*Gene Expression*, Vol. 4, 265-279, 1995
Printed in the USA. 

1052-2166/95 $5.00 + .00


# An Upstream Activator Sequence Regulates the Murine *Pgk-1* Promoter and Binds Multiple Nuclear Proteins

LESLIE C. SUTHERLAND,* RENÉ ST-ARNAUD,† AND MICHAEL W. McBURNEY*[1]

*Departments of Biology and Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Ontario, Canada K1H 8M5
†Genetic Unit, Shriners Hospital for Crippled Children, Montreal, Quebec, Canada

The murine *Pgk-1* gene is driven by a strong promoter that is regulated by a 304 bp upstream activator sequence (UAS). The activity of the UAS is high in undifferentiated embryonal carcinoma cells but declines when these cells are induced to differentiate with retinoic acid. The effect of the UAS on promoter activity is particularly striking when the activity of the *Pgk-1* promoter is assayed following its integration into the genome, suggesting that it may function by regulating chromatin strucure in the region of the core promoter. Three sites on the UAS bind nuclear proteins. Two of these sites bind factors present in both embryonal carcinoma cells and their differentiated derivatives whereas one site binds factors present only in differentiated cells. There appears to be both cooperation and antagonism in the binding of proteins to different sites in the UAS, suggesting that the activity of the *Pgk-1* promoter is determined by the constellation of proteins assembled upstream of its transcription start site.

Murine *Pgk-1* promoter  Nuclear protein binding  Upstream activator sequence

*Pgk-1* is the murine gene encoding phosphoglycerate kinase (PGK), a key enzyme in glycolysis. This enzyme must be present in all cells, so the *Pgk-1* promoter is expected to be a ubiquitously active one. However, the *Pgk-1* promoter is regulated. The *Pgk-1* gene resides on the X chromosome and becomes transcriptionally silenced when X chromosomes become inactivated in female somatic cells or in male germ cells. The promoters of the inactive *Pgk-1* genes are methylated (Keith et al., 1986; Singer-Sam et al., 1990; Tommasi et al., 1993). In vivo footprinting has indicated that the active *Pgk-1* promoter has a number of bound factors capable of yielding DNaseI or dimethylsulfate footprints whereas the inactive *Pgk-1* promoter does not footprint under the same conditions (Pfeifer et al., 1990; Pfeifer et al., 1991). These results suggest that the *Pgk-1* promoter is regulated by factors that activate rather than factors that repress its activity.

The cloned *Pgk-1* promoter is very active in driving transcription of reporter genes following transfection into mammalian cells (McBurney et al., 1991). The promoter is particularly active in pluripotent embryonal carcinoma and embryonic stem cells. Transgenic animals carrying the *lacZ* reporter gene driven by the *Pgk-1* promoter express β-galactosidase in all tissues, but there was developmental and tissue-specific regulation of this expression and unexpected cell-to-cell variation in expression levels (McBurney et al., 1994). The *Pgk-lacZ* transgene was expressed at particularly high levels in early embryonic cells that were rapidly proliferating, but expression levels were much lower in many differentiated cells.

Our initial studies on the *Pgk-1* promoter indi-

Received August 15, 1994; revision accepted January 3, 1995.
[1]Address correspondence to Michael W. McBurney, Department of Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Ontario, Canada, K1H 8M5.

cated that its activity was contained within a region from −425 bp to +80 bp relative to the first transcription start site (McBurney et al., 1991). This promoter is GC rich and contains no TATA box (Adra et al., 1987). The core promoter from −121 bp to +80 bp contains a number of putative Sp1 binding sites and had modest promoter activity. The 304 bp upstream of the core promoter had enhancer-like activity such that this region increased expression from the core promoter in an orientation- and position-independent fashion. In our initial attempts to more carefully characterize this enhancing activity we cut the 304 bp UAS into two, but neither subfragment had the activity of the intact 304 bp. These results suggest that transcription factors bind throughout the UAS and might cooperate to affect enhanced transcription. We report our analysis of the enhancer for regions that bind nuclear proteins.

## MATERIALS AND METHODS

### *Cell Culture*

P19 cells were cultured and induced to differentiate with retinoic acid (RA) as described (Rudnicki and McBurney, 1987).

Cells were transfected by the method of Chen and Okayama (1987). For transient assays, the test plasmid with the *lacZ* or CAT reporter gene was cotransfected with another plasmid encoding a different reporter gene to control for transfection efficiency. In some experiments, cells were cultured continuously in $3 \times 10^{-7}$ RA prior to or after transfection. Cultures were harvested after 48 h and analyzed for β-galactosidase and chloramphenicol acetyltransferase (CAT) activities (Norton et al., 1985). Activities of genes expressed on the test plasmids were normalized to the activities of the genes expressed on the control plasmids.

Stable transfections were carried out by cotransfecting plasmids carrying the test gene, a control gene, a selectable marker (usually *Pgk-puro* or *Pgk-neo*) (McBurney et al., 1991), and B17 (Adra et al., 1987), the 17 kbp BamH1 genomic fragment containing the first eight exons of the mouse *Pgk-1* gene. The last plasmid enhances the efficiency with which cotransfected plasmids are ligated and integrated into the genomes of P19 cells (McBurney et al., 1995a). After 24 h, the cells were harvested and selected in 2 μg/ml puromycin or 400 μg/ml G418. After 7 days in selective medium, colonies were pooled and cultured in the presence of RA before they were harvested and the levels of β-galactosidase and CAT measured.

### *Nuclear Extract*

P19 nuclei were isolated as outlined previously (Jones-Villeneuve et al., 1983), with the following modification. Cells were harvested in buffer A (10 mM Tris-HCl, pH 8.0, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT) and lysed with the addition of 0.1% NP-40. The resulting mixture was underlayed with 0.8 M sucrose and centrifuged at 1000 × *g* for 10 min. After removal of the supernatant the nuclei were washed by resuspension in at least 2 vol. of buffer A and centrifuged (1000 × *g* for 10 min).

Nuclear extract was obtained using a modification of the procedure of Dignam et al. (1983). The nuclear pellet was resuspended in buffer C (20 mM Tris-HCl, pH 8.0, 1.5 mM $MgCl_2$, 0.42 M NaCl, 25% glycerol, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF) and incubated on ice for 30 min with occasional vortexing. The mixture was then forced five times through a syringe with a 20-ga hypodermic needle. The nuclear debris was pelleted by centrifugation for 20 min in a microfuge (12,000 × *g*). The supernatant was transferred to a new tube and mixed with 0.33 g/ml of $NH_4SO_4$. After 30–60 min of rotation at 4°C the nuclear extract was centrifuged at 12,000 × *g* for 20 min. The pellet was resuspended in approximately one-half vol. of buffer D (20 mM Tris-HCl, pH 8.0, 0.1 M KCl, 20% glycerol, 0.2 mM EDTA, 0.5 mM PMSF) and dialyzed against 100 vol. of the same buffer for 3–6 h. Fractionation of P19 nuclear extract was performed as described (St-Arnaud and Moir, 1993).

### *Gel Mobility Shift and DNase1 Footprinting Assays*

DNA for protein binding assays was derived from a plasmid carrying the intact *Pgk-1* promoter (Adra et al., 1987). This plasmid (see Fig. 2) was digested with EcoR1 and Stu1 to release the 304 bp ES fragment comprising the entire UAS (−425 to −121 bp), with EcoR1 and Alu1 to release the 211 bp EA fragment (−425 to −214 bp), or Alu1 and Stu1 to release the 92 bp AS fragment (−213 to −121 bp). The R1 oligodeoxynucleotide was a double-stranded DNA with the sequence 5′-CTCGTGCAGATGGACAGCACCGCTGAGCAA-3′ and its complement, the sequence comprising the region from −166 to −137 bp of the *Pgk-1* promoter. The R2 oligodeoxynucleotide was also a double-stranded DNA with the sequence 5′-

GATCTGGACGTGACAAATGGAAGTAGCAC GTCA-3′ and its complement, a sequence that includes the region of the *Pgk-1* promoter from −201 to −174 bp.

Gel mobility shift assays (Garner and Revzin, 1981; Fried and Crothers, 1981; Singh et al., 1986) were performed as described (McBurney et al., 1991). All binding reactions were carried out at 23°C using gel-purified radiolabeled DNA.

DNaseI footprinting assays were performed essentially as described by Jones et al. (1985). P19 nuclear extract (40 μg) was incubated with 0.2 ng of $^{32}$P-labeled DNA in buffer [12.5 mM Tris-HCl, pH 7.9, 6.25 mM $MgCl_2$, 0.5 mM DTT, 5 μM $ZnSO_4$, 10% (v/v) glycerol, 0.05% (v/v) NP-40, 0.05 M KCl, 2% polyvinyl alcohol] in the presence of 5 μg of poly(dI-dC) for 15 min at room temperature. One minute before digestion, an equal volume of a mixture containing 5 mM $CaCl_2$ and 10 mM $MgCl_2$ was added to the DNA. DNaseI (Worthington Biochemical Corporation, Freehold, NJ) was used at a 1/1000 dilution of a 10 mg/ml stock for control reactions and a 1/50 dilution for reactions containing extract. Reactions were terminated by the addition of a mixture of 200 mM NaCl, 20 mM EDTA, 1% SDS, 250 μg/ml tRNA. This mixture was extracted with phenol and chloroform and precipitated with 2.5 vol. of ethanol. Pellets were rinsed with 70% ethanol and dried before being loaded onto an 8% polyacrylamide gel. Gels were dried and autoradiographed.

