Structural organization and regulation of the gene for the androgen-dependent glutathione peroxidase-like protein specific to the mouse epididymis

Norbert B. Ghyselinck, Isabelle Dufaure, Jean-Jacques Lareyre, Marie-Geneviève Mattei, Jean-Pierre Dufaure

Structural Organization and Regulation of the Gene for the Androgen-Dependent Glutathione Peroxidase-Like Protein Specific to the Mouse Epididymis

Norbert B. Ghyselinck, Isabelle Dufaure, Jean-Jacques Lareyre, Nicole Rigaudière, Marie-Geneviève Mattéi, and Jean-Pierre Dufaure

Laboratoire de Biologie Cellulaire
Centre National de la Recherche Scientifique URA 360
Université Blaise Pascal
63177 Aubière Cedex, France

Centre de Génétique Médicale (M.-G.M.)
Institut National de la Santé et de la Recherche Médicale U-242
Hôpital d'Enfants de la Timone
18385 Marseille Cedex 5, France

Genomic clones containing the gene for the glutathione peroxidase-like androgen-regulated murine epididymal protein of 24 kilodaltons (arMEP24) were isolated. A 9-kilobase DNA fragment was sequenced and found to contain the entire coding region of the gene, which is divided into five exons. The exact sizes and boundaries of the exon blocks were deduced by comparison with the cDNA sequence. One major and four weak transcription initiation sites in the epididymis were localized by primer extension. The promoter of the gene does not contain a conventional TATA box immediatly up-stream of the start site; rather, the sequence TATCA occurs at residue −35. Two CAAT boxes in opposite orientation and two putative binding sites for the transcription factor Sp1 were identified up-stream of the TATA-like box. To localize the *cis*-acting sequences responsible for androgen regulation of expression, fragments of the arMEP24 gene promoter region were cloned in front of the luciferase (LUC) reporter gene and cotransfected with an androgen receptor expression vector into CV-1 cells in a transient assay. LUC activities of CV-1 cells grown in the presence of various concentrations of 5α-dihydrotestosterone were compared to LUC activities of untreated controls. The DNA fragment containing up to 200 nucleotides up-stream from the major transcription start site was sufficient for the full promoter activity, but not for the responsiveness to androgen induction. Depending on the 5α-dihydrotestosterone concentration, a 2- to 4-fold induction of LUC activity was found if a −1797 to −167 arMEP24 gene fragment was used linked to the reporter gene driven by either the homologous promoter or the heterologous thymidine kinase promoter. Two or three copies of the imperfect palindromic sequence TGTTGAgagAGAACA, found at position −896 to −882 in the gene and ressembling the consensus steroid hormone-responsive element, are able to confer androgen regulation to the thymidine kinase promoter independently of their orientation. These findings support evidence that transcriptional regulation of the arMEP24 gene occurs via the sequence TGTTGAgagAGAACA. Homologies found in the sequence up-stream of the promoter with several putative binding sites for erythroid-specific *trans*-acting regulatory proteins are discussed. Finally, the *arMEP24 gene is located by in situ hybridization in* the [A2-A4] region of mouse chromosome 13. (Molecular Endocrinology 7: 258–272, 1993)

0888-8809/93/0258-0272$03.00/0
Molecular Endocrinology

INTRODUCTION

We previously described an androgen-dependent protein (arMEP24) specific to the caput epididymidis of the mouse (1, 2). It constitutes the major synthesized protein of this organ, where it is secreted into the epididymal fluid and binds to the spermatozoa (3, 4). The amino acid sequence of the polypeptide has been determined (5). It shows high homology with the glutathione peroxidase (GSHPx) (6, 7), which is an enzyme of considerable importance for several reasons. First, it has a protective role in removing lipid hydroperoxides

and H_2O_2 formed during normal oxidative metabolism, principally in the red blood cells (8), but also in other tissues. Secondly, in certain tissues (*e.g.* platelets, leukocytes, and lungs), GSHPx reduces, and thereby cleaves, the hydroperoxi-fatty acids formed by the action of lipoxygenases on lipid fatty acids; this generates the immediate precursors of the leukotrienes, which are involved in inflammatory responses (9). Finally, mouse spermatozoa and epididymal fluid have GHSPx activity, which is the major protective system against oxidative damage (10). This could maintain membrane flexibility and prevent the premature acrosome reaction that would otherwise occur (6).

In previous studies we have shown that arMEP24 mRNA expression is androgen responsive (1, 2, 11). Androgen-induced expression of genes is generally mediated through the androgen receptor (AR), which has been recently cloned (12–14). The AR is a member of the superfamily of ligand-inducible transcription factors, which includes steroid, thyroid hormone, and retinoic acid receptors (15–17). These ligand-inducible transcription factors can bind to specific DNA sequences and regulate the transcriptional activity of their target genes (17, 18). The DNA structures to which steroid receptors bind are imperfect palindromic sequences. The consensus sequence of the responsive element for glucocorticoid and progesterone receptors is GGTACAnnnTGTTCT (17). Although the expression of many genes is known to be regulated by dihydrotestosterone (DHT), well defined androgen-responsive elements (ARE) interacting with the DHT-AR complex have not yet been clearly identified (19). The best known example is the mouse mammary tumor virus (MMTV) promoter, which is able to confer androgen-responsiveness to a reporter gene (20–22). In the C3(1) gene (23, 24), the first intron is also able to bind the AR (19, 25) and confer weak androgen responsiveness to the thymidine kinase (tk) promoter (25, 26). More recently, a new functional ARE has been identified in the promoter of the prostate-specific antigen (PSA) gene (27). In the epididymis, several genes have been described as androgen dependent (28). Nevertheless, there is no evidence that their transcription is directely controlled by androgens via DNA *cis*-acting elements.

In this study, we report on the molecular cloning of the gene encoding arMEP24. Its organization, sequence, and chromosomal localization are described. The promoter and the 5′-flanking region contain some noteworthy structural elements, which are discussed. We also provide evidence that the imperfect palindromic sequence TGTTGAgagAGAACA, located about 900 nucleotides (nt) up-stream of the major transcription start site, is involved in androgen-regulated transcription of the gene.

RESULTS

Isolation and Structure of arMEP24 Gene

To isolate genomic clones for arMEP24, we used the cDNA clone M53 as a probe (5). About 1.5×10^6 phage plaques of a mouse BALB/c genomic library were screened, and two plaques were identified. This is in agreement with a single copy gene (1). These two overlapping clones were partially restriction mapped (Fig. 1, *top*). Southern hybridization of DNAs from λ91arMEP24 and λ151arMEP24 showed that the cDNA clone hybridized within a 9-kilobase (kb) *Hin*dIII fragment and that clone λ151 was positioned 3′ relative to λ91. To attest that the phage sequences are normally present in this orientation in the murine genome, mouse genomic DNA was cut with appropriate restriction enzymes, and the fragments were analyzed by hybridization of Southern blots to probes derived from λ91 and λ151 subclones. The genomic DNA yielded bands of the expected sizes (not illustrated), confirming the gene structure. A *Hin*dIII fragment of about 9 kb containing the whole gene was further mapped (Fig. 1, *bottom*) and then sequenced (Fig. 2). Comparison of the genomic DNA sequence with the cDNA sequence (6) indicates that the mRNA is encoded by five exons interrupted by four introns (Figs. 1 and 2). The sequences at the exon/intron boundaries are consistent with the splice junction sequences observed for other genes (29). In the first exon, an in-frame ATG codon is found 63 nt up-stream from the first codon ACC of the mature protein (4, 7) and is likely to be the translation initiation codon. This ATG is in an environment, GTCATGG, ressembling the optimal sequence for initiation of translation by eukaryotic ribosomes (30). The deduced 21 first amino acids (*underlined* in Fig. 2) form a sequence rich in hydrophobic residues, as expected for a signal peptide (31). The cleavage site between the two threonine residues is in agreement with the rule previously defined (32). According with this, arMEP24 is demonstrated to be a secretory protein (1, 3, 4) whose amino-terminal sequence was determined (4, 7) by direct Edman degradation on the mature protein (*boxed* in Fig. 2). Additionally, two B1 Alu-like repeated elements are observed in the sequence (*dotted-underlined*). The first one, which is flanked by 8 nt direct repeats (AAAGAGCT) starts at position 299 and ends at nt 510. The second element in the last exon has been described previously (6). Together they contain sequences that are reminiscent of the two intragenic RNA polymerase-III control regions (33).

Analysis of the Promoter Region

The 5′-terminus of the gene was determined by primer extension experiments in which the primer consisted of a 5′-end-labeled synthetic oligonucleotide 20 nt in length located in the first exon (Fig. 3). A sequencing ladder of a known DNA fragment run side by side with the primer-extended fragment positioned a major transcription start site 109 nt up-stream from the 5′-end of the primer. Four additional weaker initiation sites were present at 117, 129, 133, and 134 nt. Other faint bands smaller than 109 nt probably resulted from premature termination of reverse transcription (not shown). The major assigned 5′-end, which is numbered 1, was an

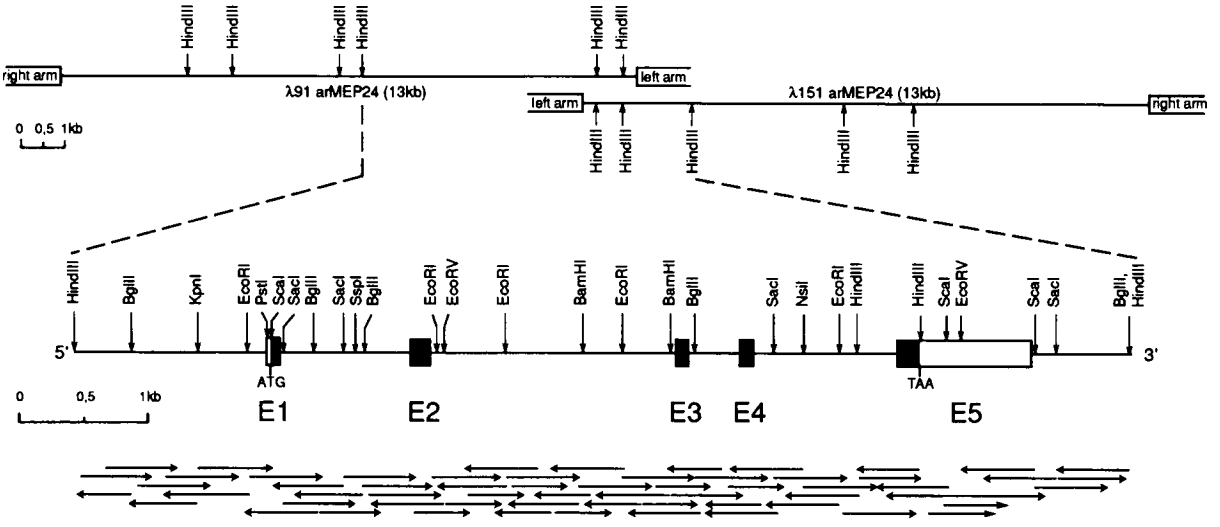

Fig. 1. Restriction Map of the arMEP24 Gene and Sequencing Strategy

Top, Partial *Hin*dIII map of phage clones λ91arMEP24 and λ151arMEP24. *Middle*, Enlarged map of the gene on which exon boxes are numbered E1-E5. The protein-coding region is indicated by the *black boxes*, while the 5′- and 3′-untranslated regions of the mRNA are indicated by *white boxes*. Initiation and stop codons are shown *below*. Restriction endonucleases sites are indicated. *Bottom*, Sequencing strategy. Fragments were subcloned into pGem7Zf(−) vector and sequenced on both strands by the dideoxy chain termination procedure. *Arrows* indicate the direction and extent of sequencing.

