

**FUNCTIONAL DOMAINS  
OF THE  
HUMAN ANDROGEN RECEPTOR**

Functionele domeinen van de humane androgeen receptor

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR  
AAN DE ERASMUS UNIVERSITEIT ROTTERDAM  
OP GEZAG VAN DE RECTOR MAGNIFICUS  
PROF. DR. P.W.C. AKKERMANS, M.A.  
EN VOLGENS BESLUIT VAN HET COLLEGE VOOR PROMOTIES.

DE OPENBARE VERDEDIGING ZAL PLAATSVINDEN OP  
WOENSDAG 14 DECEMBER 1994 OM 15.45 UUR

DOOR

GUIDO WIJNANDUS JENSTER

GEBOREN TE BOXMEER

**PROMOTIE-COMMISSIE:**

PROMOTOR: PROF. DR. J.A. GROOTEGOED

OVERIGE LEDEN: PROF. DR. J.P.H. BURBACH  
PROF. DR. F.G. GROSVELD  
PROF. DR. S.W.J. LAMBERTS

CO-PROMOTOREN: DR. A.O. BRINKMANN  
DR. IR. J. TRAPMAN



Dit proefschrift werd bewerkt binnen de Vakgroep Endocrinologie & Voortplanting,  
Faculteit der Geneeskunde en Gezondheidswetenschappen, Erasmus Universiteit Rotterdam.

Het beschreven project is gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO, Gebied Medische Wetenschappen, project: 900-546-079).

In dankbare herinnering aan mijn vader  
Voor mijn moeder  
Voor Ingrid

## Contents

### CONTENTS

	page
Abbreviations	7
<b>Chapter 1</b>	
<b>INTRODUCTION AND SCOPE OF THIS THESIS</b>	
1.1 Introduction	10
1.1.1 The nuclear receptor family	11
1.1.2 Androgens and antiandrogens	15
1.1.3 Structural and functional units as building blocks of proteins	15
1.2 Scope of this thesis	17
<b>Chapter 2</b>	
<b>POST-TRANSLATIONAL MODIFICATION</b>	
2.1 Introduction	20
2.2 Androgen receptor phosphorylation and isotypes	21
2.3 Function of androgen receptor phosphorylation	23
2.4 Activation of nuclear receptors via modulation of kinase-mediated signal transduction pathways	24
2.5 Anomalous androgen receptor mobility during SDS-PAGE	25
<b>Chapter 3</b>	
<b>ASSOCIATION WITH HEAT SHOCK PROTEINS</b>	
3.1 Introduction	28
3.2 Stress and heat shock proteins	28
3.3 The hsp associating domain in steroid receptors	29
3.4 Role of steroid receptor associated proteins	30
<b>Chapter 4</b>	
<b>SUBCELLULAR LOCALIZATION AND NUCLEAR IMPORT</b>	
4.1 Introduction	34
4.2 Nuclear localization signals in the human androgen receptor	35
4.3 Androgen receptor clustering	36
4.4 Effects of antiandrogens and various steroid hormones on androgen receptor nuclear import	38
4.5 The compartmental retention model	38

<b>Chapter 5</b>	
<b>STEROID BINDING</b>	
5.1 Introduction	42
5.2 Characteristics and 3-dimensional models of the ligand binding domain	42
5.3 Androgen receptor characteristics encoded by sequences in the ligand binding domain	45
<b>Chapter 6</b>	
<b>DIMERIZATION</b>	
6.1 Introduction	48
6.2 Dimerization domains	48
<b>Chapter 7</b>	
<b>ASSOCIATION WITH THE NUCLEAR MATRIX</b>	
7.1 Introduction	52
7.2 Steroid receptor binding to the nuclear matrix	53
7.3 Functional aspects of nuclear matrix-steroid receptor association	54
<b>Chapter 8</b>	
<b>DNA BINDING</b>	
8.1 Introduction	56
8.2 Zinc-finger structure	56
8.3 Hormone response elements	57
8.4 Modulators of androgen receptor-DNA binding	60
<b>Chapter 9</b>	
<b>TRANSCRIPTION REGULATION</b>	
9.1 Introduction	64
9.2 Transcription initiation	64
9.2.1 Assembly of the basal transcription machinery	64
9.2.2 Gene-specific transcription regulators	66
9.3 Models describing the mechanisms of action of transcription regulators	67
9.4 The diversity issue: mechanisms to achieve receptor/cell-specific gene regulation	68
9.5 Transcription activation units in steroid receptors	69
9.5.1 N-terminal transcription activation units	70
9.5.2 C-terminal transcription activation units	73
9.6 Cooperation between TAUs and between transcription regulators	74
9.6.1 Cooperation between TAUs and modulatory regions	74
9.6.2 Cooperation between transcription regulators	75
9.7 Concluding remarks	76

## Contents

### Chapter 10

#### ANDROGEN RECEPTOR AND DISEASE

10.1 Introduction	78
10.2 Androgen insensitivity syndrome	78
10.3 Kennedy's disease	80
10.4 Prostate cancer and male breast cancer	82

### Chapter 11

#### CONCLUSIONS

85

#### Appendix paper I

Domains of the human androgen receptor involved in steroid binding, transcriptional activation and subcellular localization	89
---	----

#### Appendix paper II

Nuclear import of the human androgen receptor	103
---	-----

#### Appendix paper III

Domains of the human androgen receptor and glucocorticoid receptor involved in binding to the nuclear matrix	119
--	-----

#### Appendix paper IV

Identification of two transcription activation units in the N-terminal domain of the human androgen receptor	137
--	-----

#### Appendix paper V

Changes in abundance of androgen receptor isotypes: effects of ligand treatment, glutamine-stretch variation, and mutation of putative phosphorylation sites	151
--	-----

References	168
------------	-----

Summary	189
---------	-----

Samenvatting	191
--------------	-----

List of publications	194
----------------------	-----

Curriculum vitae	195
------------------	-----

Dankwoord	196
-----------	-----

ABBREVIATIONS

AAD	acidic activation domain
AF	activation function (also referred to as transcription activation function)
AIS	androgen insensitivity syndrome
AR	androgen receptor
ARE	androgen response element
ATP	adenosine triphosphate
bp	base pairs
BSA	bovine serum albumin
CA	cyproterone acetate
cAMP	adenosine cyclic-3':5'-monophosphate
CAT	chloramphenicol acetyltransferase
cDNA	complementary deoxyribonucleic acid
COUP-TF	chicken ovalbumin upstream promoter-transcription factor
DBD	DNA binding domain
Dex	dexamethasone
DHT	5 $\alpha$ -dihydrotestosterone
DNA	deoxyribonucleic acid
DR	direct repeat
<i>dros</i>	<i>Drosophila melanogaster</i>
DRPLA	dentatorubral and pallidolusian atrophy
E <sub>2</sub>	estradiol
EcR	ecdysone receptor
EcRE	ecdysone response element
ER	estrogen receptor
ERE	estrogen response element
ERR	estrogen receptor related
FITC	fluorescein-isothiocyanate
FK506	immunosuppressant drug
GR	glucocorticoid receptor
GRE	glucocorticoid response element
GTF	general transcription factors
h (prefix)	human
HD	Huntington's disease
HRE	hormone response element
hsp	heat shock protein
ICI 176334	Casodex, a trademark of ICI Pharmaceuticals (antiandrogen)
Inr	initiator
IR	inverted repeat
kb	kilo base pairs
Kd	equilibrium dissociation constant
kDa	kilo Dalton
LBD	ligand binding domain
LNCaP	human lymph node carcinoma of the prostate (cell line)
m (prefix)	mouse
MMTV-LTR	mouse mammary tumor virus long terminal repeat
MR	mineralocorticoid receptor
mRNA	messenger ribonucleic acid
NLS	nuclear localization signal
NM	nuclear matrix
NMR	nuclear magnetic resonance
NPC	nuclear pore complex

## Abbreviations

Oct	octamer binding protein
OH-flu	hydroxy-flutamide
PAGE	polyacrylamide gel electrophoresis
pAR	plasmid encoding the androgen receptor
p (prefix)	plasmid
PCR	polymerase chain reaction
pp60 <sup>src</sup>	60 kDa transforming kinase of Rous sarcoma virus
PR	progesterone receptor
PRE	progesterone response element
PSA	prostate-specific antigen
r (prefix)	rat
R1881	17 $\alpha$ -methyl-17 $\beta$ -hydroxyestra-4,9,11-trien-3-one
R5020	17 $\alpha$ ,21-dimethyl-19-nor-pregna-4,9-diene-3,20-dione
rab	rabbit
RAF	receptor associated factor
RAR	retinoic acid receptor
RBA	relative binding affinity
RNA pol II	ribonucleic acid polymerase II
RXR	retinoid X receptor
S	Svedberg unit, sedimentation coefficient
SBMA	spinal and bulbar muscular atrophy
SCA-1	spinocerebellar ataxia type I
SDS	sodium dodecyl sulfate
SSCP	single strand conformation polymorphism
SV40	simian virus 40
T	testosterone
TAD	transcription activation domain
TAF	TBP associated factors
TAU	transcription activation unit
TBP	TATA binding protein
TF	transcription factor
Tfm	testicular feminization
TGase	transglutaminase
THR	thyroid hormone receptor
tk	thymidine kinase
TREpal	thyroid hormone response element (palindromic)
TRITC	tetramethyl rhodamine B isothiocyanate
VDR	1,25-dihydroxy-vitamin D <sub>3</sub> receptor
Zn	zinc ion



**INTRODUCTION AND SCOPE OF THIS THESIS**

1.1 INTRODUCTION

Steroid hormones (including androgens, estrogens, progestins, glucocorticoids and mineralocorticoids) are small hydrophobic molecules derived from cholesterol. They exert profound effects on cell growth, development, differentiation and homeostasis. Steroid hormones mediate their functions through specific intracellular receptors that act as hormone-dependent transcription regulators. The androgens testosterone (T) and dihydrotestosterone (DHT) are essential for sexual differentiation of the male embryo and are involved in the initiation and maintenance of spermatogenesis. Their physiological effects are mediated by the androgen receptor (AR). Upon binding of either T or DHT, the receptor is able to recognize specific DNA sequences, so-called hormone response elements (HREs). These HREs are commonly located in promoter or enhancer regions of target genes. Binding to a HRE triggers the up- or down-regulation of transcription of these genes.

The AR is a member of the large family of nuclear receptors. Structural and functional analysis has revealed a common mechanism of action of these receptors (reviewed in Evans 1988; Green & Chambon 1988; Truss & Beato 1993). The model describing this general pattern of receptor behaviour in a target cell, adapted to the AR, is shown in Figure 1.1.

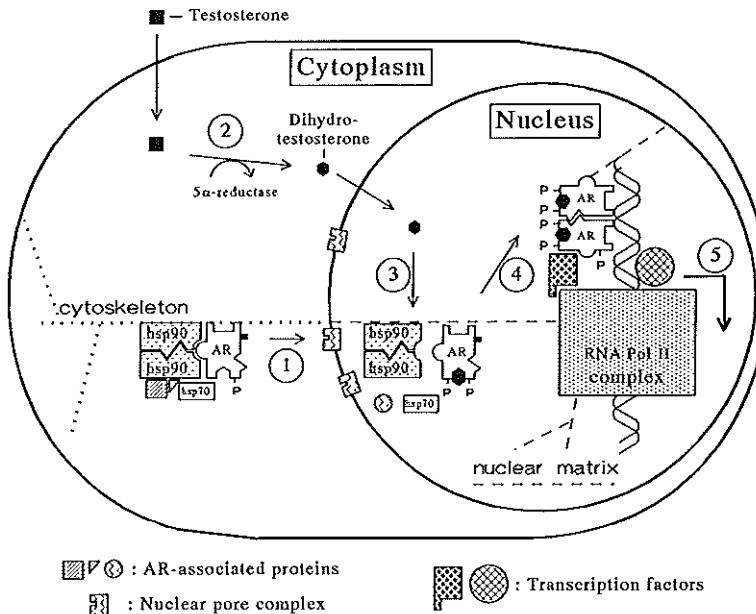


Figure 1.1 Model of androgen receptor function. The key protein in androgen action is the androgen receptor (AR), which becomes phosphorylated (P) and binds to heat shock proteins (hsp) after translation. The receptor is able to enter the nucleus (1) using an intrinsic nuclear localization signal. In some cell types, testosterone (T) can be converted into dihydrotestosterone (DHT) by the enzyme 5 $\alpha$ -reductase (2). Upon T or DHT binding, which may occur in the cytoplasm or nucleus (3), the heteromeric AR-hsp complex dissociates and the AR is able to bind to the nuclear matrix and to DNA as a dimer (4). This triggers the regulation of transcription of specific target genes (5). It is thought that receptor trafficking is facilitated by the cytoskeleton and the nuclear matrix.

After translation, the AR quickly becomes phosphorylated and associates with heat shock proteins (hsps). To be able to act as a transcription factor, the AR needs to pass the nuclear membrane, bind ligand, and interact as a dimer with the HRE (Figure 1.1).

### 1.1.1 THE NUCLEAR RECEPTOR FAMILY

The family of ligand-dependent transcription factors is characterized by a unique modular structure, with receptors classically divided into four domains: 1) an N-terminal modulating domain, 2) a centrally located DNA binding domain (DBD), 3) a hinge region, and 4) a C-terminal ligand binding domain (LBD) (Figure 1.2). The DBD of approximately 70 amino acids is highly conserved. The sequence predicts two zinc-finger motifs, each motif binding a zinc atom between four (tetrahedrally oriented) cysteine residues. The precise DBD tertiary structure has been determined by nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography of several nuclear receptor DBDs (Hård et al. 1990; Schwabe et al. 1990, 1993; Luisi et al. 1991; Katahira et al. 1992; Lee et al. 1993). The C-terminal LBD is also relatively homologous between the various family members. In particular the LBDs of the AR, glucocorticoid- (GR), progesterone- (PR), and mineralocorticoid- (MR) receptors are conserved (Figure 1.2). For full steroid binding capacity, these receptors need a domain of approximately 200 to 250 amino acids (Giguere et al. 1986; Kumar et al. 1986; Rusconi & Yamamoto 1987; Simons et al. 1989; Lees et al. 1989; Jenster et al. 1991 [Appendix paper I]).

SEQUENCE HOMOLOGY		RECEPTOR	CHROMOSOME
N-TERMINAL	DNA	LIGAND	
100	100	100	hAR X
	548 612 662	910	
<15	80	55	hPR 11
	565 629 685	933	
<15	77	50	hGR 5
	419 483 530	777	
<15	77	51	hMR 4
	601 665 736	984	
<15	57	20	hER 6
	183 247 313	553 595	

Figure 1.2 Sequence homology between the androgen receptor (AR), progesterone receptor (PR), glucocorticoid receptor (GR), mineralocorticoid receptor (MR), and estrogen receptor (ER). The hinge region (open box) is located between the DNA binding domain (stippled box) and the ligand binding domain (hatched box), and the sequence and length of this region is not conserved among the steroid receptors. The chromosomal location of the genes encoding these human receptors is indicated (Hollenberg et al. 1985; Green et al. 1986; Arriza et al. 1987; Mishari et al. 1987; Trapman et al. 1988; Faber et al. 1989).

## THE NUCLEAR RECEPTOR SUPERFAMILY

SUBFAMILIE

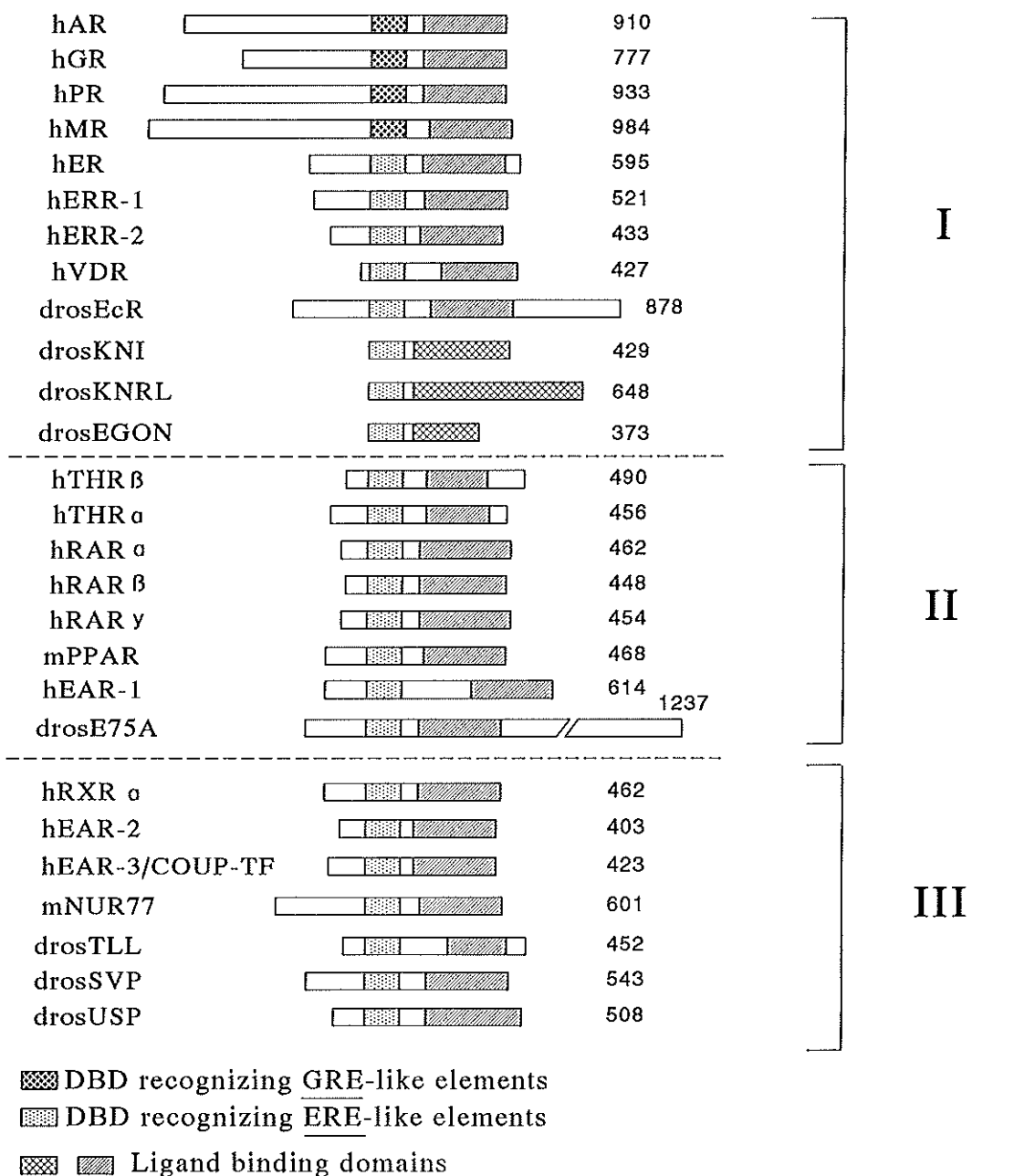


Figure 1.3 Opposite page. *Domain representation of selected members of the nuclear receptor superfamily (based on Wahli & Martinez 1991). The family is divided into three subfamilies according to the phylogenetic tree of the DNA binding domain (DBD) (Laudet et al. 1992). Based upon HRE recognition, the DBDs can be grouped into two classes: GRE-like binders, and ERE-like binders (Chapter 8.2) (Forman & Samuels 1990; Freedman 1992). The ligand binding domains (LBDs) of *Drosophila knirps* (KNI), *knirps-related* (KNRL) and *embryonic gonad* (EGON) show very little similarity to the LBDs of the other receptors, and are therefore represented in a different way.*

The N-terminal domain as well as the hinge region, are highly variable in size and amino acid composition. The N-terminal domain contains a transcription activation function (Appendix paper IV; Giguere et al. 1986; Gronemeyer et al. 1987; Kumar et al. 1987; Jenster et al. 1991 [Appendix paper I]; Simental et al. 1991; Rupprecht et al. 1993). Although the function of the hinge region is unknown, it is believed to be a spacer that flexibly links the LBD and DBD. Since the cloning of the human GR (Hollenberg et al. 1985) and human estrogen receptor (ER) (Walter et al. 1985), more than 60 different members of the family of nuclear receptors have been identified. The large family includes receptors for hormones (steroids, thyroid hormone, ecdysone), vitamins (vitamin D<sub>3</sub>, retinoic acid), and chemical agents (peroxisome proliferators), as well as a variety of receptors lacking a putative ligand. This latter group has been designated orphan receptors (O'Malley 1989; O'Malley & Conneely 1992).

Phylogenetic studies based on the DBD and LBD show a common progeny of nuclear receptors (Laudet et al. 1992; Amero et al. 1992). Through processes such as gene duplication, rearrangement, mutation, exon shuffling and transposition, it is believed that family members originated from a single precursor gene (O'Malley 1989; Dorit et al. 1990; Amero et al. 1992; Keese & Gibbs 1992; Laudet et al. 1992). Based on the phylogenetic tree for the DBD, the family can be grouped into three subfamilies (Figure 1.3) (Laudet et al. 1992). The genes encoding the various family members are located throughout the genome on different chromosomes (Figure 1.2). The AR gene is unique since it is the only member known to be located on the X-chromosome.

The full length AR cDNA has been cloned from several species:

- 1) human (Chang et al. 1988a,b; Lubahn et al. 1988a,b; Trapman et al. 1988; Faber et al. 1989; Tilley et al. 1989; Govindan 1990) (Figure 1.4),
- 2) mouse (Gaspar et al. 1990; He et al. 1990a; Charest et al. 1991; Faber et al. 1991),
- 3) rat (Chang et al. 1988a,b; Tan et al. 1988).

Parts of the AR cDNA encoding the DBD and part of the N-terminal domain have been isolated from dog, guinea pig and clawed frog (He et al. 1990a; Shibuya et al. 1993). Figure 1.4 shows the amino acid sequence of the human AR.

The DBD and the LBD are 100% identical between human, mouse and rat AR. The N-terminal domain and hinge region are highly conserved, with amino acid substitutions scattered over these domains. Most prominent are the variations in size and location of the N-terminal glutamine- and glycine stretches. The difference in length of these two repeats accounts for most of the variability in size between hAR, mAR, and rAR, but also for the variation between the different reported hAR lengths of 910, 917, 918 and 919 amino acids, respectively (Faber et al. 1989; Tilley et al. 1989; Chang et al. 1988b; Lubahn et al. 1988b) (Figure 1.5).

## Chapter 1

```

1    MEVQLGLGRVYPRPPSKTYRGAFAQNLFQSVREVIQNPGPRHPEAASAAPP
51   GASLLLLQQQQQQQQQQQQQQQQQQQQQETSPPRQQQQQGGEDGSPQAHRRG
101  PTGYLVLDDEEQPSQPQSALECHPERGCVPEPGAAVAASKGLPQQLPAPP
151  DEDDSAAPSTLSLLGPTFPGLSSCSADLKDILSEASTMQLLQQQQQEAVS
201  EGSSSGRAREASGAPTSSKDNYLGGTSTISDNAKELCKAVSVSMGLGVEA
251  LEHLSPGEQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLDDDSAGKST
301  EDTAEYSPFKGGYTKGLEGESLGCSSGSAAGSSGTLELPSTLSLYKSGAL
351  DEAAAYQSRDYINFPLALAGPPPPPPPPHPHARIKLENPLDYGSAAWAAA
401  AQCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEBGQLYGPCGGG
451  GGGGGGGGGGGGGEAGAVAPYGYTRPPQGLAGQESDFTAPDVWVYGGMVS
501  RVPYPSPTCVKSEMGPWMSYSGPYGDMRLETARDHVLPIIDYFFPPQ KT
550  CLICGDEASGCHYGALTCGSCKVFFKRAAEGKQKYL CASRNDCTIDKFRR
600  KNCPSCLRKCYE AGMTLGARKLKKLGNLKLQEEASSTTSPTEETTQ
649  KLTVSHIEGYEQPIFLNVLEAIEPGVVCAGHDNNQPDFAALLSSLNEL
699  GERQLVHVVKWAKALPGFRNLHVDDQMAVIOYSWMGLMVFAMGWRSEFTNV
749  NSRMLYFAPDLVFNFEYRMHKSRMYSQCVRMRHLSQEFGWLOITPOEFLCM
799  KALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRRFY
849  QLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFFPEMMAEIIISVQVPKLL
899  SGKVKPIYFHTQ

```

Figure 1.4 Amino acid sequence of the human AR. The amino acids encoding the DBD are boxed, whereas the underlined amino acids represent the LBD. The homopolymeric amino acid stretches in the N-terminal domain (residues 1-547) are bold. Amino acids are indicated as one-letter symbols.

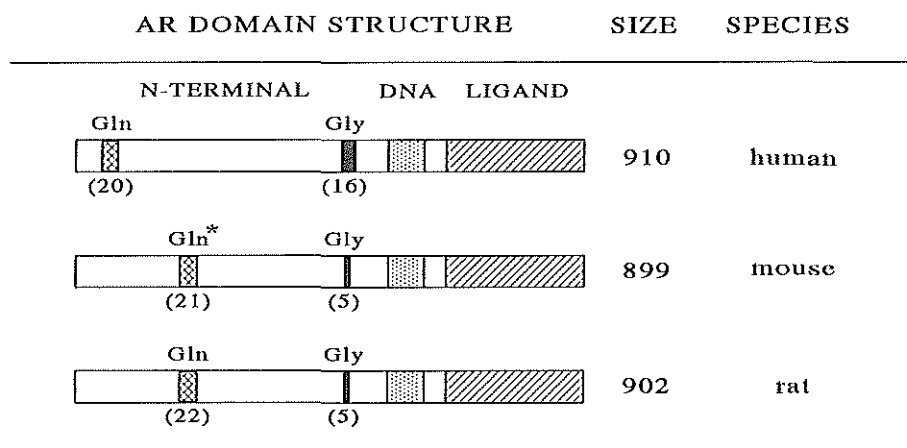


Figure 1.5 Sequence comparison between the human, mouse and rat AR. The amino acid sequences of the DBD and LBD are 100% conserved among these receptors. The mouse AR N-terminal domain is 76% homologous to the human AR N-terminal domain, whereas the rat AR N-terminal domain is 74% homologous to the human AR N-terminal domain. The position and number of residues in the glutamine- and glycine stretches are indicated. In the mouse AR, 3 histidine residues interrupt the 18 glutamine stretch (\*).

### 1.1.2 ANDROGENS AND ANTIANDROGENS

The androgens T and DHT are directly involved in the development and differentiation of the male embryo and in the initiation and maintenance of spermatogenesis. T is produced by the Leydig cells in the testis and secreted into the blood. Bound to plasma proteins, T is transported throughout the whole body. It is believed that the free steroid can reach the AR by diffusion across the cell- and nuclear membranes (Westphal 1978). Specific organs and structures (prostate, urethra, and external genitalia) depend for their growth and development on DHT. In these organs and structures, the differentially expressed  $5\alpha$ -reductase can convert T into DHT. Both androgens bind to the same AR protein with high affinity and specificity.

A problem making T and DHT impracticable in assays using cell line cultures, is their rapid conversion into inactive metabolites. Therefore, stable synthetic androgen agonists such as R1881 (17 $\alpha$ -methyl-17 $\beta$ -hydroxyestra-4,9,11-trien-3-one) and mibolerone (7 $\alpha$ ,17 $\alpha$ -dimethyl-19 nortestosterone) are being used. These agonists answer the criteria of high affinity, although receptor specificity of both compounds is not strict. Both R1881 and mibolerone bind to the AR as well as to the PR.

Antiandrogens, such as hydroxy-flutamide (OH-flu), cyproterone acetate (CA), and Casodex (ICI 176334), have the ability to compete with androgens for AR occupancy without effectuating androgen action themselves. As a result, antiandrogens inhibit androgen action and are used for treatment of several diseases and disorders, such as prostate cancer and hirsutism. In addition, AR antagonists can be useful for *in vitro* (assays using subcellular fractions) and *ex vivo* (assays using cell cultures) studies, since they appear to block specific steps in the AR activation cascade.

### 1.1.3 STRUCTURAL AND FUNCTIONAL UNITS AS BUILDING BLOCKS OF PROTEINS

In accordance with the conserved modular structure of nuclear receptors, the genomic organization of the corresponding genes is very similar. In general, the genes consist of 8 coding exons. The variable N-terminal domain is encoded by part of exon 1, the conserved DBD is encoded by exons 2 and 3, whereas exons 4 to 7 and part of exon 8 encode the LBD (Figure 1.6).

After the discovery that genes consist of exons spaced by non-coding introns, it was suggested that protein evolution could have benefited if exons coded for functional or structural units in proteins (Traut 1988). According to this exon-protein unit theory, novel genes would have been assembled from exons by duplication and intron-mediated recombination. The hypothesis implies that proteins are composed of separable functional/structural domains, and that certain protein units, encoded by discrete exons, can be identified in different proteins (Keese & Gibbs 1992; Dorit et al. 1990). This view is supported by the conserved structural/functional features of large gene superfamilies, such as the nuclear receptor family and the immunoglobulin-like family.

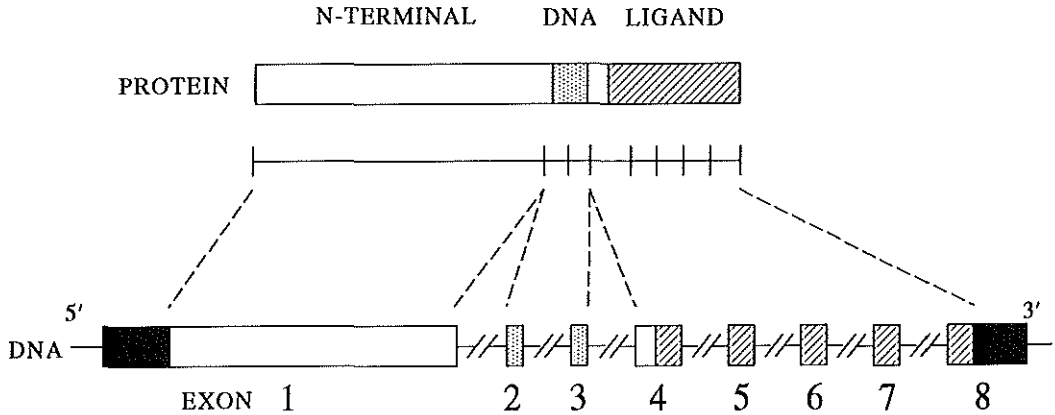


Figure 1.6 Domain structure and gene organization of nuclear hormone receptors. The genes contain 8 exons encoding the different protein domains. Black boxes represent 5' and 3' untranslated sequences.

As genomic organization for different genes encoding structurally defined proteins was determined, there were both positive and negative reports for the correspondence between exons and protein domains. However, the number of positive correspondences favoured the exon-unit theory, since the number appeared to be greater than by chance alone (Traut 1988). It should also be noticed that processes such as intron-loss/addition, junction-sliding/formation, and general evolutionary mutation, can result in degeneration of the original pattern of exons as protein units.

The exon-protein unit theory implies that exon-intron junctions would be at equivalent positions, because shuffling of exons should not alter the reading frame when they are inserted into a new location. From a large survey, the average interruptions between codons was 54%, after the first codon-nucleotide 27%, and after the second codon-nucleotide 18%, indicating the preference for one position (Traut 1988).

Based on the exon-protein unit theory, nuclear receptors (including the AR) would be build up of discrete domains, and the location, size, and boundaries of the domains should correspond with the genomic organization. Since the nuclear receptor homologs are found in organisms ranging from plants and insects to vertebrates, it is believed that evolution of the respective family began one thousand million years ago, and, therefore, some degeneration of the exon-protein unit pattern might be expected (O'Malley 1989).

In theory, an isolated receptor domain would retain its characteristics, and deletion of a specific domain would not necessarily eliminate the function of neighbouring domains. From mutational analysis combined with functionality studies of the GR, ER, and PR, it is evident that the DBD and LBD are discrete, sharply bordered units (reviewed in Muller & Renkawitz 1991; Freedman & Luisi 1993; Gronemeyer 1991). In accordance with the exon-protein unit theory, the two zinc-finger structures of the DBD are each encoded by a single exon. Similarly, the



large N-terminal domain which harbours the important transcription activation function is encoded by a single exon (Gronemeyer 1991, and references therein). In contrast, the discrete LBD is encoded by 5 exons, indicating a deviating pattern from what was to be expected according to the exon-protein unit theory (for more details see Chapter 5.3).

The single domains isolated from their receptor context retain their function. Additionally, comparable domains of different nuclear receptors can be artificially interchanged, resulting in chimeric proteins exhibiting the combined characteristics of the domains from which they were constructed (Green & Chambon 1987; Webster et al. 1988; Rupprecht et al. 1993). Even fusion to unrelated proteins, such as GAL4, *lexA*, E1A,  $\beta$ -galactosidase, *fos*, and *myc*, will transfer the receptor domain characteristic to the hybrid protein (Godowski et al. 1987a; Picard & Yamamoto 1987; Hollenberg & Evans 1988; Webster et al. 1988; Becker et al. 1989; Eilers et al. 1989; Superti-Furga et al. 1991).

Based on these findings, deletion analysis and domain transfer/fusion studies are useful and powerful methods to locate and characterize protein domains. However, there is one major drawback: loss of domain function not only can be caused by the mutation of essential amino acids that are part of the unit, but also indirectly by altered protein structure of neighbouring sequences. Only knowledge of the 3-dimensional protein folding, obtained by NMR or X-ray crystallography analysis, can provide decisive answers to such a problem. Unfortunately, both techniques are difficult and laborious, with limitations on the size of the protein (Wright 1989; Gronenborn & Clore 1990, and references therein). Up to now, only the tertiary structure of the DBD of some nuclear receptors has been elucidated. To be able to analyze other regions of nuclear receptors, the knowledge of function, location and boundaries of the different domains is essential.

### 1.2 SCOPE OF THIS THESIS

The importance to understand the detailed mechanism of action of nuclear receptors is twofold. Originally, interest in steroid hormone receptors stemmed from the central role they play in steroid hormone action. Steroid hormone functions such as regulation of embryonic development, sex differentiation, reproduction, and homeostasis, are mediated by nuclear receptors. Androgens, the male sex steroid hormones, are essential for development of the male genital tract during prenatal life, and important for development and functional maintenance of male sex organs and characteristics postnatally. The AR is the key protein that effectuates androgen action. Mutations in the AR have been implicated in male sex-linked disorders such as androgen insensitivity (AIS) and Kennedy's disease, and in prostate- and male breast cancer.

Besides the central role steroid hormone receptors play in hormone action, they constitute one of the best available model systems for the study of regulation of gene transcription. The attractiveness of the system is the fact that transcription activity of many nuclear receptors is controlled by extra-cellular compounds.

For a better understanding of androgen action in normal and diseased states, a detailed knowledge of the AR characteristics is essential. The aim of this project was to map the different functional/structural domains of the human AR and to study their specific role in AR functioning. The different domains were investigated after construction of AR mutants and subsequent expression in eukaryotic cells. The ability of these receptor proteins to associate with

## Chapter 1

heat shock proteins, to bind ligand, to enter the nucleus, to associate with the nuclear matrix, and to activate transcription, was examined. Furthermore, their migration pattern in SDS-PAGE was analyzed to provide evidence for phosphorylation. In Chapter 2 some post-translational modification events are discussed that might, or are known to modify the AR protein. Especially phosphorylation of steroid receptors is emphasized (Appendix paper V). In the absence of ligand, steroid receptors associate with hsp90. In Chapter 3 the location of the receptor domain that binds hsp90 is examined and the physiological role of this association is discussed. The signal that is responsible for AR nuclear import is described in Chapter 4 and Appendix paper II. The subcellular localization of the AR is dependent on several factors that form the basis of the compartmental retention model. The location and boundaries of the LBD is examined in Chapter 5. In Chapter 6, the domains that are involved in homodimerization of steroid receptors are described. The characterization of the domains of the AR and GR that are important for the association with the nuclear matrix is depicted in Chapter 7 and Appendix paper III. In addition, the functional aspects of the nuclear matrix interaction are discussed. In Chapter 8, the location and the possible 3-dimensional conformation of the AR DBD are presented. Moreover, different HREs that are bound by various nuclear receptors are described. Chapter 9 deals with the main nuclear receptor feature which is transcription regulation. Some basic principles of transcription initiation are described and the role of nuclear receptors in this process is discussed. The problem of how each nuclear receptor is capable of specifically regulating its own set of target genes is addressed in the diversity issue. The location and boundaries of the transcription activation units in the human AR are examined in Chapter 9 and Appendix paper IV. In Chapter 10, the site and kind of mutations in the AR that are associated with human diseases are described. Chapter 11, summarizes the location and size of the various signals and domains in the human AR and indicates the major challenges that confront future nuclear receptor research.

POST-TRANSLATIONAL MODIFICATION

## Chapter 2

### 2.1 INTRODUCTION

After translation, many proteins undergo some type of covalent modification. More than 100 different forms of post-translational modifications of the amino acid side chains are known (Han & Martinage 1992; Krishna & Wold 1993). Only few of these have been investigated or have been suggested to be involved in nuclear receptor protein alterations. These include disulfide bond formation, glycosylation, transamidation, and phosphorylation (Jackson & Tjian 1988; Chakraborti et al. 1990; Green 1993b; Kuiper & Brinkmann 1994a).

#### *Disulfide bond formation*

Paired cysteine residues allow intra- and inter-molecular disulfide bonds to form in proteins. This reversible modification might be important for the correct 3-dimensional folding of steroid hormone receptors. It has been shown that sulfhydryl alterations of the GR inactivated the receptor to a state which was unable to bind glucocorticoids (Bresnick et al. 1988; Silva & Cidlowski 1989; Chakraborti et al. 1990). Furthermore, it has been demonstrated that the oxidation or reduction state of sulfhydryl groups within the GR protein can account for much of its heterogeneity resulting in the 5-6 isoforms observed in two-dimensional gel analysis (Silva & Cidlowski 1992). These reports indicate the existence of disulfide bonds in the GR. The human AR contains 27 cysteine residues. Besides the 8 cysteine residues essential for the formation of the two zinc finger structures of the DNA binding domain (Chapter 8), the importance of the other cysteine residues in protein conformation and AR functioning is unknown.

#### *Glycosylation*

Two different types of glycosylation have been described: 1) N-linked glycosylation common to cell surface and secreted proteins, and 2) O-linked glycosylation found in secreted proteins, but also in some cytoplasmic and nuclear proteins. It is very unlikely that the AR is N-linked glycosylated since this type of protein modification mainly occurs in the endoplasmic reticulum to proteins that are directed to the Golgi apparatus, to lysosomes, to the plasma membrane or are secreted. Although less common, oligosaccharides could be linked to serine, threonine or hydroxy-lysine residues (O-linked glycosylation). Some nuclear proteins involved in transcription initiation and regulation are O-linked glycosylated (e.g. Sp1, RNA pol II, v-erbA) (Jackson & Tjian 1988; Privalsky 1990; Haltiwanger et al. 1992). Except for the v-erbA oncogene, glycosylation has not yet been shown to modify nuclear receptors.

#### *Transamidation*

Only recently, transamidation of the AR N-terminal glutamine stretch has been suggested to explain the causal link between expansion of this glutamine repeat and the development of Kennedy's disease (Chapter 10.3) (La Spada et al. 1991; Green 1993b). Transglutaminases are enzymes that catalyze an acyl transfer reaction between the carboxamide group of a peptide-bound glutamine and the amino group of a lysine residue or polyamine (Greenberg et al. 1991; Han & Martinage 1992; Krishna & Wold 1993) (Figure 2.1). The reaction results in a covalent isopeptide crosslink which is stable and resistant to proteolysis. Because of these characteristics, transglutaminase modified, crosslinked proteins are found in extracellular matrices, in fibrin networks of blood clots and in cornified features of the epidermis and its appendages (callus,

hair, nail) (Greenberg et al. 1991). More enzyme details and the relationship between the glutamine stretch, transglutaminase and Kennedy's disease are described in Chapter 10.3.

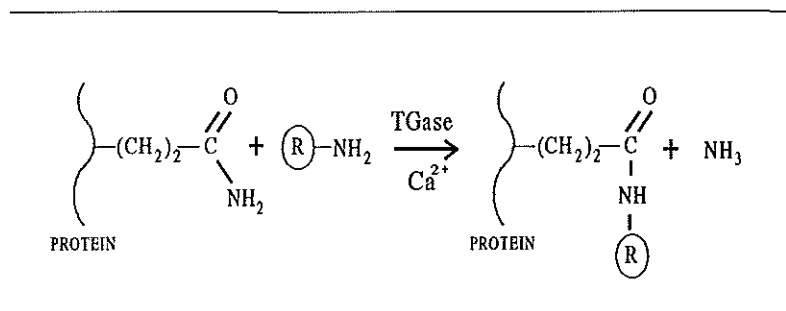


Figure 2.1 Schematic representation of the reaction catalyzed by transglutaminases (TGase). The calcium-dependent acyl transfer reaction results in an isopeptide bond between the 'protein' and the 'R' group which can also be a protein.

## 2.2 ANDROGEN RECEPTOR PHOSPHORYLATION AND ISOTYPES

The best studied steroid receptor post-translational modification process is phosphorylation. Reversible protein phosphorylation is one of the most important mechanisms of regulating intracellular processes. Through phosphorylation and dephosphorylation, protein functions can be switched on or off (Han & Martinage 1992; Krishna & Wold 1993).

Many transcription factors, including steroid hormone receptors, are phospho-proteins (Moudgil 1990; Meek & Street 1992; Hunter & Karin 1992; Ortí et al. 1992; Takimoto & Horwitz 1993; Kuiper & Brinkmann 1994a). Mutational analysis, phospho-amino acid analysis, tryptic phospho-peptide mapping, and microsequencing have identified phosphorylation sites in the GR (Bodwell et al. 1991), PR (Denner et al. 1990b; Poletti & Weigel 1993), ER (Ali et al. 1993; Le Goff et al. 1994), and VDR (Hsieh et al. 1991). Phosphorylation mainly occurs in the N-terminal domain of these receptors, although sites in the hinge region (chicken PR [Denner et al. 1990b]), hinge region and LBD (VDR [Brown & DeLuca 1991; Hsieh et al. 1991]) and in the DNA binding domain (VDR [Hsieh et al. 1991]) have also been reported. Almost all the characterized phosphorylations are on serine residues. In some cases minor phosphorylation on threonine residues (Hoeck & Groner 1990; Bodwell et al. 1991) was observed, but rarely on tyrosine residues (Migliaccio et al. 1986; Auricchio 1989). Several different kinases (protein kinase A, protein kinase C, casein kinase II, Ser-Pro directed kinases, DNA dependent kinase) have so far been implicated in phosphorylation of steroid receptors (reviewed in Moudgil 1990; Ortí et al. 1992).

With respect to the human AR, no individual phosphorylation sites have been identified in the intact AR as yet. From limited proteolysis studies of <sup>32</sup>P-labeled AR isolated from LNCaP cells, it is evident that phosphorylation occurs in the first 300 amino acids of the N-terminal domain (Kuiper et al. 1993). Phospho-amino acid analysis revealed only phosphorylation on serine residues (Kuiper & Brinkmann 1994b submitted). In the first part of the N-terminal

domain approximately 30 serine residues are present. Many of these potential phosphorylation sites comply to a consensus sequence of one of the above described kinases (Kemp & Pearson 1990; Kennelly & Krebs 1991). Analysis of AR fragments expressed in COS cells indicated that serines at position 80, 93 and 641 can be phosphorylated (Zhou et al. 1994b). Whether these serine residues are phosphorylated in the intact AR still needs to be established.

For the PR (Dalman et al. 1988; Sheridan et al. 1989; Beck et al. 1992), ER (Washburn et al. 1991), AR (Kuiper et al. 1991; Krongrad et al. 1991) and VDR (Jurutka et al. 1993) phosphorylation corresponds with a characteristic decrease in electrophoretic mobility (upshift) during SDS-PAGE analysis, resulting in different separable isotypes. Immediately following synthesis, the nascent human AR has an apparent molecular mass of 110 kDa and matures into two bands (110-112 kDa) over the ensuing 30 minutes. Alkaline phosphatase treatment resulted in the single 110 kDa protein band, showing that the upshift to 112 kDa reflects phosphorylation (Appendix paper V, Kuiper et al. 1991).

This migration upshift, resulting in different AR isotypes, can be used as a marker for AR phosphorylation. The region responsible for the doublet appearance observed in the absence of hormone, is located between amino acids 46-139 (Appendix paper V). Substitution of two Ser-Pro directed kinase consensus sites (Ser-80 and Ser-93) revealed that these potential phosphorylation sites are essential for the formation of the 112 kDa AR isotype. Together with the observation that both serine residues can be phosphorylated in an AR fragment expressed in COS cells (Zhou et al. 1994b), it seems likely that Ser-80 and Ser-93 are phosphorylated in at least part of the wild type AR population. Interestingly, deletion of the adjacent glutamine stretch (amino acids 58-77) also resulted in a single protein band. Increasing the length of the glutamine repeat caused an increase in the spacing between the two isotypes of the doublet, showing that the number of glutamine residues determines the extent of the described upshift. Two possible explanations for phosphorylation-induced AR doublet appearance in the absence of hormone can be proposed. Firstly, phosphorylation on sites Ser-80 and Ser-93 results in a conformational change in the neighbouring glutamine repeat causing the migration upshift visible as a 112 kDa isotype. Deletion of the glutamine stretch or elimination of the phosphorylation sites would both result in a single AR protein band. Secondly, Ser-80 and Ser-93 phosphorylation might retard the AR protein migration by decreasing SDS-binding to the AR. The negatively charged detergent SDS provides the negative charge to the SDS-protein complex, and therefore largely determines the migration pattern of proteins in the SDS-PAGE system. Possibly, SDS molecules that bind to the glutamine stretch will be supplanted by the negatively charged phosphate groups of the phosphorylated Ser-80 and Ser-93. Phosphorylation of AR mutants containing a shortened glutamine repeat length might supplant fewer SDS molecules, resulting in a decreased spacing between the two isotypes.

Upon incubation of cells in culture with the respective ligands, steroid receptors become hyperphosphorylated (Ortí et al. 1989; Moudgil 1990; Brown & DeLuca 1990; Ortí et al. 1992; Washburn et al. 1991; Chauchereau et al. 1991; van Laar et al. 1991; Denton et al. 1992; Beck et al. 1992; Kuiper et al. 1993). The increase in AR phosphorylation upon hormone treatment is approximately twofold (van Laar et al. 1991; Kuiper et al. 1993). This doubling is not very dramatic in contrast to other steroid receptors, and Kemppainen and co-workers (1992) even claim that a ligand-induced prolonged AR half-life is responsible for the observed additional phosphorylation.

We have shown by Western blot analysis that R1881 induced a second migration upshift resulting in three AR isotypes (110-112-114 kDa) (Appendix paper V). These ligand-induced changes could be due to additional phosphorylation and/or a re-location of phosphorylation sites. Analysis of SDS-PAGE migration patterns of AR mutants expressed in hormone-treated cells, provided evidence for the importance of Ser-80 and involvement of transcription activation in the hormone-induced upshift. Even upon R1881 treatment, the AR mutant in which Ser-80 was replaced by a glycine residue, migrated as a single protein band, suggesting that this phosphorylation site is essential for the ligand-induced AR modifications. Interestingly, the migration pattern of AR mutants, unable to activate transcription due to elimination of DNA binding or due to N-terminal deletions, revealed a decreased amount of the hormone-induced isotype compared to the other isotypes. This might indicate that part of the hormone-induced phosphorylation occurs during or following transcription activation (Appendix paper V). Similar observations have been made for the human progesterone receptor, where both hyperphosphorylation and coincident PR upshifts were partially inhibited in PR mutants unable to bind DNA or by antagonist-treatment that prevented PR DNA interaction (Bagchi et al. 1992; Takimoto et al. 1992).

A possible order of phosphorylation events is proposed in Figure 2.2. After synthesis, the AR rapidly becomes phosphorylated. Upon hormone-treatment, additional phosphorylation occurs in a ligand-dependent manner, possibly before the AR is bound to DNA. In the last step phosphorylation follows transactivation, possibly by a DNA-dependent kinase.

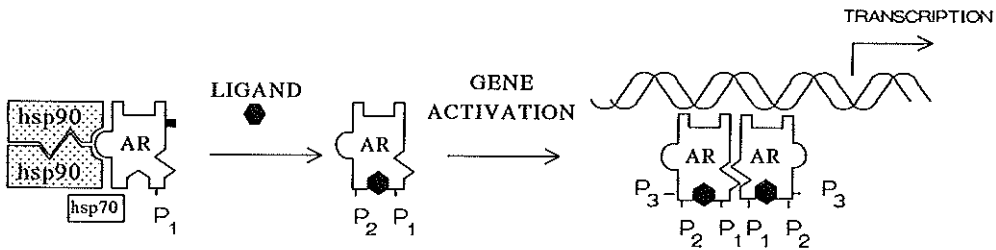


Figure 2.2 Model describing a three stage AR phosphorylation cascade. Immediately after translation, the AR can be phosphorylated (represented by the class of phosphorylation sites P1). Upon ligand binding additional phosphorylation can occur on phospho-sites represented by P2. During or after transcription regulation, the AR becomes fully phosphorylated by phosphorylation on sites represented by P3.

### 2.3 FUNCTION OF ANDROGEN RECEPTOR PHOSPHORYLATION

Virtually every step in the steroid receptor activation cascade, from hormone binding to transactivation has been postulated to be influenced by changes in receptor phosphorylation (reviewed in Orti et al. 1992; Kuiper & Brinkmann 1994a). The importance of this reversible protein modification has therefore often been emphasized. Site-directed mutagenesis of

phosphorylation sites in the human ER and human VDR revealed a reduced capacity to activate transcription of the respective mutants (Hsieh et al. 1991; Ali et al. 1993; Le Goff et al. 1994). However, functional analysis of mouse GR mutants in which known phosphorylation sites had been eliminated, showed no, or only minor changes in transactivating capacity (Mason & Housley 1993).

For the human AR, it has been shown that both the 110 and 112 kDa isoforms are able to bind hormone and can undergo the ligand-dependent transformation to a tight nuclear binding form (Kuiper et al. 1991). This indicates that the 110 to 112 kDa conversion caused by phosphorylation, is not obligatory for hormone binding or tight nuclear binding. As described above, it is suggested that hormone-induced phosphorylation of the AR occurs during or after transcription activation, implying a possible role for phosphorylation in AR transactivity (Appendix paper V). However, substitution of the proposed phosphorylation sites Ser-80 or Ser-93 had no significant effect on the AR's capacity to activate transcription of a (GRE)<sub>2</sub>tkCAT reporter gene in *ex vivo* transfection assays. Similarly, Mason & Housley (1993) have shown that elimination of mouse GR phosphorylation sites had only minor effects on transactivity. So, from the available data, it has to be concluded that for the human AR and mouse GR, the functional significance of phosphorylation is still essentially unknown.

### 2.4 ACTIVATION OF NUCLEAR RECEPTORS VIA MODULATION OF KINASE-MEDIATED SIGNAL TRANSDUCTION PATHWAYS

Over the past four years new evidence has accumulated that connects phosphorylation to nuclear receptor functioning. It was shown that modulation of kinase-mediated signal transduction pathways could affect nuclear receptor activity (reviewed in Power et al. 1992; Takimoto & Horwitz 1993; Burnstein & Cidlowski 1993). Stimulation of the cAMP-dependent kinase pathway with 8-bromo-cAMP, and the protein phosphatase inhibitor okadaic acid, mimicked progesterone action by inducing activation of a PRE-tk-CAT reporter in transiently transfected CV-1 cells (Denner et al. 1990a). Since this first report, many articles have been published on ligand-independent receptor activation, and on cooperativity between ligand-dependent activation and other signal transduction pathways (e.g. protein kinase A and protein kinase C pathways) (reviewed in Power et al. 1992; Takimoto & Horwitz 1993; Burnstein & Cidlowski 1993). Also for the AR, synergistic activation by ligand and the cAMP-dependent kinase pathway has been demonstrated. Transcription activation of the MMTV promoter in T47D cells is small in the presence of DHT or 8-bromo-cAMP separately, but simultaneous addition of DHT and 8-bromo-cAMP synergistically stimulated AR transactivity up to 30-fold (Sartorius et al. 1993). The precise mechanism by which kinase modulators influence receptor-mediated transcription is largely unknown. For the ER, the enhancing effects of protein kinase A activators on ER functioning might be explained by an increased level of ER phosphorylation (Aronica & Katzenellenbogen 1993; Le Goff et al. 1994). However, although these activators enhanced ligand-mediated PR and GR transactivity, no additional phosphorylation of the PR and GR could be detected (Moyer et al. 1993; Sartorius et al. 1993).

The exciting discovery of the existence of ligand-independent activation pathways raises a lot of questions, particularly on the mechanisms of receptor activation and on the significance for *in vivo* nuclear receptor functioning. Especially for orphan receptors, the non-classical



concept of ligand-independent activation might be applicable and could change the field of research for orphan receptor activators (O'Malley & Conneely 1992). Although the *in vivo* significance of non-classical receptor activation is still unclear, future research on nuclear receptor functioning must take this new concept into account.

### 2.5 ANOMALOUS ANDROGEN RECEPTOR MOBILITY DURING SDS-PAGE

Similar to what has been observed for several other steroid hormone receptors, there exists a significant discrepancy between the predicted molecular mass based on the AR sequence (98.4 kDa) and the AR protein size determined by SDS-PAGE (approximately 110 kDa). The region responsible for this mobility retardation of approximately 12 kDa is partially deleted in an AR mutant lacking amino acids 51 to 211. The calculated molecular mass (81.2 kDa) and the SDS-PAGE size determination (~84 kDa) of this mutant are comparable (Appendix paper V). Recently, it has been shown that the mobility retardation of the mouse GR can be attributed to an N-terminal region of approximately 200 amino acids containing all of the identified phosphorylation sites within the receptor (Hutchison et al. 1993). With the SDS-PAGE analysis of many different AR deletion mutants (Appendix papers I-V), it became evident that the first part of the AR N-terminal domain (amino acids 46-244) is responsible for the discrepancy between predicted and SDS-PAGE determined molecular mass. Comparable to the mouse GR, this region contains most, if not all AR phosphorylation sites (Kuiper et al. 1993). The basis for the molecular mass discrepancy is unknown, although it is unlikely to be caused by phosphorylation, since alkaline phosphatase treatment had no effect on the apparent 110 kDa molecular mass of the wild type AR determined by SDS-PAGE. Possibly, the region between amino acids 46 and 244 binds relatively few SDS molecules resulting in a less negative charged protein and the observed mobility retardation during SDS-PAGE. Somehow it seems impossible to attribute this mobility retardation to a smaller domain, since the SDS-PAGE molecular masses of AR mutants with smaller deletions (of approximately 50 residues) covering the region of amino acids 46 to 244, were all 10 to 12 kDa above the calculated sizes (Appendix paper V; Table 1).

**Chapter 2**

ASSOCIATION WITH HEAT SHOCK PROTEINS

### 3.1 INTRODUCTION

In the absence of ligand, steroid hormone receptors can be recovered from cells as large heteromeric complexes. These complexes have an apparent molecular mass of approximately 250-300 kDa and sediment with a sedimentation coefficient of 8-10 S in sucrose gradients. The steroid receptor in this complex is unable to bind DNA. Although the unliganded receptor might be nuclear in intact cells, upon cell fractionation into cytosolic and nuclear extracts, the receptor complex is predominantly cytosolic. Hormone treatment of cells results in the dissociation (partly or completely) of the heteromeric complex. The steroid receptor then shows a sedimentation coefficient of 4-5 S, is able to bind DNA and is predominantly present in the nuclear extract of ruptured cells (which is termed tight nuclear binding). This receptor transition from a non-DNA binding to a DNA binding state is termed transformation.

Most proteins that are part of the heteromeric complex have been characterized. Besides the steroid receptor they include heat shock proteins (hsps) with an apparent molecular mass of 90 kDa (hsp90), 70 kDa (hsp70) and 56-59 kDa (hsp56, also referred to as p56 and p59). These heat shock proteins were found to be present in AR, PR, GR and ER complexes (reviewed in Smith & Toft 1993). In addition, proteins with a molecular mass of approximately 60 kDa, 54 kDa, 50 kDa and 23 kDa were found to be associated with the chicken PR (reviewed in Smith & Toft 1993).

### 3.2 STRESS AND HEAT SHOCK PROTEINS

Hsps were originally discovered as newly synthesized proteins after the exposure of cells to increased temperatures. Later it was found that also certain chemicals and pathophysiological conditions (e.g. infections, tissue damage) induced transcription and translation of the same hsps (reviewed in Burdon 1986; Kaufmann 1990; Schlesinger 1990). Nowadays, many different proteins that are induced upon stress have been characterized in many different organisms, showing the universality of stress-response. The hsps are categorized in distinct classes based upon their molecular masses, which range from 15 to 110 kDa (reviewed in Burdon 1986; Kaufmann 1990). Interestingly, several hsps are constitutively expressed under normal physiological conditions. The above described hsp90, hsp70 and hsp56 belong to this group. The importance of hsps in cell functioning and stress-response, has been emphasized by deletion of some of the respective genes. After "knocking out" certain hsp genes, cells no longer could resist heat treatment, which ultimately resulted in cell death (Saito & Uchida 1977; Itakawa & Ryu 1979, Craig & Jacobson 1984, Picard et al. 1990).

Stress-response is believed to be a rapid, transient reprogramming of cellular activities to ensure survival during the stress period, by protection of essential cell components against damage, and by refolding or breakdown of damaged proteins. Under normal physiological conditions, several functions have been ascribed to constitutively expressed hsps. These include: 1) protein folding, 2) assembly of oligomeric protein complexes, 3) unfolding of proteins and subsequent translocation into mitochondria, endoplasmic reticulum or other cell organelles, and 4) protein degradation (reviewed in Burdon 1986; Hershko 1988; Ellis & Hemmingsen 1989; Beckmann et al. 1990; Hartl & Neupert 1990; Kaufmann 1990; Schlesinger 1990; Baker & Schatz 1991; Kelley & Georgopoulos 1992; Pratt 1993; Pratt et al. 1993). Except for protein

degradation, the other three functions are described by the so-called molecular chaperone concept. Chaperones are proteins involved in the control of the folding, transport, localization and functioning of other proteins (Kelley & Georgepoulos 1992). It is thought that in the situation of stress, the basic function of hsp is boosted by their increased expression (Burdon 1986).

