

The mouse androgen receptor

Functional analysis of the protein and characterization of the gene

Peter W. FABER,* Audrey KING,* Henri C. J. VAN ROOIJ,* Albert O. BRINKMANN,† Nico J. DE BOTH* and Jan TRAPMAN*

Departments of *Pathology and †Biochemistry II, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

Screening a mouse genomic DNA library with human androgen-receptor (hAR) cDNA probes resulted in the isolation and characterization of eight genomic fragments that contain the eight exons of the mouse androgen-receptor (mAR) gene. On the basis of similarity to the hAR gene, the nucleotide sequences of the protein-coding parts of the exons as well as the sequences of the intron/exon boundaries were determined. An open reading frame (ORF) of 2697 nucleotides, which can encode an 899-amino-acid protein, could be predicted. The structure of the mAR ORF was confirmed by sequence analysis of mAR cDNA fragments, which were obtained by PCR amplification of mouse testis cDNA, using mAR specific primers. A eukaryotic mAR expression vector was constructed and mAR was transiently expressed in COS-1 cells. The expressed protein was shown by Western blotting to be identical in size with the native mAR. Co-transfection of HeLa cells with the mAR expression plasmid and an androgen-responsive chloramphenicol acetyltransferase (CAT) reporter-gene construct showed mAR to be able to *trans*-activate the androgen-responsive promoter in a ligand-dependent manner. Transcription-initiation sites of the mAR gene were identified by S1-nuclease protection experiments, and the functional activity of the promoter region was determined by transient expression of mAR promoter–CAT–reporter-gene constructs in HeLa cells. Structural analysis revealed the promoter of the mAR gene to be devoid of TATA/CCAAT elements. In addition, the promoter region is not remarkably (G+C)-rich. Potential promoter elements consist of a consensus Sp1 binding sequence and a homopurine stretch. The polyadenylation sites of mAR mRNA were identified by sequence similarity to the corresponding sites in the hAR mRNA.

INTRODUCTION

The androgen receptor (AR) is a transcription-regulating protein that plays a pivotal role in the programming of male sexual differentiation and development. Absence or mutation of the X-chromosome-located AR gene can lead to complete androgen-insensitivity, and the affected 46,XY individual displays the external phenotype of a female [1–7]. Mutations in AR are also thought to be involved in less severe forms of aberrant male sexual development.

Structurally and functionally AR belongs to the superfamily of ligand-responsive transcription modifiers which encompasses the receptors for the steroid and thyroid hormones, retinoic acid, vitamin D₃ and several 'orphan' receptors for which a ligand has not as yet been identified [8–11]. The structural regions that these receptors have in common, namely a C-terminal ligand-binding domain, an internal DNA-binding domain, consisting of two Cys–Cys zinc finger motifs and an N-terminal hypervariable 'regulatory' domain, enable them to react to a hormonal stimulus by modulation of gene transcription through recognition of, and binding to, hormone-responsive elements (HREs) located in the control regions of target genes.

The isolation of human AR (hAR) and rat AR (rAR) cDNAs [12–17] and the elucidation of the structural organization of the hAR gene [18–20a] has provided important tools to address questions regarding AR function in the androgen signal-transduction pathway. In common with the glucocorticoid receptor and the progesterone receptor, the androgen receptor is able to regulate expression of a mouse-mammary-tumour-virus long-

terminal-repeat (MMTV LTR)-driven promoter, although perhaps less efficiently [21–23]. This indicates that the DNA motif recognized by the AR is identical with, or closely related to, the GRE and PRE consensus sequence GGTACAnnnTGTTCT.

In order to extend findings to different species, it is necessary to isolate species-specific tools. Especially for detailed studies *in vivo* concerning molecular and genetic mechanisms of AR synthesis and function, including transgenic animals, and study of the physiological effects of manipulation of the AR system, knowledge of the mouse AR (mAR) system is of high importance.

In the present study we describe the characterization of the mAR gene and cDNA. A mAR expression vector (pmAR⁰) was constructed and applied for the functional characterization of the mAR protein. In addition, the promoter of the mAR gene is structurally and functionally characterized.

MATERIALS AND METHODS

Isolation and characterization of mAR genomic DNA clones

A mouse genomic DNA library in lambda EMBL3 SP6/T7 was purchased from Clontech (Palo Alto, CA, U.S.A.). The library was plated, transferred to nitrocellulose filters (Schleicher and Schüll, Dassel, Germany) and hybridized to ³²P-labelled hAR cDNA probes (see Fig. 2a below), according to standard procedures [24,25]. Positive phages were isolated by three rounds of purification. All experiments concerning the isolation, sub-cloning and characterization of phage DNA inserts by restriction mapping and Southern-blot hybridization were performed

Abbreviations used: hAR, mAR and rAR: human, mouse and rat androgen receptor; ORF, open reading frame; CAT, chloramphenicol acetyltransferase; HRE, hormone-responsive element; MMTV, mouse mammary-tumour virus; MCS, multiple cloning site; MEM, minimal essential medium; GRE/PRE, glucocorticoid/progesterone-responsive element; cPR, chicken progesterone receptor; hER, human oestrogen receptor; LTR, long terminal repeat; UTR, untranslated region; SV40, simian virus 40; PMSF, phenylmethanesulphonyl fluoride; DTT, dithiothreitol.

These sequence data have been submitted to the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession nos. X59590 [mouse (1A) androgen-receptor cDNA], X59591 [mouse (5'-MAR) androgen-receptor cDNA and X59592 [mouse (5'-MAR, 15A, 15B, 18A, 19A, 6B, 3A, 1A) androgen-receptor cDNA].

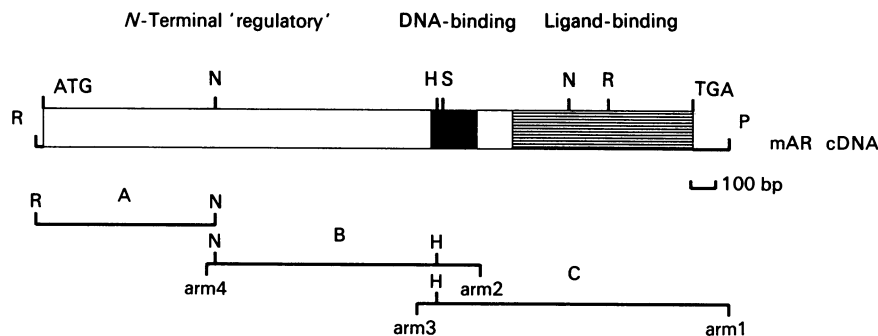


Fig. 1. Partial restriction map of mAR cDNA

Indicated are the AR domain structure as well as the three DNA fragments from which the cDNA was constructed. Key to restriction endonucleases: H, *HindIII*; N, *NcoI*; P, *PstI*; R, *EcoRI*; S, *SstI*. Arm1–Arm4 are mAR-specific primers (see the text and Fig. 3).

according to standard procedures [24,25]. Sequence analysis was done by the dideoxy-chain-termination method on both DNA strands using single-stranded M13mp18/19 DNA preparations [26] and T7 polymerase (Pharmacia, Uppsala, Sweden).

RNA preparation and S1-nuclease-protection experiments

Total cellular RNA was isolated by the guanidinium thiocyanate method [27]. For the S1-nuclease-protection assay a 0.6 kb *EcoRI–HindIII* double-stranded DNA fragment (see Fig. 5a below) was used as a probe. The fragment was end-labelled using T4 polynucleotide kinase (Gibco BRL, Grand Island, NY, U.S.A.), and approx. 1×10^5 c.p.m. of probe were annealed to 40 μ g of RNA in 80% formamide/40 mM-Pipes (pH 6.5)/0.4 M-NaCl/1 mM-EDTA overnight at 55 °C [28]. S1-nuclease (Boehringer, Mannheim, Germany) digestions were carried out for 1 h at 37 °C, and the resulting protected fragments were analysed on a denaturing 6% (w/v) polyacrylamide gel.