A residue preceded by a C (Fig. 2), the generally preferred cap site (29). The genomic DNA 5′ to the major transcription start point was analyzed for potential *cis*-acting elements that may control initiation of transcription. Several structural elements indicative of a promoter were found (Fig. 2). An unconventional TATA box (TATCA) was present at −40 to −35 in the sense strand relative to the major start site. The TATA box is believed to aid in fixing the start site of transcription, and most genes that lack the authentic TATA element have multiple start sites (34–36). Such gene promoters contain one or more putative binding sites for transcription factor Sp1 (GC box) or its complement (35, 36). This motif was found at −150 to −142 and −63 to −54. Three other GC boxes were found up-stream of the promoter region at −1023 to −1014, −1163 to −1154, and −1193 to −1184. Two CAAT boxes were present in the promoter. One motif occurred in the coding strand at −122 to −118 (CAAT), and the other was located at −78 to −74 (ATTGG) in the complementary strand. CAAT elements in this orientation have been found to be functional in other genes (37).

A series of arMEP24 putative promoter fragments was cloned in front of a reporter gene. Since the major transcription start site was found 19 nt up-stream from a *Spe*I site, we ligated the *Hin*dIII-*Spe*I (~2000 nt), *Kpn*I-*Spe*I (~600 nt) and *Eco*RI-*Spe*I (~200 nt) fragments containing this transcription start point to the firefly luciferase gene (LUC) as a transcriptional reporter (38) to obtain pMEP1-LUC to pMEP6-LUC vectors (see *Materials and Methods* or Fig. 8 for details). Two types of constructs were made, which contained the promoter sequence and flanking regions in either the sense or antisense orientation with respect to the LUC gene. A schematic representation of the different constructs is shown in Fig. 4. To study androgen regulation of the arMEP24 promoter, CV-1 cells were cotransfected with each of the pMEP-LUC constructs and with pSVARo, an AR expression vector (39). After transfection, the stimulatory effect of the different 5′-genomic fragments on the rate of LUC gene transcription in the absence and the presence of 10^{-7} M DHT was determined by performing LUC assays (Fig. 4). The luciferase activities were normalized for efficiency of transfection by cotransfection with a plasmid containing the β-galactosidase gene under the control of simian virus-40 early promoter (40). In the absence of hormone (Fig. 4, □), the construct pMEP1-LUC containing the minimal promoter region (*Eco*RI-*Spe*I, −167/+24) as well as pMEP3-LUC and pMEP5-LUC containing larger DNA fragments (*Kpn*I-*Spe*I, −577/24; *Hin*dIII-*Spe*I, −1797/24) stimulated transcription of the LUC gene several-fold compared to the background activity of the parent vector pLUC. The control plasmid ptk-LUC, which contains the herpes virus tk gene promoter, showed a transcriptional activity similar to those of pMEP1-LUC, pMEP3-LUC, and pMEP5-LUC. In contrast, none of the other pMEP-LUC vectors, containing the same genomic DNA fragments in the antisense orientation, showed significant stimulation of LUC activity. To investigate whether transcription of the promoter/LUC reporter gene constructs started at the same sites as transcription of the wild-type arMEP24 promoter, we performed primer extension experiments with total RNA extracted from the transfected CV-1 cells, using a primer derived from the published luciferase sequence (38). The 5′-start sites in RNA from transfected cells mapped at the same positions as those observed using RNA from

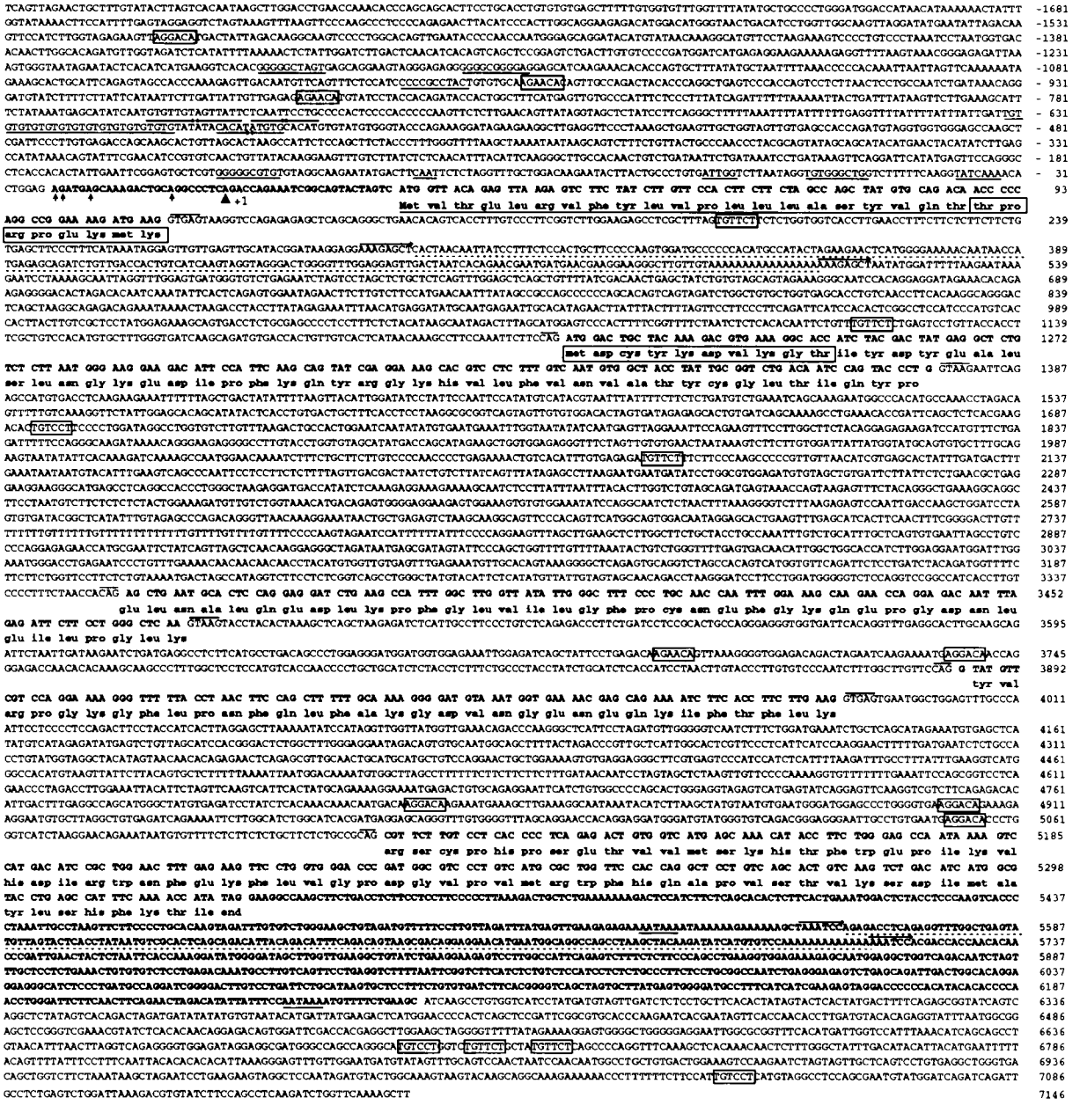

Fig. 2. Nucleotide Sequence of the arMEP24 Gene

The mouse genomic *Hin*dIII fragment of 9 kb was cloned and sequenced, as described in *Materials and Methods*. The nt are numbered according to the major transcription initiation site, as determined by primer extension. The major transcription start site is indicated by a *large arrowhead*, while minor sites are indicated with *small arrowheads*. The exon sequences, which were determined by comparison with the cDNA insert of clone M53 (5), are in *boldface*. Each exon is translated into the amino acid sequence below the nt sequence. The signal peptide is *underlined*, and the *boxed* residues are confirmed by direct peptide sequencing of purified 24-kDa protein. The consensus splice signal GTRAG/CAG flanking the introns are *overlined*. The proposed TATA-like box, CAAT boxes, GC-rich boxes, the two polyadenylation signals AATAAA, and the homopurine-homopyrimidine stretch are *underlined*. The 16-core putative ARE TGTYCT are *boxed*. Repeat and dyad structures are indicated by *arrows*, while the B1 Alu-like elements are shown by *dotted lines*.

mouse epididymis (Fig. 5). It can be concluded, therefore, that the region −167/24 indeed represents the arMEP24 gene promoter.

Evidence for an ARE

In the presence of 10^{-7} M DHT (Fig. 4, □), the constructs pMEP1-LUC and pMEP3-LUC were not induced. In contrast, pMEP5-LUC was responsive to androgens, with an average induction factor of 2, compared to 1 or 1.1 for all other constructs. The plasmid ptk-LUC was used as a negative control of stimulation. As a positive control of induction efficiency, we used a construct containing the sequence of the MMTV long terminal

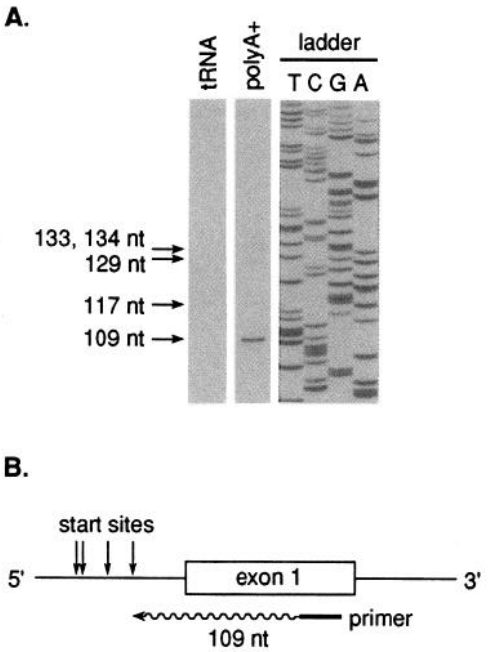

Fig. 3. Primer Extension Analysis of the 5′-End of arMEP24 mRNA

Epididymal poly(A)$^+$ RNA was hybridized with a ^{32}P-labeled primer located in the first exon and extended with AMV reverse transcriptase. The products were analyzed by polyacrylamide gel electrophoresis and autoradiography, as described in *Materials and Methods*. The sizes of the extented products, as determined from the known fragments sizes of the DNA sequence ladder, are indicated on the *left*. A, Primer extension products from yeast carrier transfer RNA (lane tRNA) and from epididymal poly(A)$^+$ RNA [lane poly(A)$^+$]. Lanes T, C, G, and A are ^{35}S-labeled DNA sequencing reaction. B, Diagram of the 5′-portion of the gene. The E1 exon sequence is indicated by an *open box*. The *thick line* and the *wavy line* stand for the synthetic primer and the extended cDNA transcripts, respectively. Transcription start points are represented by *vertical arrows*.

repeat (LTR) in front of the LUC reporter gene (pMMTV-LUC). Such MMTV-LTR-reporter gene constructs were previously described as androgen responsive in transfected cells (20–22). An induction factor of 188 was observed with pMMTV-LUC. Omission of pSVARo from the transfection mix abolished hormonal induction of expression (not shown).

