH_4Cl, 1 mM EDTA, 1 mM dithiothreitol, and 10% glycerol before use. *E. coli* initiation factors were partially purified by chromatography on DEAE-cellulose and phosphocellulose (24). Chloroplast 30 S ribosomal subunits from *E. gracilis* were prepared as described previously (25). IF-2_{chl} purified through the TSKgel DEAE-5PW HPLC column stage (1) and IF-3_{chl} purified through the TSKgel SP-5PW HPLC column stage (5) were kindly provided by Dr. Lan Ma and Qiong Lin (Department of Chemistry, University of North Carolina at Chapel Hill).

Initiation Complex Formation—Reaction mixtures using a high salt wash of *E. coli* ribosomes as the source of initiation factors were prepared and analyzed as described previously (13), except that the concentration of NH_4Cl was reduced to 40 mM. Reaction mixtures (50 μl) using the partially purified initiation factors were prepared as described (13) and contained 10 pmol of the indicated transcript, 1 unit of *E. coli* IF-1, and either *E. coli* initiation factors (0.12 unit of IF-2 and 0.3 unit of IF-3) or chloroplast initiation factors (0.3 unit of IF-2_{chl} and 1 unit of IF-3_{chl}). The units described here are as indicated previously (1, 5).

To test the ability of an oligonucleotide containing a Shine-Dalgarno sequence to inhibit initiation complex formation with various mRNAs, the indicated amount of the oligodeoxynucleotide dGGGAGUU, which is complementary to the sequence at the 3′-end of the chloroplast 16 S rRNA from *E. gracilis*, was hybridized to 0.06 A_{260} unit of chloroplast 30 S subunits by incubation at 4 °C for 16 h in 5 μl of buffer containing 25 mM Tris-HCl, pH 7.8, 50 mM NH_4Cl, 0.1 mM EDTA, 10 mM $MgCl_2$, 1 mM spermidine, and 20% glycerol. Following this incubation, the 30 S subunits were tested for the ability to participate in initiation complex formation as described (13) in reaction mixtures containing 5 pmol of the indicated mRNA.

Measurement of Off Rate of Translational Initiation Complexes—Reaction mixtures (50 μl) containing 5 pmol of the indicated message, 0.06 A_{260} unit of chloroplast 30 S subunits, and 6 pmol of [^{35}S]fMet-tRNA$_i$ were prepared as described previously (13). After incubation for 10 min at 37 °C to allow initiation complex formation, 100 pmol of nonradioactive fMet-tRNA$_i$ in 10 mM potassium succinate, pH 6.0, were added, giving a final volume of 70 μl. Incubation was continued at 37 °C, and at the indicated times, the reaction was terminated by the addition of 1 ml of cold stop buffer (50 mM Tris-HCl, pH 7.8, 40 mM NH_4Cl, and 10 mM $MgCl_2$). The amount of [^{35}S]fMet-tRNA remaining associated with 30 S complexes was determined as described previously (13).

RESULTS

Selection of Correct Translational Start Codon in RbcN Construct—In our previous work (13), a clone was prepared carrying the 5′-untranslated leader and exon 1 of the *E. gracilis* *rbcL* gene fused to a portion of the neomycin resistance gene under the control of the T7 transcriptional promoter. The mRNA transcribed from this vector (mRbcN) contains the entire 55-base untranslated leader of the *rbcL* gene (Fig. 1) and is identical to the *in vivo* RNA except for the presence of a 5′-terminal G residue that facilitates transcription by the T7 polymerase (26). This mRNA does not contain a Shine-Dalgarno sequence upstream of the start codon, but does have a very A/U-rich 5′-untranslated leader region (Fig. 1).

The RbcN mRNA directs the formation of chloroplast translational initiation complexes (fMet-tRNA·mRNA·30 S) in the presence of either *E. coli* or chloroplast initiation factors (15). Transcripts having deletions of the entire leader through the known AUG start codon are inactive in initiation complex formation, suggesting that the correct start site of this mRNA is being used for initiation complex formation (13). To verify this conclusion, the AUG start codon was mutated using saturation mutagenesis. A number of mutations in the putative AUG start codon were obtained, and two mutants were selected for additional study. The first of these mutants had three changes leading to a UAU codon in place of the original AUG start codon

```
mRbcN :-56    -50        -40        -30        -20        -10        +1
5'-   GUAUAA AUAACUGUAA AGUGAUUUUU CGAAUAUAAU UUUUAUUUUA UUAUUAAAAU AUG

mRbcN UAU :
5'-   GUAUAA AUAACUGUAA AGUGAUUUUU CGAAUAUAAU UUUUAUUUUA UUAUUAAAAU UAU

mRbcN GAC :
5'-   GUAUAA AUAACUGUAA AGUGAUUUUU CGAAUAUAAU UUUUAUUUUA UUAUUAAAAU GAC

mRbcN XΔ9 :
5'-   G[------------------ Deleted ----------------------]UCUAGAAAU AUG

mRbcN Δ20 :
5'-   G[------------- Deleted ------------] UUUUAUUUUA UUAUUAAAAU AUG

mRbcN Δ28 :
5'-   G[-------- Deleted ---------]AAUAUAAU UUUUAUUUUA UUAUUAAAAU AUG

mRbcN Δ40 :
5'-   G[--- Deleted --] AGUGAUUUUU CGAAUAUAAU UUUUAUUUUA UUAUUAAAAU AUG

mRbcN Δ1,25 :
5'-   GUAUAA AUAACUGUAA AGUGAUUUUU CGAAU [------- Deleted --------] AUG

mRbcN Δ21,40 :
5'-   GUAUAA AUAACUGUAA [----- Deleted -----] UUUUAUUUUA UUAUUAAAAU AUG

mRbcN X2 :
5'-  GUCUAGAU AACUGUAAAG UGAUUUUUCG AAUAUAAUUU UUAUUUUAUU AUCUAGAAAU AUG

mRbcN XF :
5'-  GUCUAGAU AAUAAAAUAA AAAUUAUAUU CGAAAAAUCA CUUUACAGUU AUCUAGAAAU AUG
      ↑_____________________________________↑; inverted region

mRbcN SD :
5'-   GUAUAA AUAACUGUAA AGUGAUUUUU CGAAUAUAAU UUUUAUUUUG GGAGUAAAAU AUG

mRbcN SDΔ20:
5'-   G[------------- Deleted ------------] UUUUAUUUUG GGAGUAAAAU AUG
```

