



Office de la propriété
intellectuelle
du Canada

un organisme
d'Industrie Canada

Canadian
Intellectual Property
Office

An Agency of
Industry Canada

(11) © **1,341,347**

(21) 596,763

(22) 1989/04/14

(45) 2002/03/12

(52) 195-1.22
C.L. CR 195-1.235
530-13

(51) Int.Cl. ⁷ C12N 15/12; C07K 14/205; G01N 33/566; C12Q 1/68

(19) (CA) **CANADIAN PATENT** (12)

(54) **Androgen Receptor Proteins, Recombinant DNA Molecules Coding for Such,
and Use of Such Compositions**

(72) French, Frank S., U.S.A.
Wilson, Elizabeth M., U.S.A.
Joseph, David R., U.S.A.
Lubahn, Dennis B., U.S.A.

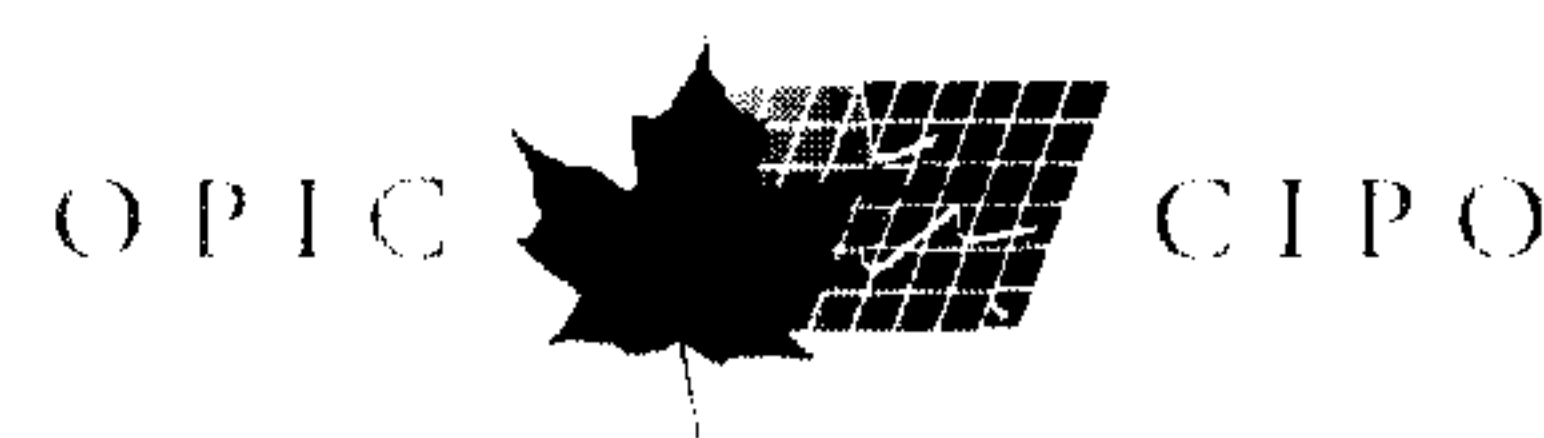
(73) University of North Carolina at Chapel Hill, U.S.A.

(30) (US) U.S.A. 182,646 1988/04/14

(57) 48 Claims

Canada

http://big.leg.gc.ca • Ottawa-Hull K1A 0C9 • *http://cipo.gc.ca*
OPIC • CIPQ 191



ABSTRACT OF THE DISCLOSURE

DNA sequences encoding human androgen receptor protein and polypeptides and proteins having substantially the same biological activity as human androgen receptor protein and the amino acid sequences of human androgen receptor protein and polypeptides and proteins having substantially the same biological activity as human androgen receptor protein are disclosed. Methods for the production and use of such compositions are also disclosed.

10

15

20

25

30

ANDROGEN RECEPTOR PROTEINS, RECOMBINANT DNA MOLECULES CODING
FOR SUCH, AND USE OF SUCH COMPOSITIONS

5 This invention was made in the course of research supported in part by grants from the National Institutes of Health (NIH HD 16910, HD 04466, and HD 18968).

TECHNICAL FIELD OF THE INVENTION

10 This invention relates to recombinant DNA molecules and their expression products. More specifically this invention relates to recombinant DNA molecules coding for androgen receptor protein, androgen receptor protein, and use of the DNA molecules and protein in investigatory, diagnostic and therapeutic applications.

15

BACKGROUND OF THE INVENTION

20 The naturally occurring androgenic hormones, testosterone and its 5 -reduced metabolite, dihydrotestosterone, are synthesized by the Leydig cells of the testes and circulate throughout the body where they diffuse into cells and bind to the androgen receptor protein ("AR"). Androgens, acting through their receptor, stimulate development of the male genitalia and accessory sex glands in the fetus, virilization and growth in the pubertal male, and maintenance of male virility and reproductive function in the adult. The

25 androgen receptor, together with other steroid hormone receptors constitute a family of trans-acting transcriptional regulatory proteins that control gene transcription through interactions with specific gene sequences.

30 When prostate cancer is found to be confined to the prostate gland, the treatment of choice is surgical removal. However, 50 to 80% of prostate cancer patients already have metastases at the time of diagnosis. Most of their tumors (70 to 80%) respond to the removal of androgen by castration or by suppression of luteinizing hormone secretion by the pituitary gland using a gonadotropin

releasing hormone analogue alone or in combination with an anti-androgen. The degree and duration of response to this treatment is highly variable (10% live < 6 months, 50% live < 3 years, and 10% live > 10 years.) Initially cancer cells regress
5 without androgen stimulation, but ultimately the growth of androgen independent tumor cells continues (35). At present it is not possible to predict on an individual basis which patient will respond to hormonal therapy and for how long. If poorly responsive patients could be identified early, they could be treated by
10 alternative forms of therapy (e.g. chemotherapy) at an earlier stage when they might be more likely to respond.

Studies on androgen receptors in prostate cancer have suggested that a positive correlation may exist between the presence of androgen receptors in cancer cells and their dependence on
15 androgenic hormone stimulation for growth. (An analogous situation exists in mammary carcinoma where there is a correlation between estrogen receptors and regression of the tumor in response to estrogen withdrawal). However, methodological problems in the measurement of androgen receptors have prevented the routine use of
20 androgen receptor assays in the diagnostic evaluation of prostate cancer. Prior to our preparation of androgen receptor antibodies, all androgen receptor assays were based on the binding of [³H]-labeled androgen. These assays have been unreliable in human prostate cancer tissue because of the extreme lability of the
25 androgen binding site and the presence of unlabeled androgen in the tissue. Endogenous androgen occupies the binding site on the receptor and dissociates very slowly (t 1/2 24-48 hr at 0C). A further problem is that biopsy samples are quite small, making it difficult to obtain sufficient tissue for [³H]-androgen binding
30 assays. Moreover, prostate cancer is heterogenous with respect to cell types. Thus within a single biopsy sample there is likely to be an uneven distribution of cells containing androgen receptors.

Development of the male phenotype and maturation of male reproductive function are dependent on the interaction of androgenic

hormones with the androgen receptor protein and the subsequent function of the receptor as a trans-acting inducer of gene expression. It has become well established over the past twenty-five years that genetic defects of the androgen receptor
5 result in a broad spectrum of developmental and functional abnormalities ranging from genetic males (46,XY) with female phenotype to phenotypically normal males with infertility. Isolation of the structural gene for the androgen receptor makes it possible to define the nature of these genomic defects in molecular
10 terms. Analysis of the functional correlates of the genetic defects may lead to a better understanding of the regulation of androgen receptor gene expression and of the mechanism of androgen action in male sexual development and function.

The androgen insensitivity syndrome, known also as testicular
15 feminization, is characterized by an inability to respond to androgen due to a defect in the androgen receptor, the protein that mediates the action of androgen within the cell. Androgen insensitivity is an inherited X-linked trait that occurs in both complete and incomplete forms. The complete form results in failure
20 of male sex differentiation during embryogenesis and absence of virilization at puberty. The result is a 46,XY genetic male with testes and male internal ducts. The testes produce normal amounts of testosterone and Mullerian inhibiting substance. Consequently development of the uterus is inhibited as in the normal male.
25 Because of the inability to respond to androgen, the external genitalia remain in the female phenotype with normal clitoris and labia. A small vagina develops from the urogenital sinus and ends in a blind pouch. At puberty feminization with breast development and female contours occur in response to testicular estrogen,
30 however, there is no growth of sexual hair even though circulating testosterone concentrations are equal to or greater than levels in the normal male.

Incomplete forms of the androgen insensitivity syndrome include a spectrum of phenotypes resulting from varying degrees of

incomplete androgen responsiveness. At one extreme, individuals have mild enlargement of the clitoris and sparse pubic hair. The opposite extreme is characterized by more complete masculinization with varying degrees of hypospadias deformity but predominantly a male phenotype. It has been reported that some adult men with severe oligospermia or azoospermia who are otherwise normal, have defects in the androgen receptor. These may include as many as 10% of infertile males.

The genetic defect eliciting this range of abnormalities is thought to be a single biochemical event at the level of the gene for the androgen receptor. The androgen receptor is a high affinity androgen binding protein that mediates the effects of testosterone and dihydrotestosterone by functioning as a trans-acting inducer of gene expression. For proper male sexual development to occur, there is a requirement for androgen and its receptor at a critical time during embryogenesis and during puberty. The majority of individuals with the androgen insensitivity syndrome have a history of affected family members, although about a third are thought to represent new mutations of this X-linked disorder. The incidence ranges from 1 in 20,000 to 60,000 male births.

In studies of families with clinical evidence of the androgen insensitivity syndrome, four main categories were recognized that range from the most severe, complete absence of receptor binding activity in a genetic male with female phenotype, to qualitatively normal receptor in affected individuals. Second in severity are affected individuals with qualitatively abnormal androgen binding by receptor present in normal levels. Examples include the failure of sodium molybdate (a reagent often used in studies on steroid receptors) to stabilize the receptor of affected individuals when molybdate is known to stabilize the wild-type receptor. Lability of the receptor under conditions that normally cause transformation has also been reported. A third group expresses a decreased amount of receptor with wild-type in vitro binding characteristics. The final grouping contains those androgen insensitivity patients in

whom no abnormality in receptor is detected. In a recent study of this form of the syndrome, the androgen receptor was as capable of binding oligonucleotides as the wild-type receptor. Indeed, with the techniques available until only recently, it has been difficult
5 in certain cases to document an androgen receptor defect in affected individuals.

Experimental methods used in assessing receptor defects in the past have relied on the ability of receptor to bind androgen with high affinity. The limitation of this methodology is that it is not
10 possible to distinguish between the lack of expression of the receptor and loss of androgen binding activity. An example of how inadequate methodology complicates diagnosis is the absence of detectable receptor binding activity in patients who are partially virilized. It is theoretically possible for a mutation to occur
15 which allows the receptor with defective androgen binding activity to induce gene transcription. Biologically active truncated forms of the glucocorticoid receptor that lack steroid binding activity but retain the DNA binding domain have been demonstrated using genetically engineered mutants.

20 Purification of the androgen receptor has been difficult to accomplish due to its low concentration and high degree of instability. Reported attempts at purification using either conventional methods of column chromatography or steroid-affinity chromatography have yielded insufficient amounts of receptor protein
25 to allow even the preparation of monoclonal antibodies.

An early report on the partial purification of the androgen receptor was disclosed by Mainwaring et al. in "The use of DNA - cellulose chromatography and isoelectric focusing for the characterization and partial purification of steroid-receptor
30 complexes," Biochem J, 134, 113-127 (1973). They used DNA-cellulose chromatography and isoelectric focusing to isolate the receptor from rat ventral prostate and determined its physiochemical properties. This group was among the first to attempt the use of steroid affinity chromatography in conjunction with conventional

chromatography, using the affinity label
17B-bromoacetoxytestosterone in receptor purification (See
Mainwaring et al., "Use of the affinity label
17B-bromoacetoxytestosterone in the purification of androgen
5 receptor proteins," Perspectives in Steroid Receptor Research,
(1980)). Partial purification of androgen receptor has also been
attempted from other tissue sources, such as ram seminal vesicles
(See Foekens et al., Molecular Cellular Endocr, 23, 173-186 (1981)
and Foekens et al., "Purification of the androgen receptor of sheep
10 seminal vesicles," Biochem Biophys Res Comm, 104, 1279-1286
(1982)). The partially purified receptor displayed characteristics
of a proteolyzed receptor, but a purification of 2,000 fold was
reported with a recovery of 33% (See Foekens et al., "Purification
of the androgen receptor of sheep seminal vesicles," Biochem Biophys
15 Res Comm, 104, 1279-1286 (1982)). Later attempts at purification
continued to combine steroid affinity chromatography with
conventional techniques, reportedly achieving significant
purification, but recoveries too low for further analysis (See Chang
et al., "Purification and characterization of androgen receptor from
20 steer seminal vesicle," Biochemistry 21, 4102-4109 (1982), Chang et
al., "Purification and characterization of the androgen receptor
from rat ventral prostate," Biochemistry 22, 6170-6175 (1983) and
Chang et al., "Affinity labeling of the androgen receptor in rat
prostate cytosol with
25 17B-[(bromoacetyl)oxy]-5-alpha-androstan-3-one," Biochemistry 23,
2527-2533 (1984)). More recent studies examine the effectiveness of
a variety of immobilized androgens for their ability to bind the
androgen receptor (See De Larminat et al., "Synthesis and evaluation
of immobilized androgens for affinity chromatography in the
30 purification of nuclear androgen receptor," The Prostate 5, 123-140
(1984) and Bruchovsky et al, "Chemical demonstration of nuclear
androgen receptor following affinity chromatography with immobilized
ligands," The Prostate 10, 207-222 (1987)). Despite these efforts,
the receptor has not been purified to homogeneity and

the quantities of purified androgen receptor obtained have been insufficient for the production of antisera.