### *In Vitro Mutagenesis*

A plasmid carrying the *Pgk-1* promoter was incubated with two oligodeoxynucleotides with the sequence 5′-AAATAATTAATTTCACGTCCT GCACGACGCGAGC-3′ and 5′-AAAAAATT AATTAGTAGCACGTCTCACTAGTCTC-3′ in a reaction that included vent buffer [10 mM KCl, 20 mM Tris-HCl, pH 8.8, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100], 100 μg/ml BSA, 4 mM $MgSO_4$, 200 mM of each dNTP, 10 ng of plasmid DNA, 50 pmol of both oligonulceotides and 8 units of Vent(exo-) enzyme (New England Biolabs), in a total reaction volume of 100 μl. The reaction proceeded at 97°C for 1 min, 66°C for 1 min, 75°C for 6 min, and was repeated for 28 cycles. The DNA was extracted with phenol and chloroform, precipitated with 0.3 M sodium acetate and ethanol, and rinsed in 70% ethanol. The pellet was resuspended in a buffer of 10 mM Tris-HCl (pH 8.0), 5 mM EDTA, and 0.5% SDS; 1 μl of a 5 μg/μl stock of proteinase K (BRL) was added to the DNA, which was incubated at 37°C for 30 min then 68°C for 10 min. The DNA was extracted with phenol and chloroform, precipitated with 0.3 M sodium acetate in ethanol, and resuspended in 50 μl of water. The DNA was then blunt ended, gene cleaned (Bio/Can Scientific), and ligated. Ligations were performed in 30 μl using 20 μl of the DNA, buffer (0.66 M Tris-HCl, pH 7.6, 50 mM $MgCl_2$, 50 mM DTT, 1 mg/ml BSA, 10 mM hexamminecobalt chloride, 2 mM ATP, 5 mM spermidine-HCl), 3 μl of 24% PEG, and 3 units of T4 DNA ligase. Incubation was at 20°C overnight. TG-1 cells were transformed by electroporation. Plasmids that gave the expected restriction fragments with AsnI (Boehringer Mannheim) were sequenced between the SphI and StuI sites. This Sph-Stu fragment was then removed from the sequenced plasmid and inserted into the same site in a plasmid carrying *Pgk*-CAT.

## RESULTS

### *An Upstream Activator Sequence Regulates* Pgk-1 *Promotor Activity*

Using transient transfection assays in P19 embryonal carcinoma cells, we previously showed that the region of the *Pgk-1* promoter downstream of −121 bp functions as a core promoter and carries a number of sequence motifs known to be binding sites for common transcription factors (McBurney et al., 1991). The region between −121 and −425 bp has enhancer-like activity in transient assays, increasing promoter activity by 10–30-fold when assayed in P19 embryonal carcinoma cells. We compared the activities of *Pgk-1* promoter constructs with and without this 304 bp upstream activator sequence (UAS) in P19 cells that had been treated with retinoic acid (RA) to induce their differentiation (Jones-Villeneuve et al., 1982). In untreated cells the longer construct with the UAS was 10 times more active than the shorter. After treatment with RA the activity of the longer promoter declined progressively until at 4 days it was less than twice as active as the shorter (Fig. 1A). Thus, the activity of the UAS decreased as the cells differentiated. When the two constructs were transfected into rat L6 myoblasts they were almost equally active (data not shown), suggesting that the factors responsible for binding the UAS and enhancing promoter activity are abundant in embryonal carcinoma cells but are rare in some other cell types.

To examine the effects of the UAS in P19 cells at longer times after RA treatment, we cotransfected cells with each construct along with a plasmid

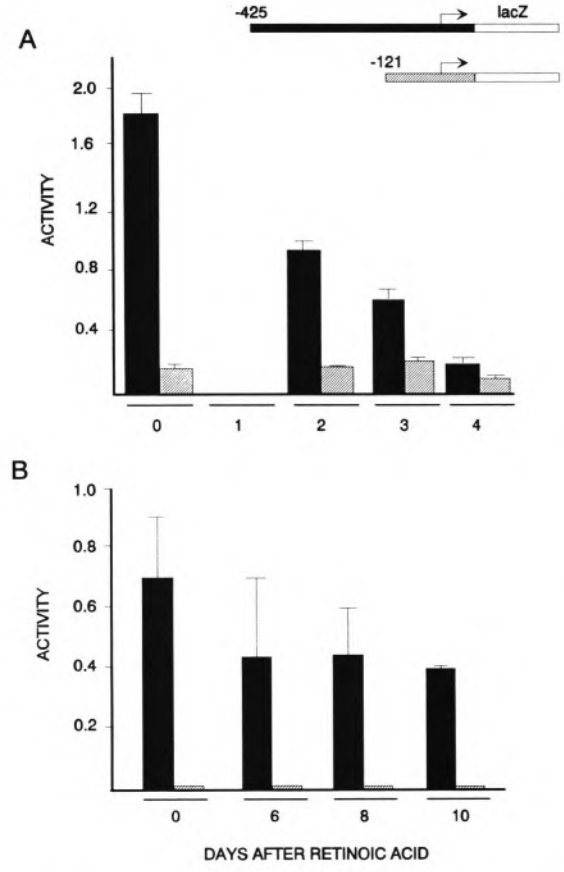



FIG. 1. Regulated expression from the *Pgk-1* promoter carrying the UAS. The *Pgk-1* promoter comprising the region spanning −425 to +80 bp (black bars) and −121 to +80 bp were used to drive the *lacZ* reporter gene. (A) These constructs were cotransfected into P19 cells along with an RSV-driven CAT construct and harvested 48 h later. The P19 cells were treated with retinoic acid (RA) for the number of days shown before cells were harvested for measurement of β-galactosidase and CAT activities. (B) The two test constructs were cotransfected along with a selectible gene, *Pgk-puro*, and stably transformed cells were selected in puromycin. Colonies of transformed cells were pooled and treated with RA for the number of days shown before being harvested for β-galactosidase and CAT activities. The activity shown is that of β-galactosidase normalized to the CAT control.

carrying a selectible gene, *Pgk-puro* (McBurney et al., 1991). Cotransfected cells were selected in puromycin. More than 1000 colonies from each transformation were pooled and the cells were treated with RA. The cells transfected with the short promoter contained undetectable promoter activity whereas the population of cells carrying the longer promoter contained abundant expression from the *Pgk-1* promoter (Fig. 1B). The expression from the longer promoter decreased by about 40% following 6–10 days of RA treatment. Because cotransfected plasmid DNAs become integrated into the host genome (McBurney et al., 1995b), we conclude from these results that the UAS is essential for activity of the *Pgk-1* promoter following its integration into host chromatin, and that once established in an active configuration only modest modulation of its expression occurs during differentiation. The levels of *Pgk-1* mRNA in differentiating cultures of P19 cells drop only slightly upon RA-induced differentiation (unpublished), indicating that the expression of the longer construct following its integration closely resembles the behavior of the endogenous gene.

### *Nuclear Proteins Bind the UAS*

The *Pgk-1* promoter is shown diagrammatically in Fig. 2. The UAS was previously defined by the restriction sites EcoR1 (−425) and Stu1 (−121). This fragment is bisected by an Alu1 site at −214 bp and neither the EcoR1-Alu1 (EA) nor Alu1-Stu1 (AS) fragment carries significant enhancer activity (McBurney et al., 1991).

Our initial search of the UAS region for DNA binding proteins from nuclear extracts of P19 cells indicated an activity binding to the region between −166 and −137 bp, labeled R1 in Fig. 2 (McBurney et al., 1991). These initial experiments involved electrophoretic mobility shift assays (EMSA) with DNA fragments derived from the UAS region. To examine the binding of nuclear proteins to the UAS more closely, we used in vitro DNase1 footprinting procedures. The 304 bp ES fragment was labeled and digested with DNase1 in the presence and absence of nuclear extract from P19 cells. Curiously, no footprint was found over the R1 region; however, a different protected region was identified on both strands of the ES probe about 11 bp upstream of the R1 region. The footprinted regions were of different sizes (Fig. 3A) but they overlapped in the 11 bp region between −193 and −183 bp (Fig. 3B). The DNA binding region identified by these footprint experiments was designated R2 and was located on the map shown in Fig. 2.

The DNase1 footprints on the R2 region were only evident in experiments using relatively large amounts of nuclear extract. We used various procedures to fractionate the nuclear extract but none of the procedures yielded material capable of protecting a region of the ES probe. One interpretation is that the footprint on the R2 region arises from the binding of more than one factor and that these factors bind in a cooperative fashion.