FIG. 1. **Sequences in 5'-untranslated leader regions of mRNA constructs.** Mutated nucleotides are *underlined*. The *Xba*I site is UCUAGA.

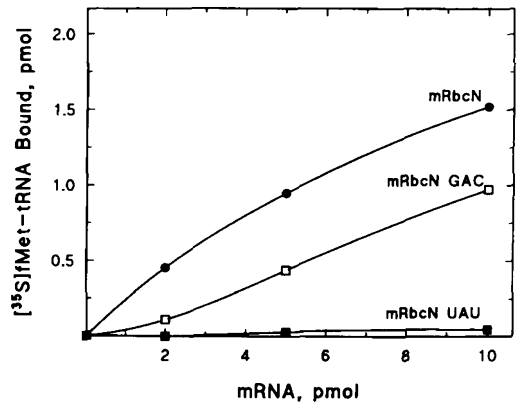

FIG. 2. **Start codon selection.** Initiation complex formation assays were performed using the indicated amounts of mRbcN (●) or its derivatives containing mutations of the start codon: mRbcN UAU (■) and mRbcN GAC (□).

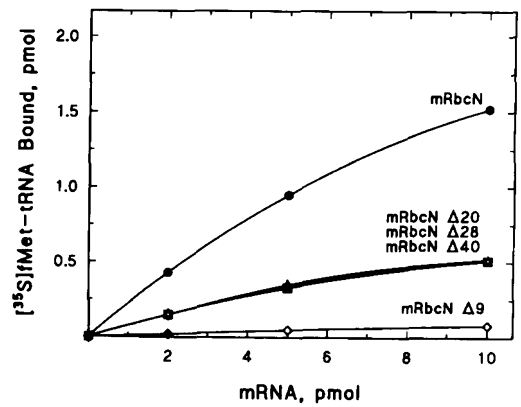

FIG. 3. **Effect of 5'-deletions on efficiency of initiation.** Initiation complex formation assays were performed using the indicated amounts of wild-type mRNA or mRNAs having 5'-end deletions: wild-type mRbcN (●), mRbcN Δ9 (◇), mRbcN Δ20 (△), mRbcN Δ28 (▽), and mRbcN Δ40 (□).

(mRbcN UAU; Fig. 1). This message was tested for the ability to participate in initiation complex formation. As indicated in Fig. 2, mRbcN UAU is essentially inactive in this assay. This result strongly confirms the previous conclusion that initiation complex formation with chloroplast 30 S subunits and mRbcN is occurring at the correct translational start site (13). In addition, this result argues that the other 18 AUG codons present in mRbcN are unable to direct initiation complex formation with chloroplast 30 S subunits.

The second mutagenic construct tested also contained a 3-base change in the AUG codon, converting it to a GAC codon (mRbcN GAC; Fig. 1). This mutation created a new AUG codon just 2 nucleotides upstream of the original AUG codon. When mRbcN GAC was tested for the ability to promote initiation complex formation, it was ~60% as efficient as the original mRNA (Fig. 2). This result again argues that the AUG codon in the 5'-region of mRbcN is the codon selected for initiation complex formation by chloroplast 30 S ribosomal subunits. The somewhat reduced efficiency of mRbcN GAC compared with the wild-type mRNA probably reflects the different context immediately surrounding the start codon (AAUGA *versus* UAUGU). There are a number of cases in which the residues immediately surrounding the AUG codon affect the efficiency of initiation complex formation in bacterial systems (27, 28) and in the eukaryotic cytoplasm (29). Further mutagenesis of the nucleotides adjacent to the start codon will be required to examine the subtle effects of start codon context on the efficiency of initiation complex formation.

