MOLECULAR AND CELLULAR BIOLOGY, Apr. 1991, p. 2081–2095
0270-7306/91/042081-15$02.00/0

Vol. 11, No. 4

Human DNA Polymerase α Gene: Sequences Controlling Expression in Cycling and Serum-Stimulated Cells

BARBARA E. PEARSON, HEINZ-PETER NASHEUER, AND TERESA S.-F. WANG*

Laboratory of Experimental Oncology, Department of Pathology, Stanford University School of Medicine, Stanford, California 94305

Received 22 October 1990/Accepted 18 January 1991

We have investigated the DNA polymerase α promoter sequence requirements for the expression of a heterologous gene in actively cycling cells and following serum addition to serum-deprived cells. An 11.4-kb genomic clone that spans the 5′ end of this gene and includes 1.62 kb of sequence upstream from the translation start site was isolated. The transcription start site was mapped at 46 ± 1 nucleotides upstream from the translation start site. The upstream sequence is GC rich and lacks a TATA sequence but has a CCAAT sequence on the opposite strand. Analysis of a set of deletion constructs in transient transfection assays demonstrated that efficient expression of the reporter in cycling cells requires 248 bp of sequence upstream from the cap site. Clustered within these 248 nucleotides are sequences similar to consensus sequences for Sp1-, Ap1-, Ap2-, and E2F-binding sites. The CCAAT sequence and the potential E2F- and Ap1-binding sites are shown to be protected from DNase I digestion by partially purified nuclear proteins. The DNA polymerase α promoter can confer upon the reporter an appropriate, late response to serum addition. No single sequence element could be shown to confer serum inducibility. Rather, multiple sequence elements appear to mediate the full serum response.

The transition of quiescent cells to active proliferation is characterized by a temporal program of gene expression (1, 15, 49). Genes that display induced expression during this program can be classified into two groups, early- and late-response genes. The expression of early-response genes does not require protein synthesis and is frequently super-induced by cycloheximide. Late-response genes generally do require prior protein synthesis, as cycloheximide can prevent their expression.

The best-characterized early-response gene is c-*fos*. Within 15 min of serum addition, the upstream serum response element (SRE) mediates transcriptional activation and deactivation that result in a transient increase in the level of c-*fos* mRNA (24, 33, 36, 53, 57, 61). Sequences in the 3′ end of the gene contribute to the short half-life of the mRNA which enables the steady-state mRNA level to reflect the transcriptional state of the gene (61). The induction of c-*fos* is closely followed by a transient increase in the expression of another early-response gene, c-*myc*. Both transcriptional and posttranscriptional mechanisms contribute to the induction of c-*myc* mRNA (6, 14, 24). There is no SRE in the c-*myc* promoter. However, the c-*myc* promoter can confer an appropriate serum response upon a heterologous gene, and this capacity may involve the E2F-binding sites in this promoter (48).

Many genes encoding enzymes and proteins involved in either nucleotide metabolism or DNA replication are late-serum-response genes. The enzyme activities or protein levels of histones, dihydrofolate reductase (dhfr), thymidine kinase (tk), thymidylate synthase, proliferating cell nuclear antigen, DNA polymerase α, primase, ligase, topoisomerase, and others increase shortly before or simultaneous with the onset of DNA synthesis. In a few cases, the isolation of cDNA and genomic clones has enabled in-depth studies of the mechanism of the late serum induction. Like expression of the early response genes, expression of histones, tk, and dhfr is induced by a combination of transcriptional and posttranscriptional mechanisms. The promoters of these three genes can also impose serum inducibility with the appropriate late kinetics upon heterologous genes (2, 22, 31, 51a, 60). However, the precise mechanism of induction is unique to each gene.

Several *cis* elements have been implicated in mediating S-phase-specific expression of histone genes (52). A 32-bp upstream sequence of the hamster H3.2 gene was suggested to confer serum-induced expression of this gene (3), and the DNA-binding activity of the H2b octamer-binding factor increases in response to serum stimulation (28). tk protein and enzyme activity levels are apparently independent of the level of tk mRNA (27). However, the level of tk mRNA does increase during the course of serum stimulation as a consequence of a transient increase in transcription rate (56), a change in nuclear processing of heterogeneous nuclear RNA, and an increase in mRNA stability (12, 25, 38). While sequences that mediate serum induction of tk gene expression have yet to be identified, the ability of dhfr sequences to direct serum-induced expression has been attributed to the upstream sequence between −322 and −113 of the human gene (22) and to the leader sequence of the murine gene (51a). The leader may impose the appropriate regulation by mediating a change in the transcription rate or the rate at which reporter mRNA is degraded or translated.

In contrast to the regulation of serum induction of histones, tk, and dhfr, serum induction of proliferating cell nuclear antigen is suggested to use an entirely posttranscriptional mechanism (10). The increase in the level of mRNA is achieved by increasing the stability of both heterogeneous nuclear RNA and mRNA.

The expression of most of the serum-inducible genes does not vary through the cell cycle of actively cycling cells. Only the early-response gene, c-*myb* (59), and the late-response histone genes are known to display cell cycle regulation. We

* Corresponding author.

have previously reported that in actively cycling cells, DNA polymerase α mRNA, protein, and in vitro activity levels are constitutively expressed through the cell cycle. However, in serum-deprived cells, these parameters are low, and the addition of serum results in their coordinate increase in parallel with the onset of DNA synthesis (62). These findings indicate that during the cellular transition from quiescence to proliferation, gene expression is regulated independently of the mechanisms in force during continuous proliferation. Furthermore, the up-regulation of the expression of late-response genes, such as DNA polymerase α, is an important function of the signal transduction cascades initiated by growth factors.

In this report, we describe the isolation of a genomic clone that includes the 5′ end of the human DNA polymerase α gene. This clone contains the promoter, and we have investigated the sequences that are required for expression in both actively cycling cells and serum-starved cells that have been induced to proliferate by the addition of serum.

MATERIALS AND METHODS

Isolation of a human DNA polymerase α genomic clone. DNA from the 4X-chromosome cell line GM1202B was digested with *Eco*RI. The sizes of the *Eco*RI fragments that hybridized to the 5′ cDNA probe *Eco*RI-*Pst*I (which contains sequences between the *Eco*RI site of the vector and the *Pst*I site of the cDNA) (63) were determined by Southern hybridization. This size range was used to construct a genomic library in Charon 35 (39) by published methods (40). This library was packaged in Promega Biotec Packagene extracts, infected into *Escherichia coli* LE392, and screened with the *Eco*RI-*Pst*I cDNA probe. Two positive clones were plaque purified and determined to contain the same genomic *Eco*RI fragment by restriction mapping. Southern hybridization of the cDNA *Eco*RI-*Pst*I fragment to restriction enzyme digests of this genomic clone delineated regions that contain exons. The 5′ *Eco*RI-*Bam*HI genomic fragment (see Fig. 1A) was subcloned into M13 for dideoxy sequencing (see Fig. 1B) (51) and into pGEM-3 for the generation of fragments for sequencing by chemical cleavage (41).

Construction of plasmids. A genomic DNA-cDNA chimeric clone was constructed to generate the fragments used to map the transcription start site. The 5′ *Eco*RI-*Pst*I cDNA fragment was inserted into the *Eco*RI and *Pst*I cloning sites of vector pT7-7 (58), and the *Eco*RI-*Nco*I fragment was replaced with the genomic *Eco*RI-*Nco*I (−1570 to +46) fragment. The resultant clone contained the upstream genomic sequence linked to the 5′-most portion of the coding sequence of DNA polymerase α (see Fig. 2A).

The following constructs were used to evaluate the upstream sequence of the DNA polymerase α gene. A negative-control plasmid, pSV0ALΔ5′, was constructed by inserting the luciferase cDNA *Hin*dIII-*Sma*I fragment from pSV2ALΔ5′ (16) between the *Hin*dIII and *Bgl*II sites of pSV0A (23, 30) with the addition of a *Bgl*II linker (5′-GAAGATCTTC-3′; New England BioLabs) at the *Sma*I site. A set of test plasmids was constructed such that the luciferase cDNA was under the control of different amounts of DNA polymerase α upstream and leader sequences. *Xho*I (5′-CCCTCGAGGG-3′; New England BioLabs) and BglII linkers were added simultaneously to the 1,560-bp *Ava*II fragment (−1515 to +45) of the DNA polymerase α genomic clone. An *Xho*I linker was added at the *Hin*dIII site of pSV0A, and the modified *Ava*II polymerase fragment was inserted in the *Xho*I and *Bgl*II sites of this vector. The plasmid in which the 3′ end of the polymerase fragment was at the *Bgl*II site was designated pDPA. 5′ deletions of pDPA were created by digestion with *Hin*dIII (−1158), *Bal*I (−525), *Sph*I (−248), *Sac*II (−116), *Eag*I (−65), and *Nar*I (−17), addition and digestion of *Xho*I linkers, and recircularization of the large vector-containing fragments. A *Bgl*II fragment of the luciferase cDNA (LΔ5′) was generated from pSV0ALΔ5′ by the addition of *Bgl*II linkers at the *Hin*dIII site and inserted into the *Bgl*II site of pDPA and its 5′-deletion constructs. For each construct, the orientation of the luciferase cDNA was selected such that its 5′ end was juxtaposed to the 3′ end of the DNA polymerase α genomic sequence. The designation pDPALΔ5′ refers to the plasmid that contains the full-length *Ava*II fragment of the DNA polymerase α upstream sequence. Deletion constructs are denoted by pDPA- followed by the abbreviation for the restriction site at which they are deleted. The deletion construct pDPA-BsLΔ5′ was derived from pDPALΔ5′, rather than pDPA, by the same strategy used for the other deletion constructs. The construct pDPA-AvLΔ5′ was created by partial digestion of pDPA-HLΔ5′ with *Ava*I, gel isolation of the full-length linear molecules, addition and digestion of *Xho*I linkers, and recircularization. The construct pSc-AvLΔ5′ was created by partial digestion of pDPA-Sc (the same as pDPA-ScLΔ5′ but without the luciferase cDNA) with *Ava*I, addition and digestion of *Bgl*II linkers, recircularization, and insertion of the luciferase cDNA at the *Bgl*II site. To construct pΔattgLΔ5′, complete *Nar*I digestion of pDPA-Sc was followed by partial *Ava*I digestion. Appropriately sized molecules were recircularized. The recombinant junctions of pDPALΔ5′ and pSV0ALΔ5′ were verified by sequencing. All derivations were verified by restriction mapping.

Three plasmids were used to generate probes for S1 nuclease analysis of RNA from transfected NIH 3T3 cells. Plasmid pGEM3/β-gal contains the 1,160-bp *Eco*RV-*Bgl*I LacZ fragment inserted into the *Sma*I site of pGEM3. The 1,100-bp *Sph*I-*Eco*RI fragment of the luciferase cDNA was inserted into the same sites of pGEM3 and designated pGEM3/luc. Plasmid pGEM3/Mpo1α contains the 550-bp *Pst*I-*Bam*HI fragment of the murine DNA polymerase α cDNA. This cDNA clone was originally isolated from a mouse cDNA library by hybridization to the conserved region (770-bp *Bam*HI-*Hin*dIII) of the human DNA polymerase α cDNA (47a).

All procedures were essentially as described previously (5, 40). All plasmids were propagated in HB101, except pGEM3/β-gal, which was propagated in a *dam* mutant strain. For transfection into mammalian cells, plasmid DNA was CsCl banded twice, phenol-chloroform extracted twice, and chloroform extracted twice.