Using EMSA the double-stranded DNA com-

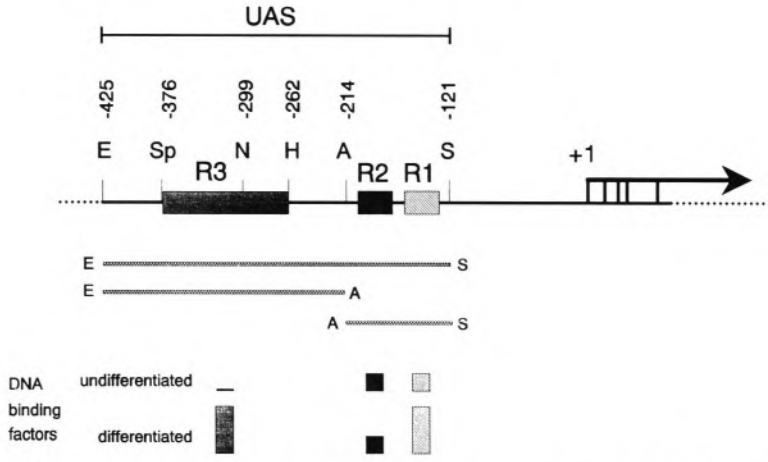



FIG. 2. Diagram of the *Pgk-1* promoter. The solid line represents the DNA sequence upstream of the *Pgk-1* transcription start sites indicated by the arrows starting at +1. The restriction enzyme sites are E, EcoR1; Sp, Sph1; N, Nar1; H Hha1; A, Alu1; and S, Stu1, for which locations are indicated in base pairs upstream of the transcription start site. The UAS refers to the enhancer region defined previously (McBurney et al., 1991). The cross-hatched bars labeled ES, EA, and AS are the probes used for EMSA and DNase1 footprinting experiments. The slashed, black, and gray boxes labeled R1, R2, and R3 refer to the regions within the UAS that bind nuclear proteins. The black box labeled R2 was defined by both DNase1 footprinting and EMSA whereas the slashed and gray boxes labeled R1 and R3 were shown to bind protein by EMSA only; no footprints were obtained over these regions. The locations of R1 and R3 were established by EMSA with oligonucleotide probes or with DNA fragments from the promoter region. The distribution of nuclear DNA binding factors is indicated below the promoter diagram where the intensity of the signal is roughly proportional to the size of the bars.

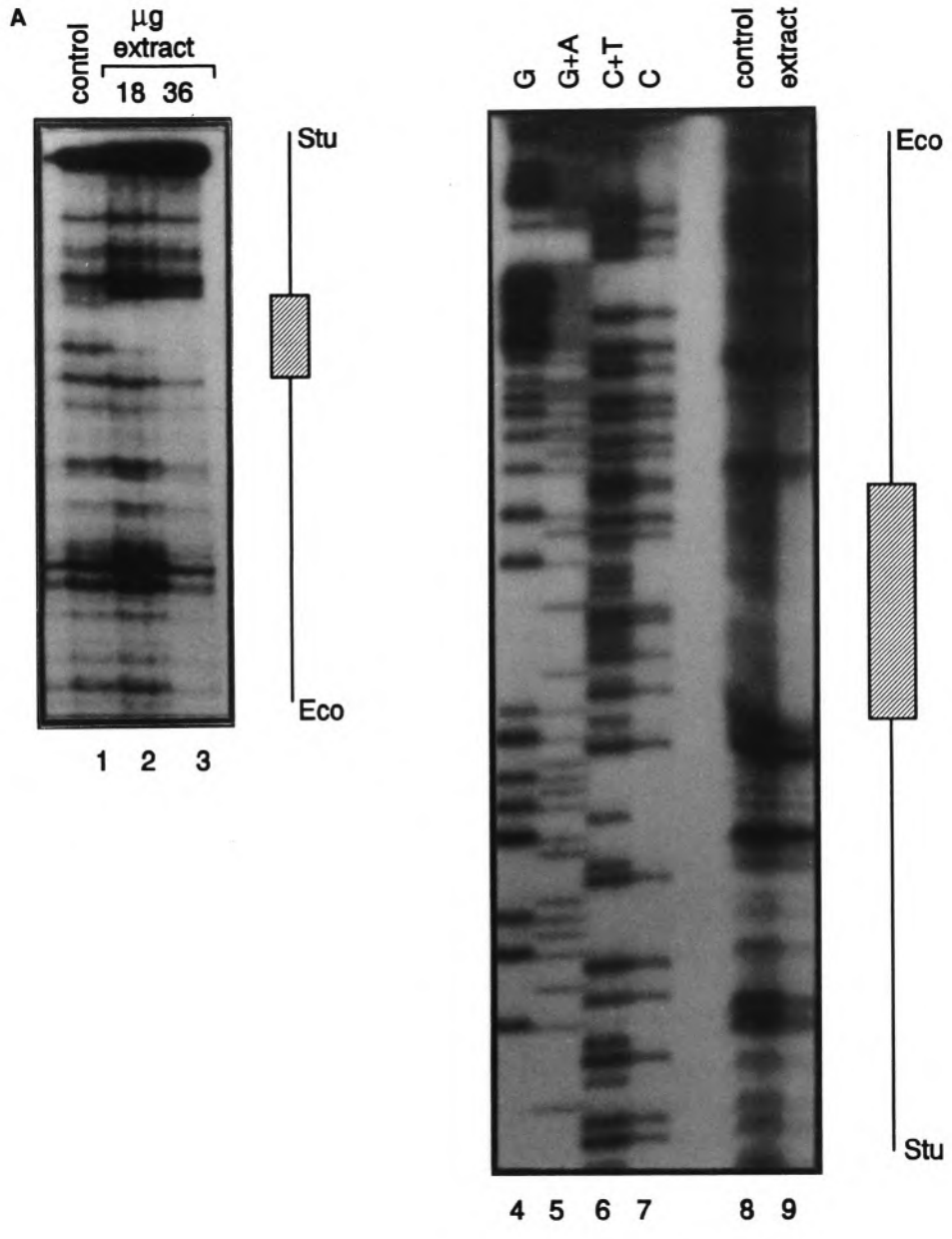
prising the R1 probe yielded four retarded bands (Fig. 4A). The upper three bands were efficiently competed by cold R1 DNA but not by cold R2 DNA. However, the fourth band was more effectively competed by R2 than R1 DNA. This fourth band runs with the same mobility as the retarded band detected with the R2 DNA probe. The R2 probe yielded one prominent band that was competed efficiently by cold R2 DNA but only inefficiently by cold R1 DNA (Fig. 4B). These results suggest that the R1 DNA is capable of binding more than a single protein and that one protein(s) binds to both R1 and R2 but binds to the latter DNA with higher affinity.

The footprint within the R2 region contains the sequence CAAATG that conforms to the consensus sequence 5′-CANNTG-3′ called the E box (Murre et al., 1989). Such regions are known to bind proteins of the basic helix-loop-helix (bHLH) class. An E box consensus is also present in the R1 oligonucleotide (CAGATG), suggesting that the common protein(s) bound to these oligonucleotides may be of the bHLH class. However, transient transfection assays involving the cotransfection of the *Pgk-1* test promoter with expression vectors encoding such bHLH proteins as myoD, Id, and E12 failed to result in modulated expression of the *Pgk-1* promoter (data not shown).

### *DNA Binding Proteins in Differentiated Cells*

The activity of the *Pgk-1* promoter is lower in P19 cells induced to differentiate with RA (Fig. 1). To determine if this modulated expression is accompanied by a change in the nuclear proteins bound to the UAS, DNA fragments from the UAS region were used in EMSA with nuclear extracts derived from P19 cells and P19 cells treated for 7 days with RA (Fig. 5). The ES probe spans the entire UAS from −425 to −121 bp and yielded two bands with both nuclear extracts. In extracts from undifferentiated cells the upper band (form I) predominated whereas in extracts from differentiated cells the lower band (form II) was more abundant.

The most significant difference between the binding activities in these two nuclear extracts was evident with the EA probe that spans −425 to −214 bp. Extracts from differentiated cells yielded an intense retarded band that was much



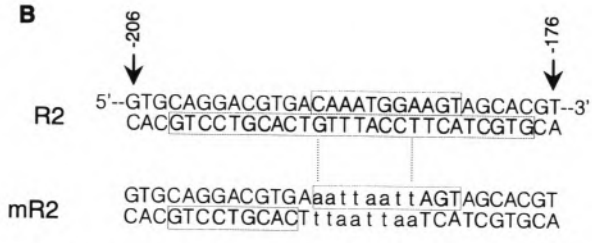




FIG. 3. DNaseI footprint on the UAS of *Pgk-1*. (A) The 304 bp ES probe was labeled with $^{32}P$ at the EcoR1 site (lanes 1–3) or at the StuI site (lanes 8 and 9). Control reactions contained no nuclear extract (lanes 1 and 8) whereas nuclear extract from P19 cells was present at the indicated amounts in lanes 2, 3, and 9. Lanes 4–7 are Maxam and Gilbert sequencing reactions. (B) DNA sequence of the region footprinted and labeled R2. The shaded regions on both the upper and lower strands are those that were protected from DNaseI digestion. The mR2 sequence is that region of the mutant UAS that contains an 8 bp alteration to the R2 region indicated by the sequence in lower case. The shaded regions are those footprinted on the mutant UAS as indicated in Fig. 9B.