Effect of Length of Untranslated Leader on Initiation Complex Formation—In previous work (15), the 55-base 5'-untranslated leader of the RbcN construct was modified by mutagenesis to introduce two *Xba*I sites (mRbcN X2; Fig. 1). The mRNA transcribed from this construct is indistinguishable from the wild-type mRNA in assays of initiation complex formation (15). Deletion of the nucleotides between these *Xba*I sites in the leader results in the synthesis of a mRNA (mRbcN XΔ9; Fig. 1) that contains only 10 nucleotides 5' to the normal start codon (9 nucleotides of the natural 5'-untranslated leader and a 5'-terminal G residue). This mRNA has little or no activity in initiation complex formation (Fig. 3) (15). There are two reasonable explanations for this observation. First, the 5'-untranslated leader may have essential sequence information that directs the chloroplast 30 S subunit to the AUG start codon. Second, there may be a minimal length in the 5'-region that is essential for ribosomal subunit binding. Since the 30 S subunit would be expected to cover ~20 nucleotides upstream of the start codon, deletion of all but 10 nucleotides 5' to the initiation codon could result in a mRNA that cannot interact with the small subunit.

To examine the importance of the length and sequence of the 5'-untranslated leader of mRbcN, several additional deletion mutants were prepared. The first deletion mutation (mRbcN Δ20; Fig. 1) has the sequence from positions −55 to −21 deleted, leaving 20 nucleotides of the 5'-untranslated leader from the wild-type sequence and the additional G residue at the 5'-end. The mRbcN Δ20 mRNA has significantly reduced activity in initiation complex formation (Fig. 3). It promotes fMet-tRNA binding to 30 S subunits only about one-third as well as the normal mRNA. The second mutation (mRbcN Δ28; Fig. 1) has residues −55 to −29 deleted. This mRNA contains a 5'-untranslated leader with 28 residues from the wild-type sequence and the additional 5'-terminal G residue. This mRNA also shows a reduction in the ability to promote initiation complex formation (Fig. 3). Indeed, it is no more efficient than the mRNA with the 5'-untranslated leader that has 20 bases present. Another deletion mutation lacks residues −55 to −41 (mRbcN Δ40; Fig. 1). This mRNA has a 5'-leader containing 40 of the 55 residues of the wild-type sequence. However, it is only 30% as efficient as the full-length mRNA in promoting fMet-tRNA binding to 30 S subunits (Fig. 3). These observations suggest that the AUG start codon can be selected in the mRNA with only 20 residues

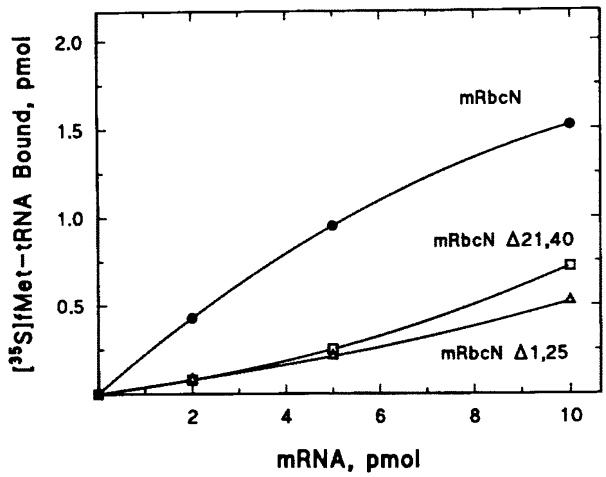

FIG. 4. **Effect of internal deletions in 5′-untranslated leader on efficiency of initiation.** Initiation complex formation assays were performed using the indicated amounts of wild-type mRNA or mRNA derivatives having internal deletions in the 5′-untranslated leader: wild-type mRbcN (●), mRbcN Δ1,25 (△), and mRbcN Δ21,40 (□).

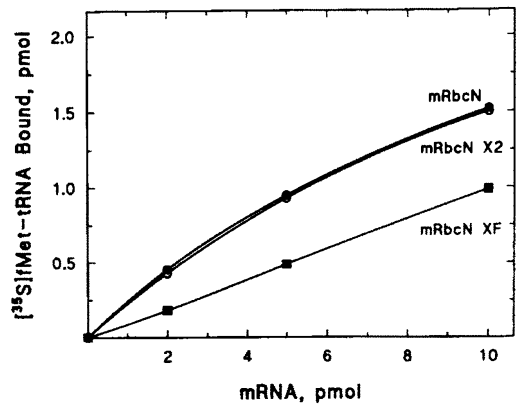

FIG. 5. **Effect of primary sequence changes in 5′-untranslated leader.** Initiation complex formation was assayed with the indicated amounts of mRbcN (●), the derivative containing two *Xba*I sites in the 5′-untranslated leader (mRbcN X2; ○), and a mRNA having a flipped sequence in the 5′-leader (mRbcN XF; ■).

of the 5′-leader present. However, maximal activity in initiation complex formation requires the full 5′-untranslated leader.