DNA amplification

Oligonucleotides were synthesized on an Applied Biosystems 381A DNA synthesizer. Amplification by the PCR [29] was performed in 100 μ l reaction mixtures containing 2% of a first-strand cDNA preparation, using 2 units of *Taq* DNA polymerase (Promega, Madison, WI, U.S.A.) under the conditions described by the manufacturer. For first-strand cDNA preparations, 100 ng of the appropriate primer was annealed to 5 μ g of total RNA, and cDNA synthesis was performed with avian-myeloblastosis-virus reverse transcriptase according to a standard protocol (Promega). Amplification was performed in a Bioexcellence DNA incubator for 30 cycles. Standard conditions were: denaturation for 1 min at 95 °C, annealing for 2 min at 60 °C and extension for 1–5 min at 70 °C. The amplification products were recovered after chloroform/3-methylbutan-1-ol extraction, electrophoresed through 1–1.5%-(w/v)-agarose gels and blotted on nitrocellulose filters for hybridization. For isolation and subcloning, the fragments were extracted with phenol/chloroform and ethanol-precipitated. The oligonucleotides used (see Fig. 3) were:

- arm1: 5'-CAGAGAAGTAGTGCAGAGTT-3'
(anti-sense, 3309–3328)
- arm2: 5'-CAGAGTCATCCCTGCTTC-3'
(anti-sense, 2062–2079)
- arm3: 5'-TTTGGACAGTACCAGGGACC-3'
(sense, 1747–1766)
- arm4: 5'-AGTGCCAAGGAGTTGTGTA-3'
(sense, 802–821)

Chloramphenicol acetyltransferase (CAT) reporter gene constructs for analysis of promoter activity

The 0.5 kb *PstI–HindIII* and the 1.3 kb *PstI–PstI* genomic DNA fragments were isolated from a subcloned 1.5 kb *EcoRI–EcoRI* fragment (see Fig. 5a below). The *PstI–PstI* fragment was cloned in both orientations in the *PstI* site of the multiple cloning site (MCS) of the promoterless pCAT-Enhancer plasmid (Promega; referred to as 'pmAR-CAT-0'). This resulted in the construction of pmAR-CAT-1 and pmAR-CAT-1 rev (reverse) respectively. The 0.5 kb *PstI–HindIII* (blunt-ended) fragment was cloned in a *PstI–AccI* (blunt-ended) pCAT-Enhancer vector, and in pCAT-Enhancer (*PstI–HindIII*), resulting in the construction of pmAR-CAT-2 and pmAR-CAT-2 rev, which contain the fragment in the normal and reverse orientation respectively.

Construction of the mAR expression plasmid pmAR⁰

Using genomic DNA fragments, and cDNA fragments obtained by PCR amplification a mAR expression plasmid, containing the complete mAR ORF, was constructed. A partial restriction map of the mAR cDNA, together with the three fragments from which it was derived, is presented in Fig. 1. Fragment A (*EcoRI–NcoI*) was derived from the 2 kb *EcoRI–EcoRI* genomic DNA fragment that contains the protein-coding region of exon 1 (see Fig. 5a below), whereas the fragments B (*NcoI–HindIII*) and C [*HindIII*-3'-untranslated region (UTR)] originated from reverse-transcriptase PCR using a first-strand cDNA preparation made with primer arm1 (Fig. 3 below) and mouse testis RNA. The fragments were produced by PCR with the primer combinations arm1 and arm3 (C) and arm2 and arm4 (B) respectively. The cDNA construct was prepared in a pGEM vector and completely sequenced. A mAR expression vector was constructed using the cDNA containing the complete mAR ORF, the simian-virus-40 (SV40) early promoter and the rabbit β -globin polyadenylation signal [21]. The resulting vector is referred to as 'pmAR⁰'.

Transfection assay

A day before transfection 3×10^5 HeLa cells were seeded into 6 cm-diameter dishes. Culture medium was Dulbecco's MEM supplemented with 5% (v/v) fetal-calf serum and antibiotics. The approx. 40%-confluent cell cultures were transfected by the calcium phosphate method [30]. For *trans*-activating studies 2.5 μ g of pmAR⁰ was used in combination with 2.5 μ g of reporter plasmid (pG 29 G-TK-CAT/pBL2-CAT) [31,32], 2.5 μ g of pCH110 (β -galactosidase expression plasmid; Pharmacia) and 2.5 μ g of carrier DNA (pTZ; Pharmacia). For promoter studies

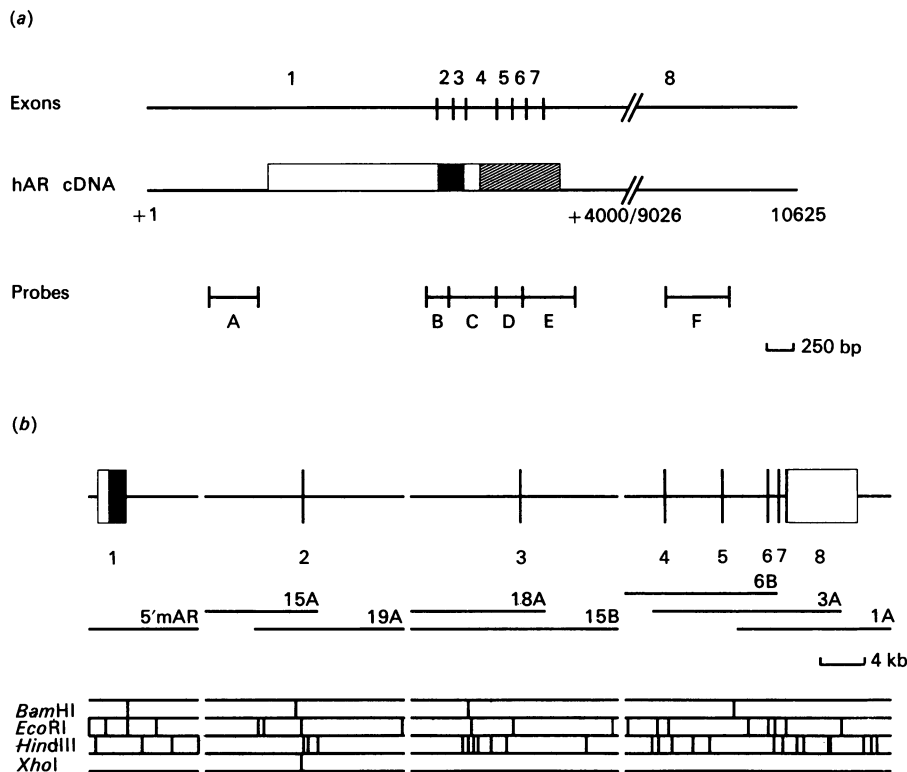


Fig. 2. Molecular cloning and characterization of the mAR gene

(a) Probes used in the screening of the mouse genomic DNA library. Depicted is a schematic representation of the full-length hAR cDNA (10625 nucleotides). The ORF is shown as a box with the black and hatched areas representing the DNA- and ligand-binding domains respectively. The distribution of the hAR cDNA sequence over the eight exons of the hAR gene is indicated above the hAR cDNA. The hybridization probes used (A–F) are indicated under the cDNA. (b) Structural organization of the mAR gene. The eight exons are numbered 1–8. The positions of the eight individual clones (5'mAR, 15A, 19A, 18A, 15B, 6B, 3A and 1A) are shown as horizontal bars. A restriction map of the genomic clones after digestion with *Bam*HI, *Eco*RI, *Hind*III and *Xho*I is presented at the bottom.

5 µg of the appropriate mAR promoter construct was used together with 2.5 µg of pCH110 and 2.5 µg of pTZ. pSV2-CAT (5 µg) was used in control experiments. All experiments were carried out in duplicate with at least two different plasmid preparations. Cell extracts were prepared at 48 h after transfection, and CAT assays were performed essentially as described in [33]. For quantification, butyryl-CoA was used, which allows the direct liquid-scintillation counting of the butyrylated chloramphenicol [34]. For immunoprecipitation and Western blotting, 2 × 10⁶ COS-1 cells were transfected using 10 µg of pmAR⁰ and 10 µg of pTZ.

Immunoprecipitation and Western-blot analysis

COS-1 cells, transfected with the mAR and hAR expression plasmids pmAR⁰ and pAR⁰ [21], mock-transfected COS-1 cells and mouse testicular tissue were lysed in 40 mM-Tris (pH 7.4)/1 mM-EDTA/10 % (v/v) glycerol/10 mM-dithiothreitol (DTT)/50 mM-NaF/0.6 mM-phenylmethanesulphonyl fluoride (PMSF)/0.1 mM-bacitracin/0.5 mM-leupeptin/1 % (v/v) Triton/0.5 % (w/v) sodium deoxycholate/0.08 % (w/v) SDS. For the immunoprecipitation the anti-hAR monoclonal antibody F39.4.1 [35] was used. The non-specific control monoclonal antibody was ER-Pr 27, which recognizes prostate-specific antigen [36]. Western blotting and immunostaining with the anti-hAR polyclonal antibody Sp061 (diluted 1:1000) were done as described in [37].