To verify whether the fragment *Hin*dIII-*Eco*RI (−1797/−167) contains a functional ARE responsible for androgen responsiveness of the pMEP5-LUC construct, it was cloned in both sense and antisense orientations in front of the heterologous tk promoter-LUC construct. Cotransfections with the AR expression vector in the presence of DHT ranging from 10^{-8}-10^{-5} M evidenced an average 2- to 4-fold increase in LUC activity independently of the orientation of the *Hin*dIII-*Eco*RI fragment (Fig. 6). This effect was not observed with the control ptk-LUC vector. From the sequence analysis of the genomic 5′-flanking region, we found the imperfect palindromic motif TGTTGAgagAGAACA at position −896 to −882, which ressembles the core consensus ARE (19). To obtain evidence that this sequence can act as a functional ARE, it was taken out of the context of the arMEP24 promoter, and two or three copies were linked to the ptk-LUC construct (see *Materials and Methods*). The results of the cotransfections with pSVARo in the presence of various DHT concentrations are shown in Fig. 7. The stimulation of transcription increased with the amount of available DHT in a dose-dependent manner, to reach a maximal efficiency at 10^{-5} M. A relative induction factor of 3 was measured with pARE3-tk-LUC containing three copies of the putative ARE in the sense orientation (Fig. 7, ■). The construct pARE2-tk-LUC, harboring two copies of the ARE in the antisense orientation, induced a 4-fold increase in transcription (□). Expression of the ptk-LUC control construct without any ARE was not inducible by DHT (Fig. 7, ○), while pMMTV-LUC expression was stimulated with an average induction factor of about 200 (●).

Other Potential Regulatory Elements

In the region up-stream of the promoter, dyad symetry structures were found at −610 to −597 (DS1), −595 to −586 (DS2), and −600 to −590 (DS3). DS1 and DS2 or DS2 and DS3 overlap (Fig. 2). Thus, only one of them can form a stem-loop structure in a single molecule. They are preceded by a $(TG)_{14}$ stretch (Fig. 2). A similar structure is present about 1500 nt down-stream of the 3′-end of the gene (not shown). These elements, which have the potential to adopt the Z-DNA conformation (see *Discussion*), are common (10^5 copies) in eukaryotic DNA (41). They are believed to regulate the expression of the adjacent genes (42). We also analyzed the 5′-flanking region for homology with potential regulatory elements by comparison with a catalog of consensus sequences (43). Several interesting motifs were found (Fig. 8), including three copies of the AP-2-binding protein (44), three motifs for the CF1 transcriptional activator (45), and one recognition sequence for Oct-4 transcription factor (46, 47). The putative PEA3-responsive element for *Ets* oncogenes, serum growth factor, and phorbol ester is found three times in the up-stream 5′-region (48–51). Most notable is the presence of six consensus sites for the GATA-1-binding protein, previously called GF-1, NF-E2, and Eryf1 (52–54), an erythroid-specific zinc finger protein that has been shown to regulate the transcription of a number of erythroid-specific genes, including the α, β, and γ-globin genes (55). An AP-1-like motif for the second erythroid-specific factor NF-E2 (56–58) is also present associated with several GT/CAC boxes similar to those observed up-stream of many erythroid promoters and regulatory elements (59, 60).

Chromosomal Localization of arMEP24 Gene

The high degree of amino acid conservation and similarity in the overall structures of arMEP24 and GSHPx

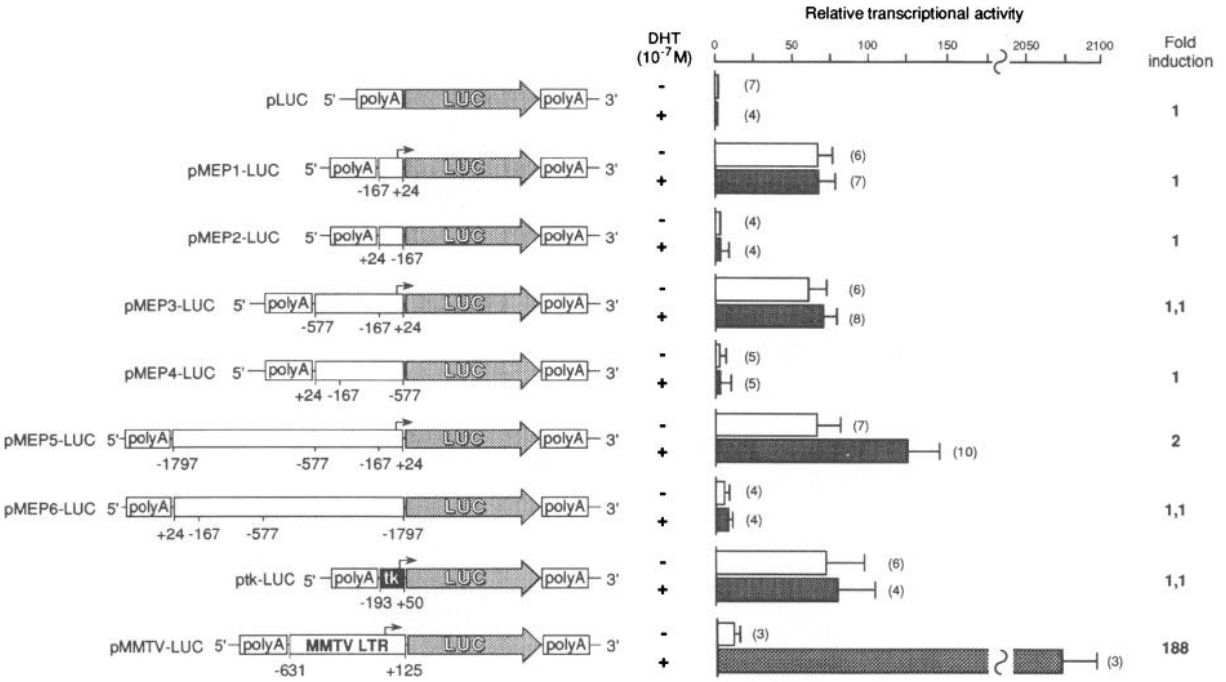

Fig. 4. Analysis of arMEP24 Gene Promoter Activity in CV-1 Cells

Different lengths of arMEP24 gene sequences up-stream of the transcription start sites were inserted into the plasmid containing a promoterless LUC reporter gene to generate pMEP-LUC vectors. Nucleotide numbering is according to Fig. 2 and previously reported sequences. *Broken arrows* represent the transcription start sites. Details of constructs are described in *Materials and Methods*. Cotransfections of pMEP-LUC constructs with pSVARo were performed in CV-1 cells in the absence (□) and presence (□) of 10^{-7} M DHT. Values were standardized relatively to β-galactosidase activities. Each *bar* is the mean of several independent assays (indicated in *parentheses*), and the *error bars* indicate the SD. Induction factors, indicated on the *right*, were calculated by dividing the relative LUC activities obtained in extracts of cells grown in medium with hormone by that obtained in extracts of cells grown in the absence of hormone.

(6, 7) strongly suggest that they evolved from a single ancestral gene by duplication. We found it of interest to determine the chromosomal localization of the arMEP24 gene. *In situ* hybridization experiments were carried out using mouse metaphases spreads. A recombinant pGEM7Zf(-) plasmid containing a 1.4-kb arMEP24 cDNA (5) was used as a probe. In the 100 metaphase cells examined after *in situ* hybridization, there were 303 silver grains associated with chromosomes, and 44 of these (15%) were located on chromosome 13. The distribution of grains on this chromosome was not random; 75% of them mapped to the [A2-A4] region of chromosome 13 (Fig. 9). These results demonstrate that the localization of the arMEP24 gene is to the [A2-A4] region of chromosome 13 in the murine genome.

DISCUSSION

We previously described the homology of arMEP24 with GSPHx (6, 7). Here we report on the structure and hormonal regulation at the molecular level of the gene encoding for arMEP24. Primer elongation of mRNA indicated the existence of a major and four weak start sites of transcription that map within a stretch of about 25 nt. The region up-stream of the multiple initiation sites contains an unusual TATCA box located −35 nt from the major one. Since *in vitro* transcription studies have shown the authentic TATA box serves to fix the site at which transcription will start (61), the fact that the arMEP24 gene has numerous transcriptional initiation sites is not surprising. Several sequences ressembling the consensus CAAT box (62) and the GC box for the Sp1-binding site (63) are present. It has been hypothesized that the presence of the latter facilitates the recognition of a weak TATA box (64). Similar GC boxes were first found in the simian virus-40 early promoter region (61), the tk gene of herpes simplex virus (65), and many so-called GC-rich housekeeping genes whose activity is required in all cells and are not subject to environmental control (66–69). Nevertheless, the arMEP24 gene promoter cannot be considered a GC-rich one. Rather, it belongs to another class of genes that have no consensus TATA element and are not GC rich. Unlike GC-rich promoters, many of these promoters are not constitutively active, but are regulated during differentiation or development and initiate transcription at only one or a few tightly clustered start sites (70). In these promoters, an initiator element may be essential for both core promoter strength and deter-