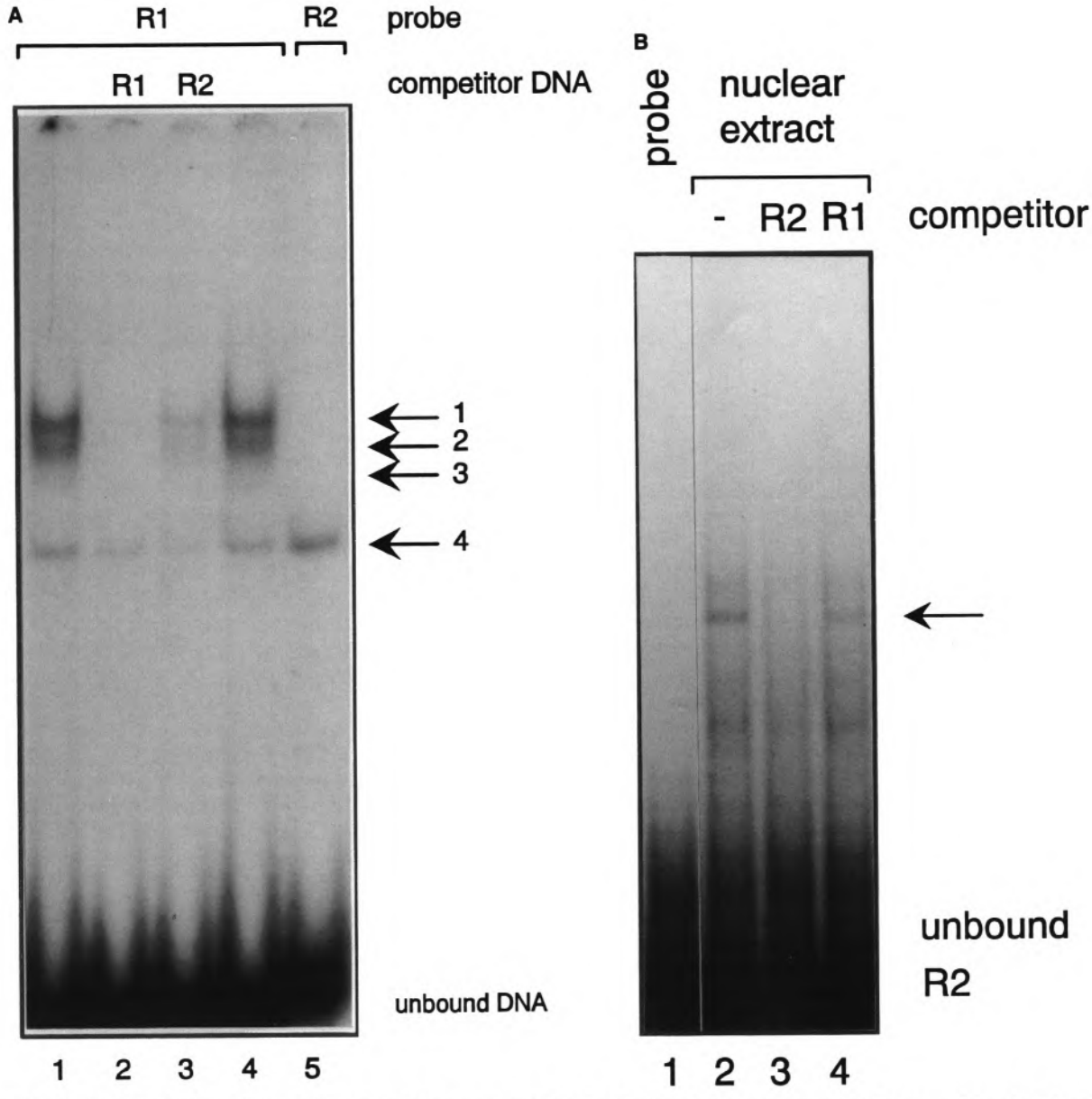



FIG. 4. Proteins that bind R1 and R2 are partially shared. (A) The double-stranded oligodeoxynucleotides comprising the R1 and R2 regions were labeled and mixed with 20 $\mu$g of P19 cell nuclear extract along with a 20-fold molar excess of competitor DNA as indicated. The arrows on the right indicate the four retarded bands that appear with the R1 probe. Note that the one band retarded with the R2 probe has the same mobility as the complex labeled 4 from the lanes carrying the R1 probe. (B) The R2 probe was labeled and mixed with 20 $\mu$g of P19 cell nuclear extract along with a 20-fold molar excess of unlabeled competitor DNA as indicated. The arrow indicates the retarded complex.

weaker in lanes containing extracts of undifferentiated cells. The band detected by the R2 DNA probe was essentially identical for the two extracts. Differentiated cell extracts yielded minor additional bands when challenged with the AS probe. When mixed with differentiated cell extract, the R1 probe yielded more intensely staining bands with different mobilities.

The factor that appears in differentiated cells and binds the EA probe was partially purified by gel filtration column chromatography of nuclear extracts from RA-treated cells (St-Arnaud and Moir, 1993). This material eluted with an apparent molecular weight of less than 65,000. This factor bound both the ES and EA probes but did not bind to the AS DNA (Fig. 6). The region to which this factor binds was delineated by the use of various subfragments for EMSA and competition and is labeled R3 in Fig. 2.

To examine the composition of the two bands

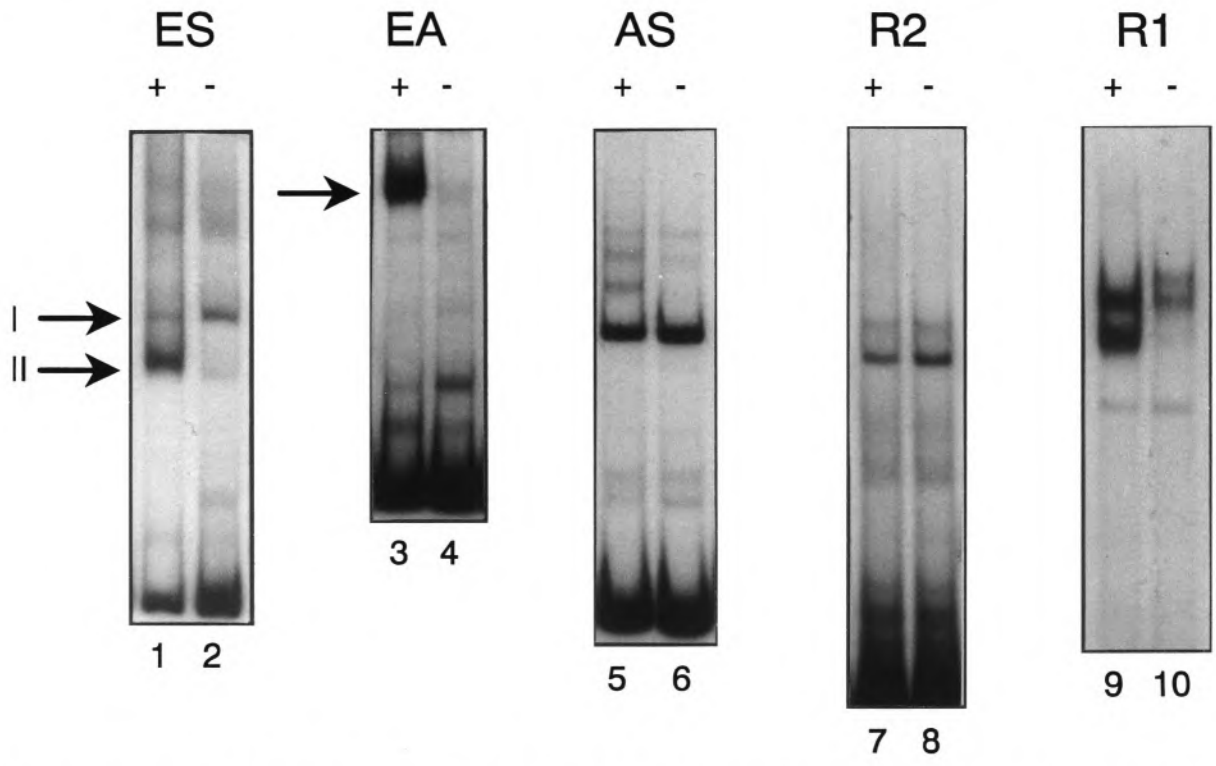



FIG. 5. Nuclear extracts from differentiated P19 cells contain a novel DNA binding factor. Gel mobility shift assays were carried out using 20 $\mu$g of nuclear extract from undifferentiated P19 cells (lanes labeled −) or P19 cells exposed to RA for 7 days (lanes labeled +). The probes used for each experiment are shown on the top of each pair of lanes and are the 304 bp ES fragment, the 211 bp EA fragment, the 92 bp AS fragment, the 34 bp R2 oligonucleotide, and the 30 bp R1 oligonucleotide (see Fig. 2). Arrows on the left indicate the two retarded complexes formed with the ES probe and the arrow to lane 3 shows the novel complex formed with the differentiated cell extract probed with the EA fragment. Gels containing the different probes were electrophoresed for different times and exposed to x-ray film for different times.

detected by the ES probe comprising the entire UAS, we performed a number of mixing and competition experiments. The ES probe yielded two bands labeled I and II (Fig. 7A). In undifferentiated cell extracts form I was abundant (Fig. 7A, lane 3) whereas form II was more intense in extracts from differentiated cells (Fig. 7A, lane 5). Cold EA DNA decreased the abundance of form II (lanes 1 and 6) whereas cold R2 DNA reduced the abundance of form I (lanes 2 and 7). Cold R1 DNA had no effect under the same conditions (data not shown). These results suggest that the two major bands identified by EMSA of the ES probe consisted of factors bound to the proximal (AS) and the distal (EA) regions of the UAS, respectively.