Primer extension and S1 nuclease analysis. Total RNA was isolated from KB or NIH 3T3 cells (9) and poly(A)$^+$ selected as described elsewhere (40). Probes used to map the transcription start site were derived from the chimeric genomic DNA-cDNA clone plasmid. The *Xho*I-digested DNA was 5′ end labeled and then digested with *Nco*I or *Eag*I. The labeled strands were isolated from a strand-separating gel. For primer extension, 5×10^4 cpm of 5′-end-labeled single-stranded antisense *Nco*I-*Xho*I primer was hybridized to 20 μg of poly(A)$^+$ RNA or 20 μg of *E. coli* tRNA and extended with avian myeloblastosis virus reverse transcriptase-XL (Life Sciences, St. Petersburg, Fla.) as described previously (8) except for the following specifications. RNA and primer were denatured at 90°C for 3 min and hybridized at 55°C for 3 h. The reaction contained 10 U of reverse transcriptase

(the enzyme stock was diluted 10-fold prior to use) (34), 1 mM rather than 0.5 mM deoxynucleoside triphosphates, 10 μg of bovine serum albumin (BSA) per ml, 0.5 mM spermidine HCl, and 4 mM $NaPP_i$, but dactinomycin was omitted.

S1 nuclease analysis was performed as described elsewhere (40) with the following modifications. To map the transcription start site, 40 μg of total RNA or 40 μg of *E. coli* tRNA was coprecipitated with 5×10^5 cpm of 5′-end-labeled single-stranded antisense *Eag*I-*Xho*I DNA from the genomic DNA-cDNA chimeric construct, denatured as for primer extension, and hybridized at 48°C. Digestion with 200 U of S1 nuclease was carried out at 37°C in 30 mM sodium acetate (pH 4.5)–250 mM sodium chloride–1 mM zinc chloride–10 μg of denatured salmon sperm DNA per ml. Products were separated on an 8% acrylamide–7 M urea gel with Maxam and Gilbert sequence ladders of the *Eag*I-*Xho*I chimeric fragment as markers.

S1 nuclease analysis of RNA isolated from transfected, serum-induced NIH 3T3 cells was as for transcription start site mapping except for the following. Three probes were prepared by digestion of pGEM3/β-gal, pGEM3/Mpo1α, and PGEM3/luc, with *Bcl*I, *Sty*I, and *Ava*II, respectively. These DNAs were dephosphorylated and digested within the vector at either *Nhe*I (pGEM3/luc and pGEM3/Mpo1) or *Hin*dIII (pGEM3/β-gal). The resultant 270-bp *Hin*dIII-*Bcl*I β-galactosidase, 643-bp *Nhe*I-*Sty*I mouse polymerase, and 761-bp *Nhe*I-*Ava*II luciferase fragments were gel isolated and 5′ end labeled. Each of the three probes (2×10^4 cpm each) was simultaneously hybridized to total RNA at 46°C and digested with 50 U of S1 nuclease, and the products were separated on a 4% polyacrylamide–7 M urea sequencing gel.

Cell culture. GM1202B cells were cultured in RPMI 1640 supplemented with 15% heat-inactivated fetal bovine serum. KB and HeLa suspension cells were cultured in suspension minimum essential medium with the addition of 10% calf serum, 20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES), and 1.6 mM sodium bicarbonate. Attached HeLa cells were cultured in Dulbecco modified Eagle medium supplemented with 10% fetal bovine serum. NIH 3T3 cells were cultured in Dulbecco modified Eagle medium supplemented with 10% calf serum. All media contained 100 U of penicillin G per ml and 100 μg of streptomycin sulfate per ml.

NIH 3T3 cells were serum stimulated as follows. At 12 to 16 h after the addition of $CaPO_4$-DNA coprecipitate, the cells were washed twice with phosphate-buffered saline (PBS) and starved in medium supplemented with 0.5% calf serum for 30 to 45 h. Electroporated cells were immediately resuspended in medium supplemented with 0.5% calf serum for starvation. Starved cells were stimulated with medium supplemented with 15% calf serum. To monitor DNA synthesis, cells were pulse-labeled with [^{3}H]thymidine for 30 min, washed twice with PBS containing 0.4 mg of thymidine per ml, lysed with 0.5 M NaOH containing 0.1 mg of carrier DNA per ml, and precipitated with 20% trichloroacetic acid. The precipitate was filtered on GF-C paper (Whatman) and counted with scintillation cocktail.

DNA transfection. For $CaPO_4$ transfection, 10-cm-diameter dishes were seeded with approximately 1×10^6 HeLa or 3×10^5 NIH 3T3 cells 1 day before transfection. $CaPO_4$ coprecipitation was performed as described previously (4). Each plate was transfected with 10 μg of test plasmid DNA, 5 μg of pRSVβ-gal (42), and 5 μg of salmon sperm DNA. Exponentially growing NIH 3T3 cells were electroporated in 0.5 ml of 1× HeBS (11) at 960 μF, 230 or 280 V, and 25°C in a Bio-Rad Gene Pulser with 30 μg of test plasmid and 10 or 30 μg of pRSVβ-gal.

Enzyme and protein assays. Cells were harvested and lysed as described elsewhere (18). Extracts were assayed for luciferase (16) and β-galactosidase (50) as described previously. Protein was quantified by the method of Bradford (7).

Partial purification of nuclear proteins. Nuclear extracts from HeLa cells were prepared (26) and nuclear protein factors were partially purified (35) as described elsewhere, with the exception that 10 mM sodium bisulfite was added to all buffers. Nuclear extract was first dialyzed to 70 mM NaCl and applied to phosphocellulose. The column was eluted in steps with 200 and 600 mM KCl. The flowthrough and the eluates were further fractionated on heparin-Sepharose. Proteins were eluted by a linear gradient from 150 to 600 mM KCl. The specific DNA-binding activity of each fraction was measured by a gel mobility shift assay (20, 21) using restriction fragments containing the element of interest.

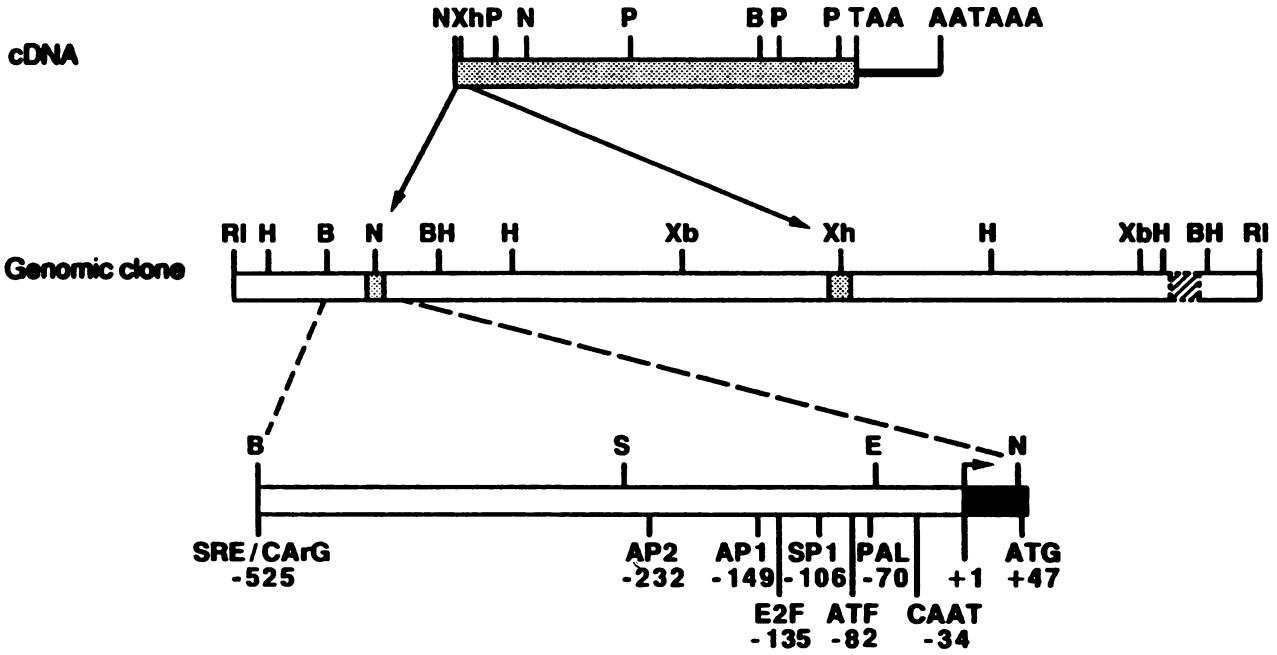
DNase I protection. The DNase I protection assays were performed as described previously (20, 21, 26) with 0.2 to 1 ng (or 5,000 to 10,000 cpm) of labeled restriction fragment, 0.5 to 2 μg of partially purified nuclear proteins, and 1 or 2 μg of poly(dI · dC) (double-stranded deoxypolymer in 10 mM Tris-HCl [pH 7.9]–1 mM EDTA–100 mM NaCl) in a 20-μl total volume. After incubation at room temperature for 20 to 30 min, the DNA-protein mixture was adjusted to 5 mM $MgCl_2$ and 1 mM $CaCl_2$, and 0.01 to 0.04 U of DNase I was added for a further 25- to 90-s incubation at room temperature. The reaction was quenched by the addition of 100 μl of stop solution (0.375% sodium dodecyl sulfate, 15 mM EDTA, 100 mM NaCl, 100 mM Tris-HCl, pH 7.6), 15 μg of tRNA, and 15 μg of proteinase K and incubated at 37°C for 15 min. Samples were extracted once with phenol-chloroform and once with chloroform, ethanol precipitated, resolubilized in sequencing gel loading buffer, and analyzed on a 6 or 8% polyacrylamide–7 M urea sequencing gel. Control protections were performed with 20 μg of BSA instead of nuclear proteins.

RESULTS

Human DNA polymerase α genomic clone. A genomic clone spanning the 5′ end of the DNA polymerase α gene was isolated from a genomic library of a human 4X-chromosome cell line (GM1202B) constructed in bacteriophage λ vector Charon 35. This clone hybridizes to a 455-bp restriction fragment, *Eco*RI-*Pst*I, of the very 5′ end of the human polymerase α cDNA (see Materials and Methods and reference 63) and is 11.4 kb. By Southern hybridization and restriction mapping, this 11.4-kb genomic sequence contains 1,571-bp of 5′-flanking sequence, the first three exons, the first two introns, and a portion of the third intron (Fig. 1A). Sequence analysis of the genomic DNA fragments that hybridized to the cDNA fragments revealed the exon-intron boundaries of the first two exons. The presence of the third exon was defined only by hybridization with cDNA fragments, and its precise boundaries were not established by sequence analysis. The first exon is 70 bp, and the second exon is 125 bp. These two exons are separated by an approximately 5,120-bp intron. The second intron is at least 3,400 bp. The entire upstream region and the first exon were sequenced (Fig. 1B).

