Nucleic Acids Research, Vol. 19, No. 24 6839–6847

# Constitutive expression of the urokinase plasminogen activator gene in murine RAW264 macrophages involves distal and 5′ non-coding sequences that are conserved between mouse and pig

A.Ian Cassady, Katryn J.Stacey, Karen A.Nimmo, Kathleen M.Murphy, Dietmar von der Ahe[1], David Pearson[1], Florence M.Botteri[1], Yoshikuni Nagamine[1] and David A.Hume*
Centre for Molecular Biology and Biotechnology, University of Queensland, Brisbane 4072, Australia and [1]Friederich-Miescher Institute, PO Box 2543, CH-4002 Basel, Switzerland

Received September 3, 1991; Revised and Accepted October 14, 1991 EMBL accession no. X63356

## ABSTRACT

**The 5′ flanking regions of the mouse and pig urokinase plasminogen activator (uPA) genes were sequenced and sequence homology interrupted by repeat elements was found to extend to −4.6kb in pig and −6.6kb in mouse. A transient transfection procedure was devised for the murine macrophage cell line RAW264. Pig uPA promoter-CAT constructs were more active than mouse constructs in this assay. This contrast may involve sequence differences within 100 bp of the transcription start site. The selective deletion of distal regions of the promoter (>2.6 kb upstream), and of a conserved element, 5′-AGGAGGAAATGAGG-TCA-3′ around −2 kb greatly reduced the activity of reporter constructs in RAW264 cells. Electrophoretic mobility shift assays using the latter sequence identified a single nuclear protein complex. This element has been referred to as PEA3/AP1-like, but the complex did not comigrate with either AP1 or known proteins that bind polypurines (including the macrophage-specific factor PU-1) and was not competed by AP1 or polypurine oligonucleotides. uPA promoters contain multiple AP1 and AP2-like DNA sequences, which were recognised by nuclear proteins expressed constitutively in RAW264 cells. They also contain multiple binding sites for NFκB but activated NFκB was not expressed in RAW264 cells. The conserved, transcribed 5′ non-coding sequences were also required for maximal gene expression. Hence, the uPA promoter contains multiple weak cis-acting elements distributed over 7.0 kb 5′ to the translation start site.**

## INTRODUCTION

Urokinase plasminogen activator (uPA) has been implicated in extracellular proteolysis in numerous cellular systems especially metastatic tumor cells and macrophages (reviewed in 1). The genes encoding uPA, and some 5′ flanking regions, have been sequenced in pig (2, 3), human (4, 5), mouse (6) and chicken (7). Sequence comparisons between mouse, human and pig genes (6) revealed the unusual property that the introns were 50−70% conserved between the three species, a level of conservation only marginally less than that between the exons. In the published sequences, there is an abrupt decline in homology in the immediate 5′ flanking region. Given the diversity of stimuli that alter the expression of uPA in various cell types, including growth factors, oncogenes, cyclic AMP agonists glucocorticoids, retinoids, phorbol esters and cycloheximide (8−14) it is rather surprising to observe so little sequence identity in the region often involved in transcriptional regulation. There are several possible explanations which could account for this observation: (i) the gene is regulated differently in different species, (ii) the control segments lie within the introns or at the 3′ end of the gene, (iii) conserved control elements lie outside the previously published 5′ sequence. In support of the third possibility, studies in pig LLC-$PK_1$ cells have suggested that essential sequences involved in transcriptional activation by calcitonin and cyclic AMP agonists lie more than 3.5 kb upstream (14). In the human uPA 5′ flanking sequence, an enhancer element that activates transcription of an enhancerless SV-40 promoter was identified between −2350 and −1824 (4). More recently, Rørth *et al.* (9) reported the existence of similar elements around −2.4 kb that conferred sensitivity to epidermal growth factor and phorbol ester in mouse keratinocytes. We have sequenced the pig uPA 5′ flanking sequence to −4.6 kb, and the mouse uPA flanking sequence to −6.6 kb and have identified extensive sequence alignment outside the previously known sequence.

The level of uPA mRNA and/or enzymic activity is increased in macrophages by many stimuli, including the major growth and differentiation factor, macrophage colony-stimulating factor (CSF-1), phorbol esters and interferon gamma (15, 11; KS, DAH; unpublished) and is reduced by others including glucocorticoids and prostaglandins (11). Studies on transcriptional regulation of gene expression in macrophages have been

* To whom correspondence should be addressed

constrained by the difficulties in introducing DNA into the cells for transient expression analysis, and few studies on the activity of any promoter in macrophages have appeared. We have developed an electroporation technique for transient transfection of the murine macrophage cell line, RAW264 (16), which expresses uPA mRNA (KS, DAH; unpublished). Transient expression assays, using mouse and pig promoters fused to the chloramphenicol acetyl transferase (CAT) gene, have been used to compare the activity of the mouse and pig uPA promoters and to identify regions that are involved in constitutive expression of the uPA gene in these cells.

## MATERIALS AND METHODS

### Isolation and sequencing of genomic clones

Construction of a pig genomic DNA library in λEMBL4 bacteriophage and isolation of a uPA genomic clone, λYN4, have been reported previously (3). The clone contains 7 kb of 5′ flanking region of the uPA gene. The sequence of 900 bp of 5′ flanking region, 6 kb of transcribed region and 300 bp of 3′ flanking region has been reported (3). In this work we extended the sequencing up to the nearest *Eco*RI site from the transcription initiation site (4.66 kb from the initiation site) using the chemical digestion method of Maxam and Gilbert (17). DNA fragments created using appropriate restriction sites were labelled either at the 5′ end with polynucleotide kinase and [$\gamma$-$^{32}$P]ATP or at the 3′ end with terminal transferase and [$\alpha$-$^{32}$P]-ddATP. In most cases the sequence was read on both strands.

Mouse genomic clones were obtained by screening a mouse genomic library, using the longest *Pst* I fragment from a mouse uPA cDNA clone pDB15 (3). To ensure that the clone contained the 5′ flanking region, these clones were further screened with an *Xba* I-*Xba* I fragment of the pig uPA gene (−2695 to −1770). As noted below, this probe subsequently proved to have regions of 90% homology with the mouse gene. The genomic library was provided by Shigeki Shibahara, and was constructed in λEMBL4 using a *Sau*3A I partial digest of BALB/c mouse liver DNA. The mouse sequence extending from the *Bgl* II site flanking the known sequence at −2187 to the *Xba* I site at −6616 was obtained by subcloning the DNA in both orientations into pBluescript (Stratagene), preparing a series of defined restriction enzyme-mediated deletions and performing dideoxy-sequencing on double-stranded plasmid templates using a Pharmacia T7 polymerase sequencing kit. Some gaps were filled by preparing defined oligonucleotide primers based on previously obtained sequence.

### Nuclear extracts

Nuclear extracts were prepared with a variation of the method of Osborn *et al.* (18). Cells were washed with phosphate buffered saline, centrifuged and resuspended in Dignam buffer A (19) containing 0.1% Nonidet P40 (100 ml/$10^7$ cells) at 4°C. Nuclei were gently pelleted (250×*g*, 10 min), and resuspended in Buffer C. After 10mins on ice, the extracts were centrifuged at 10,000×*g* for 10min and the supernatant was removed and diluted 5-fold in buffer D. The protein content of the extracts was determined using a modified Bradford assay (BioRad), and the extracts were stored frozen at −70°C.