When the partially purified factor from differentiated cells was added to nuclear extracts from undifferentiated cells, form II increased in abundance (Fig. 7B), suggesting that form II consisted of this factor bound to the UAS probe.

Differentiated cell extracts contain elevated levels of activities binding R1 (Fig. 5). To determine whether the partially purified R3 factor binding the EA fragment could bind directly to R1 or influence the binding of other factors to R1, we added this factor to the R1 probe in the presence of extracts from undifferentiated cells (Fig. 8). The partially purified factor neither bound R1 DNA nor altered the binding of other nuclear proteins to this DNA.

### *The R2 Region Appears Critical for Expression Following Stable Integration*

In EMSA assays the R2 DNA yielded one prominent retarded band, but an extensive DNAse1 footprint was detected at this site. In addition, the most abundant retarded band detected with the ES probe was competed with the R2 DNA. To assess the importance of the R2 region for *Pgk-1* promoter activity, we constructed a mutant UAS in which the central 8 bp of the R2 footprint was replaced with an AT-rich sequence

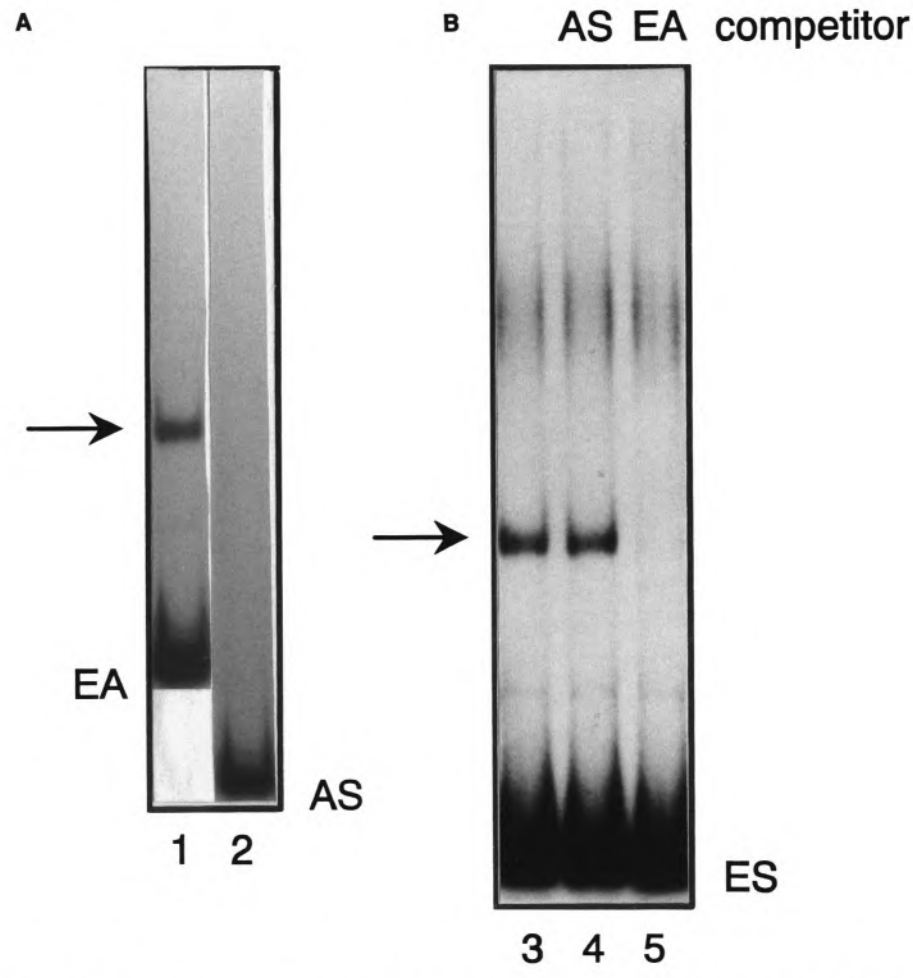



FIG. 6. Differentiated cells contain an activity that binds the R3 region of the UAS. Nuclear extract from differentiated P19 cells was partially purified by gel filtration and used in gel mobility shift assays to determine specificity of binding. (A) The 211 bp EA fragment (lane 1) and the 92 bp AS fragment (lane 2) were incubated with partially purified protein. (B) The 304 bp ES fragment was incubated with partially purified protein along with a 100-fold molar excess of competitor DNA. Lanes: 3, no competitor; 4, unlabeled AS DNA; 5, unlabeled EA DNA. Arrows indicate retarded complexes.

containing the Asn1 restriction site (Fig. 3B). Unexpectedly, the mutant ES probe bound nuclear proteins as detected by EMSA and DNase1 footprinting. The EMSA indicated that the mutant ES probe bound to nuclear proteins with lower affinity than the wild type and that the mobility of the retarded complex was different with the two probes (Fig. 9A). In competition experiments, the bands of both wild-type and mutant AS probes were competed by the wild-type R2 DNA and wild-type AS fragment (Fig. 9B). Thus, at least some of the proteins binding to the mutant R2 site appear to be the same as those that bind the wild-type R2 region. DNase1 footprinting showed a protected region on both strands of the mutant ES probe (Fig. 9C). The protected region on the coding strand encompassed the mutant 8 bp. The noncoding strand had a shorter but nevertheless distinct footprint over a region adjacent to the mutated site (Fig. 3B). Thus, the mutant 8 bp did not eliminate the binding of nuclear proteins but seemed to qualitatively alter the complex formed. It seems likely that the modified EMSA and footprints on the mutant ES probe result from binding of a reduced number of the same factors that bind the wild-type ES probe, although it is possible that the mutation has inadvertantly created a binding site for another DNA binding protein(s). The novel retarded complex seen with the mutant AS probe (Fig. 9B, open arrow) may result from the binding of a new factor to the mutated region.