RESULTS

Molecular cloning and characterization of the mAR gene

To isolate the mAR gene, a mouse genomic DNA library was screened using hAR cDNA probes. Previously we elucidated the organization of the hAR gene and showed that the hAR gene is composed of eight exons [18, 20a]. In Fig. 2(a) the hAR cDNA is schematically depicted, together with the cDNA probes used and the distribution of the hAR cDNA sequence over the exons of the hAR gene. The probes cover the complete protein-coding regions of the hAR gene exons 1–8. Screening of the mouse genomic DNA library resulted in the isolation of eight clones (Fig. 2b). Probes A and B hybridized to clone 5'mAR; clones 15A and 19A were both recognized by probes B and C; probe C hybridized to clones 15B and 18A; clones 6B and 3A could be identified by probes C, D and E; probes D, E and F recognized clone 1A. A physical map of the mAR gene was constructed on the basis of restriction-enzyme-digestion patterns combined with sequence analysis of the individual exons (see Fig. 2b). For this sequence analysis, mouse genomic DNA fragments hybridizing with hAR cDNA probes were subcloned in plasmid vectors and mapped in more detail. Small fragments containing the exons were subsequently cloned in M13 vectors and sequenced. Fig. 2(b) shows the positions of the eight mAR exons and the restriction map for *Bam*HI, *Eco*RI, *Hind*III and *Xho*I. The sequence obtained is presented in Fig. 3 and is discussed in detail below. Additional hybridization experiments using probes

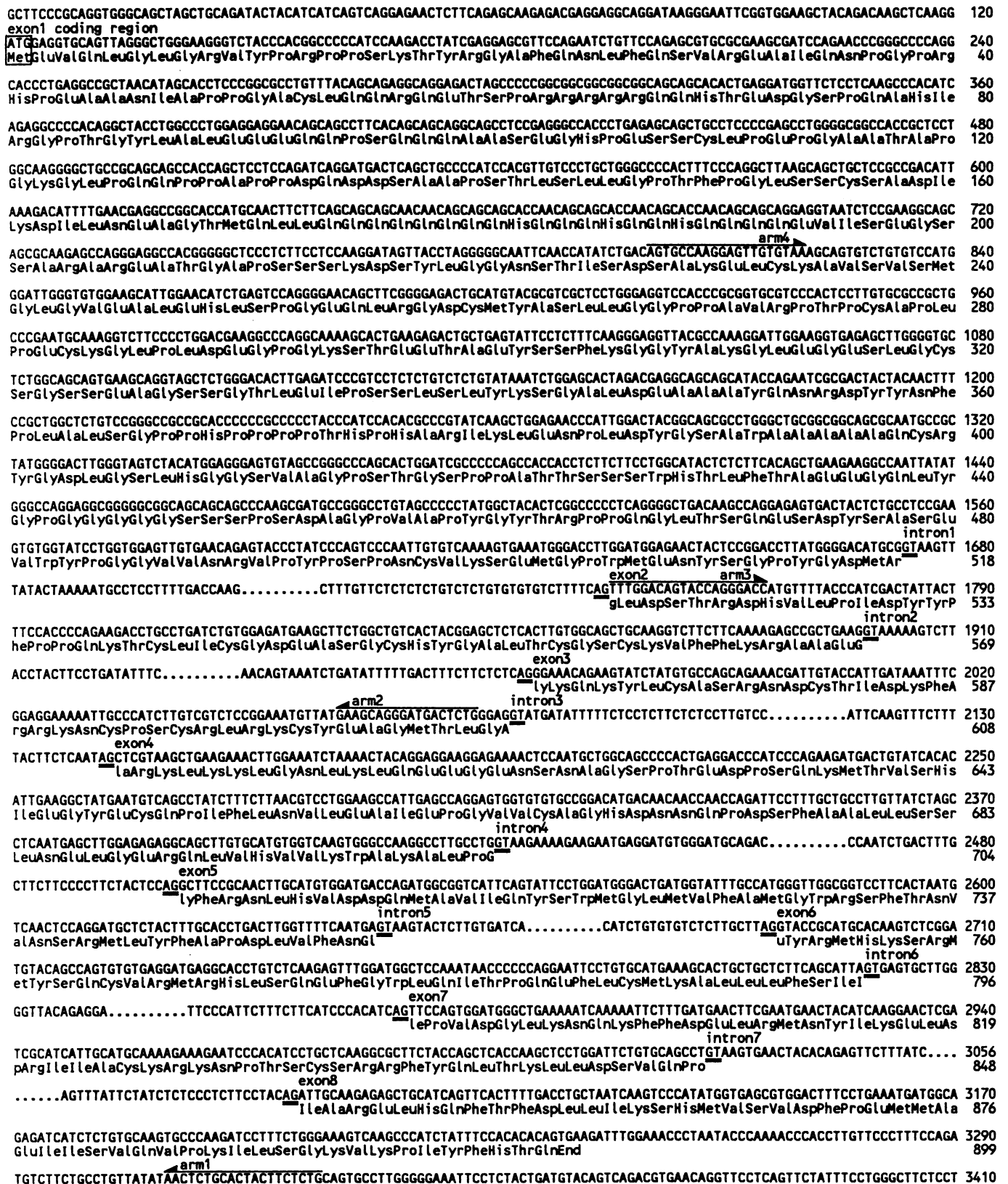


Fig. 3. Sequence analysis of the mAR gene exons

Shown is the sequence of the mAR ORF as deduced from the genomic DNA clones, together with the sequences of the intron/exon boundaries and part of the sequence of the 5'-UTR and 3'-UTR which are found in the exons 1 and 8 respectively. The GT/AG dinucleotides at the splice donor and acceptor sites are doubly underlined. The ATG translation-initiation codon is boxed, and the TGA translation-termination codon is marked 'End'. The number of each individual exon is given after the AG splice acceptor sequence; the number of the individual intron is given above the GT splice donor site. The positions of the four primers used in the construction of the full-length mAR cDNA (arm1-arm4) are indicated by horizontal arrows.