### Electrophoretic mobility shift assay (EMSA)

Oligonucleotide probes were labelled with $\alpha$-[$^{32}$P]-dATP using DNA polymerase 1 (Klenow fragment) and a 10 bp primer complementary to the 3′ end of the full length oligonucleotide. Binding reactions contained 2 μg of nuclear extract protein, 1 μg of poly d(IC), 0.5 ng of probe DNA, 10 mM EDTA and 5% glycerol in a final volume of 6 μl of buffer D. After 30 min at room temperature, samples were loaded onto polyacrylamide gels (BioRad MiniProtean II) of the desired concentration, made up in 50 mM Tris-380 mM glycine-2 mM EDTA pH 8.5. The gels were run at constant voltage in the same buffer. Bromophenol blue was added to one well (usually the probe control) and electrophoresis stopped when the dye reached the end of the gel. Gels were washed in 5% acetic acid, dried and autoradiographed.

The oligonucleotide probe sequences used were as follows: PU1: 5′-CTGAAAGAGGAACTTGGTTAGGTA-3′ (20) mB: 5′-GGCAAGGCTATTTGGGAAGGGAAAATA-3′ (21) AP1: 5′-TTCCGGCTGACTCATCAAGCG-3′ (22); supplied by Promega κB: 5′-AGGGACTTTCCGCTGGGGACTTTCCA-3′ (18) PEA3/AP1: 5′-TCTCTGGCCAGGAGGAAATGAGGTC-ATCTTGCTCT-3′ (9) AP2-like: 5′-GAATCAGCATGACAG-CCTCCAGCCAAGTAA-3′ (−2410, mouse uPA, Fig. 2). Apart from the commercial AP1 oligonucleotide, all other probes were prepared on a Milligen 7510 DNA Synthesiser within the Centre for Molecular Biology and Biotechnology, University of Queensland.

### Preparation and transfection of plasmid constructs

In the plasmid pUPACAT the 5′ flanking sequences of the mouse uPA gene from the *Xba* I site at −6616 to the *Sac* I site at +8 was placed upstream of the chloramphenicol acetyl transferase (CAT) reporter gene in pGeorge, a vector constructed from pBR322, the CAT gene, and pUC19 multiple cloning site. The puPACAT2 construct was made by inserting mouse uPA promoter sequences from −6616 to the *Xba* I site at +398 upstream of the CAT gene in pGeorge. pCAT4660 is a plasmid containing 4660 bp of pig uPA 5′ flanking region linked to the CAT gene (23). pΔPEA-3/AP1 CAT4660 was prepared using the polymerase chain reaction, and is identical to pCAT4660 except that the sequence from −2007 to −1989, including the PEA3 and AP1-like sequences, is replaced by AAT, creating a convenient restriction site. The nucleotide G at −3530 was changed to C in order to destroy the *Xho* 1 site.

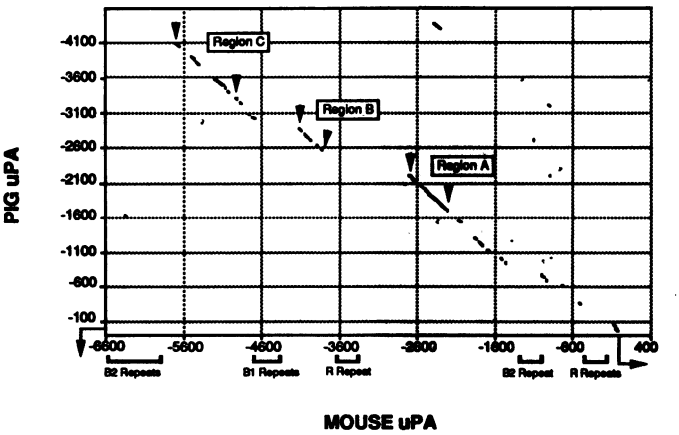


**Figure 1.** Pustell matrix plot. DNA sequences from the mouse and pig uPA 5′-flanking regions were compared using the Pustell matrix algorithm of the MacVector 3.04 application on a Macintosh IIcx (25). The areas of strong sequence similarity are indicated by regions A, B and C. The transcription start sites for both genes are marked with arrows. Areas of repeat sequence are marked; B1, B2 and R (polypurine).

RAW264 tumour cells were grown in RPMI-1640 medium plus 5% foetal calf serum. The cells were harvested in logarithmic growth phase, washed, and resuspended at *ca.* $2\times10^7$/ml in RPMI-1640 with 10% foetal calf serum and 20 mM Hepes, pH 7.1. Aliquots of 250 $\mu$l of cells were electroporated at room temperature with 10−25 $\mu$g of CsCl gradient purified plasmid DNA in an 0.4 cm electroporation cuvette at 300 V, 960 $\mu$F using a Bio-Rad Gene Pulser with a capacitance extender. The concentration of DNA used in studies of deletion mutants was adjusted so that each plasmid was present at the same molar concentration. The cells were then returned to culture and after 2−3 days were harvested and assayed for CAT activity essentially as described by Gorman *et al.* (24). Cell extract (10 to 200 $\mu$g of protein, depending on the expected activity) was incubated with 0.25 $\mu$Ci of D-threo-[dichloroacetyl-1-$^{14}$C]chloramphenicol (Amersham) at 57 mCi/mmol and 2.2 mM acetyl-CoA in a volume of 180 $\mu$l for 1 to 4 hours. The conversion of labelled substrate to acetylated chloramphenicol was quantitated using an AMBIS Radioanalyser. Conversion of chloramphenicol was kept below 20% to minimise errors due to substrate depletion. Conversions were corrected by the subtraction of results for chloramphenicol incubated with untransfected control protein extract. We chose not to cotransfect a control plasmid with the reporter constructs for three reasons; (i) the variation in CAT expression between replicates transfected with the same plasmid was small, and no reference activity was needed to control for such variation; (ii) the use of such controls assumes there is no 'cross-talk' between transfected promoters whereas we find that the activity of one reporter construct can vary inversely with that of a co-transfected construct because of competition for a limiting factor; (iii) because we are dealing with a weak promoter, especially in the truncated mouse constructs, we have to use relatively high amounts of plasmid for transfection (50−125ug/ml). At this concentration, electroporation causes DNA-dependent cellular cytotoxicity, and addition of more DNA of any form (such as a second reporter construct) reduces both CAT expression and cell viability.

## Cell Culture

All cell culture reagents were obtained from either Cytosystems (Sydney) or Commonwealth Serum Laboratories (Melbourne). The cell line RAW264 was obtained from the American Type