In transient expression assays the *Pgk-1* pro-

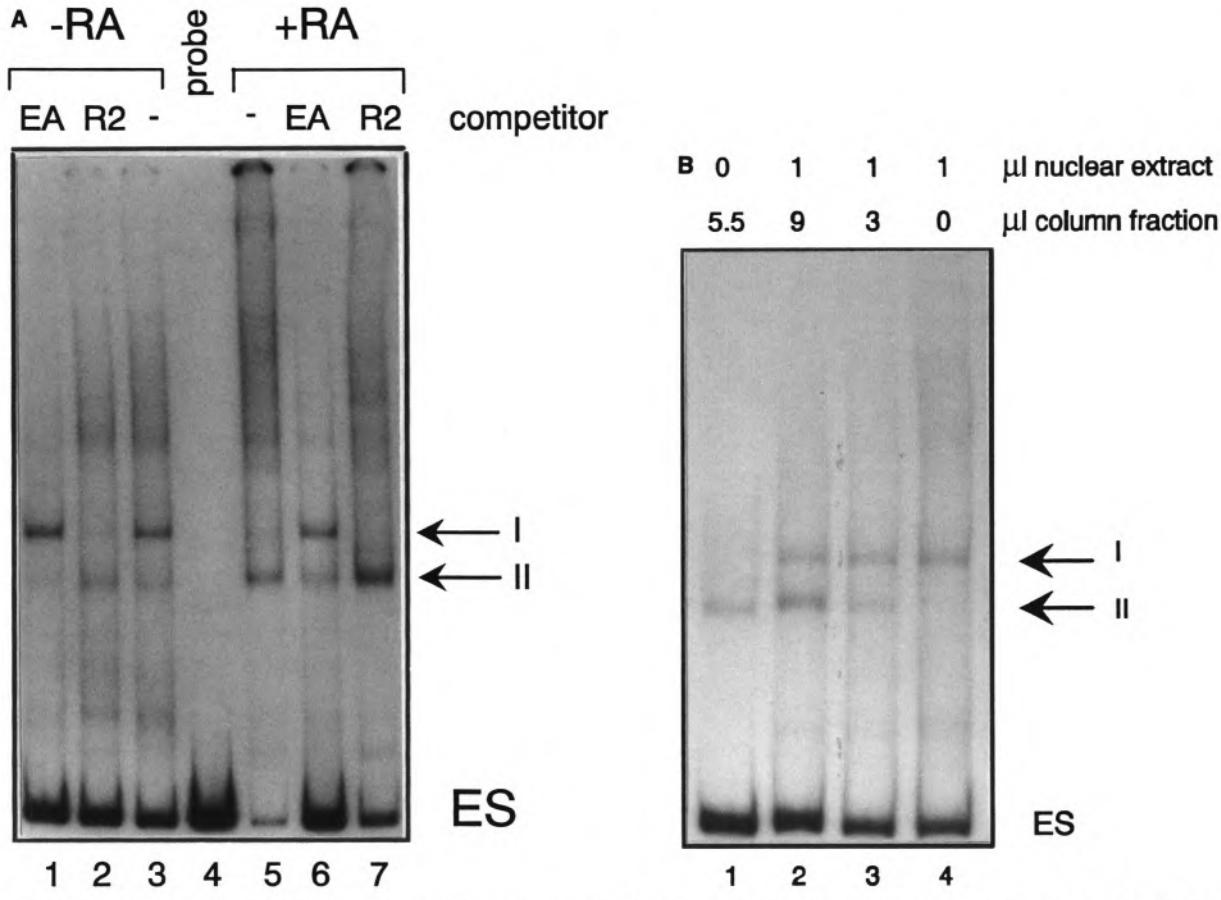



FIG. 7. Forms I and II of the ES retarded complexes are comprised of factors binding R2 and R3, respectively. (A) The labeled 304 bp ES fragment was incubated with 20 $\mu$g of nuclear extract from P19 cells (−RA, lanes 1–3) or P19 cells induced to differentiate by treatment with RA for 7 days (+RA, lanes 5–7). Lane 4 contains no extract. Lanes 1 and 6 contain a 200 molar excess of unlabeled EA DNA whereas lanes 2 and 7 contain a 200 molar excess of the R2 DNA. Arrows indicate the two retarded complexes. (B) The labeled ES fragment was incubated with 1 $\mu$l (approximately 20 $\mu$g) P19 nuclear extract and/or partially purified factor from differentiated P19 cells that binds the R3 region at the indicated concentrations. Arrows indicate the two retarded complexes.

moter carrying the mutant UAS was as active as the wild-type promoter when transfected into either undifferentiated (Fig. 10A) or differentiated (Fig. 10C) P19 cells.

When the two constructs were stably transfected into P19 cells, the mutant promoter had only 20% the activity of the wild type (Fig. 10B). Thus, it appears that the R2 region plays an important role in promoter activity when the promoter is integrated into the genome and bound as chromatin.

When the stably transfected cells were treated with RA, the difference in expression between the mutant and wild-type promoters was not significantly different (Fig. 10D). This result is consistent with the EMSA experiments showing that in differentiated cells the ES probe bound to proteins primarily in its R3 region.

## DISCUSSION

The 304 bp UAS associated with the *Pgk-1* promoter was originally identified as a promoter-specific enhancer by transient transfection experiments in P19 embryonal carcinoma cells (McBurney et al., 1991). The *Pgk-1* promoter lacking the UAS was essentially inactive following stable integration into P19 cells whereas the *Pgk-1* promoter including the UAS drove efficient expression of integrated reporter genes. Three regions of the UAS bind nuclear proteins. The footprinted R2 site appears to be critical for activity of the integrated *Pgk-1* promoter because a mutation in this region severely compromised expression from the integrated promoter but had no effect on the expression of the transiently transfected promoter constructs.

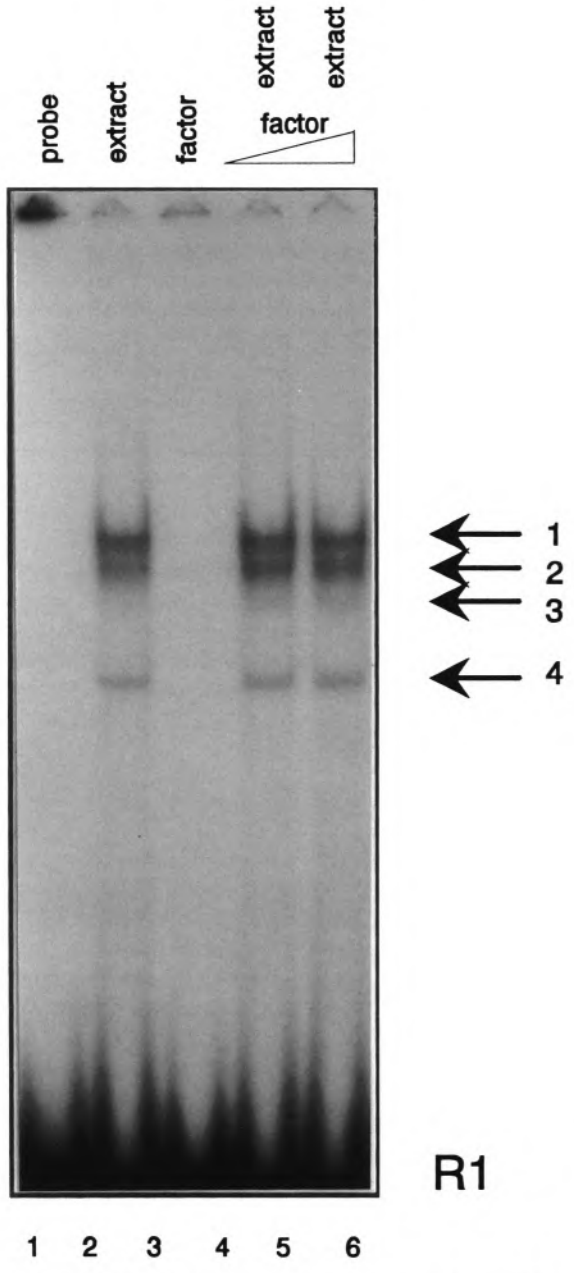
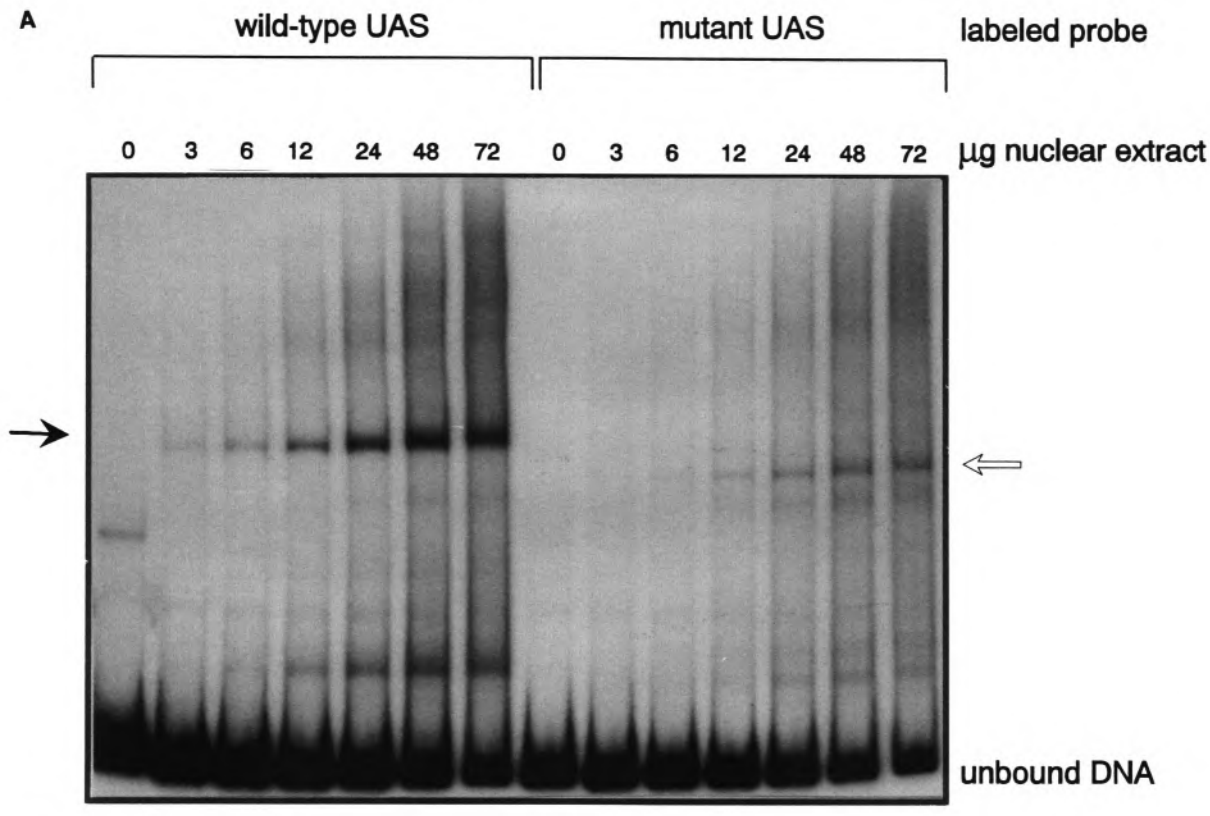



FIG. 8. The R3 binding factor does not interact with the factors binding R1. R1 DNA was labeled and mixed with P19 cell nuclear extract (lane 2) or partially purified factor that binds the R3 region (lane 3). In lanes 5 and 6, 20 $\mu$g of nuclear extract was mixed with 2 and 4 $\mu$l of partially purified extract from differentiated P19 cells. Arrows at the right indicate the four retarded complexes formed on the R1 probe.

