A characterization of the elements comprising the promoter of the mouse ribosomal protein gene RPS16

Narayanan Hariharan and Robert P. Perry

The Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA

Received February 16, 1989; Revised and Accepted May 22, 1989

ABSTRACT

The elements comprising the mouse rpS16 promoter were characterized by transfection experiments with mutant genes in which various portions of the 5' flanking region and exon I were removed or substituted with extraneous DNA sequence. These experiments were carried out with otherwise intact rpS16 genes transfected into monkey kidney (COS) cells and also with chimeric rpS16-CAT gene constructs transfected into mouse plasmacytoma cells and COS cells. The locations of the functionally important elements were generally correlated with the locations of binding sites for specific nuclear factors, which were identified by gel-mobility shift analyses and methylation interference footprints. The most upstream element, which is located approximately 165 bp from the cap site, binds the Sp1 transcription factor and augments the promoter activity by 2 to 2.5-fold. In addition, there is a complex bipartite element in the -83 to -59 region, an element in the -37 to -12 region and an element in the +9 to +29 region of exon I, all of which are essential for rpS16 expression. The rpS16 promoter has a general architecture that resembles other mouse rp promoters; however, it also possesses some distinctive characteristics.

INTRODUCTION

The biosynthesis of mammalian ribosomes involves the coordinate production of more than 70 different ribosomal proteins (rp's). In exponentially growing cells, the balanced synthesis of the various rp's is mainly attributable to the fact that the corresponding rp mRNAs are present in similar amounts and are functioning with similar translational efficiencies (1,2). In such cells the relatively uniform content of the various rp mRNAs appears to be primarily determined at the transcriptional level (Hariharan, Kelley and Perry, submitted), implying that the rp genes have similar promoter strengths. To understand the basis for this similarity, our laboratory has undertaken a detailed functional analysis of the promoters of three unlinked mouse rp genes, rpL30, rpL32 and rpS16 (3).

All three of these rp genes lack canonical TATA motifs and initiate transcription at a C residue that is embedded in a pure pyrimidine stretch flanked by GC rich sequences (4,5,6). The rpL30 and rpL32 genes have relatively short (38 and 46 bp) first exons which are entirely noncoding, whereas the rpS16 gene has a 101 bp first exon which contains the translational-start site at position +53. Recent studies of the rpL30 and rpL32 genes have localized transcriptional regulatory elements and binding sites for trans-acting factors to regions that span about 80 to 130 bp of 5' flanking sequence and about 80 bp of internal

sequence which includes the first exons and adjacent portions of the first introns (7,8,9; Hariharan, Kelley and Perry, submitted).

In the present study, we have localized the regulatory elements and factor binding sites of the rpS16 promoter. Our analysis indicates that the general architecture of the rpS16 promoter is similar to that of the other rp promoters. However, the rpS16 promoter is distinctive in that the array of elements is shifted further upstream of the cap site and includes an Sp1-binding module.

MATERIALS AND METHODS

Cell culture and transfection protocols.

Monkey kidney (COS-7) and mouse plasmacytoma (S194) cells were grown as monolayer and suspension cultures, respectively, in DME medium (GIBCO, Inc.) containing 10% fetal bovine serum (COS) or 10% horse serum (S194). Plates containing approximately 10^6 COS cells at about 70% confluency were incubated for 30 min at 37°C with 6 ml serum-free culture medium containing 3 µg/ml plasmid DNA (2:1 ratio of rpS16 and rpL30 genes) and 500 µg/ml DEAE-Dextran (Pharmacia, Inc.) (10). The DNA-containing medium was removed, and the cells incubated at 37°C for 3 hr with complete medium containing 0.2 mM chloroquine diphosphate and then for about 40 hr with complete medium. Alternatively, about 2.0 to 2.5×10^7 S194 cells, collected at a density of 5×10^5 cells/ml, were transfected with 4 µg/ml plasmid DNA in the presence of 0.5 mg/ml DEAE-Dextran, treated for 30 min with 0.1 mM chloroquine diphosphate, and then incubated in normal medium for about 48 hr (11).

Analysis of RNA and DNA.