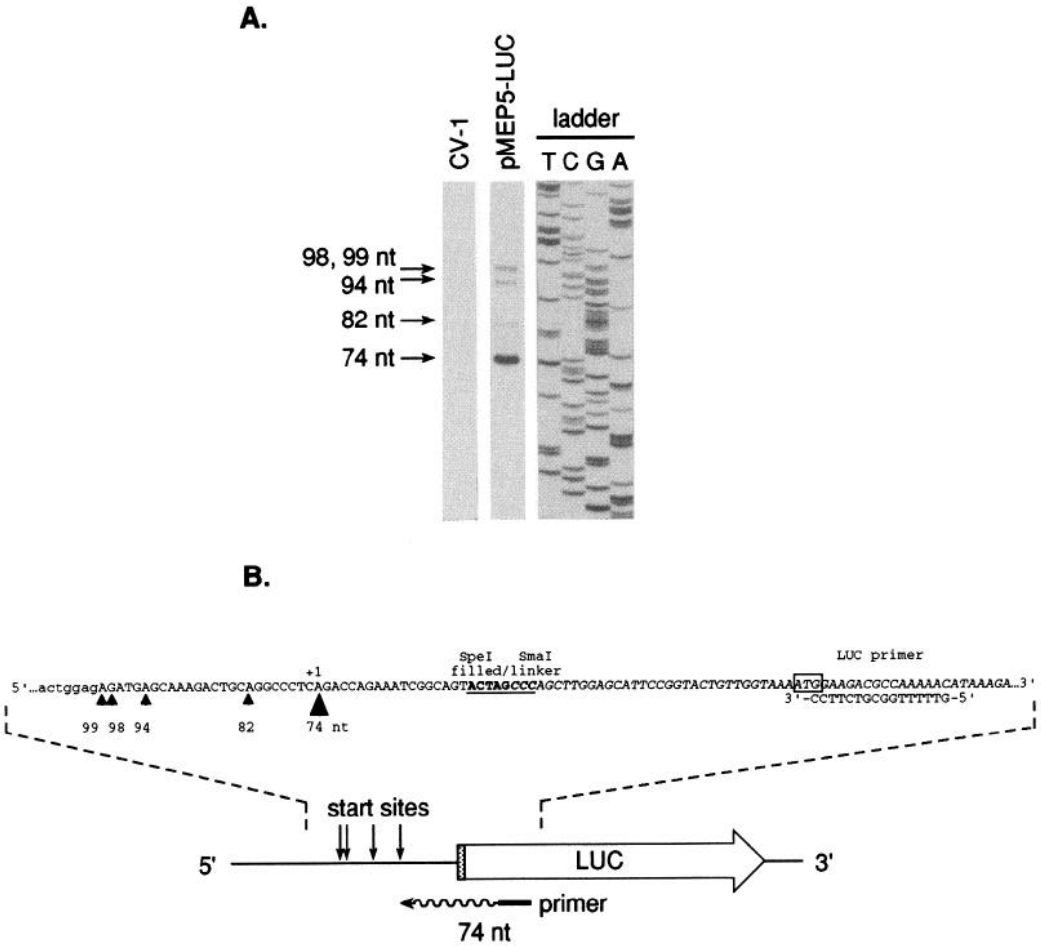

Fig. 5. Primer Extension Analysis of the 5′-End of Hybrid arMEP24/LUC mRNA in Transfected CV-1 Cells

Total RNA from transfected CV-1 cells was hybridized with a ^{32}P-labeled primer located in the coding region for LUC and extended with AMV reverse transcriptase. The products were analyzed by polyacrylamide gel electrophoresis and autoradiography, as described in *Materials and Methods*. The sizes of the extented products, as determined from the known fragments sizes of the DNA sequence ladder, are indicated on the *left*. A, Primer extension products from total RNA of untransfected CV-1 cells (lane CV-1) and from total RNA of CV-1 cells transfected with pMEP5-LUC plasmid DNA (lane pMEP5-LUC). Lanes T, C, G, and A are ^{35}S-labeled DNA-sequencing reaction. B: *Bottom*, Diagram of the promoter region of the hybrid reporter gene. The LUC-coding sequence is indicated by an *open arrowbox*. The *shaded box* represents the *Sma*I linker. The *thick line* and the *wavy line* stand for the synthetic primer and the extended cDNA transcripts, respectively. Transcription start points are represented by *vertical arrows*. *Top*, An enlargement shows the nt sequence of the hybrid gene with the primer position. Untranscribed sequences are in *lowercase letters*, arMEP24/LUC hybrid mRNA sequences are in *uppercase letters*, and LUC sequences are in *italics*. The LUC translation start codon ATG is *boxed*. The *Spe*I/*Sma*I junction is in *boldface* and *underlined*. Major (+1) and weak transcription start points are indicated by *large and small arrowheads*, respectively.

mining the actual initiation site (71). Interestingly, between position −21 to 10, the arMEP24 gene promoter shows 65% homology with the transcription start site region of the terminal deoxynucleotidyl transferase gene, in which the initiator element was first described (70). Furthermore, residues that are essential for the initiator activity are perfectly conserved (70). It has been recently hypothesized that the transcription initiation complex assembly might be nucleated by binding of transcription factors on either TATA-like or initiator elements, depending upon the relative concentrations and activities, or on both elements in concert (71). This would result in more diversified responses to distal regulatory elements. In the 5′-flanking region, most notable is a run of $(TG)_{14}$. TG repeats have been found in genomes of yeast, mouse, and man (72, 73). These sequences often show hypersensitivity toward single strand-specific nucleases (74). This might reflect the presence of conformational alterations in the DNA, which could act as general recognition signals for the nearby presence of a transcription-initiation or promoter region. By virtue of their ability to form left-handed DNA helices of Z-DNA (75), the conformational changes might facilitate or direct the binding of transcription or regulatory factors to more specific sequences on the promoter. Several DS structures are present near the $(TG)_{14}$ stretch. The rat prostatic binding protein C2 and C3(1) genes (24, 76), the rat seminal vesicle F and S

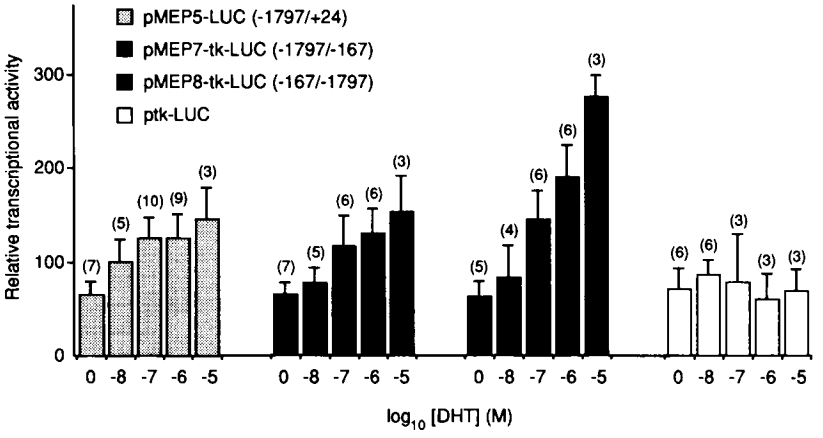

Fig. 6. Transcription-Stimulating Activity of the *Hin*dIII-*Eco*RI Fragment (−1797/−167) Transfected into CV-1 Cells in the Presence of Androgens

CV-1 cells were cotransfected with pMEP5-LUC (containing the genomic fragment −1797/24) or pMEP7-tk-LUC or pMEP8-tk-LUC (containing the genomic fragment −1797/−167 linked to the heterologous tk promoter in either the sense or antisense orientation, respectively) and pSVARo. LUC activity was measured in the presence of various DHT concentrations and normalized to β-galactosidase activity. Each *bar* represents the mean of several independent assays (indicated in *parentheses*), and the *error bars* indicate the SD.

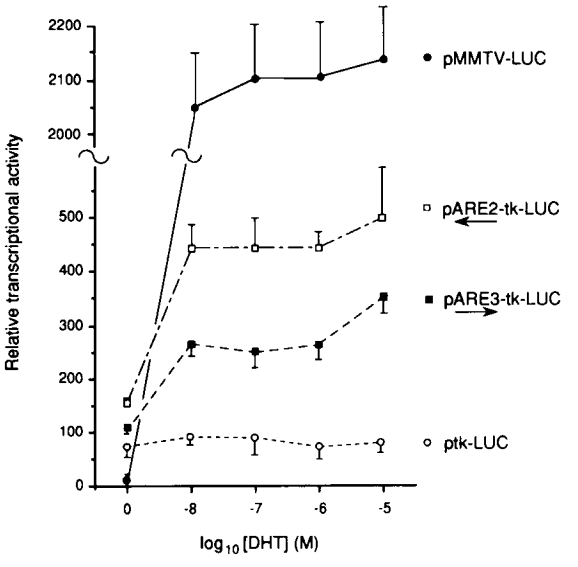

Fig. 7. Transcription-Stimulating Activity of the −896 to −882 ARE Motif Transfected into CV-1 Cells in the Presence of Androgens

CV-1 cells were cotransfected with pSVARo and pARE3-tk-LUC (■) or pARE2-tk-LUC (□) containing three or two copies of the putative ARE sequence TGTTGAgagAGAACA (−896/−882) cloned in front of the tk-promoter LUC reporter gene construct (ptk-LUC) in the sense or antisense orientation, as indicated by *arrows*. Transcriptional stimulation was measured in the absence and presence of various DHT concentrations. Constructs ptk-LUC (○) and pMMTV-LUC (●) were used as negative and positive controls, respectively. Values were standardized relative to β-galactosidase activities. Experimental details are described in *Materials and Methods*. Each *point* represents the mean of at least three independent assays, and the *error bars* indicate the SD.

genes (77, 78), the mouse renin 2 gene (79), and the rat androgen-binding protein gene (80), all androgen-controlled genes, have short inverted repeats in their promoter regions. Therefore, it is likely that the repeats of DS structures in androgen-controlled genes could be involved in the regulation of expression.

