The Human Androgen Receptor: Complementary Deoxyribonucleic Acid Cloning, Sequence Analysis and Gene Expression in Prostate

Dennis B. Lubahn, David R. Joseph, Madhabananda Sar, Jiannan Tan, Henry N. Higgs, Robert E. Larson, Frank S. French, and Elizabeth M. Wilson

Laboratories for Reproductive Biology
and the Departments of Pediatrics, Physiology (D.R.J.), Cell Biology and Anatomy (M.S.), Biology (J.-A.T.), and Biochemistry (E.M.W.)
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27599

Androgenic hormones mediate their effects on male sex differentiation and development through a high affinity receptor protein. We report here cloning of the complete coding sequence of the human androgen receptor (hAR). By sequence homology hAR is a member of the nuclear receptor family, with closest sequence identity to the progesterone, mineralocorticoid, and glucocorticoid receptors. Regions of highest homology include the DNA-binding domain and a small region within the hydrophobic ligand-binding domain. Comparison of the deduced 919 amino acid sequence of hAR (98,999 mol wt) to the 902 amino acid sequence of rat AR (98,227 mol wt) reveals identical sequences in the DNA- and hormone-binding domains, with an overall homology of 85%. In human prostate, the major androgen receptor mRNA species is 10 kilobases while a less abundant mRNA is approximately 7 kilobases. Rabbit polyclonal antibodies were raised against a synthetic peptide from the N-terminal region of hAR. Immunocytochemical analysis of human prostate tissue demonstrated that AR is localized predominantly in nuclei of glandular epithelial cells. (Molecular Endocrinology 2: 1265–1275, 1988)

INTRODUCTION

Androgens and their receptor play an important role in male physiology and pathology. The androgen receptor (AR) is a member of the steroid receptor family which is, in turn, part of a larger family of nuclear receptors (1, 2). These receptors, when bound with the appropriate steroid or other hydrophobic ligands, become transacting factors regulating transcription of specific genes (3). ARs bind the male sex steroids, dihydrotestosterone and testosterone, and regulate genes necessary for male sexual differentiation and development (4). In the genetic male, defects in the AR gene result in a spectrum of developmental abnormalities ranging from a phenotypic female to varying degrees of incomplete masculinization (5). At puberty androgens stimulate further growth and functional differentiation of the male reproductive tract. Maintenance of spermatogenesis, sperm maturation, and accessory sex gland function are dependent on androgens and their receptor (6). Human prostate cancers often remain androgen dependent in the early stages of their growth (7). The need to further understand these diverse actions of androgens led us to clone the human AR (hAR) (1).

Almost all DNA sequences encoding known members of the human steroid receptor family have been cloned and characterized. These include the glucocorticoid (8), estrogen (9, 10), progesterone (11), mineralocorticoid (12), and vitamin D receptors (13). In addition, DNA clones for nuclear receptors outside the steroid receptor group, *i.e.*, the human thyroid hormone receptors (14, 15) and human retinoic acid receptors (16–19), have been isolated. More recent isolates have included DNA clones encoding nuclear receptors for ligands not yet identified (20).

Earlier reports on cloning the hAR (1, 2) described amino acid sequences only within the DNA-binding domain (1, 2) or partial cDNA sequence (21). Proof that the isolated recombinant clones encoded AR came from the expression of AR protein with the correct steroid-binding affinity and specificity. Cloning and complete sequence analysis of cDNA for hAR, the last of the known steroid receptors to be cloned, are reported herein. In addition, AR gene expression *in vivo* is examined in human prostate by northern blot analysis of hAR mRNA and by immunocytochemical localization of receptor protein.

0888-8809/88/1265-1275$02.00/0
Molecular Endocrinology

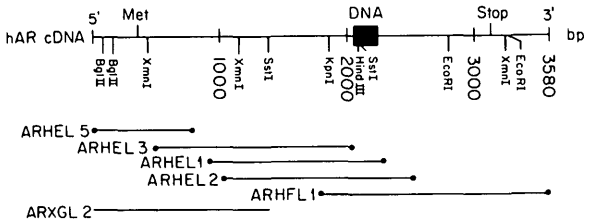


Fig. 1. Restriction Map of Overlapping hAR cDNA and Genomic Clones

Complementary clones ARHEL1, 2, 3, and 5 were isolated from a human epididymis lambda gt11 library and cDNA clone ARHFL1 was from a human foreskin fibroblast lambda gt11 library. A 15-kbp genomic clone was isolated from a lambda Charon 35, partial *Sau*3A, chromosome-sorted, X genomic DNA library. The *Sst*I genomic fragment, ARXGL2, was sequenced from its 3′-*Sst*I site through sequence corresponding to the 5′-end of cDNA clone ARHEL5. Ends of the clones are indicated with filled circles. The predicted initiation methionine (Met), the in-frame stop codon (Stop), and the DNA-binding domain (■) are indicated. Cloning strategy: cDNA fragments ARHEL1 and ARHFL1 were obtained using as probe an X-chromosome sequence (oligo B) described previously (1). Complementary DNA clones ARHEL2, ARHEL3, and a human X genomic clone (digested with *Sst*I to yield ARXGL2) were obtained using as probe a 5′ 450-bp fragment of ARHEL1. The cDNA clone ARHEL5 was obtained using a 2.7-kbp *Sst*I genomic probe that included ARXGL2. All clones were sequenced completely on both strands.

RESULTS

Human AR cDNA Sequence and Derived Protein Sequence

Two AR cDNA clones, ARHEL1 and ARHFL1, isolated as previously described (1), contained the entire DNA- and steroid-binding domains (Fig. 1). A ^{32}P-labeled 5′ 450-base pair (bp) fragment of ARHEL1 was used as probe to isolate two additional cDNA clones from a human epididymis lambda gt11 cDNA library (ARHEL2 and 3) and a genomic clone from a human X chromosome library in Charon 35 containing fragment ARXGL2. These recombinant DNA inserts yielded overlapping sequences spanning the entire N-terminal and DNA-binding domains (Fig. 1). We predicted that the hAR gene would contain the entire N-terminal region in a single exon as reported for the chicken progesterone receptor gene (22). Coding sequence at the 5′-end was confirmed with an additional human epididymal cDNA clone (ARHEL5) isolated by probing with the genomic DNA fragment (Fig. 1).

The complete sequence of the composite human AR cDNA, 3.6-kilobase pairs (kbp) in length, is shown in Fig. 2 together with the deduced amino acid sequence. It contains an open reading frame of 2757 bp encoding 919 amino acids. At the 5′-end a stop codon was observed at residue 258, upstream 105 bases from the initiation methionine at residue 363. The open reading frame ends with a TGA stop codon at residue 3120. The conserved cysteines (marked with asterisks in Fig. 2) and other conserved amino acids characterize the DNA-binding domains of nuclear receptors. A cDNA fragment in this reading frame encompassing the DNA- and hormone-binding domains was expressed to produce a protein that bound androgens with high affinity and specificity characteristic of the native AR (1, 2). The cDNA clone contained a series of adenylic acid residues at its 3′-untranslated end, but there was no efficient consensus sequence for poly(A) addition. A series of adenylic acid residues was observed also in hAR genomic DNA (data not shown) suggesting that it is not the poly(A) tail of the mRNA for hAR.