Clinical assays for the androgen receptor now include several methods. The most common is the binding of tritium-labeled hormone and measurement of binding using a charcoal adsorption assay. 5 Either a natural androgen, such as dihydrotestosterone, or synthetic androgen, such as mibolerone or methyltrienolone (R1881), can be used. An advantage of the latter in human tissue is that it is not significantly metabolized and does not bind to the serum androgen binding protein, sex steroid binding globulin. A limitation of 10 radioisotope labeling of receptor is interference caused by endogenous androgen. Although exchange assays for the androgen receptor have been described (See Carroll et al., J Steroid Biochem 21, 353-359 (1984) and Traish et al., J Steroid Biochem 23, 405-413 (1985)), their effectiveness is limited by the slow kinetics of 15 dissociation of the endogenous receptor-bound androgen.

Another method used to assess receptor status is autoradiography. In this method disclosed in Barrack et al., "Current concepts and approaches to the study of prostate cancer," 20 Progress in Clinical and Biological Research, 239, 155-187 (1987) the radioactively labeled androgen is incubated with slide-mounted tissue sections of small tissue biopsy specimens which are then frozen, sectioned and fixed. Nuclear localization of radioactivity is detected by exposure of tissue sections to x-ray film. This 25 technique requires considerable technical expertise, is labor intensive, and requires extended periods of exposure time. It is therefore of limited usefulness in general clinical assays. Another problem is high levels of background signal, i.e. a high noise/signal ratio, making it difficult to distinguish 30 receptor-bound nuclear radioactivity from unbound radioactivity distributed throughout the cells.

WU 87/05049 (Shine) discloses a method for the production of purified steroid receptor proteins, specifically estrogen receptor proteins, through the expression of recombinant DNA encoding for

such proteins in eukaryotic host cells. However, the reference does not disclose the sequence for androgen receptor protein, nor does it disclose a method for obtaining such a sequence.

5 SUMMARY OF THE INVENTION

The present invention provides a DNA sequence characterized by a structural gene coding for a polypeptide having substantially the same biological activity as androgen receptor protein. A DNA sequence
10 encoding androgen receptor protein or a protein having substantially the same biological activity as androgen receptor activity is also provided. DNA sequences may be obtained from cDNA or genomic DNA, or prepared using DNA synthesis techniques.

15 The invention further discloses cloning vehicles comprising a DNA sequence comprising a structural gene encoding a polypeptide having substantially the same biological activity as androgen receptor protein. Cloning vehicles comprising a DNA sequence encoding
20 androgen receptor protein or a protein having substantially the same biological activity as androgen receptor protein is also provided. The cloning vehicles further comprise a promoter sequence upstream of and operatively linked to the DNA sequence. In general the
25 cloning vehicles will also contain a selectable marker, and, depending on the host cell used, may contain such elements as regulatory sequences, polyadenylation signals, enhancers and RNA splice sites.

The invention further provides cells transfected or
30 transformed to produce androgen receptor protein or a protein having substantially the same biological activity as androgen receptor protein.

A further aspect of the invention provides a purified androgen receptor protein and purified
35 polypeptides and proteins having substantially the same

biological activity as androgen receptor activity, and methods for producing such proteins and polypeptides.

In defining various aspects of the invention, reference will be made to sequences of the Figures and in particular Figure 4, the detailed description of which follows under the heading "Brief Description of the Drawings".

In accordance with one embodiment of the present invention, a purified nucleic acid molecule comprises a nucleotide sequence encoding an androgen receptor protein having the amino acid sequence of Figure 4, or a fragment of the amino acid sequence which retains androgen receptor activity.

In accordance with a further embodiment of the present invention, a purified nucleic acid molecule comprises a nucleotide sequence encoding an androgen receptor protein having the amino acid sequence of Figure 5 or a fragment of the amino acid sequence which retains androgen receptor activity

In accordance with a further embodiment of the present invention, a purified nucleic acid molecule comprises a nucleotide sequence encoding a DNA-binding amino acid sequence C L I C G D E A S G C H Y G A L T C G S C K V F F K R A A E G K Q K Y L C A S R N D C T I D K F R R K N C P S C R L R K C Y E A G M.

In accordance with a further embodiment of the present invention, a purified nucleic acid molecule comprises nucleotides 2235 to 3119 of Figure 4.

In accordance with a further embodiment of the present invention, a purified nucleic acid molecule comprises a nucleotide sequence encoding an immunogenic fragment of the amino acid sequence of Figure 4.

In accordance with a further embodiment of the present invention, a purified nucleic acid molecule comprises a nucleotide sequence encoding an immunogenic fragment of the amino acid sequence of Figure 5.

In accordance with a further embodiment of the present invention, a purified androgen receptor protein comprises the amino acid sequence of Figure 4 or a fragment thereof which retains androgen receptor activity.

In accordance with a further embodiment of the present invention, a purified androgen receptor protein comprises the amino acid sequence of Figure 5 or a fragment thereof which retains androgen receptor activity.

In accordance with a further embodiment of the present invention, a DNA-binding polypeptide comprises the amino acid sequence C L I C G D E A S G C H Y G A L T C G S C K V F F K R A A E G K Q K Y L C A S R N D C T I D K F R R K N C P S C R L R K C Y E A G M.

In accordance with a further embodiment of the present invention, a steroid-binding polypeptide comprises amino acids 625 to 919 of Figure 4.

In accordance with a further embodiment of the present invention, an immunogenic polypeptide comprises an immunogenic fragment of the amino acid sequence of Figure 4.

In accordance with a further embodiment of the present invention, an immunogenic polypeptide comprises an immunogenic fragment of the amino acid sequence of Figure 5.

In accordance with a further embodiment of the present invention, there is provided a method for determining whether a subject has a normal androgen receptor protein or a mutant androgen receptor protein, the method comprises the steps of:

- (a) obtaining an androgen receptor protein sample from the subject;
- (b) conducting an assay on the protein sample to detect a normal or mutant androgen receptor protein or a normal or abnormal androgen receptor protein function.

In accordance with a further embodiment of the present invention, there is provided a method for determining whether a subject has a normal androgen receptor gene or a mutant androgen receptor gene, the method comprises the steps of:

- (a) obtaining a nucleic acid sample from the subject; and
- (b) conducting an assay on the nucleic acid sample to determine the presence of a normal androgen receptor gene or a mutant androgen receptor gene.

In accordance with a further embodiment of the present invention, a human androgen receptor protein contains the sequence

```
Trp Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg
Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg Met
Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr
Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly
Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg
Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr
Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe Thr Phe Asp Leu
Leu Ile Lys Ser His Met Val Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser
Val Gln Val Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln
```

In accordance with a further embodiment of the present invention, a cDNA which encodes a human androgen receptor protein contains the sequence

```
Trp Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg
Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg Met
Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr
Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly
Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg
Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr
Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe Thr Phe Asp Leu
Leu Ile Lys Ser His Met Val Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser
Val Gln Val Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln
```

In accordance with a further embodiment of the present invention, a replicable expression vector comprises a cDNA which encodes a human androgen receptor protein contains the sequence

Trp Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg
Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg Met
Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr
Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly
Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg
Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr
Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe Thr Phe Asp Leu
Leu Ile Lys Ser His Met Val Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser
Val Gln Val Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln

In accordance with a further embodiment of the present invention, a host cell transformed with a replicable expression vector comprises a cDNA which encodes a human androgen receptor protein contains the sequence

Trp Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg
Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg Met
Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr
Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly
Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg
Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr
Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe Thr Phe Asp Leu
Leu Ile Lys Ser His Met Val Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser
Val Gln Val Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln

In accordance with a further embodiment of the present invention, a method of producing a human androgen receptor protein comprises:

- constructing a replicable expression vector comprising a cDNA which encodes a human androgen receptor protein containing the sequence

Trp Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg
Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg Met
Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr
Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly
Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg
Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr
Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe Thr Phe Asp Leu
Leu Ile Lys Ser His Met Val Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser
Val Gln Val Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln

- transforming a host cell with the replicable expression vector;
- culturing the host cell under conditions permitting the expression of the human androgen receptor protein;
- recovering the human androgen receptor protein from the cell culture.

In accordance with a further embodiment of the present invention, a method of detecting normal or abnormal androgen receptors in humans comprises:

- obtaining a sample containing androgen receptor proteins from a subject;
- assaying the androgen receptor proteins to determine their normality or abnormality.

In accordance with a further embodiment of the present invention, a method of detecting a normal or abnormal androgen receptor gene in a human comprises:

- obtaining a nucleic acid sample from a subject;
- assaying the sample to determine the normality or abnormality of the sequence coding for the androgen receptor.

In accordance with a further embodiment of the present invention, a rat androgen receptor protein contains the sequence

```
Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser Lys Thr Tyr Arg
Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu Ala Ile Gln Asn Pro Gly Pro Arg
His Pro Glu Ala Ala Ser Ile Ala Pro Pro Gly Ala Cys Leu Gln Gln Arg Gln Glu Thr
Ser Pro Arg Arg Arg Arg Arg Gln Gln His Pro Glu Asp Gly Ser Pro Gln Ala His Ile
Arg Gly Thr Thr Gly Tyr Leu Ala Leu Glu Glu Glu Gln Gln Pro Ser Gln Gln Gln Ser
Ala Ser Glu Gly His Pro Glu Ser Gly Cys Leu Pro Glu Pro Gly Ala Ala Thr Ala Pro
Gly Lys Gly Leu Pro Gln Gln Pro Pro Ala Pro Pro Asp Gln Asp Asp Ser Ala Ala Pro
Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro Gly Leu Ser Ser Cys Ser Ala Asp Ile
Lys Asp Ile Leu Ser Glu Ala Gly Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Gln Gln
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
Gly Ser Ser Ser Val Arg Ala Arg Glu Ala Thr Gly Ala Pro Ser Ser Ser Lys Asp Ser
Tyr Leu Gly Gly Asn Ser Thr Ile Ser Asp Ser Ala Lys Glu Leu Cys Lys Ala Val Ser
Val Ser Met Gly Leu Gly Val Glu Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg
Gly Asp Cys Met Tyr Ala Ser Leu Leu Gly Gly Pro Pro Ala Val Arg Pro Thr Pro Cys
Ala Pro Leu Ala Glu Cys Lys Gly Leu Ser Leu Asp Glu Gly Pro Gly Lys Gly Thr Glu
Glu Thr Ala Glu Tyr Ser Ser Phe Lys Gly Gly Tyr Ala Lys Gly Leu Glu Gly Glu Ser
Leu Gly Cys Ser Gly Ser Ser Glu Ala Gly Ser Ser Gly Thr Leu Glu Ile Pro Ser Ser
Leu Ser Leu Tyr Lys Ser Gly Ala Val Asp Glu Ala Ala Ala Tyr Gln Asn Arg Asp Tyr
Tyr Asn Phe Pro Leu Ala Leu Ser Gly Pro Pro His Pro Pro Pro Pro Thr His Pro His
```

In accordance with a further embodiment of the present invention, a cDNA which encodes a rat androgen receptor protein contains the sequence

Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser Lys Thr Tyr Arg
Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu Ala Ile Gln Asn Pro Gly Pro Arg
His Pro Glu Ala Ala Ser Ile Ala Pro Pro Gly Ala Cys Leu Gln Gln Arg Gln Glu Thr
Ser Pro Arg Arg Arg Arg Arg Gln Gln His Pro Glu Asp Gly Ser Pro Gln Ala His Ile
Arg Gly Thr Thr Gly Tyr Leu Ala Leu Glu Glu Glu Gln Gln Pro Ser Gln Gln Gln Ser
Ala Ser Glu Gly His Pro Glu Ser Gly Cys Leu Pro Glu Pro Gly Ala Ala Thr Ala Pro
Gly Lys Gly Leu Pro Gln Gln Pro Pro Ala Pro Pro Asp Gln Asp Asp Ser Ala Ala Pro
Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro Gly Leu Ser Ser Cys Ser Ala Asp Ile
Lys Asp Ile Leu Ser Glu Ala Gly Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Gln
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Glu Val Ile Ser Glu
Gly Ser Ser Ser Val Arg Ala Arg Glu Ala Thr Gly Ala Pro Ser Ser Ser Lys Asp Ser
Tyr Leu Gly Gly Asn Ser Thr Ile Ser Asp Ser Ala Lys Glu Leu Cys Lys Ala Val Ser
Val Ser Met Gly Leu Gly Val Glu Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg
Gly Asp Cys Met Tyr Ala Ser Leu Leu Gly Gly Pro Pro Ala Val Arg Pro Thr Pro Cys
Ala Pro Leu Ala Glu Cys Lys Gly Leu Ser Leu Asp Glu Gly Pro Gly Lys Gly Thr Glu
Glu Thr Ala Glu Tyr Ser Ser Phe Lys Gly Gly Tyr Ala Lys Gly Leu Glu Gly Glu Ser
Leu Gly Cys Ser Gly Ser Ser Glu Ala Gly Ser Ser Gly Thr Leu Glu Ile Pro Ser Ser
Leu Ser Leu Tyr Lys Ser Gly Ala Val Asp Glu Ala Ala Ala Tyr Gln Asn Arg Asp Tyr
Tyr Asn Phe Pro Leu Ala Leu Ser Gly Pro Pro His Pro Pro Pro Pro Thr His Pro His

In accordance with a further embodiment of the present invention, a replicable expression vector comprises a cDNA which encodes a rat androgen receptor protein contains the sequence

Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser Lys Thr Tyr Arg
Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu Ala Ile Gln Asn Pro Gly Pro Arg
His Pro Glu Ala Ala Ser Ile Ala Pro Pro Gly Ala Cys Leu Gln Gln Arg Gln Glu Thr
Ser Pro Arg Arg Arg Arg Arg Gln Gln His Pro Glu Asp Gly Ser Pro Gln Ala His Ile
Arg Gly Thr Thr Gly Tyr Leu Ala Leu Glu Glu Glu Gln Gln Pro Ser Gln Gln Gln Ser
Ala Ser Glu Gly His Pro Glu Ser Gly Cys Leu Pro Glu Pro Gly Ala Ala Thr Ala Pro
Gly Lys Gly Leu Pro Gln Gln Pro Pro Ala Pro Pro Asp Gln Asp Asp Ser Ala Ala Pro
Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro Gly Leu Ser Ser Cys Ser Ala Asp Ile
Lys Asp Ile Leu Ser Glu Ala Gly Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Gln
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Glu Val Ile Ser Glu
Gly Ser Ser Ser Val Arg Ala Arg Glu Ala Thr Gly Ala Pro Ser Ser Ser Lys Asp Ser
Tyr Leu Gly Gly Asn Ser Thr Ile Ser Asp Ser Ala Lys Glu Leu Cys Lys Ala Val Ser
Val Ser Met Gly Leu Gly Val Glu Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg
Gly Asp Cys Met Tyr Ala Ser Leu Leu Gly Gly Pro Pro Ala Val Arg Pro Thr Pro Cys
Ala Pro Leu Ala Glu Cys Lys Gly Leu Ser Leu Asp Glu Gly Pro Gly Lys Gly Thr Glu
Glu Thr Ala Glu Tyr Ser Ser Phe Lys Gly Gly Tyr Ala Lys Gly Leu Glu Gly Glu Ser
Leu Gly Cys Ser Gly Ser Ser Glu Ala Gly Ser Ser Gly Thr Leu Glu Ile Pro Ser Ser
Leu Ser Leu Tyr Lys Ser Gly Ala Val Asp Glu Ala Ala Ala Tyr Gln Asn Arg Asp Tyr
Tyr Asn Phe Pro Leu Ala Leu Ser Gly Pro Pro His Pro Pro Pro Pro Thr His Pro His

In accordance with a further embodiment of the present invention, a host cell transformed with a replicable expression vector comprises cDNA which encodes a rat androgen receptor protein contains the sequence

Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser Lys Thr Tyr Arg
 Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu Ala Ile Gln Asn Pro Gly Pro Arg
 His Pro Glu Ala Ala Ser Ile Ala Pro Pro Gly Ala Cys Leu Gln Gln Arg Gln Glu Thr
 Ser Pro Arg Arg Arg Arg Arg Gln Gln His Pro Glu Asp Gly Ser Pro Gln Ala His Ile
 Arg Gly Thr Thr Gly Tyr Leu Ala Leu Glu Glu Glu Gln Gln Pro Ser Gln Gln Gln Ser
 Ala Ser Glu Gly His Pro Glu Ser Gly Cys Leu Pro Glu Pro Gly Ala Ala Thr Ala Pro
 Gly Lys Gly Leu Pro Gln Gln Pro Pro Ala Pro Pro Asp Gln Asp Asp Ser Ala Ala Pro
 Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro Gly Leu Ser Ser Cys Ser Ala Asp Ile
 Lys Asp Ile Leu Ser Glu Ala Gly Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Gln Gln
 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Glu Val Ile Ser Glu
 Gly Ser Ser Ser Val Arg Ala Arg Glu Ala Thr Gly Ala Pro Ser Ser Ser Lys Asp Ser
 Tyr Leu Gly Gly Asn Ser Thr Ile Ser Asp Ser Ala Lys Glu Leu Cys Lys Ala Val Ser
 Val Ser Met Gly Leu Gly Val Glu Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg
 Gly Asp Cys Met Tyr Ala Ser Leu Leu Gly Gly Pro Pro Ala Val Arg Pro Thr Pro Cys
 Ala Pro Leu Ala Glu Cys Lys Gly Leu Ser Leu Asp Glu Gly Pro Gly Lys Gly Thr Glu
 Glu Thr Ala Glu Tyr Ser Ser Phe Lys Gly Gly Tyr Ala Lys Gly Leu Glu Gly Glu Ser
 Leu Gly Cys Ser Gly Ser Ser Glu Ala Gly Ser Ser Gly Thr Leu Glu Ile Pro Ser Ser
 Leu Ser Leu Tyr Lys Ser Gly Ala Val Asp Glu Ala Ala Ala Tyr Gln Asn Arg Asp Tyr
 Tyr Asn Phe Pro Leu Ala Leu Ser Gly Pro Pro His Pro Pro Pro Pro Thr His Pro His

In accordance with a further embodiment of the present invention, a method of producing a rat androgen receptor protein comprises:

- constructing a replicable expression vector comprising a cDNA which encodes a rat androgen receptor protein containing the sequence

Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser Lys Thr Tyr Arg
 Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu Ala Ile Gln Asn Pro Gly Pro Arg
 His Pro Glu Ala Ala Ser Ile Ala Pro Pro Gly Ala Cys Leu Gln Gln Arg Gln Glu Thr
 Ser Pro Arg Arg Arg Arg Arg Gln Gln His Pro Glu Asp Gly Ser Pro Gln Ala His Ile
 Arg Gly Thr Thr Gly Tyr Leu Ala Leu Glu Glu Glu Gln Gln Pro Ser Gln Gln Gln Ser
 Ala Ser Glu Gly His Pro Glu Ser Gly Cys Leu Pro Glu Pro Gly Ala Ala Thr Ala Pro
 Gly Lys Gly Leu Pro Gln Gln Pro Pro Ala Pro Pro Asp Gln Asp Asp Ser Ala Ala Pro
 Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro Gly Leu Ser Ser Cys Ser Ala Asp Ile
 Lys Asp Ile Leu Ser Glu Ala Gly Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Gln Gln
 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Glu Val Ile Ser Glu
 Gly Ser Ser Ser Val Arg Ala Arg Glu Ala Thr Gly Ala Pro Ser Ser Ser Lys Asp Ser
 Tyr Leu Gly Gly Asn Ser Thr Ile Ser Asp Ser Ala Lys Glu Leu Cys Lys Ala Val Ser
 Val Ser Met Gly Leu Gly Val Glu Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg
 Gly Asp Cys Met Tyr Ala Ser Leu Leu Gly Gly Pro Pro Ala Val Arg Pro Thr Pro Cys
 Ala Pro Leu Ala Glu Cys Lys Gly Leu Ser Leu Asp Glu Gly Pro Gly Lys Gly Thr Glu
 Glu Thr Ala Glu Tyr Ser Ser Phe Lys Gly Gly Tyr Ala Lys Gly Leu Glu Gly Glu Ser
 Leu Gly Cys Ser Gly Ser Ser Glu Ala Gly Ser Ser Gly Thr Leu Glu Ile Pro Ser Ser
 Leu Ser Leu Tyr Lys Ser Gly Ala Val Asp Glu Ala Ala Ala Tyr Gln Asn Arg Asp Tyr
 Tyr Asn Phe Pro Leu Ala Leu Ser Gly Pro Pro His Pro Pro Pro Pro Thr His Pro His

- transforming a host cell with the replicable expression vector;
- culturing the host cell under conditions permitting the expression of the rat androgen receptor protein;
- recovering the rat androgen receptor protein from the cell culture.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a comparison of DNA-binding domains of the human androgen receptor (hAR) with members of the nuclear receptor family. (A) is a comparison of oligo A nucleotide sequence with sequences of hAR and other nuclear receptors: hPR, human progesterone receptor; hMR, human mineralocorticoid receptor; hGR, human glucocorticoid receptor; hER, human estrogen receptor; hT3R, human thyroid hormone receptor; hRAR, human retinoic acid receptor. Chromosomal locations are shown in parentheses at the left.

10 Nucleotide identity between oligo A and hAR is indicated with an asterisk. The percent homology with oligo A is in parentheses at the right of each sequence. (B) shows the structure of fibroblast clone ARHFL1 human fibroblast clone [1]). Nucleotide residues are numbered from the 5'-terminus. Restriction

15 endonuclease sites were determined by mapping or were deduced from DNA sequence. The TGA translation termination codon, determined by comparison with hPR, hMR and hGR, follows a long open reading frame containing sequences homologous to those of other steroid receptors. Arrows indicate exon boundaries in genomic clone X05AR.

20 The hatched area is the putative DNA binding domain. (C) shows a comparison of amino acid sequences of the AR DNA-binding domain with sequences of the nuclear receptor family. AR amino acid sequence was deduced from nucleotide sequence of clone ARHFL1 and is numbered beginning with the first conserved cysteine residue (+). Amino acid

25 numbers in parentheses at the left indicate the residue number of the first conserved cysteine from the references indicated above. Percent homology with hAR is indicated in parentheses on the right. The region of the DNA-binding domain from which the oligo A sequence was derived is underlined in hAR. Coding DNA of residues 1 to 31 is

30 contained within genomic clone X05AR. Abbreviations in addition to those described above are cVDR, chicken vitamin D receptor, and vERBA, erb A protein from avian erythroblastosis virus. Abbreviations for amino acid residues are:

- 10 -

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.

Figure 2 illustrates the steroid binding properties of expressed AR cDNA. (A) shows the structure of pCMVAR in the expression vector pCMV containing the human cytomegalovirus(CMV) promoter of the immediate early gene, poly(A) addition-transcription terminator region of the human growth hormone gene (hGH poly A),SV40 origin of replication (SV40 Ori), and a polylinker region for insertion of cDNAs. The plasmid pTEBR contains the ampicillin resistance gene (Amp). (B) shows saturation analysis of [³H]dihydrotestosterone binding in extracts of pCMVAR transfection of COS M6 cells. Portions of cytosol (0.1 ml, 0.3 mg/ml protein) were incubated overnight at 4°C with increasing concentrations of ³H-labeled hormone and analyzed by charcoal adsorption. Nonspecific binding increased from 18% to 37% of total bound radioactivity. (C) shows a scratched plot analysis of [³H]dihydrotestosterone binding. Error estimation was based on linear regression analysis (r=0.966). (D) illustrates the competition of unlabeled steroids for binding of 5 nM [³H]dihydrotestosterone in transfected COS M6 cell extracts. Unlabeled steroids were added at 10- and 100-fold excess of labeled hormone. Specific binding was determined as previously described.

Figure 3 is a compiled clone map of the human androgen receptor. The map shows the structure of the human androgen receptor gene and the relative positions of the nucleic acid sequences contained in the cDNA probes [A], [B], [C] and [D], human fibroblast clone [1], human epididymis clones [1] and [5], human genomic clones [1], [2], [3], [4] and [5], and rat epididymis clones [1] and [2].

Figure 4 shows the complete nucleotide sequence for human androgen receptor cDNA and the deduced amino acid sequence.

Figure 5 shows the complete nucleotide sequence of the rat androgen receptor cDNA and the predicted amino acid sequence.