### 3.3 THE HSP ASSOCIATING DOMAIN IN STEROID RECEPTORS

Much effort has been put in the identification of the steroid receptor region that contacts the hsp, especially hsp90. Knowledge on the intermolecular contact site(s) might indicate the functional role of associated hsp. Studies with steroid receptor mutants have revealed that the LBD is essential for hsp90 association (Pratt et al. 1988; Carson-Jurica et al. 1989; Chambraud et al. 1990; Housley et al. 1990; Howard et al. 1990; Cadepond et al. 1991; Schowalter et al. 1991). Since the LBD is so large (approximately 200-250 amino acids), smaller regions might be responsible for hsp binding, and it was speculated that a small conserved sequence was the site of hsp90 interaction (Danielsen et al. 1986; Housley et al. 1990; Dalman et al. 1991a). Recent studies, however, do not support the view that a relatively small receptor region is responsible for interaction with hsp90. Deletion of the entire LBD eliminated hsp90 association, but deletion of separate subregions (less than one third of the total LBD) did not affect hsp binding (Carson-Jurica et al. 1989; Chambraud et al. 1990; Cadepond et al. 1991; Schowalter et al. 1991). These observations can be explained by the assumption that there is more than one (sequence specific) contact site, or that the receptor-hsp interaction would be a kind of 'casting-mould' interaction, which is not determined by direct interactions between specific amino acids, but rather by recognition of a modular 3-dimensional conformation of the receptor by the hsp (a similar model is visualized in Figure 9.4 in Chapter 9.5.1). Besides the LBD, the nuclear localization signal of the human ER has been implicated to be necessary, but not sufficient, for complex formation (Chambraud et al. 1990).

With respect to the AR, no studies have been published on the receptor region responsible for hsp association. Few experiments have been performed which show that also for the AR, the LBD is necessary for hsp90 binding. AR mutants, lacking part of the LBD, were expressed in COS-1 cells in the absence of hormone, and immunoprecipitated from the cytosolic fraction using an antibody directed against the human AR. Figure 3.1 shows that hsp90 could be visualized in the wild type AR sample (lane 2), indicating, via co-immunoprecipitation, that the AR binds hsp90. AR17, lacking the last C-terminal 114 amino acids, did not bind hsp90 as visualized by its inability to co-immunoprecipitate hsp90 (lane 3). Control transfections of COS-1 cells with pTZ19 carrier plasmid showed that hsp70 co-immunoprecipitated non-specifically with the antibody-coated agarose beads (lane 1). From these experiments it can be concluded that the last C-terminal 114 amino acids are necessary for hsp90 binding.

Although the total number of proteins present in the heteromeric complex is still a point of discussion, the molecular ratio of hsp90, hsp70, hsp56 and steroid receptor molecules in the complex is generally agreed upon. One receptor molecule is associated with two hsp90, one hsp70 and one hsp56 molecules (Figure 3.2) (Pratt 1993; Smith & Toft 1993).

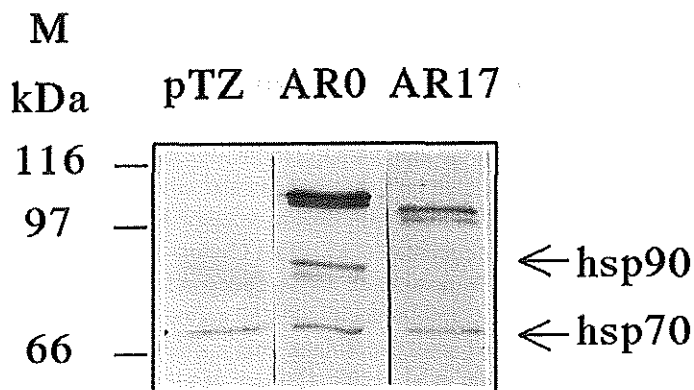


Figure 3.1 Co-immunoprecipitation of hsp90 and hsp70 with AR protein. COS-1 cells were transiently transfected with pTZ19 carrier, wild type AR (AR0), or AR17 plasmids, by the DEAE-dextran method in the absence of hormone (see experimental procedures in Appendix paper V). AR17 lacks the last C-terminal 114 amino acids. Cytosolic extracts were prepared and the AR complex was immunoprecipitated using the monoclonal antibody F39.4.1 coupled to goat-anti-mouse agarose beads. Samples were heated in sample buffer, and subjected to SDS-PAGE on a 7% gel. After separation and blotting, AR, hsp70 and hsp90 proteins were visualized using specific monoclonal antibodies F39.4.1, AC88 (anti-hsp90) and N27 (anti-hsp70) as described previously (Kuiper et al. 1991) (AC88 and N27 were generously provided by Dr. D.O.Toft and Dr. W.J.Welch).

### 3.4 ROLE OF STEROID RECEPTOR ASSOCIATED PROTEINS

The observation that different hsps are abundantly expressed in all cell types of all organisms and associate with many different proteins, implies that these chaperones are part of some sort of cellular core facility. This facility would be responsible for proper protein folding and protein transport to the correct site in the target organelle. Steroid hormone receptors would be just one group among many proteins that make use of this facility (Pratt 1993; Smith and Toft 1993). Normally, most of the *de novo* synthesized cellular proteins probably have a transient interaction with hsps, and are released once the mature conformation is obtained. Steroid hormone receptors, however, may reach a final mature conformation only after ligand binding, which could explain the continuous hsp association in the absence of hormone (Smith & Toft 1993). Not all nuclear receptors are continuously complexed with hsps in the absence of ligand. Both the thyroid hormone receptor (THR) and the retinoic acid receptor (RAR) appear not to be associated with hsp90 (Dalman et al. 1990, 1991b). Interestingly, these receptors are tight nuclear bound even in the absence of ligand (Samuels et al. 1988; Nervi et al. 1989), which indicates that hsp association (in particular hsp90) inhibits certain receptor features such as tight nuclear binding.

Hsp-receptor association has been implicated to be involved in ligand binding, intracellular receptor trafficking and nuclear import. The finding that hsp90 binds to the LBD of steroid hormone receptors suggested that hsp90 is important for the receptor in acquiring ligand binding capacity (Bresnick et al. 1989; Nemoto et al. 1990; Scherrer et al. 1990). Glucocorticoid receptors that had never been associated with hsp90 were unable to activate transcription in yeast (Picard et al. 1990). These studies might indicate the involvement of hsp90 in proper LBD folding as one of the first essential steps in the receptor activation cascade. The necessity of hsp90 association for ligand binding, however, seems to be GR specific. The PR, ER and AR are able to bind hormone even in the absence of hsp90 (Ohara-Nemoto et al. 1991; Nemoto et al. 1992; Smith & Toft 1993).

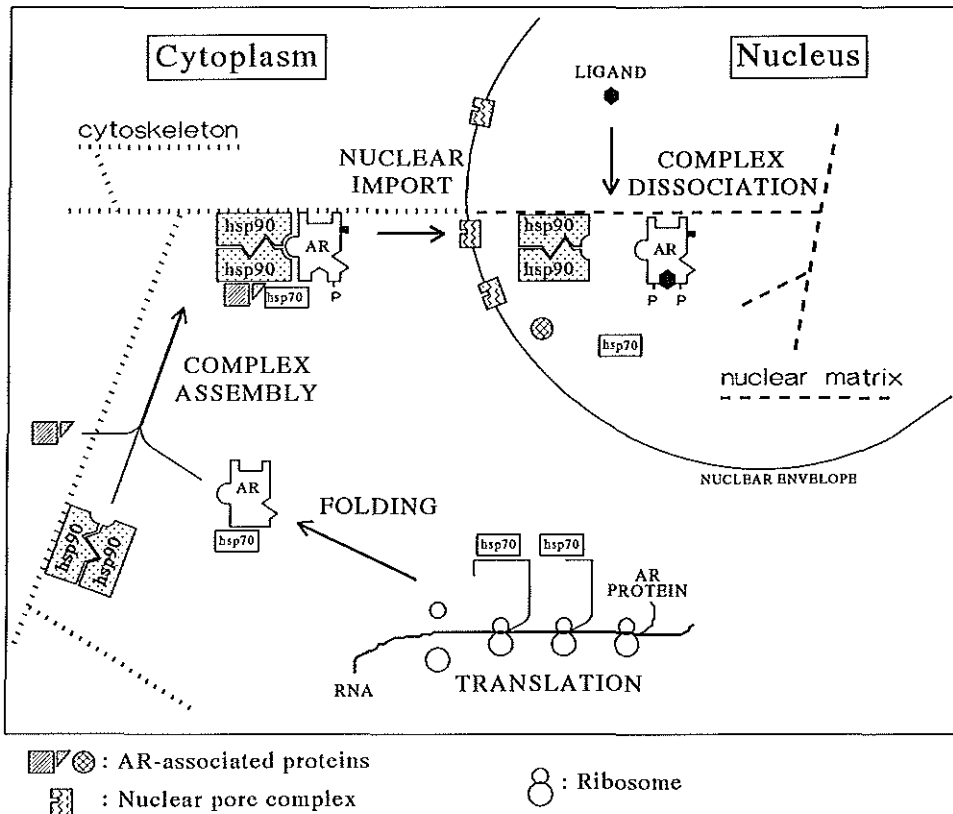


Figure 3.2 Model for chaperone function of heat shock proteins that are associated with the androgen receptor. Hsps could facilitate AR protein folding immediately after translation. Through a mechanism that involves interaction of hsps with the cytoskeleton, the AR could be transported to the nucleus and translocated across the nuclear membrane. Upon ligand binding, the heteromeric complex dissociates and the AR is released from the hsps and able to function as a transcription regulator. Ligand binding followed by complex dissociation might also occur in the cytoplasm.

## Chapter 3

The observation that hsp90 and hsp56 are associated with filamentous cytoplasmic structures (Redmond et al. 1989; Fostinis et al. 1992; Ruff et al. 1992) and are involved in intracellular trafficking of pp60<sup>v-src</sup>, and possibly also in the trafficking of the vitamin D<sub>3</sub> receptor and GR (Brugge 1986; Akner et al. 1991; Miyata & Yahara 1991; Barsony & McKoy 1992), led to the hypothesis that hsps might be part of a cellular protein trafficking system (Pratt 1993; Pratt et al. 1993). After the cloning and sequencing the cDNA encoding hsp56 it became evident that hsp56 is an immunophilin (Lebeau et al. 1992). Immunophilins are proteins that specifically bind immunosuppressive agents such as cyclosporin A, FK506 and rapamycin. All the members of this family have peptidyl-prolyl cis-trans isomerase activity that catalyses the cis-trans isomerization of proline peptide bonds, important for protein folding (reviewed in Schreiber 1991; Walsh et al. 1992; Pratt et al. 1993). The presence of a conserved stretch of 8 acidic amino acids led to the speculation that hsp56 binds to the basic nuclear localization signal (NLS) of steroid receptors, implying a functional role in nuclear import (for more details on the AR NLS see Chapter 4). This proposed model could explain the requirement of the ER NLS for detection of the complete 9 S heteromeric ER complex (Chambrud et al. 1990). Although there is no direct and decisive evidence for the existence of trafficking systems, nor for the involvement of hsps in such a facility, speculations on the few published observations results in an attractive concept (Pratt 1993; Pratt et al. 1993) (Figure 3.2).



**SUBCELLULAR LOCALIZATION AND NUCLEAR IMPORT**

### 4.1 INTRODUCTION

To be able to regulate the expression of target genes, the AR needs to be translocated to the nucleus. Proteins enter the nucleus through so called nuclear pore complexes (NPCs). NPCs are large supramolecular structures that provide aqueous channels with a functional diameter of 9-10 nm, allowing diffusion of ions, small molecules, and proteins with molecular masses up to 40-60 kDa (reviewed in Wagner et al. 1990; Garcia-Bustos et al. 1991; Nigg et al. 1991; Silver 1991). The import of larger proteins such as steroid receptors, is an active process (temperature- and ATP dependent) and requires that these proteins contain suitable nuclear localization signals (NLSs) (Peters 1986).

Besides nuclear import, export out of the nucleus is also facilitated by NPCs, indicating the possibility of a dynamic bidirectional transport. Such a nucleo-cytoplasmic shuttling has been demonstrated for the PR, ER and GR (Guiochon-Mantel et al. 1991; Dauvois et al. 1993; Madan & DeFranco 1993). Under energy-depleting conditions, the ER, PR and GR become cytoplasmic, indicating receptor export out of the nucleus and inhibition of nuclear import. In cell hybrids that originated from cells of different species of which only one species cell type contained the receptor in the nucleus, the steroid receptor could be visualized in all different nuclei of the heterokaryon. This demonstrated that protein migration from one nucleus to the other had occurred via a reversible shuttle mechanism between cytoplasm and nuclei (Guiochon-Mantel et al. 1991; Chandran & DeFranco 1992; Dauvois et al. 1993; Madan & DeFranco 1993).

The subcellular localization of steroid receptors in the absence of ligand is still a point of discussion. Although it is generally agreed upon that the ER and PR are predominantly nuclear, the localization of the unliganded GR is different in various reports, and is dependent on the cell line expressing the receptor (Gasc et al. 1989; Sanchez et al. 1990). Moreover, Brink and co-workers (1992) have shown that the subcellular localization of the unliganded GR can be influenced by the fixation and permeabilization conditions of the immunocytochemical procedure applied to visualize the GR. Contradictory findings have been published also for the AR. *In vivo* immunohistochemical studies of androgen target tissues demonstrated that the AR was nuclear in either the absence and presence of ligand (Chang et al. 1989a; Husmann et al. 1990; Sar et al. 1990; van der Kwast et al. 1991). In contrast, the AR transiently overexpressed in various cell lines demonstrated a predominant cytoplasmic localization in COS-1 cells, an evenly distribution over cytoplasm and nucleus in CHO and CV-1 cells, and a nuclear localization in HeLa cells, all in the absence of hormone (Simental et al. 1991; Jenster et al. 1993 [Appendix paper II]). Hormone treatment resulted in a nuclear AR staining in all cell lines. The differences observed between the *in vivo* immunohistochemistry and *ex vivo* transfection analysis might be due to different fixation- and staining techniques, and might be due to protein overexpression in cell lines such as COS-1 cells, that are not AR target cell types. Moreover, in the *in vivo* situations of castrated rats or patients under androgen ablation therapy, low T and DHT levels might still contribute to the nuclear localization of the AR.

Besides the nucleus, mitochondria also could be a cell organelle to which nuclear receptors are transported. The observation that [<sup>3</sup>H]-cortisol is found in the mitochondrial fraction of rat liver cells, and that cortisol treatment resulted in mitochondrial RNA synthesis, suggested a role for the GR in the regulation of expression of mitochondrial genes (Beato et al. 1969; Mansour & Nass 1970). Recently, also import of the GR protein into mitochondria was

demonstrated to occur *in vivo* and *in vitro* (Demonacos et al. 1993).

#### 4.2 NUCLEAR LOCALIZATION SIGNALS IN THE HUMAN ANDROGEN RECEPTOR

In the past 10 years, several different NLSs have been described to be responsible for the nuclear import of proteins. One common feature is the relatively high content of the basic amino acids lysine and arginine (reviewed in Garcia-Bustos et al. 1991; Silver 1991). One of the best characterized NLSs is the targeting signal of simian virus 40 (SV40) large T antigen, which consists of 7 amino acids, 5 of which are positively charged (Kalderon et al. 1984a,b; Lanford & Butel 1984). A similar basic signal is present in the hinge region of steroid receptors, and its importance for nuclear import has been demonstrated for the GR, PR, ER and AR (Picard & Yamamoto 1987; Guiochon-Mantel et al. 1989; Simental et al. 1991; Cadepond et al. 1992; Ylikomi et al. 1992; Jenster et al. 1993 [Appendix paper II]). However, in steroid receptors this SV40 large T antigen-like NLS is only part of the complete NLS. In the nuclear protein nucleoplasmin, an NLS has been identified that consists of 16 amino acids which includes a basic lysine sequence, similar to the SV40 large T antigen NLS. Two basic amino acids (Lys, Arg) followed by a 10 residues spacer precede the lysine stretch, forming a bipartite NLS (Robbins et al. 1991) (Table 4.1). This conserved bipartite NLS is present in many other nuclear proteins, including nuclear receptors (Dingwall & Laskey 1991; Friedberg 1992).

Table 4.1 Nuclear localization signals in various proteins

Protein	Species	Sequence
SV40 large T antigen	SV40	126 pKKRRKv
Nucleoplasmin	Xenopus	152 kavKRpaatkagq <b>a</b> KKKKl
Androgen receptor	human	605 crl <b>R</b> Kcyeagmtlga <b>R</b> Kl <b>K</b> Kl
Progesterone receptor	human	622 crl <b>R</b> Kccqagmvlgg <b>R</b> Kf <b>K</b> Kf
Glucocorticoid receptor	human	476 cry <b>R</b> Kclqagmnl <b>e</b> a <b>R</b> Kt <b>K</b> Kk
Mineralocorticoid receptor	human	658 crl <b>q</b> Kclqagmnlga <b>R</b> Ks <b>K</b> Kl
Estrogen receptor	human	240 crl <b>R</b> Kcyevgmmkgg <b>i</b> R <b>K</b> d <b>R</b> Rg

*Comparison of amino acid sequences of NLSs in SV40 large T antigen, nucleoplasmin, and various steroid hormone receptors. The numbers refer to the first amino acid of the sequence. Basic residues thought to be important for the signal, are shown in bold capitals. Amino acids are indicated as one-letter symbols.*

Analysis of the subcellular localization of a series of AR deletion mutants, transiently transfected into COS-1 cells, revealed that the hinge region, containing the nucleoplasmin-like NLS, was essential for nuclear import (Jenster et al. 1993 [Appendix paper II]; Zhou et al. 1994a). Deletions in the N-terminal domain or DBD had no effect on the ability of the AR to enter the nucleus. AR mutants lacking the LBD were constitutively nuclear in contrast to the predominantly cytoplasmic localization of wild type AR in the absence of ligand. This indicates

## Chapter 4

that the LBD inhibits nuclear localization of the AR in the absence of ligand when analyzed in COS-1 cells. Detailed deletion- and substitution analysis of the AR nucleoplasmin-like NLS revealed that mutation of either basic domain only partially affected nuclear import. Simultaneous mutation of both positively charged domains, resulted in an AR mutant that was exclusively cytoplasmic in the absence of hormone, and predominantly cytoplasmic with only a faint nuclear staining in the presence of R1881. This shows that the AR NLS is functionally similar to the nucleoplasmin NLS. The observation that part of the receptor population of AR mutants lacking the entire NLS was still localized in the nucleus, might indicate the presence of additional NLSs (Jenster et al. 1993 [Appendix paper II]).

For the ER and GR, other sequences besides the conserved nucleoplasmin-like NLS are involved in nuclear import. Short amino acid stretches containing a high amount of lysine and arginine residues, located in the hinge region of the ER, contribute to the capacity of the ER to enter the nucleus (Ylikomi et al. 1992). These sequences are not conserved among the steroid receptors. In the GR, a second NLS is located in the LBD (Picard & Yamamoto 1987; Cadepond et al. 1992). This signal is hormone-dependent and capable of directing a  $\beta$ -galactosidase fusion protein into the nucleus (Picard & Yamamoto 1987).

Several *ex vivo* studies provided positive evidence for the capacity of the conserved nucleoplasmin-like NLS to direct proteins into the nucleus, by linking the entire, or part of the NLS, to large extranuclear proteins such as  $\beta$ -galactosidase, pyruvate kinase and bovine serum albumin (Picard & Yamamoto 1987; Dang & Lee 1989; Addison et al. 1990; Robbins et al. 1991; Hamy et al. 1992; Schreiber et al. 1992). Recently, Zhou and co-workers (1994a) showed that the NLS (amino acids 604-624) of the hAR, when linked to pyruvate kinase, was not capable of inducing nuclear localization. The minimal region (amino acids 571-652) necessary to direct the chimeric protein to the nucleus included part of the DBD, the NLS, and most of the hinge region. This indicates the weakness of the AR-NLS and/or the necessity of neighbouring residues for proper NLS functioning.

### 4.3 ANDROGEN RECEPTOR CLUSTERING

During the subcellular localization studies, an interesting observation was made concerning the intranuclear staining pattern of AR mutants that were mutated in their DBD. In addition to the diffuse granular distribution of the wild type AR, large clusters of AR protein were present in the nucleus of cells expressing ARs mutated in the DBD (Jenster et al. 1993 [Appendix paper II]). These clusters were only observed in the presence of hormone. A similar clustering pattern was noticed for PRs mutated in their DBD (Takimoto et al. 1992), but not for comparable GR mutants (Appendix paper III). If, in addition to mutations in the DBD, the LBD was deleted, no clustering of the respective AR mutants was observed. This indicates that the LBD is necessary or responsible for the clustering phenomenon. Deletion of the NLS in combination with DBD mutations resulted in cytoplasmic clustering (Jenster et al. 1993 [Appendix paper II]; Zhou et al. 1994a).

The clustering phenomena which is dependent on the presence of the LBD and hormone, and on a mutated DBD, is unexplained up to now. Starting from the principle that upon ligand binding the AR follows a fixed pathway (like running a computer programme), one might expect that a receptor molecule, which is unable to bind DNA to stop during, or deviate from the

programme at the position of its last executable step. So, clustering may represent either a normal AR intermediate state, or ARs directed to sites for proteolysis. To investigate the possibility that cytoplasmic clusters could represent AR proteolysis in lysosomes, an AR construct and a construct encoding the lysosomal enzyme  $\alpha$ -glucosidase (Hoefsloot et al. 1990) were co-transfected in COS-1 cells (Figure 4.1). AR27 which is mutated in its NLS (deletion amino acids 620-625 and substitution of Arg-Lys  $\rightarrow$  Gly-Ala at position 608 and 609) was unable to enter the nucleus, and, due to mutations in the first zinc-finger (567 Cys $\rightarrow$ Ser; 570 Cys $\rightarrow$ Phe), unable to bind DNA, formed large clusters in the cytoplasm upon R1881 treatment (Figure 4.1). After co-transfection of the constructs, AR27 and  $\alpha$ -glucosidase were immunocytochemically detected using specific anti-AR and anti- $\alpha$ -glucosidase antibodies. These experiments demonstrated that AR27 did not co-localize with the lysosomal  $\alpha$ -glucosidase.

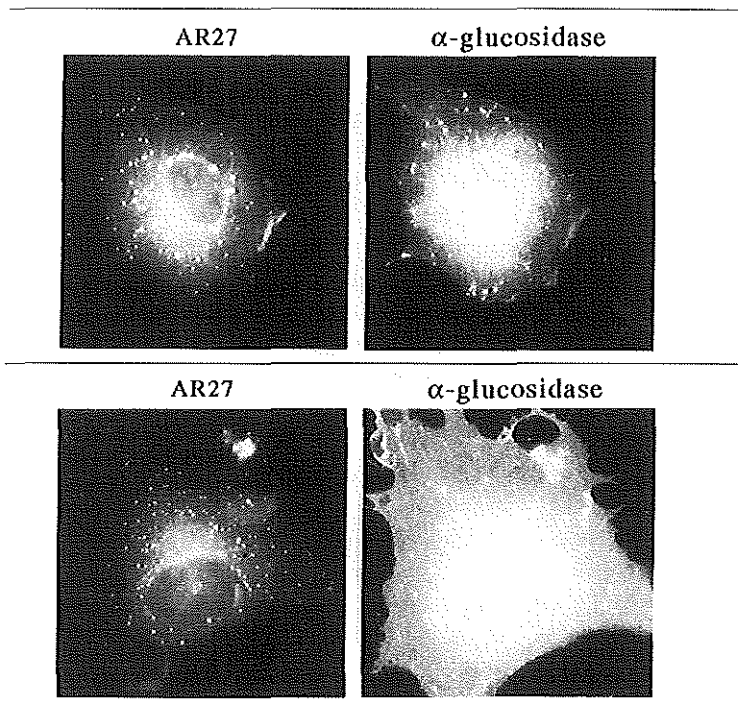


Figure 4.1 Immunocytochemical localization of AR27 and lysosomal  $\alpha$ -glucosidase. AR27 construct and an  $\alpha$ -glucosidase construct were co-transfected into COS-1 cells using the calcium-phosphate precipitation method, and the cells were incubated with 10 nM R1881 (see Materials and Methods in Jenster et al. 1991 [Appendix paper II]). AR27, which lacks its NLS and is mutated in the first zinc-finger, forms large clusters in the cytoplasm upon R1881 treatment (left photographs). AR27 was visualized using anti-AR F39.4.1 antibody and FITC-conjugated goat-anti-mouse antibody.  $\alpha$ -Glucosidase was visualized using a specific anti- $\alpha$ -glucosidase antibody and TRITC-conjugated goat-anti-rabbit antibody. Two different examples of localization of AR27 (left photographs) and  $\alpha$ -glucosidase (right photographs) protein expressed in one cell are shown. The  $\alpha$ -glucosidase expression plasmid and the anti- $\alpha$ -glucosidase antibody were kindly provided by Dr. A.J.J. Reuser (de Jonge et al. 1985; Hoefsloot et al. 1990). Magnification approximately 630 $\times$ .

### 4.4 EFFECTS OF ANTIANDROGENS AND VARIOUS STEROID HORMONES ON ANDROGEN RECEPTOR NUCLEAR IMPORT

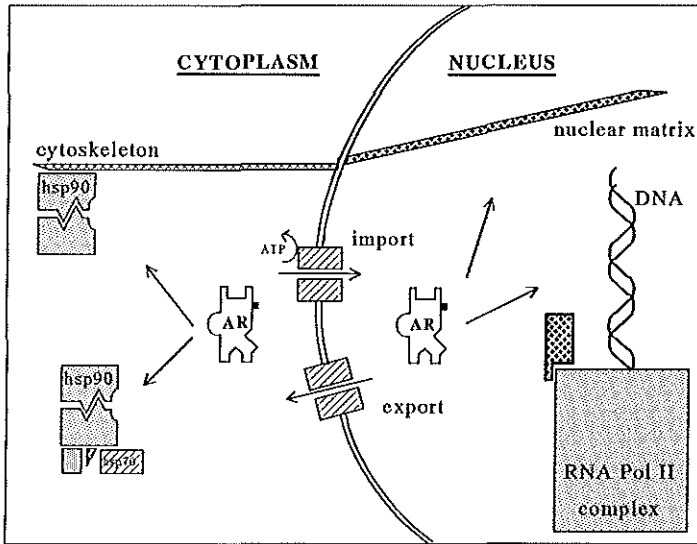
There are a number of different processes in the cascade of steroid receptor activation that may be blocked by antagonists (e.g. hsp dissociation, nuclear import, receptor dimerization, DNA binding and transactivation). Several antiandrogens (e.g. hydroxy-flutamide [OH-flu], cyproterone acetate [CA], Casodex [ICI 176334]) have been found to be able to induce AR nuclear localization (Kemppainen et al. 1992; Jenster et al. 1993 [Appendix paper II]). The concentrations of antiandrogens that resulted in a clear shift in AR immunostaining from cytoplasm to nucleus, were not sufficient for a significant induction of AR transcription activity. So, these findings indicated that the inability of these compounds to activate transcription is not caused by an inhibition of nuclear import (Jenster et al. 1993 [Appendix paper II]).

Besides androgens and antiandrogens, ligands with low affinity for the AR (dexamethasone [Dex], estradiol [E<sub>2</sub>] and promegestone [R5020, a synthetic progestagen]), also were able to induce nuclear retention (Kemppainen et al. 1992; Jenster et al. 1993 [Appendix paper II]). The concentrations of these compounds, inducing nuclear localization, were proportional to the relative binding affinities (RBAs) for the AR as determined by [<sup>3</sup>H]-R1881 AR binding competition assays (Veldscholte et al. 1990). However, the concentrations inducing nuclear retention, were far too low to saturate the AR. Already 0.01 nM R1881 resulted in a predominantly nuclear AR localization, while only 3-4% of the receptors are occupied under these conditions based on a K<sub>d</sub> of 0.3 nM. Possibly, the relative R1881 concentration in the cell is higher because the lipophilic steroid accumulates in the membranes of the cells. Furthermore, it is possible that, after the AR has been directed to the nucleus, the R1881-AR complex dissociates and the AR molecule remains nuclear for a long period, while R1881 is able to associate with a next cytoplasmic AR molecule.

### 4.5 THE COMPARTMENTAL RETENTION MODEL

Based on the dynamic shuttling of AR nuclear import and AR export out of the nucleus, there are three major factors that could influence subcellular localization: 1) accessibility of the NLS for recognition at the NPC, 2) binding to cytoplasmic or nuclear components (Laskey & Dingwall 1993), and 3) potency of the NLS. The compartmental retention model depicts the subcellular localization of the AR based on these three factors.

The balance between nuclear import and export is disturbed when the NLS in the AR is not exposed or masked by accessory proteins such as hsps. It has been observed for the GR and AR, expressed in transfected cell lines, that deletion of the LBD resulted in a ligand-independent shift from a predominantly cytoplasmic wild type receptor to a constitutively nuclear mutant (Picard & Yamamoto 1987; Jenster et al. 1991, 1993 [Appendix papers I, II]; Simental et al. 1991). This suggests that the LBD and/or the associated hsps mask the NLS and inhibit nuclear import. Deletion of the LBD, or hormone-induced conformational changes and subsequent hsp dissociation, expose the NLS, resulting in nuclear import. Since hsp90 is predominantly cytoplasmic it has the tendency to retain the associated AR in this compartment. Upon hormone binding, the AR dissociates from hsp90 and binds tightly to the nuclear matrix and to DNA, resulting in a complete shift of the nuclear import-export balance to a nuclear localization.



**FLOWCHART**

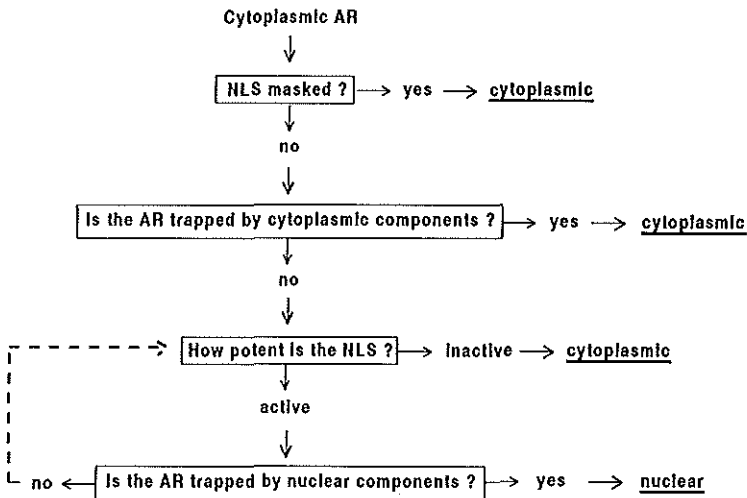


Figure 4.2 The compartmental retention model. The balance between active nuclear import and export out of the nucleus is dependent on NLS accessibility, binding to cytoplasm-specific or nucleus-specific components, and potency of the NLS. Binding of the AR to hsp90 could obstruct nuclear import by masking of the NLS or could retain the AR in the cytoplasm because hsp90 itself is predominantly cytoplasmic. If the NLS is not masked and the receptor not trapped in the cytoplasm, the potency of the NLS and possible binding to nuclear components determine subcellular localization. If the AR is not trapped by nuclear components, the potency of the NLS determines the localization of the AR (stippled arrow). In this reasoning, the export of the AR out of the nucleus is considered to be constant. Hormone treatment completely shifts the balance to a nuclear localization: the receptor is not bound to hsp90 and will be trapped in the nucleus because of binding to the nuclear matrix and DNA. The possibility of protein-protein (piggy-back) facilitated nuclear import is not taken into account.

## Chapter 4

Targeting hsp90 to the nucleus by linking it to the nucleoplasmin NLS, resulted in a complete shift from cytoplasm to nucleus of associated GRs, showing that the subcellular localization of hsp90 is an important determinant for receptor destination in the absence of hormone (Kang et al. 1994). Interestingly, these GRs were devoid of their nucleoplasmin-like NLS, indicating the possibility of a 'piggy-back' nuclear import of the GR mutants via the NLS of the associated hsp90.

Besides accessibility of the NLS and compartmental retention, the potency of the NLS might also determine AR subcellular localization. Some reports suggest that phosphorylation in or near the NLS could be involved in NLS-NPC interaction and consequently might affect nuclear import (Rihs & Peters 1989; DeFranco et al. 1991; Robbins et al. 1991). Hence, phosphorylation, possibly induced by ligand, might directly regulate the potency of an NLS. C-terminal of the AR NLS are several serine residues that could be targets for kinases. The serine at position 641 represents a consensus phosphorylation site for Ser-Pro directed kinase and a consensus site for casein kinase II. Zhou and co-workers (1994a) have shown that this serine residue was phosphorylated in an AR fragment expressed in COS cells. Nuclear import efficiency might also indirectly be regulated by the activity of the proteins that facilitate AR transport such as hsp90 and hsp56 (Chapter 3). Figure 4.2 visualizes the processes involved in AR subcellular localization and describes a possible flowchart to determine receptor localization, based on the compartmental retention model.



**STEROID BINDING**

### 5.1 INTRODUCTION

Hormone binding is the initial step in the series of sequential processes in the nuclear receptor activation cascade. Phosphorylation, hsp dissociation, dimerization, nuclear matrix association, DNA binding and transcription activation are processes that follow upon the binding of ligand to the receptor. Ligand binds with high affinity and specificity to the ligand binding domain (LBD), which is located in the C-terminal part of nuclear receptors. This domain has a length of approximately 250 amino acids, encoded by five exons (Chapter 1; Figure 1.6). The LBD is conserved among the nuclear receptor family members and it is believed that it folds into a hydrophobic pocket in which the ligand specifically binds.

### 5.2 CHARACTERISTICS AND 3-DIMENSIONAL MODELS OF THE LIGAND BINDING DOMAIN

Information on boundaries, LBD-ligand contact sites, and 3-dimensional conformation has been provided by 1) mutational analysis, 2) steroid cross-linking, 3) limited proteolytic digestion, and 4) structural comparison.

#### *Mutational analysis*

Deletion analysis revealed that for full binding capacity, the last C-terminal 200-250 amino acids of the GR, PR and AR are indispensable (Giguere et al. 1986; Rusconi & Yamamoto 1987; Dobson et al. 1989; Jenster et al. 1991 [Appendix paper I]). The ER contains approximately 60 additional amino acids located C-terminal of its LBD, that are not necessary for steroid binding (Kumar et al. 1986; Lees et al. 1989). This region has been termed the F domain (hER residues 553 to 595) and its function is unknown up to now (Pakdel et al. 1993b). Deletion analysis studies showed that almost all intra-LBD deletions abolished ligand binding. The same inhibitory effect is observed for the vast majority of different amino acid substitutions introduced in the LBD by site-directed mutagenesis and naturally occurring in patients with steroid hormone insensitivity syndromes. Most of the substitutions resulted in impaired- or loss of ligand binding (Tora et al. 1989; Wrenn & Katzenellenbogen 1993; Pakdel et al. 1993a; Sultan et al. 1993; Chen & Stallcup 1994). However, some substitutions caused an increased ligand binding affinity (Chakraborti et al. 1991), changed hormone or antihormone specificity (Veldscholte et al. 1990; Jiang et al. 1992; Culig et al. 1993; Ris-Stalpers et al. 1993), or caused changes in other LBD features such as transactivation and receptor dimerization (Fawell et al. 1990; Danielian et al. 1992; Pakdel et al. 1993a).

Many mutations in the AR DBD and LBD have been found in individuals with androgen insensitivity syndrome (AIS) (reviewed in McPhaul et al. 1991; Brinkmann & Trapman 1992; Pinsky et al. 1992; Sultan et al. 1993). The complete form of the syndrome is characterized by 46,XY karyotype, external female phenotype, intra-abdominal testes, absence of uterus and ovaries, blind ending vagina, and gynaecomastia (Chapter 10) (Griffin & Wilson 1989; Schweikert 1993). Amino acid substitutions, identified in AIS patients, that affected ligand binding are scattered throughout the LBD, although the frequency of mutations in some regions seems to be higher (McPhaul et al. 1992) (Figure 5.1A). Especially the region from amino acids 716 to 763, encoded by exon 5, contains a relatively large number of mutations. So far, there have

been no reports of mutations in the AR hinge region (encoded by the first half of exon 4). The most N-terminal and most C-terminal residue substitutions affecting ligand binding, are residues 655 and 895, providing an estimation of the LBD boundaries (Pinsky et al. 1992; McPhaul et al. 1992).

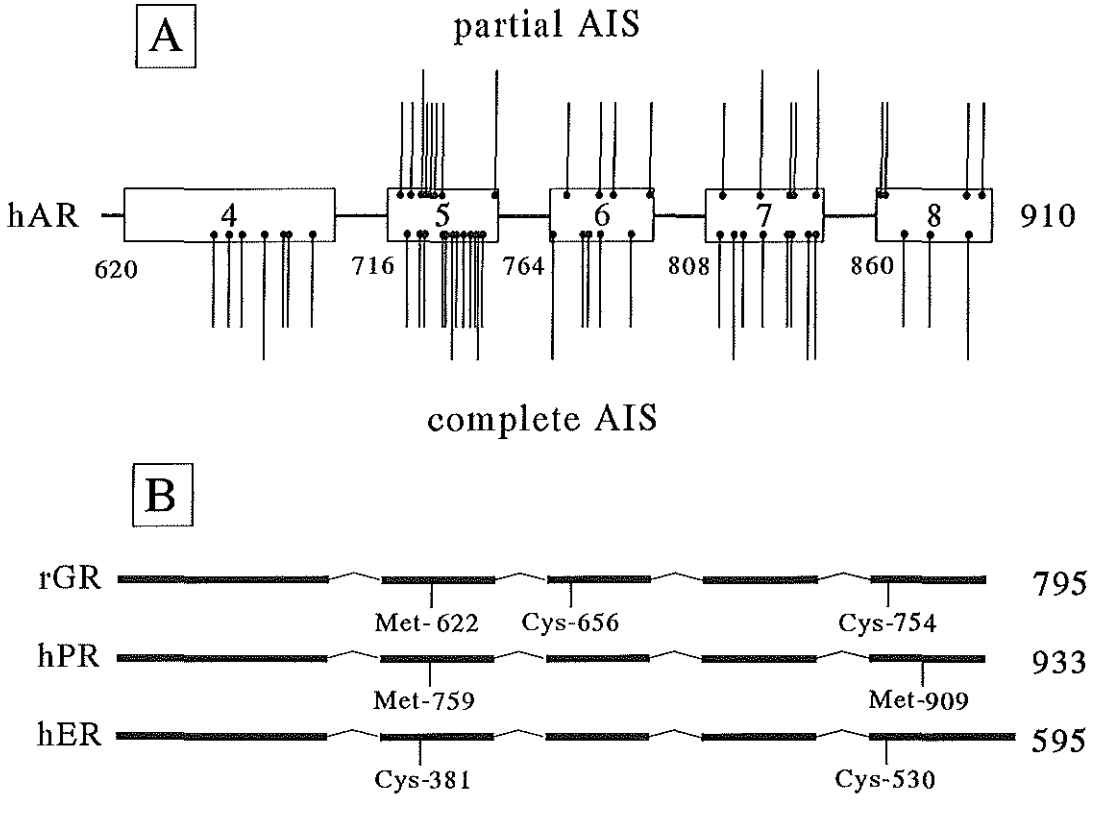


Figure 5.1 A: Location of amino acid substitutions identified in patients with androgen insensitivity syndrome (AIS). The AR LBD (encoded by exons 4 to 7 and part of exon 8) is shown (open boxes). Lines above the LBD boxes indicate the sites of the substitutions identified in patients with partial AIS. These mutations do not abolish ligand binding completely. Substitutions identified in patients with complete AIS are shown as lines below the LBD. In few cases, the same amino acid has been substituted into different amino acids in different patients (indicated as longer lines). The mutations are reviewed in McPhaul et al. 1991; Brinkmann & Trapman 1992; Pinsky et al. 1992; Sultan et al. 1993.

B: Location of amino acids identified in the rGR, hPR and hER to be covalently cross-linked to the receptor's respective ligand. Dexamethasone 21-mesylate labeled Cys-656 in the rat GR (Simons et al. 1987; Carlstedt-Duke et al. 1988) and triamcinolone acetone covalently attached to Met-622 and Cys-754 (Carlstedt-Duke et al. 1988). R5020 labeled Met-759 and Met-909 of the hPRb (Strömstedt et al. 1990). Tamoxifen aziridine alkylates Cys-530 in the hER (Harlow et al. 1989), and when this residue is mutated to alanine, it alkylates Cys-381 (Reese et al. 1992).

### *Ligand cross-linking*

More detailed information on possible LBD-steroid contact sites was provided by affinity labeling studies. Steroids containing an activated side-group (such as dexamethasone-21-mesylate, ketononestrol aziridine, and tamoxifen aziridine, and the photo-activated compounds triamcinolone acetonide, R5020, and R1881) would be expected to interact with amino acids that are in close proximity. The residue cross-linked to the labeled steroid can subsequently be identified. Affinity labeling studies of the GR, PR and ER provided evidence for residues in the LBD that are proximate to, or are themselves, ligand contact sites (Simons et al. 1987; Carlstedt-Duke et al. 1988; Harlow et al. 1989; Ratajczak et al. 1989; Reese et al. 1992; Goldstein et al. 1993) (reviewed in Carlstedt-Duke et al. 1989; Danielsen 1991). Interestingly, amino acids identified in the various steroid receptors, are located in similar regions and even are sometimes at exactly corresponding positions in the various receptors. Met-622 in the rGR, labeled by triamcinolone acetonide, corresponds exactly to Met-759 of the hPR which is labeled by R5020 (Figure 5.1B). These residues and Cys-381 in the hER are located in the LBD-region encoded by exon 5 in which a relatively large number of mutations has been identified in the AR in patients with AIS. Possibly, this part of the LBD is more important for ligand binding than other LBD regions. The C-terminal part of the LBD of the rGR, hPR and hER also contain residues that can be covalently cross-linked by the receptor's respective ligands (Figure 5.1B). Cys-754 in the rGR labeled by triamcinolone acetonide corresponds to the Cys-530 in the hER which can be labeled by tamoxifen aziridine. The corresponding residue in the hAR at the exact same position as Cys-754 in the rGR and Cys-530 in the hER is Thr-868. This Thr-868 is substituted by alanine in the AR expressed in the human prostate cancer cell line LNCaP, which results in an increased affinity of the mutant AR for progestins, estrogens and antiandrogens (Veldscholte et al. 1990; Ris-Stalpers et al. 1993).

### *Limited proteolytic digestion*

Besides deletion analysis, limited proteolytic digestion is another method that provides information on the minimal size of the ligand binding domain. In the absence of hormone, the receptor is an easy target for proteases. However, upon hormone binding, the LBD becomes partially resistant to proteolytic digestion. Resistance to proteases is different when bound to agonists or to antagonists, since different size LBD fragments are protected (Allan et al. 1992a,b; Beekman et al. 1993; Kallio et al. 1994; Kuil & Mulder 1994). For the wild type hAR, protection by R1881 resulted in a 29 kDa fragment, whereas CA, OH-flu and Casodex protected a 35 kDa LBD fragment (Kuil & Mulder 1994). The residues that determine the difference between the 29 and 35 kDa segments, are located in the hinge region. These studies showed that agonists and antagonists induced different conformational changes that, most likely, are the basis for the receptor's capacity to activate transcription (Banialmed & Tsai 1993; Beekman et al. 1993).

### *Structural comparison*

Based on structural homology and hydrophobic cluster analysis, 3-dimensional models of steroid receptor LBDs have been designed (Lemesle-Varloot et al. 1992; Goldstein et al. 1993). Although the two proposed models are very different from each other, both groups claim that the available biochemical information, described above, fits their model. This already indicates the difficulty in predicting 3-dimensional conformation of large protein domains based

on little conclusive biochemical data and on low homology to known 3-dimensional protein structures. Although the proposed models provide attractive starting points for the prediction of LBD characteristics and the sites responsible for these characteristics, the LBD's structural determination by NMR spectroscopy or X-ray diffraction is crucial and of high priority. Preliminary crystallographic studies of the THR LBD indicate that 3-dimensional structural data will soon be available (McGrath et al. 1994). Although the homology between the THR LBD and the LBD of steroid hormone receptors is not very high, structural data of the THR-LBD might form a much better basis for the prediction of the 3-dimensional conformation of the LBD of steroid hormone receptors than the models based on structural homology and hydrophobic cluster analysis.

### 5.3 ANDROGEN RECEPTOR CHARACTERISTICS ENCODED BY SEQUENCES IN THE LIGAND BINDING DOMAIN

Some of the sequences responsible for the processes that are part of the steroid receptor activation cascade, are located in the LBD. These characteristics include: 1) ligand binding, 2) hsp association (Chapter 3), 3) nuclear import (Chapter 4), 4) dimerization (Chapter 6), 5) nuclear matrix association (Chapter 7), and 6) transcription activation (Chapter 9).

According to the exon-protein unit theory, the sequences responsible for the various LBD characteristics, would each be encoded by a single exon (Traut 1988). However, it is clear from studies on ligand binding, hsp association, and transcription activation, that these characteristics are not effectuated by a sharply bordered protein unit encoded by only one of the exons 4 to 8 (Webster et al. 1989).

In general, the ligand binding domain is considered to be an inhibiting domain. Deletion results in a constitutively active transcription factor, and linking the LBD to other proteins inhibits normal activity of these proteins (Picard et al. 1988; Becker et al. 1989; Eilers et al. 1989; Superti-Furga et al. 1991). The inhibition is abolished upon ligand binding, which additionally unlocks the capacity of the LBD to enter the nucleus, dimerize, bind to the nuclear matrix, and to act as a transcription activation unit (Appendix paper III; Picard & Yamamoto 1987; Webster et al. 1988; Fawell et al. 1990). From the observation that the LBD is responsible for several receptor characteristics, it can be understood why it is so large, but it is still unknown how the small steroid hormone molecule is able to abolish inhibition of receptor characteristics by the LBD, and is able to unlock LBD features.

## Chapter 5

**DIMERIZATION**

## Chapter 6

### 6.1 INTRODUCTION

Since DNA sequences to which nuclear receptors bind (HREs) usually consist of two half-sites, it was speculated that receptors bind to DNA as protein dimers. Evidence was provided by band shift assays with labeled HREs, using co-expressed wild type receptor and receptors with a large N-terminal deletion. While the wild type receptor-HRE complex migrated slower than the N-terminal deletion receptor-HRE complex in a band shift assay, co-expression of the two receptor forms resulted in an additional intermediate retarded band, indicating heterodimerization on the HRE (Bagchi et al. 1988; Kumar & Chambon 1988; Tsai et al. 1988; Dahlman-Wright et al. 1990; Meyer et al. 1990; Perlmann et al. 1990; Palvimo et al. 1993; Wong et al. 1993). Dimerization seems to be cooperative, since receptor binding is undetectable or much weaker in gel shift assays using only half-site HREs. Binding of the second receptor to the HRE might increase the stability of the complex, due to an alteration of the DNA structure or due to protein-protein interactions between the two receptor molecules. This last possibility suggests the existence of dimerization domains in the nuclear receptors (reviewed in Forman & Samuels 1990; Gronemeyer 1991; Wright et al. 1993).

### 6.2 DIMERIZATION DOMAINS

In search for dimerization signals in the ER, deletion mapping revealed that the LBD was important for ER dimer formation (Kumar & Chambon 1988). The presence of a dimerization sequence within the ER LBD was confirmed by the identification of amino acids that were critically involved in high affinity DNA binding and dimerization, as demonstrated by band shift assays and co-immunoprecipitation experiments (Fawell et al. 1990). Moreover, a peptide with 22 amino acids, containing the sequences important for dimer formation, restored DNA binding activity to mutant ERs lacking their LBD (Lees et al. 1990). This region is highly conserved among the nuclear receptor family members, and contains a heptad repeat of hydrophobic amino acids that might form a dimerization interface. The corresponding region in the THR and RAR has been implicated in the formation of heterodimers (Forman et al. 1989; Glass et al. 1989), but also other sequences in the LBD were held responsible for THR dimerization (Spanjaard et al. 1991; Lee et al. 1992). Although there are some indications for the involvement of the LBD in PR and GR dimerization (Guiochon-Mantel et al. 1989; Dahlman-Wright et al. 1992b), the importance of the LBD sequence in PR, GR and AR dimerization still needs to be established.

A second dimerization sequence has been identified in the second zinc-finger of the steroid receptor's DBD. Mutation of five amino acids in the N-terminal part of the second zinc-finger resulted in loss of cooperative DNA binding of two GR DBDs to the HRE (Dahlman-Wright et al. 1991). This region has been termed the 'D-box' (Chapter 8, Figure 8.1) and the evidence that it is the main part of the interaction site between two GR molecules has been provided by analyzing the structure of the DBD-HRE complex by X-ray diffraction (Luisi et al. 1991). The 3-dimensional structure determination of the ER DBD, indicated that dimerization via the conserved D-box might be applicable to all steroid receptors (Schwabe et al. 1993).

Besides homodimerization, several nuclear receptors can form heterodimers on the HRE. Retinoic acid/thyroid hormone/vitamin D<sub>3</sub> receptors and orphan receptors, have been found to heterodimerize (reviewed in De Luca et al. 1991; Glass et al. 1991; Green 1993a; Giguere



1994). The protein-protein contact sites of the RXR, RAR, VDR and THR on directly repeated elements, concern various parts of the DBD. Besides the D-box, also amino acids in the first and second zinc-fingers, and a domain C-terminal of the second zinc-finger (termed T-box) can be involved (Zechel et al. 1994a,b). The recently cloned and characterized RXR plays a central role in heterodimerization, since this receptor is able to dimerize with many different nuclear receptors. Heterodimerization amplifies the number of possibilities to regulate positively or negatively the expression of specific target genes. So far, there are no reports of heterodimerization between the PR, GR, MR and AR.

Co-immunoprecipitation experiments using the GR, PR or ER have shown that these receptors can form homodimers even in the absence of DNA (Wrange et al. 1989; Fawell et al. 1990; Rodriguez et al. 1990; Cairns et al. 1991; DeMarzo et al. 1991). These studies demonstrate that receptor dimerization occurs in solution, prior to HRE-binding, and suggest that dimerization can be a regulatory step to control DNA binding as has been described for other transcription regulatory proteins (Yamamoto et al. 1988; Landschulz et al. 1989) (reviewed in Forman & Samuels 1990; Jones 1990).

For the AR, only a few studies have been carried out to investigate receptor dimer formation. These studies provided evidence for AR dimerization on the HRE and suggested a strong dimerization sequence in the DBD (Palvimo et al. 1993; Wong et al. 1993).

## Chapter 6

**ASSOCIATION WITH THE NUCLEAR MATRIX**

### 7.1 INTRODUCTION

The nucleus is a highly complex cell organelle, which contains approximately  $6 \times 10^9$  base pairs of DNA (diploid mammalian cell), RNA, and many proteins. In order to accommodate the large amount of chromatin in such a way that it remains accessible for replication and transcription, the DNA is condensed by different orders of packaging, and organized in less compact transcriptionally exposed chromatin and in highly condensed heterochromatin which is inactive in RNA synthesis (reviewed in Getzenberg et al. 1991). Besides these structural features, the correct and efficient execution of nuclear processes such as RNA synthesis, RNA processing, RNA transport, DNA replication, and DNA repair, that all take place at discrete and specific sites, calls for a nuclear skeleton as the basic scaffold structure around which DNA packaging and these nuclear processes can be organized.

Such a structure has been biochemically defined by Berezney & Coffey (1974), who obtained a residual nuclear framework from nuclei after treatment with detergent, digestion with nucleases and extraction with salt. This residual nuclear framework has been designated as the nuclear matrix (NM). The nuclear matrix consists mainly of three substructures: 1) the lamina, a fibrous network which encloses the nucleus completely and lines the inside of the nuclear membrane, 2) an internal matrix that is composed of fibres and granular structures, and 3) the residual nucleoli (reviewed in Verheijen et al. 1988; Getzenberg et al. 1990; He et al. 1990b; Berezney 1991).

Various different methods to isolate nuclear matrices have been published (reviewed in Berezney 1991; van Driel et al. 1991). All these methods have a pre-step in common that makes nuclei accessible for nucleases, after which DNA is digested and almost all of it removed. During the procedures, soluble and loosely bound nuclear components are extracted. One of the main differences in the various NM preparation protocols is the way to stabilize the matrix structure. These different stabilizing procedures (e.g. oxidizing agent sodium tetrathionate [NaTT],  $\text{Cu}^{2+}$  ions, heat treatment, embedding in agarose beads) partially determine the final NM composition, and hence form a source of controversy on the issue to what extent the various biochemically defined structures represent the *in vivo* proposed nucleoskeleton (Stuurman et al. 1992a; Appendix paper III, and references therein).

Although there is some debate about the NM composition, the observation that a variety of essential nuclear activities are matrix-associated, strongly suggests that the residual NM structures embody a scaffold network that can function as a model system to investigate nucleoskeleton-associated processes. It has been shown that specific DNA fragments bind to the NM and that these matrix associated regions (MARs) presumably anchor the chromatin loops to the matrix, thereby organizing the genome (reviewed in Zlatanova & van Holde 1992; Cook 1989, 1991). Moreover, processes such as DNA replication, DNA repair, RNA synthesis, RNA processing and RNA transport have been shown to occur associated with the NM (reviewed in Getzenberg et al. 1990; Berezney 1991). Interestingly, most of these nuclear processes are localized at discrete and specific sites. According to the finding that transcription is NM-coupled, RNA polymerase II and various transcription factors have been found to be associated with the NM (reviewed in Berezney 1991). Also various steroid receptors are attached to the NM in a ligand-dependent fashion (Appendix paper III; Kaufmann et al. 1986; Kirsch & Miller-Diener 1986; Alexander et al. 1987; Barrack 1987; Mowszowicz et al. 1988; Metzger & Korach 1990; Schuchard et al. 1991; van Steensel et al. 1991).

## 7.2 STEROID RECEPTOR BINDING TO THE NUCLEAR MATRIX

*In vitro* reconstitution experiments have shown that binding of the AR, ER or PR to the NM occurs with high affinity at a limited number of sites, and is more or less target cell specific (Barrack & Coffey 1983; Wilson & Colvard et al. 1984; Barrack 1987; Belisle et al. 1989; Metzger & Korach 1990; Schuchard et al. 1991). This indicates that the NM contains specific binding sites for steroid receptors, although little is known about the nature of these sites. Binding to the NM might be direct to its backbone or mediated by acceptor proteins as was demonstrated for the PR (Schuchard et al. 1991). Although more than 90% of the chromatin has been removed from the NM preparations, residual DNA could also still be involved in steroid receptor binding.

The AR domain responsible for binding to the NM has been located in studies with AR mutants transiently transfected into COS-1 cells (Appendix paper III). The wild type AR could be immunocytochemically visualized in permeabilized cells (referred to as nuclei) and in NM preparations only in the presence of ligand. AR mutants which were unable to bind DNA due to mutations in the DBD were also associated with the NM. Interestingly, as was observed in the AR subcellular localization studies, clusters of these AR mutants were present in nuclei and NM preparations (Chapter 4.3). The domain responsible for the tight NM association is the AR LBD. An AR mutant lacking the LBD could not be detected in permeabilized cells and only in NMs when the NaTT stabilizer was present during the first permeabilization step. Two important conclusions can be drawn from these experiments. Firstly, the process of AR tight nuclear binding is independent of DNA binding but is mediated by the LBD. Secondly, although the AR mutant which lacks the LBD is constitutively active in transcription activation assays, the receptor is not tightly bound to nuclei or NMs, which demonstrates that tight nuclear binding is not a prerequisite for transactivation. When NaTT (a weak sulfhydryl cross-linker) was added to the buffer used for the permeabilization of the transfected cells as the first step in the NM preparation, all AR mutants tested were NM-associated (Appendix paper III). This indicates that NaTT most likely cross-linked the AR mutants to the NM even when they were not tightly bound. Interestingly, this "cross-linking" seems to be selective since it does not occur to all nuclear proteins (reviewed in Stuurman et al. 1992c; Appendix paper III, and references therein), and is only observed for steroid receptors in the presence of ligand. A  $\beta$ -galactosidase fusion protein, directed to the nucleus by the SV40 large T antigen NLS, was not associated with the NM after NaTT treatment (Appendix paper III). It is thought that only proteins that are at least weakly bound to the NM will be retained in the NM preparation due to NaTT stabilization. This suggests two types of interactions with the NM: 1) a tight binding (mediated by the AR LBD), and 2) a weak interaction stabilized by NaTT involving the N-terminal domain and/or the DBD, which is only evident in AR mutants lacking the LBD.

In contrast to the AR, the wild type GR is not retained in permeabilized cells in the presence of ligand, and apparently not tightly nuclear bound. However, when NaTT was added during permeabilization, part of the GR population resisted extraction (Appendix paper III). This NaTT-stabilized interaction was abolished by deletion or mutation of the first zinc-finger of the GR DBD, indicating that the weak GR-NM association is dependent on an intact DBD. In conclusion, the liganded AR is tightly associated with the NM via its LBD, an interaction that is not observed for the GR. The liganded GR associates weakly with the NM via its DBD,

which might be similar to the weak AR-NM association, observed for an AR mutant lacking its LBD.

### 7.3 FUNCTIONAL ASPECTS OF NUCLEAR MATRIX-STEROID RECEPTOR ASSOCIATION

Although the NM has not yet been shown to be a relevant structure in intact eukaryotic cells, strong evidence has been provided for its importance in major nuclear activities. It has been suggested that DNA anchorage via MARs to the NM, resulting in a loop-domain organization, represents a level of transcription regulation (Cook 1989; Jackson et al. 1992; Zlatanova & van Holde 1992). By condensation and decondensation, chromatin loops can be made accessible or inaccessible to transcription factors, and clusters of genes may be turned on or off. This higher order of transcription control might be regulated by NM composition and could determine cell-specific and development-specific expression of genes. Interestingly, an example for such a situation has been proposed for the androgen-regulated genes encoding C<sub>3</sub> and SVS-4, which are specifically expressed in the rat ventral prostate and seminal vesicle, respectively (Getzenberg et al. 1990, and references therein). Within the same animal, both tissues contain the same genome and express ARs, but respond differentially to DHT with respect to the production of the secretory proteins C<sub>3</sub> and SVS-4. One explanation might be that the DNA is organized differently in the two tissues due to tissue-specific NM composition (Getzenberg et al. 1991). Differences in NM protein content between rat ventral prostate and seminal vesicle were confirmed by NM isolation followed by 2-dimensional gel electrophoresis (Getzenberg & Coffey 1990). These studies are consistent with the idea that transcription control is possibly determined by the composition of the NM.