Transcription start site. The previously isolated cDNA clone of human DNA polymerase α contains only 15 bp upstream of the designated translation start site (63). To locate the transcription start site and confirm the translation

A



B

```
-1571  GAATTCCAAG ATCAAGGTGC CGGCAGATTC AGTGTTTGGG GTAGGCCCAC TTTTTGGTCC ATAGATGGTG CCTTATTGCT CTGTTCTCAC ATGGAAGAAG
-1471  GGGCAAGGCA GCTCTCTGGA GCCTCTTTTC CGAGGGCACT AATCCTATTC ATGAGGGAGG GTTCCTGCTG ATGACCTAAT CACCTCCCAG AGGCCCTACC
-1371  TCCTAACACC ACCACACTGG TGATTAGGTT TCAACGCATA CATTTTATGT GGAGAGAAAC ATTCAGACCA CAGCAGTAGT GAAAGAGAAA AGAGGTTTGA
-1271  AATGCCTAGA GTTATACCCG GGCATGTGGG AAGTAATCAG CTGTGAAATT GTTAAGGTAA GGATTTGGTT CTTATTAATG ACTAGCCAAC ATGGTTCTCT
-1171  CCTGGTTGGA AAAAGCTTGA TAATAAAACT ATAAAACTTC CATCAGTTTT TAAAAGGGAA GTGTAGGAAT AAATCTGTCG TAGTACTTTA AATAGGAAGT
-1071  TATGTCTAAG TGTAAATTAC TACATGTTAT GTATCACAAC ATATGGATCA CATAACATGA TATTAAATGA TTATGAAGCT GAAAGACATT TTCTAAACAA
 -971  TCAGAAATTA AGAAACAAGT ATTGATCAAC CATGCTAGAA GACAGATTGA ATTATCTTTA TATCCTCTTT GGAGAAAATA TTACAAAATT GTTGTCACAT
 -871  GAAGAAGGGA TCAATAAGCA TGTGGGCAAT AATGTAGGAA AAAATACAAA TGTGTGGCAG GCATTATTTA ATAAAAATGT TATACCTTTT TGGAAAGAGA
 -771  TAACATGAGA TGTACACATT ACAGCCATTG TATCATACAC AAAAATTGAA TACCCTTTTT GTTTTTCATA ACATTGAAAT TTTGGAGTTC ACGTCAGGTG
 -671  TTTTGTAGAT TGTCCTACAT TTTGAATACA TCTGATTGTT TCTTCATGGT TAGATTCAGG TTAGACATTT TTTGCAAGAA GACCATATAG GCGAAGTTTT
 -571  TCCTTTTTCA TACGCATTAC ATCAGGAGGC ATATGAGATC AGATGGCCAT TTATTGGTAA TGCTACTTCT GGTTATTTGG TTAAGACGGA GCATATATCT
 -471  TTAAGTTGTT AATCACACAG CCACTGTATC CAGACTGCTG AGAAGCAAAC TATCAGGCAA AATTCAGGGC GCGCAGCAGG AGGCAGTCGC TTTCCCGGCT
 -371  CTGGGGAAAA CGATCCAACC ACGCCCACTC TCCTGCCCGC CCGCGCCTCC TTTCCGGGAA AATGGGCTCT GAGAACCAGA GCGCCCTTTC TCCCACGCGC
 -271  CCAAATCTTT TCCCATCAGC ATGCGTGCGC TTGCGCAGTC TCCAGGCGGC TCCTTTCCAG CGCCCGACCG CTTCTTTCTC CTGCTCTCCT CCGGCTTCCC
 -171  CTCAGCTCTA GCTTTTCCCT AAGGGGTCAT CACAGGGCGC CAAACGCGCG TCCGCGGGAA GGCCGGGGGC GGAGCGCGGC GCGGCGCGGC ACGTCAGTGG
  -71  CCTTCCGGCC GGAAGTCCGC AGCCTCCCGA GCCGCTGATT GGCTTTCAGG CTGGCGCCTG TCTCGGCCCC CGCGCCAGTT TTGGGCTGGT TGGCGCGGAA
  +30 TCGGGAGATT CGGGACCATG GCACCTGTGC ACGGCGACGA CTGTGAGATA GGGGCGAGTG GTGAGGGACA ATTCGCGCGC GGGGCTCCGG GTTGGGACAG
```

FIG. 1. Analysis of the DNA polymerase α genomic clone. (A) Restriction maps of the human DNA polymerase α cDNA and genomic clones. The cDNA clone is 5.4 kb. The genomic clone is 11.4 kb, and the 572-bp *Bal*I-*Nco*I fragment is expanded. Stippled boxes represent coding sequences of the cDNA and the first two exons of the genomic clone. The third exon, the precise boundaries of which are not yet defined, is represented by the striped box. Open boxes represent introns and the sequence upstream from the transcription start site. The 5′ leader sequence of the cDNA is represented by the solid box, and the 3′ noncoding region is shown by the solid line. Abbreviations: N, *Nco*I; Xh, *Xho*I; P, *Pst*I; B, *Bal*I; BH, *Bam*HI; RI, *Eco*RI; H, *Hin*dIII; Xb, *Xba*I; S, *Sph*I; E, *Eag*I; ATG, translation start codon; TAA, translation stop codon; AATAAA, polyadenylation signal. (B) DNA sequence of the 5′ portion of the DNA polymerase α clone. The first exon is underlined and in boldface. The coding portion of the first exon is in italics. Arrows indicate the location of the transcription (+1) and translation (+47) start sites.

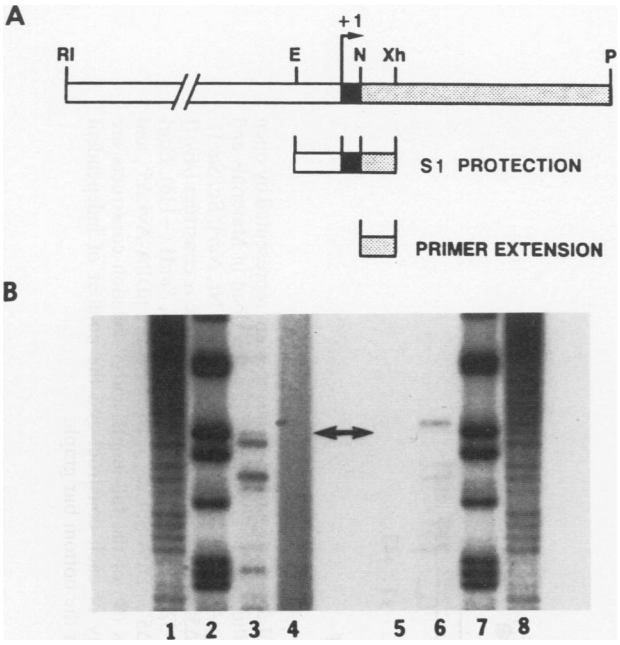


FIG. 2. Location of the transcription start site. (A) Restriction map of the genomic DNA-cDNA chimeric construct (see Materials and Methods). Abbreviations of restriction sites are as described in the legend to Fig. 1. Restriction fragment *Eag*I-*Xho*I from this chimeric clone was prepared as probe for S1 nuclease protection. Restriction fragment *Nco*I-*Xho*I was derived from the *Eag*I-*Xho*I fragment and prepared for use as a primer for primer extension. (B) Primer extension and S1 nuclease protection. Experiments were performed as described in Materials and Methods. Lanes contain the following: antisense Maxam and Gilbert sequence ladders of the *Eag*I-*Xho*I fragment, G+A (lanes 1 and 8) and C+T (lanes 2 and 7); primer extension of the *Nco*I-*Xho*I primer with poly(A)$^+$ RNA from KB cells (lane 3) or *E. coli* tRNA (lane 4); S1 nuclease protection of the *Eag*I-*Xho*I probe with total RNA from KB cells (lane 6) or *E. coli* tRNA (lane 5). The arrows point to the nucleotide designated +1, the transcription start site.

start site, S1 nuclease protection was performed. A chimeric clone was constructed to overcome the lack of appropriate restriction sites in the short first exon. This construct combines the 1,617 bp of genomic sequence upstream from the designated translation start site (*Eco*RI-*Nco*I fragment) with the first 438 bp of coding sequence from the cDNA (*Nco*I-*Pst*I fragment) (Fig. 2A). The 167-bp restriction fragment, *Eag*I to *Xho*I, of the chimeric construct was used for Maxam and Gilbert sequence ladders. S1 nuclease protection analysis showed a single major product that included 47 nucleotides upstream from the designated translation start site (Fig. 2B, lane 6).

To confirm the S1 protection result, primer extension was performed with a single-stranded 57-bp primer, isolated from the *Nco*I-*Xho*I fragment of the cDNA (Fig. 2A). Two doublets, each containing a major and a minor band, were observed (Fig. 2B, lane 3). The longest of these four extension products was a minor product and reached the sense strand nucleotide G (corresponding to a C of the antisense Maxam and Gilbert sequence ladder) at 46 nucleotides upstream from the ATG codon. The longest major product reached the sense strand nucleotide C at 45 nucleotides upstream of the ATG. For numbering, we designate this C as +2, the G as +1, and the S1 nuclease-protected C, 47 nucleotides upstream from the ATG, as −1 (Fig. 1B).

TABLE 1. Comparison of human DNA polymerase α upstream sequence and known transcriptional regulatory elements

Factor	Canonical element	Similar sequence of human DNA polymerase α[a]
CTF	-CCAAT-	-CCAAT-
ATF	-(T/G)(T/A)CGTCA-	-CACGTCA-
SP1	-GGGGCGG-	-GGGGCGG-
E2F	-TTTCGCGC-	-TTTGGCGC-
AP1	-T(T/G)AGTCA-	-GGGGTCA-
AP2	-CCCCAGGC-	-CTCCAGGC-
CArG/SRE	-CC(A/T)$_6$GG-	-CCATTTATTGG-

[a] Underlined nucleotides are those which are different from the canonical elements.

Together, S1 nuclease protection and primer extension located the preferred transcription start site of human DNA polymerase α gene between 45 and 47 nucleotides upstream from the ATG proposed to be the translation start site. No other start codons are contained in this region.

Potential transcription factor-binding sites. The 200-bp sequence immediately upstream of the preferred transcription start site has a GC content of 70%. There is no obvious TATA sequence, whereas a CCAAT sequence is located on the opposite strand (−30 to −34). Further upstream, from −70 to −57, there is a novel 14-bp palindrome (PAL), 5′-CTTCCGGCCGGAAG-3′. Several additional potential transcription factor-binding sites were identified by their similarity to consensus sequences for known transcription factor-binding sites (Table 1). From −82 to −76, the heptanucleotide -CACGTCA- is similar to the ATF-binding site found in viral and cellular transcription control regions (35). A canonical Sp1-binding site, -GGGGCGG- (29, 46), is located between −106 and −100. On the opposite strand, from −128 to −135, the sequence -TTTGGCGC- differs by only 1 base from the adenovirus E2F site (29, 54). Except for 2 mismatched nucleotides, the sequence from −149 to −143, -GGGGTCA-, is similar to the Ap1-binding site (29, 46). The sequence -CTCCAGGC- differs by 1 nucleotide from the consensus Ap2 element (29, 46) and is found between −232 and −225. Further upstream, from −525 to −515, there is the motif 5′-CCATTTATTGG-3′. Interestingly, this sequence is similar to the core sequence of the human c-*fos* SRE (61), and the CArG boxes of actin genes (19, 45). However, the polymerase α motif has 7, instead of the frequently observed 6, A/T residues between the -CC and GG-.