Total cytoplasmic RNA was isolated from COS cells and analysed by the S1 nuclease protection method as described elsewhere (8). Nuclear RNA was isolated by the hot-phenol procedure (4). The 5' end-labeled probe used for assays of rpS16 expression is diagrammed in Fig. 1 (top). For assays of rpL30 expression we used a 266 bp 5' end-labeled probe which yields a 61 nucleotide fragment when protected by properly initiated and spliced rpL30 transcripts (Hariharan, Kelley and Perry, submitted). The fragments were electrophoresed on 8 to 10% polyacrylamide-urea gels and visualized by autoradiography. For quantitation, selected exposures of the autoradiograms were densitometrically scanned under conditions of linear dose-response. Preliminary experiments established that the signals were proportional to the amount of RNA used in the assay.

Plasmid DNA was isolated from transfected cell nuclei by the Hirt procedure (12) and quantitatively assayed by Southern-blot analysis with a nick-translated fragment A probe (see below).

Chloramphenicol acetyltransferase (CAT) assays.

Extracts containing 250 µg of protein were assayed with ^{14}C -chloramphenicol as described by Gorman *et al.* (13).

Nuclear extract preparation gel-retardation assays and methylation interference footprinting.

Nuclear extracts from S194 and COS-7 cells were prepared according to Dignam *et al.* (14) with additional protease inhibitors (leupeptin and pepstatin) added to all solutions at 1 µg/ml. Gel retardation assays were carried out as described by Singh *et al.* (15). For competition experiments, the unlabeled DNA fragments were added in a 50 to 75-fold molar excess over the labeled DNA probes. Methylation interference footprints (16) were obtained with fragments that were end-labeled on the sense or antisense strands.

Plasmid construction.

All cloning procedures employed the pUC18/HB101 vector/host system. A Bam HI-Sac I fragment extending from position -400 to 107 bp beyond the poly(A) site of the rpS16 gene (6) was inserted into pUC18 by a two-step cloning procedure which destroyed the 3' Sac I site but retained an internal Sac I site at +29. A Sac I-Hinc II or Sac I-Sph I digest of this plasmid produced fragments, termed A_H or A_S, respectively, which lack 5' flanking sequences and 28 bp of exon I. An Eco RI-Sac I fragment containing the -580 to +29 region was digested with an assortment of restriction enzymes to produce an array of 5' fragments that were joined to fragment A_H at the Sac I site and blunt-end ligated into the vector Hinc II site. The 5' boundaries of constructs a, b, c, d and k were defined by restriction sites for Eco RI, Bam HI, Ava II, Dde I and Fsp I, respectively, after appropriate filling or trimming with the Klenow enzyme. Constructs e, f, g, h, i and j were made as follows. Construct d was cut at the unique Sph I site in the 5' polylinker, digested with Bal 31 nuclease (0.5 units/10 µg DNA) for 2 min at 30°C, and the reaction stopped with EGTA (17). The DNA was extracted, the ends filled by reaction with the Klenow enzyme, and then digested with Sac I. The fragments containing varying amounts of 5' flank were then joined to fragment A_H as described above. Individual clones were selected and the 5' ends determined by DNA sequencing. Construct c was used to produce internal deletion constructs l, m and n. For l, digestion with Sph I, Mnl I and Sac I produced a -179 to -147 fragment and a -97 to +29 fragment which were isolated and ligated to fragment A_S by directional cloning. Constructs m and n were produced similarly except that for m, a -179 to -170 Sph I/Alu I fragment was substituted for the -179 to -147 fragment, and for n, a -37 to +29 Fsp I/Sac I fragment was substituted for the -97 to +29 fragment. Construct o was made by ligating a synthetic oligonucleotide representing the -85 to -67 sequence to the -37 to +29 fragment and then joining this segment to fragment A_S. For construct p, a blunt-ended 32 bp Eco RI-Hinc II fragment from the polylinker of pUC18 was added to the ligation mixture; sequencing showed the orientation of this fragment to be 5'-Eco RI-Hinc II-3'. For construct q, the -85 to -67 oligonucleotide was joined to construct j at the Sph I polylinker site which is 10 bp removed from the -58 position. For construct r, an Sph I/Sac I fragment derived from construct c, which contains 179 bp of 5' flank, was digested either with Fsp I or Hae III. The Sph I-Fsp I segment (-179 to -38), the Hae III-Sac I segment (-11 to +29), and an

