Role of the Substrate Conformation and of the S1 Protein in the Cleavage Efficiency of the T4 Endoribonuclease RegB*

Received for publication, November 27, 2000
Published, JBC Papers in Press, December 15, 2000, DOI 10.1074/jbc.M010680200

Isabelle Lebars‡, Rouh-Mei Hu§‡§, Jean-Yves Lallemand‡, Marc Uzan§, and François Bontems‡‡
From the ‡Groupe de Resonance Magnétique Nucléaire, Laboratoire Département de Chimie, Synthèse Organique, Ecole Polytechnique, 91128 Palaiseau, France and §Institut Jaques Monod, 4 place Jussieu, 75005 Paris, France

The T4 endoribonuclease RegB is involved in the inactivation of the phage early messengers. It acts specifically in the middle of GGAG sequences found in early messenger intergenic regions but not GGAG sequences located in coding sequences or in late messengers. In vitro RegB activity is very low but is enhanced by a factor up to 100 by the ribosomal protein S1. In the absence of clear sequence motif distinguishing substrate and non-substrate GGAG-containing RNAs, we postulated the existence of a structural determinant. To test this hypothesis, we correlated the structure, probed by NMR spectroscopy, with the cleavage propensity of short RNA molecules derived from an artificial substrate. A kinetic analysis of the cleavage was performed in the presence and absence of S1. In the absence of S1, RegB efficiently hydrolyses substrates in which the last G of the GGAG motif is located in a short stem between two loops. Both strengthening and weakening of this structure strongly decrease the cleavage rate, indicating that this structure constitutes a positive cleavage determinant. Based on our results and those of others, we speculate that S1 favors the formation of the structure recognized by RegB and can thus be considered a "presentation protein."

The control of the pattern of gene expression is one of the main means by which a prokaryotic or eukaryotic cell reacts to a variation of its environment. The expression of a gene depends strongly on its messenger concentration, which in turn is determined by the transcription efficiency, but also by the mRNA decay rate. Messengers, indeed, may have very different degradation susceptibilities. The half-lives of Escherichia coli mRNAs, for example, range between 0.5 and 20 min and may change in response to various factors (1). This could be very important for short-lived messengers, in particular, since small variations of their half-lives will temporarily induce very strong modifications in their abundance (2).

The factors controlling mRNA half-lives are far from being completely understood and are certainly different in prokaryotic and eukaryotic organisms. In the former, in particular, the coupling of transcription and translation events together with the absence of 5′ → 3′ exonucleases imposes specific constraints on the process (3). However, in all cases, mRNA stability seems under the control not only of cis-acting but also of trans-acting factors (4). Indeed, the importance of RNA-binding proteins in the control of mRNA decay rates is becoming more and more evident. It has recently been shown, for example, that the stability of the E. coli ompA mRNA is modulated by its interactions with the host factor I (HfqI) protein, whose concentration depends, in turn, on the growth rate of the bacteria (5). Similarly, several families of proteins, such as the poly(A) or AU-rich sequence binding proteins, have been implicated in the control of eukaryotic messenger degradation (2).

In this context, the S1-RegB system appears to constitute a very interesting model. RegB, encoded in the bacteriophage T4 genome, is involved in the specific degradation of phage early messengers, including its own. Indeed, after infection by a regB- strain, the half-life of the early mRNAs is increased by a factor of three, whereas that of the middle and late mRNAs remains unaffected. Several studies have established that this effect is due to differences in the enzyme activity toward the three classes of messengers and not to its inactivation during the middle phase (6). More precisely, it has been shown that RegB specifically cleaves between the G and A of GGAG sequences in the Shine-Dalgarno regions of the early messengers. Surprisingly, a few GGAG motifs carried by early and middle mRNA and all motifs carried by late mRNA escape RegB processing, as also do most tested motifs within coding sequences. This indicates that RegB recognizes not only the GGAG sequence but also another signal, which remains unknown to date. Another striking point is that the RegB activity determined in vitro is very low in comparison with what can be inferred from the in vivo observations. Considering the specificity of the enzyme, it had been postulated that the ribosome might play a role (7). This has been verified. In vitro it was demonstrated that RegB activity can be accelerated by a factor of up to 100 (depending on the substrate) by the addition of the ribosomal protein S1 to the reaction mix (8). S1 is the largest ribosomal protein (556 amino acids). It is composed of six repetitions of a conserved domain, called the S1 motif, known to be an RNA/DNA binding motif also found in many prokaryotic and eukaryotic proteins involved in RNA metabolism. Its presence is strictly required to initiate the translation of most mRNAs (9), but its precise mechanism of action remains unknown. In our system, it seemed reasonable to postulate an interaction between S1 and the RNA, but this has not been demonstrated. In addition, the nature of S1 action has not been elucidated. It was possible to imagine a nonspecific role. S1, indeed, has been reported to be an unwinding protein (10), but this model was not completely satisfying. In particular, it did not explain the rate enhancement observed on small unstructured RNAs (8). On the contrary, we could imagine that S1 participates in enzyme specificity by selecting several targets.