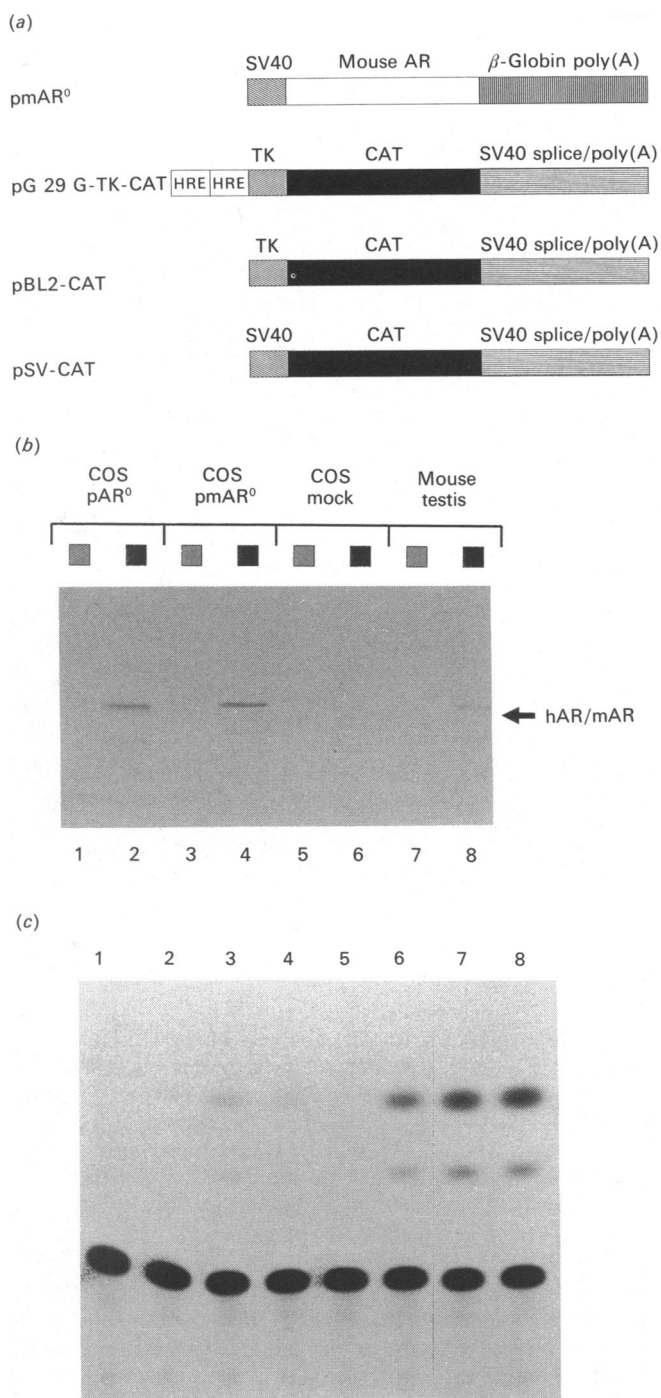


Fig. 4. Functional analysis of the mAR

(a) Constructs used for the functional assay of mAR activity. Shown are the mAR expression plasmid pmAR⁰, the androgen-responsive CAT reporter-gene construct pG 29 G-TK-CAT and the two control constructs, pBL2-CAT and pSV2-CAT. (b) Determination of the size of mAR. Shown is the Western-blot analysis of the mAR and hAR as obtained by immunoprecipitation from COS-1 cells and mouse testicular tissue. Immunoprecipitations were performed with specific (■) and non-specific monoclonal antibodies (□) using whole-cell extracts from pmAR⁰-transfected COS-1 cells (lanes 3 and 4), from pAR⁰-transfected COS-1 cells (hAR) (lanes 1 and 2), from mock-transfected COS-1 cells (lanes 5 and 6) and mouse testicular tissue (lanes 7 and 8). AR was revealed with an anti-hAR polyclonal antibody. The arrow indicates the position of AR. Molecular-mass markers were β-galactosidase (116 kDa) and phosphorylase b (97 kDa). (c) Determination of the *trans*-activating activity of mAR by co-transfection of HeLa cells with the androgen-responsive CAT reporter-gene construct pG 29 G-TK-CAT and

specific for the promoter region of the hAR gene and the polyadenylation sites of the hAR mRNA indicated their presence in clones 5' mAR and 1A, respectively, and they are depicted as such in Fig. 2(b).

The eight overlapping clones cover a 60 kb region of genomic DNA. Although all the exons of the mAR gene are present in the clones isolated, part of the information for the introns 1, 2 and 3 is lacking. The lengths of the introns as deduced from these clones are >14.5 kb (intron 1), >19.1 kb (2), >12.4 kb (3), 5.2 kb (4), 4.1 kb (5), 1.0 kb (6) and 0.7 kb (7) respectively. The physical map of the mAR gene was checked by Southern-blot analysis of mouse genomic DNA digested with the restriction enzymes *Bam*HI, *Eco*RI and *Hind*III. Blots were probed with genomic probes spanning all exons of the mAR gene. These experiments confirmed the restriction map presented in Fig. 2(b) (results not shown). In addition, the sizes of the various hybridizing fragments detected in these genomic blots indicated that the length of intron 3 exceeds 20 kb. This finding extends the minimum length of the mAR gene to at least 68 kb.

Sequence analysis of the mAR open reading frame (ORF)

As discussed above, small fragments containing exon information were sequenced to predict the structure of the mAR cDNA ORF. Protein-coding sequences as well as intron/exon boundaries were deduced on the basis of similarity to the hAR ORF and splice consensus sequences respectively. This resulted in a predicted ORF for the mAR cDNA shown in Fig. 3, together with the sequences of the intron/exon boundaries, which all conform to the GT/AG rule (doubly underlined in Fig. 3). The mAR ORF starts at the boxed ATG translation initiation codon in exon 1 and continues in exons 2–7 until encountering a termination codon in exon 8. The length of the ORF is 2697 nucleotides, which can encode a protein of 899 amino acids. All domains that characterize the steroid-receptor family are present in this ORF. A large *N*-terminal domain is followed by the DNA-binding domain consisting of the two Cys–Cys zinc fingers and the ligand-binding domain at the *C*-terminus of the protein. The *N*-terminal domain is encoded by exon 1, the information for the first and second zinc finger motif is present in the exons 2 and 3 respectively, whereas the ligand-binding domain is encoded by the exons 4–8. The 2697 nucleotides of the mAR ORF are distributed over the exons as follows: exon 1, 1553; exon 2, 152; exon 3, 117; exon 4, 288; exon 5, 145; exon 6, 131; exon 7, 158; and exon 8, 153 nucleotides respectively.

Functional analysis of mAR expressed from the molecular cloned cDNA

A mAR cDNA containing the complete ORF was constructed using conventional cloning methods combined with the PCR techniques described in the Materials and methods section. After insertion of the mAR cDNA in a pGEM vector, the 2.8 kb insert was completely sequenced and the sequence was matched with that of the deduced mAR genomic DNA structure. The sequence showed the mAR ORF to be identical with the predicted ORF as already presented in Fig. 3, and thereby confirmed the positions of the splice sites within the mAR gene (clones containing PCR artefacts were omitted). To test the functional properties of the cloned mAR, the cDNA was inserted in an expression vector

pmAR⁰ (lanes 5 and 6). Controls include pG 29 G-TK-CAT (lanes 1 and 2), a co-transfection of pmAR⁰ and pBL2-CAT (lanes 3 and 4) and pSV2-CAT (lanes 7 and 8). Cells were grown in the presence (lanes 2, 4, 6, 8) or absence (lanes 1, 3, 5, 7) of the synthetic steroid R1881. The autoradiograph displays the conversion of [¹⁴C]chloramphenicol into acetylated products.

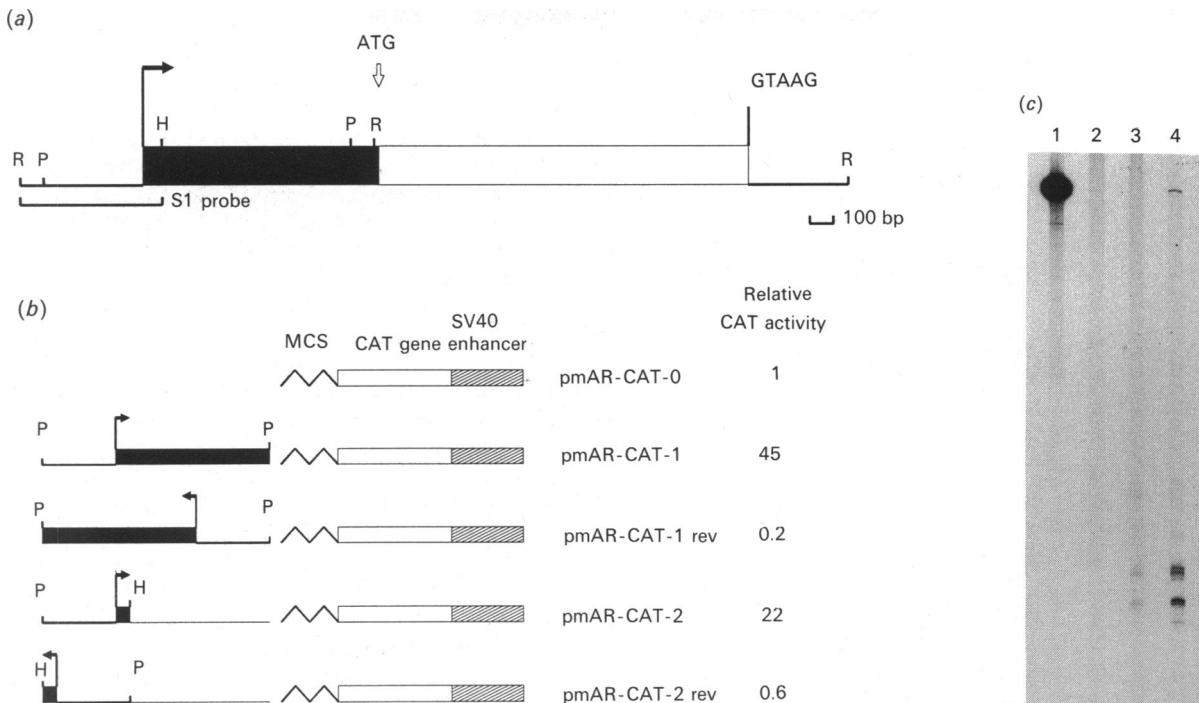


Fig. 5. Structural and functional analysis of the promoter region of the mAR gene

(a) Partial restriction map of the genomic region surrounding exon 1 of the mAR gene. Indicated are the ATG translation-initiation codon and the splice donor site GTAAG. The position of the double-stranded DNA probe *EcoRI*-*HindIII* that was used in the S1-nuclease-protection experiment is indicated by a horizontal bar. (b) Compositions of the four mAR gene promoter-CAT constructs, as well as of the plasmid pmAR-CAT-0, are shown. The four mAR promoter-CAT constructs were transiently expressed in HeLa cells and subsequently cell lysates were assayed for CAT activity. Direct counting of butyrylated chloramphenicol resulted in the relative CAT activities presented. Abbreviations: H, *HindIII*; P, *PstI*; R, *EcoRI*; MCS, multiple cloning site. (c) S1-nuclease mapping of the mAR-gene transcription start sites. The S1 probe is indicated in (a). Lane 1, probe; lane 2, RNA from mock-transfected COS-1 cells; lane 3, RNA from mouse prostatic tissue; lane 4, RNA from pmAR-CAT-1-transfected COS-1 cells.