In this study, we have shown that expression of the arMEP24 gene can be up-regulated on the level of transcription by DHT. We have evidenced a functional ARE located at position −896 to −882 (TGTTGAgagAGAACA) that can function both in the context of the arMEP24 promoter or as a separate element when cloned in front of the tk promoter. The sequence of this imperfect palindromic ARE ressembles the reverse complement consensus sequence AGAACAnnnTGTACC for binding of glucocorticoid and progesterone receptors (17), in which the order of the two half-sites is inverted. The homology is less marked with the other AREs characterized so far (21, 26, 27). The level of induction is rather low in our experiments (~2-fold). However, it is comparable to the inductions observed in a study of the effect of androgen on the transcription rate of the arMEP24 gene in run-on experiments performed on isolated nuclei from the epididymides of castrated and normal mice (Lareyre, J. J., unpublished results). Similarly, transcription of the prostate-specific kallikrein-like gene is only 2-fold induced by androgens (81). This low induction will represent at least part of the total androgen-responsive regulation mechanisms, which determines the final mRNA concentration. Obviously, the change in the rate of arMEP24 mRNA transcription is not sufficient to account for the very dramatic *in vivo* changes observed in the steady state level after castration and testosterone replacement (1, 2). Then, the role of arMEP24 mRNA turnover in gene control is also likely to be important. Indeed,

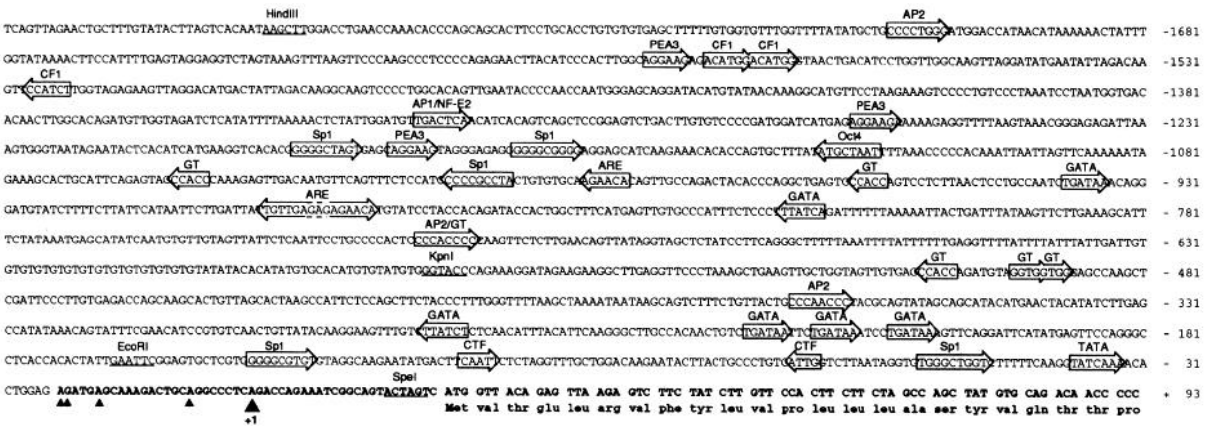

Fig. 8. Sequence Features of the arMEP24 Gene Promoter and 5′-Flanking Regions

About 2000 residues of the sequence in the promoter region extending from nt −1831 to 93 are shown. The major initiation site of transcription, numbered 1, is indicated with a *large arrowhead*. Other weak start sites are indicated by *small arrowheads*. The transcribed sequence is in *boldface*, and the translated amino acid sequence is indicated above with the three-letter code. Restriction sites used to construct pMEP-LUC vectors are *underlined*. The motifs found to be homologous to *cis*-acting elements in other promoters and/or enhancers are *boxed*, and the *arrow* indicate their orientation. Reference numbers are given in the text.

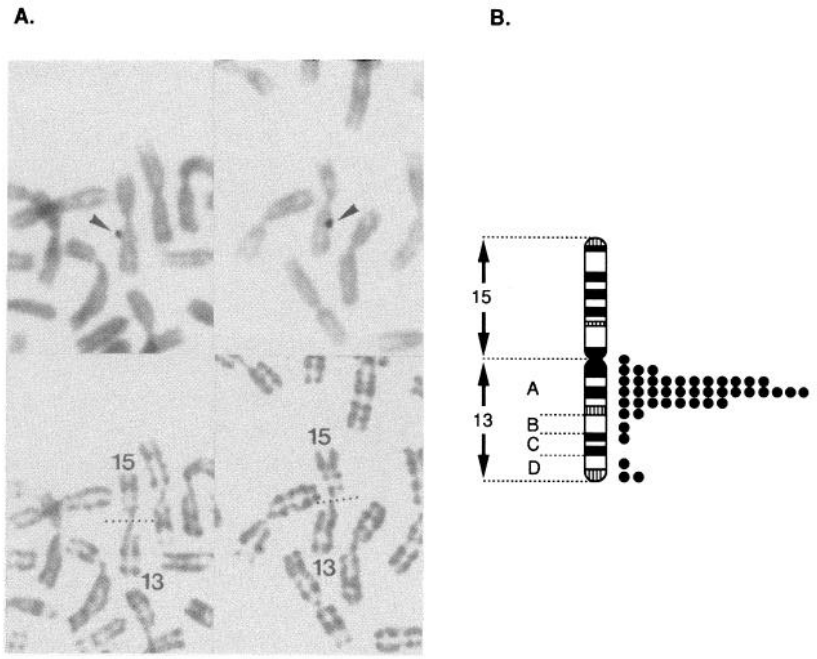

Fig. 9. Localization of the arMEP24 Gene to Mouse Chromosome 13 by *in Situ* Hybridization

A, Two partial WMP mouse metaphases, showing the specific site of hybridization to chromosome 13. *Top*, *Arrowheads* indicate silver grains on Giemsa-stained chromosomes after autoradiography. *Bottom*, Chromosomes with silver grains were subsequently identified by R-banding. B, Diagram of WMP mouse Rb (13, 15) chromosome, indicating the distribution of labeled sites.

when the changes in the steady state level of a particular mRNA under two different conditions are not matched by a comparable change in *de novo* transcription (as measured by nuclear run-on assays), changes in mRNA turnover or stability are often presumed (82). There are two limitations of this approach. First, there is no numerical estimation of half-life, only a relative comparison. Second, no account is made for changes in mRNA maturation events or nuclear transport that could influence steady state mRNA levels.

Increased attention has been paid in the last few years to the regulation of mRNA stability as an important control point of gene expression (reviewed in Refs. 82–84). Steroid hormones represent some of the earliest agents shown to control the degradation of specific mRNAs and have been shown to regulate the stability of a substantial number of mRNAs (83). Vitellogenin mRNA stabilization by estrogen is a clear example of specific regulation of gene expression by control of mRNA stability (83, 85). In mouse kidney, androgens

have little or no effect on the synthesis of three different mRNAs species whose steady state concentrations are increased 10- to 20-fold during testosterone treatment, suggesting that mRNA stabilization is a major factor in the inductions (86). Similarly, androgenic stimulation of mRNA concentrations in the prostate appears to occur predominantly via transcript stabilization (81, 87). In different reports, both 5′- and 3′-terminal sequences have been claimed to play crucial roles in mRNA stability (82, 83). Although stretches of AT-rich sequences were found in the 3′-untranslated portion of arMEP24 mRNA, a carefull analysis of the sequence failed to exhibit a conserved sequence motif with that mediating changes in mRNA turnover (84). The rate of mRNA turnover was also demonstrated to be highly dependent on the presence of the 3′-terminal poly(A) sequences, and the deadenylated mRNA apparently had a much faster rate of turnover (88). Estrogen stabilization of very low density apolipoprotein-II mRNA (89) and glucocorticoid stabilization of GH mRNA (90) result in an increase in the length of their poly(A) tails. Interestingly, expression of the cytoplasmic poly(A) polymerase is positively controlled by androgens (91). Nevertheless, modulation of the poly(A) tail length as a possible mechanism of regulation of mRNA stability does not apply for arMEP24, since no change in mRNA size was seen after induction by androgens (1). Finally, the control of mRNA degradation could also be coupled to ribosome loading and translational engagement of the mature RNA (92). For instance, the maintenance of a high density of ribosomes on vitellogenin mRNA increases its stability (93). Implication of the translational machinery in the regulation of arMEP24 expression was recently suspected (94).

In summary, it seems evident that posttranscriptional events, such as regulation of the mRNA turnover rate, have an important place in the overall scheme of arMEP24 gene expression. Their contribution to the final cellular mRNA concentration remains to be determined by other experiments. The last important question would be to determine whether the posttranscriptional effects of steroid hormones are direct. The effect of the hormone-receptor complexes may be to induce transcription of gene(s) coding for a protein(s) required for stabilization, as previously shown (83, 95).

In synthetic promoter constructs, synergistic activity has been described for the steroid receptors, with factors binding to the Sp1, Oct-1, or CACCC box consensus sequences (96–98). In the arMEP24 gene, several motifs for cell-specific or ubiquitous transcription factors are found. Their implication in the full cell-specific responsiveness of transcription to androgens is suspected. In this respect, the PEA3 motifs are of great interest, since PEA3 protein, which belongs to the *Ets* oncogene family, is specifically expressed in the epididymis and brain, but not in erythroid cell lines or hematopoietic tissues, where all of the other *Ets* oncogenes are synthesized (49). In contrast, GATA-1- and NF-E2-binding proteins are specifically restricted in these latest tissues (52, 53, 56, 57). The presence of several consensus sequences for erythroid-specific transcription factors is quiet surprising. Furthermore, the general organization in the region up-stream of the arMEP24 gene promoter (Fig. 10) is similar to that observed in the GATA-1-binding protein gene (54), the porphobilinogen deaminase gene (56, 59), the erythropoietin receptor gene (99), the α- and β-globin locus control regions (58, 60, 100, 101), and the erythrocyte GSHPx gene (102). Whether these motifs are functional in the epididymal gene is not known. They could be related to the regulatory environment of the ancestral erythroid gene from which the arMEP24 gene is probably derived. Although the two proteins show high homology (6, 7), the conserved regions between GSHPx (103) and arMEP24 are not encoded by identical exons. This may indicate a diverging evolution of the genes that can descend from a single progenitor gene. The red cell GSHPx is not secreted, and the gene does not contain any exon for a signal peptide (103). The exon encoding the N-terminal region of arMEP24, which may be critical for secretion of the peroxidase-like protein in the epididymis, could have been added to a member of the GSHPx gene family after duplication and evolution of an ancestral unit. To support this hypothesis, the arMEP24 gene is located in the [A2-A4] region of the mouse chromosome 13 in a region homologous to human chromosome 7 (104), while the human GSHPx gene is located on chromosome 3 in region 3q11–13.1 (105).

The identification of epididymal-specific regulatory elements from the arMEP24 gene will be extremely important for further analysis of hormone-regulated gene expression of GSHPx and related proteins in the epididymis.

MATERIALS AND METHODS

Isolation of Clones

The mouse BALB/c genomic library was purchased from Clontech (Palo Alto, CA). It was screened by using DNA fragments from the M53 cDNA clone (5) as probe in plaque hybridization, as described previously (1). DNA was prepared from purified clones by the liquid-lysis method (106) and was mapped with restriction endonucleases under conditions recommended by the suppliers (Boehringer Mannheim, Maylan, France; Bethesda Research Laboratories, Cergy Pontoise, France). Suitable restriction fragments were subcloned into pGem7Zf(−) vector (Promega Biotec, Madison, WI), using standard protocols (107).

Nucleotide Sequence Analysis

All of the nt sequences were determined on NaOH-denatured double stranded plasmid DNA by the dideoxy chain termination method (108), using a Pharmacia sequencing kit (St-Quentin en Yvelines, France). For gene sequence determination, each clone was sequenced at least twice on each strand. Sequences were assembled and analyzed using the BISANCE program (109) at C.I.T.I.2 (Paris, France).