Fig. 1: Schematic representation of the rps16 gene and the various mutants used for characterization of its promoter elements. At the top is the complete gene drawn to scale. Exons are shown as boxes containing Roman Numerals; introns and flanking sequences are shown as thin lines. The -580 to +120 region is expanded with the first exon shown as a solid bar and the location of the ATG translational start codon indicated; the 167 bp probe used in the S1 nuclease protection assays and the 68 nucleotide fragment that is protected by correctly initiated transcripts are also shown. The 5' and internal deletion mutants are aligned below; the zig-zag lines indicate regions that are filled with extraneous DNA sequences. Mutants *a* through *r* consist of the complete rps16 gene with 107 bp of 3' flanking sequence. At the bottom are constructs in which the rps16 promoter is linked to the CAT reporter gene (stippled bars). The relative expression as percent of maximum is shown at right. These values, obtained from data such as those illustrated in Figs. 2 and 3, represent averages of two or more independent transfection experiments.

Fig. 2: Expression of rpS16 5' flank mutants in transfected COS cells. Top panels: assay of rpS16 transcript levels by S1 nuclease protection of the probe diagrammed in Fig. 1. Arrows indicate the 68 nucleotide protected probe fragment. Bottom panels: S1 assay of transcripts from the co-transfected rpL30 gene. Each lane represents the assay of 100 μ g of cytoplasmic RNA from COS cells transfected with the indicated mutant or not transfected (COS). A. Survey of 5' deletion mutants. B. Dissection of the -179 to -38 region. C. High resolution analysis of the -99 to -76 region. D. Effects of deletions and substitutions in the -83 to -38 region. E. Effect of a substitution of the -37 to -12 sequence. The lane marked M in panel A contains a marker of *Msp* I digested pBR322 DNA labeled with the Klenow enzyme; lane 1 contains the fragment protected by 10 μ g of cytoplasmic RNA from S194 cells (Mouse).

oligonucleotide (CCAACCCAGCCGAGGAGCAGGGTGC), which substituted for the -37 to -12 region, were joined to fragment A₅.

Constructs *c*-, *d*-, *g*-, *h*-, *i*- and *j*-CAT were made as follows. Constructs *c*-, *d*-, *g*-, *h*-, *i*- and *j*- were digested with *Sac* I, the ends blunted with Klenow enzyme and the plasmids redigested with *Hind* III, which cuts in the 5' polylinker. These fragments were inserted into the *Hind* III and blunt-ended *Xba* I sites of the CAT containing plasmid p106 (7). These constructs contained the first 29 bp of the rpS16 gene linked to a site approximately 40 bp 5' of the CAT translational start codon. For constructs *g*-, *i*- and *u*-CAT, a *Hind* III-*Sac* I fragment was isolated from construct *c* and cleaved with either *Hae* II (+15), *Hpa* II (+9) or *Hae* III (-12). The 5' portions of these cleavage products were gel-purified and introduced into the p106 plasmid as described above. For the *Hae* II and *Hpa* II fragments, we used a multi-step protocol in which the *Hind* III end of the fragment was ligated to the *Hind* III site of the vector polylinker, the *Hae* II and *Hpa* II ends respectively trimmed or filled by the Klenow enzyme and then the blunt ends ligated to the blunt-ended *Xba* I site of the vector.

RESULTS

Functional characterization of the rpS16 promoter.

To identify the functionally important 5' flanking regions of the mouse rpS16 gene, we prepared a series of deletion or substitution mutants (a through r, Fig. 1) and studied their expression in transfected monkey kidney (COS-7) cells. We used a quantitative S1 nuclease protection analysis to assay the relative levels of rpS16 mRNA and to verify correct transcriptional initiation. A 167 bp S1 probe which spans the rpS16 cap site (Fig. 1) yields a 68 nucleotide fragment when protected by authentic rpS16 mRNA (Fig. 2A, lane 1). Under the conditions of these assays, the background from endogenous monkey rpS16 mRNA is negligible (Figs. 2A, C, lane 7; 2D, lane 1; 2E, lane 1). For each experiment the expression of a co-transfected intact rPL30 gene was also assayed by S1 protection analysis (bottom panels) in order to verify the equivalence of RNA yields and transfection efficiencies. Analysis of plasmid DNA in Hirt supernatant fractions from transfected cells confirmed that the cells had taken up similar quantities of the various rpS16 mutant constructs (data not shown). An initial set of experiments (Fig. 2A), indicated that all of the elements needed for efficient expression of the rpS16 gene are contained within 179 bp of 5' flanking sequence (lanes 2-4). The expression was 40 to 45% of maximum with 99 bp of 5' flank (lane 5) and essentially undetectable when only 37 bp of 5' flank was present (lane 6). The same results were obtained when either nuclear or cytoplasmic RNA was analysed.