(pmAR⁰; Fig. 4a). mAR was transiently expressed in COS-1 cells. To determine the size of the protein produced, Western-blot analysis was performed after immunoprecipitation of mAR from whole-cell extracts using an AR-specific monoclonal antibody. The blot was probed with an anti-AR polyclonal antibody. Controls included cell extracts prepared from mock-transfected COS-1 cells, COS-1 cells transfected with the hAR expression plasmid pAR⁰, cell extracts prepared from mouse testicular tissue and immunoprecipitations with a non-specific monoclonal antibody. The results of these experiments are shown in Fig. 4(b). A 110 kDa protein is specifically detected in pmAR⁰-transfected cells (compare lanes 3 and 4). Proteins of similar size are precipitated from pAR⁰-transfected cells (compare lanes 2 and 4) and mouse testicular tissue (compare lanes 4 and 8). The 110 kDa protein could not be precipitated from mock-transfected cells (lanes 5 and 6). These findings clearly show the pmAR⁰ expression vector to give rise to a full-length mAR protein. Scatchard-plot analysis revealed a dissociation constant (K_d) of 0.12 nM for methyltrienolone (R1881) binding to the expressed protein (results not shown). To test the *trans*-activating function of the cloned mAR, co-transfections with a CAT reporter-gene construct were performed in HeLa cells. In Fig. 4(a) the constructs used in these experiments are shown. pmAR⁰ is the mAR expression vector. The androgen-responsive CAT reporter gene construct is pG 29 G-TK-CAT [31], which consists of two synthetic copies of an HRE (GTTACAaacTGTTCT) upstream of the TK promoter, which is linked to the CAT reporter gene. PBL2-CAT [32] is a control construct which is comparable with pG 29 G-TK-CAT, but lacks the two HREs. pSV2-CAT is a control construct used for monitoring the transfection efficiency.

The autoradiograph of the t.l.c. analyses of the CAT assays is depicted in Fig. 4(c). The control experiments show that the synthetic androgen R1881 has no effect on CAT expression from the pG 29 G-TK-CAT reporter gene in the absence of the mAR (compare lanes 1 and 2). Similarly, in the presence of ligand, mAR has no stimulating effect on the TK promoter (compare lanes 3 and 4). In lanes 7 and 8 the activity of pSV2-CAT is illustrated, which, as expected, is not influenced by R1881. Co-transfection of pmAR⁰ and pG 29 G-TK-CAT shows a ligand-dependent increase in CAT-activity (compare lanes 5 and 6). These experiments prove that mAR is able to *trans*-activate an androgen-responsive promoter through the appropriate response elements. Quantification of the CAT assay resulted in the calculation of an induction factor of 30 (results not shown).

Determination of the transcriptional start sites and polyadenylation sites of the mAR gene

Hybridization studies using hAR probes spanning the transcription initiation sites of the hAR gene identified a region of sequence similarity in clone 5'mAR (Fig. 2b). In Fig. 5(a) a restriction map of the region surrounding the first exon of the mAR gene is shown. The ATG translation-initiation codon as well as the splice donor site at the exon 1/intron 1 boundary (GTAAG) are indicated. To determine whether the region upstream from the ATG codon contained promoter activity, a CAT reporter-gene construct containing SV40 enhancer sequences, but lacking transcription-initiation sites was used. Four CAT expression plasmids with mAR gene-promoter fragments were constructed (Fig. 5b). The original plasmid [pCAT-Enhancer (Promega)] will be referred to as 'pmAR-CAT-0'. Two

```

CTGCAGCTTGTCTTTAATGTGTCAGGAGACTCTCCCTTCTGCTTGTCTGGTGGGCCCTGG -439
GGGGAGCGGGGAGGAATACCTAAGAGCAATTTGGTAGCTGGTACTTCTAATGCCTCTTCC -379
TCCTCCAACCTCCAAGAGTCTGTTTTGGGATTTGGGTTTCAGGAATGAAATTCGCTGTGC -319
TAACTCCTCGGGAGCCGGTAGACTGTCTGTTAAAAATCGCTTCTGCTTTTGGAGCCTA -259
AAGCCCGGTTCCGAAAAACAAGTGGTATTAGGGGAAAAGGGGCTCTCAAAGGCTACAG -199
TGAGTCATTCCAGCCTTCAACCATACTACGCCAGCACTACGTTCTCTAAAGCCACTCTGC -139
GCTAGCTTGGCGGTGAGGGGAGGGGAGAAAGGAAAGGGGAGGGGAGGGGAGGGGAG - 79
AAGGAGGTGGGAAGGCAGAGAGCGCGGCTCGGGGGGGGACCGACTCAAAACTGTTC - 19
GATTTGCTTCCACTCCCGACGGCCCTCCGGAGATCCCTAGGAGCCAGCCTGTCTGGGAG + 42
AACCAGAGGGTCCGGAGCAAACTGGAGGCTGAGAGGGCATCAGAGGGGAAAAGACTGAG +102
    
```

Fig. 6. Sequence analysis of the promoter region of the mAR gene

The position of the transcription-initiation sites as determined from the S1-nuclease-protection experiments are indicated by dots. Two possible control elements, the consensus Sp1 binding site and the homopurine stretch, are underlined and doubly underlined respectively.

```

mouse CCCAAGCGCTAGTGTCTCTCTTTTGTAACTCTTGAATCTTTGTGTCTTAAATA 60
human TCCAAGCGCTAGTGTCTCTCTTTTGTAACTCTTGAATCTTTGTGTCTTAAATA 10374
mouse CATTAAAATGGCAGAACTGTTTCTTGGAAATACAGTGTGACTCTTGGTGTCTCT 120
human CATTAAAATGGCAGAACTGTTTCTTGGAAATACAGTGTGACTCTTGGTGTCTCT 10434
mouse GCGTCTGGCTTTAGAAATGTCATCCATTGCTAAATACTGGCTGTGTCTGCCAGCT 180
human GCCTCTGTCTTACAGAAATGTCATCCATTGAGTAAATAATTTGGCTACTGTCTGCCAGCT 10494
mouse AAAACTT GCCACAGCCCTCTGTTGACTGTCAGGCTCAAGTTATTTGTAACAAGAGCCC 239
human AAAACTTGGCCACATCCCTCTTATGGCTCAGGATCGAGTTATTTGTAACAAGAGACC 10554
mouse CAGAAAAGCTGCTAATGTCCTCTTATCACCATTGTTAATTTGTAAACATAAAACAAT 299
human CAGAAAAGCTGCTAATGTCCTCTTATCATTGTGTTAATTTGTAAACATAAAACAAT 10614
mouse CTAAAATTCAGATGAAATGTCATCAGAGTCTTTTTTATTAGTCTTTTTTATGGCAGTCT 359
human CTAAAATTCAGATGAAATGTCATCAGAGTCTTTTTTATTAGTCTTTTTTATGGCAGTCT
    
```

Fig. 7. Sequence analysis of the putative polyadenylation signals of mAR mRNA

The sequence as obtained from mAR genomic DNA clone 1A (Fig. 2) is compared with the sequence surrounding the well-defined polyadenylation signals of the hAR mRNA. The polyadenylation signals ATTA AAA and CATA AAA are boxed. The actual sites of poly(A) addition as determined in the hAR mRNA are indicated by dots, whereas the TG boxes after the polyadenylation signals are underlined in both sequences.