Primer Extension Analysis

Conditions for the extraction of total RNA and poly(A)$^+$ RNA purification have been described previously (1, 2). Ten pico-

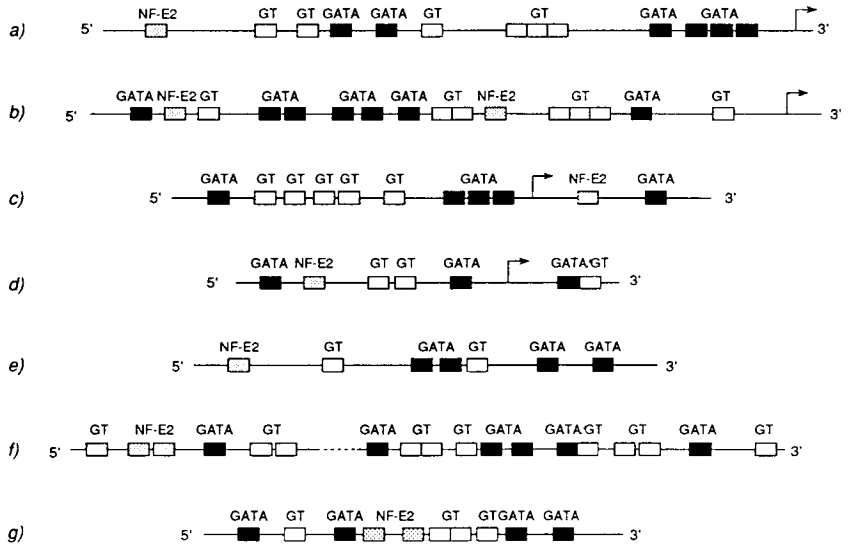

Fig. 10. Comparison of arMEP24 Promoter Region with Erythroid-Specific Regulatory Regions

Schematic representation of promoters or enhancers from arMEP24 gene (a), erythrocyte GSHPx gene (b), GATA-1-binding protein gene (c), porphobilinogen deaminase gene (d), erythropoietin receptor gene (e), β-globin (f), and α-globin (g). Control locus regions are shown (not to scale). The *black boxes* are GATA-1 consensus sequences; the *dotted boxes* are NF-E2-binding sites, and the *open boxes* named GT indicate the presence of GGTGG motifs. The *broken arrows* localize the transcription start sites. Reference numbers are given in the text.

moles of synthetic primer specific for arMEP24 mRNA (5′-CATCTTTTCCGGCCTCCCCC-3′) or specific for luciferase mRNA (5′-GTTTTTGGCGTCTTCC-3′; Eurogentec SA, Liège, Belgium) were radiolabeled using T4 kinase (Promega Biotec) in the presence of 50 μCi [τ-^{32}P]deoxy-ATP (Amersham, Les Ulis, France). One picomole was then hybridized to 10 μg poly(A)$^+$ RNA from caput epididymidis, 10 μg carrier yeast transfer RNA, or 10 μg total RNA from transfected CV-1 cells for 12 h at 35 C in 10 μl 0.04 M 1,4-piperazine diethanesulfonic acid (PIPES), pH 6.4, buffer containing 0.4 M NaCl, 1 μM EDTA, and 80% (vol/vol) formamide. Reverse transcription was performed in 20 μl 50 μM Tris-HCl (pH 8.3), 30 μM KCl, 8 μM $MgCl_2$, 6 μM dithiotreitol, 0.5 μM of each deoxy-NTP, and 50 U AMV reverse transcriptase (Promega Biotec). The samples were incubated at 42 C for 30 min, boosted by the addition of 50 U enzyme, and incubated for 1 h more. The products were analyzed on 8% (wt/vol) polyacrylamide urea gels used for sequence determination.

Construction of Plasmids

The promoterless basis plasmid pLUC, which was used for cloning promoter region fragments from the arMEP24 gene in front of the LUC gene, was derived from pSVOA (38) by inserting a *Sma*I adaptor in the unique *Hin*dIII-cloning site. Constructs pMEP1-LUC and pMEP2-LUC (*Eco*RI-*Spe*I, −167/24), pMEP3-LUC and pMEP4-LUC (*Kpn*I-*Spe*I, −577/24), and pMEP5-LUC and pMEP6-LUC (*Hin*dIII-*Spe*I, −1797/24) were generated by ligation of the appropriate blunt-ended restriction fragments into the *Sma*I site of pLUC. Insertion of the correct fragment and the orientation were checked by restriction enzyme mapping and sequencing. The tk promoter (fragment from −193 to 50) and MMTV-LTR (from −631 to 125) were obtained by polymerase chain amplification, using oligonucleotides primers 5′-CCAAGCTTCATCCCCGTGG-3′ and 5′-CCAAGCTTCGGCACGCTGTTGACGCTGT-3′ or 5′-GCAAGCTTGGCCTAGAAGTAAAAAAGGG-3′ and 5′-GCAAGCTTGGCCGTCCTGAGGGTGACCG-3′, which were directed against previously reported sequences (110, 111). DNA was amplified in 100-μl reactions containing 50 μM KCl; 1.5 μM $MgCl_2$; 10 μM Tris-HCl (pH 8.3); 0.01% gelatin (wt/vol); 0.2 μM each of dATP, dGTP, dCTP, and dTTP; 5 U *Taq* DNA polymerase; 0.5 μM of the appropriate oligonucleotide primers; and 10 ng pMMTV-chloramphenicol acetyltransferase (CAT) (20) or pBLCAT2 (112) plasmid DNA. Reactions were amplified for 25 cycles, each consisting of 1 min at 94 C, 2 min at 60 C, and 3 min at 72 C. Amplified fragments were purified on 2% (wt/vol) agarose gels, restricted with *Hin*dIII, and ligated into the *Hin*dIII site of pSVOA to generate pMMTV-LUC and ptk-LUC. Constructs were verified by DNA sequencing. Constructs pMEP7-tk-LUC and pMEP8-tk-LUC (*Hin*dIII-*Eco*RI, −1797/−167) were generated by ligation of the appropriate blunt-ended restriction fragment into the up-stream filled *Hin*dIII site of ptk-LUC. Two (pARE2-tk-LUC) or three (pARE3-tk-LUC) AREs were cloned in front of the tk promoter in the ptk-LUC vector. Oligonucleotides (positioned −896 to −882 in the arMEP24 gene) 5′-AGCTATATTGTTGAGAGAG AACATGTA-3′ and 5′-AGCTTACATGTTCTCTCTCAACAATAT-3′ were kinased, annealed, and subsequently cloned into the up-stream *Hin*dIII site of ptk-LUC. Copy number and orientation of AREs were checked by sequencing.

Cell Culture

African green monkey kidney (CV-1) cell were maintained in Dulbecco's Modified Essential Medium (DMEM) containing 5% (vol/vol) dextran-charcoal-treated fetal calf serum (Gibco-Bethesda Research Laboratories, Cergy Pontoise, France), 15 μM HEPES (pH 7.2), 2 μM glutamine, 100 U/ml penicillin, and 50 μg/ml streptomycin. They were harvested in 0.25% (wt/vol) trypsin. Cells were grown in a tissue culture incubator at 37 C, with an atmosphere of 5% CO_2 (vol/vol).

Transient DNA Transfection

Supercoiled plasmid DNA used for transfection experiments was purified by two successive CsCl density gradient centrifugations (107). For each vector, two different isolates were transfected at least twice in independent experiments. This approach minimized uncertainties in the interpretation of the

results that might stem from variability between plasmid preparations. In all cases, plasmid pCH110 containing the β-galactosidase gene under the control of the simian virus-40 early promoter was used as an internal control of the efficiency of transfection (40). Typically, 2 μg reporter plasmid, 1 μg pCH110, and 1 μg pSVARo (39), the human AR expression plasmid, were used. The total amount of DNA was increased to 15 μg by adding pGem7Zf(−). CV-1 cells were transiently transfected by the calcium phosphate precipitation method (113) without glycerol shock, using 10^6 cells in each 10-cm plate. The precipitates were left for 12 h on the cells, which were then washed with 5 ml DMEM. Cells were incubated for 24 h in 10 ml fresh medium [DMEM supplemented with glutamine, antibiotics, and 5% (vol/vol) steroid-depleted fetal calf serum] with either vehicle alone or DHT (Theramex Laboratories, Monaco) at final concentrations of 10 nM to 10 μM.

Luciferase Assays

Each 10-cm plate of transfected CV-1 cells was washed once with PBS without Ca^{2+} or Mg^{2+}, and cells were harvested in 1 ml extraction buffer (100 μM potassium phosphate, pH 7.8, and 1 μM dithiothreitol) by scraping. The cells (~5×10^6) from a single dish were pelleted by centrifugation and resuspended in 100 μl extraction buffer. Cells were lysed by three cycles of freezing in liquid nitrogen and thawing at 37 C. Cell debris was pelleted by centrifugation for 5 min. A 30-μl aliquot of each extract was tested for β-galactosidase (40). The remaining 60-μl samples were added to 350 μl 25 μM glycylglycine, pH 7.8, containing 5 μM ATP and 15 μM $MgSO_4$. The samples were placed in an LKB luminometer (LKB, Rockville, MD), and the reaction was initiated by the injection of 100 μl 1 μM luciferin (Sigma, St. Louis, MO). The peak light emission was recorded. Luciferase activities were then standardized according to the internal levels of β-galactosidase. In each experiment, ptk-LUC was transfected three times, and the mean tk-luciferase value was determined. Luciferase data from separate experiments, in which the mean tk-luciferase values were similar, were pooled. Three to 10 independent values were considered for calculation of the mean ± SEM.

Chromosome Preparation

In situ hybridization experiments were carried out using metaphase spreads from a WMP male mouse, in which all autosomes, except 19, were in the form of metacentric robertsonian translocations. Concanavalin-A-stimulated lymphocytes were cultured at 37 C for 72 h, with 5-bromodeoxyuridine added for the final 6 h of culture (60 μg/ml medium), to ensure a chromosomal R-banding of good quality. The 1.4-kb arMEP24 cDNA (5) in pGem7Zf(−) was tritium labeled by nick translation (107) to a specific activity of 1.8×10^8 dpm/μg. The radiolabeled probe was hybridized to metaphase spreads at a final concentration of 25 ng/ml hybridization solution, as previously described (114). After coating with nuclear track emulsion (NTB2, Eastman Kodak, Rochester, NY), the slides were exposed for 13 days at 4 C, then developped. To avoid any slipping of silver grains during the banding procedure, chromosomes spreads were first stained with buffered Giemsa solution, and metaphases were photographed. R-Banding was then performed by the fluorochrome-protolysis-Giemsa method, and metaphases were rephotographed before analysis.

Acknowledgments

The authors thank Dr. E. Pailhoux for many helpful discussions, and S. Guerin for her skilled technical assistance. N.G. is also gratefull to Dr. M. T. Bocquel for providing excellent advice concerning transfections and LUC assays. Plasmid pSVARo was a generous gift from Dr. A. O. Brinkmann (Rotterdam, The Netherlands), pMMTV-CAT was obtained from Prof. P. Chambon (Strasbourg, France), and pBLCAT2 from Dr. J. L. Coudert (Clermont-Ferrand, France).