In this study, we have undertaken the identification of the
molecular basis of RegB selectivity and the definition of the nature of the role of S1. In the absence of any clear sequence motif outside the GGAG, we speculate that RegB activity might be dependent on an RNA structural determinant. To test this hypothesis, we tried to correlate the structure of a series of RNA molecules, probed by NMR spectroscopy, with their cleavage susceptibility in the presence and absence of S1. Several simple substrates of RegB were obtained by the SELEX method (11). We have chosen one of them, possessing two GGAG sites, one reported to be cleaved and the other not. From a comparison of the properties of these two sites in several variants of this molecule, we propose a model of the interactions between the RNA, S1, and RegB.

MATERIALS AND METHODS

RegB and S1 Production—The S1 protein was purified from an E. coli MRE600 strain by affinity chromatography on a poly(U)-Sepharose 4B (Amersham Pharmacia Biotech) column as described (12). It was stored at 4 °C in 10 mmol.liter\(^{-1}\) Tris-Cl (pH 7.4), 6 mmol.liter\(^{-1}\) magnesium acetate, 0.1 mmol.liter\(^{-1}\) EDTA, 1 mol.liter\(^{-1}\) NH\(_4\)Cl, and 7.2 mmol.liter\(^{-1}\) β-mercaptoethanol.

The RegB gene has been introduced in a pET-7b vector (6). Its toxicity prevents the transformation of any E. coli strain possessing the T7 polymerase gene. A JM101 strain was therefore used, and RegB production was induced by infecting the culture (at around 0.7 A\(_{600}\)) with a ACE6 phage (multiplicity of 10) carrying the T7 polymerase gene (13). The protein was purified on a nickel nitrilotriacetic acid (Qiagen) column. Strikingly, its complete elution required a very high (500 mmol.liter\(^{-1}\)) imidazol concentration. After dialysis, the protein was stored at \(-20 \text{°C}\) in 20 mmol.liter\(^{-1}\) Tris-HCl (pH 8.0), 200 mmol.liter\(^{-1}\) NaCl, 0.2 mmol.liter\(^{-1}\) dithiothreitol, and 50% glycerol.

RNA Production—DNA oligonucleotide templates were purchased from Eurogentec. Phenoxyacetyl β-RNA phosphoramidite were bought from Amersham Pharmacia Biotech. RNA chemical synthesis was performed on an Amersham Pharmacia Biotech Gene Assembler Plus apparatus.

Most RNAs were produced using in vitro transcription according to the method developed by Uhlenbeck and co-workers (14). In addition, three (the S22cug, S22stem, and S22linear molecules, see under “Results”) were produced by chemical synthesis (15).

Transcription—Transcription reactions were performed in 40 mmol.liter\(^{-1}\) Tris-HCl (pH 8.0), 1 mmol.liter\(^{-1}\) spermidine, 5 mmol.liter\(^{-1}\) FAD, 0.01% Triton X-100, 80 mg.ml\(^{-1}\) 10% polyethylene glycol, 4–16 mmol.liter\(^{-1}\) MgCl\(_2\), 2 mmol.liter\(^{-1}\) each nucleotide, 400 mmol.liter\(^{-1}\) template, and 0.01 mg.ml\(^{-1}\) T7 RNA polymerase. Before large scale production, all transcriptions were tested on small volumes, and the products were analyzed by gel electrophoresis to check the quality of the matrices and to optimize the conditions (the MgCl\(_2\) concentration, in particular). In all cases, two major species were observed, one corresponding to the awaited product, the other having a much higher molecular weight. Similar observations have been reported by several authors and interpreted as the result of an RNA-degradation by RNase A-like enzymes. In parallel with each reaction, a partial digestion of the substrate, was always added to obtain a measurement of the product lengths.

The kinetics were generally followed for 2 h. Aliquots (5 μl) were typically taken at 2, 4, 6, 8, 10, 15, 20, 25, 30, 40, 50, 60, and 120 min. They were mixed with 3 μl of loading blue (8 mol.liter\(^{-1}\) urea, 20 mmol.liter\(^{-1}\) EDTA, 10% glycerol, and 0.05% bromophenol blue), heated 2 min at 98 °C, and put on ice. The substrates and products were separated by gel electrophoresis (18% polyacrylamide, 8 mol.liter\(^{-1}\) urea) and detected by autoradiography. The band intensities were quantified using a Molecular Dynamics PhosphorImager.

Most reactions were repeated two or three times. In the case of the S12 fragment, the results were so surprising (see under “Results”) that we repeated the whole process (synthesis, purification, and kinetic study twice).

Kinetic Analysis—The percentages of residual substrate and product formed at a given time were evaluated by the ratio of the corresponding band to the sum of all band intensities. Simple phenomenological models were considered to represent the reactions: S → P or S → P\(_1\) + P\(_2\), according to the number (one or two) of cleavage sites. The corresponding equations were fitted to the experimental curves. It was not always possible to obtain a correct adjustment between the model and all data points. In these cases, only the first points were taken into account.

NMR Spectroscopy—All experiments were recorded on a Bruker DRX600 spectrometer equipped with a gradient TXI probe. The spectra were processed off-line using the GIFA software (17), and the data were analyzed using either GIFC (18) or XEASY (19) programs.

H\(_2\)O samples were prepared by dissolving the RNA in 400 μl of 90% H\(_2\)O, 10% D\(_2\)O. The pH was adjusted to about 6.5 by adding a small amount of NaOH. NaN\(_2\) (0.05%) was added to prevent bacterial growth. D\(_2\)O samples were obtained by lyophilizing the corresponding H\(_2\)O samples and dissolved in D\(_2\)O. Two further lyophilizations were generally performed to remove as much H\(_2\)O as possible.

Exchangeable proton assignments were obtained from the analysis of one-dimensional and two-dimensional NOE experiments recorded in H\(_2\)O. Typically, acquisition parameters were 2048 data points over a 12-kHz spectral range, 64–128 scans, and 2-s relaxation time. Generally, 512 increments were recorded in the two-dimensional experiments. The water was always suppressed using a jump and return reading pulse (20). The carrier was positioned at the water frequency, and the delay between the two pulses was adjusted so that the excitation maximum covered the region of the imino protons.

TOCSY (21), DQF-COSY (22), and NOE spectroscopy (23) experiments were also recorded in D\(_2\)O. A MLEV17 sequence with a spin-lock during the relaxation delay.

RESULTS

Choice of the Molecules—Considering the lack of any clear sequence motif distinguishing cleaved and uncleaved GGAG-containing oligonucleotides, we hypothesized that RegB might recognize a structural determinant. To test this, we decided to compare the three-dimensional structure of good and poor substrates. One possibility would have been to study two natural sequences, such as motA (cleaved) and denV ( uncleaved). However, this presented two major difficulties. First, we have no

\(^1\) The abbreviations used are: NOE, nuclear Overhauser enhancement; TOCSY, total correlation spectroscopy; DQF-COSY, double quantum filtered correlation spectroscopy.
information about the relative position and distance between the GGAG and the putative signal. Thus, it would have been difficult to define the limits of the fragment to study. Second, the physiological role of the sequence might require the presence of structural features on its own. Differences between the cleaved and uncleaved sequences would thus have been difficult to interpret.

We therefore took advantage of a published study carried out in the aim of selecting good RegB substrates by the SELEX method (11). Two of the reported molecules (the 22nd and 40th) looked particularly attractive since both possess two GGAGs, designated in the text as the first (red) and the second (blue). The S26so, S22turn, and S22stem molecules (B) were designed to study the effect of structural perturbations on the first cleavage site; S22two, S22twoloop, and S22linear (C) were designed to study the effect of perturbations on the second site. The proposed secondary structures were calculated with the mfold software (24). They have been confirmed by NMR spectroscopy for S22cug, S22gg, S26so, S22stem, and S22linear.

**Structural Analysis**—The RNA structures of five of these molecules, the S22linear, S22cug, S22gg, S26so, and S22stem, (Fig. 1) were probed by NMR spectroscopy. The secondary structure predictions were not sufficient for our purpose; in particular, they are unable to give information concerning possible structures adopted by the loop regions.

The resonances of the imino protons (found between 9.5 and 14 ppm in spectra recorded in H_2O) constitute a good probe of the oligonucleotide secondary structure. Their observation indicates, indeed, that the corresponding protons are involved in hydrogen bonds, generally due to the formation of base pairs. In case of S22linear, a resonance at 12 ppm is observed at 5 °C showing that at least one base pair is formed at this temperature (Fig. 2). However, at 25 °C the spectrum is exempt of sharp peaks, and the only remaining signal present between 10.5 and 11 ppm is characteristic of unpaired imino protons, demonstrating the absence of structure in the conditions of the cleavage reaction (37 °C). In all other cases, resonances are observed in one-dimensional and two-dimensional spectra at both low
The spectra were recorded in H₂O by using a jump and return water suppression sequence (20). A vertical expansion or compression of the display (indicated on the right of the spectra) was applied in several cases. The spectrum recorded on the S22linear molecule at 300 K indicates that this molecule is not structured at this temperature. In contrast, all other spectra are characteristic of the involvement of imino protons in base pairs. The comparison of the spectra at low and high temperatures shows that the secondary structure is conserved at high temperature, even if several lines become very broad due to the increased exchange rate between the imino and solvent protons.

The nucleotides involved in the secondary structure can be identified by means of one-dimensional and two-dimensional NOE spectroscopy experiments recorded in H₂O. It is, in particular, easy to discriminate between the GC and AU base pairs, the latter being characterized by the presence of a strong NOE correlation between the U-imino resonance and the very narrow peak (around 7 ppm) of the A-H² proton. In all cases analyzed, the observed number of GC and AU base pairs was in agreement with the predicted structures.

The same experiments were also used when possible to assign the resonances by analysis of sequential connectivity. The S22stem molecule is predicted to form a long helical stem disrupted at the A position of the GG (Fig. 1). The two strands of the helix are joined by a short GUAA loop, believed to form a GNRA turn. The NOE pattern allows the identification of a segment that matches the predicted stem. In addition, one peak possesses a chemical shift (10 ppm) indicative of a hydrogen bond with an oxygen atom. Such a hydrogen bond is expected in the GNRA turn between the U-imino proton and an oxygen atom. Such a hydrogen bond is expected in the GNRA turn between the U-imino proton and an oxygen of the backbone. In case of S22cug and S22gg, a stem-loop-stem-loop organization is expected. The first stem has a predicted length of 5 (S22cug) or 6 (S22gg) residues, whereas the second is shorter (2 residues). The various spectra of the two molecules are nearly identical, indicating similar structure. Five (of seven) of the S22cug imino resonances were assigned. Their chaining matches that of the first predicted stem. The two remaining peaks present the characteristic features of GC base pairs but are so broad that it was impossible to observe any sequential connectivity. They probably correspond to the second short stem, its size and location between the two loops resulting in low stability and consequently in the observed peak broadening. The predicted structure of S26so is similar to that of S22cug, as are the results. Seven imino resonances (six C, one U) are observed. Their assignment was complicated by the small dispersion of the lines; however, by combining the results of different experiments performed at different temperatures, a chaining was identified corresponding to the first stem. The two remaining lines are very broad and present no sequential connectivity. As in the case of S22cug, we assumed that they correspond to the small GC/CG stem between the internal and terminal loops.

The study of the S22cug molecule was carried further by looking at the nonexchangeable protons (Fig. 3). The analysis of the aromatic-H² correlations on the one hand and of the imino- and amino-aromatic correlations on the other hand confirmed the previously obtained results but could not provide any further assignment. In fact, a close examination of the spectra recorded in D₂O reveals that all nontrivial signals involve the helix residues. This strongly suggests that the two loop regions are not structured. Interestingly, we also observed that the line width of the cytidine and uridine H²-H⁵ scalar correlations (as determined in a TOCSY experiment) varies with respect to the location of the residue (Fig. 4). These protons form a rigid system. Thus, all variations of their line shape reflect a variation of the residue dynamics (25). The five correlations corresponding to the residues of the first helix were assigned. As shown in Fig. 4, their line width increases regularly from the free end of the helix (C1) to the other end (C26). The four other correlations are located in the short helix (C8 and C21) and in the terminal and internal loops (C16 and C23, respectively). These peaks could not be assigned. However, two behave like the C26. The others are very broad, and their observation depends strongly on the temperature and on the spectrometer field (data not shown). The narrow lines correspond to residues in a single environment or, more likely here, to residues with a fast exchange rate (with respect to the chemical shift difference) between two conformations or environments. On the con-
trary, the widening or disappearance of the peaks reflects a slower (fast-intermediate or intermediate) exchange rate. The behavior of the various peaks thus confirms that the loop regions are in equilibrium between several conformations. The peak widening from the free end of the helix to the other end in particular could be interpreted as the result of an increasing perturbation of the corresponding residue environment due to the loops conformational exchange.

Kinetic Analysis of Cleavage—A kinetic analysis of the cleavage reaction has been carried out on all substrates in the presence and absence of S1. To be able to compare the cleavage rates, all experiments were done rigorously in the same conditions.

As an illustration of the quality of the data, the gel autoradiographs of the experiments with the S22, S26so, S22turn, and S22linear molecules are presented in Fig. 5. The different bands are well defined, and their sizes can be unambiguously determined by comparison with the ladder. In a few cases, a smear was observed in the middle region of the gel. Generally, it was very slight and did not interfere with the analysis. However, in case of the S22twoloop, the smear prevented the quantification of the bands. The control lanes always showed a unique spot corresponding to the full-length, indicating that there is neither degradation nor cleavage in absence of RegB. S22cug (Fig. 5A) and S22gg correspond to a fragment of the initial SELEX molecule (11). They possess two GGAG sequences, but the first was thought to be resistant. Given that the molecules are labeled in 5', two bands (corresponding to the full-length and one product) are thus expected. In fact, three
are observed whose lengths are in agreement with cutting at both sites. The unexpected cleavage at the first site was studied with the S26so (Fig. 5B), S22turn (Fig. 5C), and S22stem molecules, which differ by the number of loops. The S26so (having an internal and a terminal loop) and S22turn (in which the terminal loop is replaced by a GNRA tetra loop) are cleaved as expected, whereas the S22stem (further lacking the internal loop) appears to be resistant, even in the presence of S1. Similarly, we tried to analyze the susceptibility of the second site with the S22two, S22twoloop and S22linear series. This led to the most surprising results. S22two, in which the region comprising the first site and the internal loop has been removed, is cleaved by RegB alone but not in the presence of S1. When the site is displaced from the edge to the middle of the loop in the S22twoloop, it becomes fully resistant to cleavage. Finally, the S22linear (Fig. 5D), which corresponds to the loop region of the initial molecule, gives rise to the very low appearance of two products. Strikingly, neither of them has the expected length (8 and 9 nucleotides instead of 10). It seems that the cleavages occur immediately upstream and between the two Gs of the GGAG sequence. All these experiments were repeated two or three times to eliminate any doubt. In the case of S22linear, in particular, the whole process (RNA synthesis and purification, cleavage reaction, and analysis) was done twice.

To analyze the reaction kinetics, we chose to use simple phenomenological models compatible with the amount of data we disposed. We only considered the forward (hydrolysis) reaction because the backward (synthesis) reaction is very unlikely in the presence of water. We also neglected product inhibition. As shown later, this hypothesis is not always justified, although it has no impact on the initial rate determination. For the S26so, S22turn, and S22two molecules, a simple $S \rightarrow P$ scheme, associated to a single rate constant ($k_1$) was adopted. Two parallel reactions, $S \rightarrow P_1 (k_1)$ and $S \rightarrow P_2 (k_1')$, were considered for the S22linear substrate. Finally, for the S22cug and S22gg, we had to make an additional assumption. We assumed that the difference between the two rate constants is sufficient to allow us to consider the unique pathway $P \rightarrow S_1 \rightarrow S_2$ (i.e. to neglect the $P \rightarrow S_2$ direct reaction). In all cases, the model was over-determined, i.e. the number of experimental curves was larger than the number of parameters to be fitted. When considering the kinetic cleavage of S22cug (with and without S1) and S22gg (without S1), the experimental and fitted curves remain in good agreement even after 2 h (Fig. 6 and data not shown), indicating that our assumption concerning the rate differences is correct. However, in all other cases, the time course of the reaction is not correctly described by the model. We were always able to obtain a fit during the first 15 (S26so without S1) to 40 min (S22turn) but observed after that a slowdown of the kinetics. This suggests that the enzyme is inhibited by at least one of the reaction products. In fact, in all cases the beginning of the divergence corresponds to the formation of about 20% product.

All kinetic rate constants are reported in Fig. 7. They have been grouped in two series corresponding to the analysis of the two sites. Strikingly, we observe that the second site (in the terminal loop region) is cleaved much more efficiently than the first (at the junction of the two helices) when we look at the two molecules S22cug and S22gg. In the absence of S1, the rate constant ratio is around 15. It is lower in its presence (about 4) but remains significant. The first site is cleaved with the same efficiency in the S22cug, S22gg and S26so molecules, suggesting that the cleavage of the second site in S22cug and S22gg has only a weak influence on the properties of the first. The replacement of the terminal loop by a GNRA tetra-loop induces a drastic reduction in the rates (25-fold in the absence of S1, 7-fold in its presence). The cleavage seems completely abolished by the further elimination of the internal loop. All this suggests at first sight a strong correlation between the stability of the RNA at or near the GGAG site and the cleavage efficiency. However, the results are more puzzling when considering the second site. The simultaneous elimination of the first site and the internal loop in the S22two molecule induces a reduction in the cleavage rate (by a factor of 4) in the absence of S1 but, more surprisingly, abolishes cleavage in the presence of S1. Furthermore, the cleavage efficiency becomes null or very low when the GGAG site is in the middle of the loop (S22twoloop) or when the fragment is linearized (S22linear). In this latter case, in addition, the enzyme specificity is lost, as mentioned above.

**DISCUSSION**

Since its discovery in 1988 (7), one of the most intriguing and exciting question concerning the T4-encoded endoribonuclease RegB has been the study of its specificity. First, it has been established that the enzyme cut in the middle of a very precise sequence (GGAG). In addition, it has been shown that not all
RNA Structural Determinant Recognized by RegB

GGAG sequences are cleaved in vivo. The enzyme efficiently processes most of the GGAGs found in the intergenic (in particular Shine-Dalgarno) regions of the early T4 messengers (6, 7, 26). On the contrary, it has no activity on most of those located in the Shine-Dalgarno regions of late messengers or in coding sequences (6). By transforming an uninfected cell with two plasmids, one coding for RegB, the other for a T4 natural messenger, it was demonstrated that the pattern of RegB sensitivity observed in vivo can be reproduced in the absence of other T4 factors (6). However, this does not imply that RegB alone is able to distinguish between the two classes of GGAG sequences. It was, indeed, also shown that RegB possesses by itself a very low activity that can be accelerated by a factor up to 100 in the presence of the ribosomal protein S1 (8). In this condition it is not clear which of the two proteins carries the specificity. On the one hand, since S1 is involved in the recognition of the Shine-Dalgarno regions by the ribosome, it has been proposed that it could act as a nonspecific mediator between the RNA and RegB by increasing the enzyme concentration in the vicinity its targets through protein-protein interactions (6, 11, 26). In this way, S1 in this system would play a role similar to that of the K RNA binding domain of the RNase E (27). On the other hand, in the absence of evidence of any direct interaction between S1 and RegB, it is also possible that S1 acts as a presentation protein and modulates the efficiency of the system through specific interactions with the RNA, as recently described in the case of the Hfq1 protein for example (5).

This raises two questions. What are the respective contributions of RegB and S1 in the modulation of cleavage efficiency, and what are the signals recognized by RegB, by S1 or by both proteins on target (or non-target) RNAs?

To address these questions, few results are available in the literature concerning the enzymatic properties of the RegB-S1 system. The efficiency of RegB in the presence or absence of S1 has been evaluated in conditions similar to those used here (10 μmol liter⁻¹ RNA, 0.1 μmol liter⁻¹ RegB, 0.2 μmol liter⁻¹ S1) on two model substrates: a 112-nucleotide T7 transcript and a small decanucleotide CUUUGAGGG (8) hereafter designated JR. Interestingly, the long transcript possessed two GGAGs, one of which (located in the Shine-Dalgarno region) is cleaved. In the absence of S1, very low turnover was measured (1 fmol min⁻¹ for the decanucleotide and 1.8 fmol min⁻¹ for the long transcript, corresponding to less than 1% of the molecules cut in 60 min). This turnover is accelerated in the presence of S1 by a factor of 70–80 (80 fmol min⁻¹ for the JR decanucleotide and 120 fmol min⁻¹ for the long transcript, i.e., about 25% cleaved in 60 min). It must be pointed out that the original paper (8) is lacking one experimental parameter (the total amount of enzyme used or, equivalently, the reaction volume) preventing the conversion of turnover values to rate constants. Fortunately, these parameters could be found in the Ph.D. thesis of J. Ruckman (28), leading to the rate estimation reported in Fig. 7. Subsequently, the SELEX method was used to obtain RNA molecules that are specifically processed by RegB in the presence of S1 (11). Forty-eight sequences were reported. All but one possess a GGAG motif. In 33% of them, the GGAG is part of a GGAGGA strong Shine-Dalgarno sequence, and most (but not all) are followed by a long stretch of A or C nucleotides. Strikingly, nearly all GGAG sequences are found at the 5′ extremity of the variable region. Several of these clones have been further tested for their cleavage susceptibility in the presence or absence of S1 in conditions rather different from those used in the first study or by ourselves (4 μmol liter⁻¹ RNA, 0.1 μmol liter⁻¹ RegB, 0.8 μmol liter⁻¹ S1). Two classes of molecules could be identified. The members of the first are not appreciably cleaved by RegB alone but are very well processed in the presence of S1 (around 80% in 10 min), whereas those of the second are cleaved at the same low rate (less than 5% in 10 min) in the presence or absence of S1. A partial correlation was found between the AC richness and the S1 dependence of the rate. However, none of the observations leads to a clear signal identifying the substrates.

Compared with these data, our results raise a series of questions. First, both GGAGs of the S22cug and S22gg molecules are cleaved in our experiments, whereas only the second was found to be processed by the authors of the SELEX study. However, it must be noticed that our work uses neither the same molecules nor the same experimental conditions. The original clone 22 is circular and longer than our fragments (75 residues instead of 32). The RNA/RegB and S1/RegB ratios are higher in the SELEX procedure than in our experiments. More puzzling are the differences in cleavage efficiency.

Looking at the effect of RegB alone (in the absence of S1), the results obtained on the first site indicate a strong correlation between the secondary structure of the molecule and cleavage susceptibility. It has been suggested that the presence of stable secondary structure at or near the scissile bond would constitute a major negative determinant of RegB efficiency, which would be maximal on unstructured fragments (28). At first sight, our results seem to support this idea. Indeed, the differences between the S22cug, S22gg and S26so (which have similar rate constants) are not thought to deeply perturb the stability of the two stems flanking the first GGAG. In contrast, the replacement of the terminal loop by a GNRA turn (in the S22turn) is likely to stabilize the short stem, as most certainly will the suppression of the internal loop in the S22stem molecule. This hypothesis would also explain why the second site of
the S22gg and S22cug molecules, located in a flexible loop, is cleaved more efficiently than the first. However, the same site placed in the center of the loop (S22twoloop) or in a linear fragment (S22linear) is no longer processed. Moreover, the comparison of our rate constants with those reported by Ruckman et al. (8) indicates that the second site is cleaved in the absence of S1 much more efficiently than “good substrates” in the presence of S1. In these conditions, even if we cannot exclude the possibility that the presence of stable secondary structure impedes cleavage, our results demonstrate the existence of a positive determinant favoring interactions between RegB and S22 molecules. The particular properties of the S22linear fragment, which was designed to conserve the sequence but not the structure surrounding the second site, indicate that this structure contributes strongly to the determinant. In addition, they also suggest that the determinant may play a role not only in the efficiency but also in the specificity of the enzyme. The S22linear, indeed, is not only cleaved with low efficiency but also with wrong specificity, whereas all other molecules including the linear JR decanucleotide are processed at the correct position. A striking difference between the two linear fragments is that ours (GAGAACGGAGCACA) is rich in A and G, whereas the JR (CUUUGGGAGGG) has a single A and several CUs. It is thus tempting to suppose that the enzyme has more difficulty recognizing the correct site in the AG-rich context of the S22linear fragment but is guided either by the sequence, in the JR fragment, or by the structure, in all other S22-derived molecules. The precise nature of the structural determinant is difficult to assess with the elements we dispose of at this stage of our study, but several guesses can be made. In particular, it will be noticed that the first site of S26so and both sites of S22gg and S22cug are better cleaved than the linear fragments and share the same global structural organization. In both, the A of the GGAG is free, the last G is involved in a short (two residue) helix, and this helix is followed by a nonstructured loop region. The main difference is that the two first Gs of the GGAG are also engaged in base pairs in the first site but belong to a loop in the second. The most important parameter seems to be the presence of the short helix involving the last G. The cleavage rate is, indeed, very weak or null in all molecules lacking this element (JR, S22twoloop, S22linear). The elimination of the loop reduces the rate by a factor 2 (second site in S22two versus S22cug) to 10 (first site in S22turn versus S22gg). Its length and precise sequence seem to be of little importance, as indicated by a comparison of the first site cleavage in S22gg or S22cug and S26so. Finally, the comparison of the two sites in the S22gg and S22cug molecules suggests that the involvement of the first two Gs in base pairs induces a diminution of the rate by a factor about 15 (13 for S22cug, 16 for S22gg). It thus seems that a good RegB substrate can be formed by a short helix involving the last G of the GGAG site flanked by two nonstructured regions. However, we cannot exclude the possibility of other motifs.

The second aspect of our study concerns the role of the S1 protein. It is more difficult to analyze because we have only indirect information (the consequence of the presence of S1 on the RegB cleavage rate). Nevertheless, several remarks seem pertinent. A first puzzling point is the very low value of the rate enhancement observed at both sites of the S22cug and S22gg molecules (about 4 and 1.1 for the first and second site, respectively). However, in the presence of S1, even the first site is cleaved as efficiently as the JR molecule. The low enhancement value thus appears as a consequence of the intrinsic high cleavage rate obtained in the presence of RegB alone. Also remarkable is the fact that the GGAGs not cleaved in the absence of S1 (those of the S22stem, S22twoloop, S22linear) are not processed in its presence either. Interestingly, this is also the case for one GGAG of the long T7 transcript studied by...
Acknowledgments—We are very grateful to Drs. F. Dardel and D. Fourmy for help with the RNA transcription and Drs. M. Kochoyan, J.-L. Leroy, and K. Snoussi for help with the chemical preparation of some of the molecules. Drs. J.-B. Créchet and E. Jacquet are acknowledged for their advice on kinetic experiments. We thank Dr. R. d’Ari for very helpful comments on this manuscript.

REFERENCES

1. Coburn, G. A., and Mackie, G. A. (1999) Prog. Nucleic Acid Res. Mol. Biol. 62, 55–108
2. Ross, J. (1996) Trends Genet. 12, 171–175
3. Petersen, C. (1993) Control of Messenger RNA Stability, (Belasco, J. G., and Brawnerlan, G., eds) pp. 117–145, Academic Press, Inc., New York
4. Derrigo, M., Cestelli, A., Savettier, G. and Di Liegro, I. (2000) Int. J. Mol. Med. 11, 111–123
5. Vytvytska, O., Jakobsen, J. S., Balcunaita, G., Andersen, J. S., Baccarini, M., and von Gabain, A. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 14118–14123
6. Sanson, B., Hu, R., Troutskaia, E., Mathy, N., and Uzan, M. (2000) J. Mol. Biol. 297, 1063–1074
7. Uzan, M., Favre, R., and Brody, E. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8895–8899
8. Ruckman, J., Ringquist, S., Brody, E., and Gold, L. (1994) J. Biol. Chem. 269, 26655–26662
9. Serøensen M. A., Fricke, J., and Pedersen, J. (1998) J. Mol. Biol. 280, 561–569
10. Subramanian, A. R. (1983) Prog. Nucleic Acid Res. Mol. Biol. 28, 101–142
11. Jayasaena, V. K., Brown, D., Sitchfield, T., and Gold, L. (1996) Biochemistry 35, 2349–2356
12. Subramanian, A. R., Rienhardt, P., Kimura, M., and Suryanarayana, T. (1981) Eur. J. Biochem. 119, 245–249
13. Studier, F. W., and Moffat, B. A. (1986) J. Mol. Biol. 189, 113–130
14. Milligan, J. F., Grobe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Nucleic Acids Res. 15, 8783–8798
15. Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Werkmann, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. (1995) Nucleic Acid Res. 23, 2677–2684
16. Cazenave, C., and Uhlenbeck, O. C. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 6972–6976
17. Deluc, M.-A. (1989) Maximum Entropy and Bayesian Methods (Skilling, J., ed) pp. 295–299, Kluwer Academic Publishers Group, Dordrecht, Netherlands
18. Rouh, A., Delsuc, M.-A., Bertrand, G., and Lallemand, J.-Y. (1993) J. Magn. Reson. 102, 357–359
19. Bartels, C., Xia, T.-h., Billette, M., Gunter, P., and Wuthrich, K. (1995) J. Biomol. NMR 6, 1–10
20. Plateau, P., and Gueron, M. (1982) J. Am. Chem. Soc. 104, 7310–7311
21. Braunschweiler, L., and Ernst, R. R. (1983) J. Magn. Reson. 53, 521–528
22. Rance, M., Serøensen, O., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wuthrich, K. (1983) Biochem. Biophys. Res. Commun. 117, 479–485
23. Kumar, A., Ernst, R. R., and Wuthrich, K. (1980) Biochem. Biophys. Res. Commun. 95, 1–6
24. Zuker, M., Mathews, D. H., and Turner, D. H. (1999) RNA Biochemistry and Biotechnology (*Baciszewski, J., and Clark, B. C., eds) pp. 11–43, NATO ASI series, Kluwer Academic Publishers Group, Dordrecht, Netherlands
25. Lane, A. N., and Lefevre, J.-F. (1994) Methods Enzymol. 239, 596–619
26. Sanson, B., and Uzan, M. (1993) J. Mol. Biol. 233, 429–446
27. Kaberdin, V. R., Walsh, A. P., Jakobsen, T., McDowall, K. J., and von Gabain, A. (2000) J. Mol. Biol. 301, 257–264
28. Ruckman, J. (1993) Identification and characterization of a sequence-specific ribonuclease encoded by bacteriophage T4. Ph.D. thesis, University of Colorado

Ruckman et al. (8). Finally, although S1 seems unable to convert a nonsubstrate GGAG into a substrate, it can on the contrary completely inhibit RegB activity (S22two).

Given our first conclusion, that RegB efficiently cleaves a GGAG when the RNA molecule possesses a particular structure, it seems plausible to propose that S1 acts by promoting its formation. According to this hypothesis, a substrate will be a molecule able to adopt the correct conformation (even if it is not the stabelst), whereas a nonsubstrate will be unable to reach it. Accordingly, the second site of the S22cug and S22gg molecules is very well processed, and the rate is only marginally enhanced because the molecules already possess a near ideal conformation. On the contrary we may suppose that the S22stem, S22twoloop and S22linear are no longer able to adopt it either because they are engaged in very stable alternative structures (likely the case of S22stem) or because the sequence is no longer suitable (likely the case of S22linear).

In conclusion, our results indicate that the discrimination between a substrate and a nonsubstrate molecule depends on two positive determinants: the presence of GGAG and the possibility to adopt a particular conformation, a short stem between two loops in the present case. The efficiency of RegB alone is controlled by the structure of the molecule and can be much higher than previously thought (the S22gg second site is cleaved 400 times faster than the JR oligonucleotide). S1 probably acts by promoting the appearance of the correct RNA conformation and thus can be considered a presentation protein. In addition, it is not excluded that it could participate in the regulation of the system by discriminating between presentable and nonpresentable substrates or by inhibiting the RegB activity.