Specific protected fragments were absent in the control experiment (lane 2), whereas, in the case of mouse prostatic RNA (lane 3) and pmAR-CAT-1 transfected COS-1 RNA (lane 4), two sets of protected fragments were observed. The sequence of the region surrounding the transcription-initiation sites was determined (Fig. 6). The positions of the transcription-initiation sites is indicated with dots, and the most 5' nucleotide of the two sets of transcription initiation sites is numbered +1. The promoter region of the mAR gene lacks the canonical TATA/CCAAT elements which are found at 20–30 bp and 70–100 bp upstream from the transcription start sites of many genes. Neither is the promoter of the mAR gene exceptionally (G+C)-rich, as is often observed in TATA/CCAAT minus promoters. Potential promoter elements which are present in the mAR gene promoter are a consensus Sp1 binding site at positions –36 to –45 (underlined) and a homopurine stretch at positions –55 to –125 (doubly underlined).

Because clone 1A (Fig. 2b) hybridized to a hAR cDNA probe located at the very 3' end of the cDNA, additional hAR cDNA probes spanning the complete 6.8 kb 3'-UTR were hybridized to subcloned fragments from clone 1A. All probes gave rise to hybridization signals, but the probe spanning the polyadenylation signals of hAR mRNA produced the most intense signal (results not shown). These results suggested a moderate conservation for most of the 3'-UTR sequences and a more pronounced conservation for the sequences surrounding the polyadenylation sites. The fragment from clone 1A, which hybridizes to the 3'-terminus of hAR cDNA, was sequenced, and this sequence was matched with the corresponding hAR sequence (Fig. 7). The hAR DNA sequence extends from 10315 to 10625 from the transcription start site (cDNA), followed by a small stretch of 3'-flanking genomic sequences [20a]; the corresponding mouse sequence is numbered 1–359. The polyadenylation signals that were functionally identified in the hAR mRNA (ATTA AAA and CATA AAA) are boxed. The actual sites of poly(A) addition in the hAR mRNA are indicated by dots, and the TG boxes, usually found in the immediate downstream region of polyadenylation sites, are doubly underlined for both the human and mouse sequence.

DISCUSSION

In the present study the structural organization of the mAR gene is described. In addition, the mAR protein is functionally characterized. The mAR gene was shown to consist of eight exons. From the genomic clones an ORF of 2697 nucleotides was deduced that encodes a 899-amino-acid protein. This sequence was verified by the isolation of mAR cDNA using conventional cloning techniques coupled with PCR amplification.

The 899-amino-acid protein contains the domains (N-terminal 'regulatory', DNA-binding and ligand-binding) characteristic of the receptor family. Fig. 8 shows a schematic comparison of the sequence similarity between the mAR, the hAR and the rAR [12–17]. Total (100%) identity is observed in the DNA-binding and ligand-binding domains. In the N-terminal 'regulatory' domain, some (minor) differences were observed, resulting in 76 and 96% similarity for the hAR and rAR in these domains respectively. The overall similarity was 85% for the hAR and 98% for the rAR. Also indicated in Fig. 8 are the homopolymeric stretches of glutamine and glycine residues. As to the position within the protein, the mAR stretches resemble those of the rAR, but the composition of the glutamine stretch is different, because it is shorter by two amino acids and intermingled with three histidine residues (Gln₃-His-Gln₃-His-Gln₂-His-Gln₁).

Co-transfections of pmAR⁰ with the androgen-responsive CAT reporter gene construct pG 29 G-TK-CAT showed the

fragments, a 0.5 kb *PstI*–*HindIII* fragment and a 1.3 kb *PstI*–*PstI* fragment were cloned in front of the CAT gene in both sense and anti-sense orientation, giving rise to the constructs pmAR-CAT-1 and pmAR-CAT-1 rev in the case of the 1.3 kb *PstI*–*PstI* fragment and pmAR-CAT-2 and pmAR-CAT-2 rev in the case of the 0.5 kb *PstI*–*HindIII* fragment. All constructs were transfected into HeLa cells, together with the internal control plasmid pCH110, and 48 h after transfection cell extracts were assayed for CAT activity. Values obtained were normalized for β-galactosidase activity and the activity of the pmAR-CAT-0 plasmid was arbitrarily set to 1. The results of the experiments are given as relative CAT activities in Fig. 5(b) and show that the insertion of the two mAR gene fragments in the proper orientation produces a marked increase in CAT activity, whereas the insertion of the same fragments in the reverse orientation slightly reduces basal activity. Because pmAR-CAT-2 is active, the promoter of the mAR gene has to be located in this region. The precise positions of the transcription-initiation sites were subsequently mapped by S1-nuclease-protection experiments. RNA from mouse prostatic tissue as well as RNA from COS-1 cells transfected with the pmAR-CAT-1 promoter construct was isolated. RNA from mock-transfected COS-1 cells served as a control. The end-labelled 0.6 kb *EcoRI*–*HindIII* fragment (Fig. 5a) was used as a probe. The results are shown in Fig. 5(c).

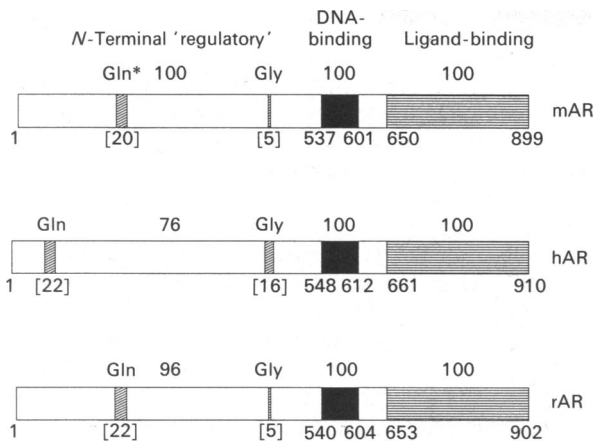


Fig. 8. Schematic comparison of the mAR, hAR and rAR

Indicated are the sequence similarities in the *N*-terminal 'regulatory' domain, the DNA-binding domain and the ligand-binding domain. Also indicated are the position and composition of the polyglutamine (Gln) and polyglycine (Gly) stretches in the *N*-terminal 'regulatory' domains. Gln* indicates the intermingling of the 17 glutamine residues with three histidine residues in the mAR polyglutamine stretch.

ability of mAR to *trans*-activate through a HRE in a ligand-dependent manner. The induction factor observed (30-fold) was comparable with that observed when the hAR expression plasmid pAR⁰ was used in a parallel experiment (P. W. Faber, unpublished work). Similar results were obtained in experiments in which the transcriptional activating properties of mAR and hAR were compared using the complete MMTV promoter in the reporter-gene construct (P. W. Faber, unpublished work). This indicates that the polyglycine stretch, which is long in hAR (16–27 residues), but virtually absent in mAR and rAR (five residues) is not essential for AR *trans*-activating function, at least not under the experimental conditions applied. The presence of a polyglutamine stretch of similar length and composition in the three receptors, although at a different position, would argue in favour of an important role of this stretch in AR function. A link between glutamine-rich protein domains and transcriptional

Table 1. Comparison of the intron lengths of mAR and hAR genes

Intron	Intron length (kb)	
	mAR	hAR
1	> 14.5	> 24.0
2	> 19.1	> 15.0
3	> 20.0	26.0
4	5.2	5.6
5	4.1	4.8
6	1.0	0.8
7	0.7	0.7

activation has been made in the transcription factor SP1 [38]. Interestingly, the recently cloned human TFIID transcription factor also contains a long polyglutamine stretch [39–41]. During the preparation of this manuscript the sequence of the mAR cDNA was reported by others [42,43]. Comparison of these cDNA sequences with the mAR sequence presented here showed the three sequences to be completely identical.

The information of the mAR cDNA is separated over exons 1–8 of the mAR gene in a manner identical with that found in the hAR gene [18–20a]. When compared with the hAR gene the lengths of the individual introns are roughly conserved between the two species (Table 1). The introns 1, 2 and 3 are large (part of the information for the introns 1 and 2 is missing for both genes, whereas intron 3 is 26 kb in the human gene and at least 20 kb in the mouse gene). The lengths of the introns 4–7 could be determined for both genes and were found to be highly comparable.