Nuclear extracts from undifferentiated P19 cells bound to only two sites, R1 and R2. Binding to R2 appears to be higher affinity because this site was footprinted with DNase1 whereas no footprint was seen on R1. The R1 and R2 sites both contain an E box and ets binding sites [GGA(A)] and appear to bind to a common protein(s). Competition between R1 and R2 for this factor indicated that R2 bound with higher affinity.

Both R1 and R2 sites seem to bind more than one protein. EMSA using the R1 probe revealed a number of bands of different mobility and differing sensitivity to competition with R2. Fractionation of nuclear extracts on heparin agarose columns yielded only one fraction with DNA binding activity, and this fraction produced a single retarded band with the R1 probe (data not shown). This result suggested that the complex banding pattern seen on R1 in EMSA arises from cooperative binding of a number of different proteins. The footprint obtained on the wild-type R2 region was extensive and became more restricted when the central 8 bp of this region was mutated. This result suggests that a number of proteins bind the R2 region, although the EMSA with the R2 probe revealed only a single retarded band. The factor binding the R3 region appears to be a single protein or a preformed complex of proteins because this activity was partially purified by gel filtration column chromatography.

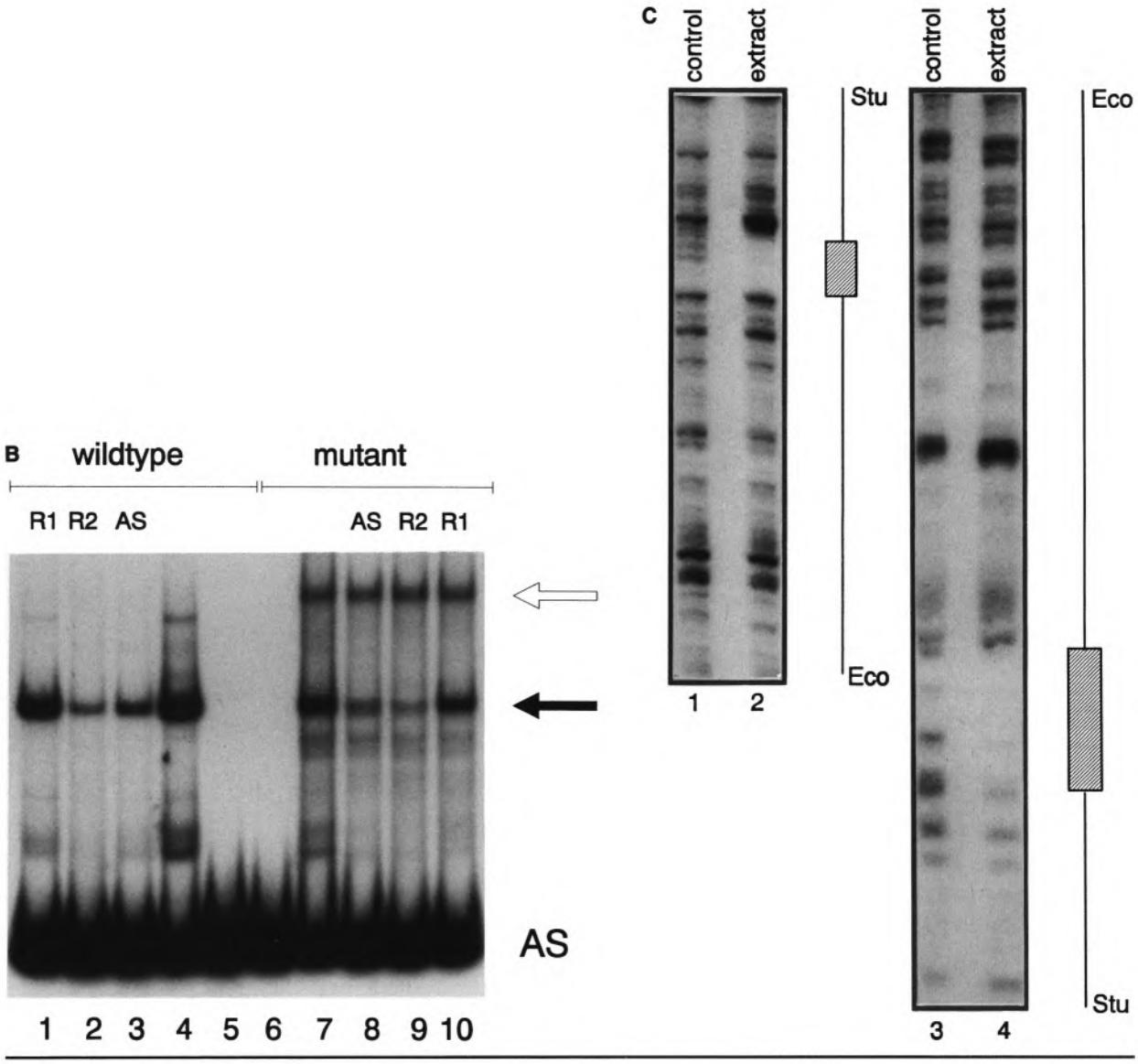
Our previous expression studies indicated that neither the EA nor the AS region of the UAS had significant enhancer activity in P19 cells (McBurney et al., 1991). We therefore used the intact UAS as a probe in EMSA to determine whether we could detect cooperation between the proximal (AS) and distal (EA) regions of the ES probe in binding to nuclear proteins. In fact, cell extracts from P19 cells contain very little protein that bound DNA from the distal (EA) region. In addition, there appeared to be mutually exclusive binding of proteins to these two regions of the UAS. The more slowly migrating band, form I, could be effectively competed by the R2 whereas the more rapidly migrating band, form II, was competed by R3. The composition of factors in form I is not clear. Because R2 can compete for factors binding R1, it is possible that a number of proteins are present in this complex. Form II seems to be relatively simple and consists of the partially purified factor bound to the R3 region. The binding of factors to the R3 region appears to be mutually exclusive, with binding of other factors to the regions containing the R1 and R2 sites, perhaps suggesting that the binding of the activating factors to the R1 and R2 regions might be stabilized by the DNA from the more distal R3 region. Alternatively, the presence of the R3 binding factor might destabilize binding of factors to the proximal region.




ABOVE AND FACING PAGE
FIG. 9. Mutation of the R2 region alters protein binding to the UAS. (A) Various amounts of nuclear extract from P19 cells were mixed with labeled ES DNA derived from the normal promoter or the promoter in which the central 8 bp of the R2 region had been mutated (see Fig. 3B). The arrow on the left indicates form I complex formed on the normal ES probe whereas the arrow on the right indicates the complex formed on the mutant ES probe. The mobility of the latter complex was slightly higher. Both mutant and wild-type ES probes yielded a weak form II band seen running in front of the form I complexes. (B) AS probes carrying the wild-type (lanes 1–5) or mutant (lanes 6–10) R2 regions were labeled and mixed with nuclear extract from P19 cells in the presence of cold competitor DNA. Competitors were R1 DNA (lanes 1 and 10), R2 DNA (lanes 2 and 9), and AS DNA (lanes 3 and 8), all of wild-type sequence. The R1 and R2 were at 30-fold molar excess whereas AS was at 10-fold excess compared to the probe. Lanes 5 and 6 are the labeled probes with no nuclear extract and lanes 4 and 7 contain nuclear extract but no cold competitor. The black arrow indicates the major retarded complex seen also in Fig. 5, lane 6. The open arrow indicates a novel retarded complex seen only with the mutant probe. (C) A DNaseI footprint was formed on the UAS containing the mutant R2 region. P19 nuclear extract (40 $\mu$g) was incubated with the end-labeled 304 bp ES fragment containing the mutated R2 sequence. Lanes 1 and 3 are control lanes with no added extract. Lanes 2 and 4 contain nuclear extract. The sequence of the footprinted region is shown shaded in Fig. 3B.

The activity of the UAS is best explained if it plays a role in chromatin configuration because it seems particularly important for promoter activity following integration into the genome. Mutation of the R2 region also reduced promoter activity but only in stably transfected cells. It is known that the accessibility of transcription factors for promoter sequences in vivo can vary depending on whether the target DNA is naked or chromatin bound (Archer et al., 1992). Although the R2 site contains an E box and P19 cells do express some of the bHLH transcription factors known to bind such sites, we have no evidence that it is bHLH factors that bind the R2 region. In fact, R1 and R2 contain DNA sequences that closely resemble the consensus binding site for the YY1 protein (Shi et al., 1991), a factor that can both activate and repress transcription. The YY1 factor is known to induce DNA bending and this may be its means of activating and repressing the c-fos promoter (Natesan and Gilman, 1993). In this regard it is interesting that the R2 mutation that had no effect on transient expression deleted the YY1 binding site but replaced it with an AT-rich sequence that may itself be bent in the absence of protein (Natesan and Gilman, 1993).

