Volume 17 Number 23 1989 Nucleic Acids Research

Bacteriophage T4 RegA protein binds to the Shine–Dalgarno region of gene 44 mRNA

Kevin R.Webster, Hedy Y.Adari* and Eleanor K.Spicer

Department of Molecular Biophysics and Biochemistry, Yale University, 333 Cedar Street, New Haven, CT 06510-8024, USA

Received July 7, 1989; Revised and Accepted October 3, 1989

ABSTRACT

We have overproduced and purified wild type regA protein, a translational repressor encoded by bacteriophage T4. The repressor activity of the cloned regA protein has been tested on four known regA target genes (T4 genes: 44, 45, rpbA and regA) using in vitro coupled transcription-translation reactions. We have demonstrated the sensitivity of two additional T4 genes coding for α-and β-glucosyltransferases to regA protein in vitro. The regA target site on the gene 44 messenger RNA has been identified through deletion analysis and RNase protection assays, using plasmids containing gene 44-lacZ fusions. The effect of regA protein on expression of 44P-β-galactosidase fusion proteins was assayed in vitro, in coupled transcription-translation reactions. Analysis of deletion mutants of gene 44-lacZ localized the regA recognition region to between nucleotides -11 and +9 of the mRNA. RNase protection assays of g44-lacZ transcripts further defined the site of regA protein interaction to between nucleotides -10 and +2 of the mRNA. This region overlaps the gene 44 Shine-Dalgarno region and the A and U of the initiation codon.

INTRODUCTION

The bacteriophage T4 regA protein controls the expression of at least 12 T4 early and middle genes, at the level of translation (1). Expression of the regA gene from a plasmid vector, in the absence of T4 infection, results in inhibition of expression of specific T4 genes both in cis (2) and in trans (3). Further studies have demonstrated that purified regA protein (M_R = 14,600) can specifically inhibit translation of target genes in in vitro protein synthesis assays (4,5). Thus regA protein does not require other phage proteins for translational repression.

Recently, the ability of regA protein to compete with ribosomes for binding to the translation initiation region of target mRNAs has been demonstrated by the studies of Winter *et al.* (5). The authors examined the effect of mutations in the rIIB mRNA on regA protein competition with ribosomes and found that mutations in the translation initiation codon had the greatest effect on competition. RNase protection assays of regA binding to rIIB mRNA further indicated that regA protein binds to a

site overlapping the rIIB initiation codon.

Although the target site on the rIIB message has been studied in some detail (5,6), the lack of a consensus target sequence in the other regA-sensitive genes makes it unclear what structures or other sequences are recognized by regA protein. To fully understand how this protein represses a subset of T4 mRNAs, detailed studies of a number of regA targets must be performed. Towards that end, we have constructed plasmids that enable inducible overproduction of the wild type regA protein and we have tested the ability of purified regA protein to repress synthesis of six T4 genes in vitro.

In addition, we have examined the region of gene 44 mRNA required for the recognition and specific binding of regA protein. As shown here, gene 44 expression is among the most sensitive to regA protein repression in in vitro assays. To define the region of g44 mRNA required for recognition by regA protein, we have constructed vectors in which the 5' end of gene 44 (nucleotides -37 to +99, relative to A of the initiation codon) was fused to the lacZ gene. Expression of the 44P-β-galactosidase fusion protein produced from the gene 44-lacZ gene is repressed by regA protein in vitro. Deletion analysis of the g44-lacZ vector has enabled localization of the regA recognition region to between nucleotides -11 and +9 of the mRNA. RNase protection studies have also localized the regA interaction site to a 12-nucleotide region of the g44 mRNA spanning nucleotides -10 to +2.

MATERIALS AND METHODS

Reagents:

Oligodeoxynucleotides were synthesized on an Applied Biosystems oligonucleotide synthesizer and were purified by HPLC on a DEAE 60-7 column (Nucleogen). Ribonucleotide triphosphates were purchased from Pharmacia (P.L. Biochemicals) and T7 RNA polymerase (300,000 units/mg) was the generous gift of D. Mueller and J. E. Coleman (Yale University).

Strains and Plasmids:

E. coli AR120 (λ cI^+, N^+) and plasmid pAS1 (7) were obtained from A. Shatzman (Smith, Kline and French). Plasmids pPL$c\alpha$gt and pPL$c\beta$gt (8) were generously provided by W. Rüger (Ruhr-Universität Bochum). pGW7, a pBR322 derivative containing a 4.0 kilobase pair fragment of λ DNA carrying the cI^{857} and N genes and the leftward promoter P_L, was kindly provided by G. G. Wilson (New England

Biolabs). pIBI20 was purchased from International Biotechnologies Inc., and M13mp18 was purchased from Bethesda Research Laboratories.

Site-directed mutagenesis of the regA90 gene:

Oligonucleotide-directed mutagenesis was carried out by annealing a 23 nucleotide primer to single-stranded DNA of a M13mp9 derivative containing the regA90 gene (4). The sequence of the oligo, 5'-GTTTCTTTTACTTTCAGAAAATC-3', was complementary to nucleotides #31-53 of the regA structural gene, with a single mismatch at nucleotide #43 in regA90. Primer annealing, chain extension and transfection of E. coli 71-18 were carried out as described by Gillam and Smith (9). RegA wild type revertants were selected by plaque lifts and filter hybridization to the mutagenic oligonucleotide labeled with (^{32}P), using Colony/Plaque Screen (NEN Research Products) membranes. The presence of the change at codon #15 from TTA to GTA was confirmed by nucleotide sequence analysis of single-stranded M13-regA DNA in a number of clones. DNA sequencing was performed by the method of Sanger et al. (10).

Purification of regA protein:

RegA protein (M_R = 14,600) was purified from AR120 cells containing plasmid pAS1regA following induction of transcription from the lambda P_L promoter by nalidixic acid treatment (11), as previously described (2). The protein purification scheme was essentially as described earlier (4). The protein was stored at concentrations of 0.7 → 1.1 mg/ml in 20 mM Tris/HCl (pH 7.5), 150 mM NaCl, 5mM $MgCl_2$, 1mM EDTA, 1mM DTT and 50% glycerol at -70°C. Concentration of regA protein beyond 1.4 mg/ml resulted in precipitation of the protein. One gram of nalidixic acid-induced cells yielded ≈2 mg of pure regA protein.

Construction of vectors for gene 44 target studies.

A.) M13 mp44. A DraI/SacI restriction fragment carrying the 5' end of gene 44 was isolated from plasmid pSP44 and cloned into M13 mp18- and pIBI20-digested with SacI and SmaI enzymes. The resulting vectors, mp44 and pIBIg44, contained nucleotides -37 to +99 of gene 44 fused in-phase with the 5' end of the E. coli lacZ gene. 5'-deletions in mp44 were then introduced at the SacI restriction site by Bal31 exonuclease digestion, following the procedure of Gray et al. (12).

B.) pg44/26: Nucleotides +27 to +99 of gene 44 were deleted from pIBIg44 by partial digestion with XmnI, which cut at position +27 in gene 44, followed by digestion at the HincII site contained within the polylinker region of pIBI20.

Religation of this vector produced pg44/26, containing nucleotides -37 to +26 of gene 44 fused to the lacZ gene on pIBI20. Subsequently, 3'-deletions initiated at the PstI restriction site in the pIBI20 polylinker region by Bal31 exonuclease digestion, produced plasmids pg44/22 and pg44/9, containing nucleotides -37 to +22 and -37 to +9 of gene 44, respectively.