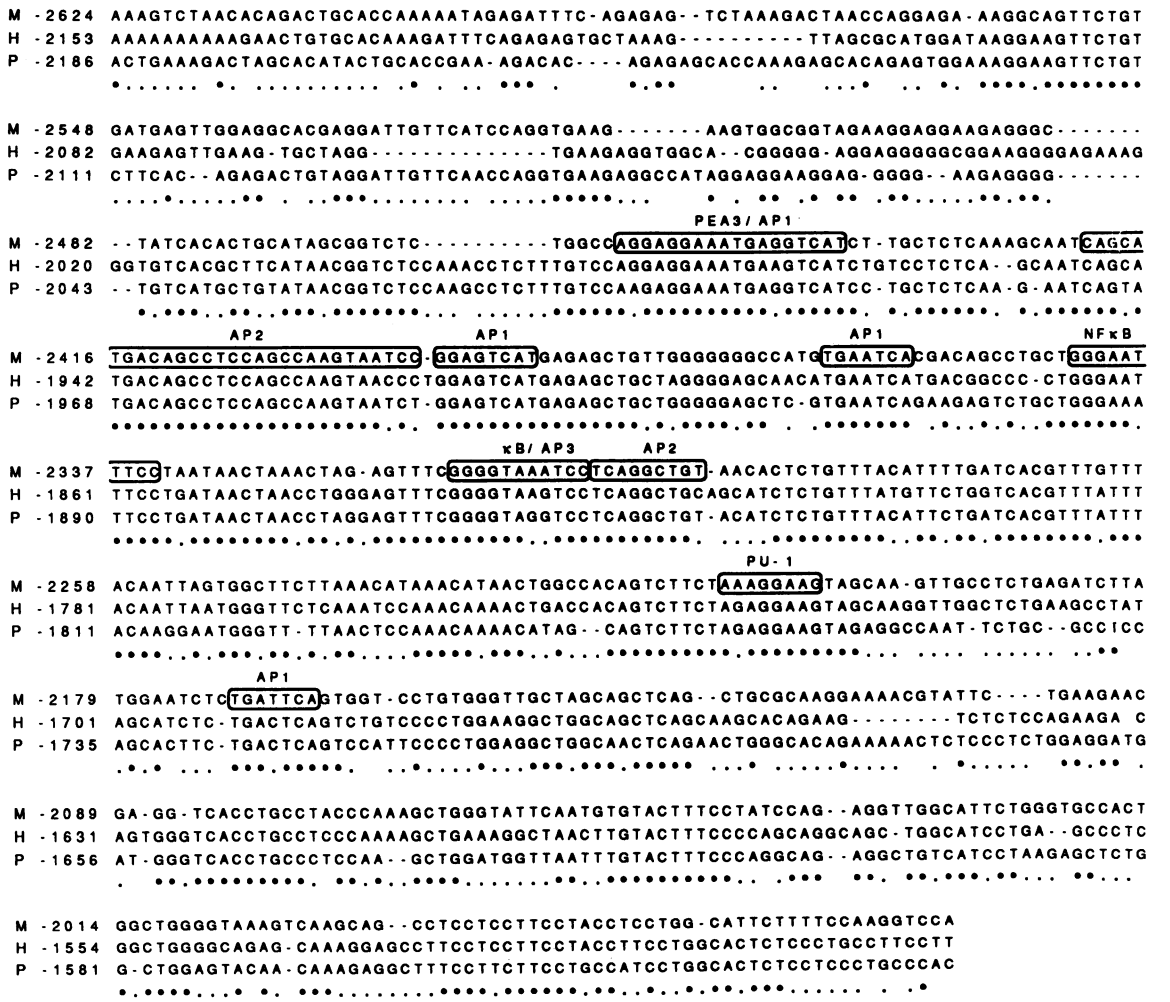


**Figure 2.** Region A sequence alignment between mouse, human and pig uPA promoters. Sequences for the mouse and pig uPA promoters were aligned with the human uPA sequence (4). Sequence identity between any two of the three sequences is marked by (·) and sequence identity between all of the three sequences is marked by (•). Consensus transcription factor binding sites are marked, and discussed in the text.

Culture Collection and was tested free of mycoplasma. Recombinant bacterial human macrophage colony-stimulating factor (CSF-1) was a generous gift from Cetus Corporation (Emeryville, Ca.). Phorbol myristate acetate and dexamethasone were obtained from Sigma (St. Louis, Mo.). Concentrated stocks in dimethyl sulphoxide were stored frozen in aliquots and diluted in complete medium immediately prior to use.

## RESULTS

### Sequence alignments

We have sequenced the mouse urokinase promoter to −6.6 kb and the pig promoter to −4.6 kb. The former extends the mouse genomic sequence 3.5 kb upstream from the sequence described by Degen *et al.* (6) and Rørth *et al.* (9). The two sequences have been deposited in EMBL Data Library (Accession Number: ) and may be obtained from us. A single variation was found between our mouse uPA sequence and the sequence of Degen *et al.* (6), an A not G at −2137, however variations were found at five positions with the sequence of Rørth *et al.* (9); our sequence contained an additional A inserted at −3176, an additional C inserted at −2989, no G present at −2956, a G not a C at −2194 and a C not a G at −2192.

Figure 1 shows a Pustell DNA matrix comparison (25) of the mouse and pig uPA promoter sequences. There is a 660 bp region of sequence conservation between the two species from −2626 to −1970 in the mouse. This region also resembles part of the human uPA flanking sequence and an alignment of the sequences from the three species is shown in Figure 2. We will refer to this segment of the promoter as Region A.

Overall, 55% of the base pairs in Region A are conserved across all three species, and >80% are conserved between any two species. In the central block of 250 base pairs, the homology is even greater (74% in all three species, >95% in any two compared). The total area of conservation covers both the enhancer and suppressor regions described previously in studies of transient expression in human tumour cells and mouse keratinocytes (4, 9).

On the boundary of Region A there is an Alu-like repeat sequence in the human promoter which extends to the boundary of the published sequence (4). In the same relative position, there is shift in the alignment between the mouse and pig sequences. Approximately 1.5kb in the mouse, and 400bp in the pig, upstream from this point there is a second block of sequence similarity between the two species (Region B) extending for around 400 bp (Figure 3). Region B is followed by a third block of homology (Region C; Figure 4), which is almost continuous in the pig promoter, but separated by a 400bp segment containing B1 repeats in the mouse (6). These repeats were assigned by comparison with the consensus B1 sequence described by Kalb *et al.* (26) and the consensus B2 repeat sequence described by Krayev *et al.* (27). Regions A, B and C in the mouse sequence were scanned for known transcription factor binding sites using IBI MacVector Version 3.5 software, and a library of sites supplied with the package. Relevant sites are annotated on Figs 2,3 and 4, and discussed elsewhere.

### Transient expression of reporter constructs in RAW264 cells

A series of deletion constructs of the mouse and pig uPA promoters linked to the chloramphenicol acetyl-transferase (CAT) gene were prepared as shown in Figure 6a. These deletions are designed to successively remove Regions A, B and C. The

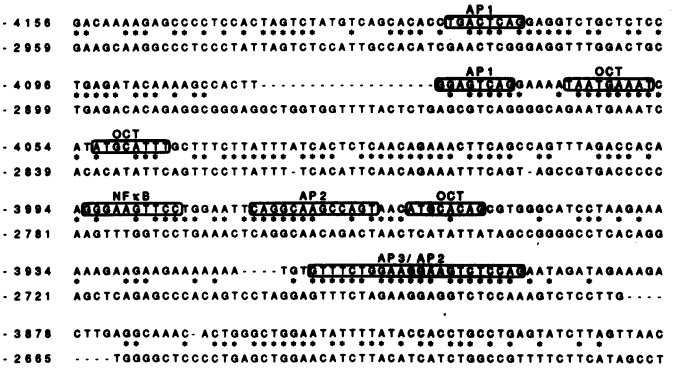


**Figure 3.** Region B sequence alignment between mouse (upper) and pig (lower) uPA promoters, annotated as in Fig 2. Consensus transcription factor binding sites are marked, and discussed in the text.

reporter constructs were transfected into the mouse macrophage cell line RAW264 using electroporation. Primary macrophages and macrophage cell-lines are difficult to transfect with conventional chemical methods (28), but a modified electroporation procedure gave highly reproducible results with RAW264 cells. In a series of experiments using a human β-actin promoter-CAT expression plasmid (29; a gift from Dr. P. Gunning) as a test system, optimum CAT expression was dependent on several factors (manuscript in preparation). The use of complete medium (*i.e.* including serum) was more effective than simpler media such as HEPES-buffered saline. CAT expression was plasmid concentration-dependent over a broad range. There was an electroporation voltage optimum around 750 V/cm at 960 μF and a temperature optimum of 25°C.