Analogous to the cytoskeleton, the NM could facilitate intra-nuclear transport and direct nuclear receptors and other transcription factors to specific nuclear sites. Positioning and anchorage of transcription units (pre-initiation complex, RNA polymerase and accessory transcription factors) to certain sites might ensure specific and efficient gene expression (Cook 1989). Moreover, binding and migration along a matrix greatly reduces the freedom of movement of proteins, accelerating protein-protein encounters and assisting in the assembly of multimeric transcription complexes.

**DNA BINDING**

## Chapter 8

### 8.1 INTRODUCTION

One of the processes that are essential for nuclear receptors and other transcription regulators to be able to modulate gene expression, is specific DNA binding. Several different structural domains are responsible for DNA binding of various proteins: helix-turn-helix domains, homeo-domains, zinc-fingers, leucine zipper-helix-loop-helix domains, and  $\beta$ -sheet motifs (reviewed in Harrison 1991; Pabo & Sauer 1992). The nuclear receptors belong to the group that use zinc ions to compose a structural module. DNA binding zinc proteins can be categorized into three classes: 1) zinc-finger proteins, containing a variable number of a 30 amino acid module with one zinc ion liganded by two cysteines and two histidines (e.g. TFIIIA, Sp1), 2) zinc-twist proteins, containing a domain of approximately 70 residues with two zinc ions each liganded by four cysteines (nuclear receptors), and 3) zinc-cluster proteins in which two zinc ions are complexed by six cysteine residues (e.g. GAL4) (reviewed in Harrison 1991; Valee et al. 1991). Although the zinc-twist and zinc-finger modules are structurally very different, the original and generally accepted term 'zinc-finger' is also being used for the nomenclature of nuclear receptor DBDs.

### 8.2 ZINC-FINGER STRUCTURE

The DBD is the best conserved domain among the nuclear receptor family members, and contains many invariant amino acids, nine of which are cysteines (Laudet et al. 1992). Mutational analysis of the GR and ER DBDs provided information on the eight cysteine residues involved in the zinc ion association, and information on DBD-dimerization (Severne et al. 1988; Dahlman-Wright et al. 1991). Furthermore, mutational analysis revealed the amino acids that determine response element specificity. These studies were based on the finding that the GR and ER DBDs bind to related, but receptor-specific, HREs: an estrogen response element (ERE) and a glucocorticoid response element (GRE) will only be bound by the respective receptor (Green & Chambon 1987; Freedman et al. 1988; Dahlman et al. 1989). Swapping of the individual zinc-fingers between the ER and GR demonstrated that the first (N-terminal) zinc-finger was primarily responsible for HRE specificity (Green et al. 1988). Subsequently, three amino acids were identified in this zinc-finger that were essential for specific HRE interaction (Danielsen et al. 1989; Mader et al. 1989; Umesono & Evans 1989; Zilliacus et al. 1991). Based on these three discriminatory residues, the nuclear receptors can be divided into at least four groups (Forman & Samuels 1990; Laudet et al. 1992) (Table 8.1).



**Table 8.1 Classification of nuclear receptors for their DNA binding specificity**

Group	P-box	Receptors
I	<b>GSckV</b>	AR, GR, PR, MR
II	<b>EGckG</b>	THR, RAR, VDR, RXR, PPAR, Nur77, EcR, E75, EAR-1, USP
III	<b>EGckS</b>	COUP-TF, EAR-2, EGON, KNI, KNRL, SVP
IV	<b>EGckA</b>	ER, ERR1, ERR2

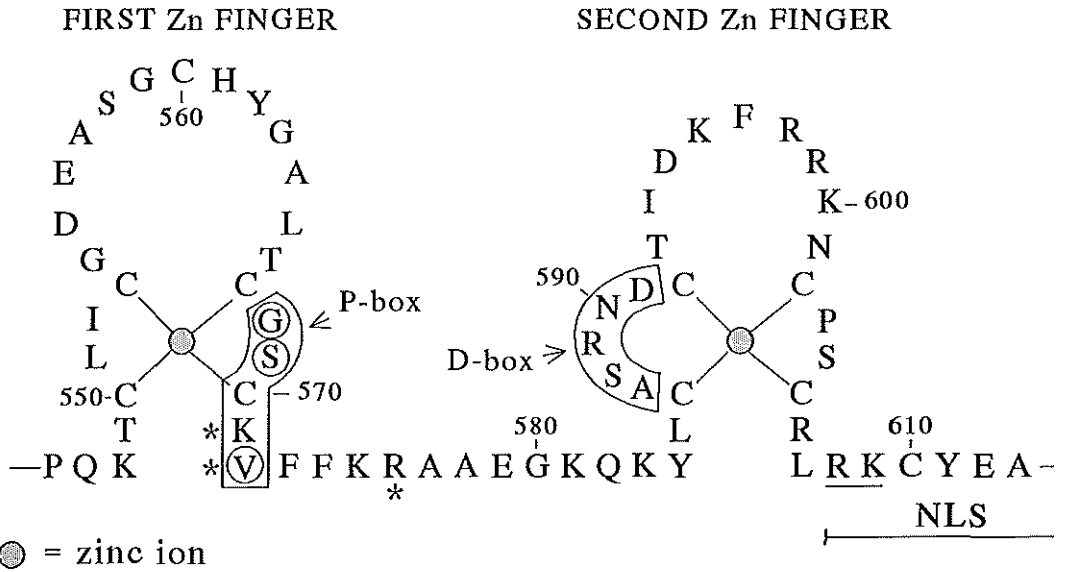
*Classification of nuclear receptors according to the sequence of their P-box which contains the three amino acids (bold capital) that determine HRE specificity. The location of the P-box in the DBD is indicated in Figure 8.1. The classification is based on that of Forman & Samuels (1990). Amino acids are indicated as one-letter symbols.*

In Figure 8.1A the sequence and different functional motifs of the AR DBD are shown. The three discriminatory residues (G, S, V) are part of the so-called P-box and located in the C-terminal part of the first zinc-finger. The D-box is located in the N-terminal part of the second zinc-finger and forms part of a dimerization interface where two DBDs make contact when bound to the HRE (Chapter 6). The elucidation of the 3-dimensional conformation of the GR, ER, RAR, and RXR DBDs has revealed the detailed protein-protein and protein-DNA contact sites (Härd et al. 1990; Schwabe et al. 1990, 1993; Luisi et al. 1991; Knegtel et al. 1993; Lee et al. 1993). The GR and ER DBDs consist of three  $\alpha$ -helices of which the two largest are oriented perpendicular to each other (Figure 8.1B). The recognition  $\alpha$ -helix containing the P-box is placed in the major groove, allowing amino acid side-chains to make sequence specific contacts to the base pairs of one HRE half-site. The second zinc-finger contains two  $\alpha$ -helices and the D-box as the dimerization interface (Figure 8.1B).

### 8.3 HORMONE RESPONSE ELEMENTS

In general, HREs consist of two half-sites, spaced by a variable amount of base pairs (bp). The half-sites partially determine receptor specificity by their specific sequence and orientation into a direct- or inverted repeat. The length of the spacer between the two half-sites is also a determinant for the kind of receptor that binds to the HRE (reviewed in De Luca 1991; Glass et al. 1991; Freedman 1992; Green 1993a; Truss & Beato 1993; Giguere 1994). The consensus ERE and GRE differ in the sequence of their half-sites (Table 8.2). Almost all HREs described so far contain half-site sequences related to the ERE. This ERE versus GRE half-site bipartition, coincides with the P-box sequence. Members of the P-box group I, bind to a GRE and are characterized by a 'GSckV' P-box motif (Table 8.1). This motif is very different from the 'EGck(G/S/A)' P-box sequences of the other three groups of which members bind to HREs containing ERE-like half-sites (Tables 8.1, 8.2) (Martinez & Wahli 1991).

# A Sequence and functional motifs of the AR DBD



# B Schematic model of the AR DBD dimer

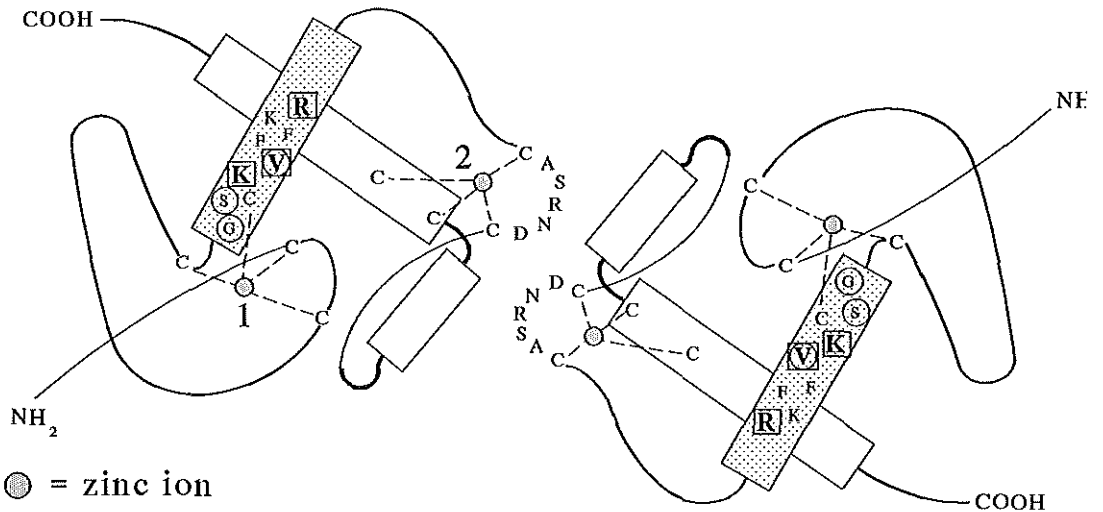


Figure 8.1 Opposite page. *Sequence and model of the AR DBD.*

A: The amino acid sequence of the AR DBD is shown in one-letter code. The domain consists of two zinc-finger modules. The first (N-terminal) module contains the P-box of which three residues (circled) determine HRE specificity. The second zinc-finger contains the first part (underlined) of the nuclear localization signal (NLS) and the D-box in which the amino acids are involved in protein-protein interaction with the second receptor molecule. Asterisks indicate the residues that most likely make base pair contacts in the HRE half-site.

B: Schematic model of two AR DBDs forming a dimer on the DNA.

The  $\alpha$ -helix (stippled box) containing the P-box and the residues (boxed) that make base pair contacts, is positioned in the major groove of a HRE half-site. The second zinc-finger (2) contains two  $\alpha$ -helices (boxes) and the D-box (residues shown) responsible for protein-protein contacts. This AR model images the position of the secondary protein structures in the AR DBD dimer, based on the GR-HRE and ER-HRE crystallographic analyses data (Luisi et al. 1991; Schwabe et al. 1993), and the schematic representation from Schwabe and co-workers for the ER (1990). The HRE should be positioned on top of the schematic dimer.

Table 8.2 Consensus HREs

HRE	Consensus		Receptor Homodimer	P-box group	Receptor Heterodimer
GRE:	<u>AGAA</u> CA >	nnn <u>TGTT</u> CT <	AR, GR, PR, MR	I I	
ERE:	<u>AGGT</u> CA >	nnn <u>TGAC</u> CT <	ER	IV	
TREpal:	<u>AGGT</u> CA >	<u>TGAC</u> CT <	THR, RAR, RXR, COUP-TF	II II III	RXR-THR RXR-RAR RAR-COUP-TF THR-COUP-TF
EcRE:	<u>AGGT</u> CA >	n <u>TGAC</u> CT <	EcR	II	
DR1:	<u>AGGT</u> CA >	n <u>AGGT</u> CA >	RXR, COUP-TF	II III	RXR-PPAR RXR-RAR
DR2:	<u>AGGT</u> CA >	nn <u>AGGT</u> CA >	COUP-TF, RAR	III	RXR-RAR
DR3:	<u>AGGT</u> CA >	nnn <u>AGGT</u> CA >	RAR, THR		RXR-VDR
DR4:	<u>AGGT</u> CA >	nnnn <u>AGGT</u> CA >	COUP-TF, RAR, THR	III	RXR-THR
DR5:	<u>AGGT</u> CA >	Annnnn <u>AGGT</u> CA >	RXR, COUP-TF, RAR, THR	II III	RXR-RAR

Consensus HREs for the various nuclear receptors. The sequence of the two half-sites of the HREs are shown, and the length of the spacer is indicated as the number of n's. The nucleotides that are underlined in the GRE and ERE, indicate the difference between these elements that determine receptor specificity. The orientation of the half-sites is indicated by arrows. In inverted repeats (IRs), the arrows face each other, whereas in direct repeats (DR), the arrows point in the same direction. On several HREs more than one combination of receptor homo- and heterodimers can be formed. In case of a receptor homodimer, the P-box group of the receptor is indicated.

## Chapter 8

The GRE is not GR specific and can also be bound by AR, PR, and MR homodimers. In search for a consensus androgen response element (ARE), Roche and co-workers (1992) screened a pool of random sequence oligonucleotides for AR DNA binding. The AR-complexed HREs were isolated, characterized and compiled to give a consensus ARE. This sequence (GG[A/T]ACAnnnTGTTCT) is very similar to the consensus GRE. No response elements were detected that did not fit the GRE consensus sequence, which emphasizes the observation that GR, AR, MR and PR may recognize the same HREs. This immediately raises the dilemma of target gene specificity in cells containing two or more of these receptors. If all receptors, present in the cell, are activated by their respective ligand, how do they recognize and specifically regulate expression of their own set of target genes? This question will be addressed in Chapter 9.4.

So far, the GR, AR, MR and PR have been observed to bind only HREs that resemble the consensus GRE sequence. This seems also true for the ER which only recognizes HREs that comply to the consensus ERE sequence. In contrast, several other nuclear receptors appear to recognize more than one type of HRE. The RXR, for example, is able to bind strongly to a DR1 and the TRE $\beta$  sequence, and binds weakly to a DR5 consensus sequence (Table 8.2) (reviewed in Giguere 1994). On top of the possibility that receptors can bind to more than one type of HRE, heterodimerization further complicates the view that receptors are committed to specific HREs. The first observations of heterodimerization with RXR as one of the partners, led to the proposal of the so-called 3-4-5 rule. The spacing of the HRE direct repeats determined the binding preference: 3 bp spacing for RXR-VDR, 4 bp for RXR-THR, and 5 bp for RXR-RAR (Table 8.2) (reviewed in Green 1993a). However, the large amount of data available today on heterodimerization between many different nuclear receptors on various HREs, show that the real picture is far more complicated (Zechel et al. 1994a,b) (reviewed in De Luca 1991; Glass et al. 1991; Green 1993a; Giguere 1994). This compilation of novel possibilities (heterodimerization, HRE half-site spacing, and half-site orientation), amplifies the number of ways to specifically regulate gene expression. So far, heterodimerization has been observed for many different nuclear receptors except for the AR, GR, MR, PR, and ER. This suggests that these latter receptors have other strategies to attain cell- and receptor-specific gene regulation.

### 8.4 MODULATORS OF ANDROGEN RECEPTOR-DNA BINDING

Different proteins are able to modulate the binding of the AR to DNA. A factor designated receptor accessory factor (RAF), potentiates DNA binding of AR and GR mutants containing the DBD and N-terminal and C-terminal flanking sequences (rat AR residues 460 to 704) (Kupfer et al. 1993). RAF has a predicted molecular mass of 130 kDa, and is present in a variety of lysates from different cell types which can enhance AR-DNA binding as determined by band-shift analysis. The N-terminal region between amino acids 460 and 521 in the rat AR (corresponding to hAR residues 467 to 527) appeared to be essential for RAF to form a heteromeric AR-RAF complex. These observations have led to the postulation that RAF could be a bridging factor (co-activator), influencing the ability of steroid receptors to activate transcription (Kupfer et al. 1993). Interestingly, the region of the AR to which RAF binds, coincides with a transcription activation unit (TAU-5), responsible for transactivity of a constitutively active AR mutant lacking its LBD (Chapter 9.5.1; Appendix paper IV).

A new class of modulators has recently been shown to inhibit AR and GR DNA binding (Burns et al. 1994; Dedhar et al. 1994). Calreticulin is a multifunctional protein that is able to bind and store  $\text{Ca}^{2+}$  in the lumen of the endoplasmic reticulum, specifically recognizes the KxGFFPR sequence of the cytoplasmic integrins, and is also found in the nucleus (Rojiani et al. 1991; Michalak et al. 1992). The presence of the highly conserved motif (KxVFKR) in the DBD of nuclear receptors, which is very similar to the integrin sequence, implied a possible binding of calreticulin to the DBD. Interestingly, the motif contains the amino acids (K,V,R) that make direct contacts with the base pairs of the HRE (Figure 8.1). Dedhar and co-workers (1994) have shown that calreticulin inhibited binding of the AR to its HRE, and subsequently repressed AR-mediated transcription activation. The same inhibiting properties were found using the GR (Burns et al. 1994). These studies indicate that calreticulin can modulate steroid receptor transactivation. However, its *in vivo* significance in AR functioning still needs to be established.

## Chapter 8

**TRANSCRIPTION REGULATION**

## Chapter 9

### 9.1 INTRODUCTION

After the AR successfully has undergone the sequential steps described in the previous chapters, it is able to perform its major task: transcription regulation. Binding to its HRE, located in promoter and/or enhancer regions of target genes, triggers or inhibits the assembly of the basal transcription machinery, resulting in the up- or down regulation of RNA synthesis. The details of the cooperation between nuclear receptors and the basal transcription machinery are largely unknown. The most recent models that illustrate this cooperativity will be discussed. Knowledge on receptor-mediated transcription regulation is essential to explain the issue of how each nuclear receptor is able to specifically regulate its own set of target genes.

### 9.2 TRANSCRIPTION INITIATION

In eukaryotes, synthesis of mRNA is carried out by RNA polymerase II. In order to accurately initiate transcription and efficiently carry out the elongation reaction, RNA polymerase II requires auxiliary proteins referred to as general transcription factors (GTFs: TFIIA, -B, -D, -E, -F, -H, -J) (reviewed in Zawel & Reinberg 1992, 1993). Together they constitute the basal transcription machinery, which is assembled in a multi-step process on the promoter region of genes. The assembly and/or stability of the basal transcription machinery is usually controlled by a second class of transcription factors. These transcription regulators bind to DNA consensus elements, located in the promoter region and/or at distant enhancers. Promoters of genes transcribed by RNA polymerase II usually contain two distinct classes of DNA elements: 1) the TATA-motif (optional) and/or the initiator site (Inr) (optional) representing the starting points for the assembly of the basal transcription machinery and the transcription start site respectively (also referred to as minimal promoter elements), and 2) consensus elements which are bound by specific DNA binding regulatory proteins, and of which the presence and arrangement are gene-specific. The combination of the number, type and position of cis-elements determines the strength and specificity of a promoter (Figure 9.1) (reviewed in Mitchell & Tjian 1989; Ptashne 1989; Gill & Tjian 1992).

#### 9.2.1 ASSEMBLY OF THE BASAL TRANSCRIPTION MACHINERY

Up to now, seven different GTFs (TFIIA, -B, -D, -E, -F, -H, -J) have been identified as auxiliary proteins that catalyze a basal level of RNA synthesis by RNA polymerase II (reviewed in Zawel & Reinberg 1993; Buratowski 1994). The assembly of the basal transcription machinery starts with the specific binding of TFIID to the TATA-box. TFIIA is able to associate with this complex and stabilizes the TFIID-DNA interaction. This complex acts as a binding site for TFIIB, which can recruit RNA polymerase II and TFIIF into the complex. Subsequently, TFII E and TFII H join the complex, and finally TFII J completes the assembly of the basal transcription machinery (Figure 9.1). These GTFs represent proteins purified from cellular extracts. Recent cloning and characterization of four of the GTFs (TFIIB, TFIID, TFII E, TFIIF) have enabled more detailed studies regarding the assembly, protein-protein interactions, and individual function of the factors that are part of the basal transcription



machinery (reviewed in Zawel & Reinberg 1993).

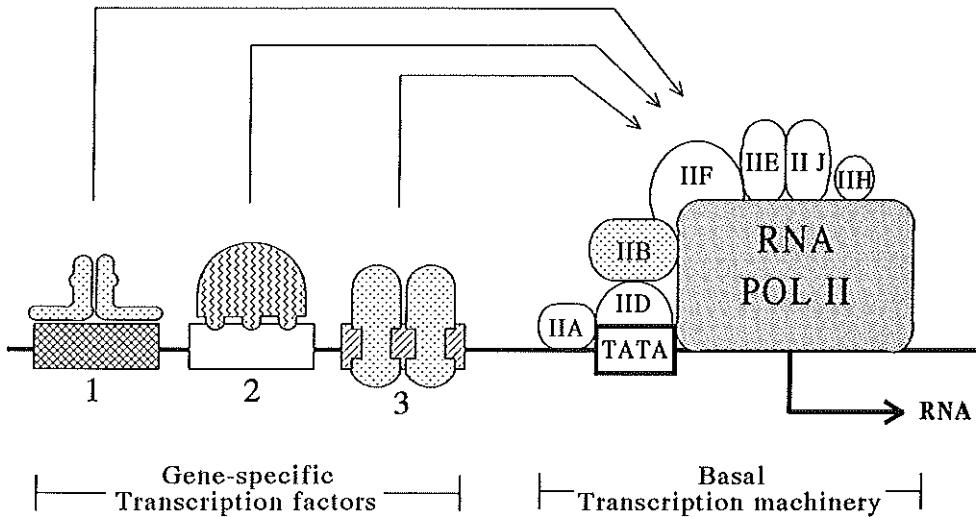


Figure 9.1. The basal transcription machinery consists of RNA polymerase II and the general transcription factors (GTFs: TFIIA, -IIB, -IID, -IIE, -IIF, -IIH, -II J) and is positioned by the binding of TFIID to the TATA-box. The assembly can be controlled by transcription regulators which bind to consensus DNA elements located in the promoter or enhancer regions (the complexes of transcription regulator [in cases 1 and 3 a dimer] with the DNA element are indicated as 1, 2, 3). The presence and arrangement of these elements is gene-specific. When a stable transcription initiation complex has been formed, RNA polymerase II starts RNA synthesis.

Especially the cloning and characterization of the TATA binding protein (TBP) provided new insight into the communication between the basal transcription machinery and the gene-specific transcription regulators (Greenblatt 1991, and references therein). TBP can replace the TFIID fraction when basal transcription is concerned. However, recombinant TBP cannot substitute for TFIID to support activated transcription, mediated by transcription regulators such as Sp1, CTF and GAL4-VP16 (reviewed in Gill & Tjian 1992). These findings, combined with the observation that the 38 kDa TBP is much smaller than the molecular mass of TFIID (more than 100 kDa), resulted in the suggestion that TFIID is a multi-protein complex, that, besides TBP, contains co-activators which link transcription regulators to the basal machinery (Lewin 1990; Peterson et al. 1990). Characterization of several polypeptides, tightly associated with TBP, provided strong evidence for the hypothesis that these TBP associated factors (TAFs) are specific bridging (adaptor) proteins that facilitate the 'cross-talk' between GTFs and transcription regulators (Kelleher et al. 1990; Dynlacht et al. 1991; Pugh & Tjian 1991; Tanese et al. 1991; reviewed in Gill & Tjian 1992).

## 9.2.2 GENE-SPECIFIC TRANSCRIPTION REGULATORS

Transcription regulators are characterized by specific DNA binding to consensus elements and their capacity to activate or inhibit transcription. They usually contain at least two independent functional domains: a DNA binding domain, and a domain responsible for the enhancement of transcription, referred to as a transcription activation unit (TAU, in the literature also referred to as TAF, AF, TAD). Nuclear receptors are a special group of transcription regulators, since they contain a ligand binding domain which makes most of these transcription factors dependent on the binding of small non-protein compounds for activity. Various structural designs of DBDs that can specifically interact with DNA have been characterized in detail for the many different families of transcription regulators (Chapter 4) (Harrison 1991; Pabo & Sauer 1992). Less well defined are the transcription activation units, which have been loosely classified in: 1) acidic activation domains (AADs), 2) glutamine-rich domains, and 3) proline-rich domains (reviewed in Ptashne 1988; Mitchell & Tjian 1989; Latchman 1990; Tjian & Maniatis 1994).

### *Acidic activation domains*

Acidic regions were originally identified in the yeast GAL4 and GCN4 factors. The TAUs of these proteins consist of a relatively short stretch of negatively charged residues which can form an amphipatic  $\alpha$ -helical structure (Giniger & Ptashne 1987). A very potent AAD is present in the herpes simplex VP16 protein. This protein by itself lacks DNA binding activity, but displays, when linked to the GAL4 DNA binding domain, a strong capacity to activate transcription (Sadowski et al. 1988). None of the above described AADs reveal any sequence homology, but can be interchanged and activate transcription when linked to a heterologous DNA binding domain (reviewed in Ptashne 1988; Mitchell & Tjian 1989).

### *Glutamine-rich domains*

A second type of TAU has been identified in the Sp1 transcription factor, in which two domains that are important for transactivation contain a relatively large number of glutamine residues (Courey & Tjian 1988; Courey et al. 1989). Proteins in which glutamine-rich domains or glutamine repeats have been recognized, are usually involved in control of gene expression and development (Bohmann et al. 1987; Duboule et al. 1987; Dynlacht et al. 1989; Gerster et al. 1990; Maity et al. 1990; Peterson et al. 1990; Tanaka & Herr 1990; Williams & Tjian 1991). In *Drosophila melanogaster*, proteins containing glutamine stretches (also referred to as opa-repeats) seem to play a central role in embryogenesis (Wharton et al. 1985; Haynes et al. 1987; Scott & Carroll 1987; Deatrick et al. 1991; Grabowski et al. 1991).

### *Proline-rich domains*

A third type of TAU consists of a proline-rich domain which was originally identified in CTF/NFI (Mermod et al. 1989). Related regions have been described in AP-2, *c-jun*, Oct-2 and SRF (reviewed in Mitchell & Tjian 1989).

Except for indefinite features like acidic, glutamine or proline-rich, none of the above described TAU-types is characterized by a common and conserved sequence-specific motif. It rather seems that TAUs, with respect to their primary amino acid sequence, are highly

transcription factor-specific. Besides the fact that TAUs are ill-defined, the mechanism by which these domains regulate transcription is poorly understood as well.

### 9.3 MODELS DESCRIBING THE MECHANISMS OF ACTION OF TRANSCRIPTION REGULATORS

Several models have been postulated that describe mechanisms in which transcription regulators communicate with the basal transcription machinery and thereby modulate gene expression (reviewed in Gill & Tjian 1992; Ing et al. 1992; Adams & Workman 1993).

#### *Model A: communication via bridging factors*

The first model proposes so-called bridging factors (co-activators) that physically link transcription regulators via protein-protein interactions to one of the GTFs (reviewed in Gill & Tjian 1992). Most evidence for the existence of bridging factors comes from the studies on TFIID and TBP, discussed previously. Recent observations support the view that these co-activators specifically mediate activation of only one type of TAU (either acidic, glutamine-rich, or proline-rich), and that the combination of co-activators, part of the TFIID fraction, might contribute to promoter- and cell specificity (Kelleher et al. 1990; Dynlacht et al. 1991; Tanese et al. 1991; Timmers & Sharp 1991; Brou et al. 1993) (reviewed in Gill & Tjian 1992; Zawel & Reinberg 1992; Tjian & Maniatis 1994). Indirect evidence for the existence of bridging factors comes from transcription interference experiments. This interference phenomenon (squenching) reflects the competition of different transcription regulators for common factors, which are essential for transcription and available at limiting concentrations in the cell (Ptashne 1988; Latchman 1990). Squenching experiments using different types of TAUs, demonstrated that acidic activation domains do not interfere with the activity of non-acidic TAUs located in the ER, which suggests bridging factor-specificity (Tasset et al. 1990). The hPR N-terminal domain has been observed to quench its own constitutive transactivation capacity in an *in vitro* transcription assay. This self-squenching did not affect basal transcription and was relieved by a partially purified cell-fraction, suggesting a bridging factor distinct from the GTFs (Shemshedini et al. 1992). Overexpression of the ER in CHO cells, or overexpression of the N-terminal TAU of the GR in yeast, interfered with endogenous cell growth, even resulting in cell death (Kushner et al. 1990; Wright et al. 1991). These latter studies suggest the squenching of a factor that is essential for cell-functioning, and that would probably be part of the basal transcription machinery which is postulated in the second model (McEwan et al. 1993).

#### *Model B: direct interaction between regulators and the basal transcription machinery*

The second model is essentially the same as the bridging factor model, except that transcription regulators directly contact the GTFs. Particularly TBP and TFIIB are the target GTFs that bind different transcription regulators. Various nuclear receptors have been shown to interact with TFIIB (Ing et al. 1992; Baniahmad et al. 1993). They include COUP-TF, PR, ER, and THR $\beta$ . Furthermore, VP16 directly binds TBP and TFIIB (Stringer et al. 1990; Lin et al. 1991). From these two models it can be concluded that TAUs are domains facilitating protein-protein interactions. These contacts, either direct or indirect, might induce and enhance the assembly of the basal transcription machinery by stabilizing the complex resulting in RNA

## Chapter 9

synthesis by RNA polymerase II.

### *Model C: nucleosome displacement*

In the third model nucleosome organization is taken into account. DNA packaging into higher order structures such as eu- and hetero-chromatin is known to strongly influence gene expression, most likely by modifying DNA accessibility to transcription factors (Chapter 7) (Getzenberg et al. 1991). The first order packaging of naked DNA into nucleosomes is involved in the regulation of transcription (reviewed in Adams & Workman 1993; Wolffe 1994). For initiation and elongation of transcription, nucleosomes need to be displaced. The presence of DNase I hypersensitive sites is usually interpreted as local alterations in chromatin structure, accompanied by nucleosome disruptions, and therefore often is used as a tool for promoter studies (reviewed in Elgin 1988). These studies revealed that transcription regulators can displace or disrupt nucleosomes, possibly making DNA accessible for other transcription regulators and the basal transcription machinery (reviewed in Adams & Workman 1993). One of the best studied promoters with respect to nucleosome organization is the MMTV-LTR. It has been proposed that GR-driven transactivation of the MMTV-LTR is achieved by a 'hit-and-run' mechanism (Rigaud et al. 1991, and references therein). GR binding to its GRE causes disruption of a nucleosome, after which NF-1 is able to associate with its DNA element thereby activating MMTV transcription (Archer et al. 1992; Truss et al. 1992a). These studies indicate that nucleosome organization can prevent transcription from promoters by precluding access of transcription factors. Nucleosome displacement could be one of the mechanisms by which transcription regulators induce gene expression (reviewed in Hager & Archer 1991).

Importantly, none of the above described models exclude one another as being the mechanism of transcription regulation. Since there is strong evidence for all three models, it seems more likely that they all contribute to the regulation of gene expression.

## **9.4 THE DIVERSITY ISSUE: MECHANISMS TO ACHIEVE RECEPTOR/CELL-SPECIFIC GENE REGULATION**

In previous paragraphs some general principles of promoter activation are described. The assembly of the basal transcription machinery is controlled by gene-specific regulators. The communication between these factors forms the basis for diversity and specificity of the control of gene expression, which has been the subject for many studies on steroid receptor-mediated gene expression.

Although the nuclear receptor family is characterized by a marked conservation of structure and function, each member regulates a specific set of genes. Besides this receptor-specificity, some of these genes will only be activated in certain cell types. The complexity of receptor- and cell-specific regulation of gene expression can elegantly be exemplified by the comparison of AR, GR, MR, and PR-mediated transcription. These four receptors recognize the same HRE but do not regulate the same set of target genes. A variety of mechanisms can achieve part of the diversity observed in the actions of these receptors: 1) cell-specific receptor expression, 2) ligand specificity, and 3) ligand availability (reviewed in Fuller 1991). For example, expression of prostate specific antigen (PSA) in LNCaP cells is induced by androgens but not by glucocorticoids (dexamethasone). However, when the GR is transfected into LNCaP

cells, dexamethasone will also stimulate expression of PSA (Cleutjens et al., submitted). Thus, androgen-specific regulation of PSA expression in prostate cells might simply be achieved by the exclusive expression of AR, whereas cell-specific PSA expression could be a result from gene-inactivation by condensation into heterochromatin in all other cell types. The same principle has been implicated in GR-specific gene expression in hepatoma cells (Strähle et al. 1989).

The diversity dilemma becomes complicated in situations where two or more different steroid hormone receptors, which can interact with the same HRE, are expressed in one cell and are all activated by their respective ligands. The specific regulation of the receptor's individual set of genes must then be dependent on the target genes promoter context. Through domain structures which are receptor-unique, nuclear receptors could interact specifically with other transcription regulators and with the basal transcription machinery, resulting in receptor-specific gene regulation. As was to be expected, the N-terminal domain of nuclear receptors which contains a transactivation function, is absolutely receptor-unique. Studies with promoters inducible with only one specific steroid receptor show that the promoter context as well as the receptor N-terminal domain determine receptor-specificity (Adler et al. 1992; Pearce & Yamamoto 1993).

Besides the fact that the AR, GR, MR and PR constitute P-box group I and recognize the same HRE (Chapter 8), the size of their N-terminal domain is extremely large as compared with all other nuclear receptors (Figure 1.3). For the thyroid hormone/retinoic acid receptors and orphan receptors, three different mechanisms have been described that increase diversity of action and explain cell and promoter discrimination: heterodimerization, variable spacing and direct/inverted orientation of the two HRE half-sites (Chapter 8.3). Since these variabilities are not observed for the AR, GR, MR, and PR, it seems reasonable to assume that their large and unique N-terminal domains compensate the dedication to only one kind of HRE (Chapter 8.3). A large N-terminal domain with multiple TAUs, could introduce diversity with respect to receptor- and cell-specificity.

Several mechanisms that contribute to the diversity observed in nuclear receptor-mediated gene expression, have not been mentioned or are discussed in previous chapters. These include: non-ligand activation (Chapter 2.4), phosphorylation (Chapter 2.2), A and B receptor isotypes (Chapter 9.6.1), and nuclear matrix composition (Chapter 7.3).

### 9.5 TRANSCRIPTION ACTIVATION UNITS IN STEROID RECEPTORS

By deletion mapping and insertional mutagenesis, transcription activation units in nuclear receptors have been located and characterized. Because of the extensive literature on this subject, this short review will be restricted to some of the work on the TAUs in the AR, GR, PR, and ER. In general, two different TAUs are present in steroid receptors: a TAU located in the N-terminal domain, and a hormone-dependent TAU in the ligand binding domain. However, for the ER and GR a weak transactivation function has also been attributed to the DBD (Hollenberg et al. 1987; Schena et al. 1989; Nardulli et al. 1991).

## 9.5.1 N-TERMINAL TRANSCRIPTION ACTIVATION UNITS

The N-terminal domains of AR, GR, PR, and ER are capable to activate transcription when linked to a heterologous DBD (from GAL4 or LexA), which shows that these receptor domains harbour a TAU by definition (Appendix paper IV; Godowski et al. 1987a; Hollenberg & Evans 1988; Berry et al. 1990; Tasset et al. 1990; Meyer et al. 1992; Shemshedini et al. 1992). Starting from the intact steroid receptor, deletion mapping has been performed to delimit the sequences essential for the N-terminal TAU (referred to as TAU-1). These studies established that the location of TAU-1 in the various steroid hormone receptors differs, and that there is little evidence for sequence- or structural homology (Figure 9.2) (Appendix paper IV; Giguere et al. 1986; Danielsen et al. 1987; Hollenberg et al. 1987; Kumar et al. 1987; Miesfeld et al. 1987; Guiochon-Mantel et al. 1988; Hollenberg & Evans 1988; Dobson et al. 1989; Imakado et al. 1991; Jenster et al. 1991 [Appendix paper I]; Simental et al. 1991; Meyer et al. 1992; Palvimo et al. 1993; Dahlman-Wright et al. 1994).

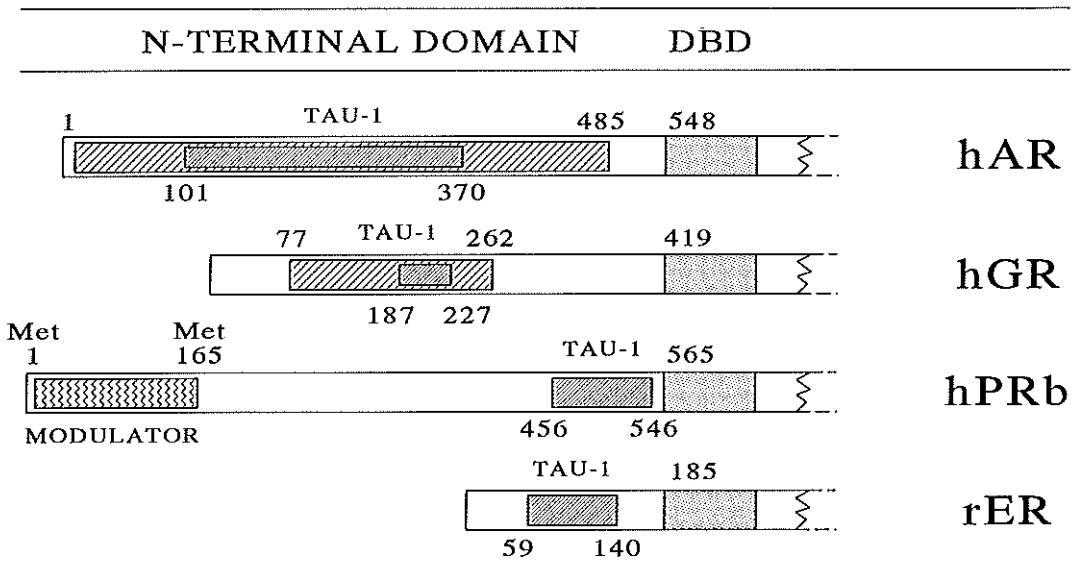


Figure 9.2 Location of different N-terminal TAUs-1 in various steroid receptors. Hatched boxes represent domains necessary for full transactivity (light shaded), and possible core sequences (dark shaded). The hPR contains a modulatory region between amino acids 1 and 165, that, by itself is incapable of inducing transactivation but enhances TAU-1 activity. The location and TAU-borders are based on Appendix paper IV (hAR), Dahlman-Wright et al. 1994 (hGR), Meyer et al. 1992 (hPRb), and Imakado et al. 1991 (rat ER).

The N-terminal TAU-1 of the hGR is characterized by a relatively high amount of acidic amino acids (Dahlman-Wright et al. 1994). In contrast, the N-terminal TAUs of the hPR and rat ER are not acidic but rich in proline residues, probably reflecting a functional and structural distinction (Imakado et al. 1991; Meyer et al. 1992). Comparison of the different TAUs by their

ability to synergize and to squelch, revealed that they belong to different classes of trans-activation domains (Tora et al. 1989; Tasset et al. 1990; Meyer et al. 1992; reviewed in Gronemeyer 1991). These results correspond with the theory that the TAUs of the AR, GR, PR and MR need to be different in order to effectuate receptor- and cell-specific gene activation. The hPRb contains a modulatory region between residues 1 and 165, that, by itself, is incapable of inducing transcription, but which enhances TAU-1 functioning. This region is present in the hPRb isoform but absent in the hPRa isoform, which is generated by an alternative translational start (Schrader & O'Malley 1972; Horwitz & Alexander 1983).

A striking finding in many of the deletion studies is that often large parts of the N-terminal domain of steroid hormone receptors can be deleted without affecting the transcription activating capacity of the receptor. This raises the question of the functional role of these specific regions. One possible answer has already been mentioned earlier by suggesting the presence of more than one TAU in the N-terminal domain of steroid hormone receptors (Chapter 9.4). Different N-terminal TAUs might be responsible for the regulation of different genes. This hypothesis was addressed in a study which aimed to delimit the TAU responsible for activity of the AR in various situations (Appendix paper IV). This study revealed that for the wild type AR, almost the entire N-terminal domain was necessary for full activity (TAU-1) when tested in HeLa cells using the (GRE)<sub>2</sub>tkCAT reporter.

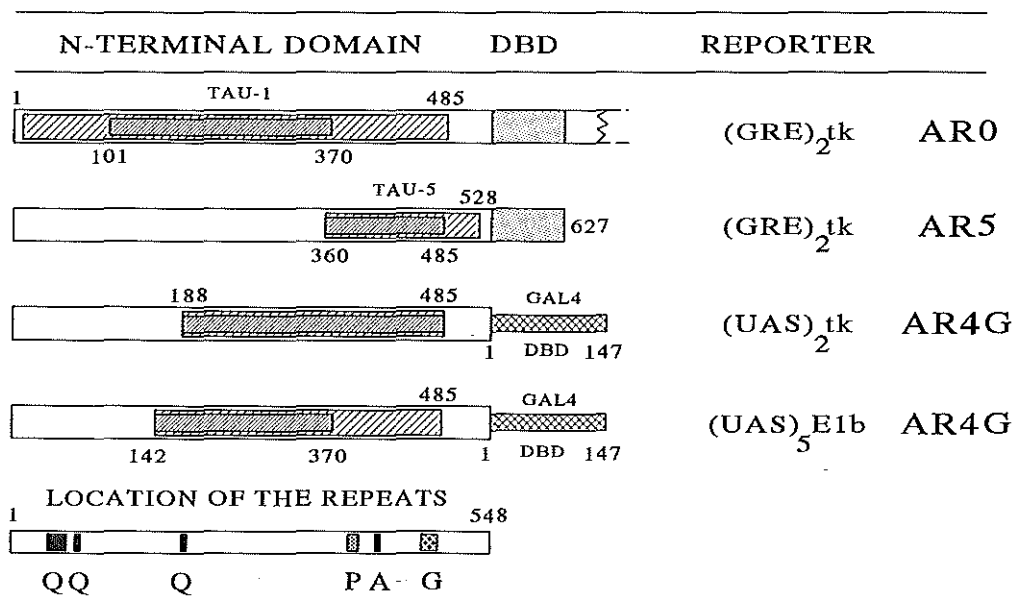


Figure 9.3 Regions in the N-terminal domain essential for transcription activation of the wild type AR (AR0) and mutant ARs. The borders of the TAU were located by analysis of a series of deletion mutants for transactivity of a certain reporter gene in HeLa cells (Appendix paper IV). Hatched boxes represent domains necessary for full activity (light shaded) or more than 50% activity (dark shaded). In addition, the location of the amino acid repeats are delineated (Q=glutamine; P=proline; A=alanine; G=glycine).

In contrast, only a small C-terminal part of the N-terminal domain (designated TAU-5), was sufficient for the constitutive active AR (lacking the LBD) to induce transcription of the same promoter in the same cell line (Figure 9.3) (Appendix paper IV). This demonstrates the capacity of the AR to use different N-terminal regions for transactivation. Furthermore, these observations suggest a determinant role for the LBD in TAU functioning. Deletion of the LBD resulted in loss of TAU-1 activity and induced the use of TAU-5.

Evidence for the role of different N-terminal regions in the activation of different kind of promoters, was provided by the analysis of a series of AR N-terminal-GAL4DBD chimeric constructs (Appendix paper IV). When transactivity was tested for two different promoters, a shift in position of the TAU responsible for full activity was observed. The location of the TAU was more C-terminal when tested on the (UAS)<sub>2</sub>tkCAT reporter compared to the (UAS)<sub>3</sub>E1b-CAT reporter (Figure 9.3). It is unlikely that these two situations represent the use of two different and separate TAUs, but rather indicate a small difference in the location of essential sequences. The same could account for the difference in location of the TAU between the wild type AR (TAU-1) and the constitutively active AR5 (TAU-5), since both regions overlap. However, the core of the two TAUs (responsible for over 50% of the activity) are separate N-terminal regions with individual characteristics. TAU-1 contains a relatively high amount of acidic amino acids, three glutamine repeats, and potential phosphorylation sites (Figure 9.3) (Faber et al. 1989; Kuiper et al. 1993; Appendix papers IV, V). TAU-5 is not acidic and harbours three different amino acid stretches: 1) a proline repeat, 2) an alanine repeat, and 3) a glycine repeat (Figure 9.3). Deletion of the Pro-stretch or Gly-stretch has no substantial effect on the activity of the hormone-regulated AR. A possible role of these repeats in TAU-5 functioning has not been investigated as yet. The largest and polymorphic glutamine repeat does play a role in transcription activation of the wild type AR. Deletion of the repeat resulted in an increased activity when tested in COS-1 cells on a reporter containing the promoter of the prostate-specific antigen (PSA) gene. The same AR mutant, however, had wild type activity when tested on a reporter containing the long terminal repeat of the mouse mammary tumor virus (MMTV-LTR) (Appendix paper V). Extension of the Gln-repeat from 20 to 48 residues resulted in a decreased AR activity on both promoters to 40-60% compared to the wild type AR (Appendix paper V; Mhatre et al. 1993). Lengthening of the Gln repeat to more than 40 residues is associated with Kennedy's disease, an X-linked neuro-degenerative disorder, characterized by a slowly progressing muscle weakness (Chapter 10.3) (La Spada et al. 1991).

Another feature of the AR TAU-1 is the impossibility to eliminate its activity by a small deletion. If the TAU-1 activity could be attributed to a small structural element, one would expect that disruption of that structure by a deletion would eliminate TAU-1 activity. Starting from an AR mutant lacking the region between residues 244 to 528, which was less than 10% active compared to the wild type AR, mutants were constructed with smaller deletions (of approximately 100 residues) covering this region. All AR mutants containing a smaller deletion, however, were quite capable of activating the transcription in the co-transfection assay (Appendix paper IV). This suggests the presence of two or more TAU-subunits, each one capable to 'take over' transactivity, or suggests that the TAU-1 structure is not characterized by essential core units. The inability to attribute TAU activity to a small structural element using AR deletion mapping studies, seems to be emphasized by the observation that up to now, no mutations (except for alterations resulting in premature translational termination) have been identified in the AR N-terminal domain of individuals with AIS (Chapter 10.2). The assumption



that it is impossible to identify essential motifs has also been implied for the hsp90 association with the LBD of steroid receptors (Chapter 3.3). Only deletion of the entire LBD eliminated hsp90 binding, but deletion of subregions did not affect hsp90 association (Carson-Jurika et al. 1989; Chambraud et al. 1990; Cadepond et al. 1991; Schowalter et al. 1991). Based on the plausible hypothesis that TAUs are protein-protein interaction sites, and taking into account all the previously mentioned TAU features, a model is proposed in which the AR N-terminal domain would be one large continuous modular structure. Different sides of this structure may form interfaces to which other proteins, such as bridging factors and GTFs can bind. These contact-interfaces together with the structural elements that shape them, would be defined as a TAU. Protein interactions might both be direct interactions between specific amino acids and of the 'casting-mould' type in which the recognition depends on the modular 3-dimensional conformation (Figure 9.4). This last type of binding (Figure 9.4; indicated by factor 1), does not need to be affected by a small deletion of any of its structural components.

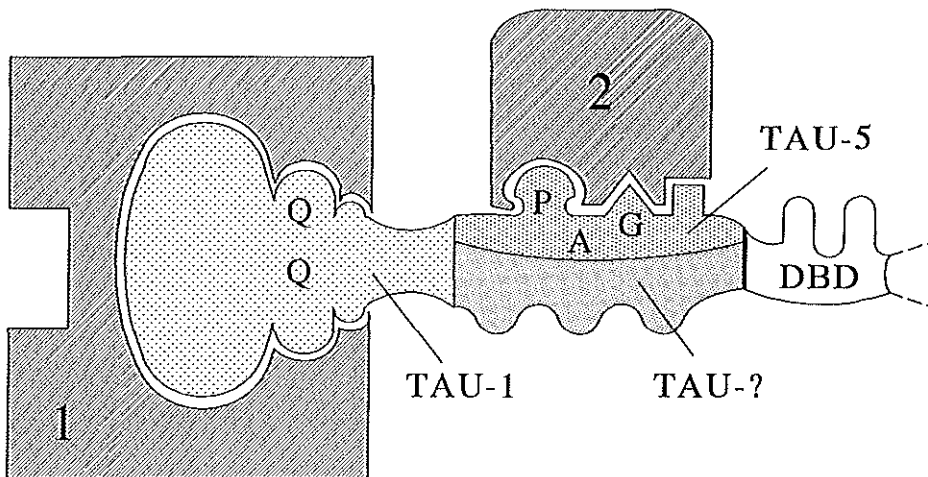


Figure 9.4 Hypothetical model describing possible protein-protein interactions between the AR N-terminal domain and other transcription factors. The N-terminal domain is delineated as one modular structure (stippled) to which on different sides proteins can interact. These contact interfaces together with the structural elements that shape the interface, would be defined as a TAU. Both residue-specific protein-protein contacts (indicated by interaction of factor 2 with the G-triangle and neighbouring box) and 'casting-mould' conformational interactions (indicated by factor 1 binding) could be possible. The single letters Q, P, A, and G represent the amino acid stretches.

### 9.5.2 C-TERMINAL TRANSCRIPTION ACTIVATION UNITS

The LBDs of the GR, PR and ER have been shown to contain a transcription activation unit (Danielsen et al. 1987; Hollenberg et al. 1987; Kumar et al. 1987; Miesfeld et al. 1987; Hollenberg & Evans 1988; Webster et al. 1988; Bocquel et al. 1989; Lees et al. 1989; Berry

et al. 1990; Tasset et al. 1990; Danielian et al. 1992; Tzukerman et al. 1994). This C-terminal TAU (referred to as TAU-2) is dependent on hormone for activity. In the LBD of the GR two different regions have been identified as possible TAUs. Hollenberg and Evans (1988) located an acidic region in the N-terminal part of the hGR LBD (amino acids 526 to 556) that functions as an activation unit. This domain is only moderately conserved among the steroid receptors, and no such sequence could be identified in the ER LBD when individual regions encoded by one of the exons 4 to 8 were tested for transactivity (Webster et al. 1989). Danielian and co-workers (1992) could specifically abolish hormone-induced transcription activation of the mouse GR by substitution of amino acids in a conserved C-terminal part of the LBD (residues 752 to 757 of the hGR, which are identical to the corresponding mouse sequences). This suggests that the conserved motif is essential for ligand-dependent transactivation and could be part of a TAU. The same motif is also required for the hormone-dependent transcription induction of the mouse ER (Danielian et al. 1992). So far, there is no evidence for the presence of a TAU in the LBD of the AR. Deletion of the N-terminal domain eliminated all AR transcription induction capacity (Appendix paper IV; Jenster et al. 1991 [Appendix paper I]; Simental et al. 1991; Palvimo et al. 1993). Chimeric AR constructs in which the AR LBD was linked to the GAL4 DBD were unable to activate transcription. However, immunocytochemical visualization of these AR mutants revealed that they formed large protein clusters in the cytoplasm, which might indicate that these mutants were incapable of activating transcription due to other defaults (see also Chapter 4.3).

### 9.6 COOPERATION BETWEEN TAUs AND BETWEEN TRANSCRIPTION REGULATORS

Characterization of TAUs in steroid receptors revealed that they are receptor-unique, of different types, and that receptors likely harbour more than one unit. These features form the basis for the specific interactions with other gene-specific transcription regulators and GTFs, allowing a receptor to regulate its own set of target genes.

#### 9.6.1 COOPERATION BETWEEN TAUs AND MODULATORY REGIONS

Intramolecular TAUs cooperate to enhance the transactivation potency of steroid hormone receptors. TAU-1 and TAU-2 have been shown to contribute synergistically to the total activity of the ER and GR (Hollenberg & Evans 1988; Tasset et al. 1990; Pham et al. 1992). In addition, the TAU-1 and TAU-2 activities of steroid receptors is determined by both cellular and promoter context (Bocquel et al. 1989; Berry et al. 1990; Tzukerman et al. 1994; reviewed in Gronemeyer 1991). These findings correspond with the TAU features and the expectation that they are involved in cell- and receptor-specific gene regulation.

Moreover, the modulatory region located in the first 164 amino acids of the hPRb adds another level of possibilities to achieve diversity of action (Meyer et al. 1992; Schrader 1993). Two distinct PR isoforms have been identified (Schrader & O'Malley 1972; Horwitz & Alexander 1983). The hPRb form is the largest and contains 164 additional amino acids in its N-terminus. The PRa is generated by an alternative translational start at codon position 165

(Figure 9.2). The two isoforms are synthesized from nine structurally different mRNAs, transcribed from two different promoters in the hPR gene (Kastner et al. 1990; Wei et al. 1990). Recent studies on the functional role of the two PR isoforms revealed that the smaller PRa dominantly inhibited PRb transactivation, which was dependent on the cell type and promoter context (Vegeto et al. 1993; Tung et al. 1993; reviewed in Schrader 1993). Interestingly, the inhibitory function of PRa did not require DNA binding and PRa could also inhibit AR, GR, and MR-mediated gene transcription.

A and b forms of the AR have also been demonstrated in human foreskin fibroblasts (Wilson & McPhaul 1994). The shorter (approximately 84-87 kDa) ARa isoform is synthesized, starting from codon-188 (ATG). The ARa expression is very low in fibroblasts compared to expression of the full length ARb. ARa is regularly observed in transiently transfected cells overexpressing the AR (Jenster et al. 1991 [Appendix paper I]; Simental et al. 1991), and the use of alternative translational starts has also been demonstrated in an AIS patient and in the Tfm mouse (Gaspar et al. 1990; Charest et al. 1991; He et al. 1991; Zoppi et al. 1993). When tested in HeLa cells using the (GRE)<sub>3</sub>tkCAT reporter gene, ARa (AR121) was capable of activating transcription (Appendix paper IV). Whether or not the ARa isoform is capable of inhibiting ARb functioning still needs to be established. So far it is unclear if the shorter AR isoform plays any physiological role.

## 9.6.2 COOPERATION BETWEEN TRANSCRIPTION REGULATORS

Besides intramolecular TAU-cooperation and concerted action of the homo- or heterodimer on the HRE, nuclear receptors have been shown to synergize with transcription regulators that bind to neighbouring elements in target promoters.

Already the presence of two or more HREs results in a synergistic transcription activation compared to the presence of only one HRE, suggesting cooperation between steroid receptor dimers (Schüle et al. 1988; Tsai et al. 1989; Baniahmad et al. 1991; Bradshaw et al. 1991; Wright & Gustafsson 1991; Rupperecht et al. 1993; reviewed in Muller et al. 1991b). This synergism has been shown in several cases to be partially or solely dependent on the receptor's TAU (Baniahmad et al. 1991; Bradshaw et al. 1991; Wright & Gustafsson 1991; Rupperecht et al. 1993). In addition to cooperation induced by two similar response elements, synergism has also been demonstrated between heterologous response elements. If one of the elements concerns a GRE, the other elements include: CACCC-box, CAAT-box, NF-1, Sp1, Oct-2A, and Oct-1 (Schüle et al. 1988; Schatt et al. 1990; Allan et al. 1991; Brüggemeyer et al. 1991; Muller & Renkawitz 1991; Muller et al. 1991a; Wieland et al. 1991; Truss et al. 1992a) (reviewed in Muller et al 1991b; Truss & Beato 1993). Besides (positive) synergism, several examples have been described for negative regulation by steroid hormone receptors: 1) interference of DNA binding due to overlapping response elements (Akerblom et al. 1988), 2) formation of an inactive GR complex containing three GR molecules on a negative GRE (Drouin et al. 1993), and 3) cross-talk between steroid receptors and the AP-1 (*fos/jun*) regulatory pathway, which seems to involve both inactivation through direct *fos/jun*-receptor interactions and DNA binding competition (reviewed in Clark & Docherty 1993; Pfahl 1993; Truss & Beato 1993).

### 9.7 CONCLUDING REMARKS

In this chapter, the basic principles are described of the regulation of TATA-box-containing promoters that are under the control of steroid hormone receptors. Models are proposed with respect to two central questions: 1) how are steroid receptors and other transcription regulators able to activate (or inhibit) transcription?, and 2) how do steroid receptors 'know' when and which genes to regulate in what types of cells? Several processes have been postulated concerning the communication between transcription regulators and the basal transcription machinery. Those protein domains of the regulators that are essential in this communication are referred to as transcription activation units. TAUs are highly regulator-specific and likely represent protein-protein interfaces. The phrases 'unique' and 'specificity' are repeatedly used with respect to the diversity issue which covers the problem of steroid hormone receptors only regulating their own set of target genes in a cell-specific pattern. Multiplicity is achieved through many variabilities, such as multiple TAUs within one receptor that are all receptor-unique, modulatory regions, and positive and negative cooperation between different transcription regulators. These features and processes in combination with many more, complicate the detailed understanding of specific gene expression, but provide enough mechanisms to achieve the diversity observed in the actions of nuclear receptors.

**ANDROGEN RECEPTOR MUTATIONS IN HUMAN DISEASE**

### 10.1 INTRODUCTION

As is clear from previous chapters, intentional mutagenesis provided a large amount of data on AR functional domains in *in vitro* and *ex vivo* (cell line transfection) situations. Information on *in vivo* AR functioning can be obtained from naturally occurring mutations in several human diseases. Mutations in the AR gene have been demonstrated in subjects with androgen insensitivity syndrome (AIS), Kennedy's disease, and patients with prostate cancer, and in males with breast cancer. These *in vivo* situations provide the opportunity to correlate receptor alteration with the impairment of its physiological function, which might confirm and add new perspectives to the understanding of AR action.

### 10.2 ANDROGEN INSENSITIVITY SYNDROME

The complete form of AIS (also referred to as testicular feminization [Tfm]) is an X-linked disorder, characterized by 46,XY karyotype, external female phenotype, intra-abdominal testes, absence of uterus and ovaries, blind ending vagina and gynecomastia (reviewed in Griffin & Wilson 1989; Schweikert 1993). Besides the complete form, mutations in the AR gene cause a variety of clinical related disorders, covered by the term 'Reifenstein syndrome' which will be referred to as partial AIS.