A similar genomic organization with conserved intron/exon boundaries has been reported for the chicken progesterone receptor (cPR) [44,45] and human oestrogen receptor (hER) [46] genes. The positions of the intron/exon boundaries are less well conserved in genes encoding more distant members of the receptor family, such as the human vitamin D₃ receptor [47] and the chicken thyroid-hormone receptor [48], indicating an early divergence during evolution between the steroid receptors and other receptors of the same family. The lengths of the introns 4–7

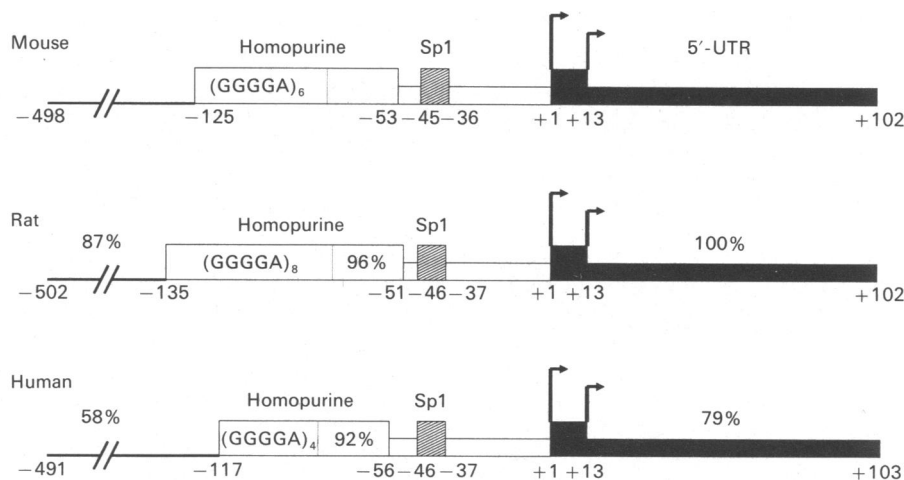


Fig. 9. Schematic comparison of the promoter regions of the mAR, rAR and hAR genes

Shown are the transcriptional start sites, the Sp1-binding site and the homopurine stretch, which was separated in a 5'- and 3'-segment, as well as part of the 5'-UTR sequences and part of the sequences upstream from the homopurine stretch. The sequence similarities are given for the 5'-UTR sequences, the 3' segment of the homopurine stretch and the sequences upstream from the homopurine stretch. In the 5'-segment of the homopurine stretch the number of GGGGA sequence motifs is given.

of the cPR and hER genes are unrelated to those observed in the mAR and hAR gene, with those of the cPR being substantially smaller and those of the hER gene being substantially larger.

Two major sites of transcription initiation were identified in a 13 bp region. These transcription-initiation sites could be defined using RNA preparations originating either from mouse prostatic tissue or from COS-1 cells transfected with mAR promoter-CAT constructs. Both promoter constructs (pmAR-CAT-1 and pmAR-CAT-2), which differed only by the length of the 5'-UTR, conferred promoter activity to a promoterless CAT reporter gene. The activity of pmAR-CAT-1, which contains an approx. 0.8 kb longer 5'-UTR than pmAR-CAT-2, was slightly higher. A similar effect has been noted using hAR promoter constructs [20a], suggesting a potentiating role of 5'-UTR sequences in transcriptional activation, an increased stability of the RNA produced or an increase in translation efficiency caused by the presence of the longer 5'-UTR.

Putative promoter elements are the consensus Sp1 binding site (GGGGCGGGAC, -36 to -45) and the homopurine stretch (-55 to -125). Sp1 is an ubiquitous transcription factor with DNA binding sites usually clustered in the promoter region of (G+C)-rich promoters, whereas homopurine stretches with various compositions have been identified in the promoter region of several genes [49]. These stretches are usually associated with a high sensitivity *in vitro* to S1-nuclease degradation and an unusual DNA structure [50]. Whether in the cell nucleus polypurine stretches are correlated with unfolding of the DNA is not known. Deletion of a polypurine/pyrimidine stretch in the promoter region of the epidermal-growth-factor-receptor gene resulted in a 3-fold decrease in promoter activity [51]. Analysis of this fragment by protein-DNA interaction experiments showed specific binding of two proteins, one of which turned out to be Sp1. A similar analysis of the promoter of the mouse *Ki-ras* gene also indicated the homopurine stretch to be involved in transcriptional activation and specific protein-DNA interactions [52].

Recently, we and others identified the promoter region of the hAR and rAR gene [20,20a,53]. A schematic comparison of the promoter regions of the three genes is given in Fig. 9. The positions of the transcription-initiation sites are identical in the three genes. The putative Sp1-binding site and the homopurine stretch are also conserved. The homopurine stretch can be divided in a 5'-fragment, consisting mainly of a multimer of the sequence motif GGGGA, and a 3'-fragment, without repeat elements. The 5'-segments of the homopurine stretch contains six, eight and four (GGGGA) blocks in the case of the mAR, rAR and hAR gene promoters respectively. More-upstream regions are moderately well conserved; the 5'-UTR shows high sequence similarity (see Fig. 9). A more detailed functional analysis will have to be conducted to evaluate the importance of the various structural elements in the AR promoter region.

The 3'-UTR of the mAR mRNA has a length of approx. 6 kb. This length is inferred from the identification of two potential polyadenylation signals in the mAR gene sequence (ATTAAA and CATAAA) at exactly the same position, where the functional polyadenylation signals of the hAR are situated [20a]. The two signal sequences differ from the canonical AATAAA hexamer sequence, but have previously been implicated in polyadenylation processes [54]. Although not strictly proven, the remarkable conservation between the hAR and mAR sequences in this region indicates that the same signals will be used for polyadenylation in the hAR and mAR transcripts. In addition, the strong sequence similarity suggests that the complete region is of high importance for polyadenylation or other, so-far-unidentified, functions of the AR gene or mRNA. The mRNA which would be produced from the complete mAR transcription unit

would have a length of approx. 10 kb {1 kb [5'-UTR]+2.7 kb [ORF]+6 kb [3'-UTR]+0.2 kb [poly(A)-tail]}, which closely corresponds to the reported length of the mAR mRNA [42,43,55].

In summary, we have defined the transcription unit of the mAR gene and functionally characterized the mAR protein product that originates from this gene. The information obtained provides the opportunity for a detailed analysis of regulation of expression of the mAR gene and mAR function in '*in vivo*' model systems. The sequence similarity between the mAR and hAR proteins and the promoter region of the mAR and hAR genes suggests a functional interchangeability for the individual protein domains as well as the regulatory regions of the promoters. A particularly interesting possibility would be either the correction of the naturally occurring mutation in the AR of the Tfm mouse [56] or the introduction of mutations, observed in the hAR in individuals with an aberrant male sexual development, in the mAR gene by homologous recombination techniques [57] to prove that these indeed correlate with the observed phenotypes. The information obtained from such experiments would be of great value for understanding of the role of AR in normal and abnormal male sexual development.

We are indebted to C. A. Berrevoets, J. A. G. M. van der Korput and J. Veldscholte for carrying out the Western and Southern blots and R1881 binding experiments and to Mrs. Paula Delfos for photography. This study was supported by a grant from the Dutch Cancer Society (KWF).