Structural Comparison of hAR and Rat AR

Human AR amino acid sequence deduced from DNA sequence was compared to that of rat AR (23). In Fig. 3A 919 amino acids of hAR are aligned with 902 amino acids of rat AR. Overall amino acid sequence homology is 85% while at the nucleotide level the homology is 83%. Within the DNA- and hormone-binding domains, amino acid sequences of the two receptors are identical (Fig. 3B). The carboxyl-terminal halves of the hAR and rat ARs, including the DNA- and hormone-binding domains, differ by only seven amino acid residues.

In contrast, the N-terminal regions of hAR and rat ARs have significant sequence divergence (Fig. 3). The overall N-terminal homology is only 77% (Fig. 3B), primarily due to repeated single amino acid motifs, which though variable in length, generally occur at the same position. For example, there are 24 repeated glycines in hAR but only five in rat AR at the same location. Similarly, five repeated glutamines in hAR correspond in position to 22 glutamines in rat AR (Fig. 3B) and eight prolines in hAR correspond to a broken sequence of seven prolines in rat AR (Fig. 3, A and B). Repeats of glutamine and arginine occur in one receptor but not in the other. Sequence homology between the repeats is 78–90% and, at one location, five consecutive alanine residues in hAR correspond to an equal number in rat AR (Fig. 3B).

Surprisingly, one group of repeated amino acids varied in length within the human species. The 21 glutamines in the hAR were observed in two cDNA clones (ARHEL5 and ARHEL3), while there were 25 repeated glutamines at the same location in the genomic clone, ARXGL2. Since these clones were isolated from independent libraries prepared from DNA of different individuals, the discrepancy in glutamine number may represent an allelic polymorphism.

Human AR Structural Homology with Other Members of the Nuclear Receptor Family

The deduced amino acid sequence of hAR was compared to other members of the nuclear receptor family. In Fig. 4A, schematic drawings of receptor structure are aligned on the basis of conserved amino acid sequences. Human AR is most similar to other receptors in its DNA-binding domain, where there are sequence homologies of 82% with the human progesterone re-

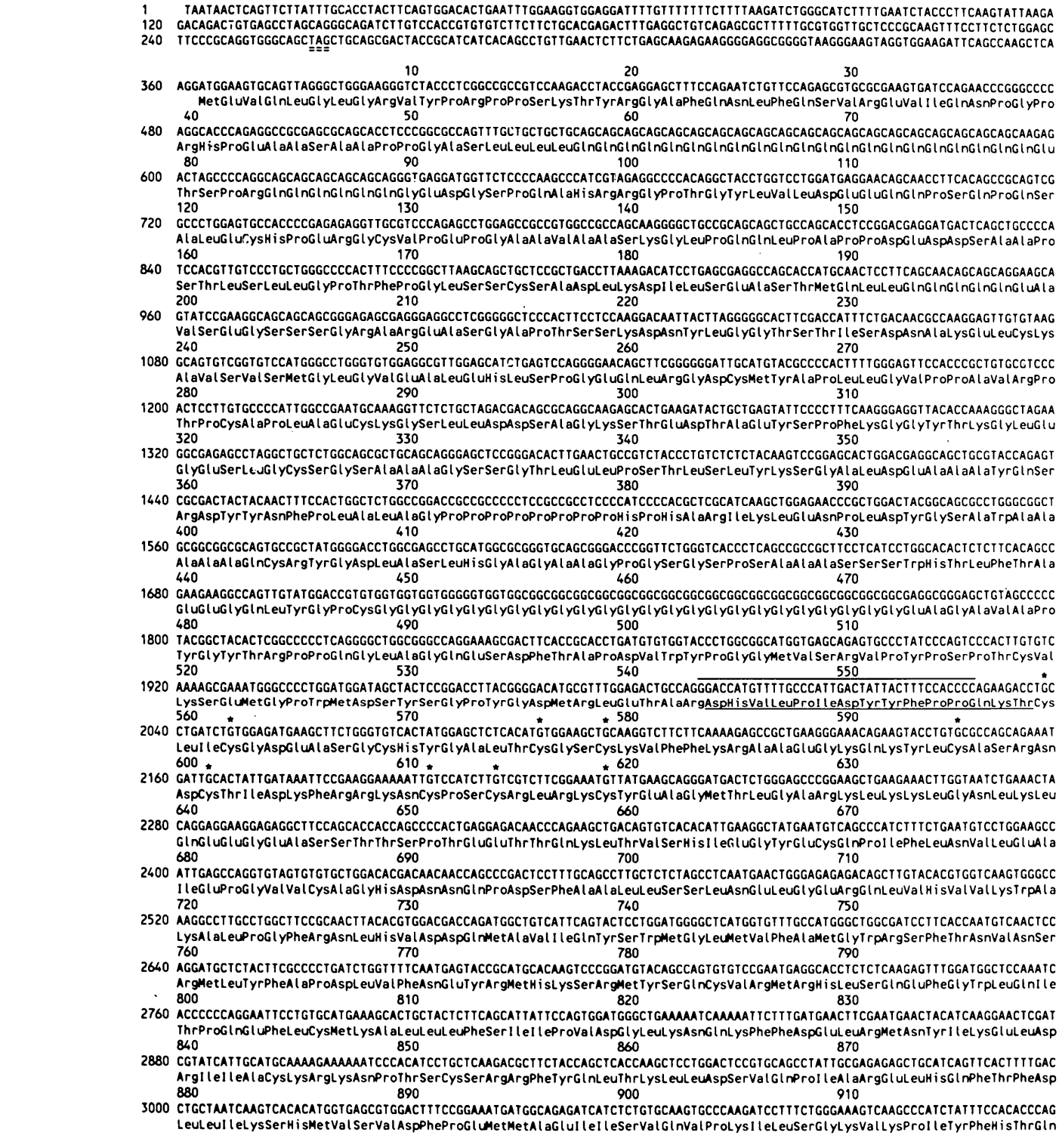


Fig. 2. Composite Nucleotide Sequence of hAR cDNA and Deduced Amino Acid Sequence

Nucleotides are numbered on the *left*. Amino acids are numbered *above* the sequence beginning with the predicted initiation methionine. An in-frame stop codon 5′ from the predicted initiation methionine is indicated by the *double dashed underline*. The 38-bp oligonucleotide, oligo B, used in initial screening of cDNA libraries (1), is indicated by the *solid overline*. The synthetic peptide to which rabbit polyclonal antibodies were raised is indicated by the *solid underline*. Conserved cysteines within the DNA-binding domain are indicated by *asterisks*. The nucleotide sequence of the hAR is available from GenBank under accession number J03180.