The RAW264 cell line is a product of Abelson leukaemia virus transformation of mouse spleen (16) and constitutively synthesizes uPA mRNA (KS, data not shown). Results in Figure. 5a show that removal of Region C from the mouse promoter caused a 3−5 fold reduction of CAT expression in transiently-transfected RAW264 cells. Further deletion of Regions B and A caused successive reductions in promoter activity of 2−3 fold.

The pig promoter constructs were up to 50-fold more active than the mouse constructs (note the different vertical scale in Figure 5b). In each of three experiments, each successive deletion reduced promoter activity. Because the pig deletions are actually in the middle of Regions B and C they cannot be directly compared with the mouse constructs. The data do, however, support the conclusion that Regions B and C play a role in uPA promoter activity in RAW264 cells. The location of the conserved regions and the positions of the deletion constructs are summarized in Figure 5c.

The mouse promoter constructs used in Figure 5a differ from those used by Rørth *et al.* (9), and from the pig promoter constructs, in lacking 5′ untranslated sequences. The CAT gene was fused to the *Sac* I site at +8 relative to the transcription start site, whereas Rørth *et al.* (9) used the *Xba* I site at +398, immediately 5′ of the start of translation. The same conserved *Xba* I site was used in the generation of the pig CAT construct. Thus, the difference in the level of expression between mouse and pig constructs in Figure 5 could be attributed to elements present between the transcription and translation start sites, between +8 and +398 in the mouse sequence. To examine this possibility a mouse reporter construct containing the additional

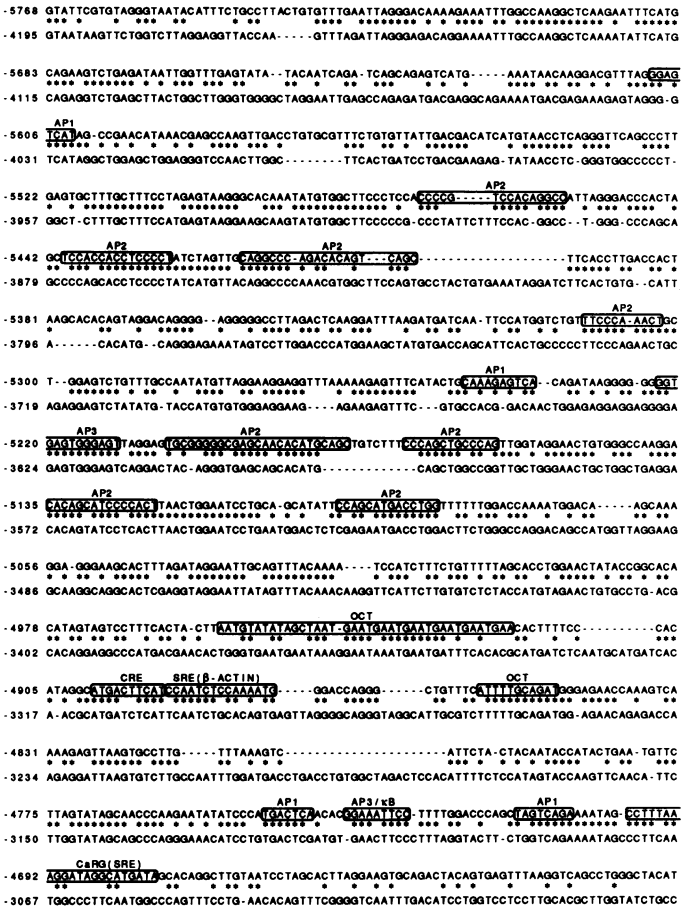

**Figure 4.** Region C sequence alignment between mouse (upper) and pig (lower) uPA promoters, annotated as in Fig 2. The sequences of the mouse and pig uPA promoters are aligned. Consensus transcription factor binding sites are marked, and discussed in the text.

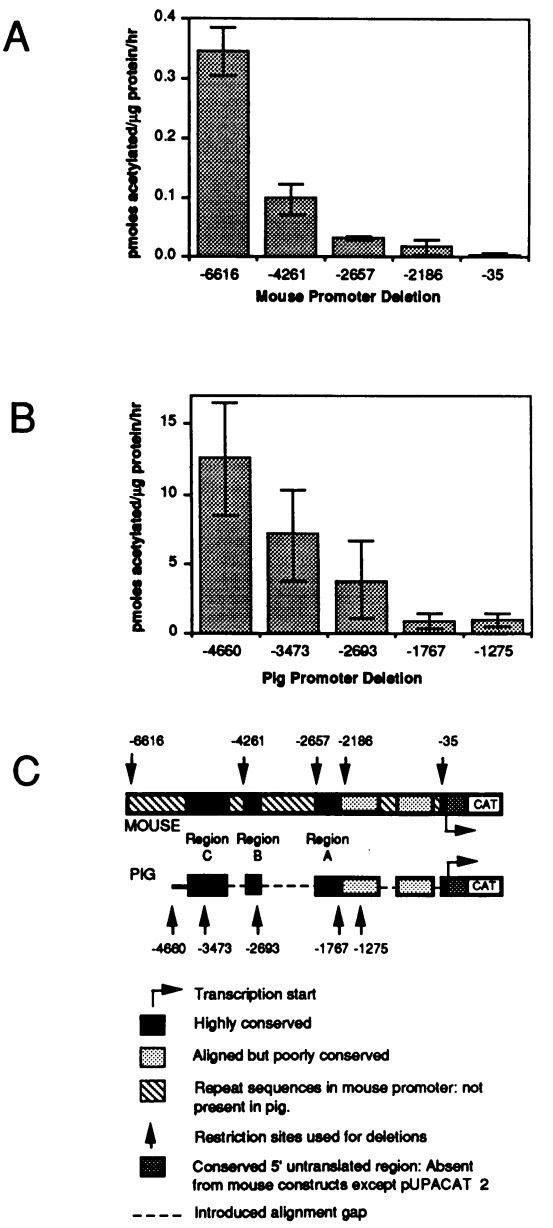


**Figure 5.** Expression of mouse and pig uPA promoter-CAT reporter gene constructs in transiently transfected RAW264 cells. A. CAT activities of RAW264 cells transfected with 20 μg of the mouse uPA promoter construct, pUPACAT (−6616) and molar equivalents of deletion constructs. Error bars show the range of values obtained in four independent experiments (each containing duplicate transfections). B. CAT activities of RAW264 cells transfected with 10 μg of of the pig uPA promoter construct, pCAT4660 and molar equivalents of deletion constructs. Error bars show the range of values obtained in three independent experiments (each containing duplicate transfections). In each experiment CAT activity dropped progessively with the first three deletions. C. Summary of mouse and pig reporter constructs. pUPACAT and mouse deletion constructs (Figure 5a) do not contain the conserved 5′ untranslated region. The 5′ untranslated region.is present in pUPACAT2 (Figure 7).

5′ non-coding sequences was prepared. The presence of this region caused a 5−10 fold increase in CAT expression compared to the shorter mouse construct, but a residual 3−5 fold difference between the pig and mouse promoters remained (Figure 6a).