Figure 6 is a photograph of a frozen section of rat ventral prostate stained with antibodies (AR-52-3-p) to the AR peptide
5 NH₂-Asp-His-Val-Leu-Pro-Ile-Asp-Tyr-Tyr-Phe-Pro-Pro-Gln-Lys-Thr in a dilution of 1 to 3000 using the avidin-biotin peroxidase technique. Androgen receptor is indicated by brown staining of nuclei in epithelial cells.

Figure 7 is a photograph showing restriction fragment length
10 polymorphisms in the human androgen receptor gene.

Figure 8 is a photograph showing a Southern blot analysis in the human androgen receptor gene in complete androgen insensitivity syndrome patients.

15 DETAILED DESCRIPTION OF THE INVENTION

In the description the following terms are employed:

Nucleotide

A monomeric unit of DNA or RNA consisting of a sugar moiety (pentose), a phosphate,, and a nitrogenous heterocyclic base. The
20 base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose) and that combination of base and sugar is a nucleotide. The base characterizes the nucleotide. The four DNA bases are adenine ("A"), guanine ("G"), cytosine ("C") and thymine ("T"). The four RNA bases are A, G, C and uracil ("U"). A and G
25 are purines, abbreviated to R, and C, T, and U are pyrimidines, abbreviated to Y.

DNA Sequence

A linear series of nucleotides connected one to the other by phosphodiester bonds between the 3' and 5' carbons of adjacent
30 pentoses.

Codon

A DNA sequence of three nucleotides (a triplet) which encodes through mRNA an amino acid, a translational start signal or a translational termination signal. For example, the nucleotide triplets TTA, TTG, CTT, CTC, CTA and CTG encode for the amino acid leucine ("Leu"), TAG, TAA and TGA are translational stop signals and ATG is a translational start signal.

Reading Frame

The grouping of codon's during translation of mRNA into amino acid sequences. During translation the proper reading frame must be maintained. For example, the sequence GCTGGTTGTAAG may be translated in three reading frames or phases, each of which affords a different amino acid sequence:

GCT GGT TGT AAG - Ala-Gly-Cys-Lys
G CTG GTT GTA AG - Leu-Val-Val
GC TGG TTG TAA A - Trp-Leu-(STOP)

Polypeptide

A linear series of amino acids connected one to the other by peptide bonds between the α -amino and carboxy groups of adjacent amino acids.

Genome

The entire DNA of a substance. It includes inter alia the structural genes encoding for the polypeptides of the substance, as well as operator, promoter and ribosome binding and interaction sequences including sequences such as the Shine-Dalgarno sequences.

Structural Gene

A DNA sequence which encodes through its template or messenger RNA ("mRNA") a sequence of amino acids characteristic of a specific polypeptide.

Transcription

The process of producing mRNA from a structural gene.

Translation

The process of producing a polypeptide from mRNA.

Expression

The process undergone by a structural gene to produce a polypeptide. It is a combination of transcription and translation.

Plasmid

- 5 A non-chromosomal double-stranded DNA sequence comprising an intact "replicon" such that the plasmid is replicated in a host cell. When the plasmid is placed within a unicellular organism, the characteristics of that organism are changed or transformed as a result of the DNA of the plasmid. For example, a plasmid carrying the gene for tetracycline resistance (Tet^R) transforms a cell
10 previously sensitive to tetracycline into one which is resistant to it. A cell transformed by a plasmid is called a "transformant".

Phage or Bacteriophage

- Bacterial virus many of which include DNA sequences
15 encapsidated in a protein envelope or coat ("capsid"). In a unicellular organism a phage may be introduced as free DNA by a process called transfection.

Cloning Vehicle

- A plasmid, phage DNA or other DNA sequences which are able to
20 replicate in a host cell, characterized by one or a small number of endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without attendant loss of an essential biological function of the DNA, e.g., replication, production of coat proteins or loss of promoter or binding sites, and which
25 contain a marker suitable for use in the identification of transformed cells, e.g., tetracycline resistance or ampicillin resistance. A cloning vehicle is often called a vector.

Cloning

The selection and propagation of a single species.

- 30 Recombinant DNA Molecule

A hybrid DNA sequence comprising at least two nucleotide sequences, the first sequence not normally being found together in nature with the second.

Expression Control Sequence

A DNA sequence of nucleotides that controls and regulates expression of structural genes when operatively linked to those genes.

5 To attain the objects of this invention it was necessary to determine the amino acid sequence and the DNA sequence of the structural gene encoding androgen receptor protein. One conventional approach would involve starting with the purified androgen receptor protein. However, as described above, significant
10 amounts of the protein for such purposes have not been obtained.

An alternative approach to circumvent the overwhelming difficulty of androgen receptor protein purification is direct isolation of the DNA encoding the messenger RNA for androgen receptor protein.

15 Our strategy for isolating AR DNA was based on evidence that the AR gene is X-linked and that no other steroid receptor gene is located on the X chromosome. Sequence data are available from cDNAs for glucocorticoid, estrogen, progesterone, mineralocorticoid and vitamin D receptors. Comparison of the derived amino acid sequences
20 has revealed a central region of high cysteine content which was found also in the v-erb A oncogene product recently identified as the thyroid hormone receptor. Within this 61-63 amino acid region is an arrangement of 9 cysteine residues that are absolutely conserved among steroid receptors thus far characterized. The
25 overall homology among sequences in this conserved region ranges between 40 and 90%. We assumed that AR would resemble other members of the steroid receptor family in the conserved DNA-binding domain.

A human X chromosomal library was screened with the synthetic oligo nucleotide probe A (Oligo A sequence = ^{5'} CTT TTG AAG AAG
30 ACC TTA CAG CCC TCA CAG GT^{3'}) of Figure 1 (A) designed as a consensus sequence from the conserved sequence of the DNA-binding domain of other steroid receptors. Screening the library with the oligo A probe resulted in several recombinants whose inserts were cloned into bacteriophage M13 DNA and sequenced. One recombinant

clone (Charon®35 X05AR) (human genomic clone [1]) contained a sequence similar to, yet distinct from, the DNA-binding domains of other steroid receptors. It had 84% sequence identity with oligo A, while other receptor DNAs were 78% to 91% homologous with the
5 consensus oligonucleotide.

From the nucleotide sequence just 5' of the DNA binding domain, oligonucleotide probe B (Oligo B sequence = ⁵GGA CCA TGT TTT GCC CAT TGA CTA TTA CTT TCC ACC CC^{3'}) was synthesized and used to screen bacteriophage lambda gt11 cDNA libraries from human
10 epididymis and cultured human foreskin fibroblasts. Recombinant phage (unamplified) screened with this oligonucleotide by in situ hybridization revealed one positive clone in each library. The epididymal clone (gt11 ARHEL1)(human epididymis clone [1]) contained the complete DNA-binding domain and approximately 1.5 kb of upstream
15 sequence, whereas the fibroblast clone (gt11 ARHFL1)(human fibroblast clone [1]) shown in Figure 1(B) contained the DNA-binding domain and 1.5 kb of downstream sequence. The DNA-binding domains of the cDNA isolates were identical to that of the genomic exon sequence.

20 Transient expression in monkey kidney cells (COS M6) demonstrated that the human foreskin fibroblast cDNA fragment encodes the steroid-binding domain of hAR. A DNA fragment (ARHFLIH-X) extending 5' to 3' from the Hind III site within the putative DNA-binding domain through the stop codon (TGA) was cloned
25 into pCMV as shown in Figure 2(A). Expression was facilitated by adding to the 5' end a consensus translation initiation sequence containing the methionine codon (ATG) in reading frame. Transfection of the recombinant construct produced a protein with high-affinity for [³H]dihydrotestosterone, Figure 2(C) saturable
30 at physiological levels of hormone. See Figure 2(B). The binding constant [$K_d = 2.7 (+ 1.4) \times 10^{-10}$ M] was nearly identical to that of native AR. The level of expressed protein, 1.3 pmol per milligram of protein, was 20 to 60 times greater than that in male reproductive tissues. Mock transfections without plasmid or

1341347

transfections with plasmid DNA lacking the AR insert yielded no specific binding of dihydrotestosterone. Figure 2(D) shows steroid specificity was identical to that of native AR, with highest affinity for dihydrotestosterone and testosterone, intermediate
5 affinity for progesterone and estradiol, and low affinity for cortisol.

Figure 3 is a clone map compiled to show the human androgen receptor gene and the nucleic acid sequences in the cDNA clones, human genomic clones, human fibroblast clones, human epididymis
10 clones, and rat epididymis clones. Human fibroblast clone [1] extended through the stop codon or the C-terminal end of the androgen receptor protein. To isolate and elucidate the sequence of the 5' or N-terminal end of the androgen receptor protein, we used a EcoRI/SstI fragment (EcoRI site was from the linker) from the 5' end
15 of human epididymis clone [1] as a probe (cDNA probe [A]), to rescreen the human X chromosomal library by standard techniques. By these techniques, human genomic clone [2] was isolated and in turn used as a probe to rescreen a human epididymis library and isolate human epididymis clone [5]. The N-terminal sequence was elucidated
20 along with the 5' flanking sequence of the androgen receptor protein and gene. Human genomic clones [3], [4] and [5] for the sequence 3' of human genomic clone [1] were obtained using cDNA probes B [a Hind III/EcoRI fragment] and C [an EcoRI fragment], by screening and isolating by standard techniques.

25 Two rat clones, rat epididymis clones [1] and [2], were isolated from a rat epididymis cDNA library using as probes the complete human epididymis clone [1] and a EcoRI/PstI fragment, cDNA probe [D], respectively. These rat clones contained the entire protein coding sequence for the rat androgen receptor, plus flanking
30 5' and 3' untranslated sequences which were used to confirm the sequence of the human androgen receptor.

The complete double-stranded sequence encoding the human androgen receptor protein was determined and the deduced amino acid sequence of the human androgen receptor protein is set forth

in Figure 4. The cDNA sequence and the amino acid sequence for the rat androgen receptor protein is set forth in Figure 5.

Recombinant DNA clones human fibroblast clone [1] isolated from human foreskin fibroblast cDNA gt11 expression library, human
5 epididymis clones [1] and [5] isolated from human epididymis cDNA
gt11 expression library were deposited in the American Type Culture
Collection with accession numbers ATCC # 40439, ATCC # 40442 and
ATCC # 40440 respectively. Human genomic clones [1], [2], [3], [4]
and [5] which were isolated from human X chromosome lambda Charon 35
10 library available as ATCC # 57750 have been deposited with the
American Type Culture Collection with accession numbers
ATCC # 40441, ATCC # 40443, ATCC # 40444, ATCC # 40445 and
ATCC # 40446 respectively.

A wide variety of host=cloning vehicle combinations may be
15 usefully employed in cloning the double stranded DNA disclosed
herein. For example, useful cloning vehicles may include
chromosomal, non-chromosomal and synthetic DNA sequences such as
various known bacterial plasmids and wider host range plasmids such
as pCMV and vectors derived from combinations of plasmids and phage
20 DNA such as plasmids which have been modified to employ phage DNA
expression control sequences. Useful hosts may include bacterial
hosts, yeasts and other fungi, animal or plant hosts, such as
Chinese Hamster Ovary cells (CHO, or monkey kidney cells (COS M6),
and other hosts. The particular selection of host-cloning vehicle
25 combinations may be made by those of skill in the art after due
consideration of factors such as the source of the DNA- i.e. genomic
or cDNA.

Cloning vehicles for use in carrying out the present invention
will further comprise a promoter operably linked to the DNA sequence
30 encoding the androgen receptor protein. In some instances it is
preferred that cloning vehicles further comprise an origin of

replication , as well as sequences which regulate and/or enhance expression levels, depending on the host cell selected.

Techniques for transforming hosts and expressing foreign cloned in them are well known in the art (See, for example, Maniatis et
5 al., *infra*). Cloning vehicles used for expressing foreign genes in bacterial hosts will generally contain a selectable marker, such as a gene for antibiotic resistance, and a promoter which functions in the host cell.

Eukaryotic microorganisms, such as the yeast *Saccharomyces*
10 *cerevisiae*, may also be used as host cells. Cloning vehicles will generally comprise a selectable marker, such as the nutritional marker TRP, which allows selection in a host strain carrying a *trp* mutation. To facilitate purification of an androgen receptor protein produced in a yeast transformant, a yeast gene encoding a
15 secreted protein may be joined to the sequence encoding androgen receptor protein.

Higher eukaryotic cells can also serve as host cells in carrying out the present invention. Cultured mammalian cells are preferred. Cloning vehicles for use in mammalian cells will
20 comprise a promoter capable of directing the transcription of a foreign gene introduced into a mammalian cell. Also contained in the expression vector is a polyadenylation signal, located downstream of the insertion site. The polyadenylation signal can be that of the cloned androgen receptor gene, or may be derived from a
25 heterologous gene.