A

```
human MEVQLGLGRVYPRPPSKTYRGAFQNLFQSVREVIQNPGPRHPEAASAAPPGASLLLLQQQ  60
rat                                    A            I     C ------  54

human QQQQQQQQQQQQQQQQQQETSP-RQQQQQQGEDGSPQAHRRGPTGYLVLDEEQQPSQPQS 119
rat   --------------  R     R RRR  HP        I  T    A E       Q   100

human ALECHPERGCVPEPGAAVAASKGLPQQLPAPPDEDDSAAPSTLSLLGPTFPGLSSCSADL 179
rat    S G   S  L      T PG      P     Q                         I 160

human KDILSEASTMQLL-----------------QQQQQEAVSEGSSSGRAREASGAPTSSKDN 222
rat         G     QQQQQQQQQQQQQQQQQ      VI      V     T   S    S 220

human YLGGTSTISDNAKELCKAVSVSMGLGVEALEHLSPGEQLRGDCMYAPLLGVPPAVRPTPC 282
rat       N     S                                    S   G          280

human APLAECKGSLLDDSAGKSTEDTAEYSPFKGGYTKGLEGESLGCSGSAAAGSSGTLELPST 342
rat           LS  EGP  G  E     S     A             SE        I  S 340

human LSLYKSGALDEAAAYQSRDYYNFPLALAGPPPPPPPPHPHARIKLENPLDYGSAWAAAAA 402
rat           V       N          S   H    T           S            400

human QCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEGQLYGPCGGGGGGGGGGGGGG 462
rat               GSV   ST   P T                    ---------------  445

human GGGGGGGGGG-----EAGAVAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPT 517
rat   -----      SSSPSD  P              S  G  S SE      V N       S 500
                                                    *  *            *
human CVKSEMGPWMDSYSGPYGDMRLETARDHVLPIDYYFPPQKTCLICGDEASGCHYGALTCG 577
rat             EN          DST                                      560
      *              *     *         *  *    *
human SCKVFFKRAAEGKQKYLCASRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKKLGNL 637
rat                                                                  620

human KLQEEGEASSTTSPTEETTQKLTVSHIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSF 697
rat           N  AG    DPS  M                                       680

human AALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNV 757
rat                                                                  740

human NSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIP 817
rat                                                                  800

human VDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFT 877
rat                                                                  860

human FDLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVKPIYFHTQ 919
rat                                               902
```

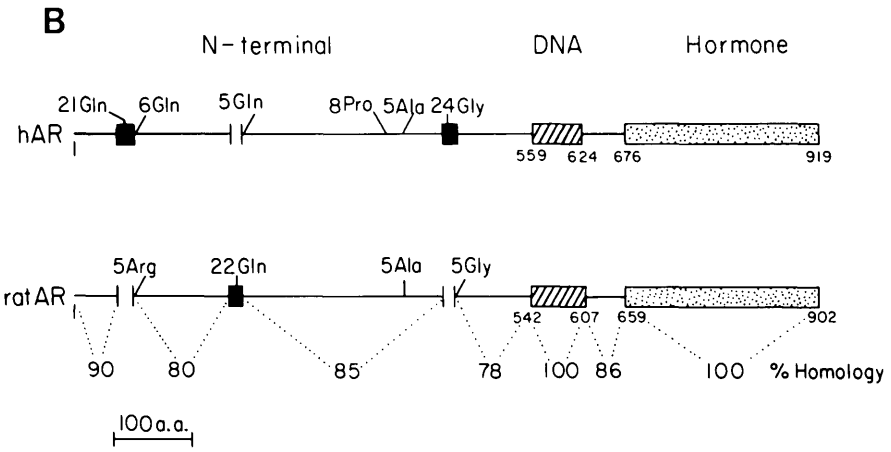


Fig. 3. Comparison of Deduced Amino Acid Sequences of hAR and Rat AR

A, Amino acid sequence alignment begins at the predicted initiation methionines. Only rat AR amino acid residues (23) that differ from the corresponding human amino acids are shown. A *hyphen* indicates a gap in the sequence where there is no corresponding

ceptor (hPR) and 79% with the human mineralocorticoid (hMR) and human glucocorticoid receptors (hGR). As described earlier (3, 24), cysteine residues thought to be critical in zinc finger formation are highly conserved. However, a histidine at residue 591 or 592 which may be involved in second finger formation in other receptors, is absent in hAR, as in the retinoic acid (16–19) and vitamin D receptors (13) and human estrogen receptor-related proteins (20). Linker regions located between the DNA- and hormone-binding domains and the N-terminal regions are variable in size and lower in sequence homology. Both rat AR and hAR contain long N-terminal regions like those of hPR (11) and hMR (12). Also like the N-terminal domains of hPR (11), hGR (8), and hMR (12), this domain in hAR has a net negative charge higher than other domains (Fig. 4B). These negatively charged and relatively hydrophilic N-terminal domains may play a role in reducing nonspecific binding to DNA (25). Hormone-binding domains of all steroid receptors are similar in size, and again hAR most closely resembles the hPR, hMR, and hGR, with homologies of around 50%. Compared with the N-terminal region (Fig. 4B), the steroid-binding domain has a greater average hydrophobicity, a structural feature consistent with its binding of hydrophobic steroids.

Despite the overall sequence variability within the hydrophobic ligand-binding domains, there is a small region of striking sequence similarity among all members of the steroid receptor family and in the retinoic acid and thyroid hormone receptors. This 23 amino acid sequence near the 5′-end of the hormone-binding domain shows sequence identity of 44–74% (Fig. 4C). If, in addition, structurally similar amino acids are included in the comparison, sequence relatedness within this region increases, ranging from 65–100% (Fig. 4C). Again, highest AR sequence similarity is with hPR, hMR, and hGR (Fig. 4C).

Expression of AR mRNA and Protein in the Human Prostate

Northern blot hybridization with a ^{32}P-labeled hAR cDNA probe revealed two species of AR mRNA in poly(A) RNA isolated from benign hyperplastic human prostate tissue (Fig. 5, lane a). The predominant 10-kilobase (kb) mRNA species was found also in human foreskin fibroblasts (Fig. 5, lane b). The 10-kb mRNA is the major AR mRNA species in rat ventral prostate (Fig. 5, lane c) and other tissues of the rat (23). A 7-kb mRNA species, prevalent in RNA extracts of human prostate (Fig. 5, lane a), was present at a relatively low level in poly(A) RNA from human foreskin fibroblasts (Fig. 5, lane b). The 7-kb mRNA species has not been observed in RNA extracts of rat tissue (23) including rat prostate (Fig. 5, lane c).