C.) pg44/9B: A synthetic oligonucleotide containing nucleotides -11 to +9 of gene 44 with flanking restriction site sequences was cloned into plasmid pIBI20 cut with EcoRI and PstI using the method of Derbyshire *et al.* (13). The sequence of the oligonucleotide was: 5'-aattcATGAGGAAATTATGATTACTtgca- 3', where upper case letters indicate gene 44 nucleotides.

In vitro protein synthesis assays:

DNA-dependent, coupled transcription-translation reactions were performed using a kit derived from E. coli (Amersham), as previously described (4). Reactions were performed in the presence of ^{35}S-methionine (6μCi) as recommended by the manufacturer except that reactions were scaled down to a 7μl total volume. Following electrophoresis of the total reaction mixture on SDS-15% polyacrylamide gels, the gels were dried and autoradiography was performed for 12-15 hours. Reactions utilized plasmid DNAs (0.85μg) except in one reaction in which a gel-purified transcript synthesized in vitro from plasmid pg44/9B linearized at a BglI site was used.

In vitro transcription and ^{32}P-end labelling of mRNA:

Transcription reactions were carried out in 250μl of 40mM Tris-HCl pH8.1, 20mM $MgCl_2$, 1mM spermidine, 5mM DTT, 50μg/ml BSA, 0.01% Triton X-100, 80 mg/ml PEG 8000 with 50nM linearized plasmid template (14). T7 RNA polymerase was added (10 units/μl final concentration) and the reaction was carried out at 37°C for 4 hours. Transcripts were purified on 6% polyacrylamide-8M urea gels. The mRNA bands were localized by UV shadowing, excised and eluted overnight (at 37°C) in 2mls of 0.5M NH_4Ac, - 1mM EDTA. mRNA was then dephosphorylated using calf intestine alkaline phosphatase and was 5' end-labeled by phosphorylation with γ ^{32}P-ATP as described by Deckman and Draper (15). ^{32}P-labeled transcripts were then repurified on 6% polyacrylamide-8M urea gels and eluted as described above.

RNase Protection Assays:

RegA protein was incubated with a gel-purified, ^{32}P-labeled transcript (75,000 cpm, Cherenkof) first for 6 min. at 25°C, then for 4 min. at 0°C in 10mM Hepes, pH7.2, 150mM NaCl, 5mM $MgCl_2$, 1mM β-mercaptoethanol, 10μg/ml yeast tRNA. In

separate experiments, a specific RNase was then added and the incubation was continued for 5 minutes on ice. An equal volume of formamide loading buffer was then added and the sample heated to 75°C for three minutes. The RNA was then loaded directly onto a 10% polyacrylamide-8M urea gel. Following electrophoresis, digestion products were visualized by autoradiography of the dried gel at room temperature. The RNA sequence was confirmed by partial digestion with RNases T_1, PhyM and U_2 (16,17).

RESULTS

Overexpression of wild type regA protein.

In our previous studies on regA protein-mediated translational repression of T4 genes in vitro (4), regA protein was purified from cells containing the overexpression vector pHA90R. RegA protein produced from these cells was found to contain a leucine for valine substitution at residue 15. To obtain overproduction of wild type regA protein, we used a synthetic oligonucleotide for site-directed mutagenesis of an M13 subclone containing the entire regA90 gene. The mutagenesis converted codon #15 of regA90 from TTA (leucine) to GTA (valine) which is the wild type codon. M13 phage containing the regA$^+$ gene were identified by nucleic acid sequencing (data not shown). The nucleotide sequence of the entire regA$^+$ gene in the M13 regA clone was determined, confirming that no other nucleotide substitutions were present. A restriction fragment carrying the regA$^+$ gene and 37 nucleotides upstream from the initiation codon was subsequently transferred from M13 to the expression vector pAS1 (7). The regA gene in the resultant plasmid, pAS1regA, is positioned downstream from the inducible phage λ leftward promoter P_L.

When AR120/pAS1regA cells were induced by nalidixic acid (11,4), induction of transcription from P_L ensued, producing high levels of regA protein (data not shown). RegA protein was purified by chromatography on DEAE-cellulose and poly(U)-agarose as previously described (4). Wild type regA protein eluted from poly(U)-agarose at a slightly higher NaCl concentration than regA90 protein (0.6M versus 0.4M). N-terminal sequencing of regA protein produced from ARI20/pAS1regA cells confirmed that residues #1-19 were identical to those of wild type regA protein. In addition, comparative HPLC tryptic peptide analysis indicated that the hydrophobicity of a single peptide was altered in comparing the regA90 and wild type proteins, in agreement with a single leucine to valine difference between the two proteins.

A.

pAS1 regA

CII r.b.s. Bam HI STOP → regA START

P_L --» GUUAUC**UAAGGA** - 10b - **AUG**GAUCCG - 24b - UU**GGA**AUGGUAAA**AUG**AUU...

Met Asp Pro - 8aa - Leu Glu Trp Met Ile

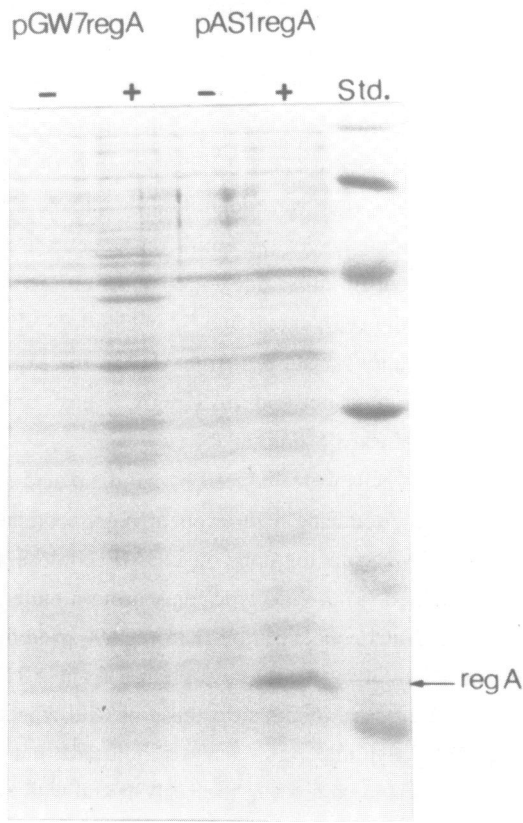
pGW7 reg A

Bam HI → regA START

P_L --» CCGAGAAUAACGAGUGGAUCCG - 24b - UU**GGA**AUGGUAAA**AUG**AUU...