A selectable marker, such as a gene that confers a selectable phenotype, is generally introduced into the cells along with the gene of interest. Preferred selectable markers include genes that confer resistance to drugs, such as neomycin, hygromycin and
30 methotrexate. Selectable markers may be introduced into the cell on a separate plasmid at the same time as the gene of interest, or they may be introduced on the same plasmid.

The copy marker of the integrated gene sequence can be increased through amplification by using certain selectable

markers. Through selection, expression levels may be substantially increased.

Androgen receptor proteins may be purified from the host cells or cell media according to the present invention using techniques well known to those in the art. Such proteins may be utilized to produce monoclonal or polyclonal antibodies according to the techniques described below.

The techniques of this invention offer considerable advances over existing technology for measurement of androgen receptor. Utilizing proteins and peptides containing the disclosed sequences monoclonal or polyclonal antibodies can be produced for use as immunochemical reagents in immunodiagnostic assays. For example, radioimmunoassays and ELISA assays can be developed utilizing these reagents which will allow detection and quantification of androgen receptor in the presence of endogenous androgen since such androgen will not interfere with antibody binding to the receptor.

Immunocytochemistry utilizing our reagents enables determination and quantification of the cellular distribution of the androgen receptor in tumor tissues, which are often heterogenous in composition. This assay offers great potential for diagnostic evaluation of prostate cancer to determine responsiveness to androgen withdrawal therapy.

In addition, the antibodies produced using the disclosed amino acid sequences can also be used in processes for the purification of androgen receptor protein produced by the above methods. One such purification process is disclosed in Logeat, F., et al., Biochemistry vol. 24 (1985), pp. 1029-1035.

Androgen receptor proteins and polypeptides synthesized from the deduced amino acid sequence can be used as immunogens for the preparation of antibodies to the androgen receptor. Peptides for such use range in length from about 3 to about 958 amino acids in length and are preferably from about 15 to about 30 amino acids in length. Shorter peptides may have significant sequence homology to

- 20 -

other steroid receptor proteins and larger peptides may contain multiple antigenic determinants; these properties could result in antibodies with cross-reactivities to other steroid receptor proteins.

5 Peptides can be synthesized from amino acid sequences in the NH₂-terminal region, the DNA-binding domain, and the carboxyl-terminal steroid binding domain. Peptide selection will be based on hydropathic plots, selecting hydrophilic regions that are more likely exposed on the receptor surface. For diagnostic
10 purposes preferred sequences will be selected from the NH₂-terminal region where there is the least homology with other steroid receptor proteins.

 Peptides for use as immunogens can be synthesized using techniques available to one of ordinary skill in the art. For
15 example, peptides corresponding to androgen receptor sequences can be synthesized using tBOC chemistry on a Biosearch Model 9500 peptide synthesizer. Peptide purity is assessed by high pressure liquid chromatography. Peptides can be conjugated to keyhole limpet hemocyanin through cysteine residues using the coupling agent
20 m-maleimido-benzoyl-N-hydroxysuccinimide ester. One can also prepare resin-bound peptides utilizing the p-(oxymethyl benzamide) handle to attach the C-terminal amino acid to solid-phase resin support.

 Proteins and peptides of this invention can be utilized for the
25 production of polyclonal or monoclonal antibodies. Methods for production of such antibodies are known to those of ordinary skill in the art and may be performed without undue experimentation. One method for the production of monoclonal antibodies is described in Kohler, G., et al., "Continuous Culture of Fused Cells Secreting
30 Antibody of Predefined Specificity," Nature, vol. 256 (1975), p. 495. Polyclonal antibodies, by way of example, can be produced by the method described below.

A

Peptide conjugates or resin-bound peptides can be injected into rabbits according to the procedure of Vaitukaitis et al., J Clin Endocrinol Metab, 33, 988-991 (1971) using a standard immunization schedule. Antisera titers can be determined in the ELISA assay.

5 For example, one androgen receptor sequence,
NH₂-Asp-His-Val-Leu-Pro-Ile-Asp-Tyr-Tyr-Phe-Pro-Pro-Gln-Lys-Thr
in the 5' region upstream from the DNA-binding domain, was used to
raise antisera in rabbits. The antisera react selectively at a
dilution of 1 to 500 with the androgen receptor both in its
10 untransformed 8-10S form and in its 4-5S transformed form. Receptor
sedimentation on sucrose gradients increases from 4 to 8-10S in the
presence of antiserum at high ionic strength and from 8-10S to
11-12S at low ionic strength sucrose gradients. In the ELISA
reaction against the peptide used as immunogen, reactivity was
15 detectable at 1 to 25,000 dilution. This antiserum at a dilution of
1 to 3000 was found effective in staining nuclear androgen receptor
in rat prostate and other male accessory sex glands (see Figure 6).

Our invention provides new molecular probes comprising
complementary DNA sequences derived from the deduced sequences
20 encoding the androgen receptor for diagnostic purposes. Such probes
may be used to detect the presence of androgen receptor mRNA in
tumor cells. Such probes may also be used for detection of androgen
receptor gene defects. Androgen receptor complementary DNA
sequences can be used as hybridization probes to detect
25 abnormalities in the androgen receptor gene or in its messenger RNA.

Androgen receptor DNA sequences disclosed and complementary RNA
sequences can be used to construct probes for use in DNA
hybridization assays. An example of one such hybridization assay
and methods for constructing probes for such assays are disclosed
30 in U.S. Patent No. 4,683,195 to Mullis et al., U.S. Patent No.
4,683,202 to Mullis, U.S. Patent No. 4,617,261 to Sheldon, III et
al., U.S. Patent No. 4,683,194 to Salki et al., and U.S. Patent No.
4,705,886 to Levenson et al.

By example, one method for detecting gene deletion utilizes Southern blotting and hybridization. DNA can be isolated from cultured skin fibroblasts or from leukocytes obtained from blood. DNA is cut with restriction enzymes, electrophoresed on an agarose gel, blotted onto nitrocellulose, and hybridized with
5 [³²P]-labeled androgen receptor DNA (see Maniatis, T. et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, 1982.)

10 In addition, small mutations can be detected utilizing methods known to one of ordinary skill in the art, from cultured skin fibroblasts of the affected individual. A cDNA library can be prepared using standard techniques. The androgen receptor clones can be isolated using a [³²P]DNA AR probe. The clones AR cDNA can then be sequenced and compared to normal AR cDNA sequences.

15 Alternatively genomic DNA can be isolated from blood leukocytes or cultured skin fibroblasts of the affected individual. The DNA is then subjected to restriction enzyme digestion, electrophoresis and is blotted onto nitrocellulose. Synthetic oligonucleotides can be used to bracket specific exons. Exon sequences are amplified using
20 the polymerase chain reaction, cloned into M13 and sequenced. The sequences are compared to normal human AR DNA sequences.

Another method of identifying small mutations or deletions takes advantage of the ability of RNase A to cleave regions of single stranded RNA in RNA:DNA hybrids. Genomic DNA isolated from
25 fibroblasts of affected individuals is hybridized with radioactive RNA probes (Promega® Biotec) prepared from wild-type androgen receptor cDNA. Mismatches due to mutations would be cleaved by RNase A and result in altered sized bands relative to wild-type on denaturing polyacrylamide gels.

30 Restriction fragment length polymorphism (RFLP) linked to the androgen receptor gene locus may be used in prenatal diagnosis and carrier detection of androgen insensitivity. For example, the presence of RFLPs in normal individuals is first established by isolating DNA from lymphocytes of at least six females (total of 12



X chromosomes). DNA can be isolated using the proteinase K procedure and fragmented using a battery of restriction enzymes. Preferred are those enzymes that contain the dinucleotide sequence CG in their recognition sequence. Southern blots are screened with
5 5-10 kb androgen receptor genomic fragments which if possible lack repetitive DNA. For those regions containing repetitive elements, total human genomic DNA can be added as competitor in the hybridization reaction. Alternatively, one can subclone selected regions to yield a probe free of repetitive elements.

10 For example, a human restriction fragment length was determined by cDNA probe (B) and Hind III restriction endonuclease using the Southern blot technique (See Figure 7). The two RFLP alleles detected are a fragment at 6.5 kb (allele) and a fragment at 3.5 kb (allele 2). Major constant fragment bands are seen at approximately
15 2 and 5 kb with minor constant bands at 0.9 and 7.5 kb. Allele 1 is present in approximately 30% of the X chromosomes of the Caucasian population. Allele 2 is present in approximately 20% of the X chromosomes of the Caucasian population. In Figure 8 Lanes A, B and D, DNA from women who are homozygous for allele 1 is shown. In
20 Figure 8 Lane C, DNA from a woman who is heterozygous for both alleles 1 and 2 is shown. Figure 8 Lane E contains DNA from a man that only possesses allele 2. This RFLP, and others determined by the clones we have isolated, will enable one to monitor the androgen receptor gene in various disease conditions described herein.

25 An example of using the androgen receptor clones to detect mutations is shown in Figure 8 where five different complete androgen insensitive patients' DNA are digested with EcoRI, electrophoresed on a Southern blot, and probed with cDNA probe B. The patient in lane B lacks a 3kb band indicating that part of the
30 androgen receptor gene is deleted. Further analysis of this and other patients DNA is possible with other AR probes and by sequencing by standard methods and comparing the abnormal sequence to the normal sequence described herein.

Other potential uses for oligonucleotide sequences disclosed, for example in construction of therapeutics to block genetic expression, will be obvious to one of ordinary skill in the art.

5

10

15

20

25

30

We claim:

1. A purified nucleic acid molecule comprising a nucleotide sequence encoding an androgen receptor protein having the amino acid sequence of Figure 4, or a fragment of said amino acid sequence which retains androgen receptor activity.
2. The nucleic acid molecule of claim 1 comprising the nucleotide sequence of Figure 4.
3. A purified nucleic acid molecule comprising a nucleotide sequence encoding an androgen receptor protein having the amino acid sequence of Figure 5 or a fragment of said amino acid sequence which retains androgen receptor activity.
4. The nucleic acid molecule of claim 3 comprising the nucleotide sequence of Figure 5.
5. A purified nucleic acid molecule comprising a nucleotide sequence encoding a DNA-binding amino acid sequence C L I C G D E A S G C H Y G A L T C G S C K V F F K R A A E G K Q K Y L C A S R N D C T I D K F R R K N C P S C R L R K C Y E A G M.
6. The nucleic acid molecule of claim 5 comprising nucleotides 2037 to 2234 of Figure 4.
7. A purified nucleic acid molecule comprising nucleotides 2235 to 3119 of Figure 4.
8. A purified nucleic acid molecule comprising a nucleotide sequence encoding an immunogenic fragment of the amino acid sequence of Figure 4.

9. The nucleic acid molecule of claim 8 wherein the immunogenic fragment comprises about 15 to about 30 consecutive amino acids of the amino acid sequence of Figure 4.
10. A purified nucleic acid molecule comprising a nucleotide sequence encoding an immunogenic fragment of the amino acid sequence of Figure 5.
11. The nucleic acid molecule of claim 10 wherein the immunogenic fragment comprises about 15 to about 30 consecutive amino acids of the amino acid sequence of Figure 5.
12. The nucleic acid molecule of any one of claims 1 to 11 wherein the nucleic acid molecule is a DNA molecule.
13. The nucleic acid molecule of any one of claims 1 to 11 wherein the nucleic is an RNA molecule.
14. A nucleic acid molecule complementary to a nucleic acid molecule of any one of claims 1 to 13.
15. A cloning vehicle comprising a nucleic acid molecule of any one of claims 1 to 14.
16. A host cell comprising a cloning vehicle of claim 15.
17. A purified androgen receptor protein comprising the amino acid sequence of Figure 4 or a fragment thereof which retains androgen receptor activity.
18. A purified androgen receptor protein comprising the amino acid sequence of Figure 5 or a fragment thereof which retains androgen receptor activity.

19. A DNA-binding polypeptide comprising the amino acid sequence C L I
C G D E A S G C H Y G A L T C G S C K V F F K R A A E G K Q K Y L C A S
R N D C T I D K F R R K N C P S C R L R K C Y E A G M.

20. A steroid-binding polypeptide comprising amino acids 625 to 919 of
Figure 4.

21. An immunogenic polypeptide comprising an immunogenic fragment of
the amino acid sequence of Figure 4.

22. The immunogenic polypeptide of claim 21 comprising about 15 to
about 30 consecutive amino acids of the amino acid sequence of Figure 4.

23. An immunogenic polypeptide comprising an immunogenic fragment of
the amino acid sequence of Figure 5.

24. The immunogenic polypeptide of claim 23 comprising about 15 to
about 30 consecutive amino acids of the amino acid sequence of Figure 5.

25. A purified androgen receptor protein isolated from a culture of a host
cell of claim 16.

26. A method for determining whether a subject has a normal androgen
receptor protein or a mutant androgen receptor protein, the method
comprising the steps of:

- (a) obtaining an androgen receptor protein sample from the subject;
- (b) conducting an assay on the protein sample to detect a normal or
mutant androgen receptor protein or a normal or abnormal androgen receptor
protein function.