There are two possible interpretations for the observations described above. First, it is possible that the inherent length of the 5′-untranslated leader in the mRNA (55 base pairs) is an important determinant for providing maximal efficiency of initiation. A second interpretation is that the sequences between positions −40 and −55 are important as a translational enhancer. To distinguish between these two possibilities, two internal deletion constructs were prepared. In the first, nucleotides between positions −1 and −25 were deleted, providing a 31-nucleotide 5′-untranslated leader carrying the residues originally between positions −40 and −55. The mRNA prepared from this construct (mRbcN Δ1,25; Fig. 1) has only ~30% of the ability of the wild-type mRNA to participate in initiation complex formation (Fig. 4). The second mutant carries an internal deletion from positions −21 to −40 in the 5′-untranslated leader (mRbcN Δ21,40; Fig. 1). The mRNA prepared from this deletion mutant contains a 35-residue 5′-untranslated leader from the *rbcN* message. This mRNA also shows a significant reduction in its ability to participate in initiation complex formation (mRbcN Δ21,40; Fig. 4). Thus, internal deletions in the 5′-untranslated leader cause a reduction in activity similar to that caused by the 5′-end deletions (Fig. 3). These results argue that the length of the 5′-untranslated leader is important for maximal activity of mRbcN in initiation.

Role of Primary Sequence of 5′-Untranslated Leader in Initiation—The results summarized above indicate that the full length of the 55-base leader of mRbcN is essential for maximal activity in initiation complex formation. However, these results do not clearly indicate whether there is essential sequence information in the 5′-untranslated leader that strongly influences the initiation process. Rather, the effects of the deletions summarized above appear to arise from changes in the length of the mRNA itself. To address the question of whether there are elements of essential sequence information in the leader, cassette mutagenesis was carried out to prepare a new mutant using the mRbcN X2 construct (Fig. 1) as the starting plasmid. The presence of the *Xba*I sites (positions −57 to −52 and −9 to −4) allowed the excision of a 48-base pair segment from the plasmid and its insertion in the opposite orientation (mRbcN XF; Fig. 1). Transcription of this plasmid results in the synthesis of a mRNA with a 5′-leader region that is the inverse complement of the mRbcN X2 sequence between positions −9 and −56. The 5′-untranslated leader from this construct has base pair changes at 24 out of 58 positions scattered throughout the leader region (*underlined* in Fig. 1) and has about the same percent A/U found in mRbcN X2. Interestingly, mRbcN XF is ~60% as active as mRbcN when tested for the ability to participate in initiation complex formation (Fig. 5). These results suggest that the primary sequence of the 5′-untranslated leader region is not the major determinant in specifying a particular AUG codon as a translational start site.

Effect of Shine-Dalgarno Sequence in 5′-Untranslated Leader—As noted above, chloroplast mRNAs in *E. gracilis* fall into two major classes. The first class of mRNAs appears to have a Shine-Dalgarno sequence comparable to that observed in prokaryotic systems. The second class of mRNAs, including mRbcN, does not have a recognizable Shine-Dalgarno sequence. It was of interest to determine the effect of a Shine-Dalgarno sequence on the ability of mRbcN to promote initiation complex formation. To examine this question, oligonucleotide-directed mutagenesis was used to replace a short stretch of nucleotides between positions −7 and −11 with the strong *E. gracilis* chloroplast Shine-Dalgarno sequence (GGGAG). The transcript designated mRbcN SD (Fig. 1) was then prepared and tested for its ability to direct translation initiation. As indicated in Fig. 6, the presence of the Shine-Dalgarno sequence 5′ to the AUG start codon enhances the ability of the mRNA to promote initiation complex formation ~1.4-fold. This prokaryote-like initiation signal is thus only modestly helpful to the mRNA in enhancing initiation. This result is in stark contrast to the critical role played by the Shine-Dalgarno sequence in prokaryotes, in which it generally enhances initiation by 10–100-fold.

To determine whether the presence of a Shine-Dalgarno sequence could offset the effect of a reduction in the length of the 5′-untranslated leader, a deletion of this construct was prepared that would direct the synthesis of a mRNA containing 20 nucleotides 5′ to the start codon (mRbcN SDΔ20; Fig. 1). The Shine-Dalgarno sequence is still positioned between positions −7 and −11 in this transcript. As indicated in Fig. 6, the deletion derivative is ~3-fold less effective in initiation complex formation compared with the full-length mRNA carrying a Shine-Dalgarno sequence. This reduction in activity is comparable to that observed with the normal RbcN mRNA and its deletion derivatives (Figs. 3 and 6). These results indicate that a long 5′-untranslated leader region is probably important for maximal efficiency in initiation whether or not a chloroplast mRNA contains a Shine-Dalgarno sequence.

Effect of Oligodeoxynucleotide Competition on Initiation

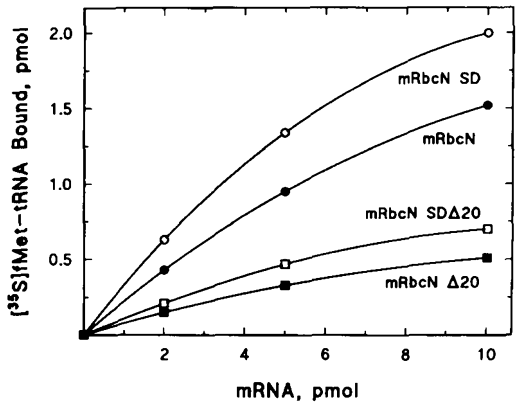

FIG. 6. **Effect of Shine-Dalgarno sequence on initiation complex formation with mRbcN.** Initiation complex formation was tested with the indicated amounts of mRbcN (●), its deletion derivative (mRbcN Δ20; ■), the derivative containing a strong Shine-Dalgarno sequence in the 5′-leader (mRbcN SD; ○), and a deletion mutant of mRbcN SD (mRbcN SDΔ20; □).