Insensitivity to steroid hormones has also been reported for glucocorticoids, mineralocorticoids, progesterone, thyroid hormone and vitamin D<sub>3</sub>, although these cases (except for thyroid hormone resistance) are not as frequent as androgen insensitivity and many of them concern partial syndromes (Keller et al. 1979) (reviewed in Hughes & O'Malley 1991; Chrousos et al. 1993; Refetoff et al. 1993). The relatively high incidence of AIS compared to the other insensitivity syndromes, can be explained by several observations. Firstly, the AR gene is the only nuclear receptor so far, found to be located on the X-chromosome, and AR mutations thus become clinically manifested in males (XY). Mutations in other receptor genes located on autosomes are clinically evident: 1) in the homozygous state, which is a rare event if the number of mutations in the population is low, or 2) in the heterozygous state as a relatively mild syndrome due to the compensating effect from the unaffected allele, or as a more severe defect as a result of a dominant negative receptor form. Such dominant negative receptor forms have been described in generalized resistance to thyroid hormone and explain its autosomal dominant inheritance and relatively high number of reported cases (Refetoff et al. 1993). Secondly, androgen action is necessary for male sexual differentiation and fertility, but is not essential for life of the individual. In contrast, estrogen, progesterone, and cortisol action have always been thought to be essential for processes during embryogenesis (such as blastocyst implantation) and therefore complete insensitivity to these hormones caused by receptor mutations would not be compatible with life. However, with respect to estrogen action, this view needs to be reconsidered since homozygous disruption of the ER gene in mice did not result in prenatal death, although the reproductive functions of these mice were severely affected (Lubahn et al. 1993).

A variety of different kind of mutations have been reported in the AR gene causing AIS (reviewed in McPhaul et al. 1991; Brinkmann & Trapman 1992; Pinsky et al. 1992; Sultan et al. 1993). They include large and single base pair deletions, insertions, and base substitutions.

Especially the subtle alterations such as missense mutations, resulting in amino acid substitutions add new information to the *in vitro* and *ex vivo* deletion mapping studies. Most mutations identified so far are located in the LBD (Figure 10.1). This is not surprising since it is known that almost any alteration in the large LBD interferes with ligand binding (Chapter 5). However, the overall distribution of mutations known today might be somewhat distorted, since originally most subjects chosen to be screened for AR mutations concerned cases with loss of androgen binding.

Amino acid substitutions affecting ligand binding are scattered throughout the LBD, although the frequency of mutations seems to be higher in some regions (McPhaul et al. 1992). The part of the LBD encoded by exon 5 (residues 716 to 763) contains a relatively large number of mutations. The first half (residues 716 to 739) of this region is characterized by mutations predominantly causing partial AIS, whereas the second half (amino acids 740 to 763) concerns almost exclusively substitutions causing the complete form (Figure 10.1). Also the first part of the LBD (residues 655 to 709), encoded by the second half of exon 4, only concerns mutations causing complete AIS.

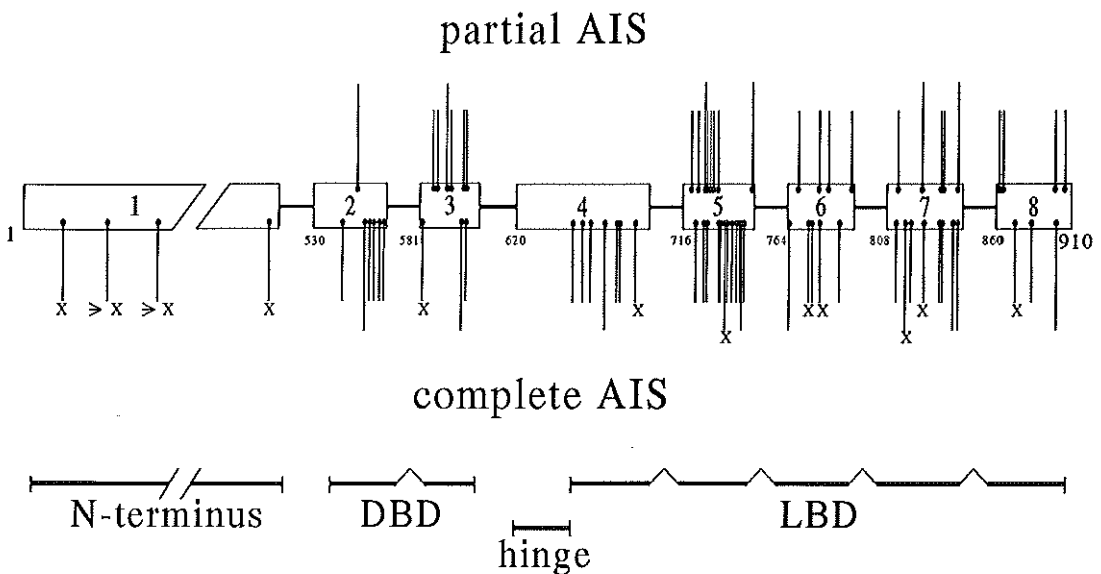


Figure 10.1 Location of amino acid substitutions, small deletions and insertions identified in subjects with complete or partial AIS. The AR amino acid sequence is depicted as boxes represented by the corresponding exons. The N-terminal domain is encoded by part of exon 1, and so far, 4 mutations have been identified in this domain that all resulted directly in a translation termination ( $\rightarrow X$ ) or resulted in a different reading frame eventually leading to a translation termination ( $\rightarrow X$ ). In some cases, the same residue has been substituted into different amino acids (or termination codons [X]) in different patients (indicated as longer lines) (mutations are reviewed in McPhaul et al. 1991; Brinkmann & Trapman 1992; Pinsky et al. 1992; Sultan et al. 1993)

## Chapter 10

Also mutations in the DBD causing AIS have been identified (Beitel et al. 1994; reviewed in Sultan et al. 1993). The correlation between the substitution of essential amino acids and the consequent development of the complete form of AIS is quite strong. Any substitution of cysteine residues involved in zinc-ion binding, or mutations of residues contacting the base pairs in the HRE, resulted in the complete form of AIS.

In the hinge region and in the N-terminal domain, no mutations have been found, except for alterations causing premature translational termination. This observation raises the question if single amino acid substitutions in these two domains will ever cause sufficient damage to the receptor to result in partial or complete AIS. The fact that the N-terminal domain and hinge region are absolutely not conserved among nuclear receptors, and the observation that small deletions in these domains could not eliminate AR functioning in transfection assays, strongly suggests that such substitutions will not cause AIS.

Premature stop-codons in various parts of the receptor mRNA introduced by base pair substitutions or mutations that result in a different reading frame, have been identified in individuals with AIS. As expected, all cases concern the complete form. Interestingly, in a subject with a premature stop-codon at codon position 60 (Gln=CAG→TAG=stop), androgen binding could still be detected, suggesting alternative initiation downstream of the termination codon (Zoppi et al. 1993). This finding, together with the identification of a 87 kDa AR form in normal fibroblasts, has led to the hypothesis of a and b AR isoforms (Chapter 9.6.1) (Wilson & McPhaul 1994).

### 10.3 KENNEDY'S DISEASE

Kennedy's disease or SBMA (spinal and bulbar muscular atrophy) is an X-linked, slowly progressing, muscle weakness which becomes manifest in adult males (Kennedy et al. 1968; Arbizu et al. 1983). The syndrome is a motor-neuron disease in which bulbospinal motor-neurons slowly degenerate. Affected males may also develop gynaecomastia and sometimes have reduced fertility, which suggested a defect in AR functioning. Analysis of the AR gene revealed an absolute correlation with an extension of the (CAG)<sub>n</sub> repeat in the coding region of exon 1 (La Spada et al. 1991). The repeat encodes a glutamine stretch which is polymorphic and encompasses between 11 and 31 residues, with an average repeat number of 21, which is most common in the (normal) human population (Caskey et al. 1992; Edwards et al. 1992; Sleddens et al. 1992). In Kennedy's disease, the amount of glutamines varies between 40 and 53, and there seems to be a correlation between the number of residues and the clinical onset of the syndrome (La Spada et al. 1991; Biancalana et al. 1992; Igarashi et al. 1992).

Lengthening of a (CAG)<sub>n</sub> repeat has also been correlated with Huntington's disease (HD), spinocerebellar ataxia type I (SCA-1), and dentatorubral and pallidolysian atrophy (DRPLA) (Orr et al. 1993; The Huntington's disease collaborative research group 1993; Koide et al. 1994; Nagafuchi et al. 1994; reviewed in Green 1993b). In fact, these syndromes have surprisingly many aspects in common: 1) all four are progressive neuron degenerative disorders, 2) in all four syndromes, the length of the (CAG)<sub>n</sub> repeat is expanded to more than 40 residues, while on average approximately 16 to 25 residues are present in the normal state, 3) in Kennedy's and Huntington's disease, and presumably in SCA-1 and DRPLA, the polymorphic repeat encodes a glutamine stretch in the corresponding proteins, and 4) the length of the repeat

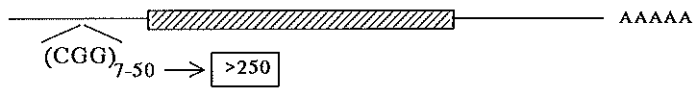


inversely correlates with the onset of the clinical features (Andrew et al. 1993; Chung et al. 1993; Duyao et al. 1993; Myers et al. 1993; Orr et al. 1993; Snell et al. 1993; The Huntington's collaborative research group 1993).

Two other disorders have been associated with triplet repeat expansions: Fragile X syndrome (FraXA) and myotonic dystrophy (DM) (Kremer et al. 1991; Verkerk et al. 1991; Brook et al. 1992; Fu et al. 1992; Mahadevan et al. 1992) (reviewed in Caskey et al. 1992; Edwards et al. 1992; Richards & Sutherland 1992). These syndromes however, differ in several aspects with the (CAG)<sub>n</sub> repeat related diseases. They concern other triplets (FraXA: CGG; DM: CTG) that are not located in a coding region and can extend in disease states to more than 1000 triplets (Figure 10.2; reviewed in Caskey et al. 1992).

---

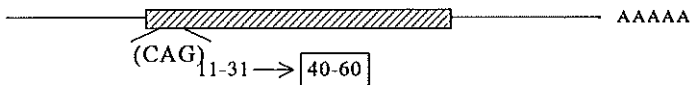
- Fragile-X Syndrome ( FMR-1 )



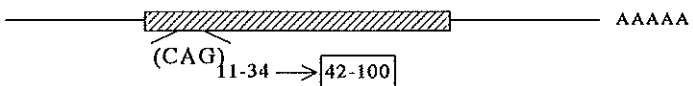
- Myotonic Dystrophy ( Myotonin Kinase )



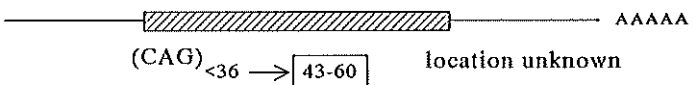
- Kennedy's Disease ( Androgen Receptor )



- Huntington's Disease ( Huntingtin )



- Spinocerebellar Ataxia type 1



- Dentatorubral and pallidolusian atrophy

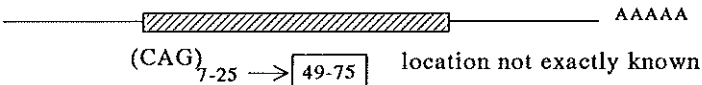


Figure 10.2 Syndromes correlated with trinucleotide repeat extension. The location, sequence and normal range of the amount of triplets are shown in the mRNA of the affected gene. The length of the repeat in the diseased state is boxed. For references see text.

## Chapter 10

The role of the glutamine repeat is unknown, although its function in AR transcription activation has been investigated (Chapter 9.5.1; Appendix paper V). Deletion of the stretch resulted in an increased transactivity, whereas an extension to 48 or 50 residues reduced AR activity to approximately 50% (Appendix paper V; Mhatre et al. 1993). These transfection results correspond to the clinical features of gynecomastia and reduced fertility in some Kennedy's disease patients, that indicate a decreased AR functioning. It could be argued that a reduced AR activity would be damaging to motor-neurons and subsequently cause Kennedy's disease. There are reports of testosterone effects on rat and human neurons and the presence of ARs in rat and human motor-neurons has been demonstrated (Araki et al. 1991; Kurz et al. 1992; Hodges et al. 1993; Matsuura et al. 1993; Zhou et al. 1993; Nakamura et al. 1994). However, individuals with partial or complete AIS and normal women (low testosterone) have no clinical features of Kennedy's disease, which implies that a decreased AR functioning itself cannot be the exclusive cause. These contra-arguments rather indicate a positive dominant effect of the (CAG)<sub>n</sub> repeat extension on motor-neuron degeneration, which is the case in HD, SCA-1 and DRPLA. The question then arises: why do women who carry the affected AR not develop the syndrome? Possibly, due to X-chromosome inactivation, only half of the respective neurons express the affected AR, which means that female carriers will retain a sufficient number of non-affected motor-neurons to prevent manifestation of clinical features. In addition, peripheral male levels of testosterone might be needed for the affected AR to become harmful for neurons. In conformity with this reasoning, Green (1993b) suggested that ARs containing an expanded Gln-stretch would be a better substrate for the enzyme transglutaminase (TGase). TGase, which is expressed in many different cell types (including neuronal tissues), catalyses the coupling of glutamine and lysine residues via their side chain groups (Chapter 2.1; Figure 2.1 for reaction). The reaction results in a covalent isopeptide crosslink that is stable and resistant to proteolysis (Greenberg et al. 1991). Since the expanded Gln-repeat might turn the AR into a TGase substrate, gradual accumulation of an undegradable isopeptide may eventually become toxic to the cells expressing the affected AR. In combination with the fact that neuronal cells and few other cell types such as Sertoli cells do not regenerate, the late onset and slowly progressive features are explained. This hypothesis is in agreement with the observation that the larger the extended repeat, the earlier the onset of clinical symptoms, and also explains the observation that up to now no regions coding for Gln-stretches of more than 38 residues were found in normal alleles (Green 1993b). The TGase hypothesis is likely also applicable to Huntington's disease, spinocerebellar ataxia type I, and dentatorubral and pallidolusian atrophy. The origin of the partially different clinical symptoms between these four syndromes, might be due to expression level, cell-specific expression and characteristics of the respective proteins carrying the expanded repeat.

### 10.4 PROSTATE CANCER AND MALE BREAST CANCER

One of the target organs for androgen action is the prostate. Its development, growth, and maintenance of functions is largely dependent on DHT, which is also true for primary prostatic tumours (reviewed in Schröder 1991; Gittes 1991; Santen 1992; Dearnaley 1994). Despite the initial androgen dependency of human prostate cancer, eventually they will become androgen-independent during androgen-ablation or antiandrogen therapy. The mechanism of this

transition is unclear, although the AR could in theory play a major role. Deletion of the LBD for example, converts the receptor into a constitutively active transcription regulator (Jenster et al. 1991 [Appendix paper I]). Moreover, mutations in the LBD could increase androgen affinity, or change hormone- or antiandrogen specificity. Examples of such receptor alterations have been reported for several steroid receptors. In female breast cancer, variant human ERs have been demonstrated, that lack part of their LBD and display a constitutive transcription activity (Fuqua et al. 1991). A single amino acid substitution in the LBD of the rat GR increased its affinity for cortisol (Chakraborti et al. 1991).

For the AR, a mutation has been identified in the LBD (868: Thr→Ala) that resulted in a changed ligand binding specificity (Veldscholte et al. 1990). Interestingly, this substitution was originally found in the prostate cancer cell line LNCaP (lymph node carcinoma of the prostate), but has recently also been demonstrated in metastatic prostate cancer tissue from a patient treated with antiandrogens (Suzuki et al. 1993). Since it has been shown that due to this substitution antiandrogens, estrogen, and progesterone are able to activate the mutant AR, it seems plausible that the antiandrogen treatment had an opposite effect and selectively induced growth of cells expressing this AR mutant.

**Table 10.1 AR mutations in prostate and male breast cancer**

**PROSTATE CANCER**

Mutation	Position	Effect	Reference
(CAG) <sub>24</sub> →(CAG) <sub>18</sub>	55	unknown	Schoenberg et al. 1994
Leu → His	692	unknown	Suzuki et al. 1993
Val → Met	706	increased transactivity by adrenal androgens and progesterone	Olig et al. 1993
Val → Met	721	wild type activity	Newmark et al. 1992 Schoenberg et al. 1994
Thr → Ala	868	LNCaP mutation: changed affinity for hormones and antiandrogens	Suzuki et al. 1993

**MALE BREAST CANCER**

Mutation	Position	Effect	Reference
Arg → Gln	598	partial AIS	Wooster et al. 1992
Arg → Lys	599	partial AIS	Lobaccaro et al. 1993b

*Mutations in the human AR identified in prostate cancer and male breast cancer.*

## Chapter 10

Another residue substitution has been identified in the AR LBD resulting in different AR characteristics. The mutation alters a valine into a methionine residue at position 706 and was found in endocrine therapy-resistant prostate cancer (Culig et al. 1993). Although the binding affinities of this AR mutant for a variety of compounds was not different from those of the wild type AR, progesterone and adrenal androgens (dehydroepiandrosterone and androstenedione) mediated a higher transcription activation through the mutant than through the wild type AR (Culig et al. 1993). Besides these two AR mutations, only three other alterations have been described in prostate cancer (Table 10.1). The effect on AR functioning has not been reported for all of them yet (Suzuki et al. 1993; Schoenberg et al. 1994). The total group of prostate cancer patients screened for AR mutations exceeds 150, and it can be concluded that this kind of alterations do not play an important role in the etiology of prostate cancer or of transition into androgen-independent growth.

Besides prostate cancer, the relatively rare male breast cancer has also been implicated as a disease possibly induced or enhanced by AR mutations. In two different cases of partial AIS, the subjects developed breast cancer (Table 10.1) (Wooster et al. 1992; Lobaccaro et al. 1993b). Up to now, it is not clear if the AR substitutions directly influenced the breast cancer development and growth, or that the cancer is a result from the increased risk of breast cancer because the AR mutation caused partial AIS and subsequent gynaecomastia (Lobaccaro et al. 1993a). Like prostate cancer, the very few mutations identified in the group of patients screened, reveals that AR mutations are not one of the main causal factors of male breast cancer.

CONCLUSIONS

From the large number of studies on AR functioning it is evident that the AR's ultimate objective is specific and efficient transcription regulation. To perform this function, the AR has 910 amino acids at its disposal that are partitioned over several functional and structural domains. For transcription activation, the receptor needs at least a transcription activation unit (TAU) located in the N-terminal domain, and a DNA binding domain (DBD) which consists of two zinc-finger modules. The AR shows androgen-dependency through a ligand binding domain which is linked to the DBD via the hinge region (Figure 11.1). Additional signals and units are certainly or likely necessary for: 1) nuclear import, 2) efficient protein folding and transport to- and into the nucleus chaperoned by hsp90, 3) intra-nuclear positioning and anchorage by nuclear matrix binding, 4) specific and cooperative DNA binding through dimerization, and 5) enhancement of transactivity by N-terminal phosphorylation.

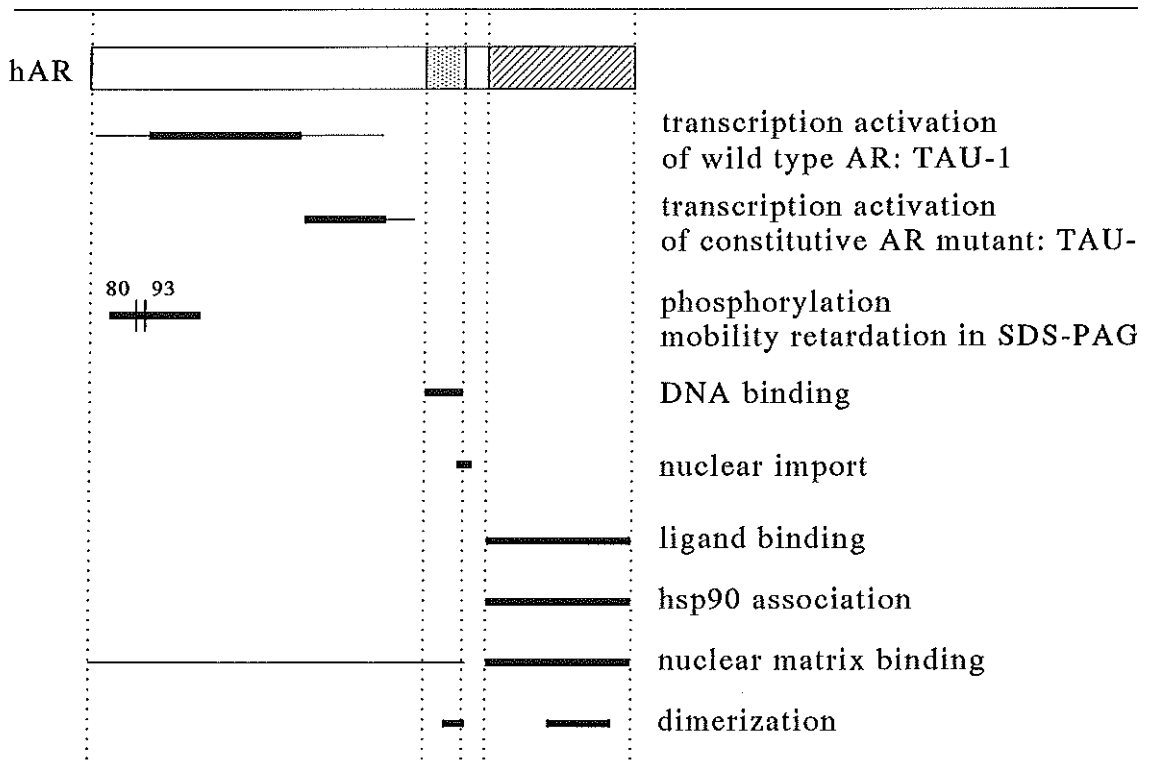


Figure 11.1 Localization and size of the various signals and domains in the human AR responsible for the different AR characteristics. Details of the different domains are discussed in the corresponding chapters.

Based on the 'exons coding for protein units' theory described in Chapter 1.1.3, it can be expected that every receptor domain responsible for a specific feature is encoded by a separate exon. This seems to be true for transcription activation and DNA binding, but many other activities are executed by the large LBD which is encoded by 5 different exons. The

observation that one large domain performs various AR activities that seem indistinguishably integrated, suggests degeneration of the original exon-protein unit pattern.

Besides the efficient execution of transcription regulation, this process is highly specific regarding AR unique target genes that are expressed in specific cell types. To comply to specificity, diversity among the steroid receptors is achieved mainly by cell-specific receptor expression, ligand-specificity and availability, and the large and unique N-terminal domain. The N-terminal domain contains different, but overlapping TAUs which might share structural elements that shape them, but at the molecule surface most likely interact specifically with different transcription factors. Multiple protein-protein interfaces (as part of TAUs) that are AR-unique, provide the conditions for gene-specific transcription regulation.

Naturally occurring AR mutations and their physiological effects confirm results from *in vitro* and cell line transfection mutational analysis studies. Many substitutions in the AR DBD and LBD disrupt these domain's respective functions. On the other hand, the absence of naturally occurring amino acid substitutions in the AR N-terminal domain, might reflect the impossibility to attribute transactivity to only a small N-terminal motif in transfection assays. This raises the question whether any small N-terminal mutation (except for those resulting in translational termination) can be disruptive enough to cause physiological AR impairment. The general conformity between the *in vivo* findings and cell line transfection studies, shows that transient transfection is a useful model system to investigate AR functioning.

Several major challenges confront the field of nuclear receptor research, among which the elucidation of the 3-dimensional conformation of the LBD and the N-terminal domain. This knowledge would provide an enormous amount of information concerning specific ligand interaction, and other LBD features such as hsp binding, nuclear matrix association and dimerization, and on protein-protein interfaces of the N-terminal domain. The latter information could be very useful to the expanding field of research on the large heteromeric protein complexes positioned on gene promoters, and on the role of the receptor therein.

## Chapter 11



DOMAINS OF THE HUMAN ANDROGEN RECEPTOR INVOLVED IN  
STEROID BINDING, TRANSCRIPTIONAL ACTIVATION AND  
SUBCELLULAR LOCALIZATION

Guido Jenster<sup>1</sup>, Hetty A.G.M. van der Korput<sup>2</sup>, Cor van Vroonhoven<sup>2</sup>,  
Theo H. van der Kwast<sup>2</sup>, Jan Trapman<sup>2</sup> and Albert O. Brinkmann<sup>1</sup>

Departments of Endocrinology & Reproduction<sup>1</sup> and Pathology<sup>2</sup>,  
Erasmus University Rotterdam, Rotterdam, The Netherlands

Modified version of Mol Endocrinol (1991) 5:1396-1404

### SUMMARY

A series of human androgen receptor (AR) deletion mutants was constructed to study the relationship between the structural domains and their different functions in the AR protein. Human AR mutants were expressed in COS-1 and HeLa cells to investigate hormone binding, transcriptional activation and subcellular localization. The wild type human AR (AR 1-910) was expressed as a 110-112 kDa doublet as revealed on immunoblots. All mutant AR proteins also migrated as doublets, except for one. This AR has a deletion from amino acid residue 51 to 211 and migrated as a single protein band, possibly due to altered post-translational modification. The AR steroid binding domain is encoded by approximately 250 amino acid residues in the C-terminal end. Deletions in this domain as well as truncation of the last 12 C-terminal amino acid residues abolished hormone binding. Co-transfection studies in HeLa cells showed that transcriptional activation of an androgen regulated reporter gene construct was induced by the wild type human AR. Mutational analysis revealed a region in the N-terminal part encoded by amino acid residues 51-211, to be essential for this transcriptional activation. Deletion of the hormone-binding domain yielded a constitutively active AR protein, indicating that in the absence of hormone this domain displays an inhibitory function. In the presence of its ligand the wild type AR was located in the cell nucleus. In the absence of androgens the receptor was mainly nuclear but cytoplasmic localization was observed as well. In contrast, an AR deletion mutant lacking part of the DNA binding domain and part of the hinge region was exclusively cytoplasmic in the absence of hormone. This mutant AR lacks a region which is highly conserved among steroid receptors and homologous to the SV40 large T antigen- and nucleoplasmic- nuclear localization signals. Transformation by hormone directed this mutant to the nucleus, indicating the presence of a second, hormone-dependent nuclear targeting mechanism.

### INTRODUCTION

Androgens are essential for sexual differentiation of the male embryo and responsible for initiation and maintenance of spermatogenesis. Testosterone and dihydrotestosterone mediate their function through the intracellular androgen receptor (AR). Upon binding of androgen, the steroid-receptor complex interacts with specific DNA sequences, located in the flanking regions of target genes, resulting in the up or down regulation of the expression of these genes.

The AR is a member of the superfamily of ligand-dependent transcription factors, which includes receptors for retinoic acid, vitamin D<sub>3</sub>, thyroid hormone and other steroid hormones (Evans 1988; Green & Chambon 1988). Mutational analysis combined with functionality studies, revealed three major functional domains within the glucocorticoid- (GR), progesterone- (PR) and estrogen- (ER) receptors (Giguere et al. 1986; Gronemeyer et al. 1987; Kumar et al. 1987; Dobson et al. 1989). The poorly conserved N-terminal region is involved in transcriptional activation. Two more conserved domains correspond to the cysteine-rich DNA-binding domain and the C-terminal ligand-binding domain. Other sequences, connected with common steroid receptor characteristics such as interaction with 90 kDa heat shock protein (Kost et al. 1989; Pratt et al. 1989; Howard et al. 1990; Pratt 1990), receptor dimerization (Tsai et al. 1988; Fawell et al. 1990) and nuclear entering (Picard & Yamamoto 1987; Guiochon-Mantel et al.

1989) also show some conservation. Structural homology between the AR and other steroid receptors indicates that similar regions are present in the AR.

Important information concerning the function of AR domains can be deduced from naturally occurring mutations in the AR in individuals with the androgen insensitivity syndrome. A number of different point mutations and deletions have illustrated the importance of the C-terminal region in functional ligand binding (Lubahn et al. 1989; Brown et al. 1990; Marcelli et al. 1990b; Ris-Stalpers et al. 1990; Sai et al. 1990).

The subcellular localization of the PR and ER is mainly nuclear, even in the absence of ligand (King & Greene 1984; Perrot-Appianat et al. 1985; Guiochon-Mantel et al. 1989). In contrast, the GR is mainly cytoplasmic in the absence, and nuclear in the presence of hormone (Fuxe et al. 1985; Picard & Yamamoto 1987; Wikström et al. 1987). Immunocytochemical analysis of the subcellular localization of the AR in normal and castrated rats shows the AR to be located in the nucleus (Husmann et al. 1990; Sar et al. 1990). Similarly, in prostate tissue derived from prostate cancer patients under androgen ablation therapy, the AR is mainly nuclear (van der Kwast et al. 1991). In contrast, recent *in vitro* mutagenesis studies with the human AR overexpressed in COS cells, showed the AR to be cytoplasmic in the absence and nuclear in the presence of androgens (Simental et al. 1991).

Different mechanisms for nuclear entering have been proposed for the GR and PR (Picard & Yamamoto 1987; Guiochon-Mantel et al. 1989). A SV40 large T antigen-like nuclear localization signal (NLS) constitutively directs the PR to the nucleus. This NLS, although present, seems not to be functional in the GR. The GR becomes nuclear after activation of a second, hormone-dependent translocation process.

In an ongoing study on AR expression and the molecular mechanism of androgen action, we are interested in defining the relationship between the structural domains and their various functions in the human AR. We have constructed a set of AR deletion mutants and expressed these transiently in COS-1 and HeLa cells. The ability of these AR mutants to bind hormone and to activate transcription was examined. In addition, the subcellular localization of these AR mutants was determined.

## RESULTS

### Expression of androgen receptor proteins

Full-length cDNA coding for the human AR was cloned into an eukaryotic expression vector (pSV-AR0, coding for AR 1-910) designed for high level expression in COS and HeLa cells (Brinkmann et al. 1989). From this expression vector, a series of deletion mutants was constructed. COS-1 cells were transfected with plasmid DNA, and expressed AR was analyzed by immunoblotting. A monoclonal antibody (F39.4.1) raised against a synthetic peptide corresponding to AR amino acid residues 301-320, was used to specifically immunoprecipitate the AR (van Laar et al. 1989; Zegers et al. 1991).

As shown in Fig. 1A, the wild type human AR (AR 1-910) migrated as a 110-112 kDa doublet, possibly reflecting different degrees of phosphorylation. A less prominent 82 kDa band may result either from internally initiated translation or from proteolytic degradation. If translation would start at the second or third AUG it would result in AR proteins of

## Appendix paper I

approximately 87 and 80 kDa, respectively. With one exception (AR $\Delta$ 51-211), all the mutant AR proteins (see Figure 2) migrated as doublets (Figure 1A). Besides being expressed as a single protein band, AR $\Delta$ 51-211 migrated with the same mobility as AR 1-653, although the calculated molecular mass of AR $\Delta$ 51-211 is approximately 13 kDa higher than that of AR 1-653 (Figure 1B). These findings indicate a possible post-translational modification event in the region 51-211.

Two mutant AR proteins (AR $\Delta$ 37-494 and AR $\Delta$ 244-360) cannot be recognized by the monoclonal antibody because they lack the peptide sequence to which the antibody is directed. Since AR $\Delta$ 37-494 and AR $\Delta$ 244-360 specifically bind R1881, we conclude that both receptor mutants are properly expressed (Figure 2).

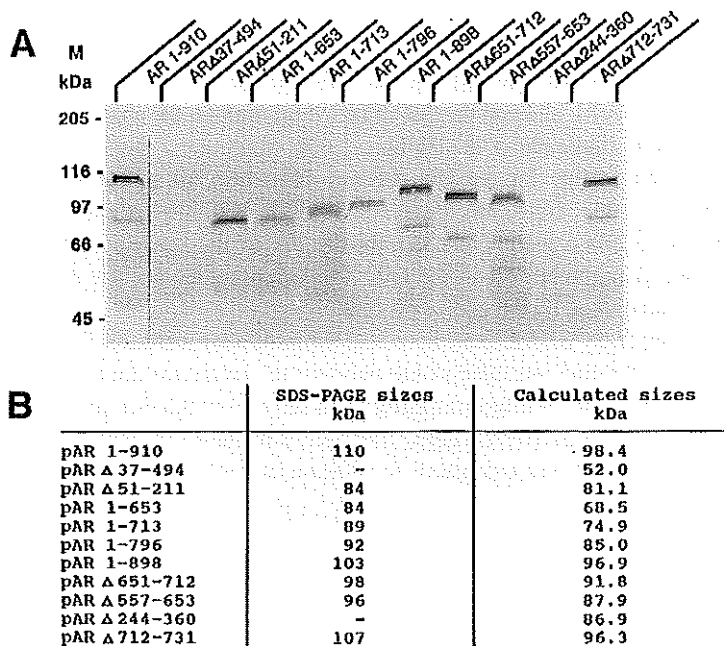


Figure 1 Immunoblot of wild type AR and AR mutants expressed in COS-1 cells.

A: AR proteins, synthesized in COS-1 cells were immunoprecipitated with the monoclonal antibody F39.4.1, separated by SDS-polyacrylamide gel electrophoresis, blotted and immunostained with the polyclonal antibody Sp061. Molecular weight markers were run on a parallel lane (M).

B: Calculated and SDS-PAGE determined molecular masses of the AR deletion mutants.

Calculated molecular masses are based on the amino acid residue content of the different AR mutants.

### Hormone binding of mutant androgen receptors

The ability of wild-type and AR mutants expressed in COS-1 cells to bind specifically the synthetic androgen R1881 was examined next. Deletion of amino acid residues 651-712 or

712-731 in the C-terminal domain of the AR abolished hormone binding completely (Figure 2). In addition, truncated receptors lacking the last 257, 197, 114 or 12 amino acid residues were unable to bind R1881. Deletions in the N-terminal and DNA-binding domain (AR $\Delta$ 557-653, AR $\Delta$ 37-494, AR $\Delta$ 51-211 and AR $\Delta$ 244-360) did not affect hormone binding. These results are in agreement with mutational analysis studies performed with GR, PR and ER, indicating that also in the AR the C-terminal part is the steroid binding domain (Giguere et al. 1986; Gronemeyer et al. 1987; Kumar et al. 1987; Dobson et al. 1989).

### Domains required for transcriptional activation

Cotransfection in HeLa cells of the AR expression plasmids and a plasmid containing an androgen-regulated reporter gene was performed, to study the ability of different AR mutants to induce transcription. The reporter plasmid contains two PRE/GREs in front of the thymidine kinase (tk) promoter linked to the chloramphenicol acetyltransferase (CAT) gene (pG29GtkCAT) (Schüle et al. 1988). Cotransfection of pAR 1-910 together with pG29GtkCAT in the presence of 10 nM R1881, resulted in a 50-fold induction of CAT expression as compared to the CAT activity of the uninduced AR 1-910 (Figure 2). C-terminal deletions of either the last 12 or the last 114 amino acid residues inactivated the receptor.

Deletion of the complete steroid binding domain (approximately 250 amino acid residues) resulted in a constitutively active AR protein. AR 1-653 was found capable to induce transcription, also in the absence of hormone to approximately the same level as the wild type AR in the presence of hormone. AR 1-713 is also constitutively active, but its level of induction of CAT activity was approximately 40% of the wild type AR activity. AR mutants with small deletions in the C-terminal domain (e.g. residues 651-712 and 712-731) which abolish R1881 binding, showed no measurable CAT induction capacity. The requirement of the DNA binding domain for transcription activation is illustrated by mutant AR $\Delta$ 557-653, which lacks this domain and was inactive. The N-terminal domain, which is completely encoded by exon 1 sequences, is of crucial importance for transcriptional activation. A mutant AR with a large deletion in the N-terminal domain (AR $\Delta$ 37-494) was still able to bind hormone but did not induce CAT activity. Mutant AR $\Delta$ 51-211 showed a strongly decreased CAT induction capacity, which was only 5% of the wild type AR level, while AR $\Delta$ 244-360 was still quite capable of inducing CAT activity (approximately 75%) (Figure 2).

### Subcellular localization

To define the mechanism and the domains involved in the nuclear localization, the different AR mutants were transiently expressed in COS and HeLa cells. After transfection, the cells were fixed and immunostained using the monoclonal antibody F39.4.1. The number of stained transfected COS-1 cells corresponded to a transfection efficiency of 5-15%. The transfection efficiency of the HeLa cells was in the range of 2-5%.

## Appendix paper I

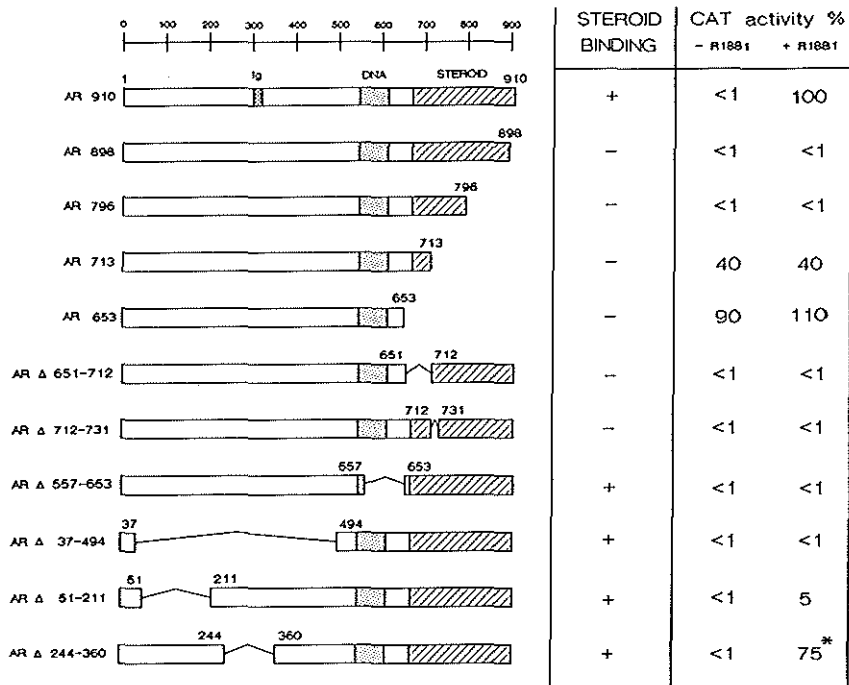


Figure 2 Steroid binding and CAT induction capacity of human AR deletion mutants.

AR proteins were expressed in COS-1 cells and hormone binding was determined by whole cell labeling with  $^3\text{H}$  R1881. Transcriptional activity was examined by cotransfection of AR expression plasmids and a GREtkCAT reporter plasmid. CAT activity was determined from cell lysates of transfected HeLa cells cultured, in the absence (- R1881) or presence (+ R1881) of 10 nM R1881. Activities were corrected for the GREtkCAT background and are presented as percentage relative to the wild type AR. Ig represents the position of the synthetic peptide used for the generation of the Sp061- and F39.4.1 antibodies. In the original manuscript, the CAT-activity of AR $\Delta$ 244-360 (\*) was reported to be <1. Reexamination of this construct revealed an insertion at codon position 244 resulting in a frame-shift and a premature stopcodon. The specific  $^3\text{H}$ -R1881 binding observed after expression of this AR mutant was due to alternative translational starts. As is shown in the figure, the correct AR $\Delta$ 244-360 mutant was capable of activating transcription to approximately 75% as compared to the wild type AR. A rectification has been submitted to the journal.

The subcellular localization of the wild type AR in the absence of hormone was mainly nuclear in transfected COS-1 cells. Nucleoli were not stained. Addition of 10 nM R1881 directed the remaining cytoplasmic AR to the nucleus, suggesting that in addition to the constitutive import of AR into the nucleus, a hormone dependent nuclear uptake process can be active (Figure 3A and 3B).

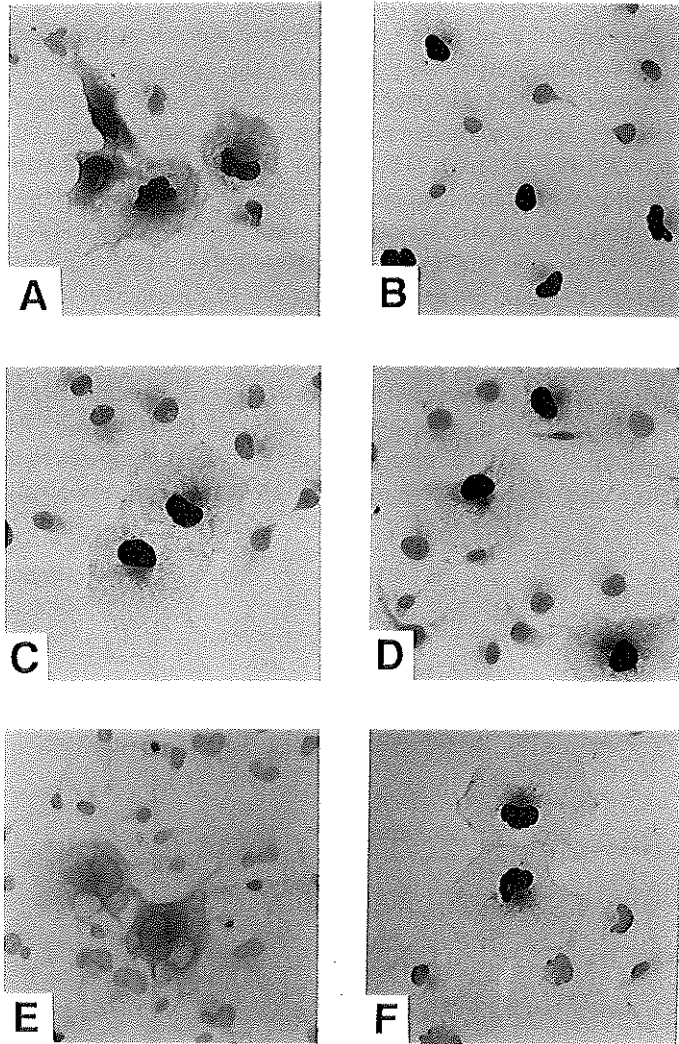


Figure 3 Immunocytochemical localization of the wild type and human AR mutants. COS-1 cells were plated directly on microscope slides and transfected using the calcium phosphate method. One day after the addition of the DNA precipitate, the cells were washed and incubated for 24 h with (right panels B, D and F) or without (left panels A, C and E) 10 nM R1881. Subsequently, the cells were fixed and stained using the peroxidase procedure as described in Materials and Methods. The different immuno-graphs show: wild type AR (A and B); constitutively active AR 1-653 (C and D); AR $\Delta$ 557-653 (E and F).

Deletion of a large part of the DNA binding domain together with part of the hinge region (AR $\Delta$ 557-653) resulted in an exclusively cytoplasmic localization of the AR protein in the absence of hormone (Figure 3E). The deleted region contains a nuclear localization signal, similar to the SV40 large T antigen nuclear entering signal (human AR amino acid residues 620-

## Appendix paper I

626: arg-lys-leu-lys-lys-leu-gly). This signal is highly conserved among steroid hormone receptors (Guiochon-Mantel et al. 1989). Interestingly, addition of hormone induced nuclear import of AR $\Delta$ 557-653 into the nucleus (Figure 3F). This confirms the observation made with the wild type receptor that the androgen receptor is capable of entering the nucleus by a hormone dependent nuclear translocation mechanism. AR mutants lacking small parts of the C-terminal domain (AR 1-898, AR $\Delta$ 651-712 and AR $\Delta$ 712-731) were located mainly in the nucleus (Figure 4). These proteins are unable to bind hormone, and addition of ligand therefore had no effect on the distribution of the AR over nucleus and cytoplasm. The constitutively active AR 1-653, which lacks the steroid binding domain, showed almost no cytoplasmic localization in transfected COS-1 cells (Figure 3C and 3D). A mutant AR, essentially identical to AR 1-653, but lacking a region of 30 amino acid residues containing the SV40 large T antigen-like NLS, was cytoplasmic in the absence and presence of hormone (data not shown). This indicates that the nuclear import of AR 1-653 is controlled by the NLS rather than that its reduced size results in a protein, small enough to pass unhindered through nuclear pores. A 51-211 deletion in the N-terminal domain did not affect the subcellular distribution, as compared to the wild type AR. Subcellular localization experiments performed with HeLa cells gave essentially identical results (data not shown).

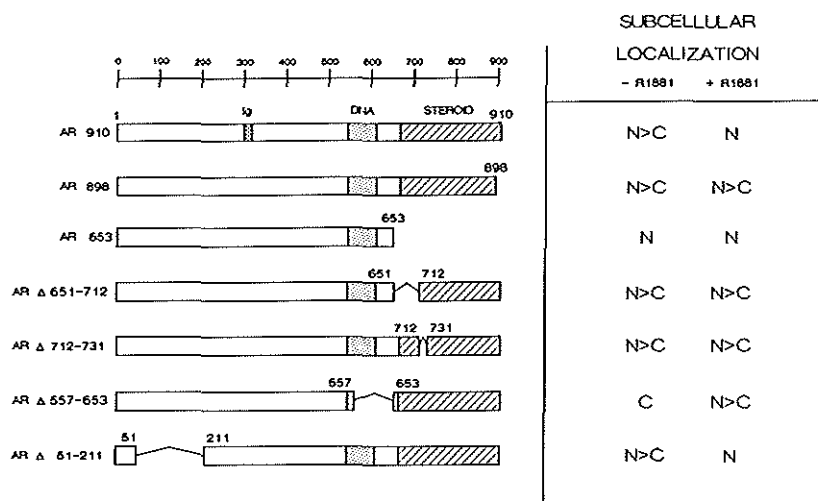


Figure 4 Subcellular localization of human AR deletion mutants. Nuclear (N) and cytoplasmic (C) localization of AR proteins in transfected COS-1 and HeLa cells in the absence (- R1881) or presence (+ R1881) of 10 nM R1881. Ig represents the position of the synthetic peptide used for the generation of the Sp061- and F39.4.1 antibodies.

The medium in which the COS-1 and HeLa cells were cultured and transfected contained 5% charcoal-stripped fetal calf serum and 45  $\mu$ M phenol red. It has been shown that high concentrations of phenol red can direct the GR to the nucleus (Picard & Yamamoto 1987) and it was determined therefore whether phenol red could interfere with the subcellular distribution of the AR. Control experiments were performed also in the complete absence of serum. COS-1



cells expressing the wild type AR or AR $\Delta$ 557-653 were grown, one day prior to fixation, either on serum- and phenol red-free medium, or on phenol red-free medium containing 5% charcoal stripped serum. Using these media no significantly different subcellular distribution of the AR was observed, as compared to the results of all other experiments which were obtained in medium containing charcoal stripped serum and phenol red (data not shown). The finding that phenol red does not influence the nuclear entering of the AR is in agreement with the transcriptional activation experiments where CAT activity of the wild type AR, estimated using phenol red containing medium, was not different from the control (only CAT reporter construct transfected) in the absence of R1881.

## DISCUSSION

Mutational analysis of the human AR through expression of AR mutants in COS-1 and HeLa cells has facilitated structural and functional characterization of the AR domains including regions for steroid binding, transcriptional activation and subcellular localization.

The wild type AR expressed in COS-1 cells migrated as a 110-112 kDa doublet when analyzed by SDS-polyacrylamide gel electrophoresis. This doublet represents AR isoforms, produced through post-translational phosphorylation (Kuiper et al. 1991). The B-form of the PR migrates as a triplet when run on SDS-polyacrylamide gel. These PR-B isoforms are also caused by differential phosphorylation (Sheridan et al. 1989). All the investigated AR deletion mutants migrated as doublets, with one exception; AR $\Delta$ 51-211 migrated as a single protein band with an apparent molecular weight of 84 kDa. It is tempting to speculate that in the deleted part of AR $\Delta$ 51-211, one or more serine, threonine or tyrosine residues are located, which are phosphorylated in the wild type AR, giving rise to the slower migrating band of the doublet. Two possible phosphorylation sites for proline directed kinase that conform to the Ser-Pro motif are located in this region (Ser-Pro at sites 80 and 93) (Faber et al. 1989). In the chicken PR, two serine residues in the N-terminal domain with this Ser-Pro motif can be phosphorylated (Denner et al. 1990b).

Besides being expressed as a single protein band, AR $\Delta$ 51-211 migrated with the same mobility as AR 1-653 and even faster than AR 1-713 when analyzed by SDS-polyacrylamide gel electrophoresis. This is unexpected since the calculated molecular mass of AR $\Delta$ 51-211 is larger than those of AR 1-653 and AR 1-713. Possibly AR $\Delta$ 51-211 has undergone less or no post-translational modification or the constitutively active AR 1-653 and AR 1-713 are subject to extra modification.

It has been shown for other steroid hormone receptors that the hormone binding domain is restricted to the approximate last 250 amino acid residues (Kumar et al. 1987; Dobson et al. 1989; Simons et al. 1989). Unlike the ER, where truncation of 61 amino acid residues does not affect hormone binding (Lees et al. 1989), the truncation of 12 amino acid residues in the AR resulted in complete loss of binding of R1881. Deletion of a small region of 19 amino acid residues (AR $\Delta$ 712-731) within the steroid binding domain also completely abolished specific R1881 binding. This small deleted region is very conserved among steroid receptors and is supposed to be a site of receptor-hsp90 interaction (Kost et al. 1989; Pratt et al. 1989; Howard et al. 1990).

Complete deletion of the steroid binding domain resulted in an AR that was constitutively

## Appendix paper I

active. Truncation of the C-terminal 257 amino acid residues of the AR yielded a protein capable of complete transcriptional activation of the CAT reporter gene in the absence of hormone. Deletion of the C-terminal 197 amino acid residues resulted in a constitutively active AR with approximately 40% activity as compared to wild type AR. These results agree with earlier published reports on the AR (Marcelli et al. 1990a; Rundlett et al. 1990) and GR (Godowski et al. 1987b; Miesfeld et al. 1987; Hollenberg & Evans 1988). Lower CAT induction levels have been found by Simental et al. (Simental et al. 1991). Their AR mutant, comparable to AR 1-653, was capable to induce transcription of a MMTV-CAT reporter gene to only 10% of the wild type AR capacity. These variations in CAT induction could be caused by differences in AR protein expression levels or by the fact that different reporter gene promoters were used. From the results we can conclude that the C-terminal steroid binding part of the AR acts as a regulatory domain, that prevents transcriptional activation in the absence of hormone.

The N-terminal domain of the human AR was found to be essential for transcriptional activation of the CAT reporter gene. Deletion of a large part of this domain (ARA37-494) completely inactivate the AR. Structural comparison of the N-terminal domain of steroid receptors showed a relatively high amount of acidic amino acid residues (Danielsen et al. 1987; Faber et al. 1989). Deletion of this acidic region in the human GR (amino acid residues 77-262, referred to as transcription activation unit [TAU-1]) strongly reduced the transcriptional activation capacity (Hollenberg et al. 1987; Hollenberg & Evans 1988). Acidic domains of variable sizes are common to several transcription regulating proteins, and at least for two yeast factors the function of these domains in transcription regulation is evident (Ptashne 1988). AR deletion mutant ARA51-211 lacks a portion of the acidic region and has a strongly reduced transcriptional activation capacity. It can be speculated that the acidic region is of high importance for transcriptional activation; however, other explanations are possible. As discussed above ARA51-211 might be different post-translationally modified. Furthermore, the deleted region in this AR mutant contains three homopolymeric glutamine stretches. Several homopolymeric amino acid stretches (viz. poly-glutamine, a poly-glycine, a poly-alanine and a poly-proline stretch) are present in the N-terminal domain of the human AR (Faber et al. 1989). The function of these amino acid repeats is at present unknown. Most proteins in which similar repeats have been identified are involved in developmental control and/or regulation of gene expression (Haynes et al. 1987; Burglin 1988). Interestingly, a large poly-glutamine stretch is also present in human TFIID (Peterson et al. 1990).

The results of the subcellular localization experiments are in agreement with the *in vivo* immunohistochemistry studies performed with rat and human tissues (Husmann et al. 1990; Sar et al. 1990). The AR is nuclear in the presence but also in the absence of androgens, although there is a significant cytoplasmic localization when hormone is absent. The nuclear localization is mainly determined by amino acid residues in the region 557-653. In the absence of hormone, ARA557-653 is almost exclusively cytoplasmic. The deleted region contains a nuclear localization signal similar to the SV40 large T antigen NLS which is conserved in all steroid hormone receptors (Guiochon-Mantel et al. 1989). It has been shown that this SV40 large T antigen-like NLS is responsible for the nuclear entering of the PR (Guiochon-Mantel et al. 1989). Interestingly, the addition of R1881 directed the cytoplasmic ARA557-653 to the nucleus, indicating the existence of a second, hormone-dependent nuclear import mechanism. Similar results have been obtained with PR deletion mutants lacking the SV40 large T antigen-like NLS

(Guiochon-Mantel et al. 1989). Picard & Yamamoto (1987) showed that the subcellular distribution of the GR is regulated by such a hormone-dependent nuclear entering mechanism. The molecular nature of the hormone-dependent nuclear localization mechanism is unclear, and it is not known where the putative signal sequence is situated in the receptor molecule. Studies with chimeric  $\beta$ -galactosidase/GR proteins have shown that the steroid binding domain of the GR is capable of directing the cytoplasmic  $\beta$ -galactosidase protein to the nucleus in a hormone-dependent fashion (Picard & Yamamoto 1987). The present investigation indicated that a similar NLS may be present in the AR and becomes activated upon hormone binding. Experiments are in progress to substantiate further the activity and localization of the second NLS in the AR.

In a recent study, Simental and co-workers (1991) presented evidence that the wild type human AR is cytoplasmic in the absence and nuclear in the presence of hormone, which is in contradiction with the observations presented here. An explanation for this discrepancy has not yet been found, but we have ruled out a possible effect of the composition of the culture medium.

In conclusion, the data presented show that: 1) the functional domain structure of the AR is comparable to that of the GR and PR, 2) at least one part of the N-terminal domain is important for the trans-activation function of the AR, 3) the doublet appearance of the AR in SDS-polyacrylamide gel electrophoresis might be due to post-translational phosphorylation of amino acid residues in the N-terminal region and 4) the AR contains at least two nuclear localization signals, of which one is hormone-dependent.

## MATERIALS AND METHODS

### Cell culture and transfection

COS-1 and HeLa cells were maintained in Eagle's minimum, essential medium supplemented with 5% fetal calf serum (FCS), non-essential amino acids and antibiotics. COS-1 and HeLa cells were transfected using the calcium phosphate precipitation method (Chen & Okayama 1987).

### AR expression vectors

A human AR cDNA expression vector (pAR 1-910) was constructed using the SV40 early promoter and the rabbit  $\beta$ -globin polyadenylation signal (Brinkmann et al. 1989). AR 1-898, AR 1-796, AR 1-713 and AR 1-653 are C-terminal truncation mutants made by introduction of a stop codon in the 3' part of the AR cDNA. For that purpose, a Xba I-linker (Promega, Madison, WI, U.S.A.) that contains an in frame stop codon, was ligated in the blunt ended Xho II, EcoR I, Stu I and Tth III sites (Figure 5). All deletion mutants were constructed by removal of internal fragments using two different restriction enzymes (Figure 5). After digestion of pAR 1-910, the linear fragment containing the vector was isolated, blunt ended and ligated. The combination of the restriction enzymes was chosen in such a way that ligation of the blunt ended plasmid would result in an in frame deletion. In cases where more than one restriction site of an enzyme was located in the human AR cDNA, partial digestions were performed (Figure 5). All deletion mutants were sequenced to verify the correct reading frame.

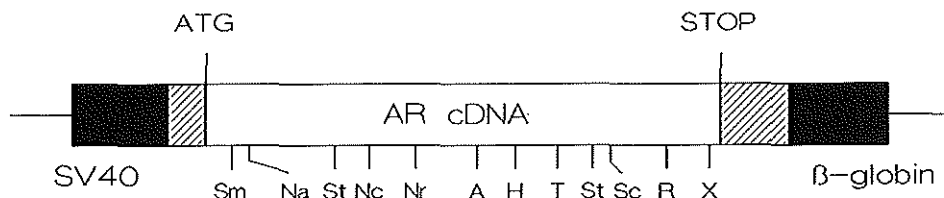


Figure 5 Schematic representation of the human AR expression vector AR 1-910. Human AR cDNA was inserted into an eukaryotic expression vector containing the SV40 early promoter and the  $\beta$ -globin polyadenylation signal (Brinkmann et al. 1989). Shaded areas represent 5' and 3' untranslated regions of the AR cDNA. Restriction sites used to construct the different deletion mutants are shown. Sm=Sma I; Na=Nar I; St=Stu I; Nc=Nco I; Nr=Nru I; A=Asp718; H=Hind III; T=Tth III; Sc=Sca I; R=EcoRI; X=Xho I.

### Western blot analysis

Transfected COS-1 cells were lysed in 40 mM Tris.HCl pH 7.0, 1 mM EDTA, 4% glycerol, 10 mM dithiothreitol and 2% SDS. The androgen receptor proteins were immunoprecipitated in whole cell lysates from approximately  $5 \cdot 10^6$  COS-1 cells, using monoclonal antibody F39.4.1 coupled to goat-anti-mouse agarose beads (Zegers et al. 1991). Western blotting and immunostaining with polyclonal antibody Sp061 (diluted 1:1000) were performed as described previously (van Laar et al. 1989).

### Hormone binding- and CAT-assays

COS-1 cells were grown and transfected in medium containing 5% dextran/charcoal-treated fetal calf serum (ccFCS). For hormone binding studies, cells were incubated with 5 nM [ $^3$ H]R1881 (methyltrienolone, S.A. 87 Ci/mmol, New England Nuclear, Dupont de Nemours, F.R.G.) for 2 h at 37°C, in the presence or absence of a 200-fold molar excess of unlabeled R1881. After washing the cells 4 times with 1% BSA in PBS, the cells were lysed by freezing and thawing 5 times. Steroid-binding protein from whole cell lysates were precipitated with protamine sulfate as described (Mulder et al. 1981), and the precipitates were counted for radioactivity. Specific binding represents the difference between counts in the absence and presence of excess unlabeled steroid, respectively. The CAT assays were performed essentially as described previously (Seed & Sheen 1988).