27. The method of claim 26 wherein the assay is an assay to detect a
normal or mutant androgen receptor protein amino acid sequence and is
selected from the group consisting of an immunoassay, an enzyme site

specific assay, an analysis of protein electrophoretic mobility or an analysis of proteolytic cleavage patterns.

28. A method for determining whether a subject has a normal androgen receptor gene or a mutant androgen receptor gene, the method comprising the steps of:

- (a) obtaining a nucleic acid sample from the subject; and
- (b) conducting an assay on the nucleic acid sample to determine the presence of a normal androgen receptor gene or a mutant androgen receptor gene.

29. The method of claim 28 wherein the assay is selected from the group consisting of probe hybridisation, direct sequencing, restriction enzyme fragment analysis and fragment electrophoretic mobility.

30. The method of claim 29 wherein the nucleic acid sample is an RNA sample and the assay is a direct sequencing assay.

31. The method of claim 30 wherein the assay comprises the steps of:

- (a) reverse transcribing the RNA sample to produce a corresponding cDNA;
- (b) performing at least one polymerase chain reaction with suitable oligonucleotide primers to amplify the androgen receptor cDNA;
- (c) obtaining the nucleotide sequence of the amplified androgen receptor cDNA; and
- (d) determining the presence or absence of an androgen receptor gene mutation in said nucleotide sequence which is associated with a disease selected from the group consisting of prostate cancer, androgen insensitivity syndrome, and infertility.

32. The method of claim 31 wherein the oligonucleotide primers comprise a number of consecutive nucleotides which are complementary to a portion of the cDNA sequence for androgen receptor protein.

33. The method of claim 28 wherein the nucleic acid sample is a DNA sample.

34. The method of claim 33 wherein the DNA sample is a genomic DNA sample and the assay comprises the steps of:

- (a) amplifying a target portion of the nucleotide sequence of the genomic DNA;
- (b) obtaining the nucleotide sequence of said amplified target position; and
- (c) determining the presence of absence of androgen receptor gene mutation in said target portion nucleotide sequence which is associated with a disease selected from the group consisting of prostate cancer, androgen insensitivity syndrome and infertility.

35. The use of a nucleic acid molecule of any one of claims 1 to 4 for administration to a subject for treatment of a disease selected from the group consisting of androgen insensitivity syndrome and infertility.

36. The use of a purified androgen receptor protein of any one of claims 17 to 18 for administration to a subject for treatment of a disease selected from the group consisting of androgen insensitivity syndrome and infertility.

37. A human androgen receptor protein containing the sequence

Trp Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg
Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg Met
Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr
Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly
Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg
Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr
Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe Thr Phe Asp Leu
Leu Ile Lys Ser His Met Val Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser
Val Gln Val Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln

38. A cDNA which encodes a human androgen receptor protein containing the sequence

Trp Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg
Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg Met
Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr
Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly
Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg
Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr
Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe Thr Phe Asp Leu
Leu Ile Lys Ser His Met Val Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser
Val Gln Val Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln

39. A replicable expression vector comprising a cDNA which encodes a human androgen receptor protein containing the sequence

Trp Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg
Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg Met
Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr
Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly
Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg
Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr
Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe Thr Phe Asp Leu
Leu Ile Lys Ser His Met Val Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser
Val Gln Val Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln

40. A host cell transformed with a replicable expression vector comprising a cDNA which encodes a human androgen receptor protein containing the sequence

Trp Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg
Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg Met
Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr
Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly
Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg
Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr
Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe Thr Phe Asp Leu
Leu Ile Lys Ser His Met Val Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser
Val Gln Val Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln

41. A method of producing a human androgen receptor protein comprising:
- constructing a replicable expression vector comprising a cDNA which encodes a human androgen receptor protein containing the sequence

Trp Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg
Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg Met
Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr
Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly
Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg
Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr
Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe Thr Phe Asp Leu
Leu Ile Lys Ser His Met Val Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser
Val Gln Val Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln

- transforming a host cell with said replicable expression vector;
- culturing said host cell under conditions permitting the expression of said human androgen receptor protein;
- recovering said human androgen receptor protein from the cell culture.

42. A method of detecting normal or abnormal androgen receptors in humans comprising:
- obtaining a sample containing androgen receptor proteins from a subject;
- assaying said androgen receptor proteins to determine their normality or abnormality.

43. A method of detecting a normal or abnormal androgen receptor gene in a human comprising:
- obtaining a nucleic acid sample from a subject;
- assaying said sample to determine the normality or abnormality of the sequence coding for the androgen receptor.

44. A rat androgen receptor protein containing the sequence

Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser Lys Thr Tyr Arg
Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu Ala Ile Gln Asn Pro Gly Pro Arg
His Pro Glu Ala Ala Ser Ile Ala Pro Pro Gly Ala Cys Leu Gln Gln Arg Gln Glu Thr
Ser Pro Arg Arg Arg Arg Arg Gln Gln His Pro Glu Asp Gly Ser Pro Gln Ala His Ile
Arg Gly Thr Thr Gly Tyr Leu Ala Leu Glu Glu Glu Gln Gln Pro Ser Gln Gln Gln Ser
Ala Ser Glu Gly His Pro Glu Ser Gly Cys Leu Pro Glu Pro Gly Ala Ala Thr Ala Pro
Gly Lys Gly Leu Pro Gln Gln Pro Pro Ala Pro Pro Asp Gln Asp Asp Ser Ala Ala Pro
Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro Gly Leu Ser Ser Cys Ser Ala Asp Ile
Lys Asp Ile Leu Ser Glu Ala Gly Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Gln Gln
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
Gly Ser Ser Ser Val Arg Ala Arg Glu Ala Thr Gly Ala Pro Ser Ser Ser Lys Asp Ser
Tyr Leu Gly Gly Asn Ser Thr Ile Ser Asp Ser Ala Lys Glu Leu Cys Lys Ala Val Ser
Val Ser Met Gly Leu Gly Val Glu Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg
Gly Asp Cys Met Tyr Ala Ser Leu Leu Gly Gly Pro Pro Ala Val Arg Pro Thr Pro Cys
Ala Pro Leu Ala Glu Cys Lys Gly Leu Ser Leu Asp Glu Gly Pro Gly Lys Gly Thr Glu
Glu Thr Ala Glu Tyr Ser Ser Phe Lys Gly Gly Tyr Ala Lys Gly Leu Glu Gly Glu Ser
Leu Gly Cys Ser Gly Ser Ser Glu Ala Gly Ser Ser Gly Thr Leu Glu Ile Pro Ser Ser
Leu Ser Leu Tyr Lys Ser Gly Ala Val Asp Glu Ala Ala Ala Tyr Gln Asn Arg Asp Tyr
Tyr Asn Phe Pro Leu Ala Leu Ser Gly Pro Pro His Pro Pro Pro Pro Thr His Pro His

45. A cDNA which encodes a rat androgen receptor protein containing the sequence

Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser Lys Thr Tyr Arg
Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu Ala Ile Gln Asn Pro Gly Pro Arg
His Pro Glu Ala Ala Ser Ile Ala Pro Pro Gly Ala Cys Leu Gln Gln Arg Gln Glu Thr
Ser Pro Arg Arg Arg Arg Arg Gln Gln His Pro Glu Asp Gly Ser Pro Gln Ala His Ile
Arg Gly Thr Thr Gly Tyr Leu Ala Leu Glu Glu Glu Gln Gln Pro Ser Gln Gln Gln Ser
Ala Ser Glu Gly His Pro Glu Ser Gly Cys Leu Pro Glu Pro Gly Ala Ala Thr Ala Pro
Gly Lys Gly Leu Pro Gln Gln Pro Pro Ala Pro Pro Asp Gln Asp Asp Ser Ala Ala Pro
Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro Gly Leu Ser Ser Cys Ser Ala Asp Ile
Lys Asp Ile Leu Ser Glu Ala Gly Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Gln Gln
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
Gly Ser Ser Ser Val Arg Ala Arg Glu Ala Thr Gly Ala Pro Ser Ser Ser Lys Asp Ser
Tyr Leu Gly Gly Asn Ser Thr Ile Ser Asp Ser Ala Lys Glu Leu Cys Lys Ala Val Ser
Val Ser Met Gly Leu Gly Val Glu Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg
Gly Asp Cys Met Tyr Ala Ser Leu Leu Gly Gly Pro Pro Ala Val Arg Pro Thr Pro Cys
Ala Pro Leu Ala Glu Cys Lys Gly Leu Ser Leu Asp Glu Gly Pro Gly Lys Gly Thr Glu
Glu Thr Ala Glu Tyr Ser Ser Phe Lys Gly Gly Tyr Ala Lys Gly Leu Glu Gly Glu Ser
Leu Gly Cys Ser Gly Ser Ser Glu Ala Gly Ser Ser Gly Thr Leu Glu Ile Pro Ser Ser
Leu Ser Leu Tyr Lys Ser Gly Ala Val Asp Glu Ala Ala Ala Tyr Gln Asn Arg Asp Tyr
Tyr Asn Phe Pro Leu Ala Leu Ser Gly Pro Pro His Pro Pro Pro Pro Thr His Pro His

46. A replicable expression vector comprising a cDNA which encodes a rat androgen receptor protein containing the sequence

Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser Lys Thr Tyr Arg
Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu Ala Ile Gln Asn Pro Gly Pro Arg
His Pro Glu Ala Ala Ser Ile Ala Pro Pro Gly Ala Cys Leu Gln Gln Arg Gln Glu Thr
Ser Pro Arg Arg Arg Arg Arg Gln Gln His Pro Glu Asp Gly Ser Pro Gln Ala His Ile
Arg Gly Thr Thr Gly Tyr Leu Ala Leu Glu Glu Glu Gln Gln Pro Ser Gln Gln Gln Ser
Ala Ser Glu Gly His Pro Glu Ser Gly Cys Leu Pro Glu Pro Gly Ala Ala Thr Ala Pro
Gly Lys Gly Leu Pro Gln Gln Pro Pro Ala Pro Pro Asp Gln Asp Asp Ser Ala Ala Pro
Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro Gly Leu Ser Ser Cys Ser Ala Asp Ile
Lys Asp Ile Leu Ser Glu Ala Gly Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Gln Gln
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
Gly Ser Ser Ser Val Arg Ala Arg Glu Ala Thr Gly Ala Pro Ser Ser Ser Lys Asp Ser
Tyr Leu Gly Gly Asn Ser Thr Ile Ser Asp Ser Ala Lys Glu Leu Cys Lys Ala Val Ser
Val Ser Met Gly Leu Gly Val Glu Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg
Gly Asp Cys Met Tyr Ala Ser Leu Leu Gly Gly Pro Pro Ala Val Arg Pro Thr Pro Cys
Ala Pro Leu Ala Glu Cys Lys Gly Leu Ser Leu Asp Glu Gly Pro Gly Lys Gly Thr Glu
Glu Thr Ala Glu Tyr Ser Ser Phe Lys Gly Gly Tyr Ala Lys Gly Leu Glu Gly Glu Ser
Leu Gly Cys Ser Gly Ser Ser Glu Ala Gly Ser Ser Gly Thr Leu Glu Ile Pro Ser Ser
Leu Ser Leu Tyr Lys Ser Gly Ala Val Asp Glu Ala Ala Ala Tyr Gln Asn Arg Asp Tyr
Tyr Asn Phe Pro Leu Ala Leu Ser Gly Pro Pro His Pro Pro Pro Pro Thr His Pro His

47. A host cell transformed with a replicable expression vector comprising a cDNA which encodes a rat androgen receptor protein containing the sequence

Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser Lys Thr Tyr Arg
Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu Ala Ile Gln Asn Pro Gly Pro Arg
His Pro Glu Ala Ala Ser Ile Ala Pro Pro Gly Ala Cys Leu Gln Gln Arg Gln Glu Thr
Ser Pro Arg Arg Arg Arg Arg Gln Gln His Pro Glu Asp Gly Ser Pro Gln Ala His Ile
Arg Gly Thr Thr Gly Tyr Leu Ala Leu Glu Glu Glu Gln Gln Pro Ser Gln Gln Gln Ser
Ala Ser Glu Gly His Pro Glu Ser Gly Cys Leu Pro Glu Pro Gly Ala Ala Thr Ala Pro
Gly Lys Gly Leu Pro Gln Gln Pro Pro Ala Pro Pro Asp Gln Asp Asp Ser Ala Ala Pro
Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro Gly Leu Ser Ser Cys Ser Ala Asp Ile
Lys Asp Ile Leu Ser Glu Ala Gly Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Gln Gln
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Glu Val Ile Ser Glu
Gly Ser Ser Ser Val Arg Ala Arg Glu Ala Thr Gly Ala Pro Ser Ser Ser Lys Asp Ser
Tyr Leu Gly Gly Asn Ser Thr Ile Ser Asp Ser Ala Lys Glu Leu Cys Lys Ala Val Ser
Val Ser Met Gly Leu Gly Val Glu Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg
Gly Asp Cys Met Tyr Ala Ser Leu Leu Gly Gly Pro Pro Ala Val Arg Pro Thr Pro Cys
Ala Pro Leu Ala Glu Cys Lys Gly Leu Ser Leu Asp Glu Gly Pro Gly Lys Gly Thr Glu
Glu Thr Ala Glu Tyr Ser Ser Phe Lys Gly Gly Tyr Ala Lys Gly Leu Glu Gly Glu Ser
Leu Gly Cys Ser Gly Ser Ser Glu Ala Gly Ser Ser Gly Thr Leu Glu Ile Pro Ser Ser
Leu Ser Leu Tyr Lys Ser Gly Ala Val Asp Glu Ala Ala Ala Tyr Gln Asn Arg Asp Tyr
Tyr Asn Phe Pro Leu Ala Leu Ser Gly Pro Pro His Pro Pro Pro Pro Thr His Pro His