Promoter delineation. To determine the limits of the upstream sequence required for the expression of human DNA polymerase α gene in cycling cells, a set of deletions of the 1.62-kb upstream sequence was constructed and tested for the capacity to direct expression of the luciferase cDNA reporter. These constructs were transiently transfected into exponentially growing HeLa cells. Plasmid pRSVβ-gal (42) was cotransfected to monitor transfection efficiency. This plasmid carries the Rous sarcoma virus long terminal repeat which directs expression of *E. coli* LacZ. In every experiment, luciferase activity (in light units) was normalized to β-galactosidase activity (50). Results of these transfection experiments are presented in Fig. 3. The luciferase activity produced by the *Ava*II construct pDPALΔ5′ (−1515 to +45) was defined as 100% activity and was the standard with which the luciferase activity produced by other constructs was compared. On average, pDPALΔ5′ yielded 60% of the

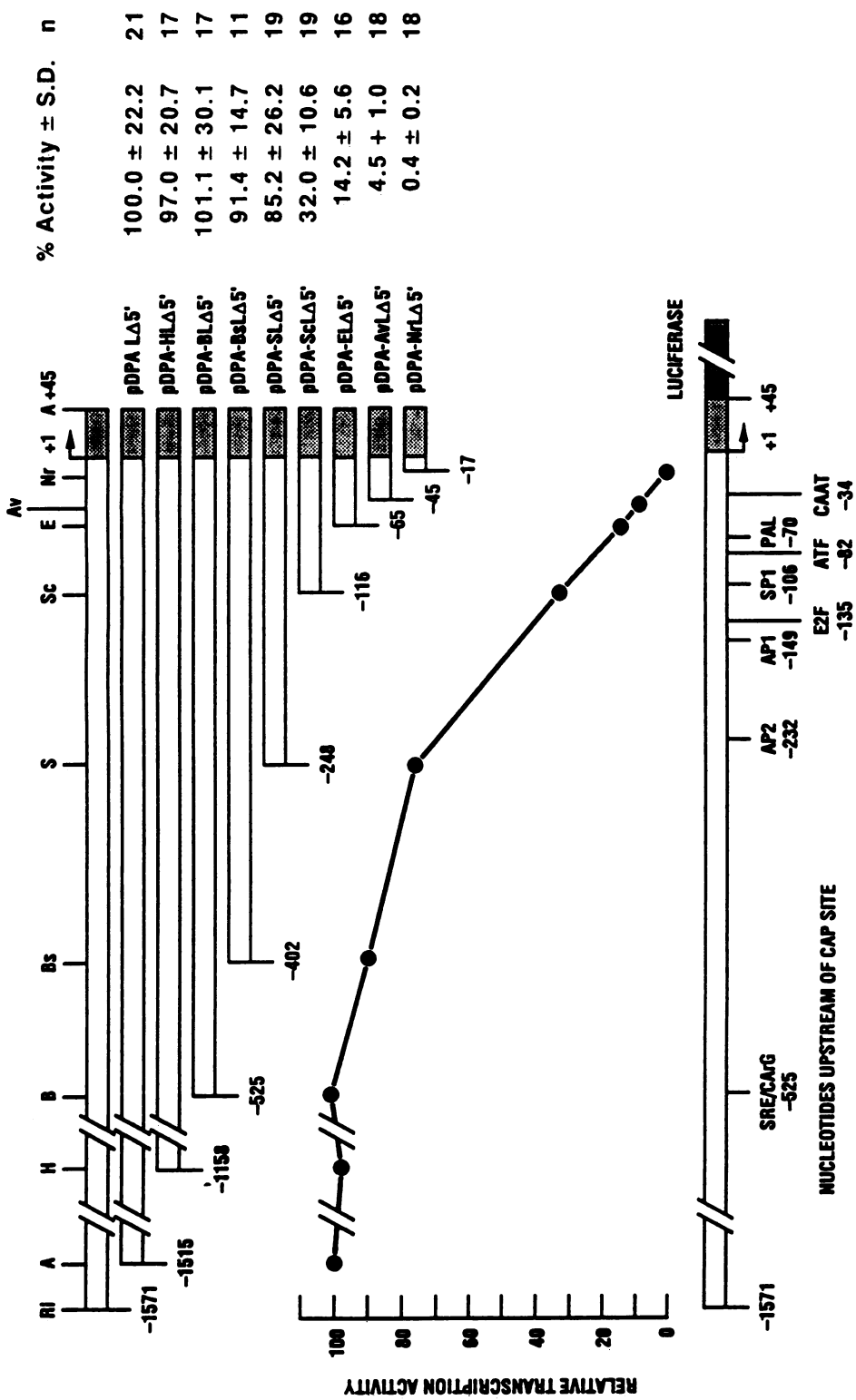


FIG. 3. Transcription activity of the upstream sequence of human DNA polymerase α gene in cycling HeLa cells. Upstream sequences are represented by open boxes, the 5′ leader sequence is shown by shaded boxes, and the luciferase reporter cDNA is shown by the solid box. Plasmids are described in Materials and Methods. Abbreviations of the restriction sites are as in the legend to Fig. 1 with the addition of the following: A, *Ava*II; Av, *Ava*I; Bs, *Bss*HII; Nr, *Nar*I; Sc, *Sac*II. The top bar graph is a schematic summary of restriction sites in the genomic upstream sequence. pDPALΔ5′ is the longest DNA polymerase α construct (*Ava*II fragment, −1515 to +45) with the luciferase reporter. Constructs with deletions at *Hin*dIII (−1158), *Bal*I (−525), *Bss*HII (−402), *Sph*I (−248), *Sac*II (−116), *Eag*I (−65), *Ava*I (−45), and *Nar*I (−17) are designated pDPA-HLΔ5′, pDPA-BLΔ5′, pDPA-BsLΔ5′, pDPA-SLΔ5′, pDPA-ScLΔ5′, pDPA-ELΔ5′, pDPA-AvLΔ5′, and pDPA-NrLΔ5′, respectively. All deletion constructs contain the luciferase reporter. Luciferase activities (●) of the full-length and 5′-deletion constructs are represented as the percentage of maximum activity. Standard deviation (S.D.) of the luciferase activity of each construct and the number of independent transfections (n) are listed. The locations of sequence elements in the upstream sequence are illustrated in the bottom bar graph.

luciferase activity produced by the simian virus 40 promoter-enhancer in pSV2ALΔ5′ (16) (data not shown). Neither the construct pSV0ALΔ5′ nor transfection with salmon sperm DNA produced any luciferase activity above the background produced by luciferase reaction buffer alone. The *Hin*dIII (−1158) and *Bal*I (−525) deletion constructs expressed luciferase at essentially the same level as did pDPALΔ5′. On average, the *Bss*HII (−402) deletion construct, which excluded the -CC(A/T)$_7$GG motif, and the *Sph*I (−248) deletion construct resulted in only a modest reduction of luciferase activity. One hundred percent activity was within the standard deviations of luciferase activity produced by these constructs. Therefore, sequences upstream of the *Sph*I site do not appear to contribute significantly to the activity of the DNA polymerase α promoter. However, an additional deletion to *Sac*II (−116) removed the sequence containing similarities to the Ap2-, Ap1-, and E2F-binding sites and reduced the luciferase activity to about 30% of that produced by pDPALΔ5′. A deletion to *Eag*I (−65) eliminated the potential Sp1- and ATF-binding sites and part of the PAL sequence. This deletion decreased the luciferase activity to 14%. The remainder of the palindrome was eliminated by an additional 20-bp deletion, to *Ava*I (−45), and decreased the luciferase activity to 4%. A deletion beyond the CCAAT sequence, to *Nar*I (−17), removed only 28 bp and reduced luciferase activity 10-fold, to 0.4%. This set of experiments identifies the sequence from −248 to +45 of the human DNA polymerase α gene as the sequence that controls the expression of this gene in cycling cells. Within this region, even small deletions result in significant reduction of reporter activity.

Nuclear proteins bind the upstream sequence *in vitro*. The binding of nuclear proteins to restriction fragments of the upstream sequence was tested by DNase I protection assays. A portion of the DNase I protection pattern of the *Sac*II-*Nco*I (−116 to +46) restriction fragment, which contains the Sp1, PAL, and CCAAT sequences, is shown in Fig. 4A. Each of these motifs was protected, although protection of the CCAAT sequence was most apparent. The protection spanning the CCAAT sequence, from −46 to −22, closely matches the deletion, from −45 to −17, that results in a 10-fold reduction of luciferase activity produced in the transient transfection assay described above. Protection of the *Bst*NI-*Eag*I (−228 to −65) fragment spans the sequences with similarity to the Ap1- and E2F-binding sites (Fig. 4B). When other restriction fragments were assayed for DNase I protection, the sequences with similarity to SRE and the Ap2-binding site were footprinted, but the ATF-like binding sequence was not (data not shown). Competition assays of gel mobility shift were performed with synthetic duplex oligonucleotides (30-mers) corresponding to the DNA polymerase α genomic sequences that have similarity to CCAAT and SRE and the binding sites of Sp1, Ap1, and Ap2. In each case, a 5- to 20-fold molar excess of the duplex oligonucleotide proportionally depleted the formation of unique nuclear protein-DNA complexes. These results indicate the specificity of these interactions (data not shown). The ability of PAL or the sequence similar to the E2F-binding site to compete for nuclear factors was not tested.

The DNA polymerase α upstream sequence confers a late serum response upon a heterologous cDNA. It is possible that the positive induction of DNA polymerase α as a response to the addition of serum to serum-starved cells (62) is mediated at least in part by an increase in the transcription rate of this gene. If so, the upstream sequence of the DNA polymerase α gene should impose this response upon a heterologous sequence. To test this hypothesis, NIH 3T3 cells were transiently transfected with the longest construct, pDPALΔ5′, serum deprived for 30 to 40 h, and then induced to proliferate by the addition of serum. NIH 3T3 cells were chosen for this experiment because they are readily brought to quiescence by serum starvation and are therefore suitable for transient transfections. At various times after serum addition, cells were harvested and extracts were prepared. Luciferase activity was normalized to β-galactosidase activity expressed from cotransfected pRSVβ-gal to correct for transfection efficiency. The results of a typical experiment are shown in Fig. 5A. DNA synthesis, as monitored by the incorporation of [^{3}H]thymidine, began 7 to 12 h after serum addition, peaked between 14 and 18 h, and began to decline by 20 h. Luciferase activity was usually higher in serum-starved cells than in cells that had been exposed to serum for 5 h. Beyond 5 h, luciferase activity, like DNA synthesis, followed a late course of induction that began between 7 and 12 h and peaked 16 to 20 h after serum addition. To test the possibility that the luciferase cDNA contains sequences that mediate the observed late serum response, cells were transfected with pRSVβ-gal and pRSVLΔ5′ (in both constructs, the Rous sarcoma virus long terminal repeat directs luciferase expression [16, 42]) and subjected to a course of serum deprivation and readdition. In this case, serum stimulation resulted in a reproducible decrease in normalized luciferase activity (Fig. 5A). Vector sequences of pRSVβ-gal, pRSVLΔ5′, and all of the polymerase constructs are the same except that polymerase constructs contain two head-to-tail simian virus 40 polyadenylation signal sequences (237 bp, *Mbo*I-*Mbo*I) between the pBR322 and polymerase sequences. Introduction of this double polyadenylation sequence into pRSVLΔ5′ accentuated the serum-induced decline of normalized luciferase activity (data not shown). These results eliminate the possibility that vector sequences impose the serum induction and suggest that the in vivo stability of luciferase RNA or protein is lower than that of β-galactosidase.

Because the efficiency of $CaPO_4$ transfection can vary from plate to plate, the response profile of the β-galactosidase activity could not be determined from these experiments. However, this information is critical for interpretation of the luciferase activity response profile. Therefore, the serum response experiments were repeated with NIH 3T3 cells that had been transfected with pDPALΔ5′ and pRSVβ-gal by electroporation. All cells were transfected simultaneously and then distributed in plates for harvest at various times. The total protein content was determined for each lysate. The specific activities of β-galactosidase and luciferase were calculated, and the specific activity of luciferase was normalized to the specific activity of β-galactosidase. The response profiles are shown in Fig. 5B. The decrease in luciferase activity between 0 and 5 h that is characteristic in serum-stimulated $CaPO_4$-transfected cells was not observed in electroporated cells. With respect to total protein, β-galactosidase activity slowly increased to a maximum of twofold by 20 h after serum addition. However, luciferase activity, when normalized to total protein, increased 1.5-fold by 7.5 h and then underwent a more rapid increase to 8-fold over the minimum by 20 h after serum addition. However, after normalization to β-galactosidase activity, the increase in luciferase activity was only fourfold. The consequence of normalizing luciferase activity to β-galactosidase activity was that the peak luciferase response was underrepresented by, at most, a factor of 2.