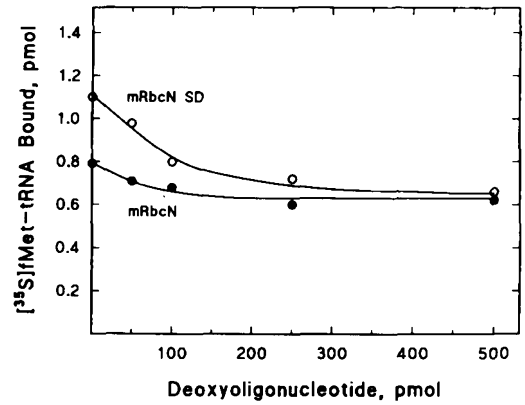

FIG. 7. **Competition of Shine-Dalgarno oligodeoxynucleotide in initiation complex formation.** Initiation complex formation was assayed using 30 S subunits that had been incubated in the presence of the indicated amounts of the Shine-Dalgarno septadeoxynucleotide (5′-GGGAGTT-3′) as described under "Experimental Procedures." The subunits were then tested for their ability to form initiation complexes using a 2-min incubation time and either mRbcN (●) or mRbcN SD (○).

Complex Formation—In prokaryotes, the Shine-Dalgarno sequence in the mRNA hydrogen-bonds to the polypyrimidine sequence located near the 3′-end of the 16 S rRNA. This interaction places certain constraints on the path of the mRNA along the small ribosomal subunit in the initiation complex. It was of interest to ask whether a mRNA that does not contain a Shine-Dalgarno sequence binds to the 30 S subunit in the same way as a mRNA having a Shine-Dalgarno sequence. In an effort to examine this question, the effect of an oligodeoxynucleotide that can bind to the polypyrimidine sequence at the 3′-end of the small subunit rRNA on the ability of mRbcN to form initiation complexes was tested. This oligodeoxynucleotide contains a GGGAG sequence that can hydrogen-bond to the 3′-end of the *E. gracilis* chloroplast 16 S rRNA. Chloroplast 30 S ribosomal subunits were incubated overnight with increasing concentrations of the oligodeoxynucleotide to allow it to bind to the small subunit. The 30 S subunit-oligodeoxynucleotide complexes were then tested for their ability to form initiation complexes with mRbcN. Small subunits incubated under these same conditions overnight in the absence of the oligodeoxynucleotide retained >80% of their activity. As indicated in Fig. 7, a 100-fold excess of the Shine-Dalgarno oligodeoxynucleotide over 30 S subunits has almost no effect on the formation of initiation complexes with mRbcN. Preincubation of 30 S subunits with a 500-fold excess of the oligodeoxynucleotide results in a small reduction (<25%) in the ability of the 30 S subunits to participate in initiation complex formation with mRbcN. This observation suggests that this chloroplast mRNA can bind to the small subunit without any close interaction with the 3′-end of the 16 S rRNA. This conclusion is in agreement with the idea that there is no Watson-Crick hydrogen bonding between the rRNA and the 5′-untranslated leader region of chloroplast mRNAs that do not contain a Shine-Dalgarno sequence.

The effect of the oligodeoxynucleotide containing the Shine-Dalgarno sequence on the ability of mRbcN SD to form initiation complexes was also tested. This mRNA has a Shine-Dalgarno sequence equivalent to that present in the competing oligodeoxynucleotide. As indicated in Fig. 7, mRbcN SD is somewhat more efficient in initiation complex formation compared with the starting mRbcN message. In the presence of a 100-fold excess of the competing oligodeoxynucleotide, the ability of the Shine-Dalgarno sequence-containing mRNA is reduced essentially to the level observed with mRbcN. At higher concentrations of the competing oligodeoxynucleotide, mRbcN and mRbcN SD behave identically. These results suggest that the small enhancement in the efficiency of initiation complex formation with the Shine-Dalgarno sequence-containing mRNA probably arises from a direct hydrogen bonding between the small subunit rRNA and the Shine-Dalgarno sequence in the mRNA. The oligodeoxynucleotide appears to be able to compete with this rRNA-mRNA interaction, leaving mRbcN SD with an ability to form initiation complexes equivalent to that observed with the wild-type leader. These data again suggest that while mRbcN SD may take advantage of the Shine-Dalgarno sequence to increase the efficiency of initiation slightly, the leader of this mRNA has an intrinsic ability to direct initiation in the absence of any classical rRNA-mRNA hydrogen bonding. It should be noted that Canonaco *et al.* (30) have observed that an oligodeoxynucleotide carrying a prokaryotic Shine-Dalgarno sequence inhibits the binding of an artificial message carrying a Shine-Dalgarno sequence to *E. coli* 30 S subunits more strongly than it inhibits the interaction of a corresponding mRNA lacking a Shine-Dalgarno sequence. Thus, the results obtained here are quite compatible with those observed in the bacterial system.