Met Ile

B



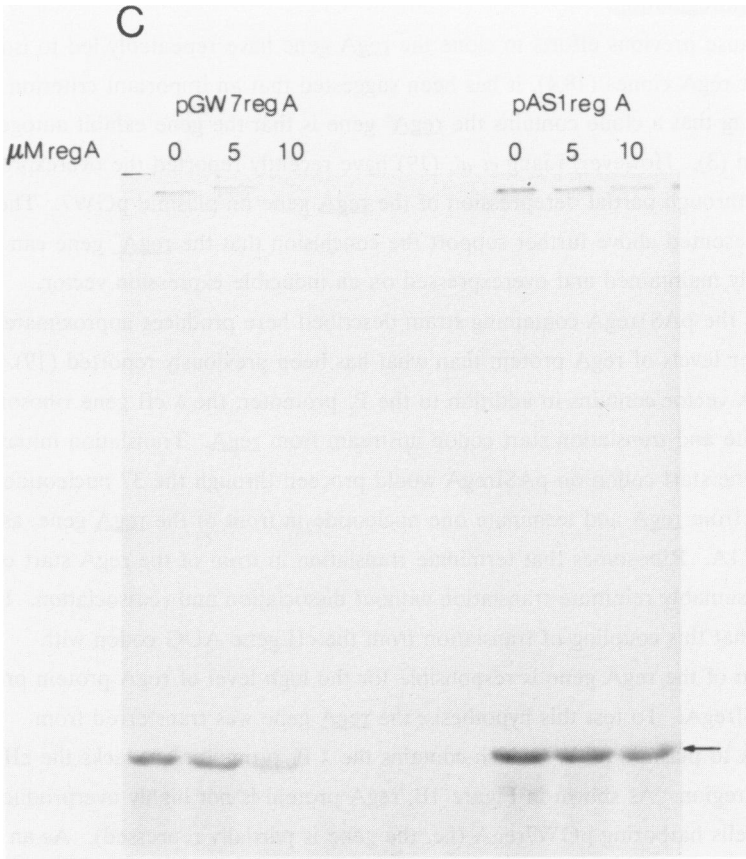


Fig. 1. A) Comparison of the translation initiation regions of mRNA from pASI and pGW7-regA plasmids. The Shine-Dalgarno regions and initiation codons are indicated by boldface type. pASIregA: Initiation of translation at the cII gene initiation codon results in the synthesis of a 14-amino acid peptide which terminates one nucleotide in front of the regA gene initiation codon. pGW7regA: Initiation of regA protein synthesis occurs at the regA gene's ribosome binding site. B) Expression of $regA^+$ in vivo: Total cell extracts of AR120 cells containing the respective plasmids before (-) and after (+) induction with nalidixic acid were analyzed by SDS-15% polyacrylamide gel electrophoresis. Proteins were visualized by Coomassie blue staining. C) RegA protein synthesis in in vitro coupled transcription-translation reactions: 0.85μg of pGW7regA and pAS1regA plasmid DNAs were used as templates in separate reactions. Purified regA protein was added (0, 5, and 10μM) to both sets of reactions, and the reactions were carried out in the presence of ^{35}S-methionine for 15 minutes. Protein products were analyzed on an SDS-15% polyacrylamide gel and visualized by autoradiography.

RegA Autoregulation:

Because previous efforts to clone the regA gene have repeatedly led to isolation of mutant regA clones (18,4), it has been suggested that an important criterion for establishing that a clone contains the regA$^+$ gene is that the gene exhibit autogenous repression (3). However, Liang *et al.* (19) have recently reported the overexpression of regA$^+$ through partial derepression of the regA gene on plasmid pGW7. The results presented above further support the conclusion that the regA$^+$ gene can be both stably maintained and overexpressed on an inducible expression vector. However, the pAS1regA-containing strain described here produces approximately 10-fold higher levels of regA protein than what has been previously reported (19). The pAS1regA vector contains in addition to the P_L promoter, the λ cII gene ribosome binding site and translation start codon upstream from regA. Translation initiating at the cII gene start codon on pASIregA would proceed through the 37 nucleotides upstream from regA and terminate one nucleotide in front of the regA gene, as shown in Figure 1A. Ribosomes that terminate translation in front of the regA start codon could presumably reinitiate translation without dissociation and reassociation. It is possible that this coupling of translation from the cII gene AUG codon with translation of the regA gene is responsible for the high level of regA protein produced from pASIregA. To test this hypothesis, the regA gene was transferred from pAS1regA to plasmid pGW7 which contains the λ P_L promoter but lacks the cII gene initiation region. As shown in Figure 1B, regA protein is not highly overproduced in induced cells harboring pGW7regA (i.e. the gene is partially repressed). As an additional test, the pGW7regA and pAS1regA plasmids were used as templates in DNA-dependent protein synthesis assays, and the effect of purified wild type regA protein on in vitro translation of regA was examined. As shown in Figure 1C, when equal amounts of plasmid DNA were used as templates, 1.7 times more (^{35}S)-labeled regA protein was synthesized from pAS1regA than from pGW7regA (as determined by densitometry) and 5μM regA protein had a 1.5 fold greater inhibitory effect on the synthesis of protein directed from pGW7regA than from pAS1regA, in 10-minute reactions. These in vitro transcription-translation assays (Figure 1C) demonstrate that purified wild type regA protein can repress the expression of the regA$^+$ gene from both pAS1regA and pGW7regA plasmids, although complete repression of synthesis from pAS1regA apparently requires higher regA protein concentrations.

In vitro repression of target genes by regA protein.

The activity of purified wild type regA protein was tested in vitro by utilizing

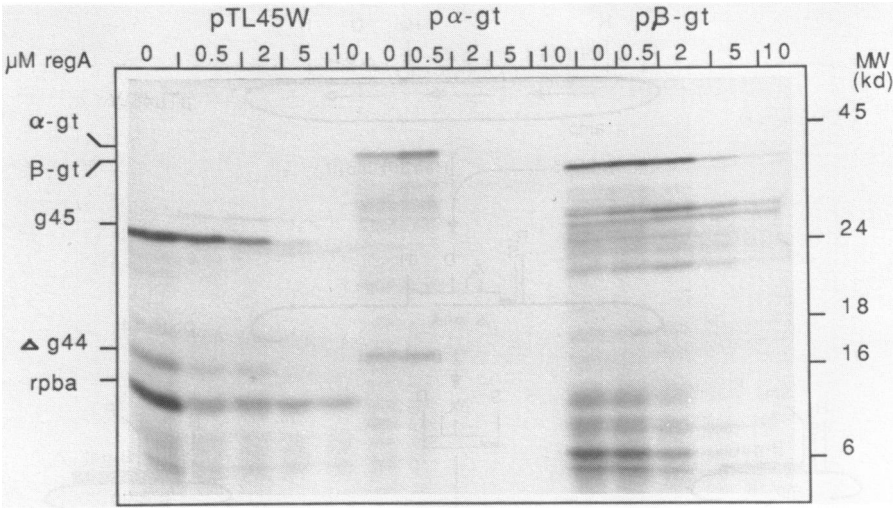


Fig. 2. Translational repressor activity of cloned regA protein. In vitro coupled transcription-translation reactions were carried out with pTL45W (Adari *et al.*, 1985), pPLcαgt, or pPLcβgt (Tomaschewski *et al.*, 1985) plasmid DNA (0.85μg each), in the presence of ^{35}S-methionine and increasing concentrations of regA protein, for 15 minutes. Products were separated on SDS-15% polyacrylamide gels and visualized by autoradiography. The band labeled rpbA appears to contain a second plasmid-encoded protein that is only weakly repressed by regA protein. The lower molecular weight bands (<16kd) produced in reactions containing pPLβgt are presumed to be due to premature translation termination of β-gt transcripts, since their synthesis is sensitive to regA protein.

coupled transcription-translation assays. When protein synthesis reactions were carried out using plasmid pTL45W (2) the products of three T4 genes - rpbA, gene 45 and Δ gene 44 (the 5' half of gene 44)- located downstream from the λ P_L promoter on pTL45W, were synthesized. As shown in Figure 2, addition of regA protein in concentrations of 1μM to 10μM to the protein synthesis reactions repressed synthesis of all three T4 proteins. A similar effect of regA protein was seen in reactions primed with plasmids containing the T4 α- and β- glucosyl transferase genes (Figure 2). All five of these genes have been shown to be regA regulated in vivo, i.e. they are overexpressed in cells infected with T4 regA$^-$ phage (1). In contrast, expression of T4 gene 32 is autogenously regulated (20) and is unaffected by regA in vivo (1). In control reactions, expression of gene 32 from plasmid pYS7 (21) was not affected by addition of 10μM regA protein (data not shown, see ref. 4) and in subsequent reactions, expression of β-galactosidase was not repressed by 5μM regA protein (see Fig. 4, below). The five T4 genes examined showed differential sensitivity to regA