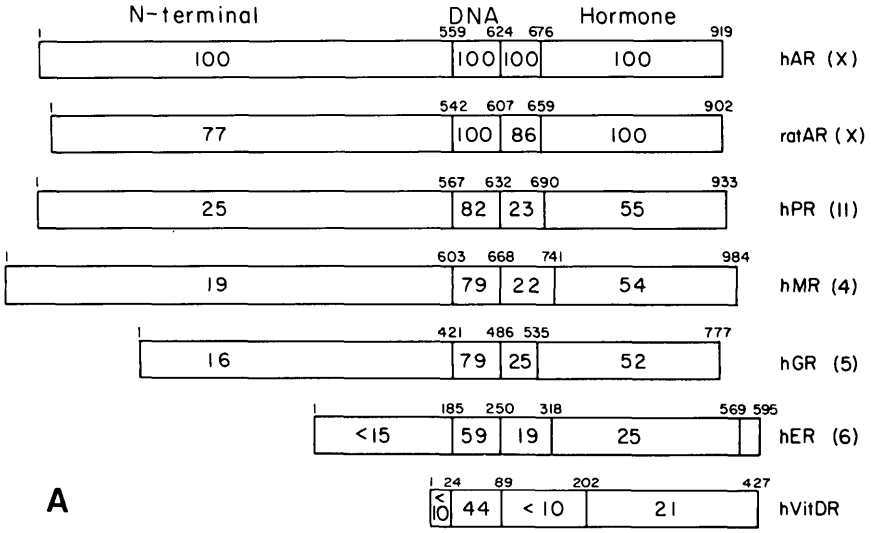
Steroid receptor protein localization within tissues and within specific cell types has been an area of intense interest, particularly the localization of sex steroid receptors in stromal and epithelial cell types (26–32). Our studies with rabbit polyclonal antibodies raised against an AR-specific peptide (23) (shown underlined in Fig. 2) indicate that AR protein is located in nuclei primarily of glandular epithelial cells of benign hyperplastic human prostate (Fig. 6). Little staining is observed in stromal cell nuclei. Immunostaining was not observed with control sections using either preimmune serum or antiserum preabsorbed with the peptide antigen. Nuclear localization of hAR by immunocytochemistry provides verification of earlier biochemical (33) and autoradiographic (34) evidence that the AR is a member of the nuclear receptor family.

DISCUSSION

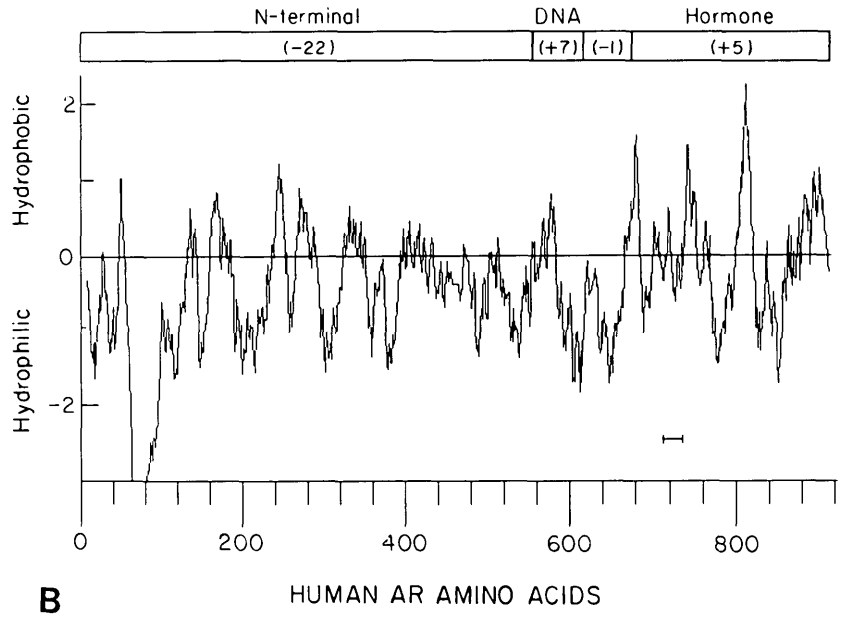
We have isolated and sequenced overlapping DNA clones, 3580 bp in length, which contain the entire coding region of the hAR cDNA. It is evident from sequence comparisons that hAR is a member of the nuclear receptor family which includes steroid hormone receptors. The deduced 919 amino acids of hAR predict a protein of 98,999 mol wt, which is somewhat smaller than the 117,000 mol wt rat receptor determined from hydrodynamic studies (35). Comparison with the derived rat AR protein sequence of 902 amino acids (predicted mol wt 98,227) suggests that the AR is highly conserved across species in its steroid- and DNA-binding domains. Major differences reside, however, in the N-terminal domain due largely to single amino acid repeats.

The discrepancy between the length of the isolated cDNA sequence (3.5 kbp) and hAR mRNA (7 and 10 kb) remains to be fully explained. The 7-kb human mRNA species probably results from differential processing of a precursor mRNA through a pathway not present in the rat (23). Based on the structure of rat AR mRNA (23), human AR mRNA could have at least 1 kb of 5′-untranslated region. The remaining difference in size between mRNA and coding sequence (3–6 kb) is likely due to a large 3′-untranslated region which is characteristic of the family of steroid hormone receptors (8, 10, 12, 13). Poly(A) residues within the 3′-noncoding region probably acted fortuitously as template for oligo(dT)-primed first-strand synthesis in preparation of the cDNA library. Further sequence analysis of the hAR

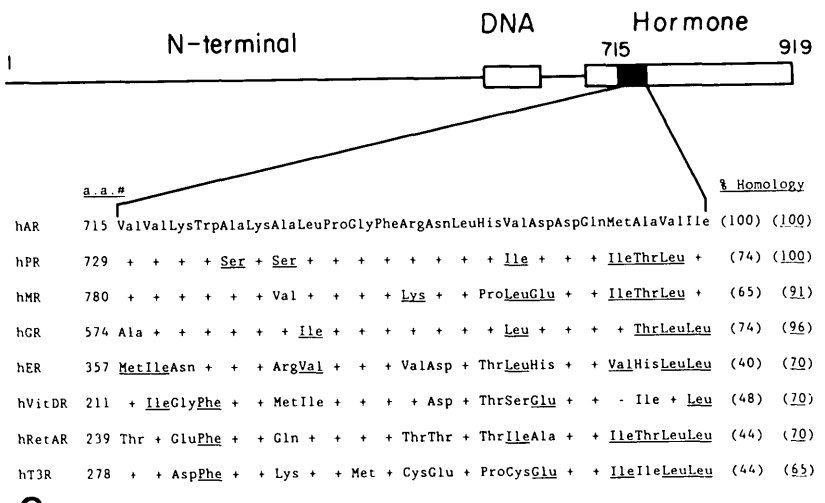
amino acid. The *asterisks* indicate conserved cysteines of the DNA-binding domain. Amino acid residue number is indicated on the *right*. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. B, The schematic drawing compares percent amino acid homology (indicated at *bottom* of figure) of individual segments (*dotted lines*) between rat AR and hAR. Amino acid residues are numbered immediately *below* the *hatched* and *stippled boxes*. Repeated amino acid motifs (>4 in length) within the N-terminal region are indicated (some by *solid boxes*).