## Immunocytochemistry

COS-1 or HeLa cells were cultured directly on microscope glass slides ( $2 \cdot 10^5$  cells/slide) and transfected using the calcium phosphate method ( $2 \mu\text{g}$  AR plasmid and  $2 \mu\text{g}$  pTZ19 carrier plasmid/slide). The COS-1 cells were plated and transfected in the presence of 5% ccFCS. 24 h after the addition of the DNA precipitate, the cells were washed and fresh medium was added. Cells were incubated with or without 10 nM R1881 for an additional 24 h, and subsequently washed with PBS and fixed in phosphate-buffered 10% formalin during 10 min. The microscope slides were washed in PBS pH 7.4 and attached cells were permeabilized in chilled ( $-20^\circ\text{C}$ ) methanol for 4 min. and chilled ( $-20^\circ\text{C}$ ) acetone for 2 min. After rehydration in PBS, the slides were incubated for 15 minutes in PBS containing 10% FCS, and then incubated with 1:1000 diluted monoclonal antibody F39.4.1 for 20 h, at  $4^\circ\text{C}$ . Excess antibody was removed by washing, followed by a 30 min. incubation of horse-radish peroxidase conjugated rabbit anti-mouse immunoglobulin (DAKO, P161), diluted 1:100 in PBS containing 4% serum. After several PBS washes, the slides were incubated with diaminobenzidine and the reaction was stopped with water. The cells were counterstained with Mayers haematoxylin for 20 seconds.

## ACKNOWLEDGMENTS

We thank Dr. R. Renkawitz for the pG29GtkCAT construct, Dr.J.A. Grootegeod for helpful discussions and for reading the manuscript and H.C.J. van Rooij for preparations of some of the AR mutants. This investigation was supported by the Netherlands Organization for Scientific Research (NWO) through GB-MW (Medical Sciences)

## Appendix paper I

**NUCLEAR IMPORT OF THE HUMAN ANDROGEN RECEPTOR**

Guido Jenster<sup>1</sup>, Jan Trapman<sup>2</sup> and Albert O. Brinkmann<sup>1</sup>

Departments of Endocrinology & Reproduction<sup>1</sup> and Pathology<sup>2</sup>,  
Erasmus University Rotterdam, Rotterdam, The Netherlands

Biochem J (1993) 293:761-768

### ABSTRACT

Nuclear import of the human androgen receptor (AR) was investigated by immunocytochemical analysis of AR deletion- and substitution mutants which were transiently expressed in COS-1 cells. The signal responsible for nuclear import is encoded by amino acid residues 608-625 and is functionally similar to the bipartite nucleoplasmin nuclear localization signal (NLS). Although the subcellular distribution of androgen receptors mutated in the DNA binding domain was unchanged as compared with the wild type AR, in the presence of ligand these mutations resulted in cluster formation of part of the receptor population. Depending on the presence or absence of the bipartite NLS, clusters were formed in the nucleus or cytoplasm, respectively.

Expression of wild type AR in different cell lines revealed a cell line specific subcellular distribution of unliganded receptor. The AR was predominantly nuclear when expressed in HeLa cells, whereas mainly a cytoplasmic staining was observed when expressed in COS-1 cells. In the presence of hormone, the AR was located in the nucleus, independent of the cell line expressing the receptor.

Antiandrogens and various steroid hormones induced the nuclear localization of the wild type AR in a dose-dependent way without activating transcription of an androgen-regulated reporter gene. This indicates that the inability of the tested compounds to activate transcription is not due to inhibited nuclear import.

### INTRODUCTION

The androgen receptor (AR) is a ligand-dependent transcription factor that belongs to the large family of steroid hormone receptors (Evans 1988; Green & Chambon 1988). Activated by their ligand, these receptors interact with specific DNA sequences located in the 5' flanking regions of target genes, resulting in the up- or down-regulation of the expression of these genes. Before exerting their function as transcription factors, these proteins have to pass the nuclear membrane. Two processes for nuclear import of proteins are commonly considered (Nigg et al. 1991; Silver 1991). Proteins either freely diffuse through the nuclear membrane or harbour a nuclear localization signal (NLS) resulting in active transport. For passive diffusion through nuclear pores, proteins should not exceed the apparent exclusion limit of approximately 60 kDa (Peters 1986). Larger proteins, such as the human AR, need specific signals for active nuclear import.

The subcellular localization of the unliganded progesterone- (PR), estradiol- (ER) and glucocorticoid (GR) receptors is still a point of discussion (Perrot-Appianat et al. 1985; Sanchez et al. 1990; Power et al. 1991b; Brink et al. 1992). Contradictory findings have been published for the AR as well. Immunohistochemical studies on the subcellular localization of the AR in normal and castrated rats demonstrated that the AR was located in the nucleus (Husmann et al. 1990; Sar et al. 1990), as is the case in comparable studies on prostate tissue from prostate cancer patients under androgen ablation therapy (van der Kwast et al. 1991). In contrast, the AR transiently overexpressed in COS cells was found to be localized in the cytoplasm (Simental et al. 1991) or distributed over cytoplasm and nucleus (Jenster et al. 1991) in the absence of ligand, and was found to be exclusively nuclear only in the presence of androgen.

One of the best characterized NLSs is the nuclear targeting signal of simian virus 40



(SV40) large T antigen, which consists of seven amino acid residues, of which five are positively charged (Kalderon et al. 1984a,b; Lanford & Butel 1984). A similar basic signal is also present in the hinge region of all steroid receptors and its importance for nuclear import has been determined for the PR, ER, GR and AR (Picard & Yamamoto 1987; Guiochon-Mantel et al. 1989; Simental et al. 1991; Ylikomi et al. 1992). Deletion of this SV40 large T antigen-like NLS, however, did not result in the exclusive cytoplasmic localization of the PR and AR in the presence of hormone, indicating the complexity of the NLS and/or the presence of additional signals.

In the nuclear protein nucleoplasmin, an NLS has been identified that consists of 16 amino acid residues, which includes a basic lysine sequence similar to the SV40 large T antigen NLS. Two basic amino acid residues (Lys, Arg) followed by a spacer of ten residues precede the lysine stretch, forming in this way a bipartite NLS (Robbins et al. 1991). This conserved, bipartite NLS is present in many other nuclear proteins, including steroid receptors (Dingwall & Laskey 1991). To identify the position of possible NLSs and to define in detail the contribution of the bipartite nucleoplasmin-like NLS in AR nuclear import, we have constructed a series of mutants of the human AR and have immunocytochemically examined the subcellular localization of these proteins. In addition, the effect of antiandrogens and various steroid hormones on nuclear import of the AR was investigated.

## RESULTS

### Cell line specific subcellular localization of the human AR

One of the possible explanations for the contradictory reports on the subcellular localization of unliganded AR may be the use of different cell lines in the various studies. To investigate the subcellular localization in different cell lines, COS-1, HeLa, CV-1 and CHO cells were transiently transfected with the pSVAR0 expression vector coding for the wild type human AR (Brinkmann et al. 1989). As shown in Figure 1, in the absence of ligand, the AR in COS-1 cells was predominantly cytoplasmic, whereas in CV-1 and CHO cells, the AR was more evenly distributed over cytoplasm and nucleus. In HeLa cells, unliganded AR was predominantly nuclear, indicating that the subcellular distribution is indeed cell line specific. Addition of 10 nM of the synthetic androgen R1881 resulted in the nuclear localization of the AR in all cell lines (Figure 1).

### Subcellular localization of AR deletion mutants

To identify the position of possible NLSs, AR deletion mutants were transiently expressed in COS-1 cells and their subcellular localization was examined. The results of these experiments are summarized in Figure 2. The deletion of part of the steroid binding domain (AR17) resulted in a receptor that was mainly located in the cytoplasm. This protein is unable to bind ligand and therefore the addition of 10 nM R1881, had no effect. AR mutants lacking even larger parts of the C-terminal domain (AR13 and AR14) were exclusively nuclear, indicating that a region present in AR17 but deleted in AR14 (amino acid residues 713-796) is responsible for the cytoplasmic retention. Both AR13 and AR14 show constitutive transactivati-

on activity (Jenster et al. 1991). Deletion up to amino acid residue 618 results in a truncated receptor (AR4) that was mainly cytoplasmic.

This observation suggests the presence of an NLS between residues 618 and 653. An AR mutant, lacking both the steroid binding domain and the DNA binding domain (AR12), was evenly distributed over the cytoplasm and nucleus. This truncated protein, with a molecular mass of approximately 60 kDa, is probably small enough to passively diffuse through the nuclear pores.

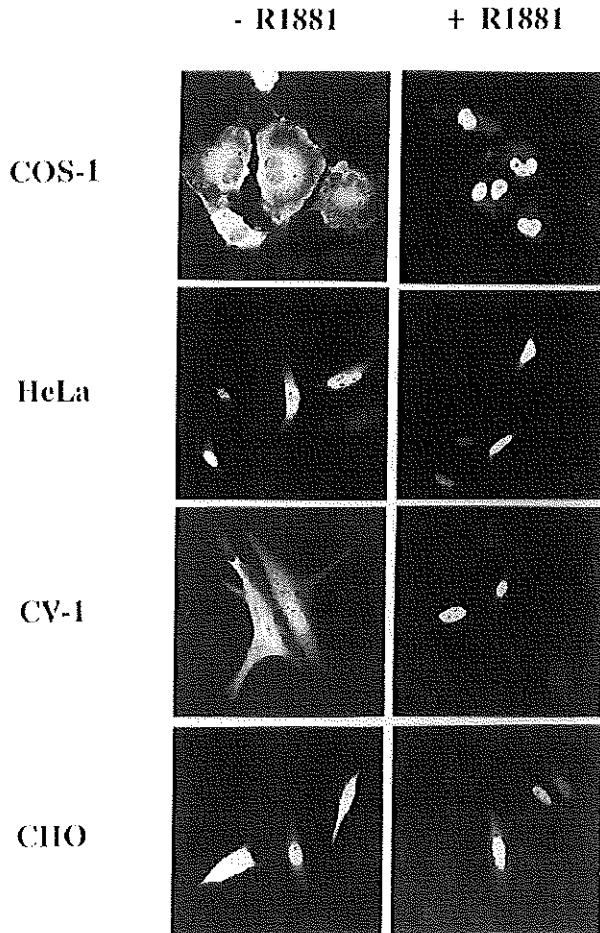


Figure 1 Immunocytochemical localization of the wild type AR in different cell lines. COS-1, HeLa, CV-1 and CHO cells were plated onto microscope slides and transfected using the calcium phosphate method. One day after the addition of the DNA precipitate, the cells were washed and incubated for 24 hours either with 10 nM R1881 (+R1881) or without ligand (-R1881). Subsequently, the cells were fixed and stained using the F39.4.1 monoclonal antibody and FITC-coupled goat anti-mouse antibodies as described in the Materials and methods section. Magnification is approximately 250x.

## Nuclear import of the AR

The second indication for the position of an NLS came from the subcellular localization of AR21, which lacks the amino acid residues 557 to 653 (including part of the DNA binding domain). We previously reported that, in the absence of ligand, AR21 was exclusively cytoplasmic; the addition of hormone directed a portion of the receptor population to the nucleus (Figure 2) (Jenster et al. 1991).

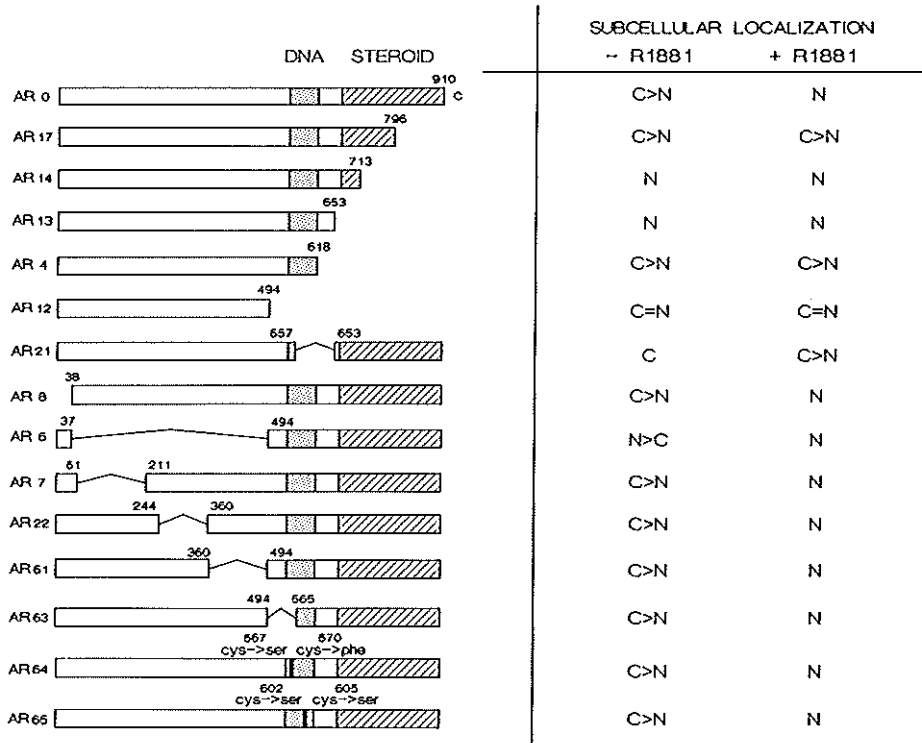


Figure 2 Subcellular localization of human AR mutants. Nuclear (N) and cytoplasmic (C) localization of AR proteins expressed in COS-1 cells in the absence (-R1881) or presence (+R1881) of 10 nM R1881. The receptor mutants were visualized using the F39.4.1 antibody, except for AR22, for which the Sp197 antibody was used.

Deletions in the N-terminal domain did not affect the AR nuclear import. Unliganded AR6, which lacks almost the entire N-terminal part, had a more nuclear than cytoplasmic distribution. Probably because of its reduced size (approximately 53 kDa), this AR mutant was able to passively diffuse through the nuclear membrane. The subcellular localizations of AR8, AR7, AR22, AR61 and AR63, all of which lack part of the N-terminal domain, were similar to the wild type AR distribution. AR64 and AR65 were also similarly distributed, compared

## Appendix paper II

with the AR0 subcellular distribution (Figure 2). In these receptor mutants, the structure of the first (AR64) or second (AR65) zinc twist is disrupted by the replacement of two of the four cysteine residues that are involved in co-ordination binding of the zinc atoms. This indicates that effective DNA binding is not involved in the subcellular distribution of the AR.

It was observed that AR21, AR63, AR64 and AR65 gave a different staining pattern in comparison with the wild type AR staining. In contrast with the diffuse granular distribution of AR0 in cytoplasm or nucleus, large clusters of AR21 protein were visible in the cytoplasm of most transfected cells in the presence of hormone (Figure 3). Ligand-induced cluster-formation was also observed for part of the receptor population of AR63, AR64 and AR65. The AR63, AR64 and AR65 clusters, however, were mainly located in the nucleus (Figure 3). These findings indicate that disruption of the DNA binding domain results in clustering of the mutated AR and that clustering is independent of the subcellular localization.

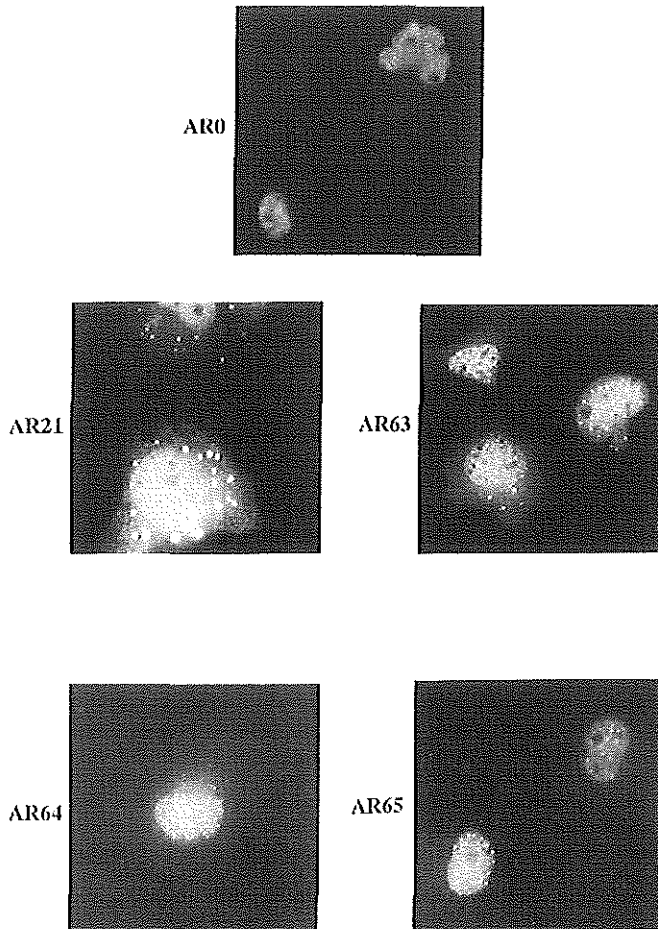


Figure 3 Staining pattern of the wild type AR (AR0) and of AR mutants expressed in COS-1 cells in the presence of 10 nM R1881.

The receptors were visualized using the F39.4.1 antibody. Magnification approximately 630x.

## Definition of an AR NLS

From the studies with the mutants AR4 and AR21, it was concluded that amino acid residues 557 to 653 are necessary for proper nuclear import. This part of the AR harbours at position 620-624 the sequence "RKLKK" which is an SV40 large T antigen NLS-like sequence. This signal can partially account for the nuclear import of the AR, PR and ER (Guiochon-Mantel et al. 1989; Simental et al. 1991; Ylikomi et al. 1992). The same region also shares sequence similarity with the nucleoplasmin NLS, which consists of two basic residues, a spacer of ten amino acid residues and a C-terminal basic domain, similar to the SV40 large T antigen NLS (Table 1). In nucleoplasmin, both basic parts are involved in nuclear targeting (Robbins et al. 1991). The NLS of nucleoplasmin is conserved among steroid receptors. To investigate if this bipartite NLS can account for the total nuclear import of the AR, both basic parts were either separately or simultaneously mutated.

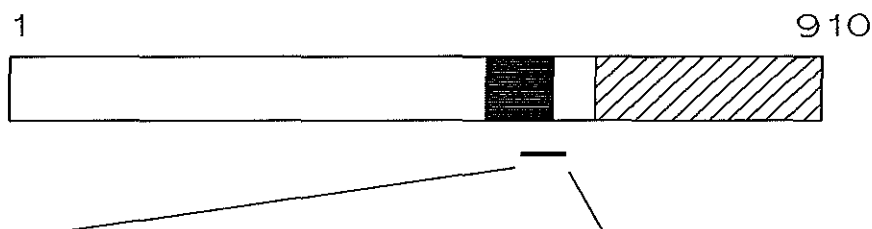
Table 1 Comparison of amino acid sequences of proven or potential nuclear localization signals in SV40 large T antigen, nucleoplasmin and various steroid hormone receptors. The numbers refer to the first amino acid residue of the sequence. Basic amino acid residues thought to be important for the signal are shown in bold capitals.

Protein	Species	Sequence
SV40 large T antigen	SV40	126 p <b>KKKRK</b>
Nucleoplasmin	Xenopus	152 kav <b>KR</b> paatkkagqa <b>KKKK</b> l
Androgen receptor	human	605 cr <b>lRK</b> cyeaagmtlg <b>aRK</b> l <b>KK</b> l
Progesterone receptor	human	622 cr <b>lRK</b> ccqagmvl <b>ggRK</b> l <b>KK</b> f
Glucocorticoid receptor	human	476 cry <b>RK</b> clqagm <b>leaRK</b> t <b>KK</b> k
Mineralocorticoid receptor	human	658 cr <b>lqK</b> clqagm <b>lgaRK</b> s <b>KK</b> l
Estrogen receptor	human	240 cr <b>lRK</b> cyevgmmkg <b>giRK</b> d <b>RR</b> g

As shown in Figures 4a and b, mutation of either basic domain affected nuclear import (AR28.1, AR28.2 and AR28.3). In the presence of ligand, however, part of the receptor population was still nuclear. Simultaneous mutation of both basic domains (AR28.31 and AR28.23) resulted in AR mutants which were cytoplasmic in the absence of ligand, and were predominantly cytoplasmic with only a faint nuclear staining in the presence of 10 nM R1881 (Figure 4a,b). The mutations introduced in the NLS of these AR mutants, in particular the replacement of the Arg-Lys by Gly-Ala in the first basic domain of the NLS, resulted in cluster formation, again indicating the role of the DNA binding domain in receptor clustering (Figure 4b). The same set of NLS mutations were introduced into the constitutively nuclear AR13, which lacks the steroid binding domain. Again, alteration of either one of the basic domains hindered nuclear import. Simultaneous modification of both basic domains in AR13 clearly

## Appendix paper II

affected nuclear transport; however, a small portion of the receptor population was still nuclear. The subcellular localization of these mutants derived from AR13, was investigated both in the absence and in the presence of ligand. As expected, hormone had no effect on the distribution (results not shown).



AR	human androgen receptor mutants		subcellular localization	
	-	+	-	+
0	608	625	C > N	N
28.1	<u>R</u> <u>K</u> c y e a g m t l g a <u>R</u> <u>K</u> <u>L</u> <u>K</u> <u>K</u> <u>L</u> g n	-----DEL-----	C	N > C
28.2	.	-----DEL-----	C	C = N
28.3	g a . . . . .	-----DEL-----	C	N > C
28.31	g a . . . . .	-----DEL-----	C	C*
28.23	g a . . . . .	-----DEL-----	C	C*

Figure 4A Mutational analysis of the nucleoplasm-like nuclear localization signal in the wild type AR. a) The wild type AR, with the NLS shown as a bar. The NLS is given below in single letter code (amino acid residues 608-627). Dots represent unchanged residues, deletions are represented as lines with the word "DEL" and the newly encoded residue is shown in case of amino acid substitutions. AR mutant numbers are on the left and the subcellular localization in the absence (-) or presence (+) of 10 nM R1881 is shown on the right as cytoplasmic (C), nuclear (N) or combinations of cytoplasmic and nuclear staining (C>N, C=N, N>C). AR28.31 and AR28.23 showed only a very faint nuclear staining in the presence of ligand indicated as C\*. Basic residues thought to be important for the function of the signal are in bold underlined capitals.

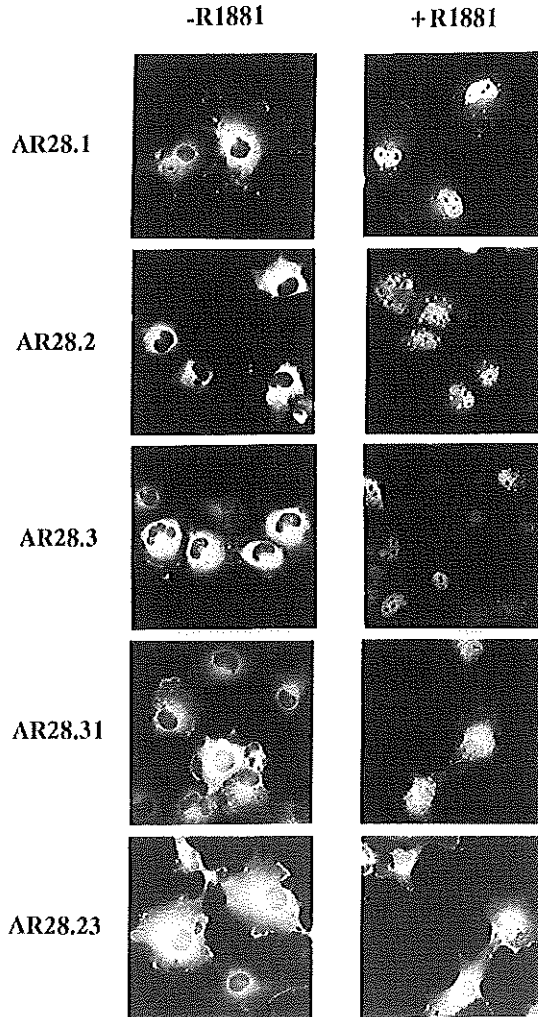


Figure 4B *Mutational analysis of the nucleoplasmin-like nuclear localization signal in the wild type AR. b) Immunofluorescence images of the AR mutants expressed in COS-1 cells in the absence (-R1881) or presence (+R1881) of 10 nM R1881. Receptor mutants were visualized using the F39.4.1 monoclonal antibody.*

### Hormone specificity of AR nuclear import

Next, the ability of antiandrogens and various steroids to induce nuclear import of the AR was investigated. COS-1 cells expressing the wild type AR were incubated for 2 hours with different concentrations (1000, 100, 10, 1, and 0.1 nM) of R5020 (promegestone), estradiol (E2), and dexamethasone or the antiandrogens hydroxyflutamide, Casodex (ICI 176334), and cyproterone acetate (CA). All compounds were capable of directing the AR to the nucleus,

although concentrations of more than 100 nM hydroxyflutamide, Casodex or dexamethasone and more than 1 nM cyproterone acetate, R5020 or estradiol were necessary (Figure 5). None of the tested compounds, however, induced transcriptional activation of an androgen regulated reporter gene (GREtkCAT) at concentrations that already resulted in a shift of the AR to the nucleus (Figure 6). This indicates that the inability of the tested compounds to induce transactivation is not due to the impairment of nuclear import.

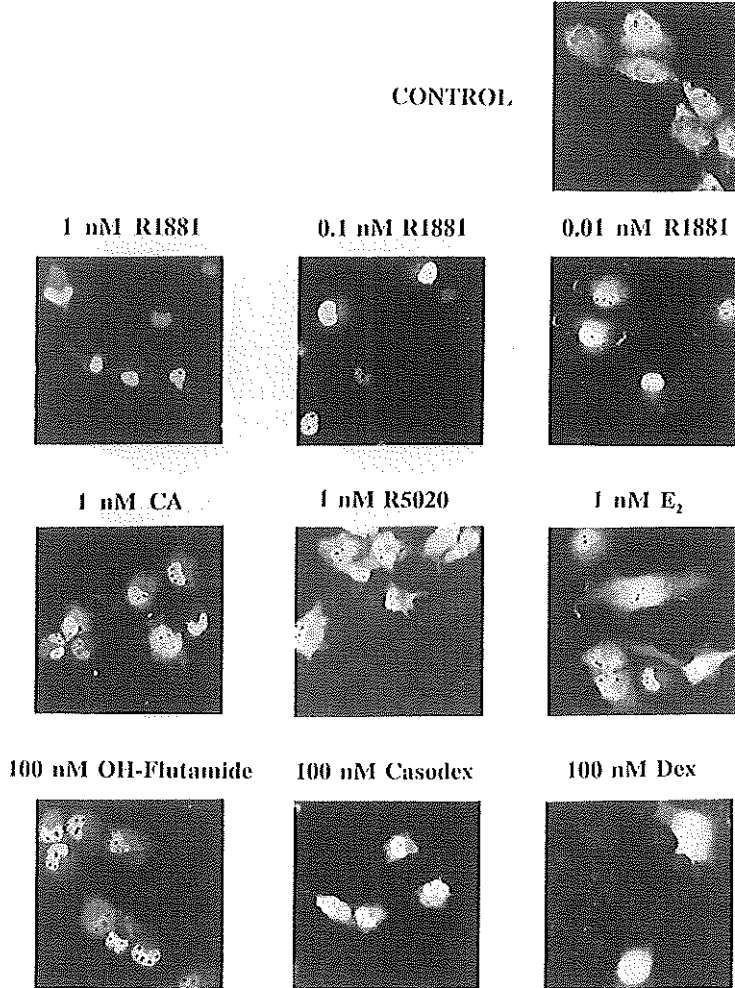


Figure 5 Subcellular localization of the wild type AR in the absence (control) or presence of various steroid hormones and antiandrogens.

*COS-1* cells that had been transfected with pSVAR0, were incubated with different concentrations from 0.01 to 1000 nM of steroids and antiandrogens for 2 h before to fixation. Immunofluorescence images are shown of the lowest concentration of steroid or antiandrogens that resulted in nuclear retention. The wild type AR was visualized using the F39.4.1 antibody. Abbreviations used: CA, cyproterone acetate; E<sub>2</sub>, estradiol, OH-flu, hydroxyflutamide; dex, dexamethasone.



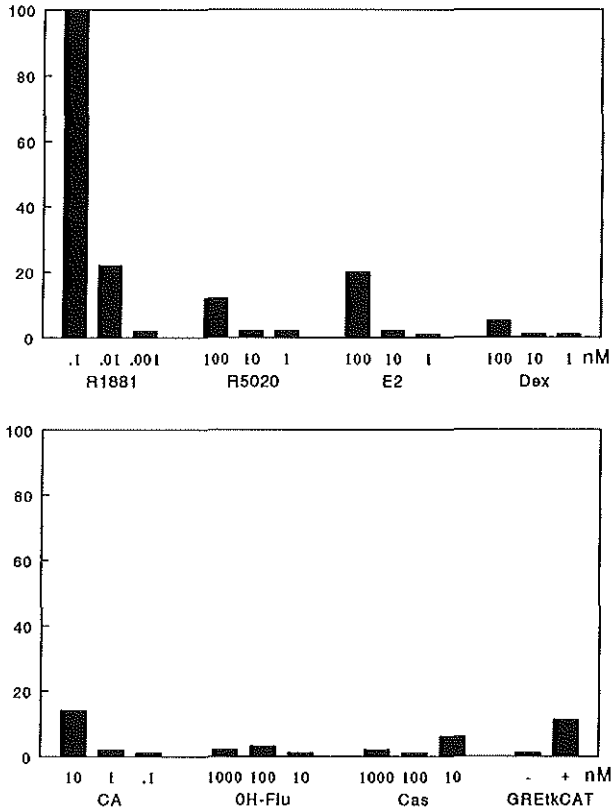


Figure 6 Transcription activation of the wild type AR induced by various steroid hormones and antiandrogens.

CAT activity was determined in COS-1 cells transfected with pSVAR0 and the pG29GtkCAT reporter plasmid (Schüle et al. 1988). One day before cell lysis, cells were incubated with different concentrations of R1881, R5020, estradiol (E2), dexamethasone (Dex), cyproterone acetate (CA), hydroxyflutamide (OH-Flu) and Casodex (Cas) as indicated. Transfection of only the reporter construct is shown as GREtkCAT in the absence of hormone (-) or in the presence of 0.1 nM R1881 (+). Activities were corrected for the assay background that was determined by mock transfection and are presented as a percentage relative to that of 0.1 nM R1881 induction.

## DISCUSSION

Nuclear localization sequences are essential for the transport of large proteins, such as the AR, into the nucleus (Nigg et al. 1991; Silver 1991). In previous publications on the nuclear import of the GR, PR and AR, a candidate signal similar to the SV40 large T antigen NLS has been described (Picard & Yamamoto 1987; Guiochon-Mantel et al. 1989; Simental et al. 1991). Deletion of this signal, however, does not always prevent nuclear import, indicating the complexity of the signal and/or the presence of additional signals. In the present study we show that the nucleoplasmin-like NLS, partially overlapping the SV40 large T antigen-like NLS, can account for nuclear import of the human AR. This nucleoplasmin NLS, composed of two basic

parts separated by a ten amino acid residue spacer, is highly conserved among steroid receptors and appears to be present in many nuclear proteins (Dingwall & Laskey 1991; Robbins et al. 1991).

Mutational analysis of the bipartite nucleoplasmin-like NLS, revealed that both basic parts contribute to nuclear targeting of the AR. Substitution of two amino acid residues in the first basic half of the NLS or deletion of the second basic domain clearly affected nuclear import, but did not result in the exclusive cytoplasmic localization of these AR mutants. Mutation of both NLS parts resulted in an AR mutant which was unable to enter the nucleus; only a very faint staining of the nucleus was observed. For the glucocorticoid receptor (GR) it has been shown that the steroid binding domain is also capable of directing a  $\beta$ -galactosidase fusion protein to the nucleus (Picard & Yamamoto 1987), indicating the presence of an additional NLS in this domain. There is no evidence for additional NLSs in the AR apart from the faint nuclear staining of AR mutants with the nucleoplasmin-like NLS deleted (AR28.23 and AR28.31) that suggest such a signal.

Recently, it has been shown that the nuclear PR becomes cytoplasmic under energy depleting conditions, demonstrating that a dynamic shuttling occurs with the receptor diffusing out of the nucleus and being actively transported back into the nucleus (Guiochon-Mantel et al. 1991). Besides the ratio of transport in to diffusion out, the balance between binding to nuclear or to cytoplasmic components could also determine the subcellular distribution of steroid receptors. Elements such as the cytoskeleton and heat shock proteins (hsp90, hsp70) might keep steroid hormone receptors cytoplasmic. However, the nuclear matrix and DNA might trap the AR in the nucleus. This could explain the observed hormone dependent AR nuclear localization in transfected COS-1 cells. Ligand induces the dissociation of the heteromeric complex of receptor with heat-shock proteins, and transforms the receptor into a high affinity DNA binding and nuclear matrix binding protein (Barrack 1987). AR13, lacking the steroid binding domain is not complexed with hsp90 and constitutively binds DNA (unpublished observations). Therefore, most of these receptors will be present in the nucleus. This finding is in agreement with the proposed model. The compartmental retention model, based upon the dynamic shuttling of steroid receptors, suggests that nuclear import and export are continuous processes and that nuclear import itself is not regulated by ligand binding. This is consistent with the observation that homologous signals are present in many nuclear proteins that are not ligand dependent. There are, however, at least two mechanisms by which ligand might indirectly regulate transport of steroid receptors across the nuclear membrane. Firstly, some reports suggest that phosphorylation in or near the NLS could be involved in NLS-nuclear pore interaction and in nuclear import (Rihs & Peters 1989; DeFranco et al. 1991; Robbins et al. 1991). Hence phosphorylation, possibly induced by ligand, might be a mechanism to regulate the efficiency of nuclear import. Secondly, besides being trapped by binding to cytoplasmic components, cytoplasmic retention could also be caused by masking of the NLS. Hormone-induced changes in structural conformation or the dissociation of the heteromeric complex of the receptor with heat-shock proteins could expose the NLS, resulting in nuclear translocation (Picard & Yamamoto 1987; Picard et al. 1988; Ylikomi et al. 1992). Evidence for the latter may be provided by the observation that AR17, lacking the last 114 amino acid residues, was predominantly cytoplasmic, whereas AR14, lacking the last 197 residues, was exclusively nuclear. Possibly, in contrast to AR14, the NLS of AR17 was masked by hsp association or by conformational interference due to the additional 83 amino acid residues. Thus, according to the

compartmental retention model, at least three possible factors could influence subcellular localization: accessibility of the NLS, translocation efficiency, and binding to nuclear or cytoplasmic components.

Disruption of the DNA binding domain did not influence hormone-induced nuclear localization of the AR. Interestingly, clusters of receptors appeared in most nuclei, which were rarely observed in cells expressing the wild type AR. Expression of different AR mutants all affected in DNA binding, revealed this clustered pattern only in the presence of ligand, indicating that it is unlikely to be a non-specific precipitation. AR21, which lacks the NLS, formed clusters in the cytoplasm.

Some components involved in trapping the AR might be cell type specific and differently distributed over cytoplasm and nucleus. In addition, differences in membrane translocation efficiency could be another explanation for the cell line specific distribution of the unliganded AR that was observed in COS-1, CV1, CHO and HeLa cells.

The mechanism of action of steroid hormone antagonists has been the subject of many studies. There are a number of different processes in the cascade of steroid receptor activation that may be blocked by antagonists, including the dissociation of the heteromeric complex of the receptor with heat-shock proteins, nuclear import, receptor dimerization, DNA-binding and the activation of transcription. For several of these processes there is substantial evidence that they can be blocked by specific antagonists, but our knowledge about the inhibition of nuclear import by antagonists is limited (Kemppainen et al. 1992). Recently it has been shown that the antiandrogen Casodex (ICI 176334), was unable to induce dissociation of hsp90, hsp70 and p59 from the AR expressed in cells derived from a lymph node carcinoma of the prostate (Veldscholte et al. 1992b). Although this prostate cancer cell line expresses a mutant AR (Thr at position 868 is substituted by Ala) (Veldscholte et al. 1990), the experiments with Casodex (ICI 176334) show that dissociation of the heteromeric hsp-receptor complex is one of the processes that can be blocked by antagonists. To investigate the possibility that antiandrogens and other steroids could interfere with AR nuclear import, different concentrations of R5020, estradiol, dexamethasone or the antiandrogens OH-flutamide, cyproterone acetate and Casodex (ICI 176334) were added to COS-1 cells that were expressing the wild type AR. All compounds were able to induce nuclear localization, although relatively high concentrations (100 nM) of OH-flutamide, Casodex or dexamethasone were necessary. The observation that all these different compounds were able to cause the translocation of the AR to the nucleus suggests that passage of the nuclear membrane itself is not a ligand-specific process. This is supported by findings that non-ligands such as dopamine and okadaic acid are also able to induce nuclear translocation of the PR and of the GR, respectively (DeFranco et al. 1991; Power et al. 1991b). Possibly, upon binding of antiandrogens or other steroids, the AR is retained in the nucleus via an interaction with DNA or with the nuclear matrix. The Casodex-induced nuclear localization of the AR seems to be a more complex process, since there is evidence that Casodex cannot induce dissociation of the heteromeric hsp-receptor complex. This suggests that the whole complex passes through the nuclear pore, or that disassembly and subsequent re-assembly occur at the nuclear pore. Consequently, DNA is not the nuclear component that traps the hsp complexed AR. Loss of cytoplasmic retention or interaction with nuclear components such as the nuclear matrix could retain the receptor in the nucleus, similar to AR mutants that are unable to bind DNA. None of the compounds that were tested induced transcriptional activation of an androgen-regulated reporter gene at the lowest concentrations that resulted in AR nuclear

## Appendix paper II

localization. This indicates that the inability of the compounds tested to induce transactivation is not caused by an inhibition of nuclear import.

In conclusion, the data presented show that: 1) the NLS responsible for nuclear import of the human AR is encoded by amino acid residues 608-625 and is functionally similar to the nucleoplasmin NLS, 2) the ratio of nuclear and cytoplasmic distribution of the unliganded AR is cell line specific, 3) disruption of the DNA binding domain resulted in the formation of AR clusters in the nucleus or in the cytoplasm, and 4) antiandrogens and various steroid hormones induced nuclear localization of the AR in a dose-dependent fashion, without affecting transcription activation.

## MATERIALS AND METHODS

### Materials

R1881 (17 $\alpha$ -methyltrienolone) and R5020 (promegestone) were purchased from NEN (Boston, MA). Cyproterone acetate was a gift from Schering (Berlin, FRG), hydroxyflutamide was from Schering U.S.A. (Bloomfield, NJ) and ICI 176334 (trademark "Casodex") was from ICI Pharmaceuticals (Macclesfield, Cheshire, U.K.). All other steroids were purchased from Steraloids (Wilton, NH, U.S.A.).

### Cell culture, transfection and chloramphenicol-acetyltransferase assay

COS-1, HeLa, CV-1 and Chinese hamster ovary (CHO) cells were maintained in Eagle's Minimum Essential Medium supplemented with 5% fetal calf serum (FCS), nonessential amino acids and antibiotics. The cells were transfected using the calcium phosphate precipitation method (Chen & Okayama 1987). AR functional activity was assessed in COS-1 cells by transient transfection as described previously (Jenster et al. 1991). The reporter plasmid contains two PRE/GREs in front of the thymidine kinase (tk) promoter linked to the chloramphenicol acetyltransferase (CAT) gene (pG29GtkCAT) (Schüle et al. 1988). CAT assays representative of at least three independent experiments are shown.

### AR expression vectors

A human AR cDNA expression vector (pAR0) was constructed using the SV40 early promoter and the rabbit  $\beta$ -globin poly-adenylation signal (Brinkmann et al. 1989). Construction of the AR mutants pAR6, pAR7, pAR13, pAR14, pAR17 and pAR21 has been described (Jenster et al. 1991). The mutant pAR12 is a C-terminal truncated mutant, made by introduction of a stop-codon into the AR cDNA. For this purpose, an XbaI linker (Promega, Madison, WI, U.S.A.) that contains an in-frame stop-codon was ligated into the blunt ended Asp718 site (Jenster et al. 1991). pAR4 was constructed from an aberrant cDNA clone that lacks exons 4 to 8. The first part of intron 3 is present, so that translation will result in a protein encoded by exons 1, 2 and 3 and 20 additional amino acid residues encoded by intron 3 (618 GGGKILFFLL-LLLLSPFSLIF 638). Deletion mutants pAR61 and pAR63 were constructed by removal of internal fragments using two different restriction enzymes. The combination of the restriction

enzymes was chosen such that ligation of the blunt ended plasmid would result in an in-frame deletion. The restriction enzymes *Nru*I and *Asp*718 were used to construct pAR61, and *Asp*718 and *Sac*I were used to construct pAR63 (Jenster et al. 1991). The mutant pAR8 was made using the pSTCGR3-795 vector, containing the cytomegalovirus (CMV) promoter and rabbit  $\beta$ -globin poly-adenylation signal (Severne et al. 1988) which was digested with *Bam*HI to remove the rat GR cDNA, and was made blunt-ended. An AR cDNA fragment was isolated from pAR0 by *Sma*I digestion and was inserted in the pSTC vector. The mutants pAR22, pAR64, pAR65 and pAR28.X were constructed by site-directed mutagenesis using polymerase chain reaction DNA amplification techniques (Higuchi et al. 1988). The following primers, introducing the deletion or substitution, were used :

pAR22 : 5' GCAGTGTCCGGTATCGATGGACTACTACAACCTTTCCA 3'

pAR64 : 5' GCTCTCACATCTGGAAGCTTCAAGGTCCTC 3'

pAR65 : 5' CCGAAGGAAGAATTCTCCATCTTCTCGTC 3'

pAR28.1 : 5' GGATGACTCTGGGCGCCGGCAATCTGAAACTACAG 3'

pAR28.2 : 5' AGCAGGGATGACTCTGGGCATGCTGAAACTACAGGAGGAAG 3'

pAR28.3 : 5' TGTCCATCTTGTTCGCTCTTGGCGCCTGTTATGAAGCAGGGATG 3'

The mutant pAR28.23 is a combination of mutants pAR28.2 and pAR28.3. The mutant pAR28.31 is a combination of pAR28.1 and pAR28.3. All the mutants were sequenced to verify the correct reading frame. In addition, the mutants were expressed in COS-1 cells and the protein was analyzed by Western blotting (Jenster et al. 1991).

### Immunocytochemistry

COS-1, HeLa, CV-1 and CHO cells were cultured directly on microscope glass slides ( $2 \times 10^5$  cells/slide) and were transfected using the calcium phosphate method ( $2 \mu\text{g}$  AR plasmid and  $5 \mu\text{g}$  pTZ19 carrier plasmid/slide). The cells were plated and transfected in medium containing 5% (w/v) dextran/charcoal-treated FCS (dccFCS). Fixation and permeabilization of the transfected cells was performed essentially as described previously (Jenster et al. 1991). After the permeabilization, the cells were rehydrated in phosphate-buffered saline (PBS) and incubated for 15 min in PBS containing 10% (v/v) FCS. Slides were incubated with 1:1000 diluted monoclonal antibody F39.4.1 (Zegers et al. 1991) or polyclonal antibody Sp197 in PBS containing 1% bovine serum albumin (BSA) for 1 h at room temperature. Excess antibody was removed by washing, followed by a 1 h incubation of FITC or TRITC conjugated goat anti-mouse or goat anti-rabbit immunoglobulin (Sigma, St.Louis, MO, U.S.A.) diluted 1:1000 in PBS containing 1% (w/v) BSA. After several PBS washes, slides were mounted with 90% (v/v) glycerol in PBS. The polyclonal antibody Sp197 (designed according to Husmann et al. 1990) contains a high titer of androgen receptor antibodies as was shown by immunoprecipitation and Western blotting. F39.4.1 and Sp197 antibody specificity has been determined by mock transfection of COS-1, HeLa, CV-1 and CHO cells, showing background staining only. AR subcellular localization was classified into five different categories (N; N>C; C=N; C>N; C) based upon intensity of fluorescence in the cytoplasm compared to the nuclear fluorescence intensity. In the case where the distribution is C=N, the fluorescence intensity is the same in both compartments. In each transfection experiment at least 50 positive cells were scored. The category that was scored most is the one that is reported. In almost all cases more than 75% of the positive cells were classified into the same category. The transfection experiments were

## Appendix paper II

performed at least three times.

### ACKNOWLEDGMENTS

We thank Dr. S. Rusconi for the rat pSTCGR3-795 expression plasmid, Dr. R. Renkawitz for the pG29GtKCAT reporter construct, B. van Steensel and Dr. R. van Driel for helpful interpretations of receptor clustering and Dr. J.A. Grootegoed for helpful discussions and for reading the manuscript. This work was supported by the Netherlands Organization for Scientific Research (NWO) through GB-MW (Medical Sciences).

Appendix paper III

---

**DOMAINS OF THE HUMAN ANDROGEN RECEPTOR AND GLUCOCORTICOID  
RECEPTOR INVOLVED IN BINDING TO THE NUCLEAR MATRIX**

Bas van Steensel<sup>1</sup>, Guido Jenster<sup>2</sup>, Klaus Damm<sup>3</sup>, Albert O. Brinkmann<sup>2</sup> and Roel van Driel<sup>1</sup>

E.C. Slater Institute<sup>1</sup>, University of Amsterdam, Amsterdam, The Netherlands; Department  
of Endocrinology & Reproduction<sup>2</sup>, Erasmus University Rotterdam, Rotterdam, The  
Netherlands; Max Planck Institute of Psychiatry<sup>3</sup>, München, Germany

J Cell Biochem, in press

Bas van Steensel and Guido Jenster made equal contributions to the work described in the article

## Appendix paper III

### ABSTRACT

Steroid receptors have been reported to bind to the nuclear matrix. The nuclear matrix is operationally defined as the residual nuclear structure that remains after extraction of most of the chromatin and all soluble and loosely bound components. To obtain insight in the molecular mechanism of the interaction of steroid receptors with the nuclear matrix we studied the binding of several deletion mutants of the human androgen receptor (hAR) and the human glucocorticoid receptor (hGR) to the nuclear matrix. Receptor binding was tested for two different nuclear matrix preparations: (i) complete matrices, in which most matrix proteins are retained during the isolation procedure and (ii) depleted matrices, which consist of only a subset of these proteins. The results show that the C-terminal domain of the hAR binds tightly to both depleted and complete matrices. In addition, at least one other domain of the hAR binds to complete matrices but not to depleted matrices. In contrast to the hAR, the hGR binds only to complete matrices. For this interaction both the DNA-binding domain and the C-terminal domain of the hGR are required, whereas the N-terminal domain is not. We conclude that specific protein domains of the hAR and the hGR are involved in binding to the nuclear matrix. In addition, our results indicate that the hAR and the hGR are attached to the nuclear matrix through different molecular interactions.

### INTRODUCTION

Steroid receptors are hormone-dependent transcription factors. They can regulate the expression of specific genes by binding to regulatory DNA sequences named hormone response elements (HREs). At least three functional domains can be distinguished in steroid receptors. These include a C-terminal steroid binding domain, a DNA binding domain (DBD) and an N-terminal domain which is required for maximal transcriptional regulatory activity (Evans, 1988; Green & Chambon, 1988).

The DBDs of the glucocorticoid receptor, mineralocorticoid receptor, progesterone receptor and androgen receptor are highly homologous. This is reflected by their ability to bind to the same HRE. However, *in vivo* these steroid receptors regulate gene expression differentially. Evidence is accumulating that this target gene specificity is determined by the interplay between steroid receptors, other transcription factors and chromatin structure (Adler et al. 1992; Hayes & Wolffe, 1992; Pearce & Yamamoto, 1993; Truss & Beato, 1993).

Many members of the steroid and thyroid hormone receptor superfamily have been reported to be bound to the nuclear matrix (reviewed in Barrack, 1987; Getzenberg et al. 1990). The nuclear matrix is operationally defined as the residual nuclear structure that remains after extraction of more than 90 % of the chromatin and all soluble and loosely-bound components (Cook, 1988; Verheijen et al. 1988; Berezney, 1991; van Driel et al. 1991). It consists of a peripheral nuclear lamina and an internal fibrogranular network. Nuclear matrices are generally isolated by treatment of cell nuclei with non-ionic detergents, digestion with nucleases and extraction with a buffer of high ionic strength (Berezney & Coffey, 1974; Kaufmann et al. 1981; He et al. 1990).

Numerous data indicate that transcription is closely associated with the nuclear matrix. Active genes are enriched in the residual DNA that remains bound to the nuclear matrix



(reviewed in Getzenberg et al. 1991). Pulse labelled, nascent RNA is associated with the nuclear matrix (Jackson et al. 1981; Ciejek et al. 1982; Jackson et al. 1993; Wansink et al. 1993). Transcriptional activity has been demonstrated in the nuclear matrix (Jackson & Cook, 1985; Razin & Yarovaya, 1985). Steroid and thyroid hormone receptors (Barrack & Coffey, 1980; Barrack, 1983; Kaufmann et al. 1986; Kirsch & Miller-Diener, 1986; Kumara-Siri et al. 1986; Alexander et al. 1987; Barrack, 1987; van Steensel et al. 1991) and a number of other transcription factors (Evan & Hancock, 1987; Klempnauer, 1988; Getzenberg & Coffey, 1990; Stein et al. 1991; Bidwell et al. 1993; van Wijnen et al. 1993) are associated with the nuclear matrix. These data indicate that the nuclear matrix may play an important role in the regulation of transcription.

The interaction of steroid receptors with the nuclear matrix is dependent on the presence of hormone (Barrack, 1987). By *in vitro* reconstitution experiments, binding of the rat androgen receptor (Barrack, 1983; Colvard & Wilson, 1984), the mouse estrogen receptor (Belisle et al. 1989; Metzger & Korach, 1990) and the chicken progesterone receptor (Schuchard et al. 1991) to the nuclear matrix was demonstrated to be saturable and of high affinity. This indicates that the nuclear matrix contains specific binding sites for steroid receptors. Little is known about the nature of these binding sites. Chick oviduct nuclear matrix contains a 10 kDa protein that has been shown to bind specifically and with high affinity to the progesterone receptor (Schuchard et al. 1991). This suggests that binding of steroid receptors to the nuclear matrix is mediated by specific acceptor proteins. On the other hand, residual nuclear matrix DNA may be involved in steroid receptor binding.

The function of the interaction of steroid receptors with the nuclear matrix is presently unknown. To obtain more insight in this interaction we set out to identify protein domains of the human androgen receptor (hAR) and the human glucocorticoid receptor (hGR) that are required for binding to the nuclear matrix.

## RESULTS

### Nuclear localization of wild type and mutant hAR and hGR

To investigate the interaction of steroid receptors with the cell nucleus and the nuclear matrix COS-1 cells were transiently transfected with expression vectors coding for wild type and mutated hAR and hGR. Binding to nuclei and nuclear matrices was examined by subjecting transfected cells to several *in situ* extractions, followed by western blot analysis to estimate the amount of receptor that remained bound to nuclei or nuclear matrices. Evidently, only receptor mutants that were located in the cell nucleus after expression in COS-1 cells could be used to study binding to the nucleus and nuclear matrix. Therefore, we first investigated the subcellular localization of wild type and mutated receptor proteins by indirect immunofluorescence microscopy, using specific antibodies against these receptors.

The subcellular distribution of the wild type and a number of mutant hGRs is shown in Figure 1a-h. The wild type hGR was both cytoplasmic and nuclear in the absence of steroid hormone (Figure 1a). After addition of the GR agonist dexamethasone (DEX), wild type and most mutant glucocorticoid receptors that were tested displayed an exclusively nuclear staining (Figure 1b-h). Two hGR deletion mutants  $\Delta 451-487$  and  $\Delta 420-777$  remained cytoplasmic in the

## Appendix paper III

presence of DEX (not shown).

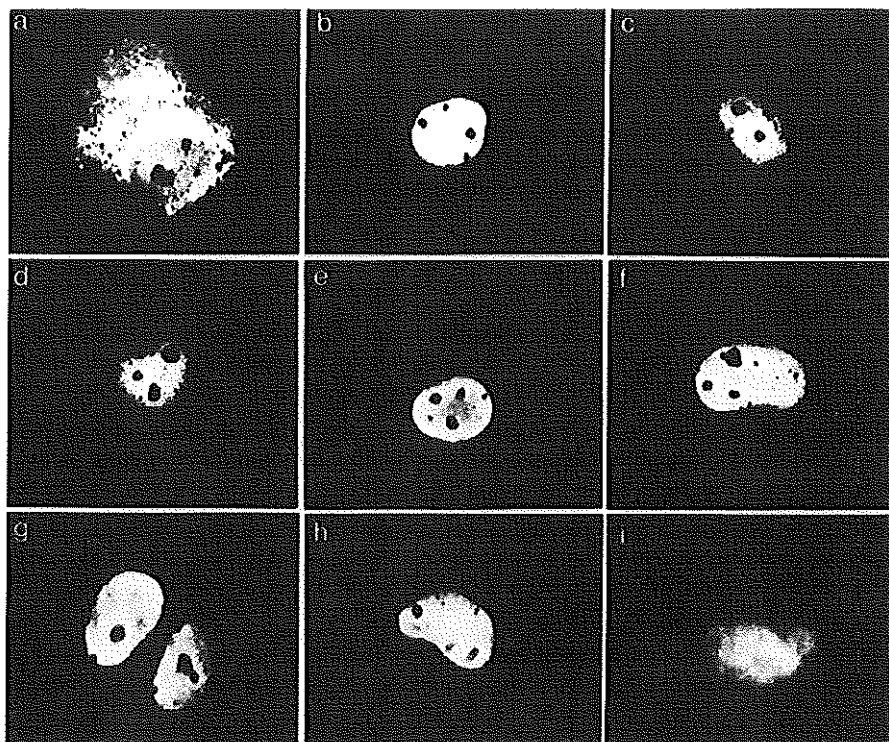


Figure 1 Subcellular localization of wild type and mutated hGR and hAR. Immunofluorescence microscopy of cells expressing various hGR and hAR mutants. COS-cells were transfected with wild type hGR (a and b), hGR mutants G438 (c),  $\Delta$ 420-451 (d), I550\* (e), I488\* (f),  $\Delta$ 9-205 (g), or  $\Delta$ 1-420 (h), or hAR mutant AR67 (i). After growing 24 hrs on microscope cover slips in the absence of hormone (a) or in the presence of  $10^{-5}$  M DEX (b-h) or  $10^{-8}$  M R1881 (i), cells were formaldehyde fixed and labeled with mouse monoclonal antibody mAb7 (a-f) or GR<sub>788-795</sub> (g, h) against the hGR, or rabbit antiserum SP066 (i) against the hAR, followed by biotinylated sheep-anti-mouse antibody and FITC labeled streptavidin (a-h) or FITC-labeled goat-anti-rabbit antibody (i).

The nuclear localization of hGR mutants I488\* and I550\* was not dependent on steroid, because these constructs lack the hormone binding domain. Based on their nuclear localization a number of hGR mutants was selected for the nuclear and nuclear matrix binding studies. These mutants are depicted in Figure 2a. Because mutant  $\Delta$ 1-420 was expressed at a very low level compared to the wild type receptor, binding studies were carried out with the almost identical mutant  $\Delta$ 9-385, which had a normal expression level.

The subcellular localization of various hAR mutants was published previously (Jenster et al. 1993). Based on these data we selected a number of mutants that were localized in the nucleus after hormone stimulation. We tested the localization of an additional hAR mutant (AR67), which lacks both the DBD and most of the N-terminal domain. After stimulation with

the AR agonist R1881, this mutant was localized mainly in the nucleus, but also some cytoplasmic labeling was present (Figure 1i). The hAR mutants that were found to be localized in the nucleus after hormone stimulation are summarized in Figure 2b.

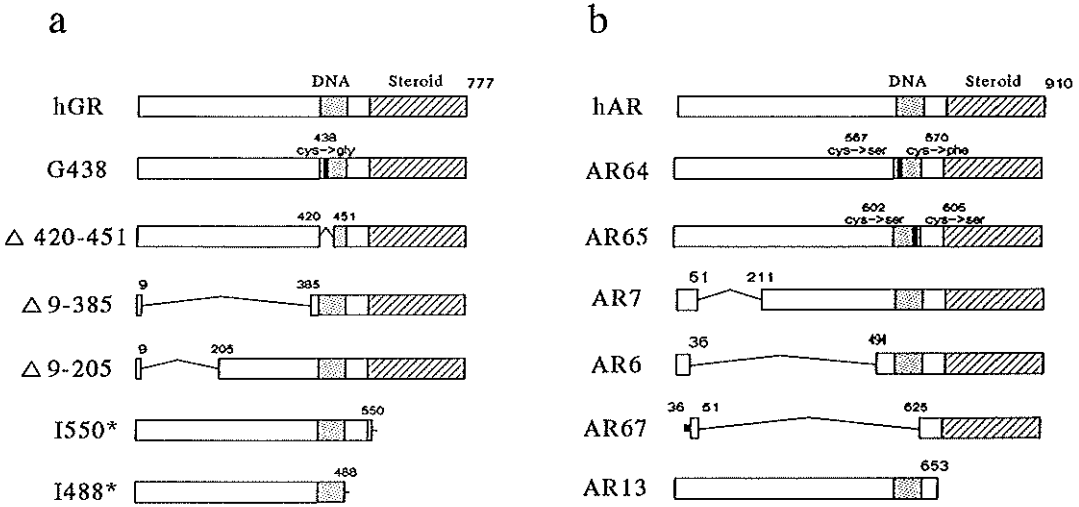


Figure 2 Scheme of receptor mutants showing predominant nuclear localization. Schematic representation of hGR (a) and hAR (b) mutants that were tested for binding to nuclei and nuclear matrices. Nomenclature of the receptor mutants was taken from the publications in which they were described first (Hollenberg et al. 1987; Hollenberg & Evans, 1988; Jenster et al. 1993).

### Isolation of nuclei and nuclear matrices and analysis of receptor binding

Binding of steroid receptors to the cell nucleus and the nuclear matrix was tested by *in situ* extraction of COS-1 cells expressing wild type or mutant hAR or hGR. Two different methods for the isolation of nuclear matrices were compared (Figure 3). In method A, cells were permeabilized with a non-ionic detergent, then stabilized with sodium tetrathionate (NaTT), digested with DNase I and extracted with 0.25 M ammonium sulphate. Method B was identical to method A except that the permeabilization and the stabilization were combined in one single step. NaTT is known to stabilize the nuclear matrix structure (Kaufmann et al. 1981; Kaufmann & Shaper, 1984) and the binding of the glucocorticoid receptor to the nuclear matrix (Kaufmann et al. 1986), most likely by introducing disulfide bridges. Components which have a weak (i.e. rapidly dissociating) interaction with the nuclear matrix may be lost during permeabilization. Immediate addition of NaTT during permeabilization may result in stabilization of some of these weak interactions. Therefore, it was expected that method B would result in retention of more proteins in the nuclear matrix fraction than method A. We analyzed receptor binding in permeabilized cells and in nuclear matrices that were obtained with both

methods.