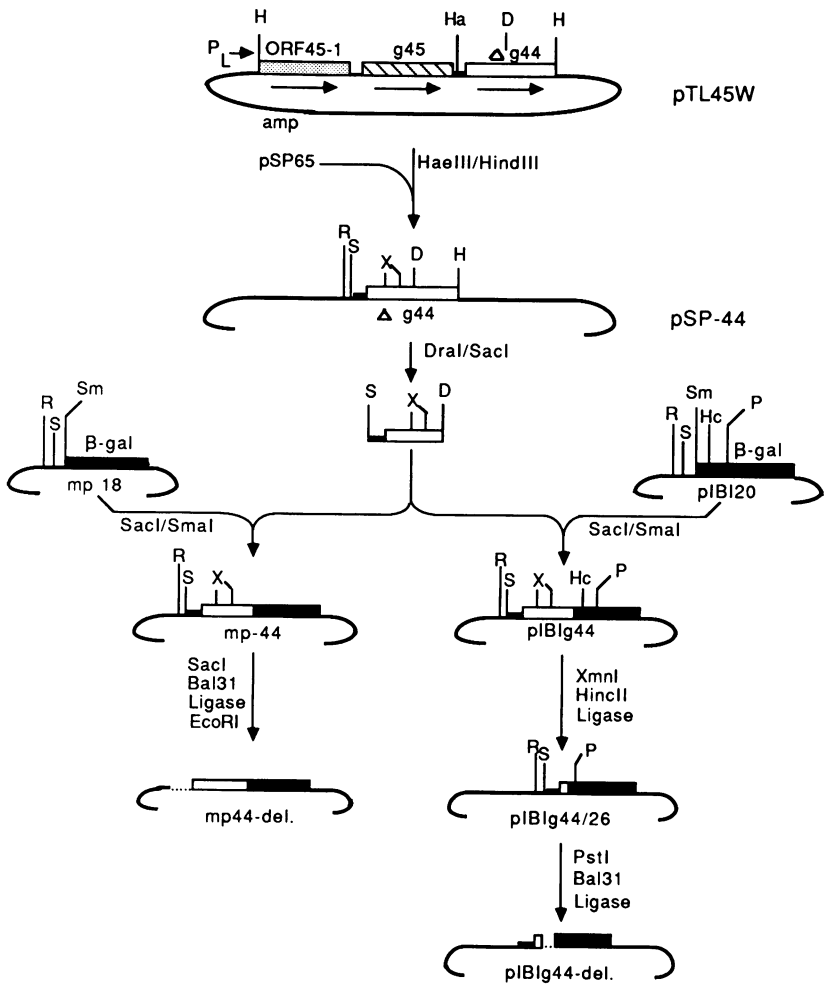


Fig. 3. Scheme used to clone the 5' region of T4 gene 44 (nucleotides -37 to +99) into M13 mp18 and pIBI20, in front of the E. coli lacZ (β-galactosidase) gene. Translation of the gene 44-lacZ mRNA from the mp44 and pIBIg44 vectors originates at the gene 44 start codon. Bal31 deletions were introduced at the 5' end of the gene 44 fragment in mp44 and separately at the 3' end of the gene 44 fragment in pIBIg44 creating two families of deletion mutants. The 5' deletion mutants in mp44 are named mp44/(5' end point position). The 3' deletion mutants in pIBIg44 are named g44/(3' end point position). Restriction enzyme abbreviations are as follows: D (DraI); Ha (HaeIII); Hc (HincII); H (HindIII); P (PstI); R (EcoRI); S (SacI); Sm (SmaI); X (XmnI).

protein in the order of Δ gene 44 >rpbA >α-gt >gene 45 >β-gt. These differences in response confirm our earlier observations of regA90 protein repression of gene 44, rpbA and gene 45 (4).

Defining the Gene 44 mRNA target.

A. Deletion analysis: To examine the nucleotides required for regA protein recognition of gene 44 mRNA, vectors were constructed in which a fragment of gene 44 containing nucleotides -37 to +99 (relative to the AUG start codon) was fused in-phase with the E. coli lacZ gene, in such a way that translation of the fusion mRNA originates at the gene 44 initiation region (Figure 3). Synthesis of the resultant fusion protein was tested for sensitivity to regA repression by using coupled in vitro transcription-translation assays. In one vector, M13 mp44, 5' deletions were made in the region upstream from the gene 44 AUG start codon while the 3' boundary of gene 44 was maintained at +99, as shown schematically in Figure 3. In a second vector, pIBIg44, 3' deletions were made in gene 44, while the 5' end was maintained at -37 (Figure 3). Deletion mutants that maintained the correct reading frame of lacZ were selected by their β-galactosidase α-complementation activity (i.e. plaques and colonies were blue when plated on isopropylthio-β-galactoside and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside). With the M13 mp44 vector, no blue plaques with deletions past nucleotide -11 were obtained, presumably because the gene 44 Shine-Dalgarno sequence is at position -9 to -6. Thus, the requirement for nucleotides -10 to -1 for regA repression could not be tested with the mp44 vector.

As shown in Figure 4, synthesis of the β-galactosidase α-peptide (M_R ≈ 19,000) (lanes 1a-c) produced by M13 mp18 was unaffected by addition of regA protein to in vitro protein synthesis reactions, whereas synthesis of the 44P-β-gal fusion protein (M_R ≈ 23,000) from M13 mp44 (lane 2a) was repressed by addition of increasing concentrations of regA protein (lanes 2b,c). In vitro protein synthesis reactions primed with mp44 DNA with deletion end points at -26 (mp44/-26) and -11 (mp44/-11) (see Figure 4, lanes 3a-c and 4a-c) indicate that in both deletion mutants synthesis of the 44P-β-gal protein remains sensitive to regA protein repression.

Synthesis of the 44P-βgal fusion protein (M_R ≈ 10,500) produced from pIBIg44 was also repressed by the addition of regA protein, as expected, whereas synthesis of β-lactamase was not repressed (Figure 5). Bal31 exonuclease deletions made in plasmid pIBIg44 produced plasmids pg44/26 and pg44/22, which contain nucleotides