A



B



C

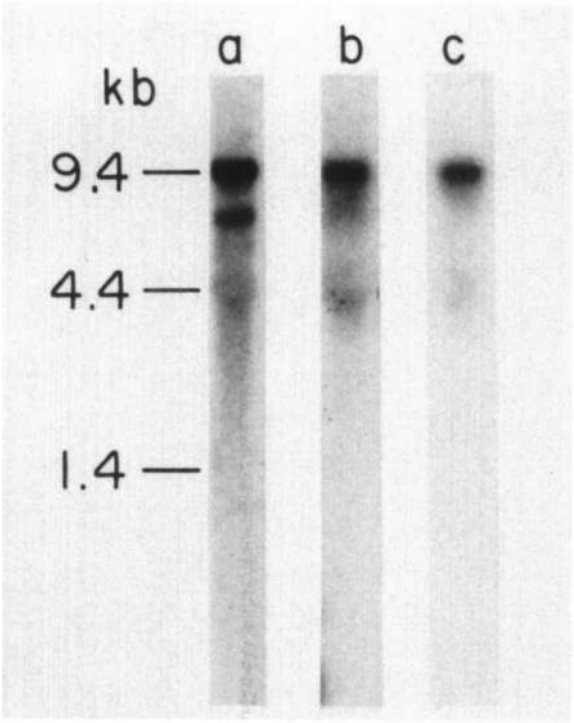


Fig. 5. Northern Blot Analysis of AR mRNA

Poly(A) RNA was isolated and analyzed as described in *Materials and Methods*. Shown are autoradiograms of northern blots containing human prostate poly(A) RNA, 10 μg (lane a), human foreskin fibroblast poly(A) RNA, 20 μg (lane b), and rat ventral prostate poly(A) RNA, 20 μg (lane c). The ^{32}P-labeled DNA probe for lane a was a 2.7-kbp *Sst*I genomic fragment containing ARXGL2 (Fig. 1). The probe for lanes b and c was a 0.7-kbp *Hin*dIII/*Eco*RI fragment of cDNA clone ARHFL1 shown in Fig. 1. The hybridizing band at 4.4 kp is thought to result from cross-hybridization of the AR probe with ribosomal RNA.

gene will elucidate the size of the 3′-noncoding sequence of hAR mRNA.

The putative translation initiation signal of hAR at nucleotide residue 363 (Fig. 2) begins a long open reading frame and is preceded by a stop codon 105 bp upstream at nucleotide residue 268. A second potential initiating methionine which is also flanked by a consensus translation initiation-like sequence (36) occurs in reading frame at nucleotide residue 927 (amino acid residue 189). This second potential in-frame initiation site would result in a protein of 76,000 mol wt, which is similar in size to a receptor protein expressed from a partial hAR cDNA clone by Chang *et al.* (2). However, this second methionine is preceded by three out-of-frame methionines that could interfere with initiation at the second site. Putative downstream second initiation sites with initiation-like sequences have been observed in human and chicken PRs (11, 37), and it has been suggested that the two initiation sites account for the A and B forms of the PR (37). At present, however, there is no compelling evidence that AR mRNA is translated *in vivo* to produce two proteins of different size.

Similarities in sequence and domain structure of nuclear receptors provide strong evidence that they evolved from a common ancestral gene. Furthermore, sequence homology among AR, PR, MR, and GR suggests that this group diverged from a common predecessor at a different time in evolution than did the estrogen and vitamin D receptors. The highly conserved 23 amino acid region within the ligand-binding domain was considered particularly noteworthy and suggested an important function for this region. Mutations within the 23 amino acid region have been shown to interfere with ligand binding (38–40), implicating this region as part of a hydrophobic ligand-binding pocket. Alternatively, this sequence may be involved in transducing a ligand-binding signal that converts the untransformed receptor to its activated DNA-binding form (41). Since all steroid receptors and receptor mutants that retain this conserved region can be found in an untransformed 9-10S non-DNA binding form (38, 42, 43), it has been speculated that the 23 amino acid sequence represents a binding site for a receptor-associated protein such as the 90-kilodalton heat shock protein (38, 41). It appears, however, that the conserved 23 amino acid region is not required for binding of an AR associated protein referred to as 8S-factor (44–46). The 8S-factor interacts with the intact 117,000 mol wt (4.5S) native receptor and its 60,000 mol wt (3.6S) proteolytic fragment, but it does not interact with the 30,000 mol wt (3S) proteolytic fragment which retains steroid-binding properties of native receptor (47).

Within the N-terminal domain, the amino acid repeat

Fig. 4. Amino Acid Homology Comparisons among Nuclear Receptor Proteins and Hydropathicity of hAR

A, Comparison of hAR deduced amino acid sequence with hPR (11); hMR (12); hGR, (8); hER, human estrogen receptor (9, 10); and hVitDR, human vitamin D receptor (13). Indicated are the percent sequence homologies with hAR (*within boxes*), amino acid residue number (*above boxes*), and chromosomal locations (*in parentheses on right*) (1, 71). B, A hydropathic plot of hAR was obtained using a protein structure analysis computer program prepared by the University of Wisconsin Genetics Computer Group (72) with the algorithm of Kyte and Doolittle (73) and an average window of 15 amino acids. Overall net charge of each receptor domain is indicated in *open boxes at the top*. The *solid line below the plot* indicates the conserved 23 amino acid sequence described below. C, Comparison of a 23 amino acid region within the steroid-binding domain of members of the nuclear receptor family. Indicated are amino acids identical to the corresponding hAR sequence (+), absence of a corresponding amino acid (−), amino acid structurally related to corresponding hAR sequence (*underlined*), and residue number of the first amino acid (*on left*). Also shown is the percent homology with hAR (*in parentheses immediately on right*) and the percent homology including structurally related amino acids (*underlined in parentheses at far right*). Groups of structurally related amino acids are (Ala Ser Thr) (Asn Gln) (Asp Glu) (Ile Leu Met Val) (Arg His Lys) (Phe Trp Tyr) according to the MicroGenie computer program. Additional abbreviations are hRetAR, human retinoic acid receptor (16–19), and hT3R, human thyroid hormone receptor (14, 15).