Internal deletion mutants (l and m, Fig. 1) were used to localize more precisely the most upstream element of the promoter. The expression of mutant l, which lacks sequences from -146 to -97, was equivalent to that of the -179 construct, whereas the expression of mutant m which lacks sequences from -169 to -97 was reduced to the level of the -99 construct (Fig. 2B, lanes 1-4). This indicates that the most upstream element lies between -179 and -147 and suggests that the sequences between -146 and -97 may be dispensable. As will be shown below, the -179 to -147 sequence contains a binding site for the Sp1 transcription factor. A mutant n, which contains the -179 to -147 region but lacks sequences between -146 and -38, is completely inactive (Fig. 2B, lane 5), indicating that the stimulatory effect of the upstream element can be realized only in association with elements located between -97 and -38.

An additional set of constructs was used to characterize the elements in the -97 to -38 region. The expression of a set of 5' deletion mutants generated by BAL-31 resection (e through j, Fig. 1) clearly demonstrated that the 5' boundary of such elements lies between -83 and -76 (Figs. 2C and D and Fig. 1). A construct containing the -85 to -67 region but lacking the -66 to -38 sequence (o, Fig. 1) was inactive, whereas a construct in which the -66 to -38 sequence was replaced by an extraneous pUC18 sequence (p, Fig. 1) was expressed at about 1/4 the level of the -83 mutant (Fig. 2D, lanes 2-4). The same low level of expression (about 10% compared to the complete rpS16 gene) was observed with a similar construct (q, Fig. 1) in which the only substituted rpS16 sequence was the segment between -66 and -59 (Fig. 2D, lane 5). These results indicate that the region between -83 and -59 is essential for the expression of the

Fig. 3: Expression of chimeric rpS16-CAT genes in transfected S194 cells. A. Analysis of 5' flank mutants. B. Analysis of exon I and cap site mutants. C. Comparison of rpS16 and rpL30 promoters: the rpS16 construct is c-CAT; the rpL30 construct contains 205 bp of 5' flank and 167 bp of internal sequence which includes the first exon, the (splicable) first intron and a noncoding portion of the second exon. Mock transfections were done with the promoterless vector p106 (7).

rpS16 gene. More than one element may be located in this region, as suggested by the reduction in expression observed when either the -83 to -76 segment (g vs. h) or the -66 to -59 segment (g) is eliminated. The roughly equivalent expression observed with constructs p and q indicates that there is probably no important regulatory element in the -58 to -38 region.

At least one additional critical 5' flanking element is located between -38 and the cap site. This was demonstrated with a construct (r, Fig. 1) in which the -37 to -12 region was replaced by an extraneous DNA sequence. This construct exhibited a markedly depressed activity when compared to a counterpart bearing the normal -37 to -12 sequence (Fig. 2E, lanes 2 and 3).

The relative contribution of some of the 5' flanking elements to the efficiency of the rpS16 promoter was also assessed with constructs in which the rpS16 promoter was linked to a chloramphenicol acetyltransferase (CAT) reporter gene (c-, d-, g-, h-, i- and j-CAT Fig. 1). The CAT gene was fused to exon I of the rpS16 gene, 29 bp downstream of the cap site and 23 bp upstream of the ATG initiation codon. With these constructs we could also study the expression of mutant genes in mouse myeloma (S194) cells because the activity of the endogenous mouse rp genes is not detected by the CAT assay. In relatively good agreement with the results described above, we observed 45% of maximum expression

Fig. 4: Nuclear factor binding analysis of rpS16 fragments covering the -179 to -38 region. The labeled fragments indicated below the panels were incubated with (+) or without (-) nuclear extract from S194 cells, electrophoresed on 5% polyacrylamide gels and visualized by autoradiography. The third lane in each set shows the effect of adding a 50 X molar excess of unlabeled DNA consisting of either a *Pvu* II-*Hind* III 340 bp fragment containing the 21 bp repeats of SV40 or an Sp1 consensus decamer GCTCCGCCCC.