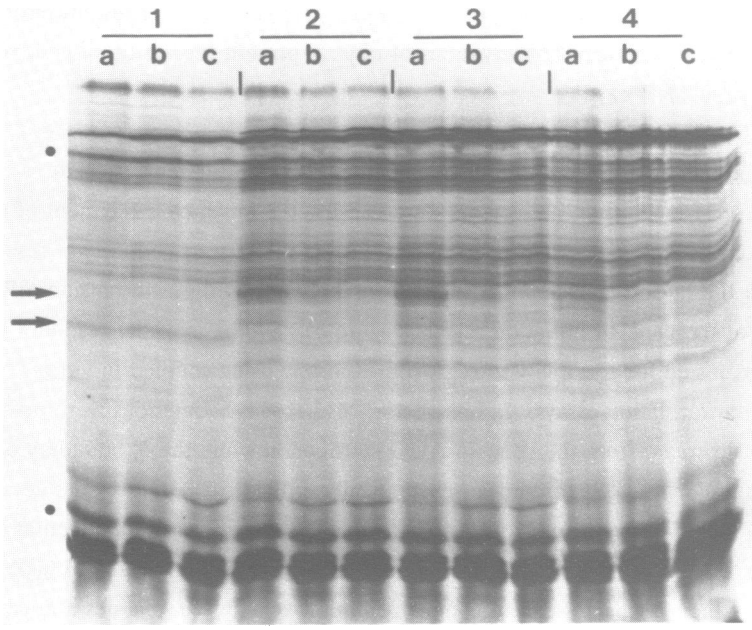


Fig. 4. Effects of regA protein on in vitro protein synthesis from M13 mp44 plasmids. DNA-directed protein synthesis reactions were carried out with replicative form DNA of phage mp18 (lanes 1), mp44 (containing -37 to +99 of gene 44) (lanes 2), mp44/-26 (-26 to +99) (lanes 3) and mp44/-11 (-11 to +99) (lanes 4) in the presence of 0 (a), 7.5 (b) and 15 (c) μM regA protein and ^{35}S-methionine. The top arrow indicates the 44P-β-gal protein made by mp44, mp44/-26, and mp44/-11 (M_R=23,000), and the bottom arrow indicates the α-peptide of β-galactosidase produced by mp18 (M_R=19,000). The top and bottom circles mark the bands presumed to be the M13 proteins II (M_R=45,000) and X (M_R=12,000), respectively (22), which are not repressed by regA protein in these reactions.

-37 to +26 and -37 to +22 of gene 44, respectively. Synthesis of fusion proteins from both plasmids was repressed by addition of 5μM regA protein to in vitro protein synthesis reactions, as shown in Figure 5. Synthesis of a smaller fusion protein produced from a deletion mutant containing nucleotides -37 to +9 of gene 44, pg44/9, also appeared to be sensitive to regA (Figure 5, lanes g44/9). The end points of the Bal31 deletions, determined by sequence analysis, and the sensitivity of the deletion mutants to regA protein repression are summarized in Figure 6.

These results indicated that nucleotides -11 to +22 of gene 44 conferred sensitivity to regA protein and further suggested that nucleotides -11 to +9 may be sufficient for regA recognition. To test this conclusion, a synthetic oligodeoxy-

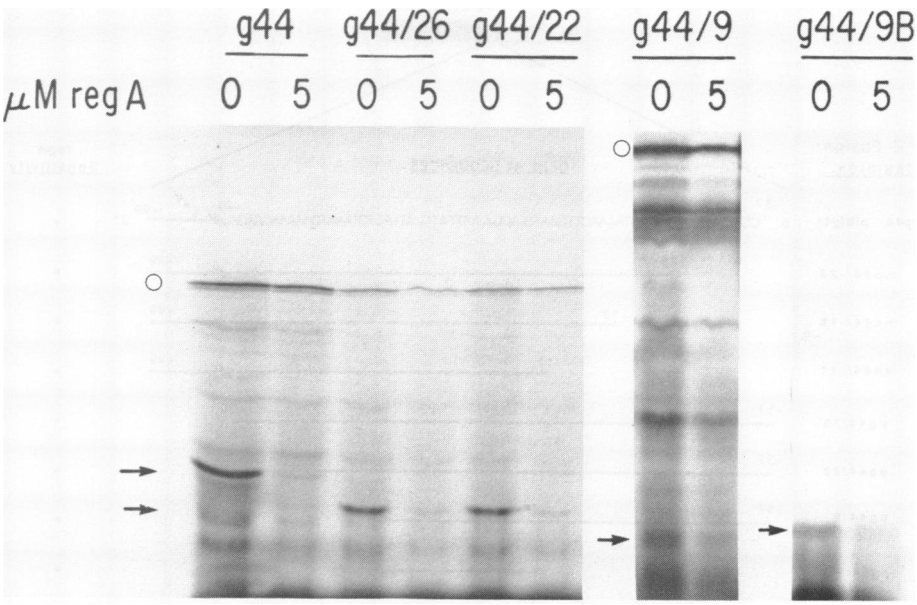


Fig. 5. RegA protein-mediated in vitro repression of protein synthesis directed by pIBIg44 plasmids. Coupled transcription-translation reactions were carried out in the presence of ^{35}S-methionine, and 0 or 5μM regA protein. Reactions were carried out with: plasmid pIBIg44 (containing bases -37 to +99 of gene 44), pg44/26 (containing bases -37 to +26 of gene 44) pg44/22 (-37 to +22), pg44/9 (-37 to +9) and a 200-base transcript synthesized in vitro from plasmid pg44/9B (-11 to +9). The individual fusion proteins (M_R≈7,000-10,500) are indicated by arrows. The open circles indicate β-lactamase (M_R=31,500), which is not repressed by regA protein. Reaction products were separated by electrophoresis on SDS-15% polyacrylamide gels and visualized by autoradiography. The figure is a composite of three gels.

nucleotide corresponding to bases -11 to +9 of gene 44 was placed at the 5' end of the lacZ gene on pIBI20, producing plasmid pg44/9B (see Materials and Methods). To test for translational repression of the small 44P-β-gal fusion protein (M_R=5,800) produced from pg44/9B, a purified in vitro-synthesized RNA transcript was used as a template in protein synthesis reactions. As shown in Figure 5, synthesis of this fusion protein also appeared to be repressed by regA protein.

B. RNase protection assay: To examine in greater detail the nucleotides involved in the interaction of regA protein with gene 44 mRNA, we determined the nucleotides protected by regA protein from digestion with RNase T_1 (which cleaves at Gp↓N [16]), RNase PhyM (which cleaves at Ap↓N, Up↓N, and Gp↓N [17]), and RNase A (which cleaves at Cp↓N and Up↓N).

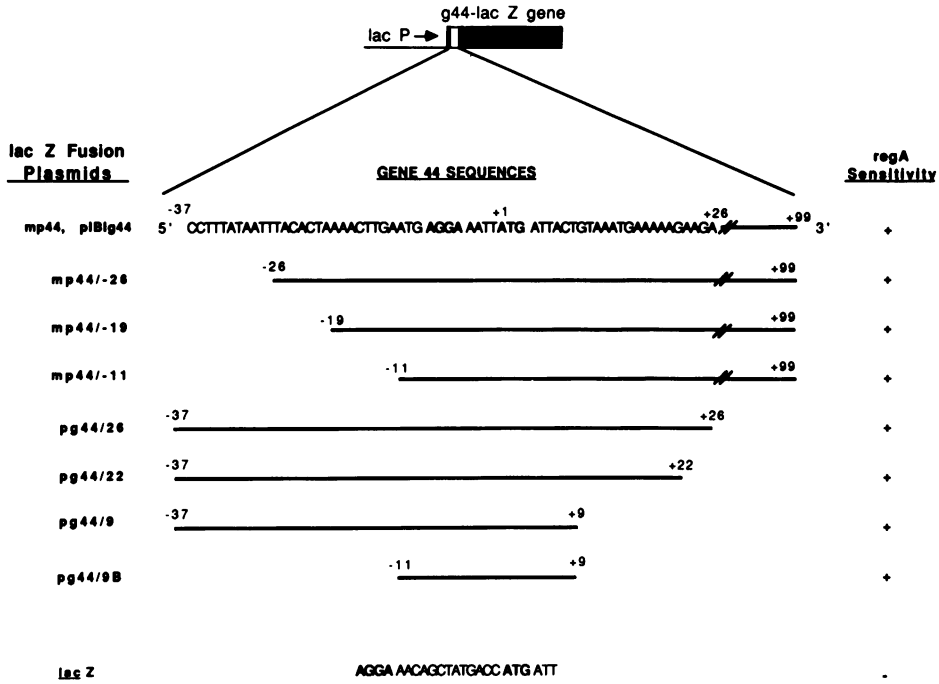


Fig. 6. Summary of the deletion analysis performed on the gene 44-lacZ plasmids. The Shine-Dalgarno region and initiation codon are indicated by boldface type. Mutants with 5' deletions to -11 and 3' deletions to +9 retain sensitivity to regA protein in vitro. A synthetic oligonucleotide containing bases -11 to +9 of gene 44 was cloned into the lacZ gene (pg44/9B) and the fusion gene retained sensitivity to regA protein. The translation initiation region of the lacZ gene, which is not repressed by regA protein, is presented for comparison.