REFERENCES

1. Pinsky, L. & Kaufman, M. (1987) *Adv. Hum. Genet.* **16**, 299-472
2. Griffin, J. E. & Wilson, J. D. (1989) in *The Metabolic Basis of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D., eds.), pp. 1919-1944, McGraw-Hill, New York
3. Brown, T. R., Lubahn, D. B., Wilson, E. M., Joseph, D. R., French, F. S. & Migeon, C. J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8151-8155
4. Sai, T., Seino, S., Chang, C., Trifiro, M., Pinsky, L., Mhatre, A., Kaufman, M., Lambert, B., Trapman, J., Brinkmann, A. O., Rosenfield, R. L. & Liao, S. (1990) *Am. J. Hum. Genet.* **46**, 1095-1100
5. Marcelli, M., Tilley, W. D., Wilson, C. M., Wilson, J. E., Griffin, J. E. & McPhaul, M. J. (1990) *J. Clin. Invest.* **85**, 1522-1528
6. Marcelli, M., Tilley, W. D., Wilson, C. M., Griffin, J. E., Wilson, J. D. & McPhaul, M. J. (1990) *Mol. Endocrinol.* **4**, 1105-1116
7. Ris-Stalpers, C., Kuiper, G. G. J. M., Faber, P. W., Schweigert, H. U., van Rooij, H. C. J., Zegers, N. D., Hodgins, M. B., Degenhart, H. J., Trapman, J. & Brinkmann, A. O. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 7866-7870
8. Green, S. & Chambon, P. (1988) *Trends Genet.* **4**, 309-314
9. Evans, R. M. (1988) *Science* **240**, 889-895
10. Beato, M. (1989) *Cell (Cambridge, Mass.)* **56**, 335-344
11. O'Malley, B. (1990) *Mol. Endocrinol.* **4**, 363-369
12. Trapman, J., Klaassen, P., Kuiper, G. G. J. M., van der Korput, J. A. G. M., Faber, P. W., van Rooij, H. C. J., Geurts van Kessel, A., Voorhorst, M. M., Mulder, E. & Brinkmann, A. O. (1988) *Biochem. Biophys. Res. Commun.* **153**, 241-248
13. Chang, C., Kokontis, J. & Liao, S. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7211-7215
14. Lubahn, D. B., Joseph, D. R., Sar, M., Tan, J.-A., Higgs, H. N., Larson, R. E., French, F. S. & Wilson, E. M. (1988) *Mol. Endocrinol.* **2**, 1265-1275
15. Tan, J., Joseph, D. R., Quarmby, V. E., Lubahn, D. B., Sar, M., French, F. S. & Wilson, E. M. (1989) *Mol. Endocrinol.* **2**, 1276-1285
16. Faber, P. W., Kuiper, G. G. J. M., van Rooij, H. C. J., van der Korput, J. A. G. M., Brinkmann, A. O. & Trapman, J. (1989) *Mol. Cell. Endocrinol.* **61**, 257-262
17. Tilley, W. D., Marcelli, M., Wilson, J. D. & McPhaul, M. J. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 327-331
18. Kuiper, G. G. J. M., Faber, P. W., van Rooij, H. C. J., van der Korput, J. A. G. M., Ris-Stalpers, C., Klaassen, P., Trapman, J. & Brinkmann, A. O. (1989) *J. Mol. Endocrinol.* **2**, R1-R4

19. Lubahn, D. B., Brown, T. R., Simental, J. A., Higgs, H. N., Migeon, C. J., Wilson, E. M. & French, F. S. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8151–8155
20. Tilley, W. D., Marcelli, M. & McPhaul, M. J. (1990) *J. Biol. Chem.* **265**, 13776–13781
- 20a. Faber, P. W., van Rooij, H. C. J., van der Korput, J. A. G. M., Baarends, W. M., Brinkmann, A. O., Groolegoed, J. A. & Trapman, J. (1991) *J. Biol. Chem.*, in the press
21. Brinkmann, A. O., Faber, P. W., van Rooij, H. C. J., Kuiper, G. G. J. M., Ris, C., Klaassen, P., van der Korput, J. A. G. M., Voorhorst, M. M., van Laar, J. H., Mulder, E. & Trapman, J. (1989) *J. Steroid Biochem.* **34**, 307–310
22. Quarmby, V. E., Kempainen, J. A., Sar, M., Lubahn, D. B., French, F. S. & Wilson, E. M. (1990) *Mol. Endocrinol.* **4**, 1399–1407
23. Rundlett, S. E., Wu, X.-P. & Miesfeld, R. L. (1990) *Mol. Endocrinol.* **4**, 708–714
24. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
25. Davis, L. G., Dibner, M. D. & Battey, J. F. (1986) *Basic Methods in Biology*, Elsevier, New York
26. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
27. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. (1977) *Biochemistry* **18**, 5294–5299
28. Favoloro, J., Treisman, R. & Kamen, R. (1980) *Methods Enzymol.* **65**, 718–749
29. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Ehrlich, H. A. (1988) *Science* **239**, 487–491
30. Chen, C. & Okayama, H. (1987) *Mol. Cell. Biol.* **7**, 2745–2752
31. Schule, R., Muller, M., Kaltschmidt, C. & Renkawitz, R. (1988) *Science* **242**, 1418–1420
32. Luckow, B. & Schutz, G. (1987) *Nucleic Acids Res.* **15**, 5490
33. Gorman, C. M., Moffat, L. F. & Horward, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044–1052
34. Seed, B. & Sheen, J. Y. (1988) *Gene* **67**, 271–277
35. Zegers, N. D., Claassen, E., Neelen, C., Mulder, E., van Laar, J. H., Voorhorst, M. M., Berrevoets, C. A., Brinkmann, A. O., van der Kwast, T. H., Ruizeveld de Winter, J. A., Trapman, J. & Boersma, W. J. A. (1991) *Biochim. Biophys. Acta* **1073**, 23–32
36. Gallee, M. P. W., Visser-de Jong, E., van der Korput, J. A. G. M., van der Kwast, Th. H., ten Kate, F. J. W., Schroeder, F. H. & Trapman, J. (1990) *Urol. Res.* **18**, 181–187
37. van Laar, J. H., Voorhorst-Ogink, M. M., Zegers, N. D., Boersma, W. J. A., Claassen, E., van der Korput, J. A. G. M., Ruizeveld-de Winter, J. A., van der Kwast, Th. H., Mulder, E., Trapman, J. & Brinkmann, A. O. (1989) *Mol. Cell. Endocrinol.* **67**, 29–38
38. Courey, A., Holtzman, D. A., Jackson, S. P. & Tjian, R. (1989) *Cell (Cambridge, Mass.)* **59**, 827–836
39. Peterson, M. G., Tanese, N., Pugh, B. F. & Tjian, R. (1990) *Science* **248**, 1625–1630
40. Kao, C. C., Lieberman, P. M., Schmidt, M. C., Zhou, Q., Pei, R. & Berk, A. J. (1990) *Science* **248**, 1646–1650
41. Hoffmann, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A., Hori-koshi, M. & Roeder, R. G. (1990) *Nature (London)* **346**, 387–390
42. He, W. W., Fischer, L. M., Sun, S., Bilhartz, D. L., Zhu, X., Young, C. Y.-F., Kelley, D. B. & Tindall, D. J. (1990) *Biochem. Biophys. Res. Commun.* **171**, 697–704
43. Gaspar, M. L., Meo, T. & Tosi, M. (1990) *Mol. Endocrinol.* **4**, 1600–1610
44. Huckaby, C. S., Conneely, O. M., Beattie, W. G., Dobson, A. D. W., Tsai, M.-J. & O'Malley, B. W. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8380–8384
45. Jeltsch, J.-M., Turcotte, B., Garnier, J.-M., Lerouge, T., Krozowski, Z., Gronemeyer, H. & Chambon, P. (1990) *J. Biol. Chem.* **265**, 3967–3974
46. Ponglikitmongkol, M., Green, S. & Chambon, P. (1988) *EMBO J.* **7**, 3385–3388
47. Zahraoui, A. & Cuny, G. (1987) *Eur. J. Biochem.* **166**, 63–69
48. Hughes, M. R., Malloy, P. J., Kieback, D. G., Kesterson, R. A., Pike, J. W., Fieldman, D. & O'Malley, B. W. (1988) *Science* **242**, 1702–1705
49. Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M. & Wohlrab, F. (1988) *FASEB J.* **2**, 2939–2949
50. Wells, R. D. (1988) *J. Biol. Chem.* **263**, 1095–1098
51. Johnson, A. C., Jinno, Y. & Merlino, G. T. (1988) *Mol. Cell. Biol.* **8**, 4174–4184
52. Hoffman, E. K., Trusko, S. P., Murphy, M. & George, D. L. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2705–2709
53. Baarends, W. M., Themmen, A. P. N., Blok, L. J., Mackenbach, P., Brinkmann, A. O., Meijer, D., Faber, P. W., Trapman, J. & Grootegoed, J. A. (1990) *Mol. Cell. Endocrinol.* **74**, 75–84
54. Sheets, M. D., Ogg, S. C. & Wickens, M. P. (1990) *Nucleic Acids Res.* **18**, 5799–5805
55. Lubahn, D. B., Joseph, D. R., Sullivan, P. M., Willard, H. F., French, F. S. & Wilson, E. M. (1988) *Science* **240**, 327–330
56. He, W. W., Young, C. Y. & Tindall, D. J. (1990) *Ann. Meet. Endocrine Soc. 72nd, Atlanta, GA*, abstr. 862
57. Capecchi, M. R. (1989) *Science* **244**, 1288–1292

Received 14 January 1991/26 February 1991; accepted 4 March 1991