The data of Rørth *et al.* (9) emphasised the role of a perfectly conserved sequence, (5′-AGGAGGAAATGAGGTCA-3′) in Region A (see Fig 2, −2440 in the mouse) in both constitutive and growth factor-inducible promoter activity in keratinocytes. This sequence was deleted from the full length pig promoter using the polymerase chain reaction (see Materials and Methods) and the resulting deletion tested for activity. As seen in Fig 6b, the deleted construct had 10-fold reduced activity compared to the control.

## Electrophoretic Mobility Shift Analysis

Rørth *et al.* (9) referred to the sequence above as a PEA3/AP1-like element. Macrophages have been shown to express a transcription factor, PU-1, which is a member of the *ets* family and binds to the core recognition sequence 5′-GAGG-AA-3′ (20) found in this element. Potential PU-1 sites are even more numerous in distal parts of the promoter. From −4970 to −5310 in mouse Region C (Figure 4), there are four polypurine tracts in which the core PU-1 (*ets* family) sequence (5′-GGAA-3′) is repeated six times. Between these tracts are several polypyrimidines, with the inverse *ets* sequence (5′-TT-CC-3′) repeated 4 times. These 300bp of mouse sequence are 65% conserved in the pig. Interestingly, of 77 base changes between the two species, 47 involve purine-purine or pyrimidine-pyrimidine substitutions suggesting there are constraints on base substitutions. In Region B (Figure 3) there are also several conserved purine strings and the *ets* core sequence is present six

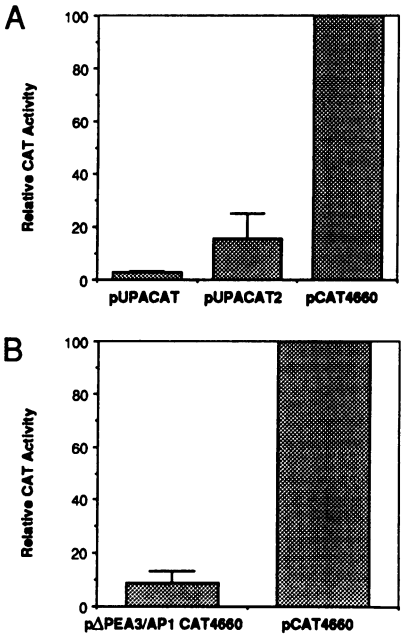


**Figure 6.** Expression of mouse and pig uPA promoter-CAT reporter gene constructs in transiently transfected RAW264 cells. A: Relative CAT activities of RAW264 cells transfected with 20 μg of pUPACAT and molar equivalents of pUPACAT2 and pCAT4660. Results were normalized to pCAT4660. Error bars show the range of values obtained in three experiments. B: Relative CAT activities of RAW264 cells transfected with 20 μg of pCAT4660 or pΔPEA3/AP1 CAT4660. Results are normalised to pCAT4660. Error bars show the range of values obtained in 3 experiments.

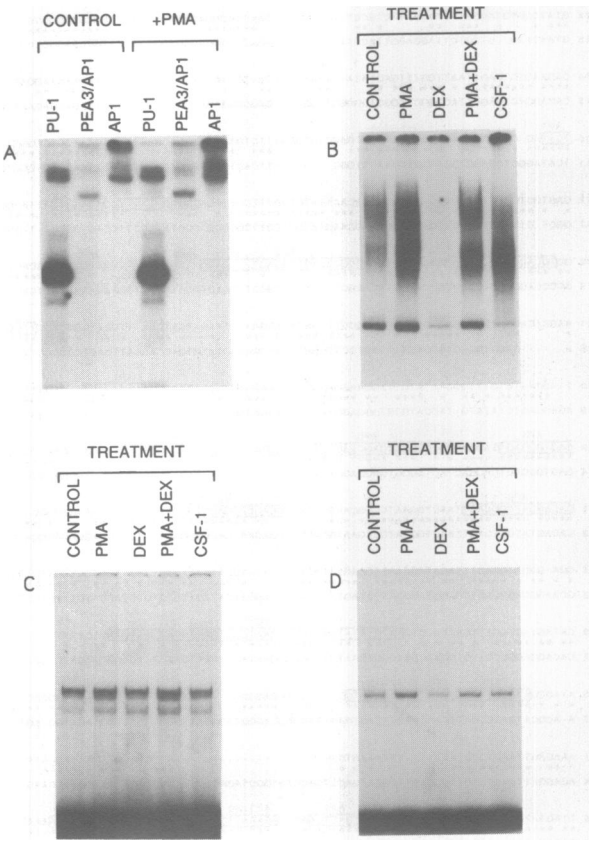


**Figure 7.** Electrophoretic mobility shift analysis of nuclear proteins expressed in RAW264 cells under various conditions. A) Nuclear proteins were incubated with the radiolabelled oligonucleotides shown (see Methods) and separated on 8% PAGE. The labelled probe has been run off to achieve greater resolution. Each probe was added at the same concentration and specific activity. B) Nuclear proteins were extracted from RAW264 cells treated for 2hr with 500 ng/ml PMA or for 6hrs with dexamethasone ($10^{-7}$ M), dexamethasone + PMA, or CSF-1 ($10^4$ U/ml) and incubated with radiolabelled AP1 oligonucleotide prior to separation on 6% PAGE. C) The same nuclear extracts as shown in Panel B were incubated with radiolabelled AP-2-like oligonucleotide from the uPA promoter (see Methods) and separated on 6% PAGE. D) The same nuclear extracts as shown in Panel B were incubated with radiolabelled κB oligonucleotide (see Methods) and separated on 6% PAGE.

times. For this reason, we performed electrophoretic mobility shift analysis of the expression of nuclear proteins in RAW264 cells that bind to the PEA3/AP1 element (Fig 7a). RAW264 cells contain two prominent nuclear activities that bound to a PU-1 oligonucleotide derived from the SV-40 early enhancer (20). The faster migrating.and more abundant complex contains a protein of approximately 32kD as determined by UV cross-linking to the radiolabelled probe and comparison to molecular weight standards (T. Dunn *et al.*; submitted for publication). The estimated Mr is consistent with its identity as PU-1 monomer (20). The nature of the slower migrating complex is not known, but we have shown elsewhere that it is also macrophage-restricted (T. Dunn *et al.*; submitted for publication). When the same extracts were incubated with an oligonucleotide spanning the PEA3/AP1-like region, no complex co-migrating with either of the PU1 bands could be seen. Instead, a single prominent band was evident. When cross-competition experiments were performed, the PU-1 oligonucleotide did not compete with the PEA3/AP1 oligonucleotide, nor was there any competition in the reverse direction (Figs 8a and 8b). We have reported elsewhere (T. Dunn *et al.*; submitted ) that RAW264 contains multiple nuclear proteins (including PU-1) that bind to the μB sequence of the immunoglobulin heavy chain enhancer (21). This sequence also contains a tandem repeat of the 5'-GGAA-3' *ets* family core recognition sequence. Hence, we performed competition experiments with the μB oligonucleotide, but again there was no competition for binding to the PEA3/AP1 oligonucleotide (Figs 8a and 8b). Cold PEA3/AP1 did compete for binding of the labelled probe, although even at 100-fold excess there was still some residual binding (Fig 8a). The PEA3/AP1 complex was resolved into two bands by running the complex to the bottom of a 6% polyacrylamide gel, but both bands had the same apparent binding affinity (not shown). Note that there is also some residual binding of the PU-1 oligonucleotide at 100-fold cold probe excess (Fig 8b). In summary, RAW264 cells contain a nuclear protein (s) that binds to the PEA3/AP1 element and is probably not an *ets* family member. In the upstream regions there are no sequences that are a better match for the PU-1 oligonucleotide than the PEA3/AP1 element. Although we have not tested each of them independently, we feel it is unlikely that any of them contributes to the expression of the gene in RAW264 cells.