A transcript containing nucleotides -37 to +26 of gene 44 in addition to flanking lacZ sequences was synthesized in vitro from the g44-lacZ gene on plasmid pg44/26 (see Figure 6). The purified 122-nucleotide transcript was end-labeled with γ-^{32}P-ATP and then treated with specific RNases in the presence or absence of regA protein. As shown in Figure 7, regA protein protected guanine residues at positions -9, -7 and -6 from digestion by RNase T_1, whereas cleavages at guanine -13 and guanine +3 were unaffected by regA protein. Low concentrations of regA protein inhibited cleavage by RNase PhyM at nucleotides between -10 and +2 (Figure 7 and 8A). Higher regA protein concentrations increased the protection of these bases from cleavage but did not result in the protection of additional cleavage sites. Similarly, digestion at nucleotides -10 and -1 by RNase A (Figure 7) was reduced in the presence of regA

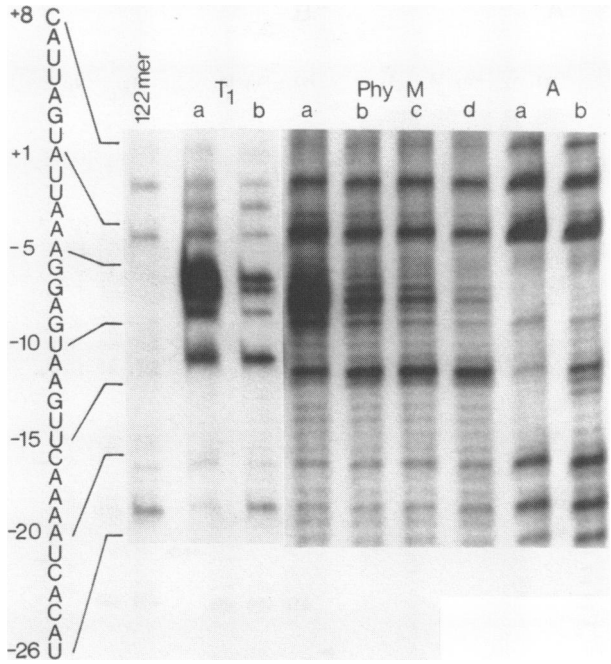


Fig. 7. Ribonuclease protection assay: A 122-base transcript produced from plasmid pg44/26, containing bases -37 to +26 of gene 44 plus flanking lacZ sequences, was 5'-end-labeled with ^{32}P and treated with RNase T_1 (0.01 unit/μl, final concentration), RNase PhyM (1 unit/μl) and RNase A ($3x10^{-4}$ unit/μl) in the presence or absence of regA protein for 5 min. at 0°C.

RNase T_1: a) no regA, b) +5μM regA protein.

RNase PhyM: a) no regA, b) +0.2 μM regA, c) +1μM regA, d) +5μM regA.

RNase A: a) no regA, b) +5μM regA.

For clarity, only bands from the translation initiation region of the transcript are shown here.The sequence of this region is given on the left. Cleavages at residues between nucleotides -10 and +2 were reduced in the presence of regA protein. Cleavages throughout the remainder of the transcript were unaffected by the presence of regA protein (see also Figure 8).

protein, as determined by densitometry of the gel autoradiogram. A number of sites which can be cleaved by RNase A and PhyM became hypersensitive in the presence of regA protein. This may have resulted from disruption of adjacent or local secondary structures by regA protein binding which would enhance cleavage by these single-strand specific RNases.

Similar protection assays were performed with a transcript synthesized from plasmid pg44/9B which contains nucleotides -11 to +9 of gene 44 and flanking lacZ

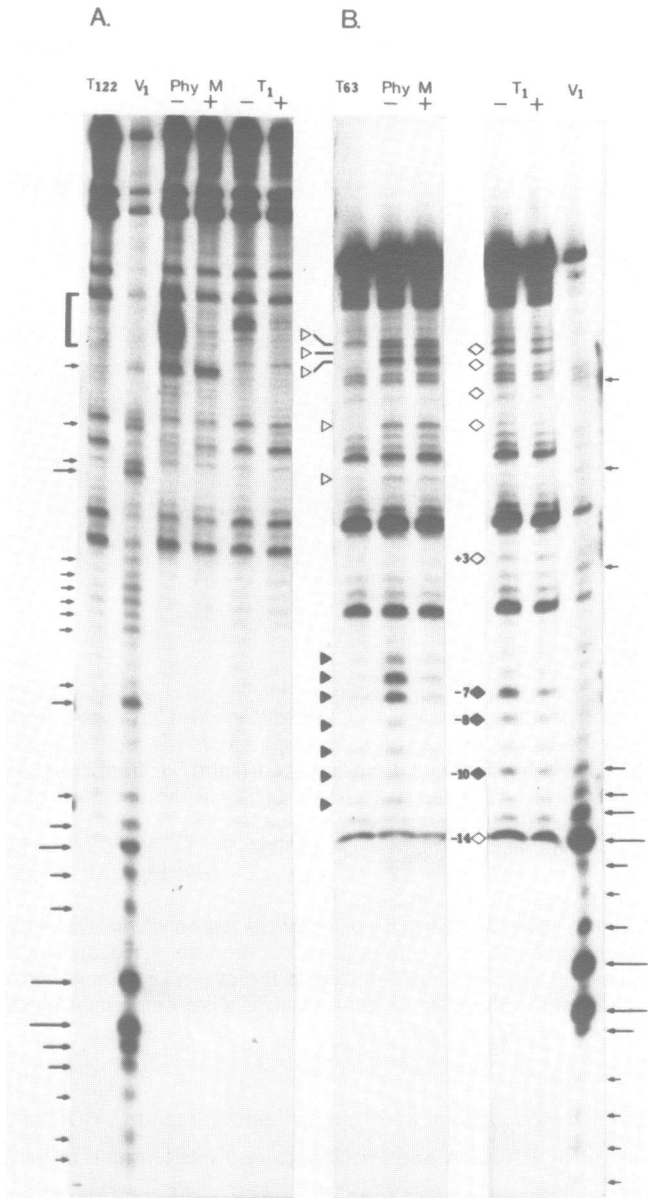


Fig. 8. RNase protection and structural analysis of two gene 44-lacZ transcripts. Protection experiments were performed in the absence (-) or presence (+) of 2μM regA protein. RNase digestion was carried out with 1 unit/μl PhyM or 0.01 unit/μl T_1 for 5 minutes at 0°C. A) The 122-nucleotide transcript produced from pg44/26 containing -37 to +26 of gene 44 and flanking lacZ sequences. The region protected

sequences (see Figure 6). Residues from -10 to -4 were protected by regA protein from digestion by RNases PhyM and T_1, (indicated in Figure 8B by filled triangles and diamonds, respectively), while cleavages at other sites in the mRNA were unaffected (open symbols, Figure 8B). Interactions with bases -3 to +2 could not be assessed due to insufficient RNAse digestion in this region of the transcript. Comparison of the cleavage sites for RNases V_1 and T_1 on both the g44/26 and g44/9B transcripts indicates that regA protein is bound to a single-stranded region of the RNA (Figure 8A and B). On both transcripts, the nucleotides protected by regA protein lie entirely within the region of the gene 44 mRNA that the deletion analysis indicated is sufficient for regA protein recognition and repression in vitro.