when the -179 to -83 region was removed and a substantial further decrease when all sequences upstream of -76 were eliminated (Fig. 3A, lanes 1-4). A residual activity of about 10% was observed with the *h*- and *j*-CAT constructs, which respectively possess 76 and 66 bp of 5' flank; however, with the *j*-CAT construct, which contains only 58 bp, expression was reduced to a negligible level (Fig. 3A; lanes 4-6). The fact that no residual activity was observed with the corresponding intact gene constructs *h* and *j* (see, for example Fig. 2C; lane 6) might be due to a greater sensitivity of the CAT assay for detecting very low levels of transcript or to subtle differences between the S194 and COS cells. In any event, these results demonstrate that for the most part, the dependence on the various 5' elements is comparable for the complete rpS16 gene and its isolated promoter region, and that the contributions of these elements are generally similar in cells of mouse lymphoid and monkey kidney origin.

Another set of CAT constructs (*g*-CAT to *u*-CAT, Fig. 1) was used to examine the importance of exon I sequences for rpS16 expression. The expression of constructs that contained all of the functionally important 5' flanking elements was reduced by about 40% when the sequences between +16 and +29 were deleted and was totally abolished when the sequences between +10 and +29 were removed (Fig. 3B, lanes 2-4). Elimination of the rpS16 cap site also completely abolished expression (Fig. 3B, lane 5). The same results were obtained in a parallel experiment with COS-7 cells (data not shown). These results

Fig. 5: Localization of factor binding sites in the -99 to -38 region of rpS16. A. Gel-mobility shift assay of a series of fragments with a fixed 3' boundary (-38) and a variable 5' boundary (indicated above lanes). B. Specificity of factor binding. A labeled -83 to -38 fragment was incubated with extract alone (lane 2) or in the presence of a 50 X molar excess of unlabeled fragment from either the -66 to +29 region (lane 3) or the -58 to +29 region (lane 4). C. Effect of substitution of -66 to -59 sequence on factor binding. -85(Δ-66-59▼) is a fragment in which the -66 to -59 region is replaced by the polylinker sequence used in construct *q*. Data shown were obtained with nuclear extracts from S194 cells; identical results were obtained with extracts from COS cells.

indicate that the region of exon I between +10 and +29 is also important for expression of the rpS16 gene. While these experiments do not directly demonstrate a transcriptional role for this exon I segment, the fact that it specifically binds nuclear factors (see below) suggests that it may indeed be an integral part of the promoter structure (see Discussion). A similar requirement for exon I sequences has also been observed for the rpL32 gene (7,9) and the rpL30 gene (Hariharan, Kelley and Perry, submitted). We have not examined the effect of removing rpS16 sequences downstream of +29. However, it is unlikely that such downstream sequences contribute substantially to the rpS16 promoter activity because the level of expression of the -179 to +29 construct (*g*-CAT) is similar to that of the rpL30 and rpL32 promoters, for which the role of internal sequences has been rigorously evaluated (8, Hariharan, Kelley and Perry, submitted). A typical example of such a comparison is shown in Fig. 3C. In this experiment the activity of the *g*-CAT construct was about 75% of that of a CAT gene linked to the complete rpL30 promoter.

In summary, we have localized several (four or five) discrete *cis*-acting elements in the rpS16 gene that contribute to its promoter activity. These elements (A through E) are diagrammed schematically in Fig. 8.

Fig. 6: Methylation interference analysis of the C region binding site. A partially methylated fragment corresponding to the -83 to -38 region, 5'-end labeled on either sense or antisense strand, was used for a gel mobility shift analysis as shown in Fig. 5. The bound (B) and unbound (UB) fragments were isolated, cleaved with piperidine, and analysed on a sequencing gel together with (G+A)-cleaved and G-cleaved marker fragments. The analysis of the sense strand is shown. Similar data were obtained with the antisense strand (not shown). The sequence at bottom shows the fully protected (full circles) and partially protected (half circles) guanine residues.

Binding sites for nuclear factors in the rpS16 promoter region.