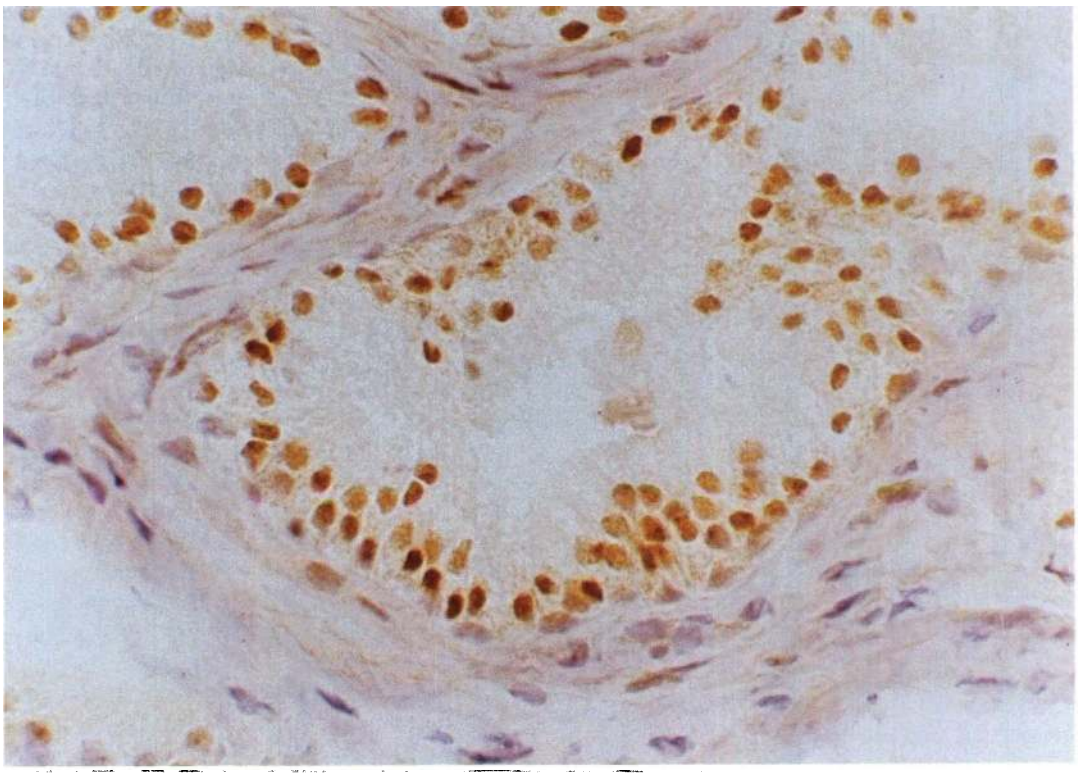


Fig. 6. Immunocytochemical Localization of AR in Human Prostate

Rabbit polyclonal antiserum raised against a hAR peptide (underlined in Fig. 2) (23) was used to immunostain androgen receptor in a fixed frozen section of prostate from a patient with benign prostatic hyperplasia using the avidin-biotin peroxidase method (70). Tissue sections were counter-stained with hematoxylin. Magnification, ×450.

motifs of differing length in hAR and rat AR have also been observed in other steroid receptors. The chicken PR has a poly-glutamic acid sequence not present in human and rabbit PR (11, 37, 48). Estrogen receptors have repeated alanines which vary in length among human, rat, mouse, chicken, and Xenopus, from none in the chicken estrogen receptor to nine alanines in the mouse estrogen receptor (9, 10, 49–52). GRs have a 19 glutamine motif in rat (53), two in human (8), and eight in mouse (41). In most instances these repeats of different length occur at the same locations within the N-terminal region. Other proteins known to have repeated amino acids include antifreeze glycoproteins that depress the blood freezing point of some arctic fish (54), rat tyrosine hydroxylase (55), and two developmentally regulated drosophila proteins (56, 57). The function of these repeat motifs and the significance of their size variation among species is not known.

A within-species sequence variation in the number of glutamines in hAR from different human subjects suggests that this N-terminal region may be a site of flux, perhaps providing a clue to the origin of this type of repeat. Individual variations point to a possible site for a frequent polymorphism, perhaps useful in carrier detection and prenatal diagnosis of the androgen insensitivity syndrome (4, 5).

Antibodies that react with hAR were obtained by immunizing rabbits with a hydrophilic peptide of unique sequence (23). The antibodies are effective probes for AR protein and demonstrate its presence in nuclei primarily in epithelial cells of hyperplastic human prostate. This tissue localization is consistent with an earlier study of human glandular benign prostatic hyperplasia using ^{3}H-androgen autoradiography (7) and with our immunocytochemical studies in rat ventral prostate where the same antibody detected ARs in epithelial cell nuclei almost exclusively (23). Other autoradiographic studies of prostate and human prostate adenocarcinoma have indicated ARs are located predominately in epithelial cells (26–28, 34).

The immunocytochemical approach should prove useful in establishing hormone receptor status in prostate cancer tissue biopsy specimens, providing a new tool for establishing potential endocrine responsiveness of prostate tumors. The availability of AR antibody and DNA probes also makes possible new approaches for detecting defects in the hAR gene and for defining structure/function relationships of the receptor protein.

MATERIALS AND METHODS

Isolation and Analysis of AR cDNA and Genomic Clones

A 38-bp synthetic oligonucleotide (oligo B, overlined in Fig. 2) was used to screen both a human fibroblast and a human

epididymal cDNA library (1). One cDNA clone was isolated from each, ARHFL1 and ARHEL1 (see Fig. 1). All other AR cDNA clones were isolated from an oligo(dT)-primed, bacteriophage lambda gt11, human epididymal cDNA library prepared as described previously (1, 58, 59). The genomic clone containing ARXGL2 (Fig. 1) was isolated from a bacteriophage lambda Charon 35, chromosome-sorted, X genomic library (partial Sau3A, XLAOXNLO1) obtained from L. L. Deaver at Los Alamos National Laboratory in Los Alamos, NM. (The X genomic library can now be obtained from the American Type Culture Collection, Rockville, MD). The AR genomic probe used to screen the human epididymal cDNA library was a 2.7-kbp *Sst*I fragment containing ARXGL2 and additional upstream sequence. Libraries were screened (60) with random-primed (Boehringer Mannheim, Indianapolis, IN) ^{32}P-labeled human DNA probes as described in the text. Hybridization (61) in 50% formamide was for 16 h at 42 C. Stringency washing was with 0.1× SSC (1× SSC, 0.15 M NaCl and 0.015 M sodium citrate) with 0.5% sodium dodecyl sulfate at 65 C for 1 h. The positive clones were isolated (62), subcloned in both orientations into M13, and sequenced by the dideoxy chain termination method (63). Oligonucleotides 17–20 bases in length were synthesized for use as sequencing primers to complete the sequence. Oligonucleotide primers were prepared for approximately 300 nucleotide intervals. Two independent M13 subclones of each fragment were sequenced on both strands. Individual AR DNA sequences were merged and compared with other sequences using the MicroGenie computer program (Beckman Instruments Inc., Palo Alto, CA).

Northern Hybridization Analysis

Human prostate poly(A) RNA was isolated (64, 65) from a patient with benign prostatic hyperplasia. RNA was treated with glyoxal (66) and fractionated on an agarose gel for blot hybridization (67). The protocol was modified by using Gene Screen Plus (New England Nuclear, Boston, MA) as the transfer membrane and 0.5% sodium dodecyl sulfate in the hybridization solutions. The filter was hybridized with the 2.7-kbp random-primed, ^{32}P-labeled, SstI genomic probe partially depicted in Fig. 1. ^{32}P-Labeled lambda *Hin*dIII and Phi X *Hae*III size markers were analyzed in parallel lanes after end labeling with ^{32}P-ATP using DNA polymerase I Klenow fragment.

Immunocytochemistry

Rabbit polyclonal antibodies were raised against an AR-specific peptide (underlined in Fig. 2) by a previously described method (23, 68). The antibody was shown to react with native AR (23). Prostate tissue was obtained from a patient with benign prostatic hyperplasia. Frozen sections (6 μm) were fixed in 4% paraformaldehyde, incubated with the immunoglobulin G fraction of immune serum (10 μg/ml) after purification on a protein-A agarose column (69) and stained by the avidin-biotin peroxidase method (70) using diaminobenzidine tetrahydrochloride as a chromogen.

Acknowledgments

We thank D. M. Fowlkes for oligonucleotide synthesis, D. G. Klapper for peptide synthesis, R. C. Jambou for assistance with the hydropathy plot, and M. O. DeShazo for typing the manuscript.

Received July 20, 1988. Accepted August 22, 1988.

Address requests for reprints to: Dr. Elizabeth M. Wilson, Laboratories for Reproductive Biology, CB#7500 MacNider Building, 202H, University of North Carolina, Chapel Hill, North Carolina 27599.

Supported by Grants HD-16910, HD-04466, HD-21744, and P30-HD-18968 (recombinant DNA and histochemistry cores) from the National Institute of Child Health and Human Development Center for Population Research, and by NIH Grants NS-17479, T32-DK-07219, and by the Andrew W. Mellon Foundation.