This conclusion focuses attention upon the AP1-like component of the PEA3/AP1 element. Since the element has been implicated in growth factor responses in keratinocytes (9) we investigated whether agonists such as macrophage-specific growth factor, CSF-1, or phorbol esters, and an antagonist of of their action, glucorticoid, caused any change in expression of the nuclear protein that bound to the PEA3/AP1 region. No change was seen,

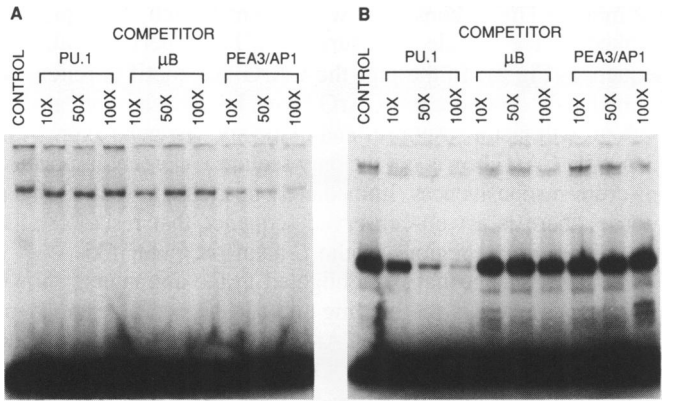


**Figure 8.** Electrophoretic mobility shift analysis of the binding specificity of nuclear Proteins Expressed in RAW264 cells. Panel A: Nuclear extracts from RAW264 cells were incubated with radiolabelled PEA3/AP1 oligonucleotide (see Methods) in the presence of the indicated excess of cold PU.1, μB or homologous probe and separated on 8% PAGE. Panel B: An identical experiment to Panel A, run back-to-back on the same apparatus, except that labelled PU.1 oligonucleotide was used.

and representative data is shown in Fig 7a, for PMA. Note in Fig 7a, that RAW264 cells did express multiple nuclear proteins that bound a consensus AP1 oligonucleotide, but these did not comigrate with the complex formed with the PEA3/AP1 element. The AP1-binding proteins of RAW264 cells are resolved more completely on 6% PAGE in Fig 7b. Untreated cells contain one well-defined complex, and two broad smears of slower-migrating binding activity, all of which can be competed by specific oligonucleotide (not shown). Unlike the PEA3/AP1 binding activity, AP1 binding proteins were acutely regulated. PMA caused a substantial increase in expression of all of the components of the two smears that bound to the consensus AP1 oligonucleotide, and steroid almost abolished the expression of all of these complexes (Fig 7b). The control for Fig 7b is provided in Figs 7a, c and d, which shows that none of these treatments altered expression of nuclear proteins that bound other oligonucleotides. The pattern of induction by CSF-1 was more puzzling, since only the faster-migrating smear was increased, whereas the more defined band actually decreased (Fig 7b). The conclusion is that the supposed AP1-like sequence (5′-TGAGGT-CA-3′) does not bind AP1 complexes generated in RAW264 cells. On the other hand, the uPA promoter does contain AP1 consensus sites; three in Region A, one in Region B and three in Region C (Figs 2,3 and 4).

In Region A, B and C we were struck by the presence in highly conserved regions of a motif containing repeats of the sequence 5′-CAGC-3′, strongly reminiscent of the AP-2 region of the SV-40 enhancer (30). One of these occurs around −2410 in the mouse promoter, in which the sequence is perfectly conserved between the three species for 29bp. We prepared an oligonucleotide covering this conserved region, and performed electrophoretic mobility shift assays as shown in Fig 7c. Two abundant, and prominent bands were observed in RAW264 cells, neither of which was altered in its abundance by CSF-1, PMA or steroid.

Finally, Region A contains a highly conserved κB consensus element, 5′-GGGAATTTCC-3′ at −2340 in the mouse. NFκB has been implicated in the macrophage-specific expression of the human immunodeficiency virus LTR (18). RAW264 cells contained only a single band that bound to an NFκB oligonucleotide based on the human immunodeficiency virus long terminal repeat. This band is also seen with a mutant oligonucleotide in which the G triplet is mutated to CTC (see 18) (data not shown). The abundance of this band was not changed by any of the treatments (Fig 7d). Slower migrating 'activated forms' (32,33) could not be detected in RAW264 cells. Such complexes were readily detectable in RAW264 cells that had been exposed to bacterial lipopolysaccharide (data not shown).

## DISCUSSION

This paper contains the first study of the transcriptional regulation of the uPA gene in macrophages using reporter gene constructs. We have presented sequence information comparing the pig and mouse uPA promoters, and have used transient transfection to confirm the biological importance of the regions of sequence homology between the two species. The sequence conservation between human and mouse in what we have called Region A was noted by Rørth *et al.* (9). Additional comparison of the pig sequence in Fig 2 highlights a block of around 250 bp that is most highly conserved. The homology blocks referred to as Regions B and C lie 5′ to the sequence reported by Rørth *et al.* (9). They lack substantial homology with Region A, as evidenced by the virtual absence of secondary alignments on the Pustell matrix (Figure 1). Surprisingly, the complex cyclic AMP-response element in the pig promoter, between −3415 and −3340 in Region C (Fig 4), identified by functional studies in pig LLC-PK1 cells (19) is very poorly conserved in the mouse sequence. There is conserved consensus cyclic AMP response element (5′-TGATCTCA-3′) immediately downstream (at −3310 in the pig; −4998 in the mouse). There could be a genuine difference in cyclic AMP responsiveness between the two species or between cell types, since uPA expression in murine macrophages is actually reduced by cyclic AMP agonists (11).

The combined PEA-3/AP-1 site (5′-AGGAAATGAGGT-CA-3′) within Region A has been implicated in phorbol ester/EGF induction of uPA in mouse keratinocytes (9) and was known to be present in the human gene. This sequence is almost perfectly conserved between mouse, pig and human (Figure 2). Deletion of this element from a 4.6 kb pig promoter caused a 10-fold reduction in expression of the reporter gene construct in RAW264 cells (Fig 6) Electrophoretic mobility shift analysis revealed that the cells contained nuclear protein(s) that formed a complex with the deleted DNA sequence. This complex is a candidate as a transcriptional regulator responsible, in part, for constitutive expression of the gene. Despite the presence in the PEA3/AP1 element of the core recognition sequence, 5′-GAGG-AA-3′, this complex did not co-migrate with either complex formed with the recognition sequence for the macrophage-specific transcription factor PU-1, a sequence has been recognised as being essential for the expression in macrophages of a mutant SV-40 virus lacking the 72 bp repeat enhancer (20). Further, in cross competition experiments, no competition was seen between the PU-1 and uPA sequences for binding to either of the complexes. Since the μB oligonucleotide, which also contains the polypurine string, did not compete for binding to the PEA3/AP1 element, we consider it unlikely that an *ets* family protein is involved in the complex.