48. A method of producing a rat androgen receptor protein comprising:
- constructing a replicable expression vector comprising a cDNA which encodes a rat androgen receptor protein containing the sequence

Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser Lys Thr Tyr Arg
Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu Ala Ile Gln Asn Pro Gly Pro Arg
His Pro Glu Ala Ala Ser Ile Ala Pro Pro Gly Ala Cys Leu Gln Gln Arg Gln Glu Thr
Ser Pro Arg Arg Arg Arg Arg Gln Gln His Pro Glu Asp Gly Ser Pro Gln Ala His Ile
Arg Gly Thr Thr Gly Tyr Leu Ala Leu Glu Glu Glu Gln Gln Pro Ser Gln Gln Gln Ser
Ala Ser Glu Gly His Pro Glu Ser Gly Cys Leu Pro Glu Pro Gly Ala Ala Thr Ala Pro
Gly Lys Gly Leu Pro Gln Gln Pro Pro Ala Pro Pro Asp Gln Asp Asp Ser Ala Ala Pro
Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro Gly Leu Ser Ser Cys Ser Ala Asp Ile
Lys Asp Ile Leu Ser Glu Ala Gly Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Gln Gln
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Glu Val Ile Ser Glu
Gly Ser Ser Ser Val Arg Ala Arg Glu Ala Thr Gly Ala Pro Ser Ser Ser Lys Asp Ser
Tyr Leu Gly Gly Asn Ser Thr Ile Ser Asp Ser Ala Lys Glu Leu Cys Lys Ala Val Ser
Val Ser Met Gly Leu Gly Val Glu Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg
Gly Asp Cys Met Tyr Ala Ser Leu Leu Gly Gly Pro Pro Ala Val Arg Pro Thr Pro Cys
Ala Pro Leu Ala Glu Cys Lys Gly Leu Ser Leu Asp Glu Gly Pro Gly Lys Gly Thr Glu
Glu Thr Ala Glu Tyr Ser Ser Phe Lys Gly Gly Tyr Ala Lys Gly Leu Glu Gly Glu Ser
Leu Gly Cys Ser Gly Ser Ser Glu Ala Gly Ser Ser Gly Thr Leu Glu Ile Pro Ser Ser
Leu Ser Leu Tyr Lys Ser Gly Ala Val Asp Glu Ala Ala Ala Tyr Gln Asn Arg Asp Tyr
Tyr Asn Phe Pro Leu Ala Leu Ser Gly Pro Pro His Pro Pro Pro Pro Thr His Pro His

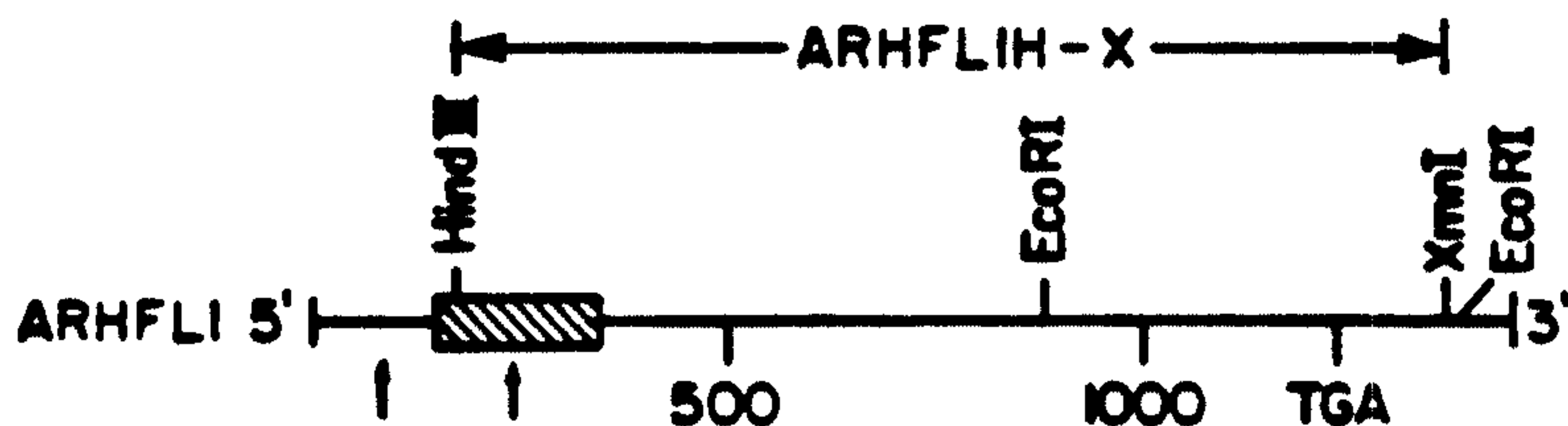
- transforming a host cell with said replicable expression vector;
- culturing said host cell under conditions permitting the expression of said rat androgen receptor protein;
- recovering said rat androgen receptor protein from the cell culture.

FIGURE 1

A.

Oligo A	Complement	5'-ACC	TGT	GAG	GGC	TGT	AAG	GTC	TTC	TTC	AAA	AG-3'	(100%)
		**	***	*	*	**	***	***	***	***	***	**	
hAR (X)		ACA	TGT	GGA	AGC	TGC	AAG	GTC	TTC	TTC	AAA	AG	(84%)
hPR (11)		ACC	TGT	GGC	AGC	TGT	AAG	GTC	TTC	TTT	AAG	AG	(88%)
hMR (4)		ACC	TGT	GGC	AGC	TGC	AAA	GTT	TTC	TTC	AAA	AG	(81%)
hGR (5)		ACT	TGT	GGA	AGC	TGT	AAA	GTT	TTC	TTC	AAA	AG	(81%)
hER (6)		TCC	TGT	GAG	GGC	TGT	AAG	GCC	TTC	TTC	AAG	AG	(91%)
hT3R (3, 17)		ACG	TGT	GAA	GGC	TGC	AAG	GGT	TTC	TTT	AGA	AG	(78%)
hRAR (17)		GCC	TGT	GAG	GGC	TGC	AAG	GGC	TTC	TTC	CGC	CG	(78%)

B.



C.

DNA-Binding Domain

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Paul J. H. H. H.

Send to J. Biol. Chem.

FIGURE 2

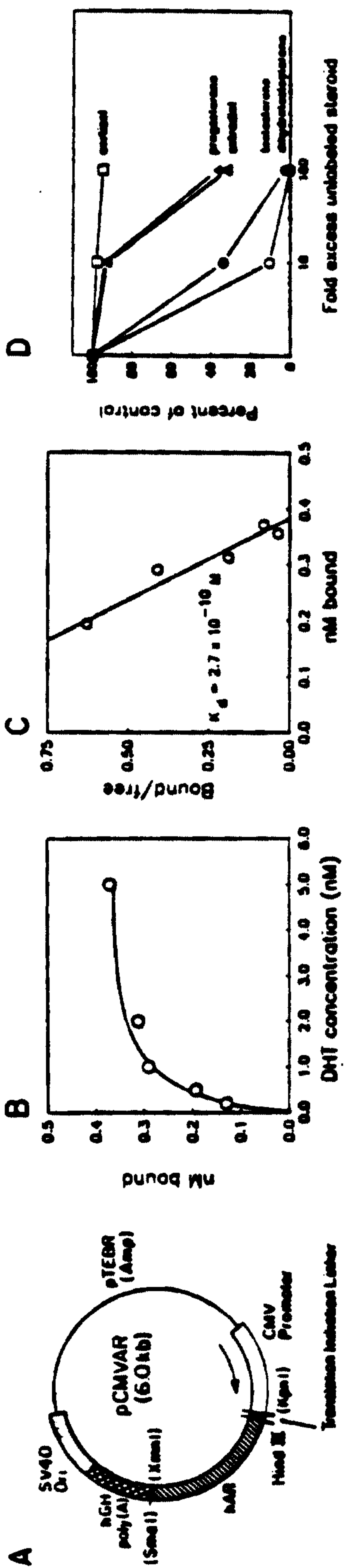
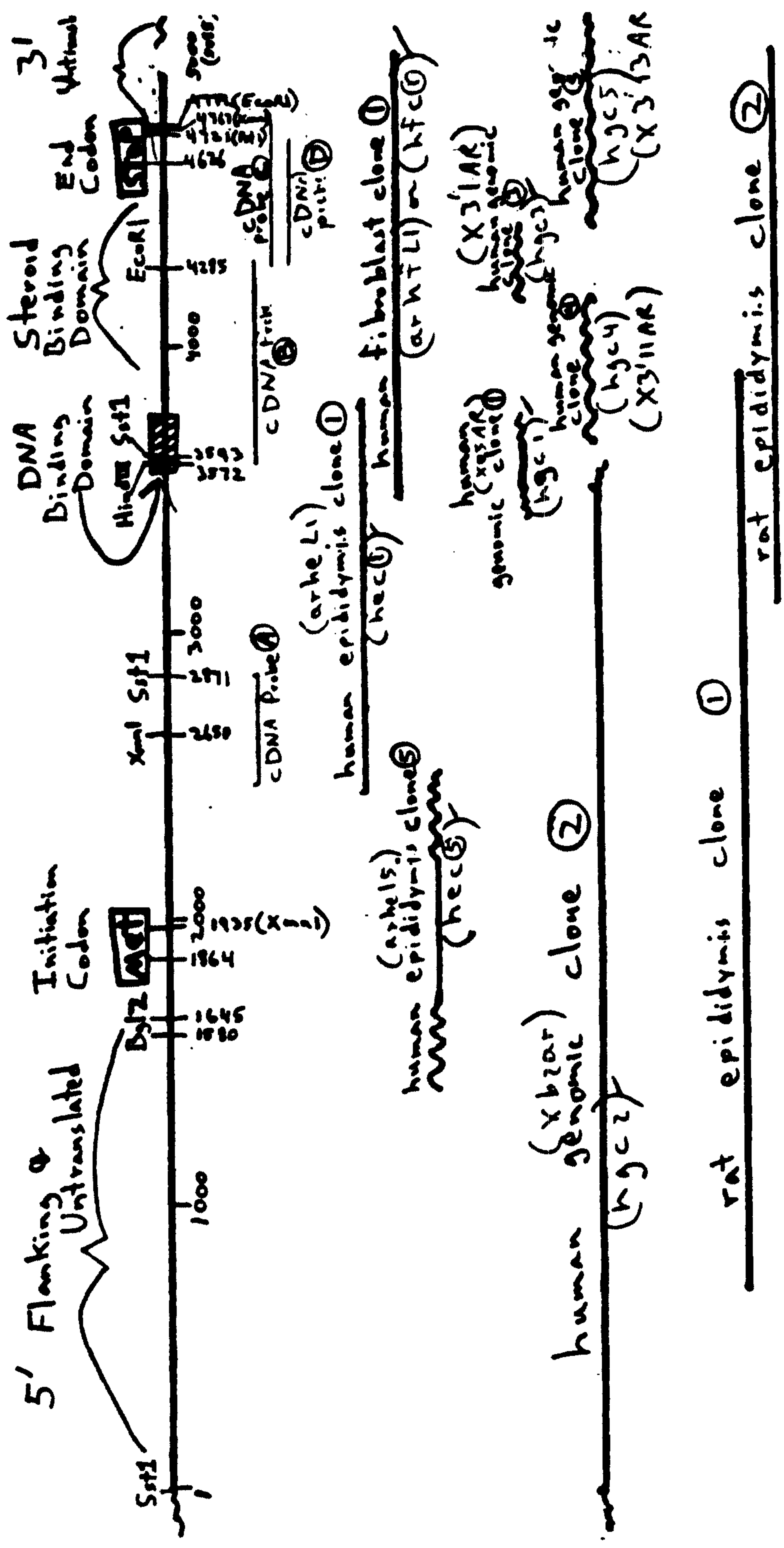


FIGURE 3
Compiled Clone Map of the Human Androgen Receptor



— = Sequenced
~ = Unsequenced

Said, M. C. B. 1987

[illegible]

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.