Proteins bound to the human *Pgk-1* promoter have been mapped by in vivo footprinting (Pfeifer



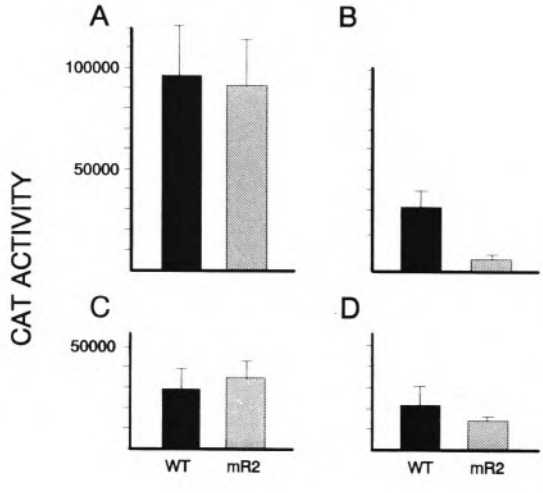




FIG. 10. The mutation altering the R2 region reduces expression from the *Pgk-1* promoter only following its integration into the genome. Reporter genes consisted of the *Pgk-1* promoter driving the CAT gene. Either the wild-type *Pgk-1* promoter (5 $\mu$g) (black bars) or the promoter carrying the R2 mutation (5 $\mu$g) (slashed bars) driving CAT were cotransfected into P19 cells along with 2 $\mu$g of RSV-*lacZ* to control for transfection efficiency and 2 $\mu$g of *Pgk-puro*, the selectible gene. (A) Transient assays into undifferentiated P19 cells. (B) Stable transformants were isolated by selection in puromycin. Populations of cells were pooled and assayed for activities. (C) Transient assays in P19 cells that had been treated with RA for 7 days prior to transfection. (D) Stable transformants isolated as in (C) except the pooled populations of cells were treated with RA for 7 days prior to measurement of activities. The results are averaged from two experiments and the bars represent the range of the data.

et al., 1990; Pfeifer and Riggs, 1991). A region corresponding to R1 appears to be footprinted in vivo. No footprint appeared in the region corresponding to R2, although the DNA sequence in this region has been conserved between mouse and man. Two in vivo footprints were obtained in the region corresponding to R3, although the DNA sequences for these footprinted regions were not well conserved (McBurney et al., 1991). Yang et al. (1988) found a number of proteins that bound the human *Pgk-1* promoter region using in vitro methods such as those described above. One of the DNaseI footprints was formed in a region corresponding to the R1 site.

At present, we do not know the nature of any of the factors detected by the UAS probe. Work of Goto et al. (1991) has established that a DNA binding factor present in testes extracts inhibits in vitro transcription initiated from the *Pgk-1* promoter. This factor binds an ets site in the R3 region. The ets-1 transcript is induced during P19 cell differentiation (Kola et al., 1993). Another transcription factor of the ets family, PEA3 (Xin et al., 1992), inhibits transcription from the *Pgk-1* promoter in cotransfection assays (A. Cowie, L. Sutherland, J. A. Hassell, and M. W. McBurney, unpublished). bFGF also inhibits *Pgk-1*-driven transcription in vitro (Nakanishi et al., 1992), although the relevant binding site was not mapped.

After this work was completed, a report appeared indicating that the *Pgk-1* promoter contains a DNA motif that confers elevated expression in cells exposed to hypoxia (Firth et al., 1994). The hypoxia response element overlaps the footprinted R2 site shown in Fig. 3. The UAS functions as a cell type specific enhancer by mediating much larger increases in expression in embryonal carcinoma cells than in differentiated cells. This cell type specificity may be based on the fact that embryonal carcinoma cells constitutively express transcription factors induced in other cells by stress. For example, embryonal carcinoma cells grown at 37°C contain high levels of heat shock element binding protein (Mezger et al., 1989). Perhaps hypoxia-inducible transcription factors are constitutively expressed in embryonal carcinoma cells and are responsible for the activity of the UAS in this cell type.

## ACKNOWLEDGEMENTS

This work was supported by grants from the Medical Research Council of Canada. M.W.McB. is a Terry Fox Cancer Research Scientist supported through the National Cancer Institute of Canada.

## REFERENCES

C. N. Adra, P. H. Boer, and M. W. McBurney (1987), Gene 60, 65–74.

T. K. Archer, P. LeFebvre, R. G. Wolford, and G. L. Hager (1992), Science 255, 1573–1576.

C. Chen and H. Okayama (1987), Mol Cell Biol 7, 2745–2752.

J. D. Dignam, P. L. Martin, B. S. Shastry, and R. G. Roeder (1983), Eukaryotic gene transcription with purified components, in Methods in Enzymology. Academic Press, Inc., New York, pp. 582–598.

J. D. Firth, B. L. Ebert, C. W. Pugh, and P. J. Ratcliffe (1994), Proc Natl Acad Sci USA 91, 6496–6500.

M. Fried and D. M. Crothers (1981), Nucleic Acids Res 9, 6505–6523.

M. M. Garner and A. Revzin (1981), Nucleic Acids Res 9, 3047–3059.

M. Goto, T. Tamura, K. Mikoshiba, Y. Masamune, and Y. Nakanishi (1991), Nucleic Acids Res 19, 3959–3963.

K. A. Jones, K. R. Yamamoto, and R. Tjian (1985), Cell 42, 559–572.

E. M. V. Jones-Villeneuve, M. McBurney, K. A. Rogers, and V. I. J. Kalnins (1982), J Cell Biol 94, 253–262.

E. M. V. Jones-Villeneuve, M. A. Rudnicki, J. F. Harris, and M. W. McBurney (1983), Mol Cell Biol 3, 2271–2279.

D. H. Keith,, J. Singer-Sam, and A. D. Riggs (1986), Mol Cell Biol 6, 4122–4125.

I. Kola, S. Brookes, A. R. Green, R. Garber, M. Tymms, T. S. Papas, and A. Seth (1993), Proc Natl Acad Sci USA 90, 7588–7592.

M. W. McBurney, S. Fournier, K. Jardine, and L. Sutherland (1995a), Somat Cell Mol Genet 20, 515–528.

M. W. McBurney, S. Fournier, P. K. Schmidt-Kastner, K. Jardine, and J. Craig (1995b), Somat Cell Mol Genet 20, 529–540.

M. W. McBurney, W. A. Staines, K. Boekelheide, D. Parry, K. Jardine, and L. Pickavance (1994), Dev Dyn 200, 278–293.

M. W. McBurney, L. C. Sutherland, C. N. Adra, B. Leclair, M. A. Rudnicki, and K. Jardine (1991), Nucleic Acids Res 19, 5755–5761.

V. Mezger, O. Bensaude, and M. Morange (1989), Mol Cell Biol 9, 3888–3896.

C. Murre, P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub, and D. Baltimore (1989), Cell 58, 537–544.

Y. Nakanishi, K. Kihara, K. Mizuno, Y. Masamune, Y. Yoshitake, and K. Nishikawa (1992), Proc Natl Acad Sci USA 89, 5216–5220.

S. Natesan and M. Z. Gilman (1993), Genes Dev 7, 2497–2509.

P. A. Norton and J. M. Coffin (1985), Mol Cell Biol 5, 281–290.

G. P. Pfeifer, R. Drouin, A. D. Riggs, and G. P. Holmquist (1991), Proc Natl Acad Sci USA 88, 1374–1378.

G. P. Pfeifer and A. D. Riggs (1991), Genes Dev 5, 1102–1113.

G. P. Pfeifer, R. L. Tanguay, S. D. Steigerwald, and A. D. Riggs (1990), Genes Dev 4, 1277–1287.

M. A. Rudnicki and M. W. McBurney (1987), Cell culture methods and induction of differentiation of embryonal carcinoma cell lines, in Teratocarcinomas and Embryonic Stem Cells, a Practical Approach (E. J. Robertson, ed.), IRL Press, London, pp. 19–49.

Y. Shi, E. Seto, L. S. Chang, and T. Shenk (1991), Cell 67, 377–388.

J. Singer-Sam, M. Grant, J. M. LeBon, K. Okuyama, V. Chapman, M. Monk, and A. D. Riggs (1990), Mol Cell Biol 10, 4987–4989.

H. Singh, R. Sen, D. Baltimore, and P. A. Sharp (1986), Nature 319, 154–158.

R. St-Arnaud and J. M. Moir (1993), Mol Cell Biol 13, 1590–1598.

S. Tommasi, J. M. LeBon, A. D. Riggs, and J. Singer-Sam (1993), Somat Cell Mol Genet 19, 529–541.

J.-H. Xin, A. Cowie, P. Lachance, and J. A. Hassell (1992), Genes Dev 6, 481–496.

T. P. Yang, J. Singer-Sam, J. C. Flores, and A. D. Riggs (1988), Somat Cell Mol Gene 14, 461–472.