DISCUSSION

Using oligonucleotide-directed mutagenesis, the regA90 gene has been converted to the regA$^+$ gene and has subsequently been cloned into the inducible expression plasmid pAS1. Because previous regA$^+$ clones have been found to be unstable (18, 4), or to produce lower levels of regA protein (i.e. through partial derepression [19]) the regA$^+$ clone used in these studies has been thoroughly characterized. Nucleotide and amino acid sequence analysis indicate that the single point mutation in the cloned regA90 gene and its protein product has been converted to wild type. Comparative tryptic peptide analysis is also consistent with the reversion of the mutation and maintenance of the integrity of the remainder of the protein. Comparison of the levels of expression of regA from pAS1regA and pGW7regA plasmids suggests that overexpression of regA is achieved through positioning the regA gene downstream from the strong λ cII gene ribosome binding site and initiation codon. Presumably, translational coupling is responsible for the effective derepression of the regA gene on pAS1regA.

by regA protein, also shown in Figure 7, is indicated by a bracket [. Structural analysis was performed using ribonuclease V_1 (specific for double-stranded RNA) (cleavages are indicated by an arrow (→), the length of which indicates the strength of cleavage) and ribonuclease T_1 (which is single-stranded specific and cleaves at Gp↓N). Note that the region protected by regA protein on both transcripts is cleaved by RNase T_1 and thus is single-stranded. Untreated transcript is labelled T_{122}.
B) A 63-nucleotide transcript produced from pg44/9B containing nucleotides -11 to +9 of gene 44 (5'- UGAGGAAAUUAUGAUUACU -3') and 43 bases of lacZ. Cleavages by PhyM (Ap↓N, Up↓N, and Gp↓N) are indicated by a triangle (△), and protected bases are indicated by filled triangles (▲). RNase T_1 cleavages (Gp↓N) are denoted by a diamond (◇) and protected bases by filled diamonds (♦). Untreated transcript is labelled T_{63}

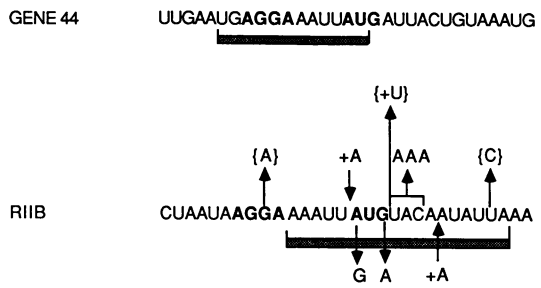
The translational repressor activity of the cloned regA protein was tested on six regA-regulated T4 genes and in each case regA protein in the range of 1 to 10 μM was sufficient for repression. In contrast, no repression of gp32 synthesis was observed with 10μM regA protein. The sensitivity of α-gt and β-gt genes to regA protein in vitro demonstrates they are true regA targets, thus, their overproduction in T4 regA$^-$ infections is not a secondary effect. The protein concentrations required for 50% repression in vitro of gene 44 (0.5μM), rpbA (0.5-1.0μM), gene 45 (2-3μM), and regA (8-10μM) agree well with those reported by Winter *et al.* (5). The concentrations of regA protein required for 50% repression of the α-gt and β-gt genes were approximately 1μM and 4-5μM, respectively. The cellular concentration of regA protein during early stages of T4 infection is estimated to be in the range of 1.5 → 7.5μM (18).

In the studies reported here, we have examined the nucleotides within the gene 44 mRNA that are required for recognition and specific binding by regA protein. The construction of vectors in which the 5' end of gene 44 was fused to the lacZ gene enabled us to examine the nucleotide sequences and/or secondary structures that confer sensitivity to regA protein repression. Deletion of gene 44 sequences to a 5' boundary of -11 and independently to a 3' boundary of +9, produced fusion genes that retained their sensitivity to repression in vitro by regA protein. In addition, an oligodeoxynucleotide containing gene 44 sequences -11 to +9 when inserted into a new context (i.e. the lacZ gene) conferred sensitivity to regA protein. This indicates that sequences within this 20-base segment spanning the translation initiation region of gene 44 mRNA are sufficient for regA protein recognition of the target.

The results of the RNase protection studies were fully consistent with the deletion analysis. RegA protein protected nucleotides -10 to +2 of a g44-lacZ mRNA containing nucleotides -37 to +26 of gene 44 from ribonuclease digestion. When a transcript containing only the 20-base fragment of gene 44 (-11 to +9) and flanking lacZ sequences was examined, specific protection by regA protein was maintained. The protected region overlaps both the Shine-Dalgarno region and the A and U of the gene 44 initiation codon. Digestion with ribonuclease V_1 and T_1 indicates that regA protein is bound to a single-stranded region in both transcripts examined. The target site in the native gene 44 mRNA is predicted by computer analysis to be single-stranded (K. Webster, unpublished data); however, there is no evidence, as yet, that a target must be single-stranded to be bound by regA protein.

Assignment of a consensus sequence which can direct the action of regA protein

A



B

GENE	TRANSLATION INITIATION REGION
GENE 44	UUGAAUGAGGAAAUUAUGAUUACUGUAAAUG
rpbA	GACUAAAGGUGUAUUAUGACUAAAAUUACUG
ALPHA-GT	GAAAGAGGAAAACUAUGCGUAUUUGCAUUU
dexA	GCGAGGAAAAUUUAAUGUUUGAUUUUAUU
RIIB	AAUAAGGAAAAUUAUGUACAAUAUUAAA
GENE 45	AUUGAAGGAAAUUACAUGAAACUGUCUAAAG
BETA-GT	ACAUAAAAGGAAUAUAUGAAAAUUGCUAUAA
GENE 1	AUUUGAGGAGAAACACAUGAAACUAAUCUUUU
REGA	CAUUGGAAUGGUAAAAUGAUUGAAAUUACUC

Fig. 9. A) Comparison of the nucleotide sequence of the translation initiation regions of gene 44 and rIIB. The regA protein binding regions identified by RNase protection of gene 44 and rIIB (5) mRNAs are indicated by boxed underlining. Mutations in rIIB leading to insensitivity to regA are indicated by arrows and mutations that preserved regA sensitivity are shown in brackets ({}) (6, 23). B) Nucleotide sequence of the translation initiation region of nine regA-sensitive genes. All genes presented here have been shown to be sensitive to regA protein repression in simplified in vitro protein synthesis assays. The genes are presented in decreasing order of sensitivity to regA protein in vitro. The Shine-Dalgarno and initiation codons are shown by boldface type, and additional sequence similarities are indicated by overlining. Sequences are from: gene 44 (24); rpbA (25); α-gt (26); dexA (27); rIIB (28), gene 45 (29); β-gt (8); gene 1 (30); regA (2, 31).

has proven difficult. Earlier genetic studies on the rIIB operon have shown that mutations within a 15-nucleotide region overlapping the initiation codon diminish regA protein's ability to repress rIIB gene expression (6,18). The same region of rIIB

mRNA was protected by regA protein from RNase digestion (5). Mutations in the rIIB gene which eliminate regA sensitivity (5, 6, 23) are shown in Fig. 9A. Mutations in the same region of the rIIB mRNA that did not affect regA repression are also shown, in brackets. In ribosome competition assays, the mutations flanking the AUG codon (UA→UAA and UAC→AAA) (Fig. 9) were found to slightly reduce regA protein binding to rIIB mRNA, whereas the initiation codon mutants A→G and G→A were virtually insensitive to regA protein (5).