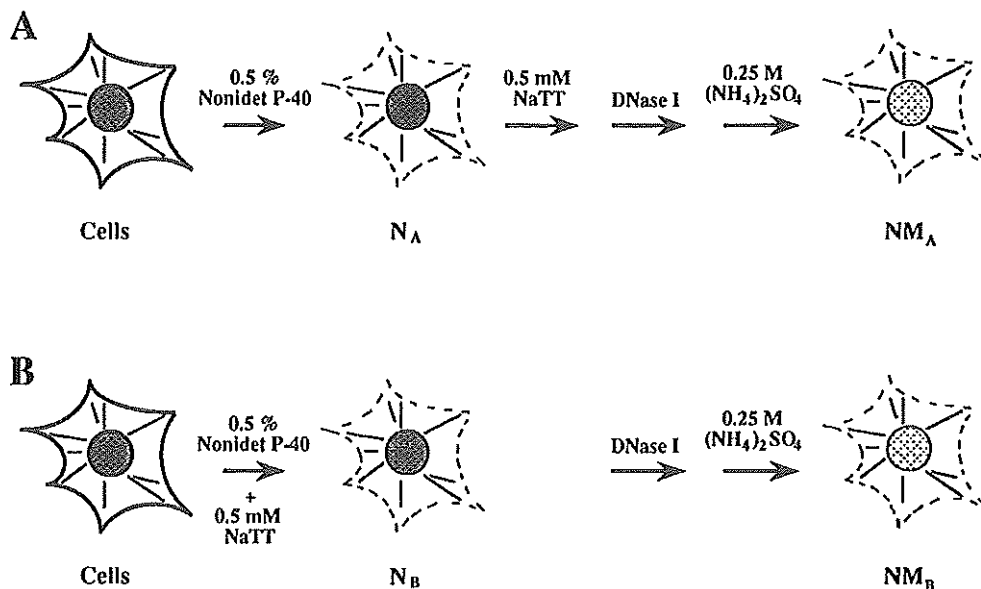


Figure 3 Scheme of the extraction methods A and B.

Schematic outline of the two *in situ* extraction methods A and B to isolate nuclear ( $N_A$  and  $N_B$ ) and nuclear matrix ( $NM_A$  and  $NM_B$ ) fractions from transfected cells. Method B was identical to method A, except that in method B the Nonidet P-40 permeabilization and the NaTT-stabilization were combined in one single incubation step (see also Materials and methods section).

Proteins from intact cells and from the insoluble fractions after extraction were separated by SDS-PAGE, blotted onto a nitrocellulose membrane and probed for the presence of receptor, using specific antibodies against hAR or hGR. The apparent molecular weight of all hGR mutants, as judged from the western blots, was close to the molecular size that can be predicted from the sequences of the constructs (data not shown). The apparent molecular weight of the wild type hAR and of the mutants AR13, AR64 and AR65 was 10-12 kDa higher than the predicted value, as was noticed previously (Jenster et al. 1991). To verify that equivalent amounts of cells and nuclear matrices were loaded, blots were also probed with a monoclonal antibody against the 65 kDa nuclear matrix protein lamin B, which served as an internal standard. Lamin B fractionates completely in the insoluble nuclear fraction (Kaufmann & Shaper, 1984). The amount of receptor that was bound to nuclei and nuclear matrices was determined semi-quantitatively as described in the *materials and methods* section.

Only receptor mutants that were located in the nucleus after steroid treatment were included in the binding studies. Therefore, receptors that remained associated with the insoluble cell fractions after permeabilization or after matrix preparation were assumed to be bound to the nucleus and the nuclear matrix, respectively. To exclude the possibility that the extractions

caused dissociation of receptors from the nucleus and subsequent trapping by residual cytoplasmic structures, we checked the localization of each receptor mutant after extraction by immunofluorescence microscopy. In no case cytoplasmic labeling was observed (data not shown). Thus, mutant receptors that were found by western blot analysis to be resistant to permeabilization and nuclear matrix extraction were bound to nuclear structures and not to cytoplasmic components (e.g. the cytoskeleton).

**Binding of wild type and mutated hAR to the nucleus and to the nuclear matrix**

When cells expressing the wild type hAR and stimulated with R1881 were permeabilized (method A), approximately 15-40 % of the receptor molecules remained bound to the nucleus (Figure 4b and Table 1).

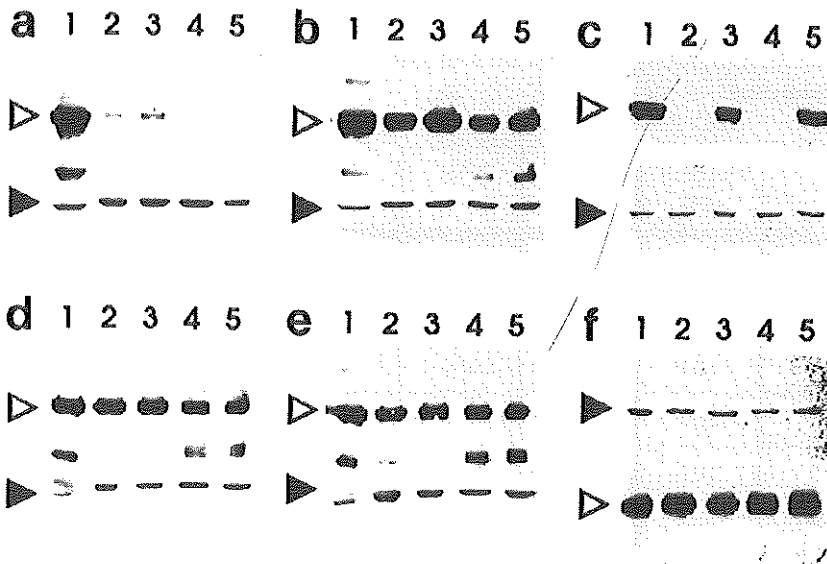


Figure 4 Western blot analysis of hAR binding to nuclei and nuclear matrices. Cells expressing wild type hAR (a and b), AR13 (c), AR64 (d), AR65 (e) or AR67 (f) were grown in the absence (a) or presence (b-f) of  $10^8$  M R1881 and extracted as described in the text and Figure 3 to obtain nuclei or nuclear matrices. Protein samples from equivalent numbers of cells, nuclei or nuclear matrices were separated by SDS-PAGE, blotted onto nitrocellulose and probed with monoclonal antibody F39.4.1 (a-e) or rabbit antiserum SP066 (f), followed by alkaline phosphatase-conjugated anti-mouse or anti-rabbit antibody, respectively. Blots were also probed with monoclonal antibody 101-B7 against laminin B, which served as an internal standard. Open arrowhead, band representing (mutant) receptor; solid arrowhead, band representing laminin B. Other bands may be hAR degradation products or represent alternative translational initiation products, because they were not observed in mock-transfected cells (data not shown). Lane 1, cells; lanes 2 and 3, nuclei; lane 4 and 5, nuclear matrices. Lanes 2 and 4, extraction method A; Lanes 3 and 5, extraction method B.

### Appendix paper III

When NaTT was added during permeabilization (method B) 40-100 % of the hAR remained bound to the nucleus. Subsequent isolation of nuclear matrices from these preparations did not result in any significant further extraction of receptor, indicating that most of the wild type hAR that is bound to the nucleus is also associated with the nuclear matrix. In the absence of hormone less than 5 % of the hAR molecules was bound to nuclei or nuclear matrices with either method A or method B (Figure 4a). Thus, binding of the hAR to the nucleus and the nuclear matrix is hormone-dependent and is enhanced two- to threefold when NaTT is present during permeabilization of the cells.

To identify the domains of the hAR that are involved in binding to the nuclear matrix, we tested the effect of several mutations on the binding of the hAR to permeabilized nuclei and nuclear matrices prepared according to method A or method B. Representative western blots are shown in Figure 4c-4f; the results of semi-quantitative analysis (see materials and methods section) of these and other blots are summarized in Table 1.

**Table 1** Semi-quantitative analysis of wild type and mutant hAR binding to nuclei and nuclear matrices

	[R1881] (M)	N <sub>A</sub>	NM <sub>A</sub>	N <sub>B</sub>	NM <sub>B</sub>
hAR	0	-	-	-	-
hAR	10 <sup>-8</sup>	++	++	+++	+++
AR64	10 <sup>-8</sup>	++	++	++	++
AR65	10 <sup>-8</sup>	++	++	++	++
AR7	10 <sup>-8</sup>	n.d.	n.d.	++	++
AR6	10 <sup>-8</sup>	++	++	++	++
AR67	10 <sup>-8</sup>	+++	+++	+++	+++
AR13	10 <sup>-8</sup>	-	-	++	++

*Semi-quantitative analysis of wild type and mutant hAR binding to nuclei and nuclear matrices.*

*The amount of receptor bound to nuclei (N<sub>A</sub> and N<sub>B</sub>) and nuclear matrices (NM<sub>A</sub> and NM<sub>B</sub>) prepared according to method A (N<sub>A</sub> and NM<sub>A</sub>) or method B (N<sub>B</sub> and NM<sub>B</sub>) was determined by semi-quantitative western blot analysis as described in the materials and methods section and was calculated as the percentage of the total amount of each mutant receptor in the cell. - = less than 5 %; + = 5-15 %; ++ = 15-40 %; +++ = 40-100 %; n.d. = not determined.*

When cells were extracted according to method A, deletion of the N-terminal domain (AR6) or disruption of either one of the two zinc fingers in the DBD (AR64 and AR65) had no significant effect on nuclear and nuclear matrix binding. However, deletion of the C-terminal

domain (AR13) resulted in complete dissociation of the hAR during permeabilization of the cells. Complementary to this, a hAR fragment lacking both the DBD and almost the entire N-terminal domain (AR67) remained bound to the cell nucleus and the nuclear matrix. We conclude that binding of the hAR to the nucleus and the nuclear matrix prepared according to method A is mediated by the C-terminal domain.

Different results were obtained with nuclear matrices isolated according to method B. With this method, all mutants tested (including AR13) remained bound to the nuclear matrix. This indicates that addition of NaTT during permeabilization affects the interaction of the hAR with the nuclear matrix not only quantitatively, but also qualitatively. The interaction of the hAR with the nuclear matrix prepared by method B is most likely mediated by more than one domain.

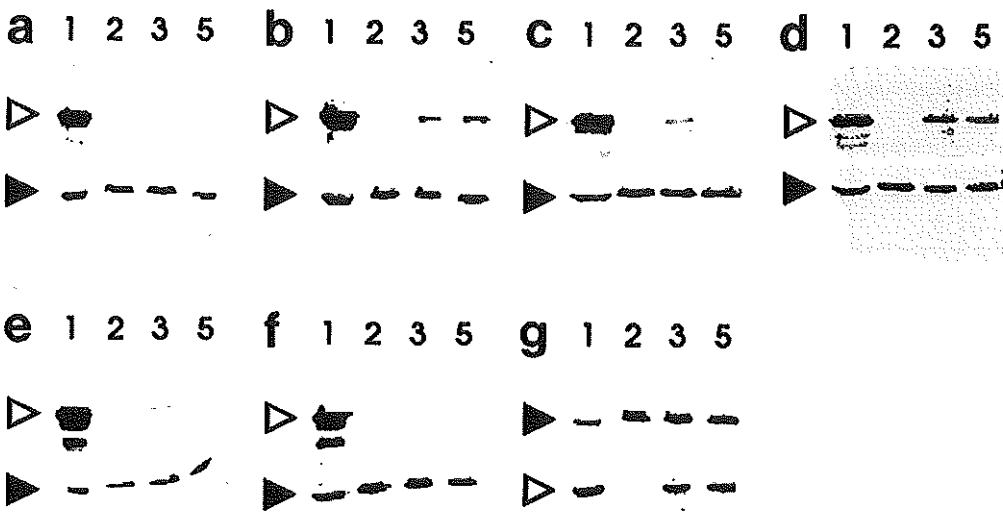


Figure 5 Western blot analysis of hGR binding to nuclei and nuclear matrices.

Cells expressing wild type hGR (a and b), I550\* (c), I488\* (d),  $\Delta$ 420-451 (e), G438 (f) or  $\Delta$ 9-385 (g) were grown in the absence (a) or presence (b-g) of  $10^{-8}$  M DEX and extracted as described in the text and Figure 3 to obtain nuclei or nuclear matrices. Protein samples from equivalent numbers of cells, nuclei or nuclear matrices were separated by SDS-PAGE, blotted onto nitrocellulose and probed with monoclonal antibody mAb7 (a-f) or GR<sub>788-795</sub> (g), followed by alkaline phosphatase-conjugated anti-mouse antibody. Blots were also probed with monoclonal antibody 101-B7 against lamin B, which served as an internal standard. Open arrowhead, band representing (mutant) receptor; solid arrowhead, band representing lamin B. Lane 1, cells; lanes 2 and 3, nuclei; lane 5, nuclear matrices. Lane 2, extraction method A; Lanes 3 and 5, extraction method B.

## Binding of wild type and mutated hGR to the nucleus and to the nuclear matrix

In contrast to the wild type hAR, the DEX-stimulated wild type hGR was completely extracted by permeabilization in the absence of NaTT (method A). Thus the hGR is not tightly bound to the nucleus under these conditions (Figure 5b). Evidently, under these conditions the hGR is also not bound to the nuclear matrix either (data not shown). However, when NaTT was added during permeabilization (method B) 15-40 % of the hGR molecules resisted extraction. The same amount of receptor was retained in nuclear matrices isolated according to method B, indicating that all nuclear bound hGR was associated with the nuclear matrix. In the absence of hormone no hGR binding to nuclei or nuclear matrices was observed (Figure 5a, Table 2). Thus, binding of the hGR to the nuclear matrix is hormone-dependent and is only observed when NaTT is present during permeabilization of the cells.

**Table 2** Semi-quantitative analysis of wild type and mutant hGR binding to nuclei and nuclear matrices

	[DEX] (M)	N <sub>A</sub>	N <sub>B</sub>	NM <sub>B</sub>
hGR	0	-	-	-
hGR	10 <sup>-8</sup>	-	++	++
G438	10 <sup>-8</sup>	-	-	-
Δ420-451	10 <sup>-8</sup>	-	-	-
Δ9-385	10 <sup>-8</sup>	-	++	++
Δ9-205	10 <sup>-8</sup>	-	++	++
I550*	10 <sup>-8</sup>	-	++	-
I488*	10 <sup>-8</sup>	-	++	+

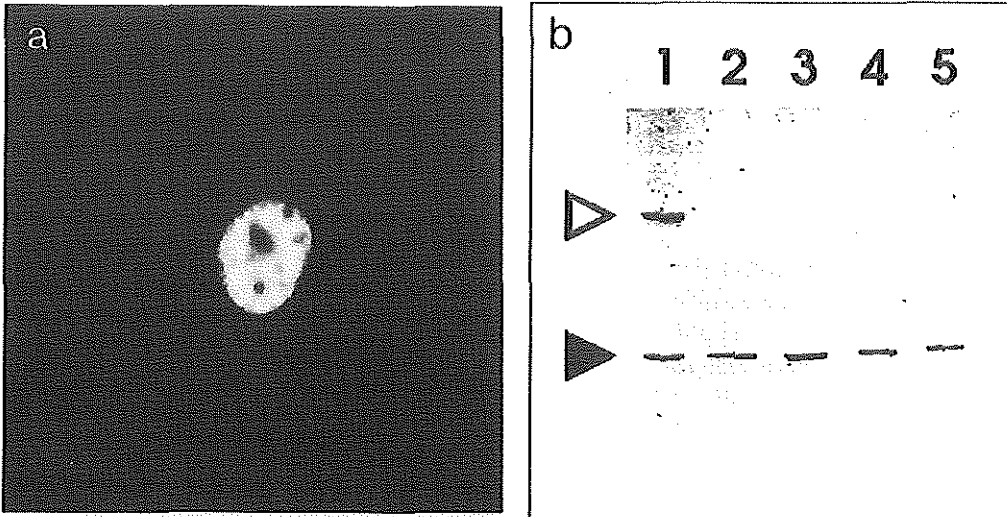
*Semi-quantitative analysis of wild type and mutant hGR binding to nuclei and nuclear matrices.*

*The amount of receptor bound to nuclei (N<sub>A</sub> and N<sub>B</sub>) and nuclear matrices (NM<sub>B</sub>) prepared according to method A (N<sub>A</sub>) or method B (N<sub>B</sub> and NM<sub>B</sub>) was determined by semi-quantitative western blot analysis as described in the materials and methods section and was calculated as the percentage of the total amount of each mutant receptor in the cell. - = less than 5 %; + = 5-15 %; ++ = 15-40 %; +++ = 40-100 %; n.d. = not determined.*

Several mutant hGRs were tested for binding to nuclei and nuclear matrices. Representative western blots are shown in Figure 5c-5f; the results of semi-quantitative analysis are summarized in Table 2. After permeabilization in the absence of NaTT (method A), none of the hGR mutants showed nuclear binding. This is consistent with the behaviour of the wild



type hGR. When cells were extracted according to method B, deletion of part or most of the N-terminal domain ( $\Delta 9-205$  and  $\Delta 9-385$ ) did not affect binding to nuclei and nuclear matrices. However, deletion of the first zinc finger of the DBD ( $\Delta 420-451$ ), or disruption of the first zinc finger by replacement of the cysteine residue at position 438 by glycine (G438), resulted in a striking decrease of receptor binding to nuclei and nuclear matrices. Deletion of the C-terminal domain (I488\* and I550\*) had no effect on nuclear binding, but resulted in a partial (I488\*) or complete (I550\*) loss of nuclear matrix binding. These results demonstrate that the intact DBD of the hGR is required for binding of the hGR to the nucleus and nuclear matrix. The C-terminal domain is not essential for binding to the nucleus, but is required for optimal binding to the nuclear matrix.



**Figure 6** Analysis of nuclear and nuclear matrix binding of a  $\beta$ galNLS fusion protein  
**a**, subcellular localization of  $\beta$ galNLS protein. COS-cells were transfected with plasmid pCHSV coding for the  $\beta$ galNLS protein, grown on cover slips, fixed in 3.7% formaldehyde and labeled with monoclonal antibody 12B3 against  $\beta$ -galactosidase, followed by FITC-labeled goat-anti-mouse antibody. No labeling was observed in cells transfected with empty expression vector.  
**b**, western blot analysis nuclear and nuclear matrix binding of  $\beta$ galNLS. Cells expressing  $\beta$ galNLS were extracted as described in the text and Figure 3 to obtain nuclei or nuclear matrices. Protein samples from equivalent numbers of cells, nuclei or nuclear matrices were separated by SDS-PAGE, blotted onto nitrocellulose and probed with monoclonal antibody 12B3, followed by alkaline phosphatase-conjugated anti-mouse antibody. The blot was also probed with monoclonal antibody 101-B7 against lamin B, which served as an internal standard. Open arrowhead, band representing  $\beta$ galNLS; solid arrowhead, band representing lamin B. Lane 1, cells; lane 2 and 3, nuclei; lane 4 and 5, nuclear matrices. Lane 2 and 4, extraction method A; Lane 3 and 5, extraction method B.

### The effect of NaTT on matrix binding of nuclear proteins

The results described above clearly show that NaTT has a stabilizing effect on the nuclear binding of steroid receptors. The observation that mutation of a single cysteine in the DBD of the hGR (mutant G438) causes a dramatic decrease in binding to NaTT-stabilized nuclei and nuclear matrices indicates that the stabilizing effect of NaTT is highly selective. Most likely, after treatment with NaTT only a specific set of nuclear proteins is associated with the nuclear matrix. We tested this further by measuring nuclear matrix binding of a protein that has no known interaction with any nuclear component and contains several SH-groups. For this purpose COS-1 cells were transfected with an expression vector coding for a fusion protein ( $\beta$ galNLS) consisting of  $\beta$ -galactosidase linked to the nuclear localization signal (NLS) of the SV40 large T antigen (Schreiber et al. 1992). The  $\beta$ galNLS protein contains 16 cysteine residues, of which the sulfhydryl groups are potential targets for oxidation by NaTT. Due to the presence of the NLS, the fusion protein is readily transported into the nucleus, as was confirmed by immunofluorescence microscopy using a monoclonal antibody against  $\beta$ -galactosidase (Figure 6a). Despite its exclusive nuclear localization, the  $\beta$ galNLS protein showed no detectable binding to nuclei or nuclear matrices, irrespective of the presence or absence of NaTT during permeabilization (Figure 6b). This observation supports the view that the nuclear interactions of the hAR and hGR presented here are not artefacts induced by random sulfhydryl crosslinking by NaTT.

## DISCUSSION

The nuclear matrix plays an important role in nuclear organization and may be involved in the regulation of transcription (de Jong et al. 1990; van Driel et al. 1991). Steroid receptors bind to the nuclear matrix after hormone stimulation (Kaufmann et al. 1986; Barrack, 1987). This binding has been reported to be saturable, indicating that specific acceptors for steroid receptors are present in the nuclear matrix (Barrack, 1983; Colvard & Wilson, 1984; Belisle et al. 1989; Metzger & Korach, 1990; Schuchard et al. 1991). These acceptors may consist of residual matrix DNA, specific matrix proteins, or both.

To obtain insight in the mechanism of binding of steroid receptors to the nuclear matrix we set out to identify protein domains in the hAR and the hGR that are involved in the interaction with the nuclear matrix. For this purpose we isolated nuclei and nuclear matrices from cells expressing wild type or mutant receptors. Binding of these receptors to nuclei and nuclear matrices was determined by semi-quantitative western blotting.

The internal fibrogranular network of the nuclear matrix is relatively labile and is in many cell types only found when a stabilization step is included in the isolation protocol. Various stabilization methods have been described, such as fixation with acrolein (DeGraaf et al. 1991), incubation at 37 or 42°C (Mirkovitch et al. 1984; McConnell et al. 1987), treatment with  $\text{Cu}^{2+}$  (Kaufmann et al. 1981; Lebkowski & Laemmli, 1982), and oxidation with sodium tetrathionate (NaTT) (Kaufmann & Shaper, 1984). The molecular events responsible for stabilization in most of these procedures are poorly understood. Stabilization by NaTT is most likely due to the formation of intra- and inter-protein S-S bridges (Kaufmann & Shaper, 1984). However, the NaTT-stabilized matrix is not a continuous network of disulfide bridged proteins

(Stuurman et al. 1992c). This implies that other interactions than disulfides are important for matrix integrity too.

In the present study we utilized NaTT to stabilize the nuclear matrix. NaTT treatment was included either during permeabilization or after permeabilization of the cells (Figure 3). When NaTT is present during permeabilization (method B) it rapidly stabilizes the nuclear matrix. We call the nuclear matrices prepared according to method B '*complete matrices*'. When NaTT is added after permeabilization (method A) some weakly associated matrix proteins may be extracted during the permeabilization step. Therefore, these nuclear matrices are called '*depleted matrices*'.

Our observations can be summarized as follows. Depleted matrices contained 15-40 % of the wild type hAR protein (Table 1). Evidently, this receptor fraction is tightly bound to the nuclear matrix. The interaction of these receptor molecules with the depleted matrix required an intact C-terminal domain. The N-terminal domain and the DNA-binding domain were not required. In contrast, complete matrices retained any of the mutant hAR proteins. We interpret these results as follows. Binding of the hAR to depleted nuclear matrices involves a tight interaction with an acceptor protein via the C-terminal hAR domain. In complete matrices at least one other interaction of the hAR with the nuclear matrix must take place, since deletion of the C-terminal domain does not abolish binding. This interaction is relatively weak, as it becomes only evident after NaTT stabilization. This interaction may be mediated by an acceptor that differs from the acceptor that binds to the C-terminal domain of the hAR. In conclusion, the C-terminal domain and at least one other domain of the hAR are involved in matrix binding.

The hGR behaved unlike the hAR. Neither wild type nor mutated hGR proteins were found in depleted matrix preparations. Complete matrices, however, contained 15-40 % of the wild type hGR protein (Table 2). These results indicate that under conditions that the matrix structure is not stabilized during permeabilization either the hGR acceptors dissociate from the matrix, or the hGR-acceptor interaction itself is a weak interaction. Treatment with NaTT evidently stabilizes such interactions. It has been reported that the glucocorticoid receptor in NaTT-stabilized rat liver nuclear matrices is covalently bound to other matrix components through intermolecular disulfide bridges (Kaufmann et al. 1986). This suggests that at least part of the stabilizing effect of NaTT results from direct crosslinking of the receptor to its nuclear matrix acceptors.

Disruption of the DBD of the hGR was sufficient to completely abolish interaction with complete matrices. Therefore, either residual matrix DNA or a matrix protein that specifically binds to the DBD may function as an acceptor for the hGR. The C-terminal domain of the hGR also contributed to matrix binding. One mutant, lacking most of this domain (I550\*), was not retained by the matrix at all, whereas another C-terminal deletion mutant (I488\*) bound only poorly. Presumably the C-terminal domain interacts with an acceptor protein. The observation that both the DBD and the C-terminal domain are required suggests that these two domains bind to the nuclear matrix in a cooperative fashion. The N-terminal domain of the hGR is not important for matrix binding under our experimental conditions.

The hAR and the hGR show different matrix binding characteristics. First, the hAR binds both to depleted and complete matrices, whereas the hGR binds only to complete matrix preparations. Second, for binding to complete matrices the hGR requires both the DBD and the C-terminal domain, whereas in the hAR disruption of either of these two domains does not affect binding to complete matrices. These differences clearly demonstrate that the hAR and the

### Appendix paper III

hGR are associated with the nuclear matrix through different molecular interactions. Most likely, the nuclear matrix contains distinct acceptors for the hAR and the hGR.

An important issue is the mechanism of stabilization of matrix protein interactions by NaTT. One might argue that the interactions of steroid receptors with the nuclear matrix are artefacts caused by non-specific formation of disulfide bridges by treatment with NaTT. A number of observations argue against this possibility. First, disruption of the DBD by a single amino acid residue substitution results in loss of binding of the hGR to the nuclear matrix, indicating that the interaction is specific. Second, we show that the fusion protein  $\beta$ galNLS does not bind to the nucleus or to the nuclear matrix in the presence of NaTT, although it is exclusively located in the nucleus and it contains 16 cysteine residues. Third, Stuurman et al. (Stuurman et al. 1992c), have demonstrated that only a very limited set of nuclear matrix proteins is actually covalently cross-linked after treatment with NaTT. Fourth, it has been shown that in NaTT-stabilized nuclear matrices the polyoma large T antigen is covalently crosslinked to a small number of proteins (Humphrey & Pigiet, 1987). These proteins are probably nearest neighbours of T antigen in the nucleus. Clearly, the stabilizing effect of NaTT is selective.

The role of different domains of the hAR and hGR in the activation of transcription has been extensively characterized (Hollenberg et al. 1987; Jenster et al. 1991). Comparison of our results with these transactivation activity data shows that there is no positive correlation between nuclear matrix binding and activation of transcription by mutant receptors. For instance, the N-terminal domains of the hAR and the hGR are required for maximal transcriptional activation (Jenster et al. 1991; Hollenberg et al. 1987), but are not essential for nuclear matrix binding. Conversely, hGR mutant I550\*, which lacks most of the C-terminal domain, still shows 40% of the wild type transactivation activity (Hollenberg & Evans, 1988), but does not bind to the nuclear matrix. Also the C-terminal domain of the hAR interacts with the nuclear matrix, but deletion of this domain does not result in a decrease in transactivation of a reporter gene construct in COS-cells (Jenster et al. 1991). This lack of correlation between matrix binding and transactivation activity suggests that matrix binding is not directly involved in the regulation of transcription by steroid receptors. However, it should be emphasized that the available data on the role of steroid receptor domains in the regulation of transcription are based on measurements using artificial reporter plasmids, which have an only poorly defined chromatin structure. It remains possible that steroid receptor-matrix interactions are required for the transcriptional activation of genes in their proper chromatin context.

We showed by immunofluorescence microscopy that all hGR mutants that were tested for nuclear matrix binding were localized in the nucleus after steroid treatment (Figure 1). These localization data are consistent with previously reported results (Cadepond et al. 1992). In general, we observed predominant nuclear localization for mutant proteins that contained the intact bipartite nuclear localization sequence (NLS). In the hGR this NLS consists of two conserved stretches of basic amino-acid residues at positions 479-480 and 491-495, separated by a spacer of ten residues (Dingwall & Laskey, 1991). Although the truncated hGR mutant I488\* lacked the C-terminal part of the bipartite NLS, we found this mutant to be located exclusively in the nucleus. In agreement with this observation Cadepond et al. (Cadepond et al. 1992) have reported the predominantly nuclear localization of a similar mutant (I491\*), which is truncated after position 491. Apparently the conserved NLS basic residues at position 491-495 are not absolutely required for nuclear import of these hGR mutants.

We also found the hAR mutant AR67 to be mainly nuclear, despite complete deletion

of the bipartite NLS at position 608-625 (Jenster et al. 1993). Because of its small molecular size (~35 kDa) this mutant protein may be able to traverse the nuclear pore without requiring a NLS (Nigg et al. 1991). The nuclear accumulation of the AR67 protein is probably the result of its interaction with the nuclear matrix.

It has been reported that disruption of the DBD of the hAR causes a dramatic change in the subnuclear distribution of this receptor (Jenster et al. 1993). After hormone stimulation the wild type hAR is homogeneously distributed throughout the nucleoplasm, whereas mutants AR64 and AR65 are concentrated in large clusters inside the nucleus. Also in nuclear matrix preparations we found a homogeneous distribution for the wild type hAR and a clustered distribution for AR64 and AR65 (data not shown). These clusters may be either large aggregates of these mutant receptors due to reduced solubility, or caused by an interaction of the hAR mutants with specific nuclear substructures, such as nuclear bodies (Stuurman et al. 1992b). In both cases the altered distribution indicates that mutant receptors AR64 and AR65 display nuclear interactions which are not manifest for the wild type receptor. Thus, caution should be taken with the interpretation of binding data of these mutants. All other mutants showed a normal wild type-like subnuclear distribution.

In conclusion, we show that specific protein domains of the hAR and hGR are involved in binding to the nuclear matrix. The binding characteristics of both receptors depend on the precise method that is used to isolate the nuclear matrix. Evidently, the interaction of certain receptor domains with the nuclear matrix is relatively labile. Our results show that the hAR and the hGR are attached to the nuclear matrix through different molecular interactions. The physiological relevance of these receptor-matrix interactions remains to be elucidated. Matrix binding of the various receptor mutants is not clearly correlated to their potency to activate transcription of an artificial reporter gene construct. However, receptor-matrix interactions may play a significant role in the activation of genes in their proper chromatin context. Alternatively, matrix binding may be involved in other receptor functions, such as inhibition of the expression of specific genes, or intra-nuclear transport or storage of receptors. Identification of the nuclear matrix acceptors that interact with steroid receptors will be an important step towards understanding the function of matrix binding.

## MATERIALS AND METHODS

### hAR and hGR expression plasmids

The hGR (Hollenberg et al. 1987; Hollenberg & Evans, 1988; Rupprecht et al. 1993) and hAR (Jenster et al. 1991; Jenster et al. 1993) expression plasmids have been published previously. Plasmid pAR67 was constructed from pAR8 (Jenster et al. 1993) in which via site-directed mutagenesis a *NarI* site was introduced in codon 625 using primer pAR28.1 (Jenster et al. 1993). *NarI* digestion resulted in the in-frame deletion of a cDNA fragment encoding residues 51-624.

### Cell culture and transfections

For hGR studies, COS-1 cells were grown in Dulbecco's Modified Eagle's Medium

## Appendix paper III

supplemented with 10% fetal calf serum. Cells were transfected by electroporation with 5  $\mu$ g receptor expression vector and 10  $\mu$ g carrier DNA (pGEM4; Promega Corp., Madison, WI) as described (Rupprecht et al. 1993). For nuclear binding studies, electroporated cells were replated in 6 cm  $\varnothing$  petridishes in Dulbecco's Modified Eagle's Medium supplemented with 10% charcoal-stripped steroid-free fetal calf serum and cultured 24 hr in the presence or absence of  $10^{-8}$  M dexamethasone (Sigma). For immunofluorescence labeling, transfected cells were replated on 16 mm glass cover slips that were pretreated with alcian blue (Brink et al. 1992), and cultured as for nuclear binding studies.

For hAR and  $\beta$ galNLS studies, cells were grown on alcian blue treated cover slips or in 6 cm  $\varnothing$  petridishes, transfected and cultured as described previously (Jenster et al. 1993). When indicated, cells were treated with  $10^{-8}$  M R1881 (17 $\alpha$ -methyltrienolone, NEN, Boston, MA) for 24 hours.

### Immunofluorescence labeling and microscopy

All incubations were carried out at room temperature. After hormone treatment for 24 hours, hAR-transfected cells grown on cover slips were either directly fixed or first extracted *in situ* to obtain nuclear or nuclear matrix fractions, which were subsequently fixed. Fixation occurred by washing twice in PBS, fixing with 3.7% formaldehyde 10 min in PBS, washing twice in PBS and permeabilization with 0.5% Nonidet P-40 in PBS for 10 min. After two wash steps in PBS, fixed and permeabilized specimen were incubated 20 min with 10% FCS in PBS, followed by incubation with the monoclonal antibody F39.4.1 (Zegers et al. 1991) (ascites fluid, diluted 1:1000) or the polyclonal antibody SP066 (Kuiper et al. 1993) for 1 hr, followed by 6 x 5 min washing in PBS, and subsequent incubation with FITC-conjugated goat-anti-mouse antibody (Sigma), or FITC-conjugated goat-anti-rabbit antibody (Sigma) for 1 hr. After 6 washing steps in PBS and 10 min incubation with 1  $\mu$ g/ml Hoechst 33258 in PBS, preparations were mounted in 1 mg/ml p-phenylene diamine (Sigma) in 75% glycerol in PBS.

hGR transfected cells, grown on cover slips, were treated with hormone for 24 hours, washed twice in PBS, fixed 10 min in 2% formaldehyde in PBS, washed twice in PBS, permeabilized 10 min in 0.5% Nonidet P-40 and washed twice in PBS. Cells were incubated 2 x 10 min in PBG (0.5% BSA, 0.1% gelatin in PBS), followed by overnight incubation with a mouse monoclonal antibody mAb7 (Okret et al. 1984) (1  $\mu$ g/ml) or monoclonal antibody GR<sub>788-795</sub> (undiluted hybridoma culture supernatant) directed against the hGR. Western blotting and immunofluorescence labeling of several hGR deletion mutants demonstrated that the epitope of mAb7 is situated in region 9-205 (data not shown). Antibody GR<sub>788-795</sub> is directed against the C-terminal amino acid residues 788-795 of the rat GR (Flach et al. 1992), which are identical to amino acid residues 770-777 of the hGR. After 6x5 min washing in PBG, cells were incubated 1 hr in PBG with biotin-conjugated sheep-anti-mouse antibody (Jackson ImmunoResearch Laboratories) followed by FITC-conjugated streptavidin (Gibco) for 1 hr. After 6x5 min washing in PBG, preparations were stained with Hoechst 33258 and mounted as described above. The same procedure was followed for labelling of permeabilized cells and nuclear matrices.

Cells transfected with the  $\beta$ galNLS fusion protein were fixed and stained as with hAR-transfected cells, using monoclonal antibody 12B3 against  $\beta$ -galactosidase.

### Permeabilization and nuclear matrix preparation

All incubations were carried out *in situ* on monolayer cultures of transfected cells.

For permeabilization, transfected and hormone treated cells were washed twice in CSK buffer (10 mM PIPES, pH 6.8; 100 mM NaCl; 300 mM sucrose; 3 mM MgCl<sub>2</sub>; 1 mM EGTA; 0.5 mM benzamidin; 1 mM PMSF; 1 μg/ml pepstatin; 1 μg/ml aprotinin), incubated 3 min with 0.5 % Nonidet P-40 in CSK with (method B) or without (method A) 0.5 mM NaTT, then 5 min with CSK with or without NaTT, respectively, and washed twice briefly with CSK. Then cells were incubated 85 min in CSK, or subjected to extractions to obtain nuclear matrices.

For nuclear matrix preparations, cells permeabilized in the absence of NaTT were incubated 10 min with 0.5 mM NaTT in CSK (method A), and cells permeabilized in the presence of NaTT were incubated 10 min in CSK (method B). Subsequently, all cells were washed once with CSK and once with digestion buffer (CSK with 50 mM NaCl instead of 100 mM NaCl), followed by incubation for 1 hr with 250 μg/ml DNase I (type IV, Sigma) in digestion buffer. Then cells were extracted with 0.25 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> in digestion buffer for 10 min and washed twice with CSK. By phase-contrast microscopy it was confirmed that more than 80% of the cells remained attached to the substratum during the *in situ* extractions. Permeabilized cells and nuclear matrices in petridishes were collected in 200 μl sample buffer (2% SDS, Tris pH 6.8, β-mercapto ethanol, glycerol), and heated at 95°C for 5 min. Total cell protein samples were obtained by washing transfected and hormone treated cells twice in CSK, after which they were collected and heated in 200 μl sample buffer. Permeabilized cells and nuclear matrices on cover slips were prepared further for immunofluorescence labeling.

### Polyacrylamide gel electrophoresis and western blotting

Protein samples from approximately 2x10<sup>5</sup> cell equivalents were separated on 8% polyacrylamide gels in the presence of SDS and transferred to nitrocellulose membranes (Towbin et al. 1979). Membranes were incubated 16 hr with 1% blocking reagent (Boehringer, Mannheim), followed by one of the antibodies mentioned above, followed by alkaline phosphatase conjugated goat-anti-mouse (Jackson ImmunoResearch Laboratories, West Grove, PA) or goat-anti-rabbit antibody (BioRad). Blots were stained using the BCIP/NBT method (BioRad). Antibodies were diluted in 0.1% gelatin, 0.5% bovine serum albumine, 0.05% Tween-20 and 300 mM NaCl in PBS. After staining, all blots were re probed with monoclonal antibody 101-B7 against lamin B (Matritech, Cambridge, MA) and alkaline phosphatase conjugated goat-anti-mouse antibody. In case of the receptor mutants I488\* and I550\*, lamin B probing was carried out on a separate, identical blot because these receptor mutants have approximately the same apparent molecular size as lamin B.

For semi-quantification of the amount of receptor bound to nuclei and nuclear matrices, western blots were made for each receptor mutant of equivalent amounts of nuclei and nuclear matrices and a dilution series of a total cell protein sample from the same experiment (not shown). Receptor band intensities of nuclei and nuclear matrices were compared by eye with the dilution series of the total cell preparation to estimate the approximate percentage of cellular receptor that was retained in nuclei and nuclear matrices. This analysis was carried out in duplicate or triplicate for each receptor mutant.

## Appendix paper III

### ACKNOWLEDGMENTS

We thank Dr. R. M. Evans for the wild type and mutant hGR plasmids, Dr. J. Ménissier-de Murcia for the pCHSV plasmid, Dr. J.A. Gustafsson for the mAb7 antibody, Dr. H. M. Westphal for the GR<sub>788-795</sub> antibody and Dr. J. Trapman for the 12B3 antibody. We thank Dr. E.R. de Kloet for stimulating discussions and Erica van Binnendijk for excellent technical assistance. This work was supported by the medical council (GB-MW) of the Dutch Organization for Scientific Research (B. v. S. and G. J.).



**IDENTIFICATION OF TWO TRANSCRIPTION ACTIVATION UNITS IN THE N-  
TERMINAL DOMAIN OF THE HUMAN ANDROGEN RECEPTOR**

Guido Jenster<sup>1</sup>, Hetty A.G.M. van der Korput<sup>2</sup>, Jan Trapman<sup>2</sup> and Albert O. Brinkmann<sup>1</sup>

Departments of Endocrinology & Reproduction<sup>1</sup> and Pathology<sup>2</sup>,  
Erasmus University Rotterdam, Rotterdam, The Netherlands

Submitted

**SUMMARY**

### SUMMARY

To locate in detail the regions in the human androgen receptor (AR) involved in transcription activation, a series of N-terminal deletions was introduced in the wild type AR and in a constitutively active AR. The different constructs were tested for their capacity to activate transcription. Almost the entire N-terminal domain (residues 1-485) was necessary for full wild type AR activity when cotransfected with the (GRE)<sub>2</sub>tkCAT reporter in HeLa cells. In contrast, a smaller part of the N-terminal domain (amino acids 360-528) was sufficient for the constitutively active AR to induce transcription of the same (GRE)<sub>2</sub>tkCAT reporter in HeLa cells. This demonstrates the capacity of the AR to use different regions in the N-terminal domain as transcription activation units (TAUs).

To obtain additional information of AR N-terminal TAUs, the GAL4 DNA binding domain was linked to either the entire, or parts of the AR N-terminal domain and co-transfected with the (UAS)<sub>2</sub>tkCAT reporter in HeLa cells. The results confirmed that the first 485 amino acid residues accommodate a transcription activation function. When the chimeric AR-GAL4 constructs were tested on a different reporter ([UAS]<sub>5</sub>E1bCAT), a small shift in position of the TAU, responsible for full transactivation, was observed.

The data presented show that the size and location of the active TAU in the human AR is variable, being dependent on the promoter context and the presence or absence of the ligand binding domain.

### INTRODUCTION

Steroid hormone receptors constitute one of the best available model systems for studying regulation of gene transcription. Based upon phylogenetic studies, the family of nuclear receptors can be grouped into three subfamilies: 1) steroid receptors to which also the androgen receptor (AR) belongs, 2) thyroid hormone and retinoid receptors, and 3) orphan receptors which lack a well defined ligand (Laudet et al. 1992). Upon testosterone or dihydrotestosterone binding, the AR undergoes several sequential processes in order to interact with cognate DNA sequences. These DNA sequences (hormone response elements [HREs]) are commonly located in the regulating regions of the target genes. The binding of the AR to the HRE results in the formation of a stable preinitiation complex near the transcription start site which allows efficient transcription initiation by RNA polymerase II. The mechanism by which steroid receptors stabilize or interact with the preinitiation complex is poorly understood. There is experimental evidence for a direct interaction of receptor with the general transcription factor IIB (TF-IIB) (Ing et al. 1992; Baniahmad et al. 1993). Furthermore, steroid receptors could indirectly associate with the preinitiation complex via so called bridging factors (co-activators), or could make promoters accessible for other transcription activators by nucleosome displacement (reviewed in Archer et al. 1992; Gill & Tjian 1992; Truss et al. 1992a; Adams & Workman 1993; Wolffe 1994).

All nuclear receptors are composed of at least four functional domains: the N-terminal domain, DNA binding domain (DBD), hinge region, and the ligand binding domain (Evans 1988; Green & Chambon 1988; Truss & Beato 1993). The DNA binding domain, composed of two "zinc-finger" structures, is, like the C-terminal ligand binding domain, highly conserved

among the steroid receptors (Laudet et al. 1992). Interestingly, although the N-terminal domain of all steroid receptors harbours a transcription activation function, its length and amino acid residue constitution is unique for each receptor (Giguere et al. 1986; Gronemeyer et al. 1987; Kumar et al. 1987; Dobson et al. 1989; Jenster et al. 1991; Simental et al. 1991; Rupprecht et al. 1993). The sizes of the N-terminal domain of the androgen-, glucocorticoid- (GR), progesterone- (PR) and mineralocorticoid receptors (MR) are approximately half of the total receptor size. Compared to other nuclear receptors, this is exceptionally large and it coincides with the observation that the AR, GR, PR, and MR are the only nuclear receptors recognizing HREs that fit consensus glucocorticoid response elements (GREs) (Forman & Samuels 1990; Laudet et al. 1992; Freedman & Luisi 1993). The specific GRE binding is reflected by the presence of a glycine, serine and valine residue in the so called P-box located in the first zinc-finger of the DNA binding domain (Danielsen et al. 1989; Mader et al. 1989; Zilliacus et al. 1991). Based on the P-box sequence, the AR, GR, PR, and MR are classified in the 'GSckV-group' (Forman & Samuels 1990; Laudet et al. 1992). The residues in the P-box determine which type of HRE half-site (consensus GRE half-site: TGTCT or estrogen response element (ERE) half-site: TGACCT) the receptor recognizes (Danielsen et al. 1989; Mader et al. 1989; Umesono & Evans 1989; Zilliacus et al. 1991). Almost all members of the thyroid hormone, retinoid- and orphan receptor subfamilies, bind HREs containing ERE-like half-sites and are classified in the 'EGckA/EGckG/EGckS' P-box groups (Forman & Samuels 1990; Laudet et al. 1992). In these subfamilies, different mechanisms have been described that enlarge diversity and could explain, at least partially, cell-, and promoter discrimination: receptor heterodimerization, variable spacing of the HRE half-sites, and direct repeat or inverted repeat orientation of the two HRE half sites (Zechel et al. 1994a,b) (reviewed in De Luca 1991; Glass et al. 1991; Green 1993a; Giguere 1994). Since these variabilities have not been found for the "GSckV class" receptors, it seems reasonable to assume that the large and unique N-terminal domains of the AR, GR, PR and MR are important in cell- and receptor specific regulation of target genes, through multiple N-terminal transcription activation units (TAUs, also referred to as TAF, AF or TAD). For some promoters, experimental evidence has been provided, showing that promoter context as well as receptor N-terminal domain could determine receptor specificity (Adler et al. 1992; Pearce & Yamamoto 1993).

Several different sequence motifs that characterize transcription activation units have been identified thus far. Acidic regions (acidic activation domains [AAD]), proline-rich domains and glutamine-rich domains (reviewed in Ptashne 1988; Mitchell & Tjian 1989; Gill & Tjian 1992; Tjian & Maniatis 1994). Although the AR N-terminal domain does not possess a significant sequence identity with other known transcription factors, it is glutamine- and proline-rich (including homopolymeric Gln-, Pro-stretches) and has a relatively high number of acidic amino acids (Faber et al. 1989).

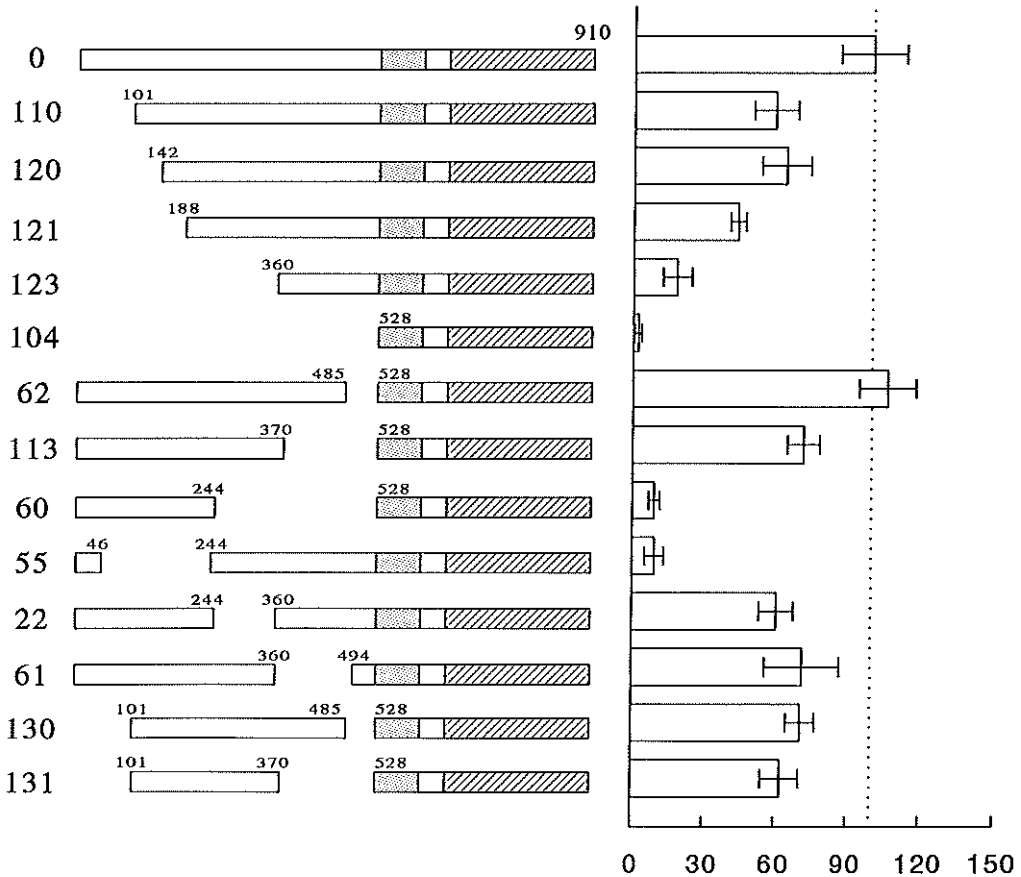
To investigate the size and location of the TAU responsible for the transactivating capacity of the human AR in different situations, a series of N-terminal deletions was introduced in the wild type AR and in a constitutively active AR. These AR mutants were tested for their ability to activate transcription of an androgen responsive reporter gene. In addition, chimeric AR-GAL4 DNA binding domain constructs were generated to obtain additional evidence for the localization of TAU regions.

## RESULTS

### Functional mapping of the TAU responsible for wild type AR transactivating capacity

A series of AR N-terminal deletion mutants was tested for transcription activity in HeLa cells using the  $(\text{GRE})_2\text{tkCAT}$  reporter (Figure 1) (pG29GtkCAT [Schüle et al. 1988]). Cotransfection of the wild type AR expression plasmid (pAR0) together with the reporter construct in the presence of 1 nM R1881 (a non-metabolizable androgen), resulted in a 20 to 40 fold induction of CAT expression as compared to the CAT activity of the unliganded AR (data not shown). N-terminal deletion of the first 100 or 141 amino acid residues (AR110, AR120 respectively) resulted in a reduction to approximately 60% activity as compared to wild type AR (Figure 1). Further deletion to amino acids 188 (AR121), 360 (AR123) or 528 (AR104) gradually reduced the transcription activation capacity of the respective AR mutants. These results suggest that almost the entire N-terminal domain is involved in transcription activation, because no sharp N-terminal border of a TAU could be marked. To locate the C-terminal boundary of the TAU responsible for the wild type AR transactivating capacity, AR62 ( $\Delta 485\text{-}528$ ), AR113 ( $\Delta 370\text{-}528$ ) and AR60 ( $\Delta 244\text{-}528$ ) were generated and tested for activity. Deletion of the last 42 residues of the AR N-terminus (AR62) had no effect on transcription activity, whereas deletion of 157 residues (AR113) reduced transactivity to approximately 70%. AR60 which lacks the second half the N-terminal domain is hardly capable of activating transcription of the  $(\text{GRE})_2\text{tkCAT}$  reporter. Importantly, the inability of AR60 to activate transcription could not be attributed to a smaller deletion. AR22 ( $\Delta 244\text{-}360$ ), AR61 ( $\Delta 360\text{-}494$ ) and AR62 ( $\Delta 485\text{-}528$ ) contain smaller deletions that together cover the large region deleted in AR60. All three AR mutants, however, were quite capable of inducing transcription. This indicates that the TAU function cannot be attributed to a single small structural element and suggests the presence of several different TAU-subunits. Deletion of the region between amino acids 46-244 in the first half of the N-terminal domain (AR55), also resulted in a major reduction of transactivating capacity. This mutant showed that although the deletion of amino acids 244-528 (AR60) strongly affected transcription activity, these sequences by themselves (AR55) are not sufficient for transactivation.

To locate the minimal part of the N-terminal domain that still retained most of the transactivation capacity, AR130 and AR131 were constructed. Although both mutants were quite capable of inducing transcription (approximately 65% compared to the wild type AR0), the only AR mutant retaining full transactivity was AR62, which again indicates that almost the entire AR N-terminus is necessary for wild type receptor activity (residues 1-485: designated TAU-1).



**Figure 1** Functional analysis of N-terminal deletion mutants of the wild type AR0. Transcriptional activity was examined by cotransfection of AR expression plasmids and a (GRE)<sub>3</sub>kCAT reporter plasmid. CAT activity was determined from cell lysates of transfected HeLa cells cultured in the absence or presence of 1 nM R1881. Activities were corrected for the (GRE)<sub>3</sub>kCAT background and the mean ( $\pm$ SE) of the R1881 treated samples of 5 independent assays are presented as a percentage relative to that of the wild type AR0.

**Functional mapping of the TAU responsible for transactivating capacity of the constitutively active AR mutant**

To obtain information on the size and location of the TAU responsible for the transactivating capacity of a constitutively active AR, a construct was generated which lacks the ligand binding domain (AR5). Compared to the wild type AR, the transcription activation capacity of AR5 was approximately 75%. R1881 had no effect on the transcription induction of the constitutively active receptor. The same series of deletions as analyzed in AR0

## Appendix paper IV

derivatives, was introduced in the constitutively active AR5. The capacity of these constructs to activate transcription was, like the wild type AR deletion series, tested in HeLa cells using the same  $(GRE)_2tkCAT$  reporter.

Surprisingly, in contrast to the results found in the ARO deletion analysis, truncation of the first 359 amino acids in the constitutively active AR (AR106), did not result in a decrease in the transactivating capacity, indicating a TAU in the 360-528 region (Figure 2). All other deletion mutants confirm the presence of a TAU in this area (designated TAU-5). Deletions of part of the 360-528 domain resulted in completely loss of, or decreased transactivating capacity (AR99, AR126, AR98 and AR117). In contrast, the mutants AR100, AR115, AR124, AR127 and AR105 in which the 360 to 528 region is present, complete, or almost complete transactivating capacity was observed. So, the smallest N-terminal region responsible for full AR5 transactivity is located between residues 360 and 528.

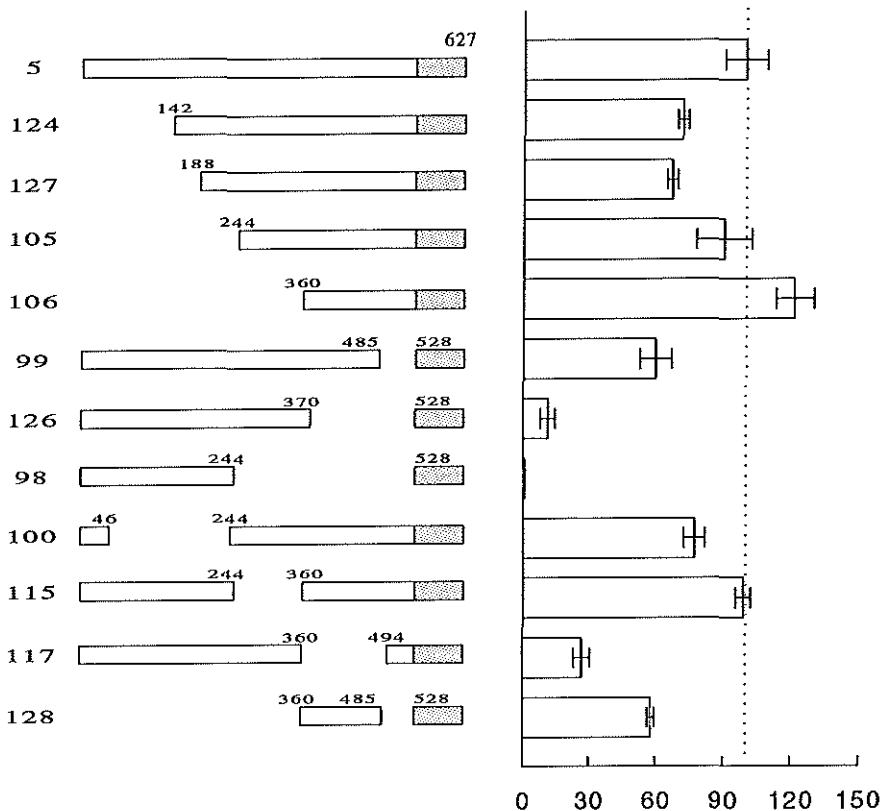


Figure 2 *Functional analysis of N-terminal deletion mutants of the constitutively active AR5. Transcriptional activity was examined by cotransfection of AR expression plasmids and a  $(GRE)_2tkCAT$  reporter plasmid. CAT activity was determined from cell lysates of transfected HeLa cells cultured in the absence of R1881. Activities were corrected for the  $(GRE)_2tkCAT$  background and the mean ( $\pm SE$ ) of 5 independent assays are presented as a percentage relative to that of AR5. AR106 is the only receptor mutant that could not be visualized by Western immunoblotting with the available anti-AR antibodies (see Materials and methods).*

### Functional mapping of the TAU responsible for transactivating capacity of the chimeric AR-GAL4 constructs.

To obtain additional evidence for the presence of a TAU in the AR N-terminal domain, a construct was generated containing the first 550 amino acids of the AR, linked to the GAL4 DNA binding domain (AR4G). In addition, a series of constructs was generated encoding several different parts of the AR N-terminal domain, linked to the GAL4 DBD. Chimeric proteins were tested for transactivating capacity in HeLa cells using the (UAS)<sub>2</sub>tkCAT reporter (Figure 3, filled bars). The (UAS)<sub>2</sub>tkCAT reporter is comparable to the (GRE)<sub>2</sub>tkCAT reporter construct used for the wild type, and constitutively active AR mutants, except that the two GRE sequences have been replaced by upstream activating sequences (UASs) to which the GAL4 DNA binding domain can bind (Banahmad et al. 1991). It could be expected that the TAU, responsible for the transactivating capacity of the constitutively active AR4G, would be located in the same region (amino acids 360-528) as TAU-5, which is responsible for the constitutively AR5 transactivating capacity. However, enlargement of the N-terminal truncation (AR94G, AR107G, AR105G, AR106G) resulted in a gradual decrease in transcription activity, when tested on the (UAS)<sub>2</sub>tkCAT reporter (Figure 3, filled bars). The region between amino acids 370 and 550 is still essential for transactivity of AR4G since the transactivating capacity of AR113G (which lacks amino acids 370-550) was strongly reduced. The TAU responsible for full AR4G activity, when tested on the (UAS)<sub>2</sub>tkCAT reporter, is located between amino acids 188 and 485 (AR91G) (Figure 3).