Deletions of the DNA polymerase α promoter reduce the

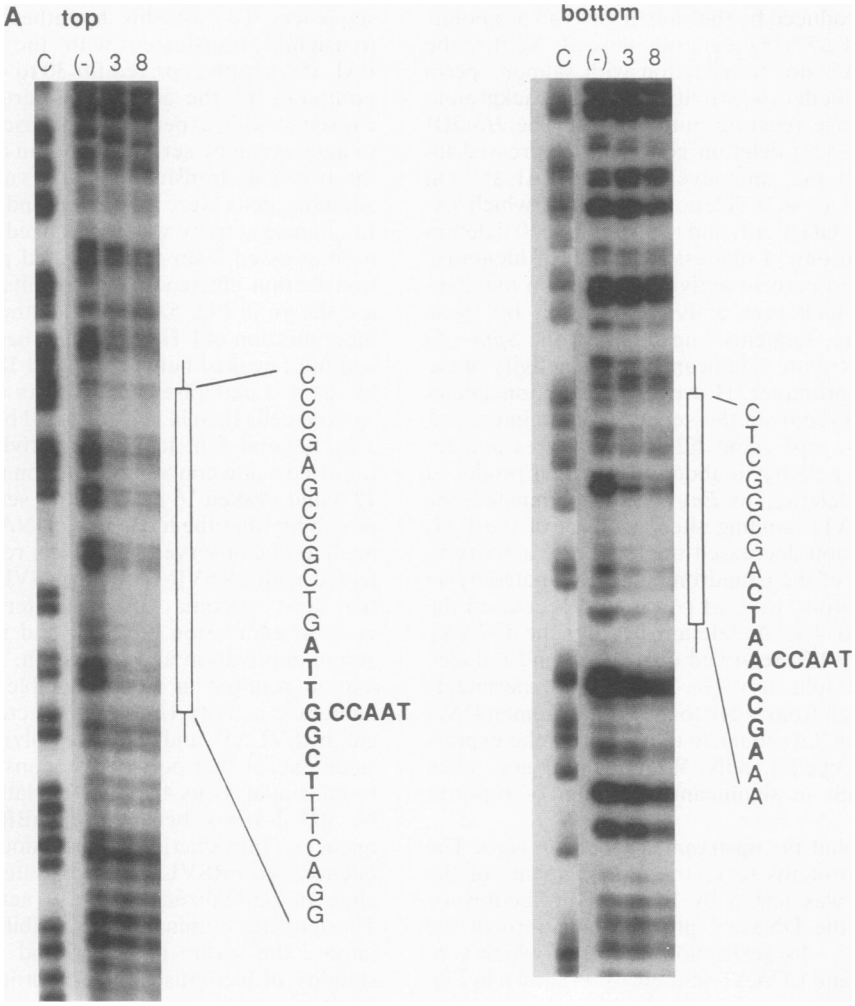


FIG. 4. DNase I protection analysis of the DNA polymerase α upstream sequence. Protections were performed as described in Materials and Methods with either the coding (top) strand or the noncoding (bottom) strand labeled. For all panels, Maxam and Gilbert sequence ladders are indicated by C and lane (−) is the control protection with BSA. (A) Footprinting of the inverted CCAAT sequence. The 161-bp *Sac*II-*Nco*I (−116 to +46) fragment was incubated with 3 μl (lanes 3) or 8 μl (lanes 8) of 6.1-μg/μl HeLa cell nuclear protein, heparin-Sepharose-purified phosphocellulose flowthrough. (B) Footprinting of the sequences with similarity to the Ap1- and E2F-binding sites. The 163-bp *Bst*NI-*Eag*I (−228 to −65) fragment was incubated with 2 μl (lanes 2) or 4 μl (lane 4) of 17.4-μg/μl HeLa cell nuclear protein, heparin-Sepharose-purified phosphocellulose 600 mM eluate.

magnitude of the serum response. If a single, short region of the promoter is necessary to direct the late response to serum, deletion of that region from the promoter would be expected to abolish the serum induction. To test for the existence of such a sequence, several of the 5′-deletion constructs, one 3′-deletion construct, and one internal-deletion construct were $CaPO_4$ transfected into NIH 3T3 cells. Their ability to confer a serum response upon the reporter was analyzed. The serum response of the longest construct, pDPALΔ5′, was always assayed in parallel for comparison, and [^{3}H]thymidine incorporation was monitored to assess the competence of the cells. Initial experiments included a number of time points, as in Fig. 5A, to follow the time course of induction of the deletion construct expression. Both of the deletion constructs pDPA-BsLΔ5′ (−402) and pDPA-ELΔ5′ (−65) responded to serum with the same kinetics as pDPALΔ5′ (data not shown). Three significant time points, 0, 5, and 18 h after serum addition, were chosen for the remainder of the experiments. As before, the minimum level of luciferase activity occurred at 5 h. The level of luciferase activity at 0 h was variable and could not be correlated with either the length of time that the $CaPO_4$-DNA coprecipitate was left on the cells or the number of hours that the cells were starved of serum. Therefore, the quiescent level of luciferase activity at 0 h and the level of induction at 18 h were normalized to the activity at 5 h. The relative luciferase activities produced by these constructs at 18 h were similar to those expressed in cycling NIH 3T3 or HeLa cells.

The serum response profiles of the various constructs are shown in Fig. 6. No correlation could be made between the relative level of expression in serum-starved cells and any specific portion of the polymerase upstream sequence. However, with one exception, successive 5′ deletions of the upstream sequence resulted in a decreasing magnitude of the 18-h response. The relative magnitude of the serum response produced by the *Bss*HII (−402) deletion construct was nearly the same as that of pDPALΔ5′. The relative magni-

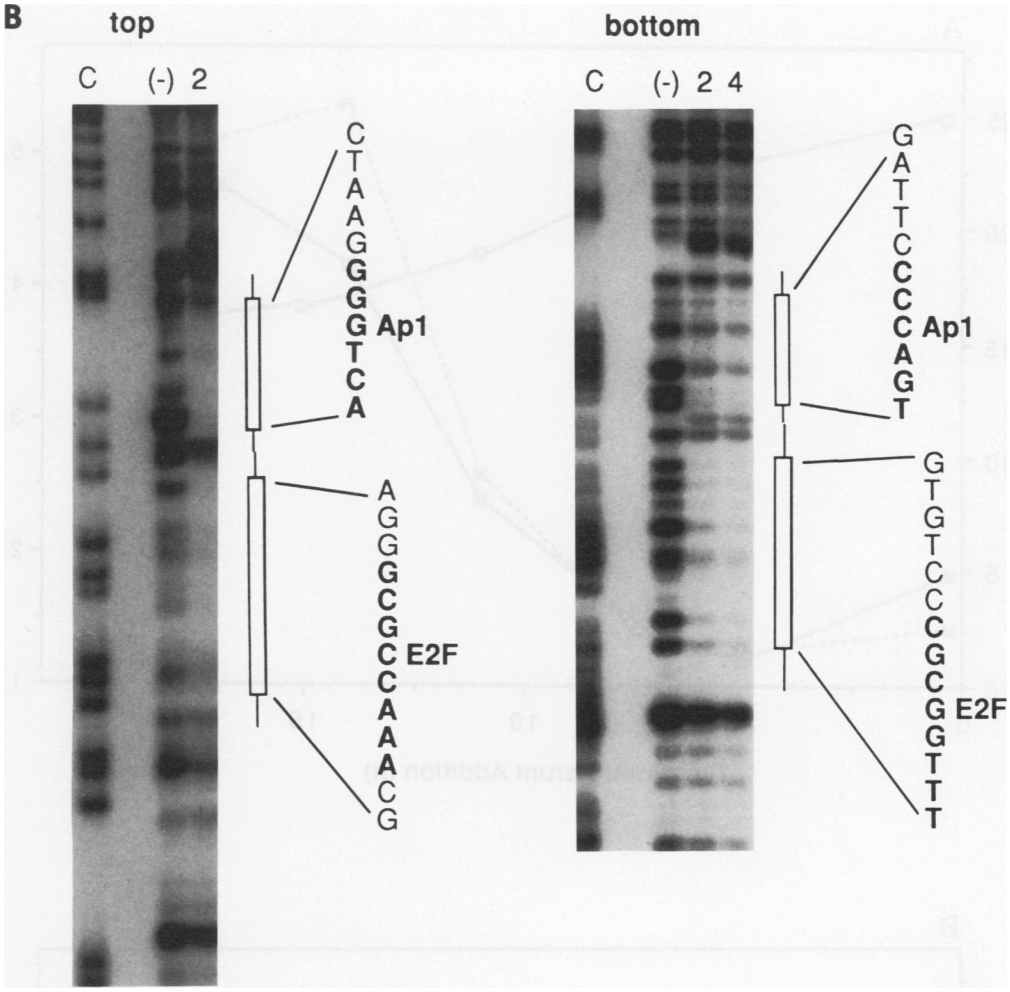


FIG. 4—*Continued.*

tude of the serum response produced by the *Sac*II (−116) deletion construct was lower. The exception mentioned above is that the response of pDPA-ELΔ5′ (−65) was greater than that of pDPA-ScLΔ5′ (−116) (5.07-fold and 3.93-fold, respectively). The *Ava*I (−45) construct produced a response that was lower than that of the pDPA-ScLΔ5′ construct. The *Nar*I (−17) construct did not appear to respond to serum.

The effect of the internal deletion that eliminated the CCAAT sequence (−45 to −17, pΔattgLΔ5′) or the 3′ deletion of the CCAAT, initiation, and leader sequences (−45 to +45, pSc-AvLΔ5′) from pDPA-ScLΔ5′ was also tested. As compared with the response of pDPA-ScLΔ5′, the deletion of only the CCAAT sequence resulted in a twofold reduction in the 18-h response and deletion of the CCAAT, initiation, and leader sequences reduced the response threefold (to 2.12 and 1.24, respectively). The 18-h luciferase activity produced by pΔattgLΔ5′ was lower than that of pDPA-ScLΔ5′, as expected. The 18-h luciferase activity produced by the 3′-deletion construct pSc-AvLΔ5′ was 24% of that of pDPALΔ5′. This activity was unexpectedly high, considering that the natural transcription start site had been removed.