Given the array of functional elements described above, it was of considerable interest to determine whether the locations of these elements coincided with the positions of nuclear factor binding sites. Gel retardation analysis was used to identify these factors in nuclear extracts prepared from both S194 and COS-7 cells. To explore the region upstream of -38 we initially used a -179 to -99 fragment, which contains element A, a -179 to -38 fragment, which contains regions A, B and C, and a fragment containing sequences from -179 to -147 and from -96 to -38, which also encompasses regions A, B and C. Element A contains the sequence GCTCCGCCCC at position -170 to -161. This sequence conforms to the consensus for high-affinity binding sites of the transcription factor Sp1 (18). All three fragments bound to a nuclear factor that could readily be competed for by the 10 bp Sp1 consensus oligomer or by the 21 bp repeat

elements of SV40, which contain multiple Sp1 binding sites (Fig. 4). These results indicate that element A does indeed contain an effective Sp1 binding site. The fragments containing the B and C elements exhibited an additional retarded band which was not eliminated by the Sp1 competitors, suggesting that this region might contain a binding site for another nuclear factor. The fact that identical results were obtained with fragments that either contained or lacked the -146 to -97 region suggests that this region is devoid of factor-binding sites.

The B-C region was systematically surveyed for binding sites with a series of fragments that have varying 5' ends (-99 to -53) and a common 3' end (-38). All fragments that contained the -66 to -38 region effectively bound to a nuclear factor in both S194 and COS cell extracts, whereas the fragments representing the -58 to -38 region did not (Fig. 5A). The specificity of this binding was demonstrated by a competition experiment: binding to a -83 to -38 fragment was eliminated by a -66 to +29 competitor, but not a -58 to +29 competitor (Fig. 5B). These results localize a binding site to the -66 to -58 region, which corresponds to the regulatory element designated as C. No factor that specifically binds to the B element was detected in these studies, despite the fact that this region (-83 to -76) is essential for rpS16 expression. Such a factor might be expected to generate an additional retarded band with the fragments that extend to -83 and -99, but no additional band was observed (compare lanes 2 or 4 vs. 8, Fig. 5A). It would also be expected to cause retardation of a fragment that contains the B region but has the C region replaced by vector sequences, and this was not detected (Fig. 5C). The C region mobility shift usually appeared as a doublet of closely spaced bands. This doublet was observed whether or not the B region was present on the labeled fragment (compare Fig. 5A, lane 8 with Fig. 5B, lane 2) and also when the labeled -83 to -38 fragment was assayed in the presence of excess unlabeled B region oligonucleotide competitor (data not shown). The fact that both retarded bands map to the C region binding site suggests that there may be two variants of the C region factor, either distinct proteins with similar binding specificities or modified forms of the same factor.

Methylation interference analysis was used to identify the guanine contact residues in the C region binding site (Fig. 6). Five contact residues were identified in the sequence CGTGACG which extends from -63 to -57. This is in perfect agreement with the results of our mobility shift analysis.

The region corresponding to element D also has a factor-binding site, as seen by the retardation (band a) of fragments containing the -40 to -12 sequence (Fig. 7A, B). This binding is specific for sequences in element D as evidenced by the fact that the binding is completely competed out by a -37 to -12 fragment but not by a -11 to +29 fragment (Fig. 7B).

Element E in exon I also exhibits factor binding as seen by the additional retarded bands (b-e) observed when the -40 to +7 fragment is extended to +15 or +29 (Fig. 7A) or when a -11 to +29 fragment is used for the binding assay (Fig. 7C). In addition to bands b through e, the -11 to +29 fragment exhibited two slowly migrating bands that partially overlap with the D element-specific band, a. All of these bands

Fig. 7: Localization of factor binding sites in the -40 to +29 region of rpS16. A. Gel-mobility shift assay with S194 cell nuclear extract of a series of fragments with a fixed 5' boundary (-40) and a variable 3' boundary (indicated above lanes). The various retarded bands are designated a through e. B. Specificity of factor binding to the -40 to -12 region. A labeled -40 to -12 fragment was incubated with extract alone (lane 2) or in the presence of a 75 X molar excess of unlabeled fragment from either the -11 to +29 region (lane 3) or the -37 to -12 region (lane 4). C. Specificity of factor(s) binding to the -11 to +29 region. A labeled -11 to +29 fragment was incubated alone (lane 2) or in the presence of a 50 X molar excess of unlabeled fragment from either the -11 to +29 region (lane 3) or the -66 to -38 region (lane 4).