The same series of chimeric proteins was tested on the (UAS)<sub>5</sub>E1bCAT construct (Figure 3, empty bars). This reporter contains five UASs upstream of the E1bTATA-box and represents a different promoter environment (Lillie & Green 1989). The transcription activation capacities of the chimeric proteins were different when tested on the (UAS)<sub>5</sub>E1bCAT reporter. A large reduction in transactivating capacity was observed between mutants AR94G ( $\Delta$ 1-142) and AR107G ( $\Delta$ 1-188) which shows that the region between amino acids 142-188 is essential for the chimeric proteins to be able to activate transcription of the (UAS)<sub>5</sub>E1bCAT reporter. In contrast, the capacities of AR94G and AR107G to activate transcription of the (UAS)<sub>2</sub>tkCAT reporter were not significantly different. This promoter-specificity, determined by AR amino acids 142-188 is also obvious for mutants AR100G and AR91G (Figure 3). On the other hand, the region between amino acids 360 and 550 is more important for the chimeric proteins to be able to activate transcription of the (UAS)<sub>2</sub>tkCAT construct. AR106G, which only contains this last part of the AR N-terminal domain, was incapable of inducing transcription when cotransfected with the (UAS)<sub>5</sub>E1bCAT construct, but retained approximately 35% activity when tested on the (UAS)<sub>2</sub>tkCAT reporter. When tested on the (UAS)<sub>5</sub>E1bCAT reporter, the minimal part of the AR N-terminal domain still capable of inducing full transactivity, is located between residues 142 and 485 (AR96G).

Appendix paper IV

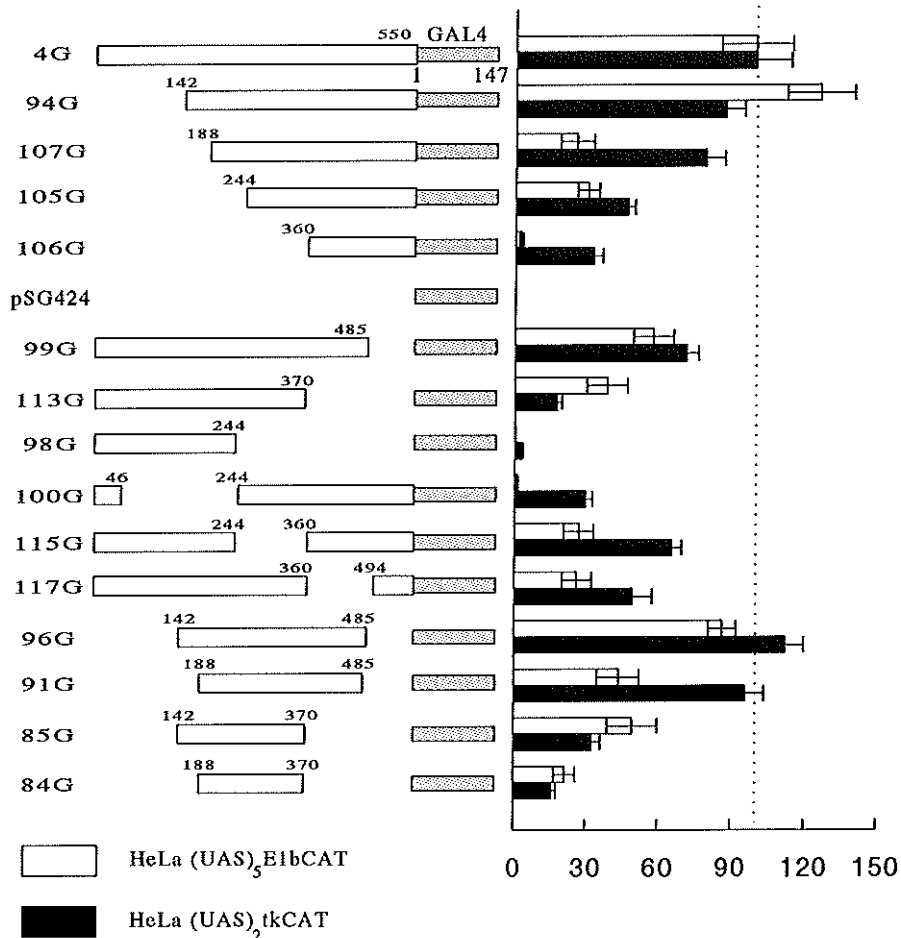


Figure 3 *Functional analysis of N-terminal deletion mutants of the chimeric AR4G*  
*Transcriptional activity was examined by cotransfection of AR-GAL4 expression plasmids and a (UAS)<sub>5</sub>E1bCAT (empty bars) or (UAS)<sub>2</sub>tkCAT (filled bars) reporter plasmid. CAT activity was determined from cell lysates of transfected HeLa cells cultured in the absence of R1881. Activities were corrected for the reporter background and the mean ( $\pm$ SE) of 5 independent assays are presented as a percentage relative to that of AR4G.*

DISCUSSION

To dissect the transcription activation properties of the human AR, a series of AR deletion mutants was analyzed to characterize and locate N-terminal regions essential for transcription activity. Deletion mapping of the wild type AR revealed that for full receptor activity, almost the entire N-terminal domain is necessary (TAU-1, residues 1-485) (Figure 4). Any deletion, except for the deletion of the last 42 residues in the AR N-terminus, affected the capacity to induce transcription when analyzed in HeLa cells using the (GRE)<sub>2</sub>tkCAT reporter. The minimal region (designated core region) that still retained over 50% transactivity compared



to AR0, was located between residues 101 and 370 (Figure 4). Interestingly, these results are clearly different from the data obtained in the deletion mapping studies of the constitutively active AR5. The same series of deletions was introduced in AR5, and the respective mutants were analyzed under identical experimental conditions (HeLa cells using the (GRE)<sub>2</sub>tkCAT reporter). These studies revealed that a region between amino acids 360 and 528 (designated TAU-5) was sufficient for full constitutively transactivity (Figure 4). This demonstrates the capability of the AR to use different N-terminal regions for transcription activation. Furthermore, these observations suggest a determinant role for the ligand binding domain in TAU functioning. Deletion of the ligand binding domain resulted in loss of TAU-1 activity and induced the use of TAU-5. These findings however, do not permit to draw conclusions with respect to the functional *in vivo* existence of separate TAUs that could explain cell- and receptor-specific gene expression, since so far, constitutively active ARs have not been identified in intact *in vivo* systems.

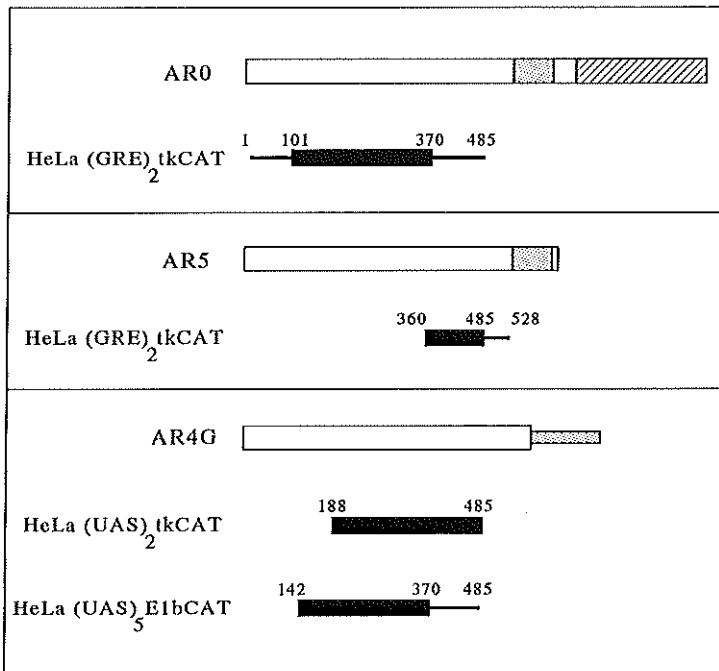


Figure 4 Summary of the regions of the AR N-terminal domain responsible for the transactivating capacity of the wild type AR and AR mutants.

Total bars represent the region necessary for full receptor activity compared to the starting receptor (AR0, AR5 or AR4G). Thick bars, as part of the total bar, represent the region responsible for 50% or more of the transcription activity (core region).

Although TAU-1 and TAU-5 overlap, the core of the two TAUs (responsible for over 50% of the activity), are separate N-terminal regions with individual characteristics. TAU-1 contains a relatively high number of acidic amino acids, three glutamine repeats of which one

is polymorphic, and potential phosphorylation sites (Faber et al. 1989; Kuiper et al. 1993). The TAU-5 core is not acidic and harbours three different amino acid stretches: 1) a proline stretch (residues 371-378), 2) an alanine stretch (residues 397-401), and 3) a glycine stretch (residues 448-463) (Faber et al. 1989). The role of the different amino acid stretches in the TAU-5 region is not known. Lengthening of the Gln repeat to more than 40 residues is associated with Kennedy's disease, an X-linked neuro-degenerative disorder, characterized by a slowly progressing muscle weakness (La Spada et al. 1991). Mhatre and co-workers (1993) have shown that the extension of the Gln-repeat to 40 or 50 residues, resulted in a decreased AR capacity to activate transcription.

The capability of the AR to use different and unique regions of its N-terminal domain as TAUs, introduces the possibility that different TAUs are responsible for the regulation of different genes, resulting in cell-specific and AR-specific gene expression. Evidence for the use of different N-terminal regions to activate different kind of promoters was provided by the analysis of a series of AR-GAL4 chimeric constructs. These proteins contained either the entire or part of the AR N-terminal domain linked to the GAL4 DNA binding domain. When the transactivating capacity was tested on two different promoters, a shift in position of the TAU responsible for full activity was observed. The location of the TAU was more C-terminal when tested on the (UAS)<sub>2</sub>tkCAT construct compared to the (UAS)<sub>3</sub>E1bCAT reporter (Figure 4). It is unlikely that these two situations represent the use of two different and separate TAUs but rather indicate a small difference in the location of essential sequences. Since the AR4G and AR5 constructs only differ in their DBD, it might be expected that the respective TAUs would be located in the same part of the N-terminal domain. When AR4G derivatives were tested on the (UAS)<sub>2</sub>tkCAT reporter and the AR5 derivatives on the comparable (GRE)<sub>2</sub>tkCAT reporter, the region between amino acids 360-528 was essential for the transactivating capacity of both AR4G and AR5. However, for full transactivating capacity, the AR4G needs a larger region (amino acids 188-485) which might indicate that the replacement of the DBD influenced the size of the N-terminal region used as the TAU.

In contrast to the DNA- and steroid binding domain, it is clear that the large and unique TAU-1 responsible for wild type AR transactivating capacity, is not a sharply bordered functional domain. TAUs have been located and characterized in the N-terminal domains of the GR and PR. The N-terminal TAU of the human GR has been delineated to the central part of the GR N-terminal domain (amino acids 77-262). In contrast to our observations, the core unit of this TAU consists of only 40 to 60 amino acids (Dahlman-Wright et al. 1994) and the same region is responsible for the wild type GR and the constitutively active GR functioning (Hollenberg & Evans 1988). The TAU, responsible for the transactivating capacity of the human PRb, is located in the last 90 amino acids of the N-terminal domain and its activity can be modulated by the first N-terminal 164 amino acids (Meyer et al. 1992). These studies established that the location of the TAU in the various steroid receptors differs, and that there is almost no evidence for sequence or structural homology. The sequences characterizing TAUs might represent interfaces that function by direct or indirect binding to general- and/or specific transcription factors (Ptashne 1988; Mitchell & Tjian 1989; Ing et al. 1992; Gill & Tjian 1992; Tjian & Maniatis 1994). However, the mechanism by which TAUs can regulate gene transcription is still largely unknown.

## MATERIALS AND METHODS

### Cell culture, transfection and CAT assay

COS-1 and HeLa cells were maintained in Eagle's Minimum Essential Medium supplemented with 5% fetal calf serum (FCS), nonessential amino acids and antibiotics. The cells were cultured in 6 cm Ø dishes ( $3 \times 10^5$  cells /dish) and transfected using the calcium-phosphate precipitation method (2.5 µg AR plasmid, 2.5 µg reporter plasmid and 5 µg pTZ19 carrier plasmid /dish) (Chen & Okayama 1987). The cells were plated and transfected in medium containing 5% dextran/charcoal-treated fetal calf serum (ccFCS). Twenty four hours prior to cell-lysis, cells were incubated with 1 nM R1881 (17 $\alpha$ -methyl-trienolone, NEN, Boston, MA). The CAT assays were performed essentially as described previously (Seed & Sheen 1988).

Transfections of complete series of AR mutants were performed 5 times in duplo, using at least two independent plasmid extracts. In each experiment, CAT activities were corrected for the reporter background and the percentage relative to that of the total number of counts in the assay was determined for each mutant. The means ( $\pm$ SEM) were calculated and the starting mutants (AR0, AR5 or AR4G) were set to 100%.

### AR expression vectors and reporter constructs

A human AR cDNA expression vector (pAR0) was constructed using the SV40 early promoter and the rabbit  $\beta$ -globin polyadenylation signal (Brinkmann et al. 1989). The AR cDNA encodes 910 amino acids of which the polymorphic glutamine stretch contains 20 Gln residues and the glycine stretch consists of 16 Gly residues. Construction of AR mutants pAR22 and pAR61 has been described (Jenster et al. 1991). The AR mutants pAR3, pAR9 (deletion codons 46-101), pAR10 (deletion codons 100-142), pAR11 (deletion codons 139-188), pAR19 (deletion codons 186-244 and an additional mutation in codon 249 resulting in Glu to Val substitution), pAR34 (deletion codons 626-654), pAR53 (deletion codons 550-593), and pAR62 (deletion codons 485-528) were constructed by site directed mutagenesis, using PCR DNA amplification techniques (Higuchi et al. 1988). In the pAR3 plasmid the unique Xba I, Cla I and Mlu I restriction sites were inserted in between the second and third codon after the ATG, resulting in five additional codons coding for amino acids SRSIR. In the AR mutants pAR9, pAR10, pAR11, pAR19, pAR34, pAR53 and pAR62, a unique Cla I site was introduced in the same matching frame in all constructs (identical to the Cla I site present in pAR22 and pAR3). In this way, parts of the AR cDNA can easily be swapped between these mutants resulting in all kinds of in frame deletions or duplications. Introduction of the Cla I site (in the format: \*\*A-TCG-AT\*) will always encode the Ser-Met or Ser-Ile amino acids. This motif was not always present on the borders of the deletion. Therefore, the mutants pAR9, pAR10, pAR34 and pAR53 contain additional modifications on the borders of the deletion (for details see primers). In several AR mutant constructs (pAR3, pAR9, pAR10, pAR34 and pAR53) it was chosen to introduce a dam methylation site (GATC) to enable Cla I to digest when the plasmid was multiplied in the dam positive *E. coli* strain DH5 $\alpha$ . This inability to digest was easily abolished when the plasmid was multiplied in the dam negative *E. coli* strain GM121.

## Appendix paper IV

The following primers, introducing the insertion or deletion were used:

AR3: 5'CAAGCTCAAGGATGGAATCTAGATCGATACGCGTGCAGTTAGGGCTG 3'

AR9: 5' AGGCACCCAGAGGCCGCATCGATCACAGGCTACCTGGTC 3'

AR10: 5' CCAAGCCCATCGTAGATCGATGCCGAGCAGCTGCCA 3'

AR11: 5' GAGCCGCCGTGGCCGCATCGATGCAACTCCTCAGC 3'

AR19: 5' ACATCCTGAGCGAGGCATCGATGGGCTGGGTGTGG 3'

AR34: 5' GCTGAAGAACTTGGATCGATTGAAGGCTATGAATG 3'

AR53: 5' CTTTCCACCCCAGAGATCGATTGATAAAATTCCGA 3'

AR62: 5' GCTGGCGGGCCAGGAATCGATGCGTTTGGAGACTG 3'

All these AR mutants were sequenced to verify the correct reading frame before they were used as starting constructs for the following AR mutants. pAR110 was constructed by removing the 2.8 kb Cla I/BamH I fragment, containing most of the AR cDNA, from pAR3 and insertion of the 2.5 kb Cla I/BamH I fragment from pAR9 into pAR3. In the same way pAR120 (combination of pAR3 and pAR10), pAR121 (combination of pAR3 and pAR11), pAR123 (combination of pAR3 and pAR22), pAR104 (combination of pAR3 and pAR62), pAR60 (combination of pAR22 and pAR62) and pAR55 (combination of pAR9 and pAR19) were constructed. pAR113 was constructed by removal of the internal fragment of pAR62 digested with Rsr II and Cla I. Ligation of the blunt-ended plasmid resulted in an in-frame deletion. pAR130 and pAR131 were constructed by combining the deletions of pAR110 and pAR62, pAR110 and pAR113, respectively.

pAR5 was constructed by ligation of a Xba I linker (Promega, Madison, WI) that contains an in-frame stop codon into the blunt-ended Cla I site of pAR34. pAR124, pAR127, pAR106, pAR99, pAR126, pAR98, pAR100, pAR115, pAR117 and pAR132 were constructed by introduction of the deletion present in pAR120, pAR121, pAR123, pAR62, pAR113, pAR60, pAR55, pAR22, pAR61 and pAR104, respectively, into pAR5. pAR105 was constructed by combining the pAR3 and pAR100 mutants as described above. pAR128 was constructed by combining the deletions of pAR106 and pAR99.

pAR0G was constructed by insertion of the Cla I digested PCR fragment encoding the GAL4 DNA binding domain (GDBD, amino acids 1-147) into the Cla I digested pAR53 vector. The orientation and sequence of the PCR insert was checked. The following primers were used on the pSG424 plasmid (Sadowski & Ptashne 1989) template to obtain the correct GAL4 DBD fragment:

GDBD-A (5' PRIMER): 5' CAAGCCTCCTGATCGATGAAGCTACTG 3'

GDBD-B (3' PRIMER): 5' CCCGGGAATTCCATCGATACAGTCAAC 3'

pAR4G was constructed by ligation of a Xba I linker that contains an in-frame stop codon into the blunt-ended Cla I site of pAR0G (multiplied in *E.coli* DH5 $\alpha$ ). Please notice that pAR0G contains two Cla I sites but that only the Cla I site 5' of the GDBD insert is dam methylated. pAR94G, pAR107G, pAR105G, pAR106G, pAR99G, pAR113G, pAR98G, pAR100G, pAR115G and pAR117G were constructed by introduction of the deletion present in pAR124, pAR127, pAR105, pAR106, pAR99, pAR126, pAR98, pAR100, pAR115 or pAR117, respectively, into pAR4G. pAR96G, pAR91G, pAR85G and pAR84G were constructed by combining the deletions of pAR94G and pAR99G, pAR107G and pAR99G, pAR94G and pAR113G, pAR107G and pAR113G, respectively.

All mutants were expressed in COS-1 cells and the AR proteins were analyzed by Western blotting (Jenster et al. 1991). Using the antibodies SP197, SP061, SP066, F52.24.4, F39.4.1 (Kuiper et al. 1993; Zegers et al. 1991; van Laar et al. 1989; Veldscholte et al. 1992a) and 2GV3/3GV2 (White et al. 1992) all mutants except AR106 could be visualized. The expression levels of the various mutants was comparable. AR106 could only be detected using the F52.24.4 antibody. This monoclonal antibody is not specific enough to use directly for development of Western blots loaded with whole cell lysates, since it recognizes many zinc-finger containing proteins. We were unable to detect AR106 after immunoprecipitation most likely because AR106 (calculated molecular weight is approximately 28 kDa) co-migrated with the light chain of the F52.24.4 antibody.

Additionally, the key mutants pAR0, pAR5, pAR0G and pAR4G were checked for nuclear localization (Jenster et al. 1993).

The pG29GtkCAT ((GRE)<sub>2</sub>tkCAT) reporter plasmid (Schüle et al. 1988) contains two progesterone/glucocorticoid-responsive elements (PRE/GRE's) in front of the thymidine kinase (tk) promoter linked to the CAT gene. The (UAS)<sub>2</sub>tkCAT reporter (kindly provided by Dr.R Renkawitz and Dr. M. Muller) is in principle the same except that the two PRE/GRE's are exchanged by two upstream activating sequences (UAS) to which the GAL4 DNA binding domain can bind (Banahmad et al. 1991). The G5E1bCAT reporter contains 5 UAS's in front of the E1b TATA-box linked to the CAT gene (Lillie & Green 1989).

### ACKNOWLEDGEMENT

We thank R. Renkawitz and M. Muller for the pG29GtkCAT and UAS-29-UAStkCAT constructs, P. Chambon for the 2GV3 and 3GV2 antibodies, J.A. Grootegoed for helpful discussions and M. Tascilar and G. Meijer for technical assistance.

This work was supported by the Netherlands Organization for Scientific Research (NWO) through GB-MW (Medical Sciences).



**CHANGES IN ABUNDANCE OF ANDROGEN RECEPTOR ISOTYPES: EFFECTS  
OF LIGAND TREATMENT, GLUTAMINE-STRETCH VARIATION,  
AND MUTATION OF PUTATIVE PHOSPHORYLATION SITES**

Guido Jenster<sup>1</sup>, Petra E. de Ruiter<sup>1</sup>, Hetty A.G.M. van der Korput<sup>2</sup>, George G.J.M. Kuiper<sup>1</sup>,  
Jan Trapman<sup>2</sup> and Albert O. Brinkmann<sup>1</sup>

Departments of Endocrinology & Reproduction<sup>1</sup> and Pathology<sup>2</sup>,  
Erasmus University Rotterdam, Rotterdam, The Netherlands

Submitted

### ABSTRACT

The SDS-polyacrylamide gel electrophoresis (SDS-PAGE) migration pattern of wild type and mutated human androgen receptors (ARs) expressed in COS-1 cells was analyzed. In the absence of hormone, the wild type AR migrated as a closely spaced 110-112 kDa doublet. Alkaline phosphatase treatment resulted in a single 110 kDa band showing that the 112 kDa upshift reflects receptor phosphorylation. Deletion of the N-terminal amino acids 46-101 or 100-142 resulted in mutant ARs migrating as single protein bands. Three consensus phosphorylation sites in this region were substituted and the resulting mutated proteins were analyzed. Two Ser-Pro directed kinase consensus sites at positions Ser-80 and Ser-93 were both necessary for the AR 112 kDa upshift. Substitution of the putative casein kinase II Ser-118 site had no effect on the AR migration pattern. Surprisingly, deletion of the glutamine repeat, located directly N-terminal of the Ser-Pro sites, resulted also in an AR single form. Lengthening of the glutamine repeat caused an increase in the spacing between the two isotopes of the doublet, showing that the number of glutamine residues determines the extent of the upshift. Hormone treatment induced an extra isotype with an apparent molecular mass of 114 kDa, resulting in a 110-112-114 kDa AR triplet. The hormone-induced upshift was dependent on the Ser-80 consensus phosphorylation site. Mutations in the DNA binding domain caused a different distribution of receptor protein over the three AR isotopes. In the presence of ligand, the amount of the hormone-induced 114 kDa isotype was reduced as compared to the 110- and 112 kDa isotopes. Reduction of the amount of hormone-induced isotype was also observed for an AR mutant with an intact DNA binding domain but incapable of activating transcription due to a deletion of a 244-528 fragment in the N-terminal domain. This suggests that part of the hormone-induced phosphorylation occurs following DNA binding and during or following transcription regulation. These findings also apply to AR mutants which lack the steroid binding domain. The constitutively active mutant AR13 migrated as three separable isotopes. Mutations in the DNA binding domain or reduction of the transactivating capacity due to a deletion in the N-terminal domain, resulted in a reduced amount of the slowest migrating isotype. Although loss of transactivating capacity seems to impair phosphorylation, substitution of the Ser-80, Ser-93 or Ser-118 by either alanine or glycine residues, had no significant effect on the capacity of the respective mutants to activate transcription of a (GRE)<sub>3</sub>tkCAT reporter in HeLa cells.

### INTRODUCTION

The androgen receptor (AR) is a member of the superfamily of ligand-dependent transcription factors, which interacts through a ligand-dependent mechanism with specific DNA sequences, thereby regulating transcription of target genes (Evans 1988; Green & Chambon 1988; Truss & Beato 1993). These nuclear receptors show a high level of molecular identity and the arrangement of the different, separate functional/structural domains is essentially the same: an N-terminal modulating domain, a central DNA binding domain, a hinge region, and a C-terminal ligand binding domain. The N-terminal region of the different members of the nuclear receptor family is highly variable in size and in amino acid composition, but their DNA and ligand binding domains have a high degree of sequence homology (Laudet et al. 1992).

Many transcription factors, including steroid hormone receptors are phosphoproteins



(reviewed in Moudgil 1990; Hunter & Karin 1992; Meek & Street 1992; Ortí et al. 1992; Takimoto & Horwitz 1993; Kuiper & Brinkmann 1994a). Mutational analysis, phospho-amino acid analysis, tryptic phospho-peptide mapping and microsequencing have identified phosphorylation sites in the glucocorticoid- (GR) (Bodwell et al. 1991), progesterone- (PR) (Denner et al. 1990b; Poletti & Weigel 1993), estrogen- (ER) (Ali et al. 1993; Le Goff et al. 1994) and vitamin D<sub>3</sub> receptor (VDR) (Hsieh et al. 1991). Phosphorylation mainly occurs in the N-terminal domain of these receptors, although sites in the hinge region (chicken PR [Denner et al. 1990b]), hinge region and ligand binding domain (VDR [Brown & DeLuca 1991; Hsieh et al. 1991]) and in the DNA binding domain (VDR [Hsieh et al. 1991]) have also been reported. Almost all characterized phosphorylation sites are on serine residues (Dalman et al. 1988; Sheridan et al. 1988; Denner et al. 1990b; Bodwell et al. 1991; Hsieh et al. 1991; Washburn et al. 1991; Denton et al. 1992; Ali et al. 1993; Poletti & Weigel 1993; Le Goff et al. 1994). In some cases minor phosphorylation on threonine residues (Hoeck & Groner 1990; Bodwell et al. 1991) has been observed, but rarely on tyrosine residues (Migliaccio et al. 1986; Auricchio 1989). With respect to the human AR, no specific phosphorylation sites have been identified as yet. From limited proteolysis studies of <sup>32</sup>P-labeled AR, it is evident that phosphorylation occurs in the first 300 amino acid residues of the N-terminal domain (Kuiper et al. 1993). Phospho amino acid analysis revealed that AR phosphorylation takes place on serine residues (G.G.J.M. Kuiper personal communications).

Several different protein kinases (protein kinase A, protein kinase C, casein kinase II, Ser-Pro directed kinases) have been described to phosphorylate steroid receptors (reviewed in Moudgil 1990; Ortí et al. 1992). In the region in which the AR is shown to be phosphorylated, several consensus sites for these kinases can be found.

Upon incubation of cells in culture with the respective ligands, steroid receptors become hyperphosphorylated (Ortí et al. 1989; Brown & DeLuca 1990; Moudgil 1990; Chachereau et al. 1991; Ortí et al. 1992; van Laar et al. 1991; Washburn et al. 1991; Beck et al. 1992; Denton et al. 1992; Kuiper et al. 1993). Since this rapid additional phosphorylation has been associated with hormone binding, nuclear import, DNA binding and transcription activation, the importance of receptor phosphorylation has often been emphasized (reviewed in Moudgil 1990; Ortí et al. 1992; Takimoto & Horwitz 1993). Mutational analysis of phosphorylation sites in the human ER and human VDR indeed showed that particular sites are important for the receptor's capacity to activate transcription (Hsieh et al. 1991; Ali et al. 1993; Le Goff et al. 1994). However, site-directed mutagenesis of the phosphorylation sites in the mouse GR had no, or only minor effect on transcription activity (Mason & Housley 1993).

For several nuclear receptors (PR [Sheridan et al. 1988; Sheridan et al. 1989; Beck et al. 1992], ER [Washburn et al. 1991], AR [Krongrad et al. 1991; Kuiper et al. 1991] and VDR [Jurutka et al. 1993]) phosphorylation is accompanied by a characteristic decrease in electrophoretic mobility (upshift) during SDS-PAGE analysis, resulting in the separation of different isotypes. Immediately following synthesis, the nascent human AR has an apparent molecular mass of 110 kDa and matures into two bands (110-112 kDa) over the ensuing 30 minutes. Alkaline phosphatase treatment restored the 110 kDa single protein band showing that the upshift to 112 kDa reflects phosphorylation (Kuiper et al. 1991).

For the human PR it has been shown that hormone-stimulated hyperphosphorylation induced additional receptor isotypes. Both hyperphosphorylation and coincident PR upshifts during SDS-PAGE were partially inhibited in PR mutants unable to bind DNA or by

**Appendix paper V**

progesterone antagonists that prevented PR-DNA interaction (Bagchi et al. 1992; Takimoto et al. 1992).

In our quest to locate and analyze functional domains of the human AR, many deletion mutants have been generated and tested for their correct length and expression level by SDS-PAGE (Jenster et al. 1991; Jenster et al. 1993). In these studies divergences from the wild type AR doublet migration pattern were observed. In order to elucidate the nature of the doublet migration pattern and use it as a marker for the identification of phosphorylation sites, deletion and substitution mutants were constructed and analyzed by SDS-PAGE and immunoblotting.

**RESULTS**

**AR migration pattern in the absence of hormone**

As shown previously, the wild type AR (length 910 amino acids) expressed in LNCaP cells (lymph node carcinoma of the prostate) or in transfected COS-1 cells, migrated as two protein bands (110-112 kDa) when analyzed by Western blotting (van Laar et al. 1990; Jenster et al. 1991; Kuiper et al. 1991).

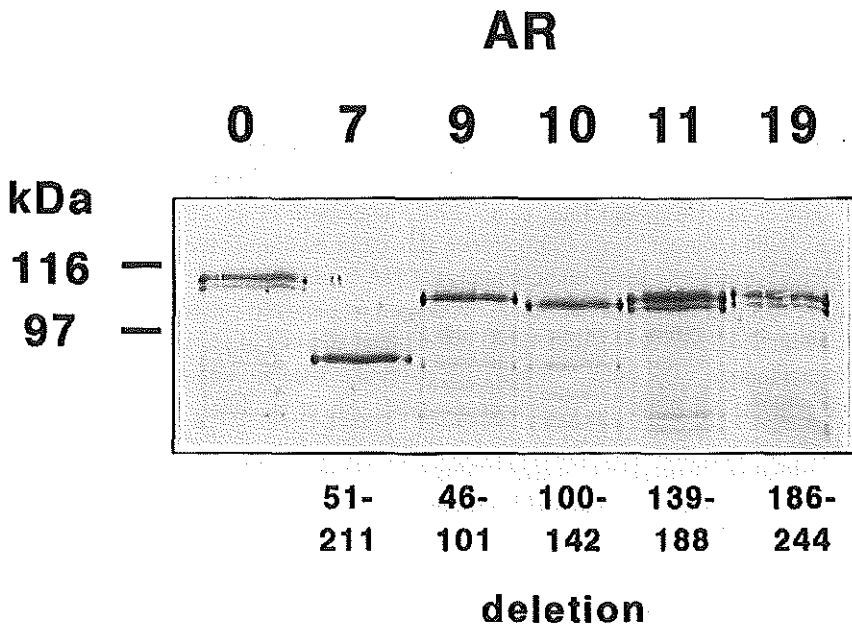


Figure 1 Immunoblot of the wild type AR and AR deletion mutants. AR proteins expressed in COS-1 cells in the absence of hormone were immunoprecipitated with the monoclonal antibody F39.4.1, separated by 7% SDS-PAGE, blotted, and immunostained with the polyclonal antibody Sp061. Molecular weight markers (kDa) were run on a parallel lane and their positions are indicated on the left. The AR mutant number and respective residues deleted are indicated.



no effect on the migration pattern of the two isotypes (Figure 3). To demonstrate that phosphorylation causes the AR mobility upshift, alkaline phosphatase (AP) treatment of cytosolic extracts was performed. As shown in Figure 3, the upper 112 kDa band of the wild type AR0 reverted into the 110 kDa isotype upon alkaline phosphatase treatment, while the 110 kDa ARS80G and ARS93A single protein bands were practically unaffected.

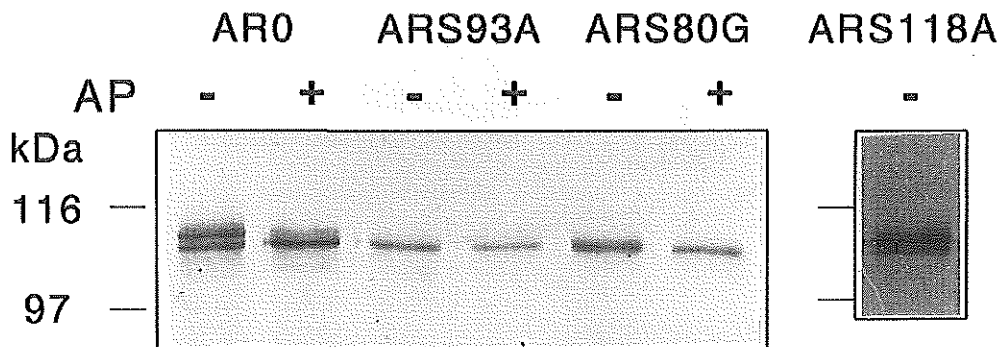


Figure 3 Immunoblot showing dephosphorylation of wild type and mutated ARs expressed in the absence of hormone.

Cytosols were prepared from COS-1 cells expressing AR proteins in the absence of hormone and incubated with (+) or without (-) alkaline phosphatase (AP). AR proteins were immunoprecipitated with the monoclonal antibody F39.4.1, separated by 7% SDS-PAGE, immunoblotted and visualized using the Sp061 antibody. Molecular mass markers (kDa) are indicated on the left.

In addition to the serine mutations, the influence of the length of the flanking glutamine stretch (Figure 2; residues 58-77) on the SDS-PAGE migration pattern of the AR was investigated. Surprisingly, deletion of the glutamine repeat resulted in an AR, migrating as a single 105 kDa band (Figure 4). Lengthening the stretch to 7, 12, 16, 20, 29 or 48 glutamine residues revealed not only an expected increase in size of the proteins, but also a gradual appearance of two bands with increasing spacing between the two AR isotypes (Figure 4).

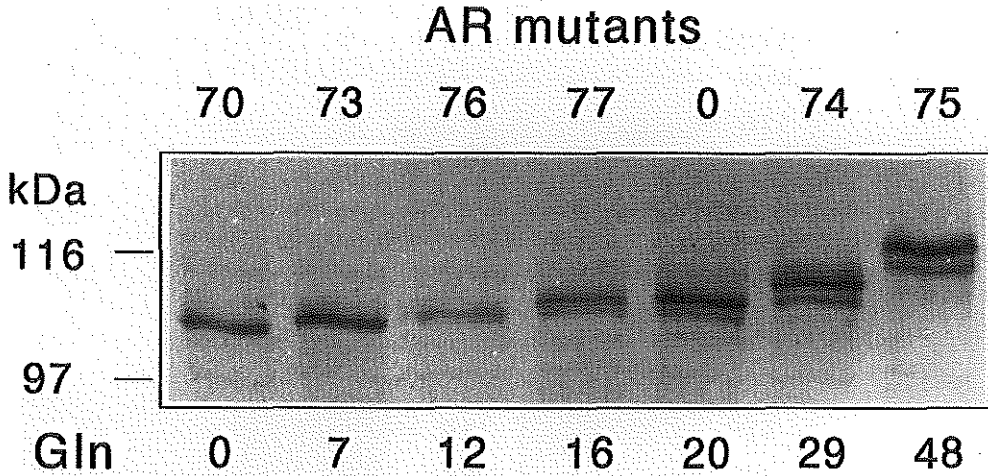


Figure 4 Immunoblot of AR proteins containing glutamine repeats with different lengths. AR proteins expressed in COS-1 cells in the absence of hormone were immunoprecipitated with the monoclonal antibody F39.4.1, separated by 7% SDS-PAGE, blotted, and immunostained with the polyclonal antibody Sp061. Molecular weight markers (kDa) were run on a parallel lane and their positions are indicated on the left. Mutant number and length of the glutamine stretch (Gln) are indicated.

#### Hormone-induced AR migration pattern

Hormone (1 nM R1881) added during 16 hours to the COS-1 cells expressing the wild type AR, induced an additional protein form, resulting in three separable receptor isotypes (Figure 5). The same hormone-induced triplet was observed for the ARS118A mutant, showing that Ser-118 is not essential for the different AR protein migration upshifts in the absence or presence of ligand (Figure 5). In contrast, the ARS80G migration pattern revealed no additional significant upshift upon hormone exposure. The single protein bands of ARS93A and AR70 ( $\Delta$ Gln stretch) both converted into a doublet upon R1881 treatment although the hormone-induced upshift of ARS93A was faint (Figure 5).

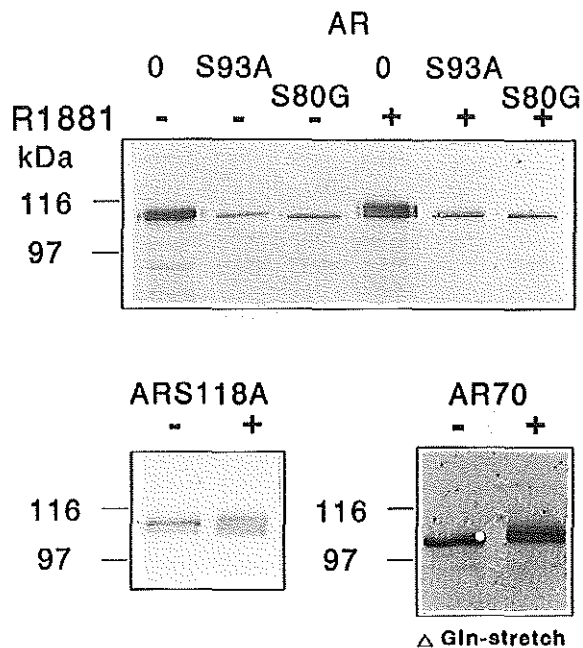


Figure 5 Hormone-induced protein upshifts of wild type AR and AR mutants. AR proteins expressed in COS-1 cells and incubated in the absence (-) or presence (+) of 1 nM R1881 for approximately 16 hours, were separated by 7% SDS-PAGE, blotted, and immunostained with the polyclonal antibody Sp061. Molecular weight markers (kDa) were run on a parallel lane and their positions are indicated on the left of each blot.

### AR isotypes and transcription activity

To pin-point the step in the AR activation cascade during which the hormone-stimulated phosphorylation takes place, a selected set of receptor mutants was analyzed in more detail. Figure 6 summarizes the subcellular localizations and the transactivation capacities of the wild type and mutant ARs studied. All of these AR mutants were exclusively or predominantly nuclear in the presence of hormone (see also Jenster et al. 1993). As has been described previously, part of the AR64 protein population formed large clusters in the nucleus upon ligand treatment (Jenster et al. 1993). AR30.3 was partially cytoplasmic because of the substitution of the two basic arginine and lysine residues which are part of the nuclear localization signal (Jenster et al. 1993) (Figure 6).

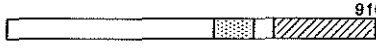
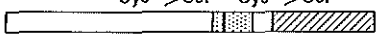
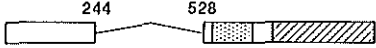

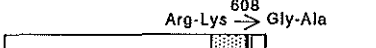
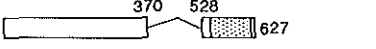
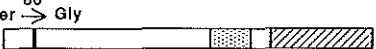


		R1881		subcellular localization		CAT activity %	
		-	+	-	+	-	+
AR0		C>N	N	<1	100		
AR64		C>N	N*	<1	<5		
AR60		C>N	N	<1	<5		
AR13			N		89 ±11		
AR30.3			N>C		<5		
AR126			N		<5		
ARS80G		C>N	N	<1	126 ±15		
ARS93A		C>N	N	<1	113 ±14		
ARS118A		C>N	N	<1	96 ±20		

Figure 6 Subcellular localization and transcription activation of the wild type AR and AR mutants. Nuclear (N) and cytoplasmic (C) localization of AR proteins expressed in COS-1 cells in the absence (-) or presence (+) of 1 nM R1881. The receptor mutants were visualized using the F39.4.1. antibody except for AR60, for which the Sp197 antibody was used. In the presence of hormone, large AR64 clusters were formed in the nucleus (indicated by N\*) (Jenster et al. 1993). CAT activity was determined in cell lysates of transfected HeLa cells cultured in the absence (-) or presence (+) of 1 nM R1881. Activities were corrected for the (GRE)<sub>3</sub>kCAT background and the percentage ( $\pm$ SEM) relative to that of the wild type AR0 of four independent assays are shown.

AR64, which is unable to bind DNA due to substitutions of two essential cysteine residues in the first zinc finger, migrated as a doublet in the absence of R1881 (Figure 7). Hormone treatment of AR64 generated less of the slowest migrating, 114 kDa form as compared to its 110- and 112 kDa isotypes, suggesting that DNA binding is necessary for part of the ligand-stimulated phosphorylation (Figure 7). In order to be able to investigate further whether DNA binding or transcription activation is needed for phosphorylation, an AR mutant was required which is still able to bind DNA, but incapable of activating transcription. Because a deletion should be far from the proposed phosphorylation sites, AR60 which lacks amino acids 244-528, was used for these experiments (Figure 6). The AR60 migration pattern showed a

Appendix paper V

doublet appearance when expressed in the absence of ligand. Hormone exposure resulted in an additional faint third isotype (partial triplet), indicating that not DNA binding but transcription activation would be the step during which part of the hormone-stimulated phosphorylation takes place (Figure 7).

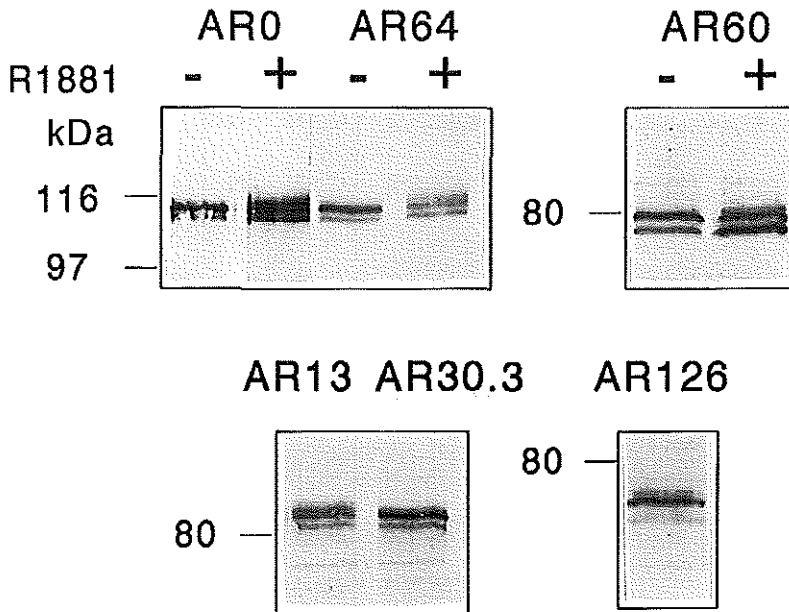


Figure 7 Hormone-induced protein upshifts of the wild type AR and AR mutants. AR proteins expressed in COS-1 cells and incubated in the absence (-) or presence (+) of 1 nM R1881 for approximately 16 hours, were separated by 7% SDS-PAGE, blotted, and immunostained with the polyclonal antibody Sp061. Molecular weight markers (kDa) were run on a parallel lane and their positions are indicated on the left for each blot.

Additional evidence was provided by experiments with the constitutively active mutant AR13, which lacks the ligand binding domain. AR13 migrated as three separable isotopes (Figure 7). Substitution of the residues Arg-Lys by Gly-Ala at position 608-609 (AR30.3), inactivated the constitutive receptor, most likely due to the inability to bind DNA (Figure 6). Analogous to the AR0 versus AR64 SDS-PAGE protein pattern, the distribution of receptor protein over the three isotopes was different for AR13 compared to AR30.3. An equal distribution over the isotopes is observed in the AR13 migration pattern. In contrast, there is considerably more of the second upshift and less of the third, slowest migrating, isotype in the AR30.3 sample (Figure 7). A similar pattern is observed for AR126. This mutant contains an intact DNA binding domain, but is hardly able to activate transcription due to a deletion in the N-terminal region (Figure 6). This again suggests that part of the hormone-induced receptor phosphorylation occurs during or following transcription activation.

To test whether substitution of the proposed phosphorylation sites would affect



transactivation, the capacity of ARS80G, ARS93A and ARS118A to induce transcription was investigated. Cotransfection experiments of these AR mutants with the (GRE)<sub>2</sub>tkCAT reporter (pG29GtkCAT [Schüle et al. 1988]) in HeLa cells showed that their transcription activation capacity was comparable to the wild type AR activity (Figure 6). Since Gln-stretches and glutamine-rich domains have been implicated in transcription activation and since the length of the poly-glutamine stretch considerably influences the isotype pattern on SDS-PAGE, AR70 (lacking the Gln-stretch) and AR75 (containing 48 glutamines) were tested for their capacity to activate transcription. The wild type AR (20 Gln), AR70 and AR75 were cotransfected with either the MMTV-CAT or the pPA2-CAT reporter construct in COS-1 cells. Both reporters can be induced by androgens when cotransfected with AR0. The MMTV-CAT contains the MMTV-LTR promoter which harbours 4 glucocorticoid response elements, while the pPA2-CAT reporter accommodates 630 base pairs of the promoter region of the prostate specific antigen (PA) gene containing 1 androgen response element (Riegman et al. 1991). Deletion of the glutamine stretch resulted in an increased activity of AR70, but only when tested on the pPA2-CAT reporter (Figure 8). Lengthening of the repeat to 48 residues decreased the transactivation capacity of AR75 on both promoters, showing that the (extended) Gln-stretch inhibits transcription.

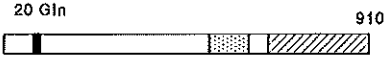
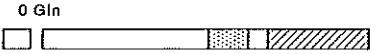
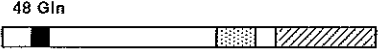
		CAT activity %			
		MMTV-CAT		pPA2-CAT	
		-	+	-	+
		R1881			
AR0		<1	100	<1	100
AR70		<1	101 ± 3	<1	136 ± 11
AR75		<1	62 ± 2	<1	44 ± 8

Figure 8 *Functional analysis of AR mutants with different lengths of the glutamine stretch. Transcriptional activity was examined by cotransfection of AR0 (20 Gln), AR70 (0 Gln) and AR75 (48 Gln) expression plasmids with the MMTV-CAT or pPA2-CAT reporter. CAT activity was determined from cell lysates of transfected COS-1 cells cultured in the absence or presence of 1 nM R1881. Activities were corrected for the reporter background and the mean ( $\pm$ SEM) of the R1881 treated samples of 5 independent assays are presented as a percentage relative to that of the wild type AR0.*

## DISCUSSION

Several reports have shown that there is a causal link between phosphorylation and the electrophoretic mobility during SDS-PAGE of the PR (Sheridan et al. 1988; Sheridan et al. 1989; Beck et al. 1992), ER (Washburn et al. 1991), AR (Krongrad et al. 1991; Kuiper et al. 1991) and VDR (Jurutka et al. 1993). Therefore, analysis of protein isotypes could indirectly

provide information on receptor phosphorylation sites and function. In the present study the SDS-PAGE protein patterns of wild type and mutated human ARs and their ability to activate transcription were investigated.

Narrowing down the region, responsible for AR doublet appearance, to residues 46-139, reduced the more than 10 possible N-terminal serine phosphorylation motifs to only three obvious consensus serine phosphorylation sites. Substitution analysis revealed that at least the two Ser-Pro directed kinase sites (Ser-80 and Ser-93) are essential for the 112 kDa AR upshift. The putative casein kinase II phosphorylation site at position 118 is not important for the AR migration pattern. Besides the potential Ser-80 and Ser-93 sites, it cannot be excluded that other phosphorylation sites are additionally essential for the 112 and 114 kDa upshifts. The observation that the unliganded AR10 (lacking amino acids 100-142) migrated as a single protein band, suggests that except for Ser-118, other serine residues in this region contribute to an upshift when phosphorylated.

Interestingly, deletion of the glutamine stretch (a repeat of 20 glutamine residues in the ARO protein), also resulted in a single receptor band. It is unlikely that deletion of the glutamine residues directly affected phosphorylation of the neighbouring phosphorylation sites and therefore resulted in a single band, since lengthening of the stretch to different sizes increased the spacing between the two AR isotypes. If phosphorylation of surrounding sites was reduced, one would rather expect the amount of the upshifted band to vary and the spacing to be constant. Two possible explanations for phosphorylation-induced AR doublet appearance in the absence of hormone are proposed. Firstly, phosphorylation on sites Ser-80 and Ser-93 results in a conformational change in the neighbouring glutamine repeat causing the migration upshift visible as a 112 kDa isotype. Deletion of the glutamine stretch or elimination of the phosphorylation sites would both result in a single AR protein band. Shortening of the glutamine stretch would reduce the extend of the conformational change and consequently the spacing between the receptor isotypes. Secondly, Ser-80 and Ser-93 phosphorylation might retard the AR protein migration by decreasing SDS-binding to the AR. The negatively charged detergent SDS provides the negative charge to the SDS-protein complex and therefore largely determines the migration pattern of proteins in the SDS-PAGE system. Possibly, SDS molecules that bind to the glutamine stretch will be supplanted by the negative phosphate of the phosphorylated Ser-80 and Ser-93. Phosphorylation of Ser-80 and Ser-93 in AR mutants containing a reduced glutamine repeat length might supplant fewer SDS molecules resulting in a decreased difference in negative charge between the two isotypes which subsequently results in a decreased spacing between the AR isotypes.

Deletion of the Gln-repeat (AR70) resulted in an increased activity when tested in COS-1 cells on a reporter containing part of the prostate specific antigen (PA) promoter (pPA2-CAT). The same AR70 mutant, however, had wild type activity when tested on the MMTV-LTR reporter (MMTV-CAT). Extension of the Gln-repeat from 20 to 48 residues resulted in a decreased AR activity on both promoters to approximately 50% as compared to the wild type AR, which confirms the observations of Mhatre and co-workers (Mhatre et al. 1993). This indicates that the glutamine stretch inhibits AR transcription activation. Lengthening of the Gln repeat to more than 40 residues is associated with Kennedy's disease, an X-linked neurodegenerative disorder, characterized by a slowly progressing muscle weakness (La Spada et al. 1991).

The androgen R1881 induced a second upshift resulting in three AR isotypes (110-112-114 kDa). These ligand-induced changes could be due to additional phosphorylation and/or a re-location of phosphorylated sites. The R1881-induced upshift is dependent on the Ser-80 consensus phosphorylation site. Substitution of the Ser-93 consensus phosphorylation site, resulted in a reduction of the amount of the ligand-induced isotype.

Interestingly, the migration pattern of AR13 consisted of three major isotypes, which might represent the triplet found for the wild type AR0 when expressed in the presence of ligand. So besides being constitutively active with respect to transcription, AR13 seems also to be constitutively phosphorylated.

Analogous to the human PR (Bagchi et al. 1992; Takimoto et al. 1992) and mouse ER (Lahooti et al. 1994), AR DNA binding is important for the hormone-induced 114 kDa upshift. The immunoblot pattern of AR64 (unable to bind DNA because of substitution of two essential cysteines in the first zinc-finger) revealed a different distribution of receptor protein over the three isotypes as compared to the AR0 triplet. There was less of the slowest, 114 kDa isotype present in the AR64 sample. A comparable observation was made when a mutation was introduced in the DNA binding domain of the constitutively active AR13 (AR30.3). The slowest migrating isotype of AR30.3 became less intense as compared to the other isotypes. Both AR64 and AR30.3 contain mutations in the DNA binding domain and subsequently are unable to activate transcription. It could still be possible that the hormone-induced upshift is not only DNA dependent but occurs predominantly during or after transcription activation. The R1881-induced migration pattern of AR60 which contains an intact DNA binding domain, but is unable to transactivate due to the deletion of amino acids 244-528, was clearly aberrant. A comparable observation was made for AR126 which was only weakly constitutively active due to a N-terminal deletion ( $\Delta$ 370-528). Like found for AR60, its slowest migrating isotype was less intense. These observations suggest that part of the hormone-induced phosphorylation is not solely dependent on DNA binding, but rather occurs during or following transcription activation. However, it cannot be excluded that due to the large N-terminal deletion, conformational changes were introduced or phosphorylation sites erased.

Interestingly, although elimination of the transactivating capacity seemed to affect hormone-induced phosphorylation, substitution of the proposed phosphorylation sites did not affect the transactivating capacity of the AR. Similarly, Mason & Housley (Mason & Housley 1993) have shown that substitution of phosphorylation sites in the mouse GR had little effect on its capacity to activate transcription. On the other hand, phosphorylation sites in the human ER and human VDR have been shown to be important for the transactivating capacity of these receptors (Hsieh et al. 1991; Ali et al. 1993; Le Goff et al. 1994).

As has been observed for several other steroid receptors, there exists a significant discrepancy between the predicted molecular mass based on the AR sequence (98.4 kDa) and the AR protein size determined by SDS-PAGE (~110 kDa) (Table 1). The region responsible for this mobility retardation of approximately 12 kDa is largely deleted in AR7 ( $\Delta$ 51-211) (Table 1). The calculated molecular mass (81.2 kDa) and its SDS-PAGE size determination (~84 kDa) are comparable. Recently, it has been shown that the mobility retardation of the mouse GR can be attributed to an N-terminal region of approximately 200 amino acids containing all identified phosphorylation sites within the receptor (Hutchison et al. 1993). With the SDS-PAGE analysis of many different AR deletion mutants (Jenster et al. 1991; Jenster et al. 1993), it became clear that the first part of the AR N-terminal domain (amino acids 46-244)

## Appendix paper V

is responsible for the discrepancy between predicted and SDS-PAGE determined molecular mass. Comparable to the mouse GR, this region contains most, if not all AR phosphorylation sites (Kuiper et al. 1993). Deletion of this region in the human AR and the mouse GR resulted in the loss of the capacity to activate transcription (Danielsen et al. 1987; Jenster et al. 1991).

**Table 1** Summary of AR mutants showing the effect of the mutation on SDS-PAGE size and migration pattern.

AR mutant	Mutation	Calculated M kDa	SDS-PAGE M kDa	Migration pattern	
				-	+ R1881
AR0 (20Gln)	wild type	98.4	110	D	T
AR7	Δ 51-211	81.2	84	S	
AR9	Δ 46-101	92.3	104	S	S
AR10	Δ100-142	94.3	103	S	D
AR11	Δ139-188	93.6	103	D	
AR19	Δ186-244	92.6	104	D	
ARS80G	S80G	98.4	110	S	S
ARS93A	S93A	98.4	110	S	pD
ARS118A	S118A	98.4	110	D	T
AR64	C567S;C570F	98.4	110	D	pT
AR60	Δ244-528	70.2	76	D	pT
AR13	Δ653-910	68.5	84	T	T
AR30.3	Δ653-910	68.5	84	pT	pT
	RK608GA				
AR126	Δ370-528	50.4	60	pT	pT
	Δ627-910				
AR70 (0Gln)	Δ58-77	95.9	105	S	T
AR73 (7Gln)		96.8	107	"D"	
AR76 (12Gln)		97.4	108	"D"	
AR77 (16Gln)		98.0	109	"D"	
AR74 (29Gln)		99.6	112	D	
AR75 (48Gln)		102.0	114	D	

*Calculated molecular mass (M) is based on the amino acid residue constitution and SDS-PAGE M is an estimation of the size determined after gel electrophoresis. Migration patterns are indicated as S (single form), D (doublet), "D" (less spaced doublet), pD (partial doublet), T (triplet) and pT (partial triplet) in the absence (-) or presence (+) of 1 nM R1881.*

The basis for the molecular mass discrepancy is unknown although it is unlikely that it is caused by phosphorylation, since alkaline phosphatase treatment had no effect on the apparent 110 kDa molecular mass of the wild type AR determined by SDS-PAGE. Possibly, the region between amino acids 46 and 244 binds relatively few SDS molecules resulting in a less negatively charged protein and the observed mobility retardation during SDS-PAGE. So far, it

was not possible to attribute the mobility retardation to a smaller domain, since the SDS-PAGE molecular masses of AR9, AR10, AR11 and AR19 were all 10-12 kDa above the calculated sizes (Table 1).

It becomes evident that the N-terminal region between residues 46-244 in the hAR is of high interest since it is 1) essential for transcription activation (Jenster et al. 1991), 2) contains a glutamine stretch which is involved in the capacity of the AR to activate transcription, 3) harbours most of the AR phosphorylation sites (Kuiper et al. 1993), and 4) is responsible for the mobility retardation.

## EXPERIMENTAL PROCEDURES

### Cell culture, transfection and Western blotting

COS-1 cells were maintained and transfected in Eagle's Minimum Essential Medium (MEM) supplemented with 5% fetal calf serum, nonessential amino acids and antibiotics (medium). The cells were cultured in T150 cm<sup>2</sup> flasks ( $3 \times 10^6$  cells/flask) and transfected using the DEAE-dextran method. The COS-1 cells were incubated with 15 ml MEM containing 20 µg AR expression plasmid and 0.2 mg DEAE-dextran. After 6 hours, cells were treated with 15 ml MEM, 25% (v/v) glycerol for 1 minute, washed twice with MEM and re-incubated with medium for 2 days. Sixteen hours prior to cell lysis, medium was replaced and 1 nM R1881 (17 $\alpha$ -methyl-17 $\beta$  hydroxyestra-4,9,11-trien-3-one, NEN, Boston, MA) was added. Cells transfected in T150 cm<sup>2</sup> flasks were lysed in 1 ml lysisbuffer (40 mM Tris.HCl pH 7.0, 1 mM EDTA, 4% (v/v) glycerol, 10 mM dithiotreitol, 2% (w/v) SDS, 0.6 mM PMSF and 0.5 mM bacitracin). In most cases, part of the cell-lysis mix (approximately 20 µl) was mixed with 5 times sample buffer (200 mM Tris.HCl pH 6.8, 25% (v/v) glycerol, 10% (w/v) SDS, 50 mM dithiotreitol, 1% (w/v) bromophenolblue), heated, and loaded on a 7% SDS-PAGE gel as described previously (Kuiper et al. 1991). Sometimes AR protein was first immunoprecipitated using monoclonal antibody F39.4.1 coupled to goat anti-mouse agarose beads (Zegers et al. 1991). Western blotting and immunostaining were performed as described previously (Kuiper et al. 1991). The Sp061 (Zegers et al. 1991) and Sp197 (Kuiper et al. 1993) polyclonal anti-AR antibodies were used to visualize the different AR proteins.