The relative serum responses of the various polymerase promoter constructs suggest that the dominant late serum response control lies within the sequence from −65 to −17. That more than one late serum response element is contained within this region is suggested by the intermediate magnitude of the serum response produced by the −45 construct. Sequences outside of this region also appear to influence the magnitude of the late serum induction of DNA polymerase α

Induction of luciferase transcript in response to serum. In our previous report (62), serum addition to serum-starved cells resulted in an increase in DNA polymerase α enzyme activity that paralleled an increase in the level of polymerase mRNA. It was therefore of interest to determine whether the serum induction of luciferase activity is paralleled by an increase in luciferase RNA. This question was addressed by electroporation of NIH 3T3 cells with polymerase promoter deletion constructs and pRSVβ-gal followed by serum starvation and readdition. Transfection by electroporation allowed direct comparison of luciferase and DNA polymerase α RNAs because the difference in transfection efficiencies between plates is minimal. The quantities of luciferase, β-galactosidase, and endogenous polymerase α RNAs at 0, 5, and 20 h after the addition of serum were determined by S1 nuclease analysis (Fig. 7). The cDNA probe for the endogenous mouse DNA polymerase α specifically detected processed transcripts. However, the probes for luciferase

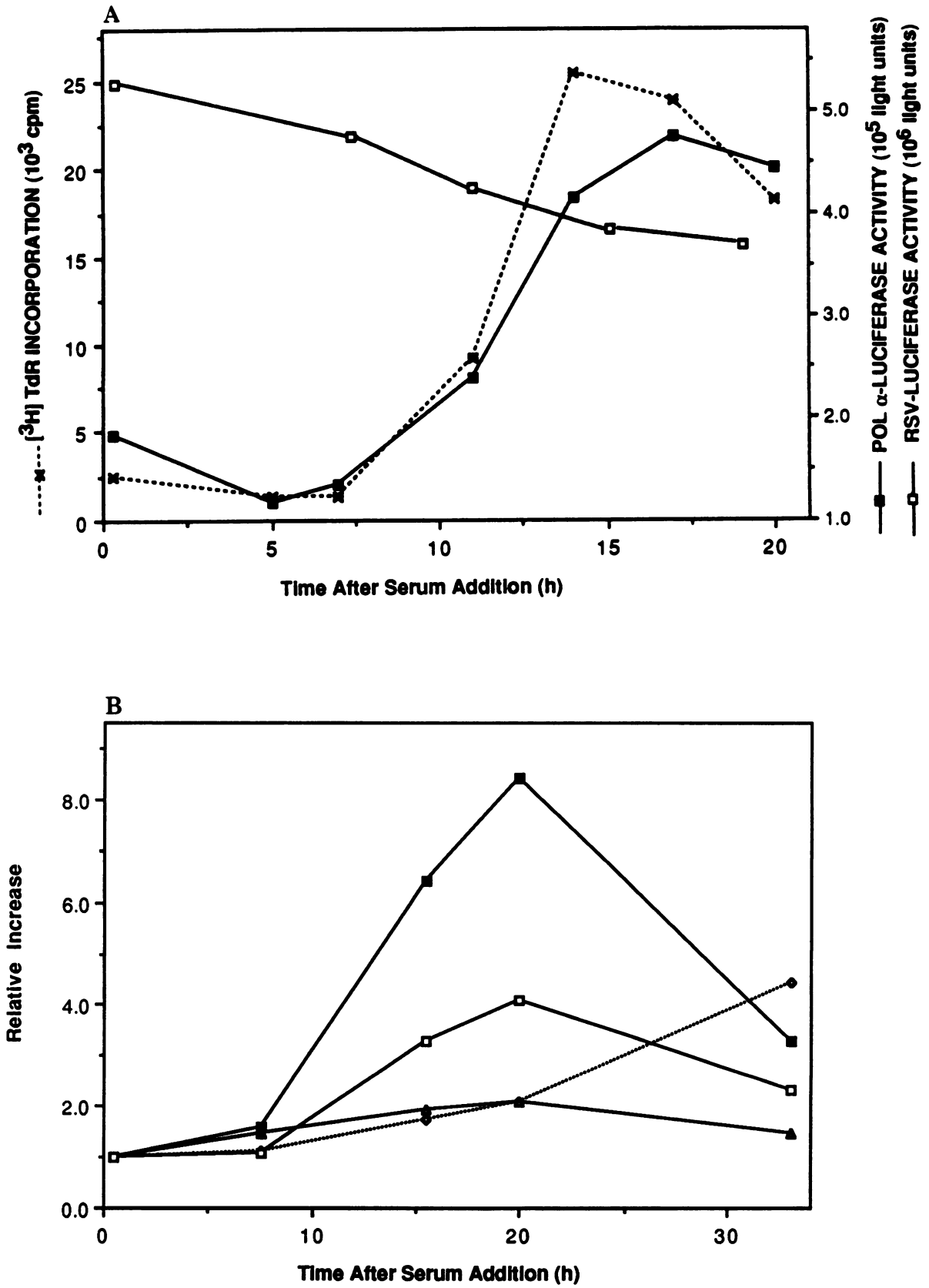


FIG. 5. The DNA polymerase α upstream sequence confers a late serum response upon the luciferase reporter. Experiments were performed as described in Materials and Methods. (A) Serum response profiles of luciferase activity from $CaPO_4$-transfected cells. Pol α-luciferase activity and RSV-luciferase activity represent the luciferase activities derived from the test plasmids pDPALΔ5′ and pRSVLΔ5′,

and β-galactosidase could not distinguish between processed and unprocessed RNA because the simian virus 40 splice site provided by the vector was outside of the regions of homology to these probes. Therefore, the protection of these probes from S1 nuclease provides information about the level of the total luciferase or β-galactosidase RNA and not specifically the level of the mature messages. Cell extracts were prepared in parallel for the determination of ^{3}H-thymidine incorporation (data not shown).

Upon the readdition of serum, β-galactosidase RNA was induced. Without further characterization of this response, pRSVβ-gal was considered to be an inappropriate control, so luciferase RNA levels were normalized to the endogenous mouse polymerase α mRNA levels. Therefore, no quantitative comparison could be made between the level of luciferase RNA and the level of luciferase activity, which was normalized to β-galactosidase activity (Fig. 6).

The induction profiles of luciferase RNA as expressed from the constructs pDPALΔ5′ and pDPA-ScLΔ5′ (Fig. 7, lanes 1 to 3 and 4 to 6) closely followed the induction profile of the endogenous DNA polymerase α mRNA. Luciferase RNA expressed from the 3′-deletion construct pSc-AvLΔ5′ (Fig. 7, lanes 7 to 9) was also induced in parallel with the endogenous polymerase mRNA and to a level similar to that expressed from pDPA-ScLΔ5′, the −116 deletion construct. Finally, the construct pDPA-AvLΔ5′ (−45) produced an unexpectedly high level of luciferase RNA in serum-starved cells (Fig. 7, lane 10). This level had decreased by 5 h of serum stimulation and then underwent a modest increase between 5 and 18 h (Fig. 7, lanes 11 and 12).

These results demonstrate that the *Ava*II fragment (−1515 to +45) of the DNA polymerase α upstream sequence can confer an appropriate response to serum addition upon a reporter gene at the level of RNA. Furthermore, the *Sac*II-*Ava*II fragment (−116 to +45) of the polymerase gene also imposes appropriate serum induction kinetics upon the expression of reporter RNA, although to a lower magnitude than directed by the *Ava*II fragment. A larger 5′ deletion (pDPA-AvLΔ5′) and a 3′ deletion (pSc-AvLΔ5′) of the DNA polymerase α upstream sequence not only lower the magnitude of the serum induction of reporter RNA but may also alter the kinetics of this response.

DISCUSSION

We have isolated a human genomic clone that contains the promoter of the DNA polymerase α gene. By locating the start site of this gene, we have confirmed our previous conclusion that the cDNA clone contains the translational start ATG and therefore the entire coding sequence (63). The promoter of this gene is contained within 248 bp upstream of the transcription start site. Deletion through this sequence reduces promoter activity in cycling cells, serum-starved cells, and cells stimulated with serum. Furthermore, deletion through this region reduces the magnitude of the induced expression observed in response to serum stimulation of serum-starved cells.

The DNA polymerase α promoter sequence is typical of mammalian promoters in that it is GC rich and contains a number of potential transcription factor-binding sites (46). Although there is no apparent TATA sequence, transcription initiation occurs within a small region, predominantly within 45 to 47 nucleotides upstream of the translation start. Immediately upstream of the proposed transcription initiation site, 5 of 6 nucleotides match the TATA-less initiator signal of murine lymphocyte-specific terminal deoxynucleotidyltransferase (55). Of 11 nucleotides spanning the polymerase initiation site, 7 match the dhfr initiation signal (44). Therefore, initiation of transcription of the DNA polymerase α gene may occur by a mechanism similar to that of terminal deoxynucleotidyltransferase or dhfr.

Our studies of the DNA polymerase α promoter in cycling cells indicate that the sequence motifs, similar to consensus sequences of transcription factor-binding sites, may control the transcription rate of this gene. The inverted CCAAT sequence has precedent in both the human and herpes simplex virus tk promoters (17, 43). In the DNA polymerase α promoter, a 28-bp sequence (−45 to −17) includes the inverted CCAAT and enhances transcriptional activity in cycling cells 10-fold over the minimal activity of the *Nar*I (−17) deletion construct (Fig. 3). Nuclear protein binds this region specifically (Fig. 4A). While the identity of this protein has not been established, it could be the previously characterized CTF/NF-I or CP1 (46). Nuclear proteins also bind to the potential Ap2, Ap1, E2F, and Sp1 transcription factor-binding sites and the unique PAL sequence in vitro, and deletion of sequences containing these sites reduces in vivo promoter activity. The possibility that other sequence elements within these regions provide the control functions has not been ruled out, and the capacity of purified Ap2, Ap1, Sp1, or CTF proteins to bind the DNA polymerase α promoter sequences has not been tested. However, purified E2F does bind with high affinity to a synthetic oligomer of the sequence found in the DNA polymerase α promoter, -TTTGGCGC- (48). The sequence with similarity to SRE at −525 of the polymerase gene can be deleted without detriment to the in vivo transcriptional activity in actively cycling HeLa cells (Fig. 3) and yet is readily footprinted by nuclear proteins in vitro (data not shown). Therefore, the function of this polymerase motif is unclear. Further analysis of finer deletions, point mutations, and linker-scanner mutations should clarify the contribution of the potential transcription factor-binding sites to the activity of the DNA polymerase α promoter in cycling cells.

In response to the addition of serum to serum-starved cells, the DNA polymerase α genomic sequence from −1515 to +45 directs an increase in expression of the luciferase reporter activity to an appropriate magnitude and with the appropriate kinetics for the induction of endogenous DNA polymerase α activity and protein (Fig. 5A) (62). The induction could be mediated by transcriptional or posttranscriptional mechanisms.

Evidence that the upstream sequence of the DNA polymerase α gene directs a transcriptional response to serum addition is provided by analysis of the serum responses of polymerase promoter deletion constructs. Complete elimination of the serum induction should have resulted in a decreasing luciferase activity response profile like that di-

respectively, and normalized to β-galactosidase activity derived from the cotransfected control plasmid, pRSVβ-gal. DNA synthesis is represented by [^{3}H]TdR ([^{3}H]thymidine) incorporation. (B) Relative serum response profiles from electroporated cells for total protein (◇), luciferase specific activity (■), β-galactosidase specific activity (△), and luciferase specific activity normalized to β-galactosidase specific activity (□). Approximately 6.3 × 10^6 cells in 0.5 ml were transfected with 10 μg of pDPALΔ5′ and 30 μg of pRSVβ-gal.