The data in Fig 7 show that the complex that binds the PEA3/AP1 element in RAW264 cells is also quite distinct from those that bind an AP1 oligonucleotide. As noted by Rørth *et*

*al.* (9) the AP1-like sequence of the uPA enhancer, 5'-TGAGGTCA-3' is also closely related to a cyclic AMP-response element, 5'-TCACGTCA-3'. There is evidence that *c-jun* proteins, which form part of the AP-1 dimer with *c-fos*, can also form dimers with cyclic AMP response element binding (CREB) proteins (9). Given the constitutive expression of multiple AP-1 binding activities in RAW264 cells, we can presume that *jun* proteins are expressed. The involvement of *jun* proteins in the complex that bind the sequence TGAGGTCA therefore seems plausible, but a definitive conclusion awaits immunological characterisation. Whatever the nature of these two complexes, it should be noted that in keratinocytes proteins that bound to this sequence were induced by phorbol ester and growth factor (9), whereas in RAW264 cells they are present constitutively and completely unaffected by agonists.

The data in Fig 7d suggest, in contradiction to earlier claims that macrophage express activated NF$\varkappa$B (18), RAW264 cells contain only the so-called constitutive $\varkappa$B-binding protein, expression of which is not correlated with enhancer activity. Hence, the conserved $\varkappa$B element in Region A is unlikely to play a role in constitutive expression of uPA in RAW264 cells. Surprisingly, phorbol ester had no effect on the expression of activated forms of NF$\varkappa$B where, in the same extracts, there was clear induction of AP1 binding proteins (Fig 7b). The p50 and p65 subunits of NF$\varkappa$B are found in most cell types complexed to a cytoplamic inhibitor protein, I$\varkappa$B (31). The complex is dissociated upon treatment of the cells with phorbol esters, an effect that can by mimicked by direct phosphorylation with protein kinase C (32). In primary macrophages and macrophage cell lines including RAW264, NF$\varkappa$B is activated by bacterial lipopolysaccharide (33, and unpublished) which does not induce uPA. Nor does it have any effect on expression of uPA promoter-reporter constructs in RAW264 cells (data not shown). For these reasons we think it unlikely that the $\varkappa$B consensus sequence in Region A functions as an NF$\varkappa$B-dependent enhancer element in macrophages. The observation leaves open the question of why NF$\varkappa$B is not activated by PMA in RAW264 cells. We have confirmed this finding in primary macrophages (unpublished) raising the possibility that the activity of this factor is regulated differently in this cell type.

Despite the presence of multiple presumptive binding sites for transcription factors shown in Figs 7 and 8 , and evidence from the deletion study that at least one of their recognition sequences was essential for maximal gene expression, truncated promoter constructs containing Region A alone were poorly active in RAW264 cells. Maximal activity required both Region B and Region C (deleted progressively by the *Hind* III and *Eco*R V deletions respectively in Fig 5). Regions B and C contain additional consensus phorbol ester response elements, including a conserved AP1 consensus site at −4750, that would presumably be recognised by the AP1 that is constitutively expressed in RAW264 cells (Fig 4). Regions B and C also contain multiple repeats of AP2-like sequences, which can also be recognised by binding activities expressed constitutively in RAW264 cells (Fig 7c). In fact, the AP2-like sequences around −4720 (Region C) in the mouse is immediately adjacent to an AP3-like sequence (5'-GGAAATTCC-3') at −4740, which is, in turn, next to the conserved AP1 site at −4750. The overall structure is strikingly reminiscent of the SV-40 72 bp repeat (30). One other feature of Region C that may be relevant to expression of uPA is the presence of a serum-response element or CArG box, the 10 bp palindrome 5'-CCTTTAAAGG-3', from −4699 to −4690 in the mouse. This is conserved with one mismatch in the pig and flanked on the 5' side by a further 12 bp of perfectly aligned sequence (Fig 4). In the pig, the CArG box motif is reiterated immediately downstream. CArG boxes have been implicated in both growth factor responses and tissue-specific gene expression (34). The CArG box is not the only sequence element implicated in serum-responsiveness. Immediately upstream at −4880 in the mouse, there is a well-conserved sequence that resembles the serum-responsive regions of the $\beta$-actin promoter (35).

The last element that is highlighted in the alignments shown in Figs 2,3 and 4 is the binding site for the octamer family of transcription factors, 5'-ATGCAAAT-3' (36), recognised by the ubiquitous transcription factor Oct-1. The sequence at −4060 in the mouse (region B) is the 5'-TAATGARAT-3' sequence found in the Herpes simplex virus immediate early enhancer, which is known to bind Oct-1 (37). Macrophages, including RAW264 cells, contain both Oct-1 and a tissue-specific octamer binding protein, Oct-2 (36, T. Dunn *et al.*; submitted for publication).

The data in Fig 7 demonstrate that the transcribed 5' non-coding region plays a role in uPA gene expression. Sequences immediately downstream of the transcription start site have been shown to affect expression of transfected constructs in other systems (38), due to the presence of elements that control transcription initiation (39), transcription elongation (40) and mRNA stability or processing (41,42). There are also examples of enhancers within introns, for example in the $\beta$-actin and adenosine deaminase genes (35, 43). In other cases, the requirement seems to be for a splicing event, and the sequence of the intron upstream of the translation start site is irrelevant (34). Our data do not distinguish between these potential mechanisms.

Even when transcribed 5' non-coding sequences were included in both constructs, the pig promoter construct was more active in RAW264 cells than the mouse promoter (Fig 7). From the data in Fig 6, a pig construct that lacks Regions A, B and C (−1767) is 50-fold more active than an equivalent mouse construct (−2186). Assuming that the transcribed 5' non-coding region acts independently, this observation suggests there is some functional difference in the proximal promoter, a region that is poorly conserved between the two species (Fig 1). In the pig, the proximal promoter consists of a TATA box, three GC-rich Sp1-binding sites and a CCAAT box. The CCAAT box and two of the Sp1 sites are displaced to −450 bp in the mouse promoter by an intervening polypurine repeat starting at −60 bp. In previous studies using pig LLC-PK$_1$ cells, removal of the CCAAT box and the two more distal Sp1 sites from the pig uPA promoter greatly reduced basal activity in an *in vitro* transcription assay and abolished inducibility by cyclic AMP (14). Thus, the relatively low activity of the mouse promoter could be caused by the separation of these proximal elements.

The alignment between the mouse and pig sequences ceases around −5.8 kb and −4.2 kb respectively. The extreme 5' end of the mouse promoter sequence contains two head to head B2 type repeats based on the consensus sequence for similar elements in mouse uPA introns (6) suggesting that there is another shift in the sequence alignment between the two species at this point. This raises the possibility that more blocks of sequence conservation exist further upstream. As noted in the introduction, the introns of uPA genes are almost as highly conserved between species as the exons. The sequence data we have presented show that, after a gap of around 2 kb immediately upstream of the

transcription start site, sequence conservation reappears and extends for at least a further 2.5 kb (interrupted by repeat elements in the mouse). One obvious possibility is that there is another gene in this upstream region. We used the program MacVector to perform an analysis of open reading frames (ORFs) in the mouse promoter. The analysis reveals the existence of prolonged (>200 bp) continuous coding capacity that correlates well with the regions of sequence conservation with the pig gene (which also contains prolonged ORFs on the reverse strand). Region B contains consensus TATA boxes in both species (marked on Fig 3) although these do not align. A search for consensus splice donor acceptor sites did not reveal convincing intron-exon boundaries that are conserved between mouse and pig, and we could not detect any transcripts that hybridised with uPA promoter in RAW264 cells (unpublished). Furthermore, a search of the GenBank database failed to reveal any homology between the uPA promoter and any known coding sequence, and no mouse ORF could be aligned with a pig ORF at the protein level. These analyses do not eliminate the possibility that Regions B and C contain 5′ non-coding or leader sequences of another gene extending further upstream.