In comparing the RNase protected regions of rIIB and gene 44 mRNAs, the differences in the sequences are perhaps more striking then the similarities. In the case of gene 44, the Shine Dalgarno region is protected from RNase digestion but the AUG start codon is not fully protected. In contrast, the AUG codon in rIIB mRNA is of central importance and the Shine-Dalgarno region is apparently not involved in regA protein recognition. Nevertheless, regA protein binding to either the Shine-Dalgarno or the AUG start codon region would presumably block ribosome initiation complex formation. However, the sequence(s) that specifies regA protein binding to the two targets is not readily apparent. Comparison of the rIIB and gene 44 sequences with other known regA target genes, shown in Fig. 9B, shows some similarities (indicated by overlining). However, the significance of these similarities in unclear.

Although the regA target(s) is still not clearly defined, the short region of gene 44 mRNA required for recognition is clearly different from that of some other well characterized mRNA targets that form specific RNA structures. For example, ribosome protein S4 recognition of α-mRNA involves interaction with a pseudo-knot structure (15,32) and R17 coat protein repression of the replicase gene involves interaction with a stem and loop structure (33). The small gene 44 mRNA target, on the other hand, does not appear to form a hairpin structure, or pseudo-knot. In this respect regA protein-RNA interactions may be similar to fd gene V protein binding to gene II mRNA, where a 16 nucleotide sequence without notable secondary structure is required for nucleation of gene V protein binding (34). Elucidation of the role of specific nucleotides within the gene 44 and rIIB mRNAs and characterization of the structure of these local RNA regions are essential for understanding the mechanism of regA protein-RNA recognition.

ACKNOWLEDGEMENTS

The authors wish to thank J. DeLuca for excellent technical assistance and to thank W. Chen for performing one of the in vitro protein synthesis assays. We thank W. Rüger (Ruhr Universität) for kindly providing pPLcαgt and pPLcβgt plasmids, D. Mueller for the gift of T7 RNA polymerase, and C. Martin, D. Mueller, K.R. Williams and T.-C. Lin for many helpful discussions. We also thank J. Elliott for his generous help with the HPLC analysis of regA peptides. We are especially grateful to W.H. Konigsberg for his support of this work and for his continuing interest in this project. We thank S. Herzog for her skillful preparation of this manuscript. This work has been supported by National Institutes of Health Grants GM13109 (to E.K.S.) and GM12607 (to W.H.K.).

*Present address: Department of Physiology, Tufts Medical School, 136 Harrison Street, Boston, MA, USA

REFERENCES

1. Wiberg, J.S. and Karam, J.D. (1983) in Bacteriophage T4, Mathews, C.K., Kutter, E.M., Mosig, G. and Berget, P.B. Eds., pp. 193-201. American Society for Microbiology, Washington, D.C.
2. Adari, H.Y., Rose, K., Williams, K.R., Konigsberg, W.H., Lin, T.-C. and Spicer, E.K. (1985) Proc. Natl. Acad. Sci., **82**:1901-1905.
3. Miller, E.S., Karam, J., Dawson, M., Trojanowska, M., Gauss, P. and Gold, L. (1987) J. Mol. Biol., **194**:397-410.
4. Adari, H.Y. and Spicer, E.K. (1986) Proteins. **1**, 116-124.
5. Winter, R.B., Morrissey, L., Gauss, P., Gold, L., Hsu, T. and Karam, J. (1987) Proc. Natl. Acad. Sci., **84**:7822-7826.
6. Karam, J., Gold, L., Singer, B.S., and Dawson, M. (1981). Proc. Natl. Acad Sci. **78**, 4669- 4673.
7. Rosenberg, M., Ho, Y.-S. and Shatzman, A. (1983) Methods Enzymol., **101**:123-138.
8. Tomaschewski, J., Gram, H., Crabb, J.W. and Rüger, W. (1985) Nucl. Acids. Res., **13**:7551-7568.
9. Gillam, S. and Smith, M. (1979) Gene **8**, 81-97.
10. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci., **74**:5463-5467.
11. Mott, J.E., Grant, R.A., Ho, Y.-S. and Platt, T. (1985) Proc. Natl. Acad. Sci., **82**:88-92.
12. Gray, H.B., Ostrander, D.A., Hodnett, J.L., Legerski, R.J. and Robberson, D.L. (1975) Nucl. Acids Res., **2**:1459-1492.
13. Derbyshire, K.M., Salvo, J.J. and Grindley, N.D.F. (1986) Gene, **46**:145-152.

14. Milligan, J.F., Groebe, D.R., Witherell, G.W., and Uhlenbeck, O.C. (1987) Nucl. Acids Res. 15:8783-8798.
15. Deckman, I.C. and Draper, D.E. (1987) J. Mol. Biol., 196:323-332.
16. Donis-Keller, H., Maxam, A.M. and Gilbert, W. (1977) Nucl. Acids Res. 4, 2527-2538.
17. Donis-Keller, H. (1980) Nucl. Acids Res. 8, 3133-3142.
18. Miller, E.S., Winter, R.B., Campbell, K.M., Power, S.D. and Gold, L. (1985) J. Biol. Chem., 260:13053-13059.
19. Liang, Y.L., Wei, R., Hsu, T., Alford, C., Dawson, M. and Karam, J. (1988) Genetics 119:743-749.
20. Russel, M., Gold, L., Morrissett, H. and O'Farrell, P.Z. (1976) J. Biol. Chem. 251:7263-7270.
21. Shamoo, Y., Adari, H.Y., Konigsberg, W.H., Williams, K.R., and Chase, J. (1986) Proc. Natl. Acad. Sci. 83:8844-8848.
22. Fulford, W. and Model, P. (1984) J. Mol. Biol., 173:211-226.
23. Campbell, K.M., Stormo, G.D. and Gold, L. (1983) in Gene Function in Prokaryotes, Beckwith, J, Davis, J. and Gallant, J. Eds. pp. 185-209. Cold Spring Harbor Laboratory, N.Y.
24. Spicer, E.K., Nossal, N.G. and Williams, K.R. (1984) J. Biol. Chem. 259: 15425-15432.
25. Hsu, T. Wei, R., Dawson, M., and Karam, J.D. (1987) J. Virol., 61: 366-374.
26. Gram, H. and Ruger, W. (1985) EMBO J. 4: 257-264.
27. Gauss, P., Gayle, M., Winter, R.B. and Gold, L. (1987) Mol. Gen. Genet., 206 24-34.
28. Pribnow, D., Sigurdson, D.C., Gold L., Singer, B.S. and Napoli, C. (1981) J. Mol. Biol. 149: 337-376.
29. Spicer, E.K., Noble, J.A., Nossal, N.G., Konigsberg, W.H. and Williams, K.R., (1982) J. Biol. Chem. 257: 8972-8979.
30. Broida, J. and Abelson, J. (1985), J. Mol. Biol. 185: 545-563.
31. Trojanowska, M., Miller, E.S., Karam, J., Stormo, G. and Gold, L. (1984) Nucleic Acids Res., 12, 5979-5993.
32. Tang, C.K. and Draper, D.E. (1989) Cell, 57: 531-536.
33. Carey, J., Cameron, V., deHaseth, P.L. and Uhlenbeck, O.C. (1983) Biochemistry, 22: 2601-2610.
34. Michel, B. and Zinder, N. (1989) Proc. Natl. Acad. Sci. 86: 4002-4006.