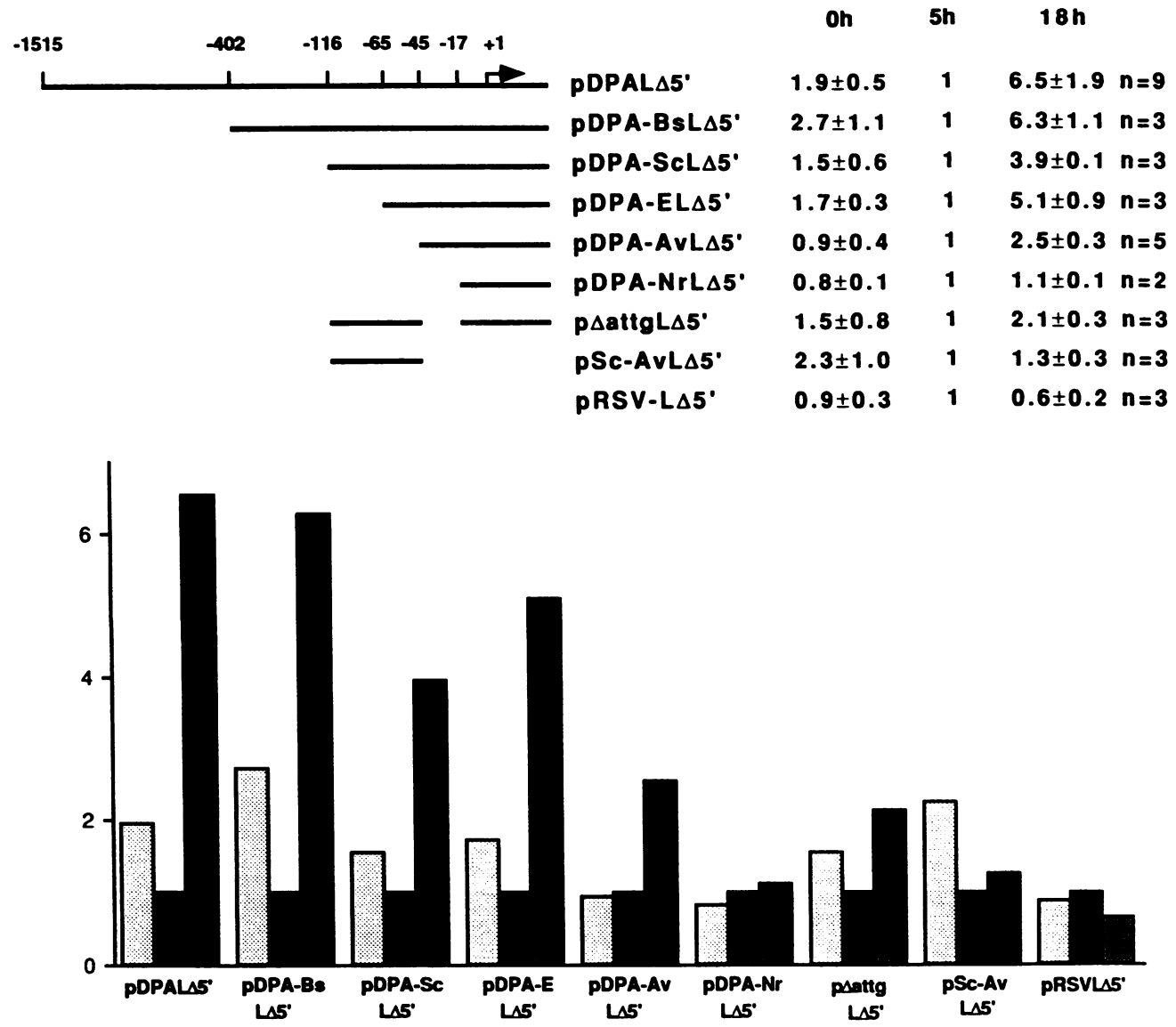


FIG. 6. Relative serum response profiles of luciferase activity produced by DNA polymerase α deletion constructs. NIH 3T3 cells were transfected by $CaPO_4$ coprecipitation. At the top of the figure, the portion of the polymerase upstream sequence contained in each transfected construct is represented by the line drawing. The name of each construct is followed by the average luciferase activity ± standard deviation at 0 and 18 h after serum addition relative to the luciferase activity produced by the indicated construct at 5 h after serum addition. The number (n) of independent experiments in which a construct was tested is indicated. In each experiment, each construct was tested in duplicate. All luciferase activities were normalized to β-galactosidase activity before comparison with activities at other time points. At the bottom of the figure, the luciferase activity produced by each construct at 0, 5, and 18 h (bars from left to right in each group) after serum addition relative to the value at 5 h is depicted in the bar graph.

rected by the Rous sarcoma virus long terminal repeat (Fig. 5A). While none of the polymerase 5′-deletion constructs produced a similarly decreasing serum response (Fig. 6), successive deletion through the polymerase promoter does alter the magnitude of the response. These results suggest that individual sequence regions upstream of the transcription start site can modulate the serum response, probably by modulating the transcription rate. However, no single sequence element was found to direct the full response.

The *Bss*HII (−402) deletion construct produces nearly the same magnitude of serum induction as the longest (*Ava*II, −1515) construct does (Fig. 6). This result minimizes the contribution of the sequence with similarity to SRE (−525) to the control of the late serum induction. This is not surprising, considering that the canonical SRE is associated with the immediately-early serum response of c-*fos* (61) and the muscle-specific expression of α-actins (47). However, over the course of serum induction of human IMR90 cells, an increase in the level of in vitro nuclear protein-binding activity to this motif was observed with the kinetics of a late response (data not shown). This contradiction may reflect a difference in the utilization of sequence elements between IMR90 and NIH 3T3 cells or that the sensitivity of the in vivo assay used in NIH 3T3 cells is insufficient.

The magnitude of the serum induction produced from the *Eag*I (−65) deletion construct is greater than that of the

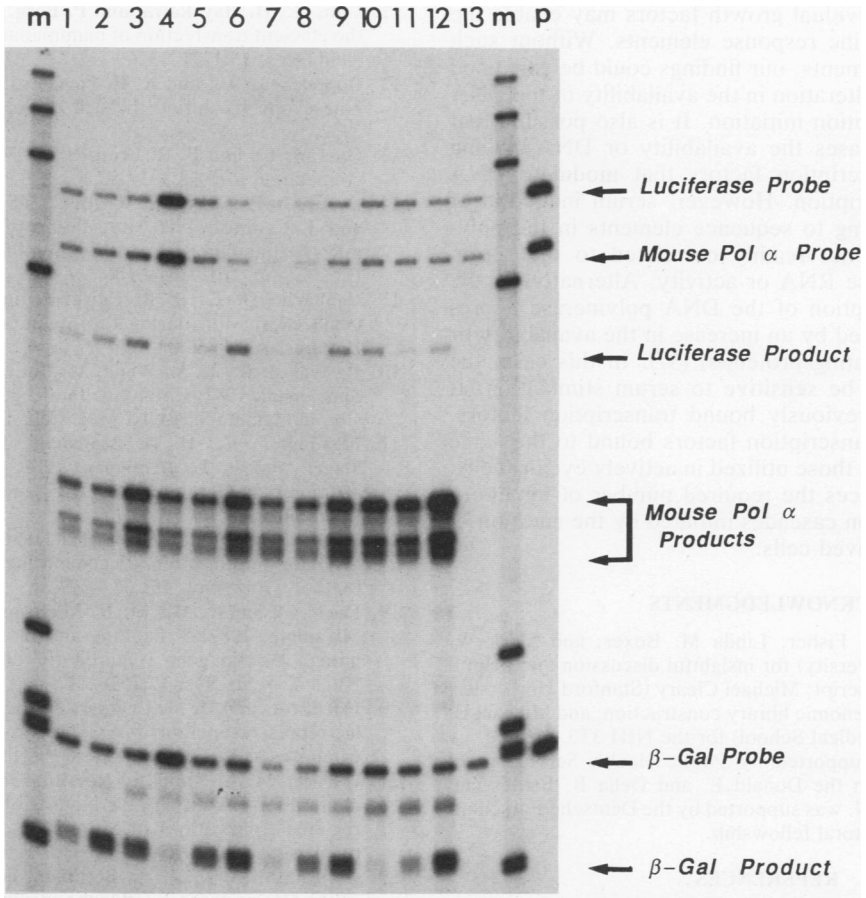


FIG. 7. Level of luciferase RNA produced by DNA polymerase α constructs in response to serum addition. For each polymerase construct tested, two 0.5-ml aliquots of NIH 3T3 cells (2.5×10^7/ml) were electroporated at 280 V with 30 μg of test plasmid and 10 μg of pRSVβ-gal and the two aliquots were pooled before distribution into plates. Electroporated cells were serum starved and stimulated. Total RNA was prepared at 0, 5, and 20 h after serum addition and analyzed for S1 nuclease protection of luciferase, mouse polymerase α, and β-galactosidase cDNA probes. Lanes 1, 2, and 3, 0-, 5-, and 20-h time points for cells transfected with pDPALΔ5′; lanes 4, 5, and 6, 0-, 5-, and 20-h time points for cells transfected with pDPA-ScLΔ5′; lanes 7, 8, and 9, 0-, 5-, and 20-h time points for cells transfected with pSc-AvLΔ5′; lanes 10, 11, and 12, 0-, 5-, and 20-h time points for cells transfected with pDPA-AvLΔ5′; lane 13, S1 nuclease analysis of tRNA; lanes m, 5′-end-labeled φx174 *Hae*III markers; lane p, cDNA probes. Lanes 1 to 9 utilized 20 μg of RNA, and lanes 10 to 13 utilized 40 μg of RNA.

*Sac*II (−116) deletion construct (Fig. 6). This finding may reflect the presence of sequence elements within the *Sac*II-*Eag*I fragment that negatively regulate the magnitude of the response. The dominant sequence elements that positively influence the magnitude of the late serum response appear to be located between −65 and −17. Additional positive late serum response sequences may be distributed throughout the promoter region. Two candidate positive regulatory elements upstream of −116 are the potential Ap1 and E2F transcription factor-binding sites. Ap1-binding sites have been hypothesized to participate in late gene induction (13), and E2F binding to the c-*myc* promoter has been shown to influence that gene's serum response (48). Another possible positive regulatory element is the inverted CCAAT sequence. Internal deletion of this region (pΔattgLΔ5′) reduces the magnitude of the serum induction. However, no change in the extent of in vitro binding of nuclear protein to this sequence was observed during serum induction of IMR90 cells (data not shown). Again, this in vitro result may reflect a difference between cell lines, especially since serum regulation of the interaction between nuclear DNA-binding proteins isolated from BALB/c 3T3 cells and a CCAAT sequence of the human tk gene has been reported (32). As is the case for the sequence elements that control expression in actively cycling cells, precise definition of the sequence elements modulating serum induction will require further analysis by point and linker-scanner mutations.

Luciferase RNA expressed from the polymerase promoter sequence is induced with kinetics similar to that of the endogenous DNA polymerase α mRNA (Fig. 7, lanes 1 to 3). Promoter deletions alter the magnitude and possibly the kinetics of the RNA response. These data support our conclusion that the polymerase α promoter controls, at least in part, the late serum response of this gene. However, studies using a transfected control that is not serum induced at the level of RNA would allow a more detailed analysis of the mechanism of the late serum induction of DNA polymerase α.

The sequence elements in the upstream region of the polymerase α gene that modulate serum response are either the same as those required for promoter activity in cycling cells or distinct and intermingled with those elements. Fu-

ture studies with individual growth factors may enable the identification of specific response elements. Without such specific sequence elements, our findings could be explained by a serum-induced alteration in the availability of the DNA template for transcription initiation. It is also possible that serum addition increases the availability or DNA-binding activity of the transcription factors that modulate DNA polymerase α transcription. However, serum induction of nuclear protein binding to sequence elements in the polymerase promoter is not readily correlated to the serum induction of luciferase RNA or activity. Alternatively, the induction of transcription of the DNA polymerase α promoter may be mediated by an increase in the availability or activity of a coactivating protein(s) (37). In this case, the coactivator(s) would be sensitive to serum stimulation. It could act through previously bound transcription factors, possibly the same transcription factors bound to the same sequence elements as those utilized in actively cycling cells. This hypothesis reduces the required number of targets of the signal transduction cascades initiated by the addition of serum to serum-deprived cells.

ACKNOWLEDGMENTS

We thank Robert P. Fisher, Linda M. Boxer, and Mark W. Walberg (Stanford University) for insightful discussion and helpful comments on the manuscript; Michael Cleary (Stanford University) for technical advice in genomic library construction; and Michael E. Greenberg (Harvard Medical School) for the NIH 3T3 cell line.

This research was supported by Public Health Service grant CA14835 and gifts from the Donald E. and Delia B. Baxter and Beekhuis Funds. H.-P.N. was supported by the Deutsche Forschungsgemeinschaft postdoctoral fellowship.