Sent: 4/2/19

ATCCTCAGGAGCCAGCCTGCTGGGAGAACAGAGGGCTCCGGAGCAAACTGGAGGCTGAGAGGGCATCAGAGGGGAAAAGACTGASTTAGCCACTCCAGTGCCATACAGAAGCTTAAGGG	120
ACATACCAACGCCAGGACCGCCAGCCAGCAGACCAACCGCTGTGCAGAGCCCGGGCTTCAGAGCCGCCCCAGAGACTGCCCTTTCCCTTTCGGTGAGTTTCAAAGCTGCCGAGC	240
CTCGGAGGAAGCAAGAAAGATTCTGGGTAGGACTACGACTGCTTTTGTCTCTCTCTCTTCACTCCCTTCACTCCCTTGGGTGGCTTCCCTTGACGGCAATAGGCAAGCTTCTGTGGCCAG	360
CTCTCTCCCTACCAACCAAGCTCTGCCAGCAGTGTGCACAGAGGCAACTCTCTTTGGCTCAAGAGCAGACAGCTTGTGCGCAATGGAAAGGAGGCTTTTGGGAGCCGAGAGACTGAGG	480
AGCAACACGACGCTGGAGACTCCTTGATCCAGGCTCTCCCTCTGCACTTCTCTACTGCCCGGCTTCACTGCTGTGTGCGCACTAGATTGAAGAAATGAAGAGACATGGGGCTTCA	600
GTAGTCAAGAGCAAAACAAAGCAAAAGAAAAACAAAAGAAATAGCCCACTTCTTATTGTGCACTGCTTTCAGTGGACATGACTTTGGAAAGCAGATGATTTTGCTTCTCCCACTCA	720
AGCTTTGAGCATCTTTAATCTGTGCTCTCAAGTATTAGGAGCAAACTGTGAACATAGCAGGCGAGATCTGTCTACGCGGTGCCCTTCTCTTACAGGAGACTTTGAGGTATCTGGGCG	840
TCCCCCTCTCCCTGCAAGTTTTCCTCTGAGCTTCCCGAGCTGGGAGCTAGCTGCAGATACATCATCAGTCAGTACAGAACTTTCAGAGCAAGACAGCAGGAGCAGGATAAG	960

[illegible]

ACCCAAACCACCTTGTTCCTTTTCAGATGCTTCTGCTGTTATATAACTCTGCACTACTTCTCTGCATTGCTTGGGGAAATTCTCTACTGATGTACAGTCTGTCTATGAACATGT 3840
TCCCCAAGTTCATATTCCTGGGCTTTTCCTTCTTTCTTTTCTTCTCTGCTCTTTTACCTCCCATGGACATTTGAATCCGCTGGGTGTGTGCTCTCTGCTGTGTTTGAAGT 3960
TTTGTGTATTTCTCAAGTCTGTGATCATCTTCTTGTGGCCAGTGTCAACTGTGCTTGTATAGCACTGTGCTGTGTGCCAACCAAGCAAATGTTTACTCACCTTATGCCATGGCAA 4080
GTTTACAGAGCTATAAGTATCTTGGGAAGAACAAACAGAGAGTAAAAAAACC(A)₄₅

The open reading frame

The sequence shown begins 120 bp downstream from the 5'-end of cDNA clone rARep1. The open reading frame begins at nucleotide 994 and ends with a stop codon TGA at nucleotide 3700. Underlined is a methionine codon (ATG, nucleotide residue 578) that immediately precedes a stop codon (TAG, nucleotide residue 602), and both of these precede the predicted methionine initiation codon (ATG, nucleotide residue 994). Also underlined is a stop codon (TAA) in reading frame at residue 988, 5' to the predicted initiation methionine at residue 994. Several amino acid repeat sequences are underlined as is a sequence resembling a poly(A) addition signal at nucleotide 4129 (AAGAAA). On the 3'-end of the cDNA insert rARep3 are 45 adenylic acid residues which begin 23 nucleotides after the poly(A) addition-like signal. An asterisk highlights each of the nine conserved cysteines of the DNA-binding domain. The sequence of amino acids enclosed in a box was used in peptide synthesis for polyclonal antibody production. The nucleotide sequence of the rat AR is available from GenBank under accession number M20133.

Sent 1/2

FIGURE 6

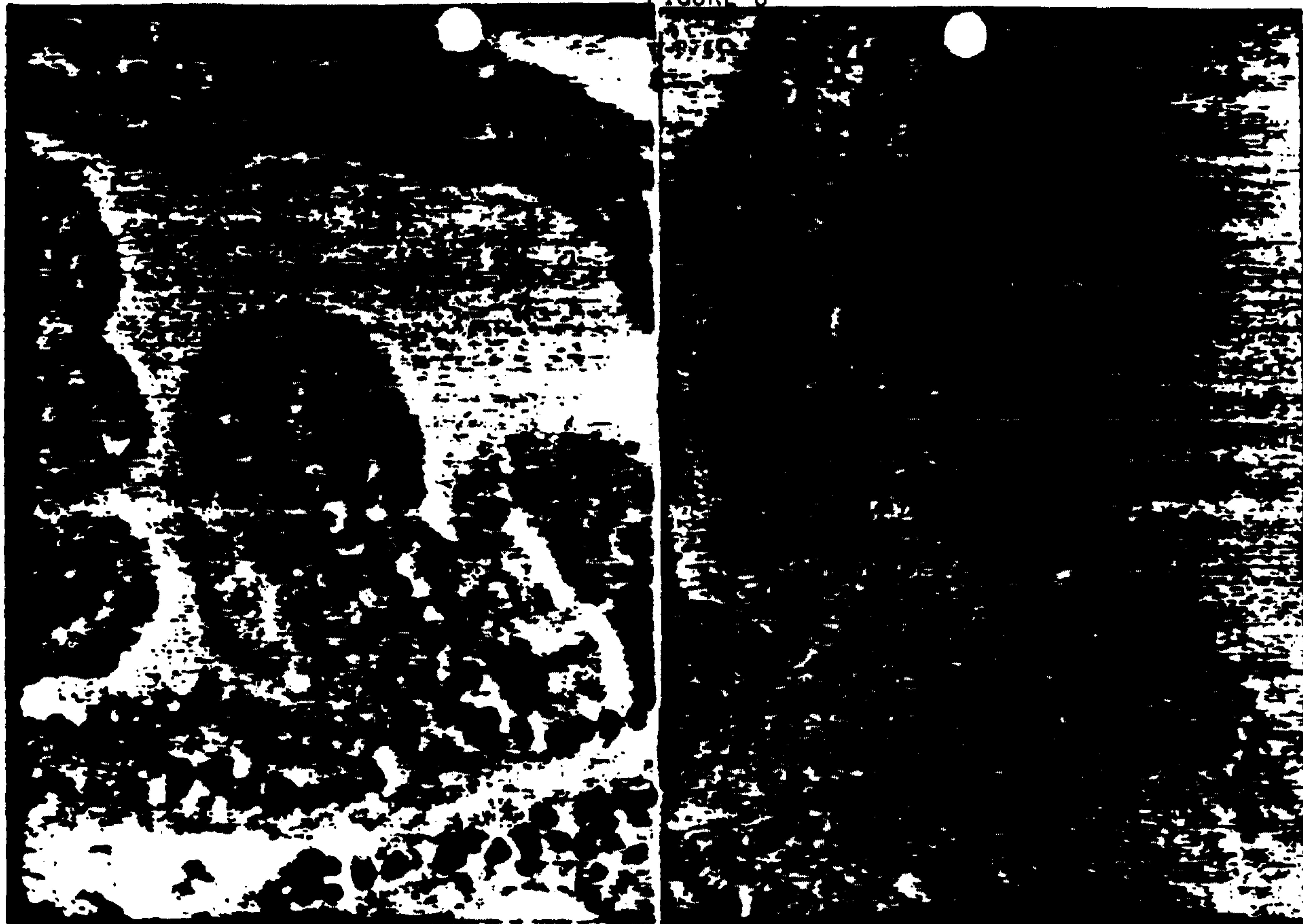
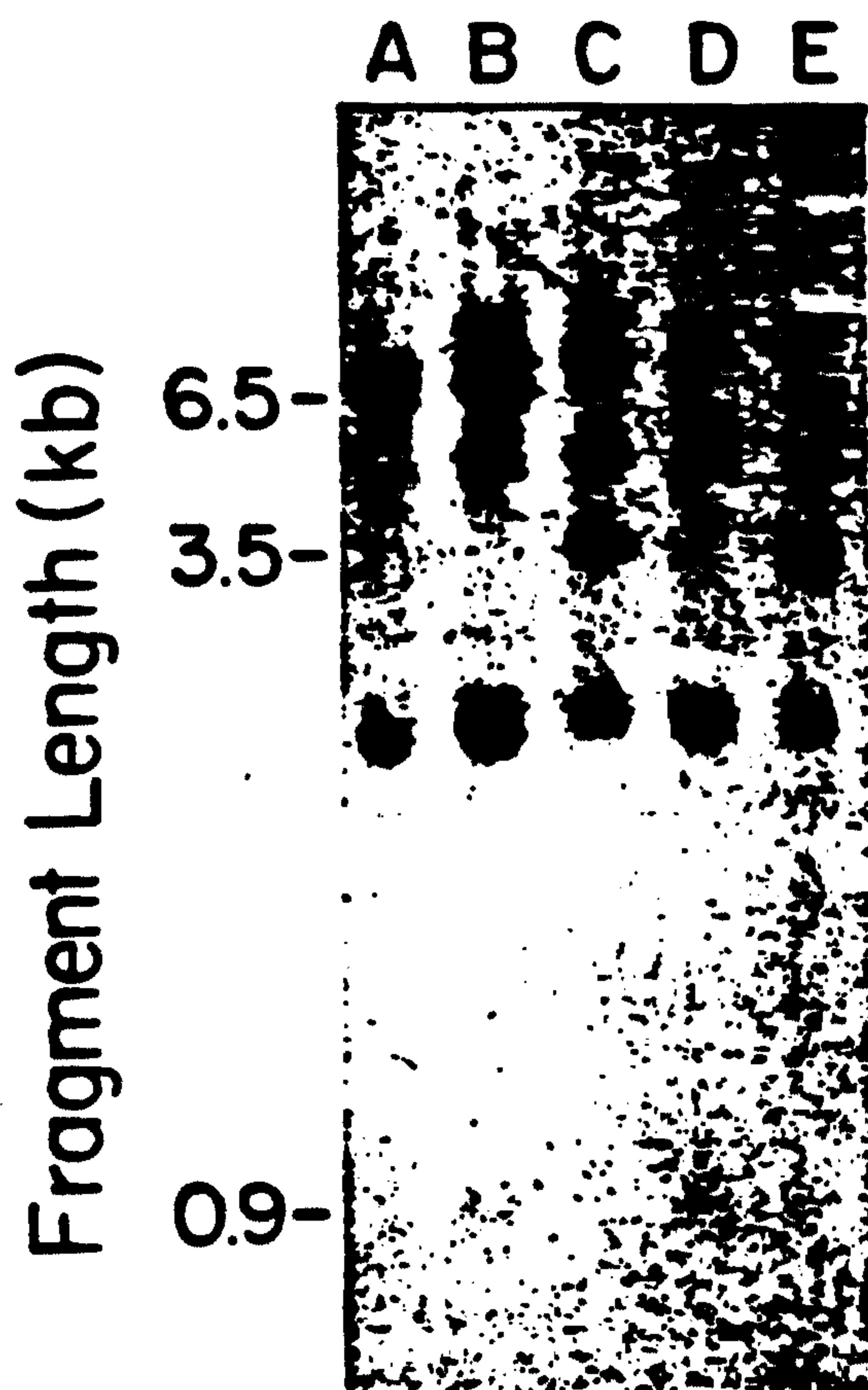


Figure 7. Frozen section of rat ventral prostate stained with antibodies (AR-52-3-p) to the AR peptide $\text{NH}_2\text{-Asp-His-Val-Leu-Pro-Ile-Asp-Tyr-Tyr-Phe-Pro-Pro-Gln-Lys-Thr}$ in a dilution of 1 to 3000 using the avidin-biotin peroxidase technique. Androgen receptor is indicated by brown staining of nuclei in epithelial cells. Immuno-staining was performed as previously described (60).

Send to: *[Signature]*

FIGURE 7

Restriction Fragment Length Polymorphism In the Human Androgen Receptor Gene



Sim 3/4/88

8/8

1341347

FIGURE 8

Southern Blot Analysis of Complete Androgen Insensitivity Syndrome Patients



Sim 3/4/80