Stability of Initiation Complexes Formed with Various mRNAs—The data provided above suggest that the 5′-untranslated leader of mRbcN contains essential information that directs the 30 S ribosomal subunit to the AUG start codon. The data also argue that a section of the leader >20 nucleotides is important for maximal initiation and that the addition of a Shine-Dalgarno sequence has a small enhancing effect on the ability of this mRNA to participate in initiation. There are several possible stages during initiation complex formation in which variations in the 5′-leader sequence might influence the initiation process. The formation of the initiation complex occurs by a rather complicated sequence of events, many of which may be important for efficient initiation (8). Changes in the 5′-untranslated leader might also lead to initiation complexes that are inherently less stable, resulting in a more rapid rate of dissociation. In an effort to determine whether the differences in the amount of initiation complex observed with several of the mRNAs used here resulted from changes in the rate of formation of the complexes, efforts were made to examine the time courses for initiation complex formation with a number of the mRNAs. However, the amount of initiation complex observed in every case had reached a maximal value within 2 min after assembly of the reaction mixtures and remained constant for ~20 min. Since the 2-min time point was the earliest one that could be measured accurately, potential changes in the rate of initiation complex formation could not be measured directly

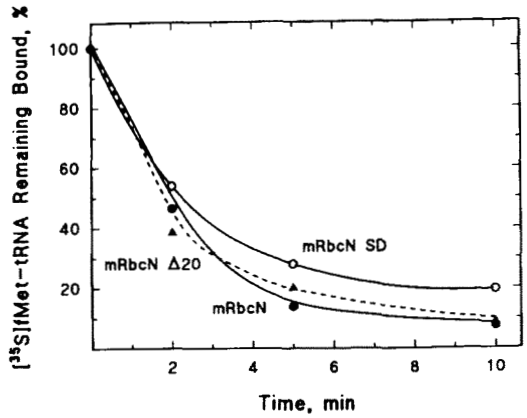

FIG. 8. **Rate of dissociation of initiation complexes formed with various mRNAs.** Initiation complexes were formed as described under "Experimental Procedures" using 5 pmol of mRbcN (●), mRbcN SD (○), or mRbcN Δ20 (▲). After formation of the complexes, unlabeled fMet-tRNA was added, and the amount of [^{35}S]fMet-tRNA remaining in 30 S initiation complexes was determined at the indicated times.

with the various mRNAs available. The data obtained indicate that the measurements reported here represent equilibrium values. Differences in the amount of initiation complexes observed with the various mRNAs probably are related to the equilibrium constants governing the formation of the various complexes.

Although it was not practical to obtain a direct measure of the forward rate constants for initiation complex formation, it was possible to gain some insight into the stability of the various complexes as measured by the off rate of the bound fMet-tRNA. To observe the off rate, 30 S initiation complexes were assembled with the appropriate mRNA and labeled fMet-tRNA. The reaction mixtures were then diluted in the presence of unlabeled fMet-tRNA, and the amount of labeled complex remaining as a function of time was measured. The unlabeled fMet-tRNA was added in a 17-fold molar excess to mask the reassociation reaction. The dissociation of the labeled initiation complex formed with mRbcN can clearly be observed following the addition of the unlabeled competitor (Fig. 8). With this mRNA, nearly half of the original complex is dissociated within 5 min, and <10% remains after 10 min of incubation. The stability of the complex formed with the mRbcN Δ20 deletion mutant is essentially the same as that formed with the natural mRNA. This observation suggests that it is the formation of the initiation complex rather than its dissociation that controls the amount of complex present at equilibrium. A similar analysis was carried out with mRbcN SD (Fig. 8). Again, the off rate observed is essentially identical to that found with mRbcN, indicating that the somewhat enhanced initiation complex formation observed with this mRNA probably arises from changes in the rate of formation of the complex rather than from a slower rate of dissociation of the complex once assembled.

Role of Chloroplast Initiation Factors in Initiation Complex Formation—The experiments described above were performed using chloroplast 30 S ribosomal subunits and initiation factors from *E. coli*. In general, it is believed that the recognition of the translational start site is carried out by the small subunit of the ribosome rather than by the initiation factors in prokaryote-type systems (7, 8). To make sure that the observations described above were not significantly affected by the use of the bacterial initiation factors, several of the constructs were tested for activity in initiation complex formation using IF-2_{chl} and IF-3_{chl} rather than the crude *E. coli* initiation factor preparation. As indicated in Table II, initiation complex formation could be observed using the chloroplast initiation factors and

TABLE II
Initiation complex formation with chloroplast initiation factors

mRNA	Initiation complex formation[a]	
	Chloroplast factors	*E. coli* factors
mRbcN	0.30	0.38
mRbcN SD	0.70	0.74
mRbcN Δ20	0.07	0.11
mRbcN SDΔ20	0.15	0.14

[a] Picomoles of fMet-tRNA bound to 30 S subunits.

mRbcN. When mRbcN SD was tested in this assay, an ~2-fold enhancement in activity was observed compared to that obtained with mRbcN (Table II). This result is essentially the same as that observed with the purified *E. coli* factors. Finally, the deletion derivatives mRbcN Δ20 and mRbcN SDΔ20 were tested and showed about the same -fold reduction in activity as observed in the previous experiments (Fig. 6 and Table II). These observations indicate that the results obtained with the bacterial factors are a direct reflection of the ability of the chloroplast translational system to use the mRNA derivatives tested here.