In summary, our analysis has revealed that the constitutive activity of the uPA promoter in RAW264 cells cannot be ascribed to a single tissue-specific element, but rather results from an interaction between multiple elements spread over at least 6.6 kb of 5′ flanking sequence in the mouse and 4.6 kb in the pig. These elements are separated from the transcription start site by a long stretch of unconserved DNA that might serve a spacer function, diluting the effect of a powerful enhancer. The existence of such a large and complex promoter may be a reflection of the multiples roles of uPA in development, and its ability to respond to numerous different external stimuli in different cell types.

## ACKNOWLEDGEMENTS

Much of the work in this report was performed separately in the two laboratories, and the authors contributed equally. The Australian authors wish to acknowledge that much of the groundwork that made this collaboration possible, in particular the isolation of the genomic DNA clones, was performed by our colleagues in Basel.

## REFERENCES

1. Danø,K. Andreasen,P.A., Grøndahl-Hansen,J., Kristensen,P., Neilsen,L.S. and Skriver,L. (1985) *Adv. Cancer Res.*, **44**, 139–266.
2. Nagamine,Y., Sudol,M. and Reich,E. (1983) *Cell*, **32**, 1181–1190.
3. Nagamine,Y., Pearson,D., Altus,M.S. and Reich,E. (1984) *Nucleic Acids Res.*, **12**, 9525–9541.
4. Verde,P., Boast,S., Franzè,A., Robbiati,F. and Blasi,F. (1988) *Nucleic Acids Res.*, **16**, 10699–10716.
5. Riccio,A., Grimaldi,G., Verde,P., Sebastio,G., Boast,S. and Blasi,F. (1985) *Nucleic Acids Res.*, **13**, 2759–2771.
6. Degen,S.J.F., Heckel,J.L., Reich,E. and Degen,J.L. (1987) *Biochemistry*, **26**, 8270–8279.
7. Leslie,N.D., Kessler,C.A., Bell,S.M. and Degen,J.L. (1990) *J. Biol. Chem.*, **265**, 1339–1344.
8. Grimaldi,G., Di Fiore,P., Locatelli,E.K., Falco,J. and Blasi,F. (1986) *EMBO J.*, **5**, 855–861.
9. Rørth,P., Nerlov,C., Blasi,F. and Johnsen,M. (1990) *Nucleic Acids Res.*, **18**, 5009–5017.
10. Bell,S.M., Brackenbury,R.W., Leslie,N.D. and Degen,J.L. (1990) *J. Biol. Chem.*, **265**, 1333–1338.
11. Hume,D.A. and Gordon,S. (1984) *Exp. Cell Res.*, **105**, 347–355.
12. Hamilton,J.A. (1982) *Arthritis Rheum.*, **25**,432–440.
13. Medcalf,R.L., Van den Berg,E. and Schleuning,W-D. (1988) *J. Cell Biol.*, **106**, 971–978.
14. von der Ahe,D., Pearson,D., Nakagawa,J., Rajput,B. and Nagamine,Y. (1988) *Nucleic Acids. Res.*, **16**, 7527–7544.
15. Collart,M.A., Belin,D., Vassalli,J.-D., de Kossodo,S. and Vassalli,P. (1986) *J. Exp. Med.*, **164**, 2113–2118.
16. Raschke,W., Baird,S., Ralph,P. and Nakoinz,I. (1978) *Cell*, **15**, 261–267.
17. Maxam,A.M. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560–564.
18. Osborn L., Kunkel,S. and Nabel,G.J. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 2336–2340.
19. Dignam,J.D., Lebovitz,R.M., and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489.
20. Klemsz,M.J., McKercher,S.R., Celada,A., Van Beveren, C. and Maki,R.A. (1990) *Cell*, **61**, 113–124.
21. Nelsen,B., Kadesch,T. and Sen,R. (1990) *Mol. Cell. Biol.*, **10**, 3145–3154
22. Lee,W., Mitchell,P. and Tijan,R. (1987) *Cell*, **47**, 741–752
23. von der Ahe,D., Pearson,D., and Nagamine,Y. (1990) *Nucleic Acids Res.*, **18**, 1991–1999.
24. Gorman,C. Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
25. Pustell,J.M. (1988) *Nucleic Acids Res.*, **16**, 1813–1820.
26. Kalb,V.F., Glasser,S., King,D. and Lingrel,J.B. (1983) *Nucleic Acids Res.*, **11**, 2177–2184.
27. Krayev,A.S., Markusheva,T.V., Kramerov,D.A., Pyskov,A.P., Scrybin, K.G., Bayev,A.A. and Georgiev,G.P. (1982) *Nucleic Acids Res.*, **10**, 7461–7475.
28. Wu,L., Anaraki,F., Morahan,P.S. and Leary,K. (1990) *J. Leuk. Biol.*, **48**, 229–236.
29. Ng,S.-Y., Gunning,P., Eddy,R., Ponte,P., Leavitt,J., Shows,T. and Kedes,L. (1985) *Mol. Cell. Biol.*, **5**, 2720–2732.
30. Wildeman, A.G. (1988) *Biochem. Cel Biol.* **66**:567–577
31. Nolan, G.P., Ghosh,S., Liou,H-C, Tempst,P. and Baltimore,D. (1991) *Cell*, **64**: 961–969
32. Ghosh, S. and Baltimore,D. (1990) *Nature* **344**, 678–682.
33. Nabel,G. and Baltimore,D. (1987) *Nature*, **326**, 711–713.
34. Tuil,D., Clergue,N., Montarras,D., Pinset,C., Kahn,A. and Phan-Dinh-Tuy,F. (1990) *J. Mol. Biol.*, **213**, 677–686.
35. Kawamoto,T., Makino,K., Orita,S., Nakata,A. and Kakunaga,T. (1989) *Nucleic Acids Res.*, **17**, 523–537.
36. Herr,W., Sturm,R.A., Clerc,R.G., Corcoran,L.M., Baltimore,D., Sharp,P.A., Ingraham,H.A., Rosenfeld,M.G., Finney,M., Ruvkun,G.and Horvitz,H.R. (1988.) *Genes Dev.*, **2**, 1513–1516.
37. Latchman, D.S., Partridge,J.F., Estridge,J.K. and Kemp,L.M. (1989) *Nulceic Acids Res.*, **17**, 8533–8542.
38. Kriegler,M. (1990) *Meth. Enzymol.*, **185**, 512–527.
39. Ayer,D.E. and Dynan,W.S. (1990) *Mol. Cell. Biol.*, **10**, 3635–3645.
40. Chen,Z., Harless,M.L., Wright,D.A. and Kellems,R.E. (1990) *Mol. Cell. Biol.*, **10**, 4555–4564.
42. Haile,D.J., Hentze,M.W., Rouault,T.A., Harford,J.B. and Klauser,R.D. (1989) *Mol. Cell. Biol.*, **9**, 5055–5061.
43. Braddock,M, Thorburn,A.M., Chambers,A., Elliott,G.D., Anderson,G.J., Kingsman,A.M. and Kingsman,S.M. (1990) *Cell*, **62**, 1123–1133.
44. Aronow,B., Lattier,D., Silbiger,R., Dusing,M., Hutton,J., Jones,G., Stock,J., McNeish,J., Potter,S., Witte,D. and Wiginton,D. (1989) *Genes Dev.*, **3**, 1384–1400.