In Vitro Selection of RNA Specifically Cleaved by Bacteriophage T4 RegB Endonuclease†

Vineetha K. Jayasena,‡ David Brown,‡ Timur Shtatland,‡ and Larry Gold*,‡,§

Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, Colorado 80309-0347, and NeXstar Pharmaceuticals, Inc., 2860 Wilderness Place, Suite 200, Boulder, Colorado 80301

Received August 10, 1995; Revised Manuscript Received December 5, 1995⊗

ABSTRACT: T4 RegB endonuclease specifically cleaves at -GGAG- sites in several early T4 messages, rendering them nonfunctional. Not all -GGAG- sites are processed equally by RegB; those found at the Shine−Dalgarno sequences and in intercistronic regions are processed with higher efficiency than the -GGAG- sites located in coding regions. The low activity of RegB observed *in vitro* is enhanced by 1−2 orders of magnitude by the *Escherichia coli* ribosomal protein S1. We have used SELEX (systematic evolution of ligands by exponential enrichment) on a combinatorial RNA library to obtain molecules that are specifically cleaved by T4 RegB endonuclease in the presence of S1. The sequences obtained contain the required -GGAG- tetranucleotide and were unusually enriched in adenosine and cytosine nucleotides. No consensus structure or sequence motif other than -GGAG- was conserved among the selected molecules. The majority of the RNAs are entirely dependent on S1 for RegB-catalyzed cleavage; however, a few RNAs are found to be S1 independent but are cleaved by RegB with much lower rates.

Regulation of gene expression achieved through the control of mRNA stability is now well documented in both eukaryotes and prokaryotes [reviewed in Brawerman (1987), Belasco and Higgins (1988), Deutscher (1988), and Peterson (1992)]. In eukaryotes, terminal structural elements such as 5′ capped structures and poly(A) tails confer stability to the mRNA. In many prokaryotes, transcription termination signals capable of folding into stem−loop structures located in the 3′ noncoding regions have been implicated in protecting mRNA from nonspecific degradation by exonucleases (Hayashi & Hayashi, 1985; Wong & Chang, 1986). For other messages, such as the T4 gene 32 mRNA (Gorski et al., 1985), 5′ noncoding regions encompassing the ribosomal binding site and the first few nucleotides of the open reading frame confer stability, presumably through interactions with protein factors. The stability of some messages may be determined by more than one structural element and may also be regulated by multiple mechanisms (Peterson, 1992).

There are a few known examples of proteins that regulate translation by cleaving a specific sequence or structure within an mRNA. RNase III specifically recognizes an internal stem−loop structure in several messages, including its own, and processes it by an endonucleolytic cleavage (Bradwell et al., 1989; Schmeissner et al., 1984; Portier et al., 1987). This rate-limiting step is followed by rapid degradation of the message, presumably by exonucleolytic attack on the newly exposed 3′ end. In *Escherichia coli*, the major exonucleolytic degradation pathway of mRNA occurs through 3′ to 5′ exonucleases such as RNase II and polynucleotide phosphorylase (Donovan & Kushner, 1986). 5′ exonucleases are yet to be discovered in *Escherichia coli*, although they have been shown to exist in eukaryotic systems (Sterns & Maupin, 1987). Structures that serve as targets for RNase III typically occur in untranslated regions of the message. However, degradation of mRNA initiated by one or more endonucleolytic cleavages in coding regions may also occur as in the case of the *puf*LMX message in *Rhodobacter capsulatus* (Chen & Belasco, 1990; Klug & Cohen, 1990).

In bacteriophage T4, an endonuclease activity that specifically cleaves several early T4 messages, including its own, within the Shine−Dalgarno sequence has been previously described (Uzan et al., 1988). This nuclease recognizes the 5′-GGAG-3′ sequence in the Shine−Delgarno region of these messages and specifically cleaves between the second G and the A in the tetranucleotide sequence. This endonuclease activity is found to be encoded by the T4 *regB* gene (Ruckman et al., 1989). T4 RegB endonuclease not only renders the mRNA directly nonfunctional due to its inability to bind to the ribosome but also shortens the chemical half-life of some of the processed messages (Sanson & Uzan, 1993). The latter function is attributed to RegB-dependent cleavage events occuring in the AU-rich region immediately downstream of the processed -GGAG-, probably carried out by a different endonuclease.

Clearly, the recognition site has to be more complex than merely the tetranucleotide sequence. Although -GGAG- sequences in ribosomal binding sites and in intercistronic regions of polycistronic transcripts are strong processing sites, the same sequence in the coding regions is generally processed with much lower efficiency. In at least one instance, a -GGAG- sequence within the Shine−Dalgarno region of the T4 *denV* mRNA is found not to be cleaved by RegB (Uzan et al., 1988). The reason for this bias is not clear. Purified preparations of RegB have low rates of cleavage *in vitro* compared to its activity *in vivo* (Ruckman et al., 1995). However, the activity of RegB *in vitro* is enhanced by 1−2 orders of magnitude by ribosomal particles.

† This research was supported by research grants from the National Institutes of Health to L.G. (GM28685 and GM19963) and funds from NeXstar Pharmaceuticals, Inc.

* Corresponding author.
‡ University of Colorado.
§ NeXstar Pharmaceuticals, Inc.

⊗ Abstract published in *Advance ACS Abstracts,* February 1, 1996.

0006-2960/96/0435-2349$12.00/0

The stimulatory effect is entirely contributed by S1 protein in the 30S particles (Ruckman et al., 1995). Whether S1 mediates this effect through protein−protein interactions with RegB or by interacting with the RNA is not known. The existence of nonspecific interactions between S1 and nucleic acids is well-established and has been extensively investigated (Bear et al., 1976; Draper et al., 1977; Mulsch et al., 1981; Thomas et al., 1978). Although S1 is able to interact with any DNA or RNA, it binds with higher affinity to oligonucleotides with pyrimidines (Draper et al., 1977).