DISCUSSION

In *E. coli*, several features of the mRNA appear to play an important role in translational initiation. The first and most striking feature is the Shine-Dalgarno sequence. Mutations in this region have a drastic effect on the efficiency of initiation (31–33). In addition, statistical analysis indicates that nucleotides −20 to +13 are not random, suggesting that this whole region has a modulating effect on the efficiency of initiation (32, 34, 35). Finally, it is clear that secondary structural features can play a critical role in initiation. Start sites with little or no secondary structure are generally much more efficient than those in which the Shine-Dalgarno sequence or AUG codon is base-paired (36, 37).

In this work, a number of mutants were constructed to delineate the region of the 5′-untranslated leader in the *E. gracilis* chloroplast *rbcL* mRNA that is important for specifying the AUG start codon. This mRNA does not have a Shine-Dalgarno sequence and hence cannot use the major bacterial determinant for specifying the start codon. The start site of the *rbcL* mRNA is probably specified by several features, all of which contribute to the selection process. One component that appears to be important is a 5′-untranslated leader of at least 20 nucleotides. This observation is quite reasonable considering data showing that the 30 S ribosomal subunit probably covers ~20 residues 5′ to the start codon (38). These 20 residues do not have to represent immediate upstream sequences since an internal deletion mutant shows almost the same activity as 5′-end deletion mutants. The results presented here indicate that maximal initiation complex formation is dependent on the presence of the full-length 55-base leader sequence. This chloroplast mRNA may thus have two areas through which it contacts the 30 S subunit: one near the AUG codon and the other involving more distal regions of the mRNA. This observation is compatible with those obtained in prokaryotes, where it is believed that regions of the mRNA beyond the ribosome-binding site contribute to the interaction of the mRNA with the 30 S subunit (31, 36, 39). For example, Borisova *et al.* (40) observed that 53 nucleotides upstream of the AUG codon are required for initiation complex formation with the replicase cistron of MS2. It has been suggested (31) that during initiation, the 30 S subunit interacts with a larger region of the mRNA than is protected in the final complex formed. The ribosome-binding site is thus only a part of the true initiation region (41).

Details are not yet available on all of the precise contacts between the prokaryotic 30 S subunit and mRNA. However, it

has been suggested that the mRNA interacts with the 30 S subunit via a U-shaped trough or channel (42). It is likely that this channel is formed from parts of both the rRNA and the ribosomal proteins. In *E. coli*, at least two proteins (S1 and S21) along with several stretches of the rRNA have been implicated in facilitating the binding of the mRNA to the 30 S subunit. Recent evidence (42) indicates that the mRNA does not fill the entire channel, but, depending on its specific sequence and structure, makes contacts with groups lining the channel.

The analysis of the translational start site of the *rbcL* mRNA presented here argues that the leader region does not contain a specific sequence that is absolutely required for determining the start site of translation. For the *E. gracilis* chloroplast 30 S subunit, the major interactions between the mRNA and the subunit may involve the ribose phosphate backbone of the mRNA rather than the bases themselves. Thus, the overall length and structure (or lack of structure) of this region may be the essential features present. An analysis of the effects of structural elements in the mRNA on initiation is presented in the following paper (44).

Finally, it should be noted that the insertion of a strong Shine-Dalgarno sequence 5′ to the start codon leads to only a modest improvement in the ability of the RbcN mRNA to participate in initiation complex formation. This result is quite different from *E. coli* messages, which show a much stronger requirement for the presence of a Shine-Dalgarno sequence in the mRNA (7). It is also compatible with the presence of two classes of mRNAs in *E. gracilis* chloroplasts, some having Shine-Dalgarno sequences and others lacking this sequence. The observation that the competing Shine-Dalgarno oligodeoxynucleotide overcame the slight advantage in initiation observed with the mRbcN SD argues that when a Shine-Dalgarno sequence is present, it probably results in a slightly altered placement of the mRNA in the RNA-binding trough on the 30 S subunit. This idea is compatible with the proposal that the RNA-binding channel is rather wide. Certain messenger RNAs may sit in this trough in a slightly different way than do other mRNAs (43).

Acknowledgments—We thank Dr. Lan Ma for the gift of *E. gracilis* IF-2_{chl} and Qiong Lin for the gift of IF-3_{chl}.

REFERENCES

1. Ma, L., and Spremulli, L. L. (1990) *J. Biol. Chem.* **265,** 13560–13565
2. Ma, L., and Spremulli, L. L. (1992) *J. Biol. Chem.* **267,** 18356–18360
3. Gold, J. C., and Spremulli, L. L. (1985) *J. Biol. Chem.* **260,** 14897–14900
4. Kraus, B. L., and Spremulli, L. L. (1986) *J. Biol. Chem.* **261,** 4781–4784
5. Wang, C.-C., and Spremulli, L. L. (1991) *J. Biol. Chem.* **266,** 17079–17083
6. Shine, J., and Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U. S. A.* **71,** 1342–1346
7. Gold, L. (1988) *Annu. Rev. Biochem.* **57,** 199–233
8. Gualerzi, C. O., and Pon, C. L. (1990) *Biochemistry* **29,** 5881–5889
9. Kozak, M. (1989) *J. Cell Biol.* **108,** 229–241
10. Bonham-Smith, P. C., and Bourque, D. P. (1989) *Nucleic Acids Res.* **17,** 2057–2080
11. Copertino, D. W., Christopher, D. A., and Hallick, R. B. (1991) *Nucleic Acids Res.* **19,** 6491–6497
12. Christopher, D. A., and Hallick, R. B. (1989) *Nucleic Acids Res.* **17,** 7591–7608
13. Wang, C.-C., Roney, W. B., Alston, R. L., and Spremulli, L. L. (1989) *Nucleic Acids Res.* **17,** 9735–9747
14. Steege, D. A., Graves, M. C., and Spremulli, L. L. (1982) *J. Biol. Chem.* **257,** 10430–10439
15. Roney, W. B., Ma, L., Wang, C.-C., and Spremulli, L. L. (1991) *NATO ASI Ser. Ser. H Cell Biol.* **55,** 197–205
16. Dretzen, G., Bellard, M., Sassone-Corsi, P., and Chambon, P. (1981) *Anal. Biochem.* **112,** 295–298
17. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Yanisch-Perron, C., Viera, J., and Messing, J. (1985) *Gene (Amst.)* **33,** 103–119
19. Hanahan, D. (1983) *J. Mol. Biol.* **166,** 557–580
20. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 488–492
21. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol.* **154,** 367–382
22. McClary, J. A., Witney, F., and Geisselsoder, J. (1989) *BioTechniques* **7,** 282–289
23. Graves, M. C., and Spremulli, L. L. (1983) *Arch. Biochem. Biophys.* **222,** 192–199
24. Suttle, D. P., Haralson, M. A., and Ravel, J. M. (1973) *Biochem. Biophys. Res. Commun.* **51,** 376–382
25. Graves, M. C., Breitenberger, C. A., and Spremulli, L. L. (1980) *Arch. Biochem. Biophys.* **204,** 444–454
26. Gingrich, J. C., and Hallick, R. B. (1985) *J. Biol. Chem.* **260,** 16162–16168
27. Eckhardt, H., and Lührmann, R. (1981) *Biochemistry* **20,** 2075–2080
28. Schmitt, M., Kyriatsoulis, A., and Gassen, G. H. (1982) *Eur. J. Biochem.* **125,** 389–394
29. Kozak, M. (1986) *Cell* **44,** 283–292
30. Canonaco, M. A., Gualerzi, C. O., and Pon, C. L. (1989) *Eur. J. Biochem.* **182,** 501–506
31. Gren, E. J. (1984) *Biochimie (Paris)* **66,** 1–29
32. Jacob, W. F., Santer, M., and Dahlberg, A. E. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 4757–4761
33. Hui, A., and de Boer, H. A. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 4762–4766
34. Schneider, T. D., Stormo, G. D., Gold, L., and Ehrenfeucht, A. (1986) *J. Mol. Biol.* **188,** 415–431
35. Stormo, G. D., Schneider, T. D., and Gold, L. M. (1982) *Nucleic Acids Res.* **10,** 2971–2996
36. de Smit, M. H., and van Duin, J. (1990) *Prog. Nucleic Acid Res. Mol. Biol.* **38,** 1–35
37. de Smit, M. H., and van Duin, J. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 7668–7672
38. Steitz, J. A. (1969) *Nature* **224,** 957–964
39. McCarthy, J., and Gualerzi, C. (1990) *Trends Genet.* **6,** 78–85
40. Borisova, G. P., Volkova, T. M., Berzin, V., Rosenthal, G., and Gren, E. J. (1979) *Nucleic Acids Res.* **6,** 1761–1774
41. Evstafieva, A. G., Shatsky, I. N., Bogdanova, A. A., Semenkov, Y. P., and Vasiliev, V. D. (1983) *EMBO J.* **2,** 799–804
42. Gualerzi C. O., Teana, A. L., Spurio, R., Canonaco, M. A., Severini, M., and Pon, C. L. (1990) in *The Ribosome: Structure, Evolution & Genetics* (Hill, W. E., Dahlberg, A., Garrett, R. A., Moore, P. B., Schlessinger, D., and Warner, J. R., eds) pp. 281–291, American Society for Microbiology Press, Washington, D. C.
43. Bhangu, R., and Wollenzien, P. (1992) *Biochemistry* **31,** 5937–5944
44. Koo, J. S., and Spremulli, L. L. (1994) *J. Biol. Chem.* **269,** 7501–7508