Received September 17, 1992. Revision received November 30, 1992. Accepted November 30, 1992.

Address requests for reprints to: Jean-Pierre Dufaure, Laboratoire de Biologie Cellulaire, CNRS URA360, 24 avenue des Landais, 63177 Aubière Cedex, France.

Present address: Dr. Norbert B. Ghyselinck, Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, INSERM U184, Institut de Chimie Biologique, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg Cedex, France.

This work was supported by a grant from CNRS (URA 360: Expression et Régulation du Génome) and financial assistance from the Fondation pour la Recherche Medicale Française and INSERM (Contrat 89 4004).

The nt sequence reported in this paper has been submitted to the GenbankTM/EMBL data Bank with accession no. M68896.

REFERENCES

1. Ghyselinck NB, Jimenez C, Lefrançois AM, Dufaure JP 1990 Molecular cloning of a cDNA for androgen-regulated proteins secreted by the mouse epididymis. Mol Endocrinol 4:4–15
2. Faure J, Ghyselinck NB, Jimenez C, Dufaure JP 1991 Specific distribution of mRNA for 24 kDa proteins in the mouse epididymis as revealed by *in situ* hybridization: developmental expression and regulation in the adult. Biol Reprod 44:13–22
3. Jimenez C, Ghyselinck NB, Depeiges A, Dufaure JP 1990 Immunochemical localization and association with spermatozoa of androgen-regulated proteins of Mr 24000 secreted by the mouse epididymis. Biol Cell 6:171–174
4. Jimenez C, Lefrançois AM, Ghyselinck NB, Dufaure JP 1992 Characterization and hormonal regulation of 24kDa proteins synthesis by the murine adult epididymis. J Endocrinol 133:197–203
5. Ghyselinck NB, Dufaure JP 1990 A mouse cDNA sequence for epididymal androgen-regulated proteins related to glutathione peroxidase. Nucleic Acids Res 18:7144
6. Ghyselinck NB, Jimenez C, Dufaure JP 1991 Sequence homology of androgen regulated epididymal proteins with glutathione peroxidase. J Reprod Fertil 93:461–466
7. Ghyselinck NB, Rigaudière N, Dufaure JP 1991 Une protéine androgéno dépendante sécrétée par la tâte de l'épididyme de souris présente des homologies élevées avec différentes glutathion peroxydases. C R Acad Sci [D] (Paris) 313:1–6
8. Flohé L 1982 Glutathione peroxidase brought into focus. In: Pryor WA (ed) Free Radicals in Biology. Academic Press, New York, vol 5:223–253
9. Samuelson B 1983 Leukotrienes: mediators of immediate hypersensitivity reactions and inflammation. Science 220:568–575
10. Alvarez JG, Storey BT 1984 Lipid peroxidation and the reactions of superoxide and hydrogen peroxide in mouse spermatozoa. Biol Reprod 30:833–841
11. Ghyselinck NB, Jimenez C, Courty Y, Dufaure JP 1989 Androgen dependent messenger RNA(s) related to secretory proteins in the mouse epididymis. J Reprod Fertil 85:631–639
12. Chang C, Kokontis J, Liao S 1988 Structural analysis of complementary DNA and amino acid sequence of human and rat androgen receptors. Proc Natl Acad Sci USA 85:7211–7215

13. Lubahn DB, Joseph DR, Sar M, Tan J, Higgs HN, Larson RE, French FS, Wilson EM 1988 The human androgen receptor: complementary deoxyribonucleic acid cloning, sequence analysis and expression in the prostate. Mol Endocrinol 2:1265–1275
14. Trapman J, Klaassen P, Kuiper GGJ M, van der Korput JAG M, Faber PW, van Rooij HCJ, Geurts van Kessel A, Voorhorst MM, Mulder E, Brinkmann AO 1988 Cloning, structure and expression of a cDNA encoding the human androgen receptor. Biochem Biophys Res Commun 153:241–248
15. Evans RM 1988 The steroid and thyroid hormone receptor superfamily. Science 240:889–895
16. Green S, Chambon P 1988 Nuclear receptors enhance our understanding of transcription regulation. Trends Genet 4:309–314
17. Beato M 1989 Gene regulation by steroid hormones. Cell 56:335–344
18. Yamamoto KR 1985 Steroid receptor regulated transcription of specific genes and gene networks. Annu Rev Genet 19:209–252
19. Claessens F, Rushmere NK, Davies P, Celis L, Peeters B, Rombauts WA 1990 Sequence-specific binding of androgen-receptor complexes to prostatic binding protein genes. Mol Cell Endocrinol 74:203–212
20. Cato ACB, Henderson D, Ponta H 1987 The hormone response element of the mouse mammary tumor virus DNA mediates the progestin and androgen induction of transcription in the proviral long terminal repeat region. EMBO J 6:363–368
21. Ham J, Thompson A, Needham M, Webb P, Parker MG 1988 Characterization of response elements for androgens, glucocorticoids and progestins in mouse mammary tumour virus. Nucleic Acids Res 16:5263–5276
22. Otten AD, Sanders MM, McKnight GS 1988 The MMTV LTR promoter is induced by progesterone and dihydrotestosterone but not not by estrogen. Mol Endocrinol 2:143–147
23. Parker MG, White R, Hurst H, Needham M, Tilly R 1983 Prostatic steroid binding protein: isolation and characterization of the C3 genes. J Biol Chem 258:12–15
24. Hurst HC, Parker MG 1983 Rat prostatic steroid binding protein: DNA sequence and transcript maps of the two C3 genes. EMBO J 2:769–774
25. Rushmere NK, Parker MG, Davies P 1987 Androgen receptor-binding regions of an androgen-responsive gene. Mol Cell Endocrinol 51:259–265
26. Claessens F, Celis L, Peeters B, Heyns W, Verhoeven G, Rombauts WA 1988 Functional characterization of an androgen response element in the first intron of the C3(1) gene of prostatic binding protein. Biochem Biophys Res Commun 164:833–840
27. Riegman PHJ, Vliestra RJ, van der Korput JAG M, Brinkmann AO, Trapman J 1991 The promoter of the prostate-specific antigen gene contains a functional androgen responsive element. Mol Endocrinol 5:1921–1930
28. Brooks DE 1987 Developmental expression and androgenic regulation of the messenger RNA for major secretory proteins of the rat epididymis. Mol Cell Endocrinol 53:59–66
29. Breathnach R, Chambon P 1981 Organization and expression of eukaryotic split genes coding for proteins. Annu Rev Biochem 50:349–383
30. Kozak M 1986 Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eucaryotic ribosomes. Cell 44:283–292
31. Gierasch LM 1989 Signal sequences. Biochemistry 28:923–930
32. von Heijne G 1986 A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14:4683–4690
33. Geiduschek EP, Tocchini-Valentini GP 1988 Transcription by RNA polymerase III. Annu Rev Biochem 57:873–914
34. Basler K, Oesch B, Scott M, Westaway D, Wälchli M, Groth DF, McKinley MP, Prusiner SB, Weissmann C 1986 Scrapie and cellular PrP isoforms are encoded by the same chromosomal gene. Cell 46:417–428
35. Ishii S, Imamoto F, Yamanashi Y, Toyoshima K, Yamamoto T 1987 Characterization of the promoter region of c-*erb*B-2 protooncogene. Proc Natl Acad Sci USA 84:4374–4378
36. Ishii S, Xu YH, Stratton RH, Rowe BA, Merlino GT, Pastan I 1985 Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene. Proc Natl Acad Sci USA 82:4920–4924
37. Dorn A, Bollekins J, Staub A, Benoist C, Mathis D 1987 A multiplicity of CCAAT box-binding proteins. Cell 50:863–872
38. de Wet JR, Wood KV, DeLuc M, Helinski DR, Subramani S 1987 Firefly luciferase gene: structure and expression in mammalian cells. Mol Cell Biol 7:725–737
39. Brinkmann AO, Faber PW, van Rooij HCJ, Kuiper GGJ M, Ris C, Klaassen P, van der Korput JAG M, Voorhorst MM, van Laar JH, Mulder E, Trapman J 1989 The human androgen receptor: domain structure, genomic organization and regulation of expression. J Steroid Biochem 34:307–310
40. Herbomel P, Bourachot B, Yaniv M 1984 Two distinct enhancers with different cell specificities coexist in the regulatory region of polyoma. Cell 39:653–662
41. Hamada H, Petrino MG, Kakunaga T 1982 A novel repeated element with Z-DNA forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc Natl Acad Sci USA 79:6465–6469
42. Lancillotti F, Lopex MC, Arias P, Alonso C 1987 Z-DNA in transcriptionnally active chromosomes. Proc Natl Acad Sci USA 84:1560–1564
43. Faisst S, Meyer S 1992 Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res 20:3–26
44. Williams T, Tjian R 1991 Analysis of the DNA-binding and activation properties of the human transcription factor AP-2. Genes Dev 5:670–682
45. Riggs KJ, Merrell KT, Wilson G, Calame K 1991 Common Factor 1 is a transcriptional activator which binds the c-*myc* promoter, the skeletal α-actin promoter, and the immunoglobulin heavy-chain enhancer. Mol Cell Biol 11:1765–1769
46. Okamoto K, Okazawa H, Okuda A, Sakai M, Muramatsu M, Hamada H 1990 A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell 60:461–472
47. Schöler HR, Dressler GR, Balling R, Rohdewohld H, Gruss P 1990 Oct-4: a germline-specific transcription factor mapping to the mouse *t*-complex. EMBO J 9:2185–2195
48. Gutman A, Wasylyk B 1990 The collagenase gene promoter contains a TPA and oncogene-responsive unit encompassing the PEA3 and AP-1 binding sites. EMBO J 9:2241–2246
49. Xin JH, Cowie A, Lachance P, Hassell JA 1992 Molecular cloning and characterization of PEA3, a new member of the *Ets* oncogene family that is differentially expressed in mouse embryonic cells. Genes Dev 6:481–496
50. Wasylyk C, Kerckaert JP, Wasylyk B 1992 A novel modulator domain of Ets transcription factors. Genes Dev 6:965–974
51. Macleod K, Leprince D, Stehelin D 1992 The *Ets* gene family. Trends Biochem Sci 17:251–256
52. Orkin SH 1990 Globin gene regulation and switching: circa 1990. Cell 63:665–672
53. Yamamoto M, Ko LJ, Leonard MW, Beug H, Orkin SH, Engel JD 1990 Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. Genes Dev 4:1650–1662
54. Hannon R, Evans T, Felsenfeld G, Gould H 1991 Struc-