REFERENCES

1. Lubahn DB, Joseph DR, Sullivan PM, Willard HF, French FS, Wilson EM 1988 Cloning of human androgen receptor complementary DNA and localization to the X chromosome. Science 240:327–330
2. Chang C, Kokontis J, Liao S 1988 Molecular cloning of human and rat complementary DNA encoding androgen receptors. Science 240:324–326
3. Evans RM 1988 The steroid and thyroid hormone receptor superfamily. Science 240:889–895
4. Wilson JD, Griffin JE, Leshin M, MacDonald PC 1983 The androgen resistance syndromes: 5α-reductase deficiency, testicular feminization, and related disorders. In: Stanbury JB, Wyngaarden JB, Fredrickson DS, Goldstein JL, Brown MS (eds) The Metabolic Basis of Inherited Disease. McGraw-Hill, Inc, New York, pp 1001–1026
5. Brown TR, Migeon CJ 1987 Androgen insensitivity syndromes: paradox of phenotypic feminization with male genotype and normal testicular androgen secretion. In: Hormone Resistance and other Endocrine Paradoxes. Cohen MP, Foa PP (eds) Springer-Verlag, New York, pp 157–203
6. Hansson V, Ritzen EM, French FS, Nayfeh SN 1975 Androgen transport and receptor mechanisms in testis and epididymis. In: Greep RO, Astwood EB, Hamilton DW, Geiger SR (eds) Endocrinology, vol V. Male Reproductive System. Handbook of Physiology, American Physiological Society, Washington, DC, pp 173–201
7. Barrack ER, Tindall DJ 1987 A critical evaluation of the use of androgen receptor assays to predict the androgen responsiveness of prostatic cancer. In: Coffey DS, Bruchovsky N, Gardner Jr WA, Resnick MI, Karr JP (eds) Current Concepts and Approaches to the Study of Prostate Cancer. Alan R Liss Inc, New York, pp 155–187
8. Hollenberg SM, Weinberger C, Ong ES, Cerelli G, Oro A, Lebo R, Thompson EB, Rosenfeld MG, Evans RM 1985 Primary structure and expression of a functional human glucocorticoid receptor cDNA. Nature 318:635–641
9. Greene GL, Gilna P, Waterfield M, Baker A, Hort Y, Shine J 1986 Sequence and expression of human estrogen receptor complementary DNA. Science 231:1150–1154
10. Green S, Walter P, Kumar V, Krust A, Bornert J, Argos P, Chambon P 1986 Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. Nature 320:134–139
11. Misrahi M, Atger M, d'Auriol L, Loosfelt H, Meriel C, Fridlansky F, Guiochon-Mantel A, Galibert F, Milgrom E 1987 Complete amino acid sequence of the human progesterone receptor deduced from cloned cDNA. Biochem Biophys Res Commun 143:740–748
12. Arriza JL, Weinberger C, Cerelli G, Glaser TM, Handelin BL, Housman DE, Evans RM 1987 Cloning of human mineralocorticoid receptor complementary DNA: structural and functional kinship with the glucocorticoid receptor. Science 237:268–275
13. Baker AR, McDonnell DP, Hughes M, Crisp TM, Mangelsdorf DJ, Haussler MR, Pike JW, Shine J, O'Malley BW 1988 Cloning and expression of full-length cDNA encoding human vitamin D receptor. Proc Natl Acad Sci USA 85:3294–3298
14. Weinberger C, Thompson CC, Ong ES, Lebo R, Gruol DJ, Evans RM 1986 The c-erb-A gene encodes a thyroid hormone receptor. Nature 324:641–646
15. Benbrook D, Pfahl M 1987 A novel thyroid hormone receptor encoded by a cDNA clone from a human testis library. Science 238:788–791