are eliminated by unlabeled -11 to +29 competitor, but not by a -66 to -38 competitor (Fig. 7C), attesting to the specificity of the binding reactions. The reason for the multiplicity of retarded bands seen with the E region fragment is presently unclear. It might be due to the binding of multiple factors that individually recognize sequences in the +8 to +29 region or that associate *via* protein-protein interactions. Alternatively, it might represent the binding of several different derivatives of a single factor. The factors binding to the C,D and E regions are clearly different from each other, as demonstrated by the failure of the -58 to +29 fragment to compete for C region binding (Fig. 5B, lane 4) and by the lack of competition between the D region and the -11 to +29 E region fragment (Fig. 7B, lane 3).

DISCUSSION

The foregoing experiments have revealed the general characteristics of the mouse rpS16 promoter. This promoter consists of several discrete *cis*-acting elements which specifically bind to at least four different nuclear factors (Fig. 8). The most upstream element A, located approximately 165 bp from the cap site, binds the Sp1 transcription factor. Although this element is not absolutely required for promoter activity, the promoter operates at only 40 to 45% of maximum efficiency without it. A second element,

Fig. 8: Summary of functional analysis of rpS16 promoter elements and the localization of nuclear factor binding sites. The promoter region is shown with the 5' flank as a line and exon I as a solid rectangle; the transcriptional start site (cap site) is indicated by a horizontal arrow. The elements localized by transfection experiments with mutant genes are designated A through E and enclosed in brackets. Regions with no apparent function are represented by dotted lines. The regions containing nuclear factor binding sites are shown as open rectangles; X signifies that this region may bind more than one factor.

B-C, is located 59 to 83 bp upstream of the cap site. Constructs that lack B-C are basically inactive, whether or not element A is present. This element appears to be complex, consisting of an upstream portion, B (-83 to -76), and a downstream portion C (-66 to -59). Constructs that lack either the B or the C portion are poorly expressed (not more than 10% of the maximum level). Curiously, we were able to detect nuclear factor binding only to the C portion of this complex element. If there is another factor that associates with the B region, its affinity is apparently insufficient to permit detection by our binding assay. In the living cell, protein-protein interactions with other factors might conceivably help stabilize a relatively weak interaction between such a factor and the B element. A third element D, located 12 to 37 bp upstream of the cap site, specifically binds a nuclear factor and is also required for efficient expression of the rpS16 gene. This element may function similarly to the TATA motif of other eukaryotic genes, although in rpS16, like many other mouse rp genes studied to date, this motif is highly degenerate (AAAAAT at positions -28 to -23).

An additional element, E, is located in exon I between positions +11 and +29. The fact that this element specifically binds one or more nuclear factors has led us to suppose that it has a transcriptional function, presumably as an integral part of the rpS16 promoter. The presence of a transcriptional regulatory element downstream of the cap site may be a common feature of mammalian rp genes. The mouse rp genes encoding L32 and L30 also contain sequences downstream of the cap site that bind nuclear factors and are critical for promoter activity (7,8,9; Hariharan, Kelley and Perry, submitted). In the case of rpL32, these sequences were shown to be important for the expression of the complete rp gene as well as for the activity of the isolated promoter region; moreover, the transcriptional role of internal rpL32 sequences was directly demonstrated by nuclear run-on assays.

There are some sequences in the region between A and E that apparently do not contribute to the rpS16 promoter activity and correspondingly do not contain any binding sites for nuclear factors. As demonstrated by construct J, a large portion of the segment separating elements A and B (-146 to -97) can be deleted without any significant effect on rpS16 expression, indicating that this segment is probably

dispensable. Although we cannot rigorously exclude the possibility that when element A is brought closer to the cap site, it precisely compensates for the loss of an element in the -146 to -97 region, the absence of factor-binding sites in this region tends to argue against the existence of such an element. The sequence separating elements C and D (-58 to -38) may also be devoid of any important regulatory elements, as indicated by the roughly equivalent expression of constructs *p* and *q*. The fact that deletion of this region significantly reduces rpS16 expression compared to a substitution by extraneous DNA sequence, suggests that the spacing between C and D is important for proper functioning of the promoter. We may suppose by analogy to other promoters (19) that interactions between factors binding to the various elements is an essential feature of rp promoter function.