REFERENCES

1. **Almendral, J. M., D. Sommer, H. Macdonald-Bravo, J. Burckhardt, J. Perera, and R. Bravo.** 1988. Complexity of the early genetic response to growth factors in mouse fibroblasts. Mol. Cell. Biol. **8**:2140–2148.
2. **Artishevsky, A., A. Grafsky, and A. S. Lee.** 1985. Isolation of a mammalian sequence capable of conferring cell cycle regulation to a heterologous gene. Science **230**:1061–1063.
3. **Artishevsky, A., S. Wooden, A. Sharma, E. Resendez, Jr., and A. S. Lee.** 1987. Cell cycle regulatory sequences in a hamster histone promoter and their interactions with cellular factors. Nature (London) **328**:823–827.
4. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Siedman, and K. Struhl (ed.).** 1988. Current protocols in molecular cloning. Greene Publishing Associates, Brooklyn, N.Y.
5. **Berger, S. L., and A. R. Kimmel (ed.).** 1987. Methods in Enzymology, vol. 152. Academic Press, Inc., San Diego, Calif.
6. **Blanchard, J.-M., M. Piechaczyk, C. Dani, J.-C. Chambard, A. Franchi, J. Pouyssegur, and P. Jeanteur.** 1985. c-*myc* gene is transcribed at high rate in G_0-arrested fibroblasts and is post-transcriptionally regulated in response to growth factors. Nature (London) **317**:443–445.
7. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
8. **Calzone, F. J., R. J. Britten, and E. H. Davidson.** 1987. Mapping of gene transcripts by nuclease protection assays and cDNA primer extension. Methods Enzymol. **152**:611–632.
9. **Cathala, G., J.-F. Savouret, B. Mendez, B. L. West, M. Karin, J. A. Martial, and J. D. Baxter.** 1983. Laboratory methods: a method for isolation of intact translationally active ribonucleic acid. DNA **2**:329–335.
10. **Chang, C.-D., L. Ottavio, S. Travali, K. E. Lipson, and R. Baserga.** 1990. Transcriptional and posttranscriptional regulation of the proliferating cell nuclear antigen gene. Mol. Cell. Biol. **10**:3289–3296.
11. **Chu, G., H. Hayakawa, and P. Berg.** 1987. Electroporation for the efficient transfection of mammalian cells with DNA. Nucleic Acids Res. **15**:1311–1326.
12. **Coppock, D. L., and A. B. Pardee.** 1987. Control of thymidine kinase mRNA during the cell cycle. Mol. Cell. Biol. **7**:2925–2932.
13. **Curran, T., and B. R. Franza, Jr.** 1988. Fos and jun: the Ap1 connection. Cell **55**:395–397.
14. **Dean, M., R. A. Levine, W. Ran, M. S. Kindy, G. E. Sonenshein, and J. Campisi.** 1986. Regulation of *c-myc* transcription and mRNA abundance by serum growth factors and cell contact. J. Biol. Chem. **261**:9161–9166.
15. **Denhardt, D. T., D. R. Edwards, and C. L. J. Parfett.** 1986. Gene expression during the mammalian cell cycle. Biochim. Biophys. Acta **865**:83–125.
16. **De Wet, J. R., K. V. Wood, M. DeLuca, D. R. Helinski, and S. Subramani.** 1987. Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. **7**:725–737.
17. **Flemington, E., H. D. Bradshaw, Jr., V. Traina-Dorge, V. Slagel, and P. L. Deininger.** 1987. Sequence, structure and promoter characterization of the human thymidine kinase gene. Gene **52**:267–277.
18. **Fordis, M. C., and B. H. Howard.** 1986. Use of the CAT reporter gene for optimization of gene transfer into eukaryotic cells. Methods Enzymol. **151**:383–397.
19. **Frederickson, R. M., M. R. Micheau, A. Iwamoto, and N. G. Miyamoto.** 1989. 5′ flanking and first intron sequences of the human β-actin gene required for efficient promoter activity. Nucleic Acids Res. **17**:253–271.
20. **Fried, M., and D. M. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acid Res. **9**:6505–6523.
21. **Garner, M. M., and A. Revzin.** 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions. Applications to components of E. coli lactose operon regulatory system. Nucleic Acids Res. **9**:3047–3060.
22. **Goldsmith, M. E., C. A. Beckman, and K. H. Cowan.** 1986. 5′ nucleotide sequences influence serum-modulated expression of a human dihydrofolate reductase minigene. Mol. Cell. Biol. **6**:878–886.
23. **Gorman, C., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
24. **Greenberg, M. E., and E. B. Ziff.** 1984. Stimulation of 3T3 cells induces transcription of the c-fos proto-oncogene. Nature (London) **311**:433–437.
25. **Gudas, J. M., G. B. Knight, and A. B. Pardee.** 1988. Nuclear posttranscriptional processing of thymidine kinase mRNA at the onset of DNA synthesis. Proc. Natl. Acad. Sci. USA **85**:4705–4709.
26. **Hennighausen, L., and H. Lubon.** 1987. Interaction of protein with DNA in vitro. Methods Enzymol. **152**:721–735.
27. **Ito, M., and S. E. Conrad.** 1990. Independent regulation of thymidine kinase mRNA and enzyme levels in serum-stimulated cells. J. Biol. Chem. **265**:6954–6960.
28. **Ito, M., A. Sharma, A. S. Lee, and R. Maxon.** 1989. Cell cycle regulation of H2b histone octamer DNA-binding activity in Chinese hamster lung fibroblasts. Mol. Cell. Biol. **9**:869–873.
29. **Jones, N. C., P. W. J. Rigby, and E. B. Ziff.** 1988. Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. Genes Dev. **2**:267–281.
30. **Kadesch, T., and P. Berg.** 1986. Effects of the position of the simian virus 40 enhancer on expression of multiple transcription units in a single plasmid. Mol. Cell. Biol. **6**:2593–2601.
31. **Kim, Y. K., S. Wells, Y.-F. C. Lau, and A. S. Lee.** 1988. Sequences contained within the promoter of the human thymidine kinase gene can direct cell-cycle regulation of heterologous fusion genes. Proc. Natl. Acad. Sci. USA **85**:5894–5898.
32. **Knight, G. B., J. M. Gudas, and A. B. Pardee.** 1987. Cell-cycle-specific interaction of nuclear DNA-binding proteins with a CCAAT sequence from the human thymidine kinase gene. Proc. Natl. Acad. Sci. USA **84**:8350–8354.

33. **Konig, H., H. Ponta, U. Rahmsdorf, M. Buscher, A. Schonthal, H. J. Rahmsdorf, and P. Herrlich.** 1989. Autoregulation of fos: the dyad symmetry element as the major target of repression. EMBO J. **8**:2559–2566.
34. **Krug, M. S., and S. L. Berger.** 1987. First-strand cDNA synthesis primed with oligo(dT). Methods Enzymol. **152**:316–325.
35. **Lee, K. A. W., T.-Y. Hai, L. SivaRaman, B. Thimmappaya, H. C. Hurst, N. C. Jones, and M. R. Green.** 1987. A cellular protein, activating transcription factor, activates transcription of multiple E1A-inducible adenovirus early promoters. Proc. Natl. Acad. Sci. USA **84**:8355–8359.
36. **Leung, S., and N. G. Miyamoto.** 1989. Point mutational analysis of the human c-fos serum response factor binding site. Nucleic Acids Res. **17**:1177–1195.
37. **Lewin, B.** 1990. Commitment and activation at Pol II promoters: a tail of protein-protein interactions. Cell **61**:1161–1164.
38. **Lieberman, H. B., P.-F. Lin, D.-B. Yeh, and F. H. Ruddle.** 1988. Transcriptional and posttranscriptional mechanisms regulate murine thymidine kinase gene expression in serum-stimulated cells. Mol. Cell. Biol. **8**:5280–5291.
39. **Loenen, W. A. M., and F. R. Blattner.** 1983. Lambda charon vectors (Ch32, 33, 34 and 35) adapted for DNA cloning in recombination-deficient hosts. Gene **26**:171–179.
40. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
42. **McGregor, G. R., A. E. Moss, J. F. Burke, and T. C. Caskey.** 1987. Histochemical staining of clonal mammalian cell lines expressing E. coli beta galactosidase indicates heterogeneous expression of the bacterial gene. Somatic Cell Mol. Genet. **13**:253–265.
43. **McKnight, S. L., R. C. Kingsbury, A. Spence, and M. Smith.** 1984. The distal transcription signals of the herpesvirus tk gene share a common hexanucleotide control sequence. Cell **37**:253–262.
44. **Means, A. L., and P. J. Farnham.** 1990. Transcription initiation from the dihydrofolate reductase promoter is positioned by HIP1 binding at the initiation site. Mol. Cell. Biol. **10**:653–661.
45. **Minty, A., and L. Kedes.** 1986. Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionarily conserved repeated motif. Mol. Cell. Biol. **6**:2125–2136.
46. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245**:371–378.
47. **Miwa, T., and L. Kedes.** 1987. Duplicated CArG box domains have positive and mutually dependent regulatory roles in expression of the human α-cardiac actin gene. Mol. Cell. Biol. **7**:2803–2813.
47a. **Moore, A., and T. Wang.** Unpublished data.
48. **Mudryj, M., S. W. Hiebert, and J. R. Nevins.** 1990. A role for the adenovirus inducible E2F transcription factor in a proliferation dependent signal transduction pathway. EMBO J. **9**:2179–2184.
49. **Pardee, A. B.** 1989. G_1 events and regulation of cell proliferation. Science **246**:603–608.
50. **Rosenthal, N.** 1987. Identification of regulatory elements of cloned genes with functional assays. Methods Enzymol. **152**: 704–720.
51. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
51a. **Schimke, R.** Personal communication.
52. **Schumperli, D.** 1988. Multilevel regulation of replication-dependent histone genes. Trends Genet. **4**:187–190.
53. **Shaw, P. E., S. Frasch, and A. Nordheim.** 1989. Repression of c-fos transcription is mediated through $p67^{SRF}$ bound to the SRE. EMBO J. **8**:2567–2574.
54. **SivaRaman, L., and B. Thimmappaya.** 1987. Two promoter-specific host factors interact with adjacent sequences in an EIA-inducible adenovirus promoter. Proc. Natl. Acad. Sci. USA **84**:6112–6116.
55. **Smale, S. T., and D. Baltimore.** 1989. The "initiator" as a transcription control element. Cell **57**:103–113.
56. **Stewart, C. J., M. Ito, and S. E. Conrad.** 1987. Evidence for transcriptional and post-transcriptional control of the cellular thymidine kinase gene. Mol. Cell. Biol. **7**:1156–1163.
57. **Subramaniam, M., L. J. Schmidt, C. E. Crutchfield III, and M. J. Getz.** 1989. Negative regulation of serum-responsive enhancer elements. Nature (London) **340**:64–70.
58. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82**:1074–1078.
59. **Thompson, C. B., P. B. Challoner, P. E. Neiman, and M. Groudine.** 1986. Expression of the c-*myb* proto-oncogene during cellular proliferation. Nature (London) **319**:374–380.
60. **Travali, S., K. E. Lipson, D. Jaskulski, E. Lauret, and R. Baserga.** 1988. Role of the promoter in the regulation of the thymidine kinase gene. Mol. Cell. Biol. **8**:1551–1557.
61. **Treisman, R.** 1985. Transient accumulation of c-fos RNA following serum stimulation requires a conserved 5′ element and c-fos 3′ sequences. Cell **42**:889–902.
62. **Wahl, A. F., A. M. Geis, B. H. Spain, S. W. Wong, D. Korn, and T. S.-F. Wang.** 1988. Gene expression of human DNA polymerase α during cell proliferation and the cell cycle. Mol. Cell. Biol. **8**:5016–5025.
63. **Wong, S. W., A. F. Wahl, P.-M. Yuan, N. Arai, B. E. Pearson, K.-I. Arai, D. Korn, M. W. Hunkapiller, and T. S.-F. Wang.** 1988. Human DNA polymerase α gene expression is cell proliferation dependent and its primary structure is similar to both prokaryotic and eukaryotic replicative DNA polymerases. EMBO J. **7**:37–47.