ture and promoter activity of the gene for the erythroid transcription factor GATA-1. Proc Natl Acad Sci USA 88:3004–3008
55. Evans T, Reitman M, Felsenfeld G 1988 An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. Proc Natl Acad Sci USA 85:5976–5980
56. Mignotte V, Wall L, deBoer E, Grosveld F, Romeo PH 1989a Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. Nucleic Acids Res 17:37–54
57. Romeo PH, Prandini MH, Joulin V, Mignotte V, Prenant M, Vainchenker W, Marguerie G, Uzan G 1990 Magakaryocytic and erythrocytic lineages share specific transcription factors. Nature 344:447–449
58. Talbot D, Grosveld F 1991 The 5′ HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. EMBO J 10:1391–1398
59. Mignotte V, Eleouet JF, Raich N, Romeo PH 1989b *Cis*- and *trans*-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. Proc Natl Acad Sci USA 86:6548–6552
60. Philipsen S, Talbot D, Fraser P, Grosveld F 1990 The β-globin dominant control region: hypersensitive site 2. EMBO J 9:2159–2167
61. Fromm M, Berg P 1982 Deletion mapping of DNA regions required for SV40 early region promoter function *in vivo*. J Mol Appl Genet 1:457–481
62. Benoist C, O'Hare K, Breathnach R, Chambon P 1980 The ovalbumin gene-sequence of putative control regions. Nucleic Acids Res 8:127–142
63. Kadonaga JT, Jones KA, Tjian R 1986 Promoter-specific activation of RNA polymerase II transcription by Sp1. Trends Biochem Sci 11:20–23
64. Vigneron M, Barrera-Saldana HA, Bat D, Everett RE, Chambon P 1984 Effect of the 21-bp repeat upstream element on *in vitro* transcription from the early and late SV40 promoters. EMBO J 3:2373–2382
65. McKnight SL, Kingsbury RC, Spence A, Smith M 1984 The distal transcription signals of the herpes virus tk gene share a common hexanucleotide control sequence. Cell 37:253–262
66. Valerio D, Duyvesten MGC, Dekker BMM, Weeda G, Berkvens TM, van der Voorn L, van Ormondt H, van der Eb AJ 1985 Adenosine deaminase: characterization of a gene with a remarkable promoter. EMBO J 4:437–443
67. Patel PI, Framson PE, Caskey CT, Chinault AC 1986 Fine structure of the human hypoxanthine phosphoribosyltransferase gene. Mol Cell Biol 6:393–403
68. Masters JN, Attardi G 1985 Discrete human dihydrofolate reductase gene transcripts present in polysomal RNA map with their 5′ ends several hundred nucleotides upstream of the main mRNA start site. Mol Cell Biol 5:493–500
69. Mitchell PJ, Carothers AM, Han JH, Harding JD, Kas E, Venolia L, Chasin LA 1986 Multiple transcription start sites. DNase I-hypersensitive sites, and an opposite strand exon in the 5′ region of the CHO dhfr gene. Mol Cell Biol 6:425–440
70. Smale ST, Baltimore D 1989 The 'initiator' as a transcription control element. Cell 57:103–113
71. Roeder RG 1991 The complexities od eukaryotic transcription initiation: regulation of preinitiation complex assembly. Trends Biochem Sci 16:402–408
72. Miesfeld R, Krystal M, Arnheim N 1981 A member of a new repeated sequence family which is conserved throughout eucaryotic evolution is found between the human d et b globin genes. Nucleic Acids Res 9:5931–5947
73. Rogers J 1983 CACA sequences–the ends and the means? Nature 305:101–102
74. Schon E, Evans T, Welsh J, Efstratiadis A 1983 Conformation of promoter DNA: fine mapping od S1-hypersensitive sites. Cell 35:837–848
75. Rich A, Nordheim A, Wang A 1984 The chemistry and biology of left-handed Z-DNA. Annu Rev Biochem 53:791–846
76. Delaey B, Dirckx L, Decourt JL, Claessens F, Peeters B, Rombauts WA 1987 Rat prostatic binding protein: the complete sequence of the C2 gene and its flanking regions. Nucleic Acids Res 15:1627–1641
77. Harris SE, Mansson PE, Jully DB, Burkhart B 1983 Seminal vesicle secretion IV gene: allelic difference due to a series of 20-base-pair direct tandem repeats within an intron. Proc Natl Acad Sci USA 80:6460–6464
78. Williams L, McDonald C, Higgins S 1985 Sequence organization of rat seminal vesicle F gene: location of transcriptional start point and sequence comparison with six other androgen regulated genes. Nucleic Acids Res 13:659–672
79. Panthier JJ, Dreyfus M, Tronik-LeRouse D, Rougeon F 1984 Mouse kidney and submaxillary gland renin genes differ in their 5′ putative regulatory sequences. Proc Natl Acad Sci USA 81:5489–5493
80. Joseph DR, Hall SH, Conti M, French FS 1988 The gene structure of rat androgen-binding protein: identification of potential regulatory DNA elements of a FSH-regulated protein. Mol Endocrinol 2:3–13
81. Wolf DA, Schulz P, Fittler F 1992 Transcriptional regulation of prostate kallikrein-like genes by androgen. Mol Endocrinol 6:753–762
82. Atwater JA, Wisdom R, Verma IM 1990 Regulated mRNA stability. Annu Rev Genet 24:519–541
83. Nielsen DA, Shapiro DJ 1990 Insights into hormonal control of messenger RNA stability. Mol Endocrinol 4:953–957
84. Raghow R 1987 Regulation of messenger RNA turnover in eukaryotes. Trends Biol Sci 12:358–360
85. Brock ML, Shapiro DJ 1983 Estrogen stabizes vitellogenin mRNA against cytoplasmic degradation. Cell 34:207–214
86. Berger FG, Loose D, Meisner H, Watson G 1986 Androgen induction of messenger RNA concentrations in mouse kidney is post-transcriptional. Biochemistry 25:1170–1175
87. Page MJ, Parker MG 1982 Effects of androgens on the transcription of rat prostatic binding protein gens. Mol Cell Endocrinol 27:343–355
88. Brawerman G 1981 The role of poly(A) sequence in mammalian messenger RNA. CRC Crit Rev Biochem 10:1–38
89. Cochrane AW, Deely RG 1988 Estrogen-dependent activation of the avian very low density apolipoprotein II and vitellogenin genes. transient alterations in mRNA polyadenylation and stability during induction. J Mol Biol 203:555–567
90. Paek I, Axel R 1987 Glucocorticoids enhance stability of growth hormone mRNA. Mol Cell Biol 7:1496–1507
91. Raju VS, Reddy PRK 1984 Androgen regulation of poly(A) polymerase in the ventral prostate of rat. Steroids 43:407–414
92. Yen TJ, Gay DA, Pachter JS, Cleveland DW 1988 Autoregulated changes in stability of polyribosome-bound β-tubulin mRNAs are specified by the first 13 translated nucleotides. Mol Cell Biol 8:1224–1235
93. Blume JE, Shapiro DJ 1989 Ribosome loading, but not protein synthesis, is required for estrogen stabilization in *Xenopus laevis* vitellogenin mRNA. Nucleic Acids Res 17:9003–9014
94. Rigaudière N, Ghyselinck NB, Faure J, Dufaure JP 1992 Regulation of the epididymal glutathione peroxidase-like protein in the mouse: dependence upon androgens and testicular factors. Mol Cell Endocrinol 89:67–77
95. Vannice JL, Taylor JM, Ringold GM 1984 Glucocorticoid-mediated induction of α1-acid glycoprotein: evidence for

hormone-regulated RNA processing. Proc Natl Acad Sci USA 81:4241–4245
96. Schüle R, Muller M, Kaltschmidt C, Renkawitz R 1988 Many transcription factors interact synergistically with steroid receptors. Science 242:1418–1420
97. Strähle U, Schmid W, Schutz G 1988 Synergistic action of the glucocorticoid receptor with transcription factors. EMBO J 7:3389–3395
98. Allan GF, Ing NH, Tsai SY, Srinivasan G, Weigel NL, Thompson EB, Tsai MJ, O'Malley BW 1991 Synergism between steroid response and promoter elements during cell-free translation. J Biol Chem 266:5905–5910
99. Heberlein C, Fischer KD, Stoffel M, Nowock J, Ford A, Tessmer U, Stocking C 1992 The gene for erythropoietin receptor is expressed in multipotential hematopoietic and embryonal stem cells: evidence for differentiation stage-specific regulation. Mol Cell Biol 12:1815–1826
100. Reddy PMS, Shen CKJ 1991 Protein-DNA interactions *in vivo* of an erythroid-specific, human β-globin locus enhancer. Proc Natl Acad Sci USA 88:8676–8680
101. Strauss EC, Andrews NC, Higgs DR, Orkin SH 1992 *in vivo* footprinting of the human α-globin locus upstream regulatory element by guanine and adenine ligation-mediated polymerase chain reaction. Mol Cell Biol 12:2135–2142
102. Moscow JA, Morrow CS, He R, Mullenbach GT, Cowan KH 1992 Structure and function of the 5′-flanking sequence of the human cytosolic selenium-dependent glutathione peroxidase gene (*hgpx1*). J Biol Chem 267:5949–5958
103. Chambers I, Frampton J, Goldfarb P, Affara N, McBain W, Harrison PR 1986 The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination' codon, TGA. EMBO J 5:1221–1227
104. Justice MJ, Stephenson DA 1991 Mouse chromosome 13. Mammalian Genome 1:S205–220
105. Chada S, LeBeau MM, Casey L, Newburger PE 1990 Isolation and chromosomal localization of the human glutathione peroxidase gene. Genomics 6:268–271
106. Grossberger D 1987 Minipreps of DNA from bacteriophage lambda. Nucleic Acids Res 15:6737
107. Sambrook KJ, Fritsch EF, Maniatis T 1989 Molecular Cloning–A Laboratory Manual, ed 2. Cold Spring Harbor Laboratory, Cold Spring Harbor
108. Sanger P, Nicklen S, Coulson AR 1977 DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467
109. Dessen P, Fondrat C, Valencien C, Mugnier C 1990 BISANCE: a french service for access to biomolecular sequence databases. Comp Appl Biosci 6:355–356
110. McKnight 1980 The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. Nucleic Acid Res 8:5949–5964
111. Fasel N, Pearson K, Buetti E, Diggelmann H 1982 The region of mouse mammary tumor virus DNA containing the long terminal repeat includes a long coding sequence and signals for hormonally regulated transcription. EMBO J 1:3–7
112. Luckow B, Schütz G 1987 Cat constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. Nucleic Acids Res 15:5490
113. Chen C, Okayama H 1987 High efficiency transformation of mammalian cells by plasmid DNA. Mol Cell Biol 7:2745–2752
114. Mattei MG, Philip N, Passage E, Moisan JP, Mandel JL, Mattei JF 1985 DNA probe localization at 18p113 band by *in situ* hybridization and identification of a small supernumerary chromosome. Hum Genet 69:268–271