### Alkaline phosphatase treatment

Cells transfected in T150 cm<sup>2</sup> flasks were suspended in 1 ml TGD buffer (40 mM Tris.HCl pH 8.5, 10% (v/v) glycerol, 1 mM dithiotreitol, 0.6 mM PMSF, 0.5 mM bacitracin, 0.2 mg/ml leupeptin). Cytosols (400 µl) were prepared and incubated for 60 minutes at 37 °C with 200 units of calf alkaline phosphatase (Boehringer Mannheim) as described previously (Kuiper et al. 1991).

### Chloramphenicol acetyl transferase (CAT) assays and subcellular localization

AR functional activity was assessed in HeLa or COS-1 cells by transient transfection using the calcium-phosphate precipitation method (Chen & Okayama 1987). Transfections of

## Appendix paper V

AR mutants were performed 4 or 5 times in duplo, using at least two independent plasmid extracts (Jenster et al. 1991). In each experiment, CAT activities were corrected for the reporter background and the percentage relative to that of the wild type AR (set to 100%) was determined. The means ( $\pm$  SEM) were calculated. CAT assays were performed as described previously (Jenster et al. 1991). For AR subcellular localization, COS-1 cells were cultured directly on microscope glass slides and transfected using the calcium-phosphate precipitation method. Fixation, permeabilization and immunocytochemistry of the transfected cells were performed essentially as described previously (Jenster et al. 1991; Jenster et al. 1993). The anti-AR monoclonal F39.4.1 and polyclonal Sp197 antibodies were used to visualize the different AR proteins.

### AR expression vectors

A human cDNA expression vector (pAR0) was constructed using the SV40 early promoter and the rabbit  $\beta$ -globin polyadenylation signal (Brinkmann et al. 1989). The AR cDNA encodes 910 amino acids of which the polymorphic glutamine-stretch contains 20 Gln residues and the glycine-stretch consists of 16 Gly residues. Construction of pAR7, pAR13, pAR22 and pAR64 has been described previously (Jenster et al. 1991; Jenster et al. 1993). pAR30.3 contains the same residue substitutions as pAR28.3 and was constructed in the same way (Jenster et al. 1993). The mutants pAR9 (deletion codons 46-101), pAR10 (deletion codons 100-142), pAR11 (deletion codons 139-188), pAR19 (deletion codons 186-244 and an additional mutation in codon 249 resulting in Glu to Val substitution), pAR34 (deletion codons 626-654), pAR62 (deletion codon 485-528), pARS80G (=pAR79), pARS93A (=pAR78), pARS118A (=pAR96) and pAR70 were constructed by site-directed mutagenesis, using PCR DNA amplification techniques (Higuchi et al. 1988). The following primers, introducing the deletion or substitution, were used:

AR9: 5' AGGCACCCAGAGGCCGCATCGATCACAGGCTACCTGGTC 3'  
AR10: 5' CCAAGCCCATCGTAGATCGATGCCGCAGCAGCTGCCA 3'  
AR11: 5' GAGCCGCCGTGGCCGCATCGATGCAACTCCTTCAGC 3'  
AR19: 5' ACATCCTGAGCGAGGCATCGATGGGCTGGGTGTGG 3'  
AR34: 5' GCTGAAGAACTTGGATCGATTGAAGGCTATGAATG 3'  
AR62: 5' GCTGGCGGGCCAGGAATCGATGCGTTTGGAGACTG 3'  
AR70: 5' GCCAGTTTGCTGCTGCTCGAGACTAGCCCCAGGCAG 3'  
AR78: 5' GAGGATGGTGCTCCCCAAGCC 3'  
AR79: 5' CAACAGACTGGCCCCAGGCAG 3'  
AR96: 5' CACAGCCGCAGGGCGCCCTGGAGTG 3'

pARS118A (pAR96) lacked one CAG codon at position 83 resulting in the deletion of one glutamine residue in a short glutamine stretch of normally 6 residues (Figure 2) (Faber et al. 1989).

In the AR mutants pAR9, pAR10, pAR11, pAR19, pAR34, and pAR62, a unique Cla I site was introduced in the same matching frame in all constructs (identical to the Cla I site present in pAR22 [Jenster et al. 1993]). In this way parts of the AR cDNA can easily be swapped between these mutants resulting in all kinds of in frame deletion or duplications.

pAR60 was constructed by removing the 1.74 kb Cla I/BamH I fragment, containing most of the AR cDNA, from pAR22 (deletion of codons 244-360) and insertion of the 1.24 kb Cla I/BamH I fragment from pAR62 into pAR22. pAR5 (deletion codons 627-910) was constructed by ligation of a Xba I linker (Promega, Madison, WI) that contains an in-frame stop codon into the blunt-ended Cla I site of pAR34. pAR113 (deletion codons 370-528) was constructed by removal of the internal fragment of pAR62 digested with Rsr II and Cla I. Ligation of the blunt-ended plasmid resulted in an in-frame deletion.

From pAR113, pAR126 was constructed by introducing the 370-528 deletion present in pAR113 in the pAR5 plasmid.

pAR75 was constructed by PCR amplification of part of the first exon containing the (CAG)<sub>47</sub>CAA stretch on genomic DNA isolated from a patient with Kennedy's disease and exchanged with the pAR0 sequences using the Sma I and Afl II restriction sites. pAR73, pAR76, pAR77 and pAR74 were constructed in the same way using genomic DNA with the respective (CAG)<sub>n</sub> repeat length from androgen insensitivity patients or cell line materials. All AR mutants were sequenced to verify the correct reading frame.

### ACKNOWLEDGEMENT

We thank Dr. R. Renkawitz for the pG29GtkCAT reporter, Dr. J.A. Grootegoed for helpful discussions and Dr. A. Danek for providing material of a patient with Kennedy's disease. This work was supported by the Netherlands Organization for Scientific Research (NWO) through GB-MW (Medical Sciences).

## References

## REFERENCES

- Adams CC, Workman JL 1993 Nucleosome displacement in transcription. *Cell* 72:305-308
- Addison C, Jenkins JR, Stürzbecher H-W 1990 The p53 nuclear localisation signal is structurally linked to a p34<sup>cdc2</sup> kinase motif. *Oncogene* 5:423-426
- Adler AJ, Danielsen M, Robins DM 1992 Androgen-specific gene activation via a consensus glucocorticoid response element is determined by interaction with nonreceptor factors. *Proc Natl Acad Sci USA* 89:11660-11663
- Akerblom IE, Slater EP, Beato M, Baxter JD, Mellon PL 1988 Negative regulation by glucocorticoids through interference with a cAMP responsive enhancer. *Science* 241:350-353
- Akner G, Mossberg K, Wikström A-C, Sundqvist K-G, Gustafsson J-A 1991 Evidence for colocalization of glucocorticoid receptor with cytoplasmic microtubules in human gingival fibroblasts, using two different monoclonal anti-GR antibodies, confocal laser scanning microscopy and image analysis. *J Steroid Biochem Mol Biol* 39:419-432
- Alexander RB, Greene GL, Barrack ER 1987 Estrogen receptors in the nuclear matrix: direct demonstration using monoclonal antireceptor antibody. *Endocrinology* 120:1851-1857
- Ali S, Metzger D, Bornert J-M, Chambon P 1993 Modulation of transcriptional activation by ligand-dependent phosphorylation of the human oestrogen receptor A/B region. *EMBO J* 12:1153-1160
- Allan GF, Ing NH, Tsai SY, Srinivasan G, Weigel NL, Thompson EB, Tsai M-J, O'Malley BW 1991 Synergism between steroid response and promoter elements during cell-free transcription. *J Biol Chem* 266:5905-5910
- Allan GF, Leng X, Tsai SY, Weigel NL, Edwards DP, Tsai M-J, O'Malley BW 1992a Hormone and antihormone induce distinct conformational changes which are central to steroid receptor activation. *J Biol Chem* 267:19513-19520
- Allan GF, Tsai SY, Tsai M-J, O'Malley BW 1992b Ligand-dependent conformational changes in the progesterone receptor are necessary for events that follow DNA binding. *Proc Natl Acad Sci USA* 89:11750-11754
- Amero SA, Kretsinger RH, Moncrief ND, Yamamoto KR, Pearson WR 1992 The origin of nuclear receptor proteins: a single precursor distinct from other transcription factors. *Mol Endocrinol* 6:3-7
- Andreeva M, Markova D, Loidl P, Djondjurov L 1992 Intranuclear compartmentalization of transcribed and nontranscribed *c-myc* sequences in Namalva-S cells. *Eur J Biochem* 207:887-894
- Andrew SE, Goldberg YP, Kremer B, Telenius H, Theilmann J, Adam S, Starr E, Squitieri F, Lin B, Kalchman MA, Graham RK, Hayden MR 1993 The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nature Genet* 4:398-403
- Apostolopoulos JJ, La Scala MJ, Shen P, Howlett GJ 1989 The effect of triiodothyronine on the association of the rat apolipoprotein A-I, C-III and A-IV genes with the nuclear matrix. *Biochem Biophys Res Commun* 161:576-582
- Araki I, Harada Y, Kuno M 1991 Target-dependent hormonal control of neuron size in the rat spinal nucleus of the bulbocavernosus. *J Neurosci* 11:3025-3033
- Arbizu T, Santamaria J, Gomez JM, Quilez A, Serra JP 1983 A family with adult spinal and bulbar muscular atrophy, X-linked inheritance and associated testicular failure. *J Neurol Sci* 59:371-382
- Archer TK, Lefebvre P, Wolford RG, Hager GL 1992 Transcription factor loading on the MMTV promoter: a bimodal mechanism for promoter activation. *Science* 255:1573-1576
- Aronica SM, Katzenellenbogen BS 1993 Stimulation of estrogen receptor-mediated transcription and alteration in the phosphorylation state of the rat uterine estrogen receptor by estrogen, cyclic adenosine monophosphate, and insulin-like growth factor-1. *Mol Endocrinol* 7:743-752
- Arriza JL, Weinberger C, Cerelli G, Glaser TM, Handelin BL, Housman DE, Evans RM 1987 Cloning of human mineralocorticoid receptor complementary DNA: structural and functional kinship with the glucocorticoid receptor. *Science* 237:268-275
- Auricchio F 1989 Phosphorylation of steroid receptors. *J Steroid Biochem* 32:613-622
- Bagchi MK, Elliston JF, Tsai SY, Edwards DP, Tsai M-J, O'Malley BW 1988 Steroid hormone-dependent interaction of human progesterone receptor with its target enhancer element. *Mol Endocrinol* 2:1221-1229
- Bagchi MK, Tsai SY, Tsai M-J, O'Malley BW 1992 Ligand and DNA-dependent phosphorylation of human progesterone receptor *in vitro*. *Proc Natl Acad Sci USA* 89:2664-2668
- Baker KP, Schatz G 1991 Mitochondrial proteins essential for viability mediate protein import into yeast mitochondria. *Nature* 349:205-208
- Baniahmad C, Muller M, Altschmied J, Renkawitz R 1991 Co-operative binding of the glucocorticoid receptor



- DNA binding domain is one of at least two mechanisms for synergism. *J Mol Biol* 222:155-165
- Baniahmad A, Ha I, Reinberg D, Tsai S, Tsai M-J, O'Malley BW 1993 Interaction of human thyroid hormone receptor  $\beta$  with transcription factor TFIIB may mediate target gene derepression and activation by thyroid hormone. *Proc Natl Acad Sci USA* 90:8832-8836
- Baniahmad A, Tsai M-J 1993 Mechanisms of transcriptional activation by steroid hormone receptors. *J Cell Biochem* 51:151-156
- Barrack ER, Coffey DS 1980 The specific binding of estrogens and androgens to the nuclear matrix of sex hormone responsive tissues. *J Biol Chem* 255:7265-7275
- Barrack ER 1983 The nuclear matrix of the prostate contains acceptor sites for androgen receptors. *Endocrinology* 113:430-432
- Barrack ER, Coffey DS 1983 The role of nuclear structure in hormone action. In: *Steroid hormone receptors: structure and function*. Editors: Eriksson H, Gustafsson J-A. Elsevier Science Publishers, BV, New York, pp 221-229
- Barrack ER 1987 Steroid hormone receptor localization in the nuclear matrix: interaction with acceptor sites. *J Steroid Biochem* 27:115-121
- Barsony J, McKoy W 1992 Molybdate increases intracellular 3',5'-guanosine cyclic monophosphate and stabilizes vitamin D receptor association with tubulin-containing filaments. *J Biol Chem* 267:24457-24465
- Beato M, Biesewig D, Braendle W, Sekeris CE 1969 On the mechanism of hormone action: subcellular distribution and binding of (1,2- $^3$ H) cortisol in rat liver. *Biochim Biophys Acta* 192:494-507
- Beck CA, Weigel NL, Edwards DP 1992 Effects of hormone and cellular modulators of protein phosphorylation on transcriptional activity, DNA binding, and phosphorylation of human progesterone receptors. *Mol Endocrinol* 6:607-620
- Becker DM, Hollenberg SM, Ricciardi RP 1989 Fusion of adenovirus E1A to the glucocorticoid receptor by high-resolution deletion cloning creates a hormonally inducible viral transactivator. *Mol Cell Biol* 9:3878-3887
- Beckmann RP, Mizzen LA, Welch WJ 1990 Interaction of hsp70 with newly synthesized proteins: implications for protein folding and assembly. *Nature* 248:850-854
- Beekman JA, Allan GF, Tsai SY, Tsai M-J, O'Malley BW 1993 Transcriptional activation by the estrogen receptor requires a conformational change in the ligand binding domain. *Mol Endocrinol* 7:1266-1274
- Beitel LK, Prior L, Vasilioni DM, Gottlieb B, Kaufman M, Lumbroso R, Alvarado C, McGillivray B, Trifiro M, Pinsky L 1994 Complete androgen insensitivity due to mutations in the probable  $\alpha$ -helical segment of the DNA-binding domain in the human androgen receptor. *Human Mol Genet* 3:21-27
- Belisle S, Bellabarba D, Lehoux J-G 1989 Estradiol receptor-nuclear interactions in aging mouse uteri: the role of DNA and nuclear matrix. *J Steroid Biochem* 32:41-43
- Berezney R & Coffey DS 1974 Identification of a nuclear protein matrix. *Biochem Biophys Res Commun* 60:1410-1417
- Berezney R 1991 The nuclear matrix: a heuristic model for investigating genome organization and function in the cell nucleus. *J Cell Biochem* 47:109-123
- Berry M, Metzger D, Chambon P 1990 Role of the two activating domains of the oestrogen receptor in the cell-type and promoter-context dependent agonistic activity of the anti-oestrogen 4-hydroxytamoxifen. *EMBO J* 9:2811-2818
- Biancalana V, Serville F, Pommier J, Julien J, Hanauer A, Mandel JL 1992 Moderate instability of the trinucleotide repeat in spino bulbar muscular atrophy. *Human Mol Genet* 1:255-258
- Bidwell JP, van Wijnen AJ, Fey EG, Dworetzky S, Penman S, Stein JL, Lian JB, Stein GS 1993 Osteocalcin gene promoter-binding factors are tissue-specific nuclear matrix components. *Proc Natl Acad Sci USA* 90:3162-3166
- Bocquel MT, Kumar V, Stricker C, Chambon P, Gronemeyer H 1989 The contribution of the N- and C-terminal regions of steroid receptors to activation of transcription is both receptor and cell-specific. *Nucleic Acid Res* 17:2581-2595
- Bodwell JE, Ortí E, Coull JM, Pappin DJC, Smith LI, Swift F 1991 Identification of phosphorylated sites in the mouse glucocorticoid receptor. *J Biol Chem* 266:7549-7555
- Bohmann D, Bos TJ, Admon A, Nishimura T, Vogt PK, Tjian R 1987 Human proto-oncogene *c-jun* encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. *Science* 238:1386-1392
- Bradshaw MS, Tsai SY, Leng X, Dobson ADW, Conneely OM, O'Malley BW, Tsai M-J 1991 Studies on the mechanism of functional cooperativity between progesterone and estrogen receptors. *J Biol Chem* 266:16684-16690

## References

- Bresnick EH, Sanchez ER, Harison RW, Pratt WB 1988 Hydrogen peroxide stabilizes the steroid-binding state of rat liver glucocorticoid receptors by promoting disulfide bond formation. *Biochemistry* 27:2866-2872
- Bresnick EH, Dalman FC, Sanchez ER, Pratt WB 1989 Evidence that the 90-kDa heat shock protein is necessary for the steroid binding conformation of the L cell glucocorticoid receptor. *J Biol Chem* 264:4992-4997
- Brink M, Humbel BM, de Kloet ER, van Driel R 1992 The unliganded glucocorticoid receptor is localized in the nucleus, not in the cytoplasm. *Endocrinology* 130:3575-3581
- Brinkmann AO, Faber PW, van Rooij HCJ, Kuiper GGJM, Ris C, Klaassen P, van der Korput JAGM, Voorhorst MM, van Laar JH, Mulder E, Trapman J 1989 The human androgen receptor: domain structure, genomic organization and regulation of expression. *J Steroid Biochem* 34:307-310
- Brinkmann AO, Trapman J 1992 Androgen receptor mutants that affect normal growth and development. In: *Cancer surveys. Growth regulation by nuclear hormone receptors.* vol 14:95-111
- Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion J-P, Hudson T, Sohn R, Zemelman B, Snell RG, Rundle SA, Crow S, Davies J, Shelbourne P, Buxton J, Jones C, Juvonen V, Johnson K, Harper PS, Shaw DJ, Housman DE 1992 Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* 68:799-808
- Brou C, Chaudhary S, Davidson I, Lutz Y, Wu J, Egly J-M, Tora L, Chambon P 1993 Distinct TFIID complexes mediate the effect of different transcriptional activators. *EMBO J* 12:489-499
- Brown TR, Lubahn DB, Wilson EM, French FS, Migeon CJ, Corden JL 1990 Functional characterization of naturally occurring mutant androgen receptors from subjects with complete androgen insensitivity. *Mol Endocrinol* 4:1759-1772
- Brown TA, DeLuca HF 1990 Phosphorylation of the 1,25-dihydroxyvitamin D<sub>3</sub> receptor. *J Biol Chem* 265:10025-10029
- Brown TA, DeLuca HF 1991 Sites of phosphorylation and photoaffinity labeling of the 1,25-dihydroxyvitamin D<sub>3</sub> receptor. *Arch Biochem Biophys* 286:466-472
- Brugge JS 1986 Interaction of Rous sarcoma virus protein pp60<sup>src</sup> with cellular proteins pp50 and pp90. *Curr Opin Microbiol Immunol* 123:363-372
- Brügge U, Kalff M, Franke S, Scheiderei C, Beato M 1991 Ubiquitous transcription factor OTF-1 mediates induction of the MMTV promoter through synergistic interaction with hormone receptors. *Cell* 64:565-572
- Buratowski S 1994 The basics of basal transcription by RNA polymerase II. *Cell* 77:1-3
- Burdon RH 1986 Heat shock and heat shock proteins. *Biochem J* 240:313-324
- Burglin TR 1988 The yeast regulatory gene PHO<sub>2</sub> encodes a homeobox. *Cell* 53:339-340
- Burns K, Duggan B, Atkinson EA, Fanulski KS, Nemer M, Bleackley RC, Michalak M 1994 Modulation of gene expression by calreticulin binding to the glucocorticoid receptor. *Nature* 367:476-480
- Burnstein KL, Cidlowski JA 1993 Multiple mechanisms for regulation of steroid hormone action. *J Cell Biochem* 51:130-134
- Buttayan R, Olsson CA 1986 Prediction of transcriptional activity based on gene association with the nuclear matrix. *Biochem Biophys Res Commun* 138:1334-1340
- Cadepond F, Schweizer-Groyer G, Segard-Maurel I, Jibard N, Hollenberg SM, Giguere V, Evans RM, Baulieu E-E 1991 Heat shock protein 90 as a critical factor in maintaining glucocorticosteroid receptor in a nonfunctional state. *J Biol Chem* 266:5834-5841
- Cadepond F, Casc J-M, Delahaye F, Jibard N, Schweizer-Groyer G, Segard-Maurel I, Evans R, Baulieu E-E 1992 Hormonal regulation of the nuclear localization signals of the human glucocorticosteroid receptor. *Exp Cell Res* 201:99-108
- Cairns W, Cairns C, Pongratz I, Poellinger L, Okret S 1991 Assembly of a glucocorticoid receptor complex prior to DNA binding enhances its specific interaction with a glucocorticoid response element. *J Biol Chem* 266:11221-11226
- Carlstedt-Duke J, Strömstedt P-E, Persson B, Cederlund E, Gustafsson J-A, Jörnvall H 1988 Identification of hormone-interacting amino acid residues within the steroid-binding domain of the glucocorticoid receptor in relation to other steroid hormone receptors. *J Biol Chem* 263:6842-6846
- Carlstedt-Duke J, Strömstedt P-E, Dahlman K, Rae C, Berkenstam A, Haggood J, Jörnvall H, Gustafsson J-A 1989 Structural analysis of the glucocorticoid receptor protein. In: *Carlstedt-Duke J, Eriksson H, Gustafsson J-A (eds) The steroid/thyroid hormone receptor family and gene regulation.* Birkhäuser verlag, Basel, vol 4:93-108
- Carson-Jurika MA, Lee AT, Dobson AW, Conneely OM, Schrader WT, O'Malley BW 1989 Interaction of the chicken progesterone receptor with heat shock protein (hsp) 90. *J Steroid Biochem* 34:1-9

- Caskey CT, Pizzuti A, Fu Y-H, Fenwick Jr RG, Nelson DL 1992 Triplet repeat mutations in human disease. *Science* 256:784-789
- Chakraborti PK, Hoeck W, Groner B, Simons Jr SS 1990 Localization of the vicinal dithiols involved in steroid binding to the rat glucocorticoid receptor. *Endocrinology* 127:2530-2539
- Chakraborti PK, Garabedian MJ, Yamamoto KR, Simons Jr SS 1991 Creation of "super" glucocorticoid receptors by point mutations in the steroid binding domain. *J Biol Chem* 266:22075-22078
- Chambraud B, Berry M, Redeuilh G, Chambon P, Baulieu E-E 1990 Several regions of human estrogen receptor are involved in the formation of receptor-heat shock protein 90 complexes. *J Biol Chem* 265:20686-20691
- Chandran UR, DeFranco DB 1992 Internuclear migration of chicken progesterone receptor, but not simian virus-40 large tumor antigen, in transient heterokaryons. *Mol Endocrinol* 6:837-844
- Chang C, Kokontis J, Liao S 1988a Molecular cloning of human and rat complementary DNA encoding androgen receptors. *Science* 240:324-326
- Chang C, Kokontis J, Liao S 1988b Structural analysis of complementary DNA and amino acid sequences of human and rat androgen receptors. *Proc Natl Acad Sci USA* 85:7211-7215
- Chang C, Chodak G, Sarac E, Takeda H, Liao S 1989a Prostate androgen receptor: immunohistochemical localization and mRNA characterization. *J Steroid Biochem* 34:311-313
- Charest NJ, Zhou Z-X, Lubahn DB, Olsen KL, Wilson EM, French FS 1991 A frameshift mutation destabilizes androgen receptor messenger RNA in the Tfm mouse. *Mol Endocrinol* 5:573-581
- Chauchereau A, Loosfelt H, Milgrom E 1991 Phosphorylation of transfected wild type and mutated progesterone receptors. *J Biol Chem* 266:18280-18286
- Chen C, Okayama H 1987 High-efficiency transformation of mammalian cells by plasmid DNA. *Mol Cell Biol* 7:2745-2752
- Chen D, Stallcup MR 1994 The hormone-binding role of 2 cysteines near the C-terminus of the mouse glucocorticoid receptor. *J Biol Chem* 269:7914-7918
- Chrousos GP, Detera-Wadleigh SD, Karl M 1993 Syndromes of glucocorticoid resistance. *Ann Intern Med* 119:1113-1124
- Chung M-y, Ranum LPW, Duvick LA, Servadio A, Zoghbi HY, Orr HT 1993 Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. *Nature Genet* 5:254-258
- Ciejek AM, Nordstrom JL, Tsai M-J, O'Malley BW 1982 Ribonucleic acid precursors are associated with the chick oviduct nuclear matrix. *Biochemistry* 21:4945-4953
- Ciejek EM, Tsai M-J, O'Malley BW 1983 Actively transcribed genes are associated with the nuclear matrix. *Nature* 306:607-609
- Clark AR, Docherty K 1993 Negative regulation of transcription in eukaryotes. *Biochem J* 296:521-541
- Colvard DS, Wilson EM 1984 Androgen receptor binding to the nuclear matrix *in vitro* and its inhibition by 8S androgen receptor promoting factor. *Biochemistry* 23:3479-3486
- Cook PR 1988 The nucleoskeleton: artifact, passive framework or active site? *J Cell Sci* 90:1-6
- Cook PR 1989 The nucleoskeleton and the topology of transcription. *Eur J Biochem* 185:487-501
- Cook PR 1991 The nucleoskeleton and the topology of replication. *Cell* 66:627-635
- Courey AJ, Tjian R 1988 Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* 55:887-898
- Courey AJ, Holtzman DA, Jackson SP, Tjian R 1989 Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. *Cell* 59:827-836
- Craig EA, Jacobsen K 1984 Mutations of the heat inducible 70 kilodalton genes of yeast confer temperature sensitive growth. *Cell* 38:841-849
- Culig Z, Hobisch A, Cronauer MV, Cato ACB, Hittmair A, Radmayr C, Eberle J, Bartsch G, Klocker H 1993 Mutant androgen receptor detected in an advanced-stage prostatic carcinoma is activated by adrenal androgen and progesterone. *Mol Endocrinol* 7:1541-1550
- Dahlman K, Strömstedt P-E, Rae C, Jönvall JJ, Carlstedt-Duke J, Gustafsson J-A 1989 High level expression in *Escherichia coli* of the DNA binding-domain of the glucocorticoid receptor in a functional form utilizing domain-specific cleavage of a fusion protein. *J Biol Chem* 264:804-809
- Dahlman-Wright K, Siltala-Roos H, Carlstedt-Duke J, Gustafsson J-A 1990 Protein-protein interactions facilitate DNA binding by the glucocorticoid receptor DNA-binding domain. *J Biol Chem* 265:14030-14035
- Dahlman-Wright K, Wright A, Gustafsson J-A, Carlstedt-Duke J 1991 Interaction of the glucocorticoid receptor DNA-binding domain with DNA as a dimer is mediated by a short segment of five amino acids. *J Biol Chem* 266:3107-3112
- Dahlman-Wright K, Wright APH, Gustafsson J-A 1992b Determinants of high-affinity DNA binding by the

## References

- glucocorticoid receptor: evaluation of receptor domains outside the DNA-binding domain. *Biochemistry* 31:9040-9044
- Dahlman-Wright K, Almlöf T, McEwan IJ, Gustafsson J-A, Wright APH 1994 Delineation of a small region within the major transactivation domain of the human glucocorticoid receptor that mediates transactivation of gene expression. *Proc Natl Acad Sci USA* 91:1619-1623
- Dalman FC, Sanchez ER, Perini F, Pratt WB 1988 Localization of phosphorylation sites with respect to the functional domains of the mouse L cell glucocorticoid receptor. *J Biol Chem* 263:12259-12267
- Dalman FC, Koenig RJ, Perdew GH, Massa E, Pratt WB 1990 In contrast to the glucocorticoid receptor, the thyroid hormone receptor is translated in the DNA binding state and is not associated with hsp90. *J Biol Chem* 265:3615-3618
- Dalman FC, Scherrer LC, Taylor LP, Akil H, Pratt WB 1991a Localization of the 90-kDa heat shock protein-binding site within the hormone-binding domain of the glucocorticoid receptor by peptide competition. *J Biol Chem* 266:3482-3490
- Dalman FC, Sturzenbecker LJ, Levin AA, Lucas DA, Perdew GH, Petkovitch M, Chambon P, Grippo JF, Pratt WB 1991b Retinoic acid receptor belongs to a subclass of nuclear receptors that do not form "docking" complexes with hsp90. *Biochemistry* 30:5605-5608
- Dang CV, Lee WMF 1989 Nuclear and nucleolar targeting sequences of c-erb-A, c-myc, N-myc, p53, hsp70, and HIV tat proteins. *J Biol Chem* 264:18019-18023
- Danielian PS, White R, Lees JA, Parker MG 1992 Identification of a conserved region required for hormone dependent transcriptional activation by steroid hormone receptors. *EMBO J* 11:1025-1033
- Danielsen M, Northrop JP, Ringold GM 1986 The mouse glucocorticoid receptor: mapping of functional domains by cloning, sequencing and expression of wild-type and mutant receptor proteins. *EMBO J* 5:2513-2522
- Danielsen M, Northrop JP, Jonklaas J, Ringold GM 1987 Domains of the glucocorticoid receptor involved in specific and nonspecific deoxyribonucleic acid binding, hormone activation, and transcriptional enhancement. *Mol Endocrinol* 1:816-822
- Danielsen M, Hinck L, Ringold GM 1989 Two amino acids within the knuckle of the first zinc finger specify DNA response element activation by the glucocorticoid receptor. *Cell* 57:1131-1138
- Danielsen M 1991 Structure and function of the glucocorticoid receptor. In: Parker MG (ed) *Nuclear hormone receptors*. Academic press, London, pp 39-78
- Dauvios S, White R, Parker MG 1993 The antiestrogen ICI 182780 disrupts estrogen receptor nucleocytoplasmic shuttling. *J Cell Sci* 106:1377-1388
- Dearnaley DP 1994 Cancer of the prostate. *Br Med J* 308:780-784
- Deartrick J, Daly M, Randsholt NB, Brock HW 1991 The complex genetic locus polyhomeotic in *Drosophila melanogaster* potentially encodes two homologous zinc-finger proteins. *Gene* 105:185-195
- Dedhar S, Rennie PS, Shago M, Hagesteijn C-YL, Yang H, Filmus J, Hawley RG, Bruchovsky N, Cheng H, Matusik RJ, Giguere V 1994 Inhibition of nuclear hormone receptor activity by calreticulin. *Nature* 367:480-483
- DeFranco DB, Qi M, Borror KC, Garabedian MJ, Brautigan DL 1991 Protein phosphatase types 1 and/or 2a regulate nucleocytoplasmic shuttling of glucocorticoid receptors. *Mol Endocrinol* 5:1215-1228
- de Graaf A, van Bergen en Henegouwen PMP, Meijne AML, van Driel R, Verkleij AJ 1991 Ultrastructural localization of nuclear matrix proteins in HeLa cells using silver-enhanced ultra-small gold probes. *J Histochem Cytochem* 39:1035-1045
- de Jong L, van Driel R, Stuurman N, Meijne AML, van Renswoude J 1990 Principles of nuclear organization. *Cell Biol Int Rep* 14:1051-1074
- de Jonge AJ, de Smit S, Kroos MA, Reuser AJ 1985 Cotransfer of synthetic human genes into mouse cells using isolated metaphase chromosomes or cellular DNA. *Hum Genet* 69:32-38
- De Luca LM 1991 Retinoids and their receptors in differentiation, embryogenesis, and neoplasia. *FASEB J* 5:2924-2933
- DeMarzo AM, Beck CA, Onate SA, Edwards DP 1991 Dimerization of mammalian progesterone receptors occurs in the absence of DNA and is related to the release of the 90-kDa heat shock protein. *Proc Natl Acad Sci USA* 88:72-76
- Demonacos C, Tsawdaroglou NC, Djordjevic-Markovic R, Papalopoulou M, Galanopoulos V, Papadogeorgaki S, Sekeris CE 1993 Import of the glucocorticoid receptor into rat liver mitochondria *in vivo* and *in vitro*. *J Steroid Biochem Mol Biol* 46:401-413
- Denner LA, Weigel NL, Maxwell BL, Schrader WT, O'Malley BW 1990a Regulation of progesterone receptor-mediated transcription by phosphorylation. *Science* 250:1740-1743
- Denner LA, Schrader WT, O'Malley BW, Weigel NL 1990b Hormonal regulation and identification of chicken

- progesterone receptor phosphorylation sites. *J Biol Chem* 265:16548-16555
- Denton RR, Koszewski NJ, Notides AC 1992 Estrogen receptor phosphorylation. Hormonal dependence and consequence on specific DNA binding. *J Biol Chem* 267:7263-7268
- Dingwall C, Laskey RA 1991 Nuclear targeting sequences- a consensus? *Trends Biochem Sci* 16:478-481
- Dobson ADW, Conneely OM, Beattie W, Maxwell BL, Mak P, Tsai M-J, Schrader WT, O'Malley BW 1989 Mutational analysis of the chicken progesterone receptor. *J Biol Chem* 264:4207-4211
- Dorit RL, Schoenbach L, Gilbert W 1990 How big is the universe of exons? *Science* 250:1377-1382
- Drouin J, Sun YL, Chamberland M, Gauthier Y, De Léan A, Nemer M, Schmidt TJ 1993 Novel glucocorticoid receptor complex with DNA element of the hormone-repressed POMC gene. *EMBO J* 12:145-156
- Duboule D, Haenlin M, Galliot B, Mohier E 1987 DNA sequences homologous to the *Drosophila* opa repeat are present in murine mRNAs that are differentially expressed in fetuses and adult tissues. *Mol Cell Biol* 7:2003-2006
- Duyao M et al. 1993 Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nature Genet* 4:387-392
- Dynlacht BD, Attardi LD, Admon A, Freeman M, Tjian R 1989 Functional analysis of NTF-1, a developmentally regulated *Drosophila* transcription factor that binds neuronal cis elements. *Genes Dev* 3:1677-1688
- Dynlacht BD, Hoey T, Tjian R 1991 Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* 6:563-576
- Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R 1992 Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 12:241-253
- Eilers M, Picard D, Yamamoto KR, Bishop JM 1989 Chimaeras of myc oncoprotein and steroid receptors cause hormone-dependent transformation of cells. *Nature* 340:66-68
- Elgin SCR 1988 The formation and function of DNase I hypersensitive sites in the process of gene activation. *J Biol Chem* 263:19259-19262
- Ellis RJ, Hemmingsen SM 1989 Molecular chaperones: proteins essential for the biogenesis of some macromolecular structures. *Trends Biochem Sci* 14:339-342
- Evan GI, Hancock DC 1987 Studies on the interaction of the human *c-myc* protein with cell nuclei: *p62c-myc* as a member of a discrete subset of nuclear protein. *Cell* 43:253-261
- Evans RM 1988 The steroid and thyroid hormone receptor superfamily. *Science* 240:889-895
- Faber PW, Kuiper GGJM, van Rooij H CJ, van der Korput JAGM, Brinkmann AO, Trapman J 1989 The N-terminal domain of the human androgen receptor is encoded by one large exon. *Mol Cell Endocrinol* 61:257-262
- Faber PW, King A, van Rooij H CJ, Brinkmann AO, de Both NJ, Trapman J 1991 The mouse androgen receptor: functional analysis of the protein and characterization of the gene. *Biochem J* 278:269-278
- Fawell SE, Lees JA, White R, Parker MG 1990 Characterization and colocalization of steroid binding and dimerization activities in the mouse estrogen receptor. *Cell* 60:953-962
- Fisher L, Catz D, Kelley D 1993 An androgen receptor mRNA isoform associated with hormone-induced cell proliferation. *Proc Natl Acad Sci USA* 90:8254-8258
- Flach H, Kaiser U, Westphal HM 1992 Monoclonal antipeptide antibodies to the glucocorticoid receptor. *J Steroid Biochem Mol Biol* 42:462-474
- Forman BM, Yang C-R, Au M, Cassanova J, Ghysdael J, Samuels HH 1989 A domain containing leucine zipper-like motifs mediate novel *in vivo* interactions between the thyroid hormone and retinoic acid receptors. *Mol Endocrinol* 3:1610-1626
- Forman BM, Samuels HH 1990 Interactions among a subfamily of nuclear hormone receptors: the regulatory zipper model. *Mol Endocrinol* 4:1293-1301
- Postinis Y, Theodoropoulos PA, Gravanis A, Stourmaras C 1992 Heat shock protein hsp90 and its association with the cytoskeleton: a morphological study. *Biochem Cell Biol* 70:779-786
- Freedman LP, Luisi BF, Korszun ZR, Basavappa R, Sigler PJ, Yamamoto KR 1988 The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. *Nature* 334:543-546
- Freedman LP 1992 Anatomy of the steroid receptor zinc finger region. *Endocr Rev* 13:129-145
- Freedman LP, Luisi BF 1993 On the mechanism of DNA binding by nuclear hormone receptors: a structural and functional perspective. *J Cell Biochem* 51:140-150
- Friedberg EC 1992 Nuclear targeting sequences. *Trends Biochem Sci* 17:347
- Fu Y-H, Pizzuti A, Fenwick Jr RG, King J, Rajnarayan S, Dunne PW, Dubel J, Nasser GA, Ashizawa T, de Jong P, Wieringa B, Korneluk R, Peryman MB, Epstein HF, Caskey CT 1992 An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 255:1256-1258

## References

- Fuller PJ 1991 The steroid receptor superfamily: mechanisms of diversity. *FASEB J* 5:3092-3099
- Fuqua SAW, Fitzgerald SD, Chamness GC, Tandon AK, MacDonnell DP, Nawaz Z, O'Malley BW, McGuire WL 1991 Variant human breast tumor estrogen receptor with constitutive transcriptional activity. *Cancer Res* 51:105-109
- Fuxe K, Wikström A-C, Okret S, Agnati LF, Härfstrand A, Yu Z-Y, Granholm L, Zoli M, Vale W, Gustafsson J-A 1985 Mapping the glucocorticoid receptor immunoreactive neurons in the rat tel- and diencephalon using monoclonal antibody against rat liver glucocorticoid receptor. *Endocrinology* 117:1803-1812
- Garcia-Bustos J, Heitman J, Hall MN 1991 Nuclear protein localization. *Biochim Biophys Acta* 1071:83-101
- Gasc J-M, Delahaye F, Baulieu E-E 1989 Compared intracellular localization of the glucocorticosteroid and progesterone receptors: an immunocytochemical study. *Exp Cell Res* 181:492-504
- Gaspar ML, Meo T, Tosi M 1990 Structure and size distribution of the androgen receptor mRNA in wild type and Tfm/y mutant mice. *Mol Endocrinol* 4:1600-1610
- Gerster T, Balmaceda CG, Roeder RG 1990 The cell type-specific octamer transcription factor OTF-2 has two domains required for the activation of transcription. *EMBO J* 9:1635-1643
- Getzenberg RH, Pienta KJ, Coffey DS 1990 The tissue matrix: cell dynamics and hormone action. *Endocr Rev* 11:399-417
- Getzenberg RH, Coffey DS 1990 Tissue specificity of the hormonal response in sex accessory tissues is associated with nuclear matrix protein patterns. *Mol Endocrinol* 4:1336-1342
- Getzenberg RH, Pienta KJ, Ward WS, Coffey DS 1991 Nuclear structure and the three-dimensional organization of DNA. *J Cell Biochem* 47:289-299
- Giguere V, Hollenberg SM, Rosenfeld MG, Evans RM 1986 Functional domains of the human glucocorticoid receptor. *Cell* 46:645-652
- Giguere V 1994 Retinoic acid receptors and cellular retinoid binding proteins: complex interplay in retinoid signaling. *Endocr Rev* 15:61-79
- Gill G, Tjian R 1992 Eukaryotic coactivators associated with the TATA box binding protein. *Curr Opin Genet Dev* 2:236-242
- Giniger E, Ptashne M 1987 Transcription in yeast activated by a putative amphipathic alpha helix linked to a DNA binding unit. *Nature* 330:670-672
- Gittes RF 1991 Carcinoma of the prostate. *N Engl J Med* 324:236-244
- Glass CK, Lipkin SM, Devary OV, Rosenfeld MG 1989 Positive and negative regulation of gene transcription by a retinoic acid-thyroid hormone receptor heterodimer. *Cell* 59:697-708
- Glass CK, DiRenzo J, Kurokawa R, Han Z 1991 Regulation of gene expression by retinoic acid receptors. *DNA Cell Biol* 10:623-638
- Godowski PJ, Picard D, Yamamoto KR 1987a Signal transduction and transcriptional regulation by glucocorticoid receptor-lexA fusion proteins. *Science* 241:812-816
- Godowski PJ, Rusconi S, Miesfeld R, Yamamoto KR 1987b Glucocorticoid receptor mutants that are constitutive activators of transcriptional enhancement. *Nature* 325:365-368
- Goldstein RA, Katzenellenbogen JA, Luthy-Schulten ZA, Seielstad DA, Wolynes PG 1993 Three-dimensional model for the hormone binding domains of steroid receptors. *Proc Natl Acad Sci USA* 90:9949-9953
- Govindan MV 1990 Specific region in hormone binding domain is essential for hormone binding and transactivation by human androgen receptor. *Mol Endocrinol* 4:417-427
- Grabowski DT, Carney JP, Kelley MR 1991 An adult male specific gene in *Drosophila* containing the repetitive element opa. *Biochim Biophys Acta* 1090:115-118
- Green S, Walter P, Kumar V, Krust A, Bornert J-M, Argos P, Chambon P 1986 Human oestrogen receptor cDNA: sequence expression and homology to v-erb-A. *Nature* 320:134-139
- Green S, Chambon P 1987 Oestradiol induction of a glucocorticoid-responsive gene by a chimaeric receptor. *Nature* 325:75-78
- Green S, Chambon P 1988 Nuclear receptors enhance our understanding of transcription regulation. *Trends Genet* 4:309-314
- Green S, Kumar V, Theulaz I, Wahli W, Chambon P 1988 The N-terminal DNA-binding 'zinc finger' of the oestrogen and glucocorticoid receptors determines target gene specificity. *EMBO J* 7:3037-3044
- Green S 1993a Promiscuous liaisons. *Nature* 361:590-591
- Green H 1993b Human genetic diseases due to codon reiteration: relationship to an evolutionary mechanism. *Cell* 74:955-956
- Greenberg CS, Birckbichler PJ, Rice RH 1991 Transglutaminases: multifunctional cross-linking enzymes that stabilize tissues. *FASEB J* 5:3071-3077
- Greenblatt J 1991 Roles of TFIID in transcriptional activation by RNA polymerase II. *Cell* 66:1067-1070

- Griffin JE, Wilson JD 1989 The androgen resistance syndromes: 5 $\alpha$ -reductase deficiency, testicular feminization and related disorders. In: Scriver CR, Baudat AL, Sly WS, Valle D (eds) *The Metabolic Basis of Inherited Disease*. McGraw-Hill, New York, pp 1919-1944
- Gronemeyer H, Turcotte B, Quirin-Stricker C, Bocquel MT, Meyer ME, Krozowski Z, Jeltsch JM, Lerouge T, Garnier JM, Chambon P 1987 The chicken progesterone receptor: sequence, expression and functional analysis. *EMBO J* 6:3985-3994
- Gronemeyer H 1991 Transcription activation by estrogen and progesterone receptors. *Annu Rev Genet* 25:89-123
- Gronenborn AM, Clore GM 1990 Protein structure determination in solution by two-dimensional and three-dimensional nuclear magnetic resonance spectroscopy. *Anal Chem* 62:2-15
- Guiochon-Mantel A, Loosfelt H, Ragot T, Bailly A, Atger M, Misrahi M, Perricaudet M, Milgrom E 1988 Receptors bound to antiprogestin form abortive complexes with hormone responsive elements. *Nature* 336:695-698
- Guiochon-Mantel A, Loosfelt H, Lescop P, Sar S, Atger M, Perrot-Applanat M, Milgrom E 1989 Mechanisms of nuclear localization of the progesterone receptor: evidence for interaction between monomers. *Cell* 57:1147-1154
- Guiochon-Mantel A, Lescop P, Christin-Maitre S, Loosfelt H, Perrot-Applanat M, Milgrom E 1991 Nucleocytoplasmic shuttling of the progesterone receptor. *EMBO J* 10:3851-3859
- Hager GL, Archer TK 1991 The interactions of steroid receptors with chromatin. In: Parker MG (ed) *Nuclear hormone receptors*. Academic press, London, pp 217-234
- Hamy F, Helbecque N, Henichart J-P 1992 Comparison between synthetic nuclear localization signal peptides from the steroid/thyroid hormone receptors superfamily. *Biochem Biophys Res Commun* 182:289-293
- Han K-K, Martinage A 1992 Post-translational chemical modification(s) of proteins. *Int J Biochem* 24:19-28
- Hård T, Kellenbach E, Boelens R, Maler BA, Dahlman K, Freedman LP, Carlstedt-Duke J, Yamamoto KR, Gustafsson J-A, Kaptein R 1990 Solution structure of the glucocorticoid receptor DNA-binding domain. *Science* 249:157-160
- Harlow KW, Smith DN, Katzenellenbogen JA, Greene GL, Katzenellenbogen BS 1989 Identification of cysteine 530 as the covalent attachment site of an affinity-labeling estrogen (ketononestrol aziridine) and antiestrogen (tamoxifen aziridine) in the human estrogen receptor. *J Biol Chem* 264:17476-17485
- Harrison SC 1991 A structural taxonomy of DNA-binding domains. *Nature* 353:715-719
- Hartl F-U, Neupert W 1990 Protein sorting to mitochondria: evolutionary conservation of folding and assembly. *Science* 247:930-938
- Hayes JJ, Wolffe AP 1992 The interaction of transcription factors with nucleosomal DNA. *Bioessays* 14:597-603
- Haynes SR, Rebbert ML, Mozer BA, Forquignon F, Dawid IB 1987 Pen repeat sequences are GGN clusters and encode a glycine-rich domain in a *Drosophila* cDNA homologous to the rat helix destabilizing protein. *Proc Natl Acad Sci USA* 84:1819-1823
- He WH, Fisher LM, Sun S, Bilhartz DL, Zhu X, Young CY-F, Kelly DB, Tindall DJ 1990a Molecular cloning of androgen receptors from divergent species with a polymerase chain reaction technique: complete cDNA sequence of the mouse androgen receptor and isolation of androgen receptor cDNA probes from dog, guinea pig and clawed frog. *Biochem Biophys Res Commun* 171:697-704
- He D, Nickerson JA, Penman S 1990b Core filaments of the nuclear matrix. *J Cell Biol* 110:569-580
- He WH, Kumar MV, Tindall DJ 1991 A frame-shift mutation in the androgen receptor gene causes complete androgen insensitivity in the testicular-feminized mouse. *Nucleic Acids Res* 19:2373-2378
- Hersko A 1988 Ubiquitin-mediated protein degradation. *J Biol Chem* 263:15237-15240
- Higuchi R, Krummel B, Saiki RK 1988 A general method of *in vitro* preparation and specific mutagenesis of DNA fragment: study of protein and DNA interactions. *Nucleic Acids Res* 16:7351-7367
- Hodges LL, Jordan CL, Breedlove SM 1993 Hormone-sensitive periods for the control of motoneuron number and soma size in the dorsolateral nucleus of the rat spinal cord. *Brain Res* 602:187-190
- Hoeck W, Groner B 1990 Hormone-dependent phosphorylation of the glucocorticoid receptor occurs mainly in the amino-terminal transactivation domain. *J Biol Chem* 265:5403-5408
- Hoefsloot LH, Willemsen R, Kroos MA, Hoogeveen-Westerveld M, Hermans MMP, van der Ploeg AT, Oostra BA, Reuser AJJ 1990 Expression and routing of human lysosomal  $\alpha$ -glucosidase in transiently transfected mammalian cells. *Biochem J* 272:485-492
- Hollenberg SM, Weinberger C, Ong ES, Cerelli G, Oro A, Lebo R, Thompson EB, Rosenfeld MG, Evans RM 1985 Primary structure and expression of a functional human glucocorticoid receptor cDNA. *Nature* 318:635-641
- Hollenberg SM, Giguere V, Segui P, Evans RM 1987 Colocalization of DNA-binding and transcriptional

## References

- activation functions in the human glucocorticoid receptor. *Cell* 49:39-46
- Hollenberg SM, Evans RM 1988 Multiple and cooperative trans-activation domains of the human glucocorticoid receptor. *Cell* 55:899-906
- Horwitz KB, Alexander PS 1983 *In situ* photolinked nuclear progesterone receptors of human breast cancer cells: subunit molecular weights after transformation and translocation. *Endocrinology* 113:2195-2201
- Housley PR, Sanchez ER, Danielsen M, Ringold GM, Pratt WB 1990 Evidence that the conserved region in the steroid binding domain of the glucocorticoid receptor is required for both optimal binding of hsp90 and protection from proteolytic cleavage. *J Biol Chem* 265:12778-12781
- Howard KJ, Holley SJ, Yamamoto KR, Distelhorst CW 1990 Mapping the hsp90 binding region of the glucocorticoid receptor. *J Biol Chem* 265:11928-11935
- Hsieh JC, Jurutka PW, Galligan MA, Terpening CM, Haussler CA, Samuels DS, Shimizu Y, Shimizu N, Haussler MR 1991 Human vitamin D receptor is selectively phosphorylated by protein kinase C on serine 51, a residue crucial to its trans-activation function. *Proc Natl Acad Sci USA* 88:9315-9319
- Hughes MR, O'Malley BW 1991 Genetic defects of receptors involved in disease. In: Parker MG (ed) *Nuclear hormone receptors*. Academic press, London, pp 321-353
- Humphrey GW, Pigiet V 1987 Protein disulfide crosslinking stabilizes a polyoma large T antigen-host protein complex on the nuclear matrix. *Exp Cell Res* 171:122-136
- Hunter T, Karin M 1992 The regulation of transcription by phosphorylation. *Cell* 70:375-387
- Husmann DA, Wilson CM, McPhaul MJ, Tilley WD, Wilson JD 1990 Antipeptide antibodies to two distinct regions of the androgen receptor localize the receptor protein to the nuclei of target cells in the rat and human prostate. *Endocrinology* 126:2359-2368
- Hutchison KA, Dalman FC, Hoeck W, Groner B, Pratt WB 1993 Localization of the  $\approx 12$  kDa Mr discrepancy in gel migration of the mouse glucocorticoid receptor to the major phosphorylated cyanogen bromide fragment in the transactivating domain. *J Steroid Biochem Mol Biol* 46:681-686
- Igarashi S, Tanno Y, Onodera O, Yamazaki M, Sato S, Ishikawa A, Miyatani N, Nagashima M, Ishikawa Y, Sahashi K, Ibi T, Miyatake T, Tsuji S 1992 Strong correlation between the number of CAG repeats in androgen receptor genes and the clinical onset of features of spinal and bulbar muscular atrophy. *Neurology* 42:2300-2302
- Imakado S, Koike S, Kondo S, Sakai M, Muramatsu M 1991 The N-terminal transactivation domain of rat estrogen receptor is localized in a hydrophobic domain of eighty amino acids. *J Biochem* 109:684-689
- Ing NH, Beekman JM, Tsai SY, Tsai M-J, O'Malley BW 1992 Members of the steroid hormone receptor superfamily interact with TFIIB (S300-II). *J Biol Chem* 267:17617-17623
- Itakawa H, Ryu HL 1979 *J Bacteriol* 138:339-344
- Jackson DA, McCready SJ, Cook PR 1981 RNA is synthesized at the nuclear cage. *Nature* 292:552-555
- Jackson DA, Cook PR 1985 Transcription occurs at a nucleoskeleton. *EMBO J* 4:919-925
- Jackson DA, Dolle A, Robertson G, Cook PR 1992 The attachments of chromatin loops to the nucleoskeleton. *Cell Biol Int Rep* 16:687-696
- Jackson DA, Hassan AB, Errington RJ, Cook PR 1993 Visualization of focal sites of transcription within human nuclei. *EMBO J* 12:1059-1065
- Jenster G, van der Korput JAGM, van Vroonhoven C, van der Kwast ThH, Trapman J, Brinkmann AO 1991 Domains of the human androgen receptor involved in steroid binding, transcriptional activation, and subcellular localization. *Mol Endocrinol* 5:1396-1404
- Jenster G, Trapman J, Brinkmann AO 1993 Nuclear import of the human androgen receptor. *Biochem J* 293:761-768
- Jiang S-Y, Langan-Fahey SM, Stella AL, McCague R, Jordan VC 1992 Point mutation of estrogen receptor (ER) in the ligand-binding domain changes the pharmacology of antiestrogens in ER-negative breast cancer cells stably expressing complementary DNAs for ER. *Mol Endocrinol* 6:2167-2174
- Jones N 1990 Transcriptional regulation by dimerization: two sides to an incestuous relationship. *Cell* 61:9-11
- Jost J-P, Seldran M 1984 Association of transcriptionally active vitellogenin II gene with the nuclear matrix of chicken liver. *EMBO J* 3:2005-2008
- Jurutka PW, Terpening CM, Haussler MR 1993 The 1,25-dihydroxy-vitamin D<sub>3</sub> receptor is phosphorylated in response to 1,25-dihydroxy-vitamin D<sub>3</sub> and 22-oxacalcitriol in rat osteoblasts, and by casein kinase II, *in vitro*. *Biochemistry* 32:8184-8192
- Kalderon D, Richardson WD, Markham AT, Smith AE 1984a Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* 311:33-38
- Kalderon D, Roberts BL, Richardson WD, Smith AE 1984b A short amino acid sequence able to specify nuclear location. *Cell* 39:499-509
- Kallio PJ, Jänne OA, Palvimo JJ 1994 Agonists, but not antagonists, alter the conformation of the hormone-



- binding domain of androgen receptor. *Endocrinology* 134:998-1001
- Kang KI, Devin J, Cadepond F, Jibard N, Guiochon-Mantel A, Baulieu E-E, Catelli M-G 1994 *In vivo* functional protein-protein interaction: nuclear targeted hsp90 shifts cytoplasmic steroid receptor mutants into the nucleus. *Proc Natl Acad Sci USA* 91:340-344
- Kastner P, Krust A, Turcotte B, Stropp U, Tora L, Gronemeyer H, Chambon P 1990 Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms A and B. *EMBO J* 9:1603-1614
- Katahira M, Knegtel R, Schilthuis J, Boelens R, Eib D, van der Saag P, Kaptein R 1992 The structure of the human retinoic acid receptor-beta DNA-binding domain determined by NMR. *Nucleic Acids Symp Ser.* 27:65-66
- Kaufmann SH, Coffey DS, Shaper JH 1981 Considerations in the isolation of rat liver nuclear matrix, nuclear envelope, and pore complex lamina. *Exp Cell Res* 132:105-123
- Kaufmann SH, Shaper JH 1984 A subset of non-histone nuclear proteins reversibly stabilized by the sulfhydryl cross-linking reagent tetrathionate. *Exp Cell Res* 155:477-495
- Kaufmann SH, Okret S, Wikström A-C, Gustafsson J-A, Shaper JH 1986 Binding of the glucocorticoid receptor to the rat liver nuclear matrix. *J Biol Chem* 261:11962-11967
- Kaufmann SHE 1990 Heat shock proteins and the immune response. *Immunol Today* 11:129-136
- Keese PK, Gibbs A 1992 Origins of genes: "Big bang" or continuous creation?. *Proc Natl Acad Sci USA* 89:9489-9493
- Kelleher RJ, Flanagan PM, Kornberg RD 1990 A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* 61:1209-1215
- Keller DW, Wiest WG, Askin FB, Johnson LW, Strickler RC 1979 Pseudocorpus luteum insufficiency: a local defect of progesterone action on endometrial stroma. *J Clin Endocrinol Metab* 48:127-132
- Kelley WL, Georgopoulos C 1992 Chaperones and protein folding. *Curr Opin Cell Biol* 4:984-991
- Kempainen JA, Lane MV, Sar M, Wilson EM 1992 Androgen receptor phosphorylation, turnover, nuclear transport, and transcriptional activation. *J Biol Chem* 267:968-974
- Kemp BE, Pearson RB 1990 Protein kinase recognition sequence motifs. *Trends Biochem Sci* 15:342-346
- Kennedy WR, Alter M, Sung JH 1968 Progressive proximal spinal and bulbar muscular atrophy of late onset: a sex-linked recessive trait. *Neurology* 18:671-680
- Kennelly PJ, Krebs EG 1991 Consensus sequences as substrate specificity determinants for protein kinases and protein phosphatases. *J Biol Chem* 266:15555-15558
- King WJ, Greene GL 1984 Monoclonal antibodies localize estrogen receptor in the nuclei of target cells. *Nature* 307:745-749
- Kirsch TM, Miller-Diener 1986 The nuclear matrix is the site of glucocorticoid receptor complex action in the nucleus. *Biochem Biophys Res Commun* 137:640-648
- Klempnauer K-H 1988 Interaction of *myb* proteins with nuclear matrix *in vitro*. *Oncogene* 2:545-551
- Knegtel RMA, Katahira M, Schilthuis JG, Bonvin AMJJ, Boelens R, Eib D, van der Saag PT, Kaptein R 1993 The solution structure of the human retinoic acid receptor- $\beta$  DNA-binding domain. *J Biomol NMR* 3:1-17
- Koide R, Ikeuchi T, Onodera O, Tanaka H, Igarashi S, Endo K, Takahashi H, Kondo R, Ishikawa A, Hayashi T, Saito M, Tomoda A, Miike T, Naito H, Ikuta F, Tsuji S 1994 Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). *Nature Genet* 6:9-13
- Kost SL, Smith DF, Sullivan WP, Welch WJ, Toft DO 1989 Binding of heat shock proteins to the avian progesterone receptor. *Mol Cell Biol* 9:3829-3838
- Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR, Richards RI 1991 Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)<sub>n</sub>. *Science* 252:1711-1714
- Krishna RG, Wold F 1993 Post-translational modification of proteins. *Adv Enzymol Relat Areas Mol Biol* 67:265-298
- Krongrad A, Wilson CM, Wilson JD, Allman DR, McPhaul MJ 1991 Androgen increases androgen receptor protein while decreasing receptor mRNA in LNCaP cells. *Mol Cell Endocrinol* 76:79-88
- Kuil CW, Mulder E 1994 Mechanism of antiandrogen action: conformational changes of the receptor. *Mol Cell Endocrinol* 102:R1-R6
- Kuiper GGJM, de Ruitter PE, Grootegoed JA, Brinkmann AO 1991 Synthesis and post-translational modification of the androgen receptor in LNCaP cells. *Mol Cell Endocrinol* 80:65-73
- Kuiper GGJM, de Ruitter PE, Trapman J, Boