RNA (1995), *1*:79–88. Cambridge University Press. Printed in the USA.

Translational control of maturation-protein synthesis in phage MS2: A role for the kinetics of RNA folding?

HERMAN GROENEVELD, KARINE THIMON, and JAN VAN DUIN

Leiden Institute of Chemistry, Department of Biochemistry, Gorlaeus Laboratories, University of Leiden, P.O. Box 9502, 2300 RA Leiden, The Netherlands

ABSTRACT

The gene for the maturation (A) protein of the single-stranded RNA coliphage MS2 is preceded by an untranslated leader of 130 nt. Secondary structure of the leader was deduced by phylogenetic comparison and by probing with enzymes and chemicals. The RNA folds into a cloverleaf, i.e., three stem-loop structures enclosed by a long-distance interaction (LDI). This LDI is essential for translational control. Its 3′ moiety contains the Shine–Dalgarno region of the A-protein gene, whereas its complement is located 80 nt *upstream*, i.e., about 30 nt from the 5′-terminus of the RNA chain. Mutational analysis shows that this base pairing represses expression of the A-protein gene. We present a model in which translational starts can only take place on nonequilibrated RNA, in which base pairing between the complementary regions has not yet taken place. We suggest that this pairing is kinetically delayed by the intervening sequence, which contains the three hairpins of the cloverleaf. The model is mainly based on the observation that reducing the length of the intervening sequence reduces expression, whereas increasing the length has the opposite effect. In addition, further stabilization of the LDI by a stronger base pair does not lead to a decrease in A-protein synthesis. Such a decrease is predicted to occur if translation would be controlled by the equilibrium structure of the leader RNA. These and other observations fit a kinetic model of translational control by RNA folding.

Keywords: kinetics of RNA folding; RNA phage; RNA secondary structure; translation

INTRODUCTION

MS2 is a single-stranded RNA coliphage belonging to group I. Its RNA is of messenger-RNA polarity and contains the information for four proteins (Fig. 1A). Replicase assembles with several host proteins to form an active holoenzyme capable of synthesizing minus and plus strands. The lysis protein inserts into the cytoplasmic membrane, leading to cell lysis. The coat and maturation proteins are structural components of the icosahedral virion. Per virion there are 180 copies of the coat protein and only one copy of the maturation protein. The maturation or A-protein is necessary for the infection process. It contacts the F-pili of male *Escherichia coli* bacteria, which leads to one proteolytic hit in the A-protein and subsequent release of the RNA from the virion (see van Duin, 1988 for a review).

Reprint requests to: Jan van Duin, Leiden Institute of Chemistry, Department of Biochemistry, Gorlaeus Laboratories, University of Leiden, P.O. Box 9502, 2300 RA Leiden, The Netherlands; e-mail: duin_j@rulgca.leidenuniv.nl.

The timing and rates of gene expression in the RNA phages are controlled at the level of translation. In the past, control of coat, lysis, and replicase synthesis have been studied extensively (Horiuchi, 1975; Berkhout et al., 1987; Adhin & van Duin, 1989; de Smit & van Duin, 1990a, 1990b, 1993; Witherell et al., 1991; van Himbergen et al., 1993 and references therein). A major role in control is played by RNA folding, which can deny access of ribosomes to start regions. For instance, the lysis and replicase cistrons are not accessible to ribosomes in the unperturbed state of the RNA. Their expression is set in motion by translation of the coat-protein gene, which distorts the RNA structure and allows for their translation. As a consequence, replicase and lysis gene expression are translationally coupled to the coat gene.

To avoid that phage RNA serves at the same time as template for ribosomes and replicase, the two enzymes compete for a common site on the RNA at the start of the coat gene (Kolakofsky & Weissmann, 1971). This competition scheme evidently can only function if there are no independent ribosome entry sites for ly-

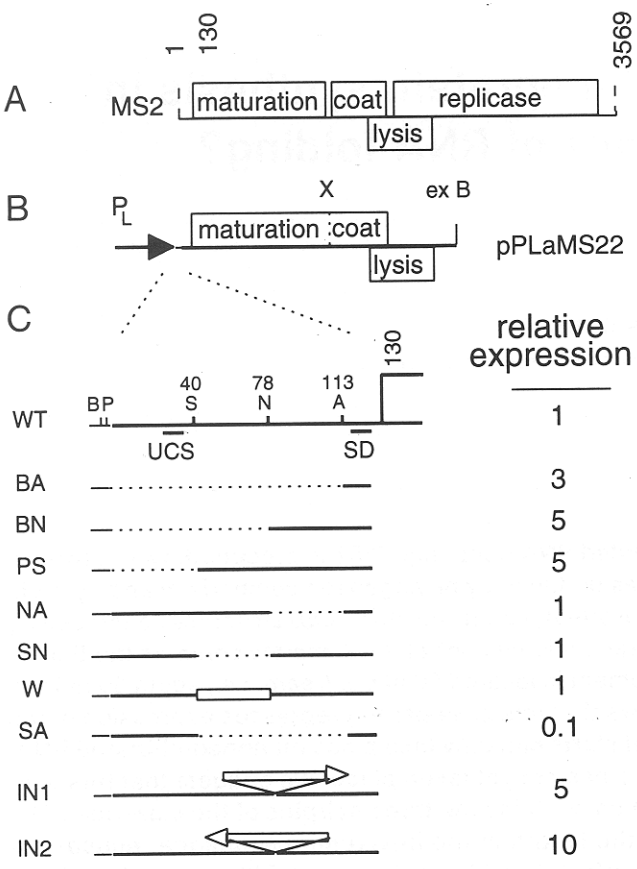

FIGURE 1. **A:** Genetic map of group I RNA bacteriophage MS2. **B:** Relevant part of the expression-vector pPLaMS22 (described in the Materials and methods) showing the translational fusion between maturation and coat-protein genes. **C:** Diagram of the leader sequence and average relative expression levels of deletion and insertion clones discussed in the text. Dotted lines indicate deletions; open arrows are insertions; open bar in construct W is a replacement by an unrelated sequence; UCS, upstream complementary sequence; SD, Shine–Dalgarno sequence; B = *Bam* HI, P = *Pst* I, S = *Sac* I, N = *Nco* I, A = *Avr* II, X = *Xba* I, exB = filled-in *Bam* HI site.

sis, replicase, and A-protein. Translational coupling of lysis and replicase to the coat gene thus provides for a single ribosome entry site for these three genes. This leaves only the control of the ribosome entry at the A-protein gene unsettled.

Not much is known about the regulation of the A-protein, except that its synthesis is kept down by secondary structure of the RNA; treatment with formaldehyde yielded about a 10-fold increase in A-protein synthesis in vitro (Robertson & Lodish, 1970). In this paper we examine the control of A-protein gene expression. To do so we have studied the structure of the 5′ untranslated region of MS2 RNA by phylogenetic comparison, chemical and enzymatic probing, and computer-assisted RNA folding. The region folds into a cloverleaf structure with a long-distance interaction (LDI) between the Shine–Dalgarno (SD) region of the A-protein gene and a region 80 nt *upstream*. This interaction is shown to repress translation. It should be noted that the inhibitory sequence is present *before* the ribosome binding site is transcribed. Our results suggest that formation of the LDI allows for a short time lag during which ribosomes can start translation at the maturation gene on the nonequilibrated RNA, i.e., RNA in the process of folding. Equilibrated RNA molecules are supposed not to bind ribosomes to a measurable degree. This regulatory system ensures both a low rate of A-protein synthesis and the exclusion of ribosomes from full-sized replicatable RNA.

RESULTS

Outline of the RNA structure and the regulatory mechanism

In Figure 2A we present the secondary structure for the 5′ untranslated leader of the group-I phage MS2. The first 28 nt fold into a separate hairpin that is followed by a small domain containing three local hairpins designated here as the W(est), S(outh), and E(ast) arms. The domain is closed by an LDI consisting of the SD region of the A-protein gene and an upstream complementary sequence (UCS). We envisage that some time will lapse between the synthesis of the SD sequence by the replicase and its pairing with the already present UCS. Only during this timespan the A-protein start would be accessible to ribosomes.

Phylogenetic support for the structural model

We have determined the sequences of the 5′ untranslated leaders of group I phages fr (Adhin et al., 1990), M12, JP501, and of the related group II phage KU1 (this work). Together with the published sequences of MS2 (Fiers et al., 1976) and GA (group II) (Inokuchi et al., 1986), this yields six sequences to be used for comparative analysis. The sequence of the untranslated leader is strongly conserved in group I. No substitutions were found in JP501, and the differences with fr and M12 are shown in Figure 2A. There is support for arm E, and to a lesser extent for the S arm. The LDI is consistent with the fr sequence, although there is only one classical covariation (one A-U pair in MS2 is G-C in fr). Sequence differences in the W arm are inconclusive but not inconsistent with the proposed structure.

A feature often encountered in comparative analysis of phage RNA structure is the movement of mismatches through a helical segment. An example is the S arm of M12 and MS2, where the bulge has shifted one position (see also Skripkin et al., 1990 and Olsthoorn et al., 1994).

In Figure 2B we present the structure model for the leader of phage KU1 in comparison with the changes found in the related phage GA. There is support for the arms E and S, whereas the evidence for the LDI is similar to what was found for group I; one straight covaria-

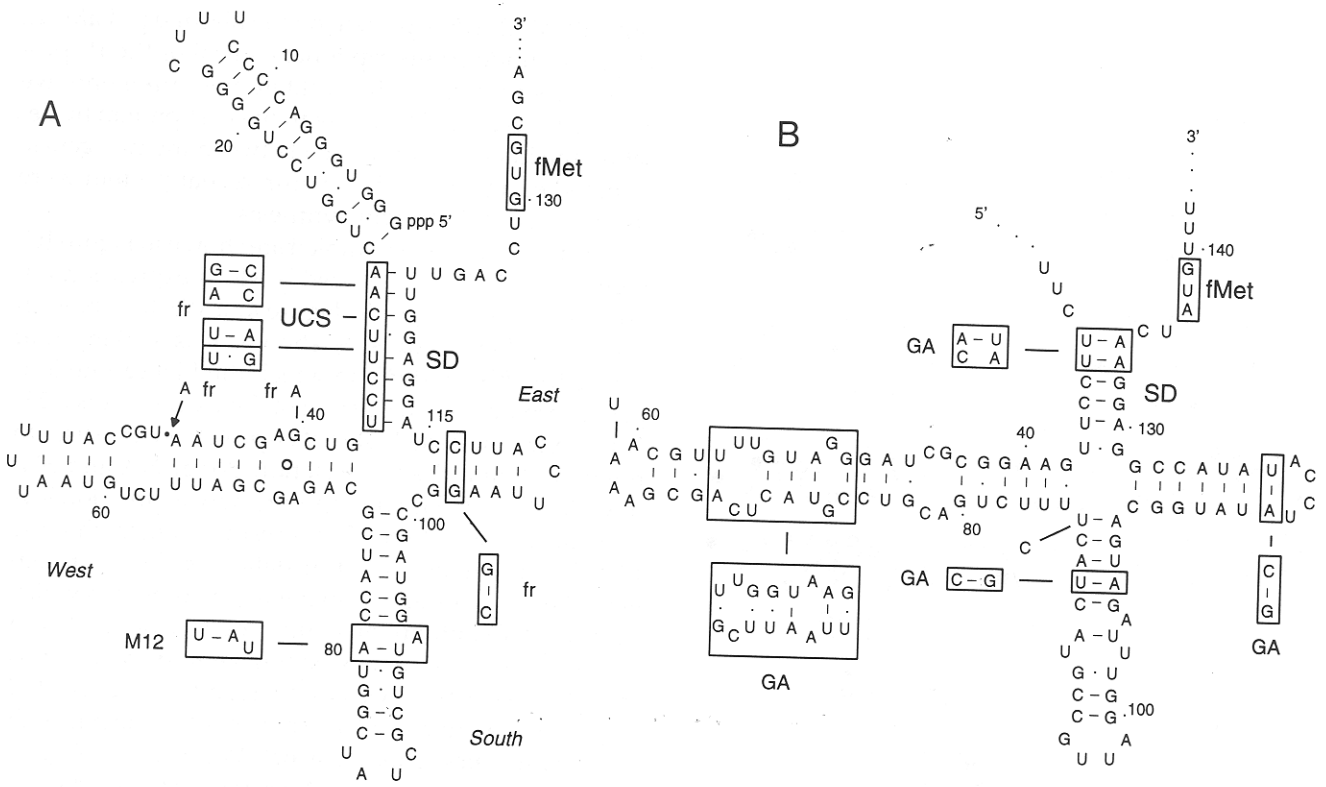

FIGURE 2. A: Proposed RNA structure of the 5′ untranslated leader of MS2 RNA. The start codon at position 130, the UCS at positions 29–36, and the base differences with the group I phages fr and M12 are boxed. **B:** Proposed RNA structure of the 5′ untranslated leader of group II phage KU1, showing differences with group II phage GA. Additional substitutions in GA are found at positions 61, 83, and 86.

tion, combined with a U-A to C·A change. Fortunately, comparing the complete structures for group I with those for group II yields more decisive evidence. Similar cloverleaves can be drawn for both groups. The E and S arms contain convincing comparative evidence, and this also holds for the LDI region. The sequences of the W arm diverge quite a bit. Although the helices for MS2 and KU1 look very similar, strict phylogenetic support is absent.

Evidence from chemical and enzymatic probing

The structure of MS2 RNA in the region under scrutiny here was also probed with dimethylsulfate (DMS), diethylpyrocarbonate (DEP), and 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-*p*-toluene sulfonate (CMCT) and with RNases T1, T2, and V1. The results of these experiments are summarized in Figure 3. They support our model. Chemical modifications and T1 and T2 cleavages occur mostly at loop positions. There is also modification of the bulge nucleotide in the S arm.

The A residues of the G·A;A·G double pair in the W arm are modified by DMS (reactive with A N-1) but not by DEP (reactive with A N-7). Such a G·A;A·G double pair is rather stable (Santa Lucia et al., 1990) and NMR studies have shown that the conformation of G·A in the double pair and in the GNRA tetraloop is similar, with hydrogen bonds from the G exocyclic amino proton to the A N-7, and from the A exocyclic aminoproton to the G N-3 (Heus & Pardi, 1991; Santa Lucia & Turner, 1993). Our modification pattern, in particular modification by DMS and absence of sensitivity to DEP is consistent with the presence of the G·A;A·G double pair in the W arm.

Furthermore, the stem regions appear not sensitive to modification or cleavage by T1 or T2 except for two weak DMS hits at the boundary of the LDI. As expected, there are T2 cuts in the loops in W, S, and E arms. Some of these occur in the stem close to the loops, presumably due to breathing of the weak A-U pairs. There are two unexpected observations. One is the absence of reactivity at the internal loop of 6 nt in the W arm; this region has one cut by RNase V1, which suggests stacking of the nucleotide in the internal loop. The other is an RNase V1 cut in the loop of the S arm. This may be indicative of a tertiary interaction.

Computer and literature predictions

Neither the literature (Fiers et al., 1975; Iserentant & Fiers, 1979) nor the older computer programs such as

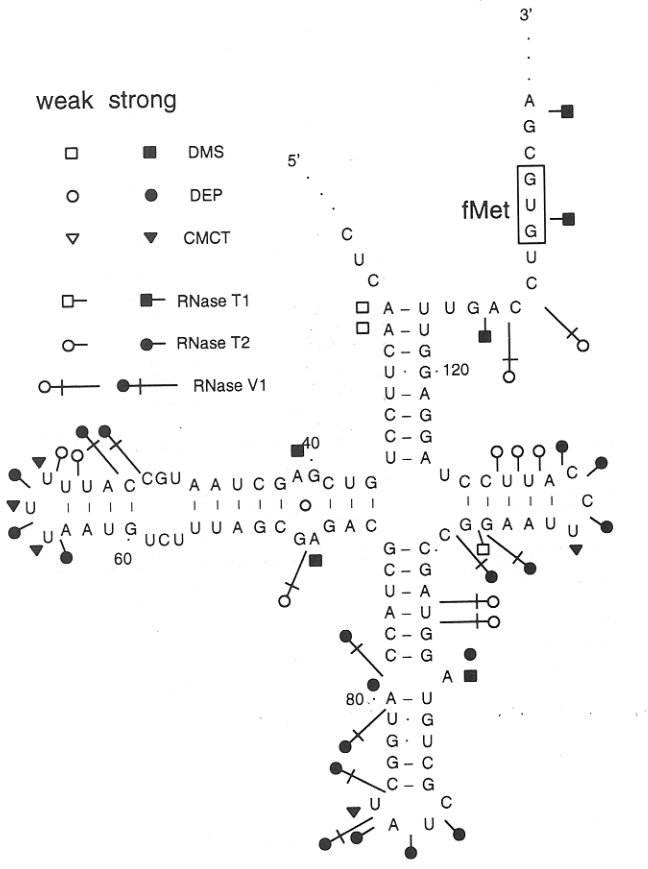

FIGURE 3. Structure-probing results of the first 135 nt of MS2 RNA. Filled symbols represent strong hits; open symbols weak hits. DMS (dimethylsulfate), DEP (diethylpyrocarbonate), and CMCT (1-cyclohexyl-3-[2-morpholinoethyl] carbodiimide metho-*p*-toluene sulfonate).

PC-Fold and GCG-Fold predict the structure presented in Figure 2A. However, M-Fold (GCG software, Genetics Computer Group, Madison, Wisconsin), which contains an updated set of energy parameters, does predict the structure as we present it. We conclude that our model is supported by phylogenetic evidence, structure probing, and computer prediction.

Relation between RNA structure and translation of the A-protein gene

Deletion studies

We and others have previously shown that intramolecular base pairing of the start region of a gene can have profound negative effects on the frequency of translational initiation (Hall et al., 1982; Gold, 1988; de Smit & van Duin, 1990a, 1990b, and references therein). It therefore seemed reasonable to assume that the LDI formed by base pairing of the UCS with the SD region down-regulates A-protein synthesis. To substantiate this assumption we made several deletions involving different parts of the various arms and assayed their effect on A-protein synthesis. The studies were carried out with MS2 cDNA present on the plasmid pPLaMS22 (Fig. 1B), under transcriptional control of the P_L promoter from phage λ. To facilitate measurements we made a translational fusion with the coat protein by deleting the intercistronic region between the two genes. Antibodies against SDS-denatured coat protein were used to detect A-protein synthesis.

The deletions that we studied are shown in Figure 1C, together with their measured levels of expression relative to the wild type. In Figure 4, a Western blot, illustrating these results, is shown. It is striking that clones BA, BN, and PS, in which the LDI can no longer be formed (due to deletion of the UCS sequence 29–36), display increased rates of A-protein synthesis. On the other hand, the two deletion clones NA and SN, in which the UCS is still present show expression levels similar to wild type. These results suggest that the LDI down-regulates the translation of the A-protein gene. We also carried out an initial experiment in which the potential contribution of the sequence of the W arm was assessed. The *Sac* I–*Nco* I fragment, present in this stem-loop structure, was exchanged for an unrelated sequence, capable of forming a regular hairpin with seven base pairs. This construct, W (Figs. 1C, 5), produced the same amount of A-protein as the wild type.

Targeted substitutions

More reliable information on the negative influence of the UCS can be obtained by introducing specific substitutions, that decrease its potential for base pairing with the SD region. Because one side of the interaction is formed by the SD sequence, only mutations in the UCS sequence are interpretable. An example is clone MM (Fig. 6A), in which multiple mutations destabilize the LDI and indeed result in an increase in expres-

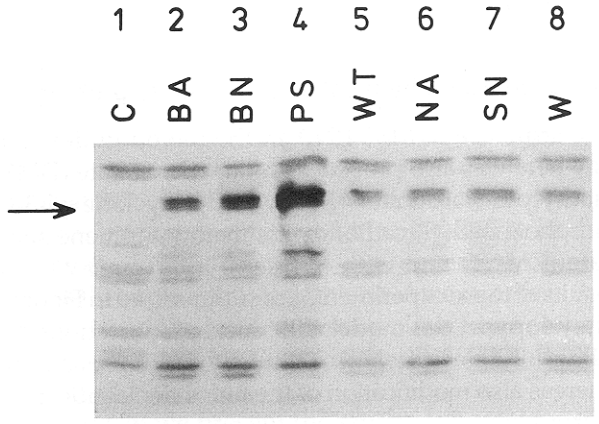

FIGURE 4. Western blot showing the differential synthesis of the A-coat fusion protein (indicated by an arrow) in the deletion mutants. The average relative values, determined from three to five independent experiments by comparing serial dilutions are given in Figure 1. C is a control containing the empty vector.

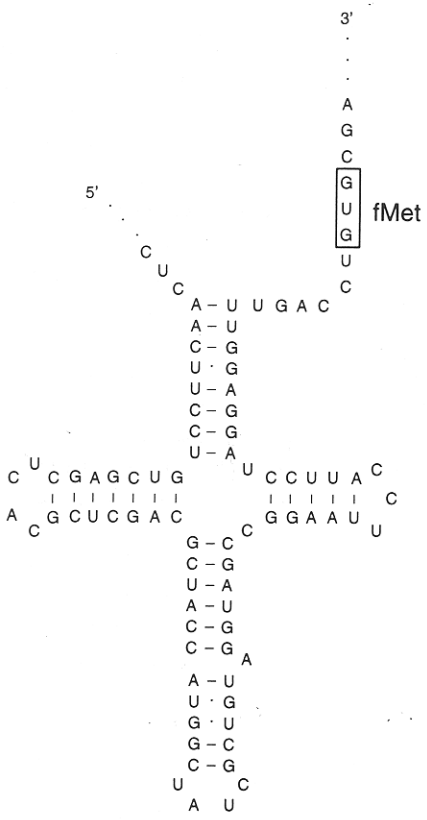

FIGURE 5. Proposed RNA structure of mutant W.

sion level by fivefold. More specific base substitutions are represented by the clones A30U and CC3435AA (Fig. 6A), carrying one and two transversions, respectively, in the UCS. These changes also lead to about a fivefold increase in A-protein yield (Table 1). We conclude that the UCS pairs with the SD region and reduces translation of the maturation-protein gene.

We checked the mRNA stability in most of the mutants used in this study, by primer extension on total

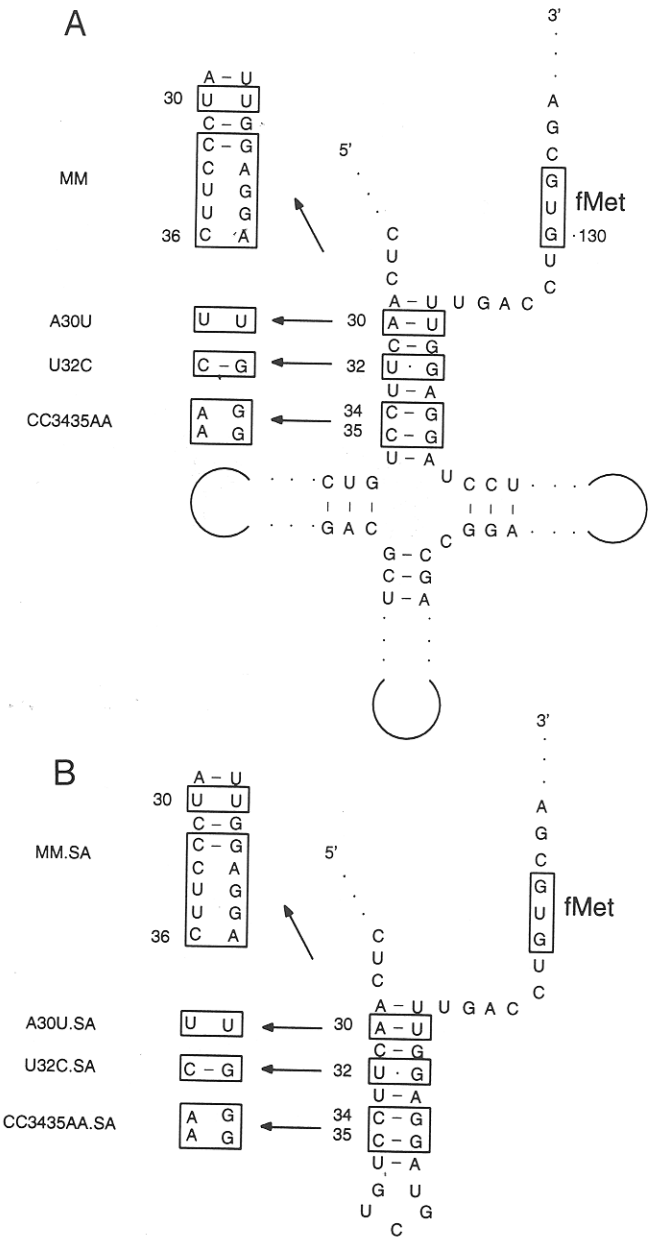

FIGURE 6. Substitutions present in clones MM, A30U, U32C, and CC3435AA. **A:** Derivatives of the wild-type leader. **B:** Derivatives of the single hairpin leader (clone SA). Expression levels of these clones are shown in Table 1. As a consequence of the cloning procedure, the sequence of the loop of mutants MM.SA and CC3435AA.SA is 5′ GUCGCU 3′.

TABLE 1. Average change in A-protein expression resulting from mutations in the leader, in wild-type leader, and in clone SA (shown in Fig. 6B).

Mutation	Relative expression
Wild type	1
MM	5
A30U	5
U32C	1
CC3435AA	5
SA	0.1
MM.SA	1
A30U.SA	0.5
U32C.SA	Not detectable
CC3435AA.SA	3

RNA isolates. The mobilities of the full-length transcripts corresponded to the predicted sizes, whereas the amount of RNA was about the same in all of the clones. No significant variation in the band pattern resulting from premature stops of the reverse transcriptase reaction could be detected (results not shown). This suggests that the differences in A-protein levels derive from changes in the initiation frequency rather than from differential degradation of the transcripts.

Role of kinetics of RNA folding in A-protein expression

Although the evidence presented above strongly suggests that the LDI negatively affects A-protein synthesis, there are several observations that lead us to believe that this is not the whole story. For instance, if we suppose that the cloverleaf is formed very rapidly during MS2 RNA plus-strand synthesis, then the ribosome will always be confronted with the complete structure and initiation will essentially be determined by the breathing of the equilibrated structure, as outlined by de Smit and van Duin (1990a). If, on the other hand, formation of the LDI is relatively slow, then there exists a time window during which the translational start site is freely accessible to ribosomes. In this scenario most, if not all, of the A-protein would be synthesized on nascent chains in which the final folding has not yet taken place.

There are several indications in support of this idea. For example, in mutant CC3435AA the stability of the LDI can be calculated to decrease by 5 kcal/mol (Freier et al., 1986). It is known that in structure-restricted ribosome-binding sites translation goes up (down) 10-fold per 1.5 kcal/mol of structure destabilization (stabilization) (de Smit & van Duin, 1990a, 1994). Accordingly, for mutant CC3435AA, we expect an increase in A-protein yield by several orders of magnitude, but this is not found. Furthermore, when the LDI is further stabilized by converting the U·G pair into C-G, as realized in mutant U32C (Fig. 6A), this stabilization of 3 kcal/mol is expected to decrease the A-protein yield by two orders of magnitude. In fact, this mutation has no effect on A-protein synthesis (Fig. 7, lanes 3, 5; Table 1).

To investigate the possibility that the time required for formation of the LDI is important in determining the A-protein level, we constructed clone SA (Fig. 1C), containing a deletion that results in the replacement of the three hairpins by a 5 nt loop (Fig. 6B). Because the two complementary sequences are now separated by only 5, instead of 80 nt we expect that it will take much less time to form the LDI and that the opportunity of the ribosome to contact the SD sequence will accordingly be decreased. Indeed, this clone makes 10 times less A-protein than wild-type, consistent with the kinetic model (Fig. 7, lane 10; Table 1).

We envisage that in the wild-type situation A-protein is synthesized from nonequilibrated strands only, whereas in clone SA the distance between the complementary regions is so small that the protein is mainly, if not exclusively, translated from the equilibrium structure. If so, the prediction is that stabilization of the LDI in clone SA will further suppress translation, whereas in the wild-type structure it will not. Indeed, the U32C mutation has no effect on expression in the wild-type cloverleaf, but it does further decrease translation in clone SA (Figs. 6B, 8; Table 1). Similarly, destabilizations of the LDI should have a relatively larger effect in SA than in the wild-type construct, containing the three arms. This is best shown for the double mutant CC3435AA, which raises expression about 5-fold in the wild-type cloverleaf but roughly 30-fold in the truncated construct SA (Table 1; Fig. 7).

In a final experiment, the S arm was extended. The *Nco* I site was opened and a 100-nt fragment of M13 was inserted in two different orientations (clones IN1 and IN2, see Fig. 1C). This change is expected to have little influence on the stability of the LDI, but a rather pronounced effect on the kinetics of its formation. As shown in Figure 1C, the insert causes a 5-fold increase

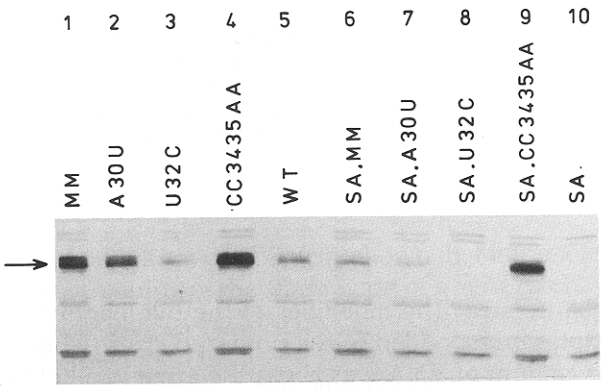

FIGURE 7. Representative Western blot showing effect of substitutions in wild-type leader (lanes 1–5) and single hairpin (SA) leader (lanes 6–10) on A-protein expression (arrow). Expression values were determined by comparing serial dilutions and are given in Table 1. To obtain a realistic estimate, three to five independent experiments were performed and the numbers shown in Table 1 are the averaged values.

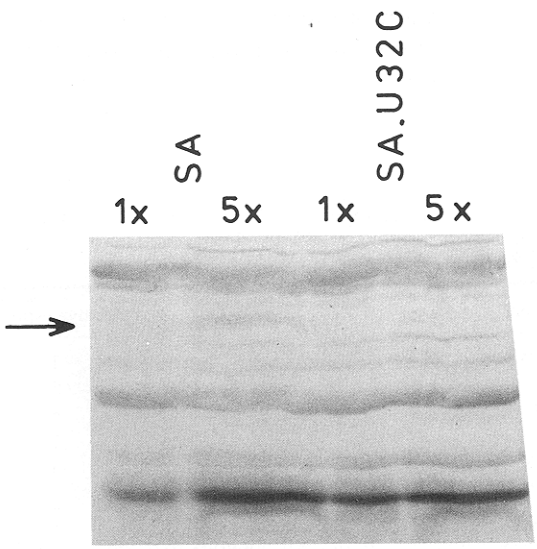

FIGURE 8. Western blot of clones SA and U32C.SA. Lanes 2 and 4 contain five times the amount of cell lysate to visualize the difference in expression between the two clones.

in A-protein for IN1 and a 10-fold increase for clone IN2. This result provides further support for a kinetic component in the control of A-protein synthesis.

DISCUSSION

In this paper we propose a model for the translational control of the A-protein gene in group I phage MS2. As a first step in the analysis, we have determined the secondary structure of the RNA around the start of the cistron. Based on phylogenetic comparison and structure probing we have come to the conclusion that the RNA sequence of interest folds into a cloverleaf, formed by three adjacent local hairpins enclosed by an LDI (Fig. 2A). The crucial element is the LDI. Its 3′ moiety contains the 7-nt-long SD sequence of the A-protein. The 5′ complement is located 80 nt *upstream* and can form an uninterrupted stem structure of eight base pairs. It is important to realize that on nascent chains the SD complement is present before the SD region is produced.

As a second step in the analysis, we have made a series of deletions and measured their capacity to synthesize the A-protein. All deletions that include the upstream part of the LDI produce increased amounts of A-protein, whereas deletions that do not include this region show an unaltered expression. Similarly, single-base substitutions in the UCS that impair its potential to basepair with the SD region (mutants MM, A30U, and CC3435AA) show increased A-protein yields, whereas changes in, for instance, the W arm of the cloverleaf did not show any effect (Fig. 5). These results suggest that the LDI acts as a negative control element by shielding the SD region.

Given these data, there are two ways in which the LDI can shut down the A-protein gene. In the equilibrium model, the time of formation of the interaction is shorter than the time it takes a ribosome to bind the start site. If so, all A-protein synthesis must originate from the breathing of the structure, i.e., the thermodynamic equilibrium between folded and unfolded RNA, as described by de Smit and van Duin (1990a). In this model, the ribosome is confronted with the equilibrated structure. In the kinetic model, the formation of the inhibitory LDI is considered to take more time than binding of the ribosome to the start region. In this view, each nascent strand would allow a ribosome a certain time to initiate. After this time has lapsed, the LDI would form and the cistron would be permanently closed, except for the, probably negligible, contribution by breathing.

We believe that the experimental evidence is in favor of the kinetic model. First, the destabilizations we have introduced in the LDI are predicted by the equilibrium model to lead to a much larger rise in expression than what we actually observe. It could be argued, though, that we have already reached the maximum in A-protein translation after a single destabilizing mutation. Therefore, we also further stabilized the LDI by substituting a U·G for a C-G pair (U32C). The fact that this substitution did not decrease translation argues strongly against the equilibrium model, where such a change is expected to lead to a 10–100-fold reduction in expression. In the kinetic model, the U32C change is not predicted to exert a measurable influence because this mutation will hardly affect the kinetics of formation of the LDI, whereas it will substantially increase its final stability. The validity of this argument is underscored by the fact that the U32C change is indeed able to decrease translation when the role of kinetics is minimized, i.e., in construct SA, where the three arms have been replaced by a 5-nt loop.

There is one more argument that supports the kinetic model. Deleting the three arms is supposed to considerably shorten the time it takes to form the LDI (which now basically forms the stem of a regular hairpin), whereas substantial insertions are expected to increase the time. In line with the expectation, the three-arm deletion (SA) decreases expression 10 times, whereas the insertion of 100 nt in the S arm (clones IN1 and IN2) stimulates translation 5–10-fold.

Although each argument by itself has a certain weakness (one could for instance argue that the insertions distort the LDI, or that a loop of 5 nt would confer a higher stability to the LDI than the three arms do), we think the collective weight of the arguments is in favor of the kinetic model.

Kinetics of ribosome binding is fast enough to capture nonequilibrated RNA

The time of formation of single hairpins is in the microsecond to millisecond range (Tinoco et al., 1990; Fernandez, 1992). The kinetic model we present requires that the initial binding step of 30S subunits to mRNA is at least as fast as the folding of the RNA. Our previous work on the relationship between RNA secondary structure and translational yield shows implicitly that this is true (de Smit & van Duin, 1990a). That analysis makes clear that ribosomes can withdraw mRNA from the helix-coil equilibrium by binding to the unfolded form (Fig. 9), i.e., ribosome binding is quick compared to the folding reaction. In the model system used (the start of the MS2 coat-protein) the ribosome-binding site is contained within a simple stem-loop structure. If anything, the cloverleaf studied here is expected to fold slower than the single hairpin at the start of the coat gene. Therefore, ribosome binding to the not yet equilibrated cloverleaf at the start of the A-protein gene is a realistic possibility.

Additional regulatory elements

So far, we have ignored the fact that the RBS of the A-protein gene of MS2 extends from nt 110 to 145

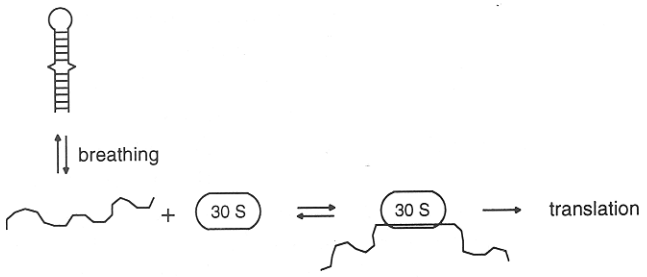

FIGURE 9. Competition between hairpin formation and ribosome binding during translation initiation (after de Smit & van Duin, 1990a).

(Steitz, 1975). The RNA structure around the start codon may therefore also influence translation. According to our structure model (H. Groeneveld & J. van Duin, in prep.) this region forms a weak interaction with a sequence 230 nt downstream. This pairing may further quench what is left of A-protein expression after formation of the LDI with the UCS, which we report here as the major source of translational control.

Comparison with literature data

Robertson and Lodish (1970) reported that the replicative intermediate (RI) fraction isolated (and therefore equilibrated) from phage-infected cells was more active than full-sized phage RNA in directing ribosomes to the A-cistron start. This finding is not readily explained by our model. One suggestion offered by Lodish (1975) was increased A-gene dosage in the RI fraction. An alternative could be RNA breakdown. To explain the data of Robertson and Lodish, Fiers et al. (1975) designed a model in which A-protein translation was controlled by an LDI between the cistron start and a sequence some 700 nt downstream. This interaction is incompatible with our deduced structure and does not stand up in the comparative analysis with the other known sequences of group I and group II phages.

Role for kinetics of RNA folding in other processes

A few studies have been published on the potential role of kinetics of formation of RNA structures in various processes.

One example is duplex formation between the primer RNA of plasmid ColE1 and antisense RNA1, which controls replication (Polisky et al., 1990). Mutations in the primer, outside the region of complementarity and thus not predicted to affect the final base pairing between RNA1 and the primer, still increase replication. The authors propose that the final folding of the primer RNA is reached via a discrete set of intermediates. The primer is suggested to be accessible to RNA1 only in one of the intermediates. The mutations that increase ColE1 replication are suggested to affect the kinetics of the folding process and thus the interaction potential with RNA1.

Another example comes from a study on splicing in HeLa-cell extracts (Eperon et al., 1988). The authors show that a sequence complementary to the splice site inhibits splicing more effectively when the distance between the two complements is reduced. This indicates that the kinetics of RNA folding can play a role in alternative splicings.

Several instances are known in which an LDI controls the translation of mRNA and in all these cases one might expect the speed of RNA folding to play a role in this sort of control (Carter-Muenchau & Wolf, 1989; Petersen, 1989; Wikström et al., 1992; van Himbergen et al., 1993; Yuzawa et al., 1993). Only for the *E. coli gnd* gene, which codes for the enzyme 6-phosphogluconate dehydrogenase, data are available on the effect of variation in the distance between the SD-box and a downstream inhibitory complementary sequence, centered around codon 73 (Carter-Muenchau & Wolf, 1989). Gradually diminishing the distance to 38 codons has no effect, whereas shortening the intervening sequence to 11 codons leads to a 50–150-fold repression of translation. These results fit a kinetic model in which placing the internal complement near the SD shortens the time of formation of the LDI.

Finally, there is a recent study from Ma et al. (1995) in which the effect is analyzed of mutations in a hairpin structure that prevents translation of the IS10 transposase messenger transcribed from fortuitous host promoters. Some of these mutations, though not strongly destabilizing that hairpin, still raise IS10 translation. The authors present evidence that these mutations affect the rate at which the inhibitory stem is formed.

MATERIALS AND METHODS

Escherichia coli strains

K12 strain KA797 (*ara*, Δ(*lac-proAB*), *thi*, Sm^r, F^+) was used for the amplification of the RNA bacteriophages MS2, JP501, M12, and KU1. K12 strain JM101 (*lac*, *pro*, *supE*, *thi/F*, *traD36*, *proAB*, $lacI^q$, *lacZ* Δ*M15*) was used for amplification of the M13 vectors (Messing & Vieira, 1982). Single-stranded DNA for targeted mutagenesis was produced in K12 strain BW313 (*dut*, *ung*, *thi*, *relA*, *spoT1/F'lysA*) (Kunkel, 1985). Expression studies were performed in K12 strain M5219 ($M72\ trpA_{am}$, $lacZ_{am}$, $Sm^r/\lambda dbio_{252}$, cI_{857} Δ*H1*) (Remaut et al., 1981).

Preparation of phage RNA

MS2, JP501, M12, and KU1 were grown in *E. coli* strain KA797. The phage particles were precipitated from the supernatant of the culture by ammonium-sulfate and purified on a CsCl gradient, as described (Voorma et al., 1971). Freon extraction was omitted. The RNA was isolated by phenol ex-

traction, followed by ethanol precipitation, and was stored in 0.1 mM EDTA at −20 °C.

RNA probing

The procedure as described by Christiansen et al. (1990) was followed, with some minor modifications. Primer extension on modified RNA was performed as described by Schmidt et al. (1987).

Sequence of the 5′ regions of M12, JP501, and KU1

The sequence of JP501, M12, and KU1 was determined after construction of cDNA clones as will be described elsewhere (H. Groeneveld & J. van Duin, manuscript in prep.). M12 was a gift of R.N. Konings. JP501 and KU1 were provided by Drs. Furuse and Hirashima. These phages are part of the Watanabe collection of the Keio University.

Construction of the plasmid pPLaMS22 containing the A-protein–coat-protein fusion

The *Sal* I–*Hin*d III fragment of the plasmid pMS2-7 (Devos et al., 1979), containing MS2-cDNA nucleotides 1,365–3,569, was used to replace the *Sal* I–*Hin*d III fragment present in the polylinker of M13mp19 (Messing & Vieira, 1982). An *Xba* I–*Bam* HI fragment of this new M13 construct—the *Xba* I site is present in the M13 polylinker—was cloned in place of the *Xba* I–*Bam* HI fragment of pPLc236/2, a plasmid derived from pPLc236 (Remaut et al., 1981), containing the MS2-cDNA *Xba* I(1305)–*Bam* HI(2057) fragment behind the P_L promoter. In this way the last codon of the maturation gene was fused in frame to the 11th codon of the coat gene. The fusion-point of the two reading-frames was sequenced. Subsequently, the clone was extended to include MS2 nucleotides 1–103. The MS2 cDNA (1–2,057) was placed in pPLa2311 (Remaut et al., 1981) yielding the plasmid pPLaMS22, used in the expression studies (Fig. 1B).

Construction of the mutants

To facilitate construction of the deletion mutants, position A45 in the MS2 sequence was changed into a C, to create an extra *Sac* I site. This change has no effect on translation of the A-protein gene.

Mutants BA, BN, PS, NA, SN, and SA resulted from the deletion of the *Bam* HI–*Avr* II, *Bam* HI–*Nco* I, *Pst* I–Sac I, *Nco* I–*Avr* II, *Sac* I–*Nco* I, and *Sac* I–*Avr* II fragments, respectively, after removal of the sticky ends with T4 DNA polymerase. Mutant W was made by replacing the *Sac* I(40)–*Nco* I(78) fragment by a synthetic *Sac* I–*Nco* I fragment, yielding the construct shown in Figure 5. Mutants IN1 and IN2 were made by cloning the M13 *Afl* III fragment (3,620–3,720) in either orientation in the *Nco* I site of pPlaMS22. Mutants MM, A30U, U32C, and CC3435AA were obtained by site-directed mutagenesis on M13mp19 (Kunkel, 1985), carrying the *Bam* HI–*Xba* I fragment of MS2 cDNA. A30U.SA and U32C.SA were made by site-directed mutagenesis on SA, whereas MM.SA and CC3435AA.SA were made in the same way as SA, but were derived from MM and CC3435AA instead of the wild-type pPlaMS22.

Expression studies

Cultures were grown at 28 °C to an OD_{650} of 0.2. Then the P_L-promoter was induced by raising the temperature to 42 °C and samples of the cultures were taken after 45 min. The pelleted cells were resuspended in 50 μL of a buffer containing 40 mM Tris-HCl, pH 8.0, 80 mM EDTA, and 1.6 mM 2-mercaptoethanol, in the presence of 80 μg/mL RNase, 80 μg/mL DNase, and 20 μg/mL lysozyme, and left on ice for 30 min. Subsequently, the "freeze-thaw" procedure (Beremand & Blumenthal, 1979) was applied. After centrifugation the pellet-fraction containing the A-protein–coat protein fusion was dissolved in denaturing buffer (Laemmli, 1970) and the samples were boiled for 5 min. This fractionation removed much of the cells' soluble protein. Following electrophoresis the protein was transferred to nitrocellulose (Towbin et al., 1979). For immunodetection an antiserum directed against SDS-denatured MS2 coat protein was used. Quantitative determination of expression levels was performed by comparing the intensities of the bands of serial dilutions.

ACKNOWLEDGMENTS

We thank M. de Smit and A. Gultyaev for stimulating discussions and R.N. Konings for phage M12. K. Furuse and A. Hirashima generously provided us with related phages JP501 and KU1, which are part of the Watanabe collection at Keio University. This work was in part supported by the H.C.M. program of the E.C. (ERB4050PL921475). Karine Thimon was a summer exchange student from the University of Paris VII, participating in the Erasmus program.

Received December 28, 1994; returned for revision January 13, 1995; revised manuscript received January 23, 1995

REFERENCES

Adhin MR, Avots A, Berzin V, Overbeek GP, van Duin J. 1990. Complete nucleotide sequence of the group I bacteriophage fr. *Biochim Biophys Acta 1050*:104–109.

Adhin MR, van Duin J. 1989. Translational regulation of the lysis gene in RNA bacteriophage fr requires a UUG initiation codon. *Mol Gen Genet 218*:137–142.

Beremand MN, Blumenthal TH. 1979. Overlapping genes in RNA phage: A new protein implicated in lysis. *Cell 18*:257–266.

Berkhout B, Schmidt BF, van Strien A, van Boom J, van Westrenen J, van Duin J. 1987. Lysis gene of MS2 is activated by translation termination at the overlapping coat gene. *J Mol Biol 195*:517–524.

Carter-Muenchau P, Wolf E Jr. 1989. Growth-rate-dependent regulation of 6-phosphogluconate dehydrogenase level mediated by an anti-Shine–Dalgarno sequence located within the *Escherichia coli gnd* structural gene. *Proc Natl Acad Sci USA 86*:1138–1142.

Christiansen J, Egebjerg J, Larsen N, Garrett RA. 1990. Analysis of rRNA structure: Experimental and theoretical considerations. In: Spedding G, ed. *Ribosomes and protein synthesis: A practical approach*. Oxford: IRL Press. pp 229–252.

de Smit MH, van Duin J. 1990a. Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. *Proc Natl Acad Sci USA 87*:7668–7672.

de Smit MH, van Duin J. 1990b. Control of prokaryotic translational initiation by mRNA secondary structure. *Prog Nucleic Acid Res Mol Biol 38*:1–35.

de Smit MH, van Duin J. 1993. Translational initiation at the coat-protein gene of phage MS2: Native upstream RNA relieves inhibition by local secondary structure. *Mol Microbiol 9*:1079–1088.

de Smit MH, van Duin J. 1994. Control of translation by mRNA secondary structure in *Escherichia coli*. A quantitative analysis of literature data. *J Mol Biol 244*:144–150.

Devos R, van Emmelo J, Contreras R, Fiers W. 1979. Construction and characterization of a plasmid containing a nearly full-size DNA copy of bacteriophage MS2 RNA. *J Mol Biol 128*:595–619.

Eperon LP, Graham IR, Griffiths AD, Eperon IC. 1988. Effects of RNA secondary structure on alternative splicing of pre-mRNA: Is folding limited to a region behind the transcribing RNA polymerase? *Cell 54*:393–401.

Fernandez A. 1992. A parallel computation revealing the role of the in vivo environment in shaping the catalytic structure of a mitochondrial RNA transcript. *J Theor Biol 157*:487–503.

Fiers W, Contreras R, Duerinck F, Haegeman C, Iserentant D, Merregaert J, Min Jou W, Molemans F, Raeymakers A, Vandenberghe A, Volckaert C, Isebaert M. 1976. Complete nucleotide sequence of bacteriophage MS2 RNA: Primary and secondary structure of the replicase gene. *Nature (Lond) 260*:500–507.

Fiers W, Contreras R, Duerinck F, Haegeman G, Merregaert J, Min Jou W, Raeymakers A, Volckaert G, Ysebaert M, van de Kerckhove J, Nolf F, van Montagu M. 1975. A-protein gene of bacteriophage MS2. *Nature 256*:273–278.

Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T, Turner DH. 1986. Improved free energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci USA 83*:9373–9377.

Gold L. 1988. Posttranscriptional regulatory mechanisms in *E. coli*. *Annu Rev Biochem 57*:199–233.

Hall MN, Gabay J, Debarbouille M, Schwartz M. 1982. A role for mRNA secondary structure in the control of translation initiation. *Nature 295*:616–618.

Heus HA, Pardi A. 1991. Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science 253*:191–194.

Horiuchi K. 1975. Genetic studies of RNA phages. In: Zinder ND, ed. *RNA phages*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 29–50.

Inokuchi Y, Takahashi R, Hirose T, Inayama S, Jacobson AB, Hirashima A. 1986. The complete nucleotide sequence of group II RNA coliphage GA. *J Biochem 99*:1169–1180.

Iserentant D, Fiers W. 1979. Secondary structure of the 5′ end of bacteriophage MS2 RNA. *Eur J Biochem 102*:595–604.

Kolakofsky D, Weissmann C. 1971. Qβ replicase as a repressor of Qβ RNA-directed protein synthesis. *Biochim Biophys 246*:596–599.

Kunkel TA. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci USA 82*:488–492.

Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond) 227*:680–685.

Lodish HF. 1975. Regulation of in vitro protein synthesis by bacteriophage RNA by RNA tertiary structure. In: Zinder ND, ed. *RNA phages*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 301–318.

Ma CK, Rayner JC, Kolesnikow T, Simons E, Simons RW. 1995. Inhibition of translation by mRNA secondary structure: The rate of structure formation is critical. *Mol Microbiol*. Forthcoming.

Messing J, Vieira J. 1982. A new pair of M13 vectors for selecting either DNA strand of double digested restriction fragment. *Gene 19*:269–276.

Olsthoorn RCL, Licis N, van Duin J. 1994. Leeway and constraints in the forced evolution of a regulatory RNA helix. *EMBO J 13*:2660–2668.

Petersen C. 1989. Long range translational coupling in the rplJL-rpoBC operon of *Escherichia coli*. *J Mol Biol 206*:323–332.

Polisky B, Zhang XJ, Fitzwater T. 1990. Mutations affecting primer RNA interaction with the replication repressor RNA I in plasmid ColE1: Potential RNA folding pathway mutants. *EMBO J 9*:295–304.

Remaut E, Stanssens P, Fiers W. 1981. Plasmid vectors for high efficiency expression controlled by the P_L promoter of coliphage lambda. *Gene 15*:81–93.

Robertson HD, Lodish HF. 1970. Messenger characteristics of nascent bacteriophage RNA. *Proc Natl Acad Sci USA 67*:710–716.

Santa Lucia J Jr, Kierzek R, Turner DH. 1990. Effects of GA mismatches on the structure and thermodynamics of RNA internal loops. *Biochemistry 29*:8813–8819.

Santa Lucia J Jr, Turner DH. 1993. Structure of $(rGGCGAGCC)_2$ in solution from NMR and restrained molecular dynamics. *Biochemistry 32*:12612–12623.

Schmidt BF, Berkhout B, Overbeek GP, van Strien A, van Duin J. 1987. Determination of the RNA secondary structure that regulates lysis expression in bacteriophage MS2. *J Mol Biol 195*:505–516.

Skripkin EA, Adhin MR, van Duin J. 1990. Secondary structure of the central region of bacteriophage MS2 RNA. *J Mol Biol 211*:447–463.

Steitz JA. 1975. Ribosome recognition of initiator regions in the RNA bacteriophage genome. In: Zinder ND, ed. *RNA phages*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 319–352.

Tinoco I Jr, Puglisi JD, Wyatt JR. 1990. RNA folding. In: Eckstein F, Lilley DMJ, eds. *Nucleic acids and molecular biology, vol 4*. Berlin: Springer Verlag. pp 205–226.

Towbin H, Staehelin T, Gordon J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc Natl Acad Sci USA 76*:4350–4354.

van Duin J. 1988. Single-stranded RNA bacteriophages. In: Calendar R, ed. *The bacteriophages, vol 1*. New York: Plenum Press. pp 117–167.

van Himbergen J, van Geffen B, van Duin J. 1993. Translational control by a long range RNA–RNA interaction; a basepair substitution analysis. *Nucleic Acids Res 21*:1713–1717.

Voorma HO, Benne R, den Hartog TJA. 1971. Binding of aminoacyl-tRNA to ribosomes programmed with bacteriophage MS2. *Eur J Biochem 18*:451–462.

Wikström PM, Lind LK, Berg DE, Björk GR. 1992. Importance of mRNA folding and startcodon accessibility in the expression of genes in a ribosomal protein operon of *Escherichia coli*. *J Mol Biol 224*:949–966.

Witherell GW, Gott JM, Uhlenbeck OC. 1991. Specific interaction between RNA phage coat protein and RNA. *Prog Nucleic Acid Res Mol Biol 40*:185–220.

Yuzawa H, Nagai H, Mori H, Yura T. 1993. Heat induction of σ^{32} synthesis mediated by mRNA secondary structure: A primary step of the heat shock response in *Escherichia coli*. *Nucleic Acids Res 21*:5449–5455.