The three mouse rp promoters analysed to date, rpL30, rpL32 and rpS16, have a similar architecture. Each is composed of at least five discrete *cis*-acting elements distributed over a ~200 bp region spanning the cap site. The array of elements in rpS16 is shifted about 50 to 90 bp upstream compared to those in rpL30 and rpL32. Conceivably, this shift might be related to the relative proximity between the start sites for transcription and translation. In rpS16 these sites are only 52 bp apart, whereas in rpL30 and rpL32 they are separated by 178 and 840 bp, respectively. Although it is theoretically possible for a DNA sequence to serve both a transcriptional and a coding function, the constraints on the evolution of such a sequence would be formidable.

Most of the *cis*-acting elements have been shown to specifically bind distinct nuclear factors, and we presume that these factors are needed for the proper functioning of the elements. Although rpL30 and rpL32 have at least two elements in common, there is presently no clear evidence for any overlap with the set of rpS16 elements. An example of an rpS16 distinctive element is A, which binds the Sp1 transcription factor and is responsible for a 2.2- to 2.5-fold stimulation of promoter activity. An Sp1-binding element is not present in either the rpL30 or rpL32 promoters, although these genes also contain upstream elements that boost transcriptional activity over a basal level. It is noteworthy, therefore, that the different rp genes are transcribed at relatively uniform rates despite the fact that their promoters are composed of different sets of modular elements. The "promoter strengths" probably derive from the particular combinatorial arrangement of the multiple elements rather than from the nature of the elements, themselves. This feature may be especially well-suited to genes that have to be expressed efficiently in all types of tissue. To date we have not detected any marked qualitative or quantitative differences in rp nuclear factors among proliferating cells derived from different tissues and different mammalian species, suggesting that this general promoter architecture may be evolutionarily conserved in the higher eukaryotes.

ACKNOWLEDGEMENTS

This research was supported by grants from the National Science Foundation (DCB84-13609), the National Institutes of Health (AI-17330, CA-06927) and an appropriation from the Commonwealth of Pennsylvania.

REFERENCES

1. Meyuhas, O. and Perry, R. P. (1980) *Gene* **10**, 113-129.
2. Meyuhas, O., Thompson, E. A. and Perry, R. P. (1987) *Mol. Cell. Biol.* **7**, 2691-2699.
3. Wiedemann, L. M., D'Eustachio, P., Kelley, D. E. and Perry, R. P. (1987) *Somatic Cell and Molecular Genetics* **13**, 77-80.
4. Dudov, K. and Perry, R. P. (1984) *Cell* **37**, 457-468.
5. Wiedemann, L. M. and Perry, R. P. (1984) *Mol. Cell. Biol.* **4**, 2518-2528.
6. Wagner, M. and Perry, R. P. (1985) *Mol. Cell. Biol.* **5**, 3560-3576.
7. Atchison, M. L., Meyuhas, O. and Perry, R. P. (1989) *Mol. Cell. Biol.* **9**, 2067-2074.
8. Chung, S. and Perry, R. P. (1989) *Mol. Cell. Biol.* **9**, 2075-2082.
9. Moura-Neto, R., Dudov, K. P. and Perry, R. P. (1989) *Proc. Natl. Acad. Sci. USA* (in press).
10. Bienz, M. and Pelham, H. R. B. (1986) *Cell* **45**, 753-760.
11. Grosschedl, R. and Baltimore, D. (1985) *Cell* **41**, 885-897.
12. Hirt, B. (1967) *J. Mol. Biol.* **26**, 365-369.
13. Gorman, C. M., Moffat, L. F. and Howard, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044-1051.
14. Dignam, J. D., Lebovitz, R. M. and Roeder, R. G. (1983) *Nucl. Acids. Res.* **11**, 1475-1489.
15. Singh, H., Sen, R., Baltimore, D. and Sharp, P. A. (1986) *Nature* **319**, 154-158.
16. Siebenlist, U. and Gilbert, W. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 122-126.
17. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, Cold Spring Harbor Laboratory, pp. 1-545.
18. Kadonaga, J., Jones, K. and Tjian, R. (1986) *Trends Biochem. Sci.* **11**, 20-23.
19. McKnight, S. and Tjian, R. (1986) *Cell* **46**, 795-805.