16. Petkovich M, Brand N, Krust A, Chambon P 1987 A human retinoic acid receptor which belongs to the family of nuclear receptors. Nature 330:444–450
17. Giguere V, Ong ES, Segui P, Evans RM 1987 Identification of a receptor for the morphogen retinoic acid. Nature 330:624–629
18. Brand N, Petkovich M, Krust A, Chambon P, de The H, Marchio A, Tiollais P, Dejean A 1988 Identification of a second human retinoic acid receptor. Nature 332:850–853
19. Benbrook D, Lernhardt E, Pfahl M 1988 A new retinoic acid receptor identified from a hepatocellular carcinoma. Nature 333:669–672
20. Giguere V, Yang N, Segui P, Evans R 1988 Identification of a new class of steroid hormone receptors. Nature 331:91–94
21. Trapman J, Klaassen P, Kuiper GGJM, van der Korput JAGM, 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
22. Huckaby C, Conneely OM, Beattie WG, Dobson ADW, Tsai M, O'Malley B 1987 Structure of the chromosomal chicken progesterone receptor gene. Proc Natl Acad Sci USA 84:8380–8384
23. Tan J, Joseph DR, Quarmby VE, Lubahn DB, Sar M, French FS, Wilson EM 1988 The rat androgen receptor: primary structure, autoregulation of its messenger ribonucleic acid, and immunocytochemical localization of the receptor protein. Mol Endocrinol 2:1276–1285
24. Evans RM, Hollenberg SM 1988 Zinc fingers: gilt by association. Cell 52:1–3
25. Danielsen M, Northrop JP, Jonklaas J, Ringold GM 1987 Domains of the glucocorticoid receptor involved in specific and nonspecific deoxyribonucleic acid binding, hormone activation, and transcriptional enhancement. Mol Endocrinol 1:816–822
26. Beckman WC, Mickey DD, Fried FA 1985 Autoradiographic localization of estrogen and androgen target cells in human and rat prostate carcinoma. J Urol 133:724–728
27. Peters CA, Barrack ER, Morphologic localization of androgen receptors in human prostate: a new method of steroid receptor autoradiography. Program of the 67th Annual Meeting of The Endocrine Society, Baltimore, MD, 1985, p 108 (Abstract)
28. Peters CA, Barrack ER 1987 A new method for labeling and autoradiographic localization of androgen receptors. J Histochem Cytochem 35:755–762
29. Korach KS, Horigome T, Tomooka Y, Yamashita S, Newbold R, McLachlan JA 1988 Immunodetection of estrogen receptor in epithelial and stromal tissues of neonatal mouse uterus. Proc Natl Acad Sci USA 85:3334–3337
30. Bigsby RM, Cunha GR 1986 Estrogen stimulation of deoxyribonucleic acid synthesis in uterine epithelial cells which lack estrogen receptors. Endocrinology 119:390–396
31. Bruner-Lorand J, Mechaber D, Zwick A, Hechter O, Eychenne B, Baulieu EE, Robel P 1984 Characteristics of separated epithelial and stromal subfractions of prostate: I. Rat ventral prostate. Prostate 5:231–254
32. Cunha GR, Lung B 1979 The importance of stroma in morphogenesis and functional activity of urogenital epithelium. In Vitro 15:50–71
33. Bruchovsky N, Wilson JD 1968 The intranuclear binding of testosterone and 5α-androstan-17β-ol-3-one by rat prostate. J Biol Chem 243:5953–5960
34. Sar M, Liao S, Stumpf WE 1970 Nuclear concentration of androgens in rat seminal vesicles and prostate demonstrated by dry-mount autoradiography. Endocrinology 86:1008–1011
35. Wilson EM, French FS 1979 Effects of proteases and protease inhibitors on the 4.5S and 8S androgen receptor. J Biol Chem 254:6310–6319
36. Kozak M 1986 Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44:283–292
37. Conneely OM,Dobson ADW, Tsai M, Beattie WG, Toft DO, Huckaby CS, Zarucki T, Schrader WT, O'Malley BW 1987 Sequence and expression of a functional chicken progesterone receptor. Mol Endocrinol 1:517–525
38. Pratt WB, Jolly DJ, Pratt DV, Hollenberg SM, Giguere V, Cadepond FM, Schweizer-Groyer G, Catelli M-G, Evans RM, Baulieu E-E 1988 A region in the steroid binding domain determines formation of the non-DNA-binding, 9S glucocorticoid receptor complex. J Biol Chem. 263:267–273
39. Kumar V, Green S, Stack G, Berry M, Jin J-R, Chambon P 1987 Functional domains of the human estrogen receptor. Cell 51:941–951
40. Giguere V, Hollenberg SM, Rosenfeld MG, Evans R 1986 Functional domains of the human glucocorticoid receptor. Cell 46:645–652
41. Danielsen M, Northrop J, Ringold G 1986 The mouse glucocorticoid receptor: mapping of functional domains by cloning, sequencing and expression of wild-type and mutant receptor proteins. EMBO J J 5:2513–2522
42. Joab I, Radanyi C, Renoir M, Buchou T, Catelli M-G, Binart N, Mester J, Baulieu E-E 1984 Common non-hormone binding component in non-transformed chick oviduct receptors of four steroid hormones. Nature 308:850–853
43. Catelli MG, Binart N, Jung-Testas I, Renoir JM, Baulieu EE, Feramisco JR, Welch WJ 1985 The common 90-kd protein component of non-transformed '8S'steroid receptors is a heat-shock protein. EMBO J 4:3131–3135
44. Colvard DS, Wilson EM 1981 Identification of an 8S androgen receptor-promoting factor that converts the 4.5S form of the androgen receptor to 8S. Endocrinology 109:496–504
45. Colvard DS, Wilson EM 1984 Androgen receptor binding to nuclear matrix *in vitro* and its inhibition by 8S androgen receptor-promoting factor. Biochemistry 23:3479–3486
46. Wilson EM 1985 Interconversion of androgen receptor forms by divalent cations and 8S androgen receptor-promoting factor effects of Zn^{2+}, Cd^{2+}, Ca^{2+}, and Mg^{2+}. J Biol Chem 260:8683–8689
47. Wilson EM, Colvard DS 1984 Factors that influence the interaction of androgen receptors with nuclei and nuclear matrix. In: Catt KJ, Dufau ML (eds) Hormone Action and Testicular Function. Annals of the New York Academy of Science, New York, 438:85–100
48. Loosfelt H, Atger M, Misrahi M, Guiochon-Mantel A, Meriel C, Logeat F, Benarous R, Milgrom E 1986 Cloning and sequence analysis of rabbit progesterone-receptor complementary DNA. Proc Natl Acad Sci USA 83:9045–9049
49. Koike S, Sakai M, Muramatsu M 1987 Molecular cloning and characterization of rat estrogen receptor cDNA. Nucleic Acid Res 15:2499–2513
50. White R, Lees JA, Needham M, Ham J, Parker M 1987 Structural organization and expression of the mouse estrogen receptor. Mol Endocrinol 1:735–744
51. Krust A, Green S, Argos P, Kumar V, Walter P, Bornert J-M, Chambon P 1986 The chicken oestrogen receptor sequence: homology with v-erbA and the human oestrogen and glucocorticoid receptors. EMBO J 5:891–897
52. Weiler IJ, Lew D, Shaprio DJ 1987 The Xenopus Laevis estrogen receptor: sequence homology with human and avian receptors and identification of multiple estrogen receptor messenger ribonucleic acids. Mol Endocrinol 1:355–362
53. Miesfeld R, Rusconi S, Godowski PJ, Maler BA, Okret S, Wikstrom A-C, Gustafsson J-A, Yamamoto KR 1986 Genetic complementation of a glucocorticoid receptor deficiency by expression of cloned receptor cDNA. Cell 46:389–399
54. Davies P, Roach A, Hew C-L 1982 DNA sequence coding

for an antifreeze protein precursor from winter flounder. Proc Natl Acad Sci USA 79:335–339
55. Grima B, Lamouroux A, Blanot F, Biguet NF, Mallet J 1985 Complete coding sequence of rat tyrosine hydroxylase mRNA. Proc Natl Acad Sci USA 82:617–621
56. Wharton KA, Yedvobnick B, Finnerty VG, Artavanis-Tsakonas SA 1985 Opa: a novel family of transcribed repeats shared by the notch locus and other developmentally regulated loci in D. melanogaster. Cell 40:55–62
57. Poole SJ, Kauvar LM, Drees B, Kornberg T 1985 The engrailed locus of Drosophila: structural analysis of an embryonic transcript. Cell 40:37–43
58. Young RA, Davis RW 1983 Efficient isolation of genes by using antibody probes. Proc Natl Acad Sci USA 80:1194–1198
59. Gubler U, Hoffman BJ 1983 A simple and very efficient method for generating cDNA libraries. Gene 25:263–269
60. Benton WD, Davis RW 1977 Screening lambda gt recombinant clones by hybridization to single plaques in situ. Science 196:180–182
61. Wahl GM, Stern M, Stark GR 1979 Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulfate. Proc Natl Acad Sci USA 76:3683–3687
62. Kaslow DC 1986 A rapid biochemical method for purifying lambda DNA from phase lysates. Nucleic Acids Res 14:6767
63. Sanger F, Nicklen S, Coulson AR 1977 DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467
64. Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ 1979 Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294–5299
65. Aviv H, Leder P 1972 Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc Natl Acad Sci USA 69:1408–1412
66. Carmichael GG, McMaster GK 1980 The analysis of nucleic acids in gels using glyoxal and acridine orange. Methods Enzymol 65:380–391
67. Thomas PS 1980 Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc Natl Acad Sci USA 77:5201–5205
68. Smith DF, Lubahn DB, McCormick DJ, Wilson EM, Toft DO 1988 The production of antibodies against the conserved cysteine region of steroid receptors and their use in characterizing the avian progesterone receptor. Endocrinology 122:2816–2825
69. Goding JW 1976 Conjugation of antibodies with fluorochromes: modifications to the standard methods. J Immunol Methods 13:215–226
70. Sar M 1985 Application of avidin-biotin complex technique to the localization of estradiol receptor in target tissues using monoclonal antibodies. In: Bullock GR, Petrusz P (eds) Techniques in Immunocytochemistry. Academic Press, New York, vol 3:43–54
71. Robertson M 1987 Towards a biochemistry of morphogenesis. Nature 330:420–421
72. Devereux I, Haeberli P, Smithies O 1984 A comprehensive set of sequence analysis programs for the VAX. Nucleic Acid Res 12:387–395
73. Kyte J, Doolittle RF 1982 A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132