SELEX[1] (systematic evolution of ligands by exponential enrichment) is a technique that has been successfully used to generate high-affinity ligands to various proteins, peptides and small molecules [reviewed in Gold et al. (1995) and Gold (1995)]. Recently SELEX has been used to generate RNA molecules that specifically undergo autolytic cleavage reactions in the presence of lead ions (Pan & Uhlenbeck, 1992). Here, we have used a similar technique as an approach to understand sequence or structural elements required for RegB processing of RNA and to understand the role of S1 in stimulating this activity. The selection generated many relatively unstructured ligands carrying the expected -GGAG- sequence and a domain rich in A and C nucleotides.

EXPERIMENTAL PROCEDURES

Enzymes. T7 RNA polymerase was kindly provided by NeXstar Pharmaceuticals, Boulder, CO. T4 RNA ligase and T4 polynucleotide kinase were purchased from New England Biolabs. Taq DNA polymerase and AMV reverse transcriptase were obtained from Promega. RNase T1 was purchased from Boehringer Mannheim.

Other Proteins. RegB was used for a stock prepared previously (Ruckman et al., 1995). Purified S1 was generously provided by Alap Subramanian of the Max Planck Institute, Berlin, Germany.

Nucleic Acids. The DNA template for the starting RNA pool, 5′-TTTCGAATTCAAAGCTGTTTAGCTAC[30N]-CAGCATGCTCGACAGGCATCG-3′, was purchased from Operon Technologies, Inc. The 5′ PCR primer, 5′-CCGAAGCTTAATACGACTCACTATAGGGC-GATGCCTGTCGAGCATGCTG-3′, contains the T7 promoter and a *Hin*dIII site. The 3′ PCR primer, 5′-TTTCGAATTCAAAGCTGTTTAGCTAC-3′, carries an *Eco*RI site and was also used to prime cDNA synthesis from the selected RNA. Both PCR primers were synthesized on an Applied Biosystems Model 394 DNA synthesizer by using standard cyanoethyl phosphoramidites and purified by gel electrophoresis under denaturing conditions. The primers hybridize to the DNA template as shown in Figure 1A.

RNA was prepared by *in vitro* transcription of PCR-amplified DNA template containing a contiguous 30-nucleotide-long random region. PCR amplification was carried out in a buffer containing 10 mM Tris-HCl (pH 8.4), 50 mM KCl, 7.5 mM $MgCl_2$, 50 μg/mL BSA), 1 mM dNTPs, 5 μM each primer, and 50 units/mL Taq DNA polymerase in a 400-μL reaction volume. In the first round of selection, 1 nmol from the PCR-amplified synthetic DNA template was used for the *in vitro* transcription of RNA. In the subsequent rounds, approximately one-fifth of the PCR-amplified DNA was transcribed in 40 mM Tris-HCl (pH 8.1), 6 mM $MgCl_2$, 5 mM DTT, 1 mM spermidine, 80 mg/mL PEG 8000, 0.002% Triton X-100, 2 mM NTPs, 50 μCi [α-^{32}P]UTP (800 Ci/mmol), and 1 unit/μL T7 RNA polymerase in a 100-μL reaction volume for 4 h at 37 °C. A 10-fold molar excess of GMP over GTP was included in the reaction in order to generate RNA with predominantly 5′-monophosphate ends suitable for ligation into circular products (Milligan & Uhlenbeck, 1989). The DNA template was destroyed by incubating the reaction mixture with 20 units of DNase I at 37 °C for 15 min. The RNAs were purified on 8% polyacrylamide gels under denaturing conditions.

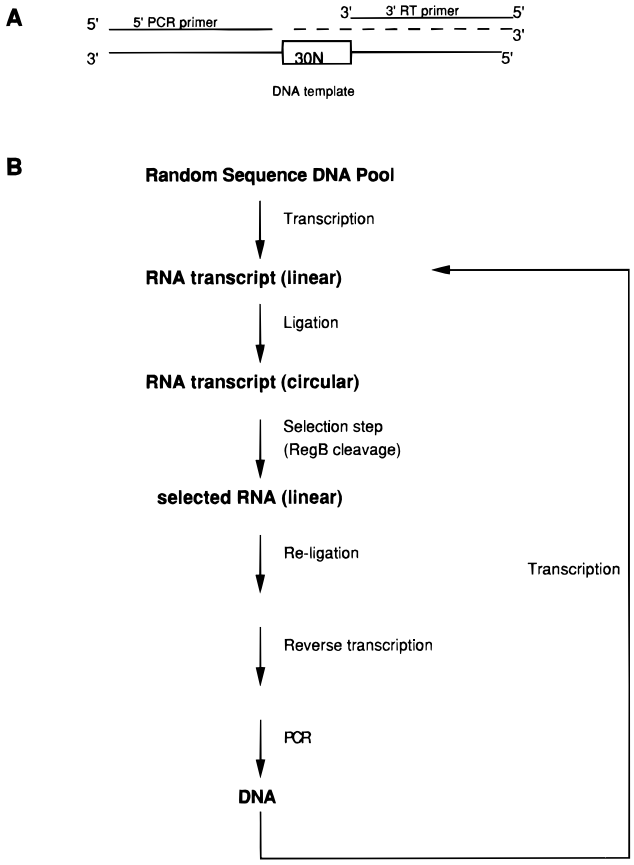

FIGURE 1: (A) Hybridization of the primers to the DNA template used in the selection. (B) Schematic representation of the selection protocol.

SELEX Protocol. Modification of a protocol described previously to generate circularly permuted tRNA molecules was employed (Pan et al., 1991). An outline of the SELEX protocol used in the study is shown in Figure 1. In order to obtain intramolecular circles, the transcribed RNA was ligated in 50 mM HEPES (pH 7.5), 3 mM DTT, 10 mM $MgCl_2$, 10 μg/mL BSA, 10% DMSO, 1 mM ATP, and 1 unit/μL T4 RNA ligase for 2 h at 37 °C. Circular products were isolated from end-to-end ligated products and unligated RNA by gel electrophoresis on 8% polyacrylamide gels under denaturing conditions.

The selection step was carried out using 4 μM circular RNA in the presence of RegB and S1 proteins in 50 mM Tris-HCl (pH 8.4), 1 mM DTT, and 0.1 mM EDTA at 37 °C. The complexity of the starting RNA repertoire was about 10^{14} unique sequences. Initially, RegB and S1 were used at

[1] Abbreviations: SELEX, systematic evolution of ligands by exponential enrichment; RT, reverse transcriptase; PCR, polymerase chain reaction; BSA, bovine serium albumin; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; PEG, poly(ethylene glycol); Tris,tris(hydroxymethyl)aminomethane.

0.3 and 2.4 μM concentrations, respectively, for incubation times of 2 h. After four rounds of selection, RegB and S1 concentrations were reduced to 0.1 and 0.8 μM, respectively, and the incubation times were progressively shortened to 5 min to increase the selection stringency. RegB-cleaved product (linear RNA) was partitioned from unreacted substrate (circular) by gel electrophoresis on 8% polyacrylamide gels under denaturing conditions.

The selected RNA was religated in order to regenerate the complete randomized region. RegB cleaves RNA leaving the phosphate group on the 3′ end of the processed molecule (Uzan et al., 1988). The selected RNAs were incubated in 70 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM DTT, 1 mM ATP, and 30 units of T_4 polynucleotide kinase in a 20-μL reaction volume for 1 h at 37 °C in order to phosphorylate the 5′ end and remove the phosphate group on the 3′ end by virtue of its 3′ phosphatase activity (Cameron & Uhlenbeck, 1977) in a single reaction step. The resulting RNA with a phosphate group on the 5′ end and a free -OH group on the 3′ end was then religated into the circular form essentially using the same ligation reaction conditions as described earlier. Complementary DNA from these selected RNAs was made by AMV reverse transcriptase in 50 mM Tris-HCl (pH 8.3), 60 mM NaCl, 10 mM DTT, 6 mM $Mg(OAc)_2$, 1 mM dNTPs, and 5 μM 3′ primer, and the resulting cDNAs were subsequently amplified by PCR as described above. The selection was carried out for 12 rounds, the round 12 selected pool was cloned, and the individual clones were sequenced and analyzed using standard protocols.

Assay of RegB Activity. RNAs were internally labeled by including [α-^{32}P]GTP in the transcription reaction and the RNA was ligated into circular molecules as described above. RegB activity was assayed under the same conditions used for the selection. RNA was heated to 70 °C briefly, allowed to cool to room temperature, and reacted with 100 nM RegB and 800 nM S1 protein in a buffer containing 50 mM Tris-HCl (pH 8.4), 1 mM DTT, and 0.1 mM EDTA in a 20-μL reaction volume at a final RNA concentration of 4 μM, for specified amounts of time at 37 °C, as mentioned in the figure legends. Reactions were stopped by adding deionized formamide to a final concentration of 50% (v/v), and the samples were run on 8% polyacrylamide gels containing 8 M urea to separate cleaved products (linear) from unreacted starting material (circular). The gels were analyzed by a Molecular Dynamics phosphoimager and the percentage of cleavage was determined by calculating the ratio of radioactivity in the product band relative to the total amount of radioactivity in both circular and linear bands.

RNase T1 Digestion. In order to determine the cleavage site, partial digestions by RNase T1 or alkali were carried out as described (Donis-Keller et al., 1977). Unlabeled circular RNA was cleaved with RegB in the presence of S1 under the reaction conditions described earlier. The linearized RNA was gel-purified and radiolabeled at the 5′ end with [γ-^{32}P]ATP by T4 polynucleotide kinase. Approximately 4 pmol of this RNA was reacted with 2.5 or 5 units of RNase T1 in a buffer containing 20 mM sodium citrate (pH 5.0), 1 mM EDTA, 4.8 M urea, and 0.5 μg/μL tRNA at 50 °C for 15 min. Alkaline hydrolysis ladders of the same RNAs were generated by incubating 4 pmol of 5′-radiolabeled RNA in a buffer containing 50 mM Na_2CO_3 (pH 9.0), 1 mM EDTA, and 0.3 μg/μL tRNA at 90 °C for 15 min. Both reactions were stopped by adding formamide to a final concentration of 50% (v/v), and the samples were run on 10% sequencing gels.

Binding Studies. Equilibrium dissociation constants (K_ds) between several selected RNAs and S1, RegB, or both proteins were determined using the nitrocellulose filter binding assay (Carey et al., 1983). Internally labeled, circularized RNA was incubated with S1 in the absence or presence of RegB or with RegB alone in RegB reaction buffer for 1 min at 37 °C. The concentrations of the proteins used ranged from 750 to 0.5 nM and total binding reaction volumes were 20 μL. After incubation, the reaction mixtures were vacuum-filtered through nitrocellulose filters (0.45-μm Millipore 25-mm HAWP) that had been pre-wet with 0.5 mL of RegB reaction buffer. The filters were washed with 5 mL of RegB buffer and the bound RNA was quantitated by scintillation counting, providing a measure of the percent RNA bound at varying protein concentrations. For each experiment, the amount of RNA retained on the filter in the absence of protein was also determined and was used for background correction. The percentage of RNA bound was plotted against the log protein concentration to estimate the K_d value.

RESULTS

The sequences of ligands obtained by cloning and sequencing of the round 12 selected pool are shown in Table 1. Although several of the clones occurred twice, the majority occurred only once. Out of the 42 clones obtained, all but 3 possess the -GGAG- sequence that was identified earlier as the required sequence for RegB cleavage. In the majority of the clones, the -GGAG- is found to occur within 3−4 nucleotides of the 5′ end of the randomized region; in some cases, the first G in the tetranuclotide sequence is actually provided by the 5′ fixed region. There is a strong bias for the existence of a G (in 54% of the clones) for the position immediately 3′ to the -GGAG- tetranucleotide; C is significantly underrepresented at this position (5%), and A is slightly so (13%). In contrast, G is not found in any of the clones at the position immediately 5′ to the -GGAG- tetranucleotide. About 33% of the clones, including those with the highest rates of cleavage, carry a GGAGGA sequence representing a strong Shine−Dalgarno sequence.

The most striking feature of the selected clones is their unusually high content of A and C nucleotides. Clones such as 7, 13, 16, and 42 consist almost entirely of A/C residues. Only six clones have short (5−6 nucleotides) A/C-rich stretches. We have grouped the sequences into two main classes on the basis of their A/C richness. Sequences that belong in class I carry relatively long (>10 nucleotides) stretches of A/C-rich regions, whereas class II sequences have short stretches of A/C nucleotides. An additional feature of the class I ligands is that the majority of them possess the -GGAG- sequence at or close to the 5′ end of the randomized region. In the class II ligands the -GGAG- tetranucleotide is generally found at or near the 3′ end of the randomized region. Sequence analysis failed to reveal any consensus structure pattern or any other sequence motif for these clones. Computer folding of the RNA sequences (Zuker, 1989; Jaeger et al., 1989) predicts that in most cases the -GGAG- lies in partially unpaired regions with the A nucleotide almost always unpaired. In the computer-predicted structures, the -GGAG- sequences were not usually observed in completely paired regions.

Table 1: Sequences Obtained by Cloning and Sequencing after 12 Rounds of Selection for Cleavage by T4 RegB Endonuclease in the Presence of S1[a]

	----5′ FIXED REGION-----	------RANDOMIZED REGION----	------3′FIXED REGION---
Class I			
1,27	gggcgaugccugucgagcaugcug	GUGGAGUCCAAAAGUCGUAAACCCGCCUGA	guagcuaaacagcuuugaauucgaaa
2		AAGGAGUAUAGAAUACGCUUGCACCACAAC	
3		CACGGAGCAAAGAAGAUUAACCACACACGC	
4		g GAGGAAAGAAGUUCAACCAACCUACACCAC	
5		g GAGUCCAGAGAGAGUAACCAACCUUCCACC	
7		g GAGGACAAACGAAAACACACACCGACCAAC	
9		AAGGAGGAUAAAAGUCUAACAACCCACGC	
11		g GAGGAGAAAGAUGUCAACCAACUUACCCAC	
13		g GAGGAAAAAACCAAUGAACAAACCACCCAC	
15		CACGGAGUCCAACACGUUAACCAACCGCAC	
17,21		g GAGGCCAGAAUAACAAAUCACCUGCACCGC	
18		g GAGUAUAAAAAGUCAAACCAGUCCACUGCC	
21		g GAGGCCAGAAUAACAAAUCACCUGCACCGC	
22		GGAGCGAGAAAACGG▼AGCACAACCCAGCCA	
25		CAAGGAGGACAAACGAAUAACCACAACCAACCAC	
26		AUGGAGGACAACGAUACAACCACACCACGC	
28		CGGAGGAUAGAACACAUACCAACCCAACAC	
30		CACGGAGAGGAAAAAUCAACAACGACCACG	
31		CACGGAGACCAAAACGAUACAACACGCAC	
33		AUGGAGAAGAAAAGAUCAACACACCACCAC	
35		g GAGGACAGAAAUUCACACACUACUCACCAC	
36		g GAGUCCAAAAGGUCAACCAUACACCCCACC	
37		AAGGAGUAUAGAAUACGCUUGCACCACAAC	
39		g GAGGACAGAAGACAAUCACACAACCCACA	
45,48		g GAGAAUAAAGAACAUUCAACACCUACCCAC	
46		AUGGAGGACCGAAGUAACACUAACCCGCAC	
47		g GAGGAGAAGUGUACAACCCAACCAACCCAC	
12		GCAACAGCACGGAGUACAGAACAACCACGC	
16		CAACCCACGGAGUACAAACCGACCACCAAC	
42		CACCCAACCACAAACACCACGGAGGCCAGA	
Class II			
20,32		UUGACCAAACUAGCCAUACAGGCGUCGGAG	
29		AACCAAUUAUCAGUCAAGAACUGCACGGAG	
40		AAGAAUAAACGUCUCAAGGAGACCACGG▼AG	
41		GUACGAAAUUUGCAGAACACCUGCGUCGGA g	
44		UACCACAGAGAUUAAUACACGGAGGACAGA	
6		ACGGAUACACCAGG▼ACAGUAACGCAUGCUC	
Non-cleavers			
24		UGCUUGGGGAAUAAUGAUCGGACCCCGCA	
43		ACGGAAUAAAGGGAAAGUAACGCAUGCGUC	

[a] The randomized region and 5′ and 3′ fixed regions are indicated. The GGAG sequences are underlined and the cleavage sites that have been determined experimentally are represented by arrowheads.

Four clones, 22, 40, 11, and 47, each carry two -GGAG- motifs. Since only a single cleavage event was selected for, we expected these clones to cleave specifically at one -GGAG- site and not the other. In the case of clones 11 and 47, the two GGAG sites are overlapping and cleavage at one site probably eliminates cleavage at the other. In order to determine the cleavage sites in clones 22 and 40, we carried out partial RNase T1 digestions of RegB-processed RNA. Clone 6, which undergoes a low level of RegB cleavage (data not shown) although it lacks a -GGAG- motif, was also tested. Partial RNase T1 digestion patterns of clones 6, 22, and 40 along with their respective alkaline hydrolysis ladders are shown in Figure 2. As expected, each clone displays a single processing site, as evidenced by the unique digestion pattern. RegB-mediated cleavage in clone 6 occurs at the -GGAC- site as indicated by an arrowhead in Table 1. The unique processing sites of clones 22 and 40 are similarly indicated by arrowheads. In the two latter cases, the surrounding sequences were carefully studied to understand why one -GGAG- site was chosen over the other for cleavage by RegB.

Computer RNA folding of clone 22 predicts that the -GGAG- that gets cleaved exists in a completely unpaired loop region, while the -GGAG- site that is not processed occurs in a helical region of the RNA. Furthermore, in this clone, the unprocessed -GGAG- is flanked on the 5′ end by a G nucleotide and on the 3′ end by a C. Nucleotide G is not found immediately 5′ to any of the cleavable -GGAG- sites, and a C nucleotide immediately 3′ to it is significantly underrepresented. This suggests that the primary sequence flanking the -GGAG- tetranucleotide and the local secondary structure of the surrounding region may play a role in specifying RegB cleavage of its substrate. The situation is not as clear in the case of clone 40. Computer folding predicts that the -GGAG- site which is not cleaved exists in a completely base-paired region, but the cleavable -GGAG- site can exist either in an unpaired loop region or in a partially unpaired region. However, these are only theoretical predictions based on minimum energy calculations, and chemical protection data are required for a better understanding of the local folded structure around the -GGAG- sequence.

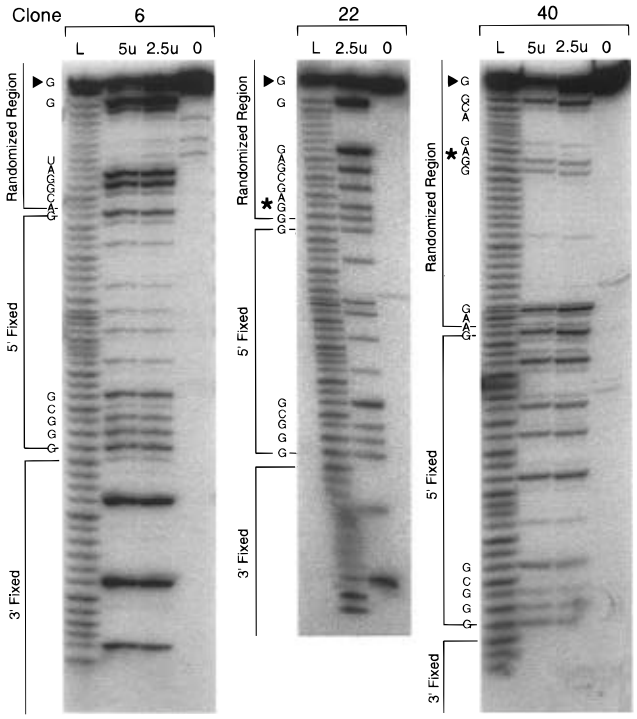

FIGURE 2: Partial RNase T1 digestion patterns of RNA from clones 6, 22, and 40. L, partial alkaline hydrolysis; O, in the absence of RNase T1; 2.5u or 5u indicates the number of RNase T1 units used in the digestion. The arrowhead represents the site of cleavage by RegB. In the case of clones 22 and 40, the GGAG sites that are not processed are indicated by an *asterisk*. 5′ and 3′ fixed regions are indicated by boxes.

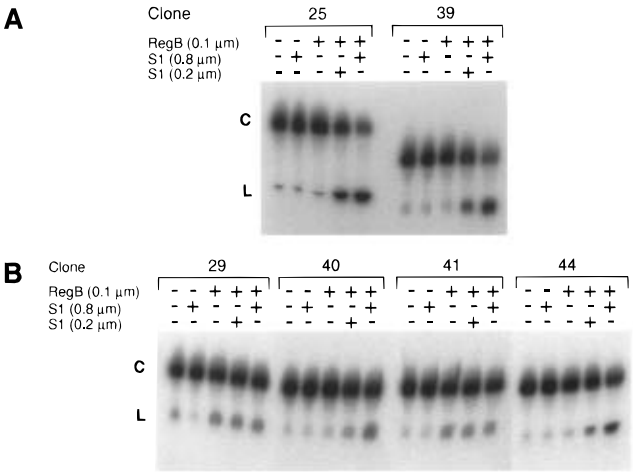

FIGURE 3: Effect of S1 on RegB-catalyzed cleavage on RNAs from selected class I (panel A) and class II (panel B) clones. As described under Experimental Procedures, 4 μM circular RNA was reacted in the presence (+) or absence (−) of RegB and S1 proteins at the final concentrations indicated for 15 min at 37 °C. The linear products (L) are separated from circular (C) RNA by gel electrophoresis.

We have investigated the cleavage requirements for several clones that carry either long or short stretches of A/C nucleotides. Clones 24 and 43, neither carrying a GGAG motif, showed no appreciable reaction above background (data not shown). These sequences probably resulted from random nicking of the circular RNA in the last round of selection. Figure 3 shows the cleavage requirements for class I clones 25 and 39 (each carrying a relatively long stretch of A/C nucleotides) and class II clones 29, 40, 41, and 44 (carrying short stretches). RegB-catalyzed cleavage is

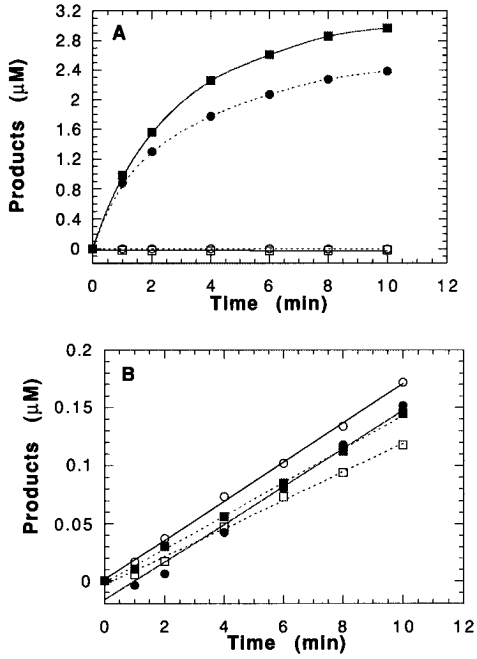

FIGURE 4: Kinetics of the RegB-catalyzed cleavage reaction for RNAs from (A) class I clones 25 (squares) and 28 (circles) and (B) class II clones 32 (squares) and 41 (circles). The cleavage reaction was carried out as described under Experimental Procedures, in the presence (filled symbols) or absence (open symbols) of 800 nM S1 protein. The slope of the initial linear phase of the reaction was used to obtain the rate of reaction under the specified reaction conditions.

entirely dependent upon the presence of S1 for clones 25 and 39 and two of the class II clones, 40 and 44 (Figure 3, compare lane 3 versus lanes 4 and 5). In the absence of S1, RegB does not catalyze any appreciable reaction above background nicking in these clones. Increasing the concentration of S1 yields higher product formation (compare lanes 4 and 5). The kinetics of the cleavage reaction using two class I clones, 25 and 28, in the presence or absence of S1 protein is shown in Figure 4, panel A. Neither clone shows any cleavage above background with RegB alone. However, addition of S1 causes a significant increase in product formation. We have followed the kinetics of RegB cleavage in the presence of S1 for several class I clones, and the initial velocities observed under the conditions employed in the selection are listed in Table 2. These clones undergo cleavage with about 10−30-fold higher initial velocities than the class II clones.

In contrast, S1 does not stimulate RegB cleavage in clones 29 and 41. Both clones carry short (5−6 nucleotides) stretches of A and C nucleotides. The kinetics of RegB cleavage for class II clones 32 and 41 in the presence or absence of S1 protein is shown in Figure 4, panel B. There is little or no difference in the rates of cleavage in the presence or absence of S1, indicating that S1 does not stimulate the cleavage reaction for either clone. Interestingly, in the absence of S1, RegB can catalyze cleavage in these clones, albeit at much lower rates, whereas it appears to be completely unable to do so in the class I clones. Thus, S1 is required for RegB-catalyzed cleavage of all the tested class I RNAs, which comprise about 83% of the selected clones. These RNAs undergo RegB-catalyzed cleavage reactions with about 10−30-fold higher rates compared to class II RNAs and generally carry relatively long stretches of A and C nucleotides. The correlation between A/C richness and

Table 2: Comparison of Reaction Rates of RNAs from Selected Clones[a]

clone number	S1 dependence for cleavage	observed initial rate[b] (μM/min)
25	+	0.78
28	+	0.65
16	+	0.42
35	+	0.39
21	+	0.25
44	+	0.23
36	+	0.16
1	+	0.15
18	+	0.14
39	+	0.12
6	−	0.04
29	−	0.02
32	−	0.02
41	−	0.02
random pool		0.18×10^{-3}

[a] The cleavage reactions were carried out under the same conditions used in the selection process using 4 μM circular RNA, 100 nM RegB, and 800 nM S1 at 37 °C. The rates were estimated from the slope of the initial linear phase of the reaction profile. [b] Rates were measured at 0.1 μM RegB and 0.8 μM S1.

the rate of cleavage or the ability of the reaction to be stimulated by S1 is attractive but not absolute, as shown in the cases of clones 44 and 40 (Figure 3, lower panel, and Table 2). Both RNAs carry very short regions of A and C nucleotides but require S1 for cleavage.

The position of the -GGAG- appears to play a signiicant role in determining the target site, since 69% of the selected clones carry the target sequence at the 5′ end of the randomized region. However, it does not seem to be sufficient for higher rates of cleavage. Clone 16, in which the -GGAG- is located close to the middle of the randomized region, has a high rate of cleavage by RegB (Table 2). Another example of clone 42, which also undergoes very strong cleavage by RegB (data not shown). However, class II clones with very low rates of cleavage generally have -GGAG- positioned at the 3′ end of the randomized region.

We performed equilibrium binding experiments to determine if there is a correlation between cleavage efficiency and binding affinities to RegB, S1, or the combination of the two proteins. Binding experiments were carried out with RNAs from clones 25, 39, and 41 as well as the initial random population to characterize the binding properties of the class I and class II RNAs and compare them with the initial pool from which they were selected. In the absence of RegB, binding of S1 to clones 25 and 39 is slightly better than the other two RNAs, with an apparent K_d of over 500 nM (Figure 5, panel A). Interestingly, binding of S1 to clone 41 is even poorer than to the random pool. There is no detectable binding of RegB in the absence of S1 to any of the RNAs over the concentration range tested (5−250 nM). In the presence of both proteins, binding to clones 25 and 39 is significantly enhanced to an apparent K_d of about 50 nM. In contrast, the presence of both proteins has little effect on the binding of clone 41 and the random RNA pool.

Under the conditions used in the binding experiment, a significant amount of the reactant RNAs would be converted into products, as we were not able to prevent the cleavage reaction following the binding step. In order to determine whether the product RNAs bind to S1 in the presence or absence of RegB, we performed the same binding experiment

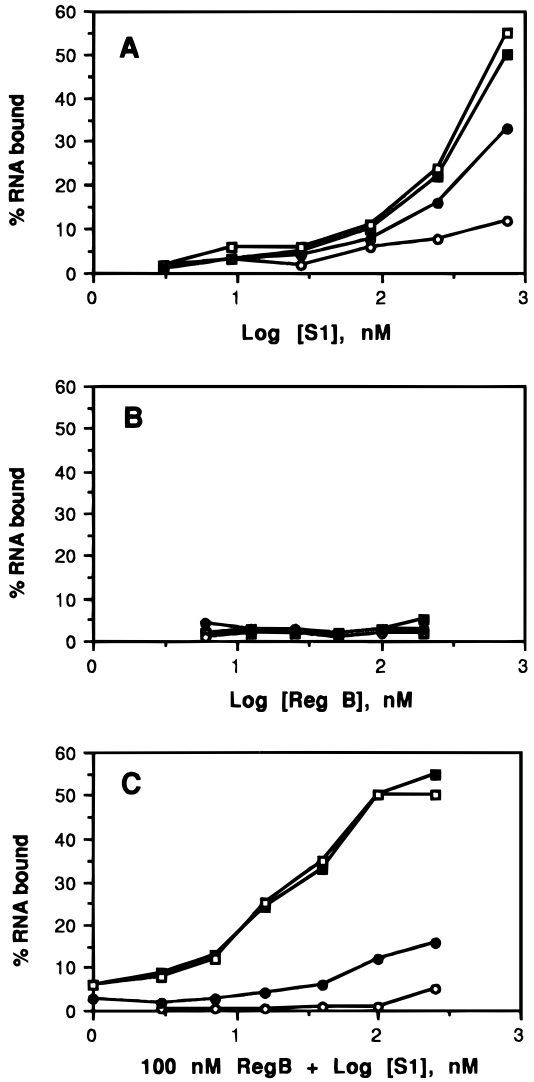

FIGURE 5: Binding analysis of RNAs from selected clones for S1 alone (panel A), RegB alone (panel B), and S1 in the presence of a constant 100 nM RegB (panel C). The protein concentrations used were 3×10^{-9} to 7.5×10^{-7} M S1 (panel A), 6.25×10^{-9} to 2×10^{-7} M RegB (panel B), and 4×10^{-10} to 2.5×10^{-7} M S1 at 100 nM constant RegB (panel C) with 1×10^{-10} M RNA in each case. The RNAs tested were from the initial random pool (filled circles), clone 41 (open circles), clone 25 (filled squares), and clone 39 (open squares).

using purified cleavage products of RegB cleavage. We found that the product RNAs bind to S1 and the combination of both proteins with binding properties very similar to the substrate RNAs (data not shown). Presumably, the affinities of the reactants for RegB and S1 are similar to those of the products. Substrate binding by one or more of the protein components precedes RNA cleavage, and cleavage results in the disruption of the -GGAG- recognition sequence. Our binding experiments have shown that the binding characteristics remain unaltered following the cleavage step. The most likely explanation for this observation is that the RNA regions independent of the -GGAG- may be responsible for binding in the presence of both RegB and S1. Although the region of RNA responsive to protein binding has not been elucidated, the A/C-rich region predominant in class I clones would be a likely candidate.

We were interested in finding out if richness for A/C nucleotides is a general feature associated with RegB cleavage. We have searched the nucleotide sequences for

A/C-rich regions in the vicinity of the ribosomal binding sites of several T4 messenger RNAs that are known to be processed by RegB. We could not detect relatively long stretches of A/C-rich regions near the RegB processing sites of these mRNAs. However, many early T4 messages carry AU-rich sequences (about 6−10 nucleotides long) immediately downstream of the processed GGAG sites. RegB-dependent secondary cleavage events within this AU-rich region have been reported to occur in gene 43 and *comC*α messages of T4 (Sanson & Uzan, 1993).

DISCUSSION

The overwhelming majority of the clones obtained in the selection carried the -GGAG- tetranucleotide, indicating that this sequence is preferentially cleaved by RegB, probably with a GGAGGA consensus. As this consensus constitutes a strong Shine−Dalgarno sequence, it may explain the observation that -GGAG- sequences in ribosomal binding sites are processed by RegB much more efficiently than the same sequences in the coding regions of many early messages (Uzan et al., 1988). The only other sequence that was selected was -GGAC-, although the rate of cleavage at this sequence is much lower. The other striking feature is that the class I clones, which make up the huge majority of the selected clones, possess fairly long stretches of A and C nucleotides. Computer folding of the RNA sequences predicts that in most cases the -GGAG- tetranucleotide lies in partially unpaired regions. Thus, it would seem that one criterion for efficient RegB cleavage is that the -GGAG- be in at least a partially unstructured conformation. The relatively long stretches consisting of A and C nucleotides that are evident in most of our clones could also be contributing to a partially unpaired conformation near the -GGAG- site.

Comparison of the observed rates shows that the existence of a -GGAG- sequence leads to a rate acceleration of about 10^2 over the starting RNA repertoire. A further rate enhancement up to a factor of about 30 is apparently achieved from the nucleotides surrounding the -GGAG- sequence. There is a very strong correlation between the rate of cleavage and the ability of the cleavage reaction to be stimulated by S1 (Table 2). The RNAs in which RegB-catalyzed cleavage is stimulated by S1 have higher rates, and RNAs in which the cleavage event is independent of S1 have cleavage rates lower by more than an order of magnitude. The latter category of RNAs carry short stretches of A/C nucleotides, whereas the former category of RNAs typically have relatively long A/C-rich regions.

It is possible that the preference of RNA ligase may have had some influence on the outcome of the selection. This is true for the second ligation step following the cleavage event, after which the ends are different for each RNA molecule. The molecules that had ends which are not efficiently ligated by RNA ligase would not have survived the selection. The primers used are common to all the RNA molecules and would probably not have contributed to the selection process.

The turnover numbers for RegB calculated from these reactions are very low compared with other protein enzymes (Ruckman et al., 1995; this work). Clone 25, which is the fastest cleaving ligand so far tested as a substrate for RegB, produces a turnover number of 7.8/min. However, very low turnover numbers are not uncommon among enzymes involved in RNA cleavage. Examples include *Bacillus subtilis* RNase M5 (Pace et al., 1984) and RNase T_2 (Yasuda & Inoue, 1981).

The selection criteria that we have used in this experiment do not exclude the enrichment of sequences that are capable of undergoing self-cleavage independent of RegB. However, this selection did not yield any RNAs capable of undergoing a protein-independent cleavage event. Since sequences which are capable of self-cleavage would be expected to occur in the initial RNA repertoire at a much lower probability than those that are cleavable by RegB, they may have been competed out by the latter in the early rounds of selection. Presumably, the class II RNAs having very low rates of RegB cleavage would eventually be competed out by the dominating class I sequences if the selection had been allowed to continue further. The existence of a protein-independent cleavage observed in another selection carried out under different conditions is being currently investigated.

The stimulatory effect of S1 on RegB cleavage of RNAs that belong in class I is clearly evident, but the data from our selection experiment are insufficient to shed light on the mechanism by which S1 exerts its stimulatory action. S1 may be directly binding to the RNA and mediating the cleavage reaction through protein−RNA interactions. Interactions between S1 and RNA have been well-documented (see introduction). Recently, SELEX has been employed to generate high-affinity ligands against purified S1 protein. These ligands contain a pseudoknot structure with well-conserved nucleotides in the loop regions; the ligands bind tightly to S1 with K_ds of 2 nM (Ringquist et al., 1995). Our binding studies do not indicate significantly higher binding between S1 and the selected RNAs over the initial RNA pool. However, this does not exclude the possibility that nonspecific interactions between RNA and S1 could stimulate the cleavage reaction. S1 has been shown to be able to unwind RNA (Bear et al., 1976; Szer et al., 1976) and has been proposed to facilitate interactions between mRNA and the ribosome by unwinding double-stranded regions in the message (Thomas & Szer, 1982; Van Dieijen et al., 1976). Although subsequent experiments suggest that the unwinding activity may not play a significant role in translation or replication of Qβ RNA (Cole et al., 1982; Suryanarayana & Subramanian, 1983), it may have other important biological implications. In this case, S1 may bind to the RNA and unwind secondary structures sufficiently for RegB to recognize its target sequence.

Alternatively, S1 could be stimulating RegB activity directly through protein−protein interactions. S1 is a fairly large protein of about 61 kDa, carrying multiple domains for binding to ribosomal proteins as well as to nucleic acids. *E. coli* ribosomal protein S1 is noted for being able to form complexes with proteins other than those of ribosomal origin, having very diverse functions. S1 is an integral part of the four-subunit complex of Qβ replicase (Wahba et al., 1974; Inouye et al., 1974) and plays a functional role in the recognition and replication of the phage (+) strand (Kamen et al., 1972). S1 has also been shown to form a distinct complex with Redβ, a protein involved in phage λ replication, and appears to exert a negative regulatory function (Muniappa & Mythili, 1993; Venkatesh & Radding, 1993). Our binding experiments with RegB show the absence of specific binding to the selected clones over the initial randomized

pool. However, binding is dramatically improved in the presence of S1. It has been proposed that T4 RegB endonuclease is probably too small (19 kDa in size) to carry both a catalytic domain and a RNA binding domain, and it may have evolved to make use of the RNA binding property of the readily available host S1 protein in order to carry out its biological function (Ruckman et al., 1995; Sanson & Uzan, 1993). However, a direct physical interaction between the two proteins has not yet been observed.

ACKNOWLEDGMENT

We are grateful to Drs. Sean Eddy and Gary Stormo for carrying out computer analysis of our sequences. We thank Dr. Alap Subramanian for repeated gifts of purified S1 protein. We also thank Drs. Judy Ruckman and Sumedha Jayasena for many stimulating discussions and helpful editorial comments. This work was funded by NIH grants GM 19963 and GM 28685 and also partially supported by NeXstar Pharmaceuticals.

REFERENCES

Bardwell, J. C. A, Regnier, P., Chen, S.-M., Nakamura, Y., Grunberg-Manago, M., & Court, D. L. (1989) *EMBO J. 8*, 3401−3407.
Bear, D. G., Ng, R., Derveer, D., Johnson, N. P., Schleich, T., & Noller, H. F. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 1824−1828.
Belasco, J. G., & Higgins, C. F. (1988) *Gene 72*, 15−23.
Brawerman, G. (1987) *Cell 48*, 5−6.
Cameron, V., & Uhlenbeck, O. C. (1977) *Biochemistry 16*, 5120−5126.
Carey, J., Cameron, V., de Hasath, P. C., & Uhlenbeck, O. C. (1983) *Biochemistry 22*, 2601−2607.
Chen, C. A., & Belasco, J. G. (1990) *J. Bacteriol. 172*, 4578−4586.
Cole, P. E., Sawchyn, I., & Guerrier-Takada, C. (1982) *J. Biol. Chem. 257*, 12929−12934.
Deutscher, M. P. (1988) *Trends Biochem. Sci.* 136−139.
Donis-Keller, H., Maxam, A. M., & Gilbert, W. (1977) *Nucleic Acids Res. 4*, 2527−2538.
Donovan, W. P., & Kushner, S. R. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6427−6430.
Draper, D. E., Pratt, C. W., & von Hippel, P. H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4786−4790.
Gold, L. (1995) *J. Biol. Chem. 270*, 13581−13584.
Gold, L., Polisky, B., Uhlenbeck, O., & Yarus, M. (1995) *Annu. Rev. Biochem. 64*, 763−797.
Gorski, K., Roch, J. M., Prentki, P., & Krisch, H. M. (1985) *Cell 43*, 461−469.
Hayashi, M. N., & Hayashi, M. (1985) *Nucleic Acids Res 13*, 5937−5948.
Inouye, H., Pollack, Y., & Petre, J. (1974) *Eur. J. Biochem. 45*, 109−117.
Jaeger, J. A., Turner, D. H., & Zuker, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7706−7710.
Kamen, R., Kondo, M., Romer, W., & Weissmann, C. (1972) *Eur. J. Biochem. 31*, 44−51.
Klug, G., & Cohen, S. N. (1990) *J. Bacteriol. 172*, 5140−5146.
Milligan, J. F., & Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51−62.
Mulsch, A., Colpan, M., Wollny, E., Gassen, H. G. & Riesner, D. (1981) *Nucleic Acids Res. 9*, 2367−2385.
Muniappa, K., & Mythili, E. (1993) *Biochem. Mol. Biol. Intl. 31*, 1−11.
Pace, B., Stahl, D. A., & Pace, N. R. (1984) *J. Biol. Chem. 259*, 11454−11458.
Pan, T., & Uhlenbeck, O. C. (1992) *Biochemistry 31*, 3887−3895.
Pan, T., Gutell, R. R., & Uhlenbeck, O. C. (1991) *Science 254*, 1361−1364.
Peterson, C. (1992) *Mol. Microbiol. 6*, 277−282.
Portier, C., Dondon, L., Grunberg-Manago, M., & Regnier, P. (1987) *EMBO J. 6*, 2165−2170.
Ringquist, S., Jones, T., Snyder, E. E., Gibson, T., Boni, I., & Gold, L. (1995) *Biochemistry 34*, 3640−3648.
Ruckman, J., Parma, D., Tuerk, C., Hall, D. H., & Gold, L. (1989) *New Biol. 1*, 54−65.
Ruckman, J., Ringquist, S., Brody, E., & Gold, L. (1995) *J. Biol. Chem. 269*, 26655−26662.
Sanson, B., & Uzan, M. (1993) *J. Mol. Biol. 233*, 429−446.
Schmeissner, U., McKenny, K., Rosenberg, M., & Court, D. (1984) *J. Mol. Biol. 176*, 39−53.
Stevens, A., & Maupin, M. K. (1987) *Nucleic Acids Res. 15*, 695−708.
Suryanarayana, T., & Subramanian, A. R. (1983) *Biochemistry 22*, 2715−2719.
Thomas, J. O., & Szer, W. (1982) *Prog. Nucleic Acid Res. Mol. Biol. 27*, 157−187.
Thomas, J. O., Kolb, A., & Szer, W. (1978) *J. Mol. Biol. 123*, 163−176.
Uzan, M., Favre, R., & Brody, E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8895−8899.
van Dieijen, G., van Knippenberg, P. H., & van Duin, J. (1976) *Eur. J. Biochem. 64*, 511−518.
Venkatesh, T. V., & Radding, C. (1993) *J. Bacteriol. 175*, 1844−1846.
Wahba, A. J., Miller, M. J., Niveleau, A., Landers, T. A., Carmichael, G. G., Weber, K., Hawley, D. A., & Slobin, L. I. (1974) *J. Biol. Chem. 249*, 3314−3316.
Wong, H., & Chang, S. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3233−3237.
Yasuda, T., & Inoue, Y. (1981) *Eur. J. Biochem. 114*, 229−234.
Zuker, M. (1989) *Science 244*, 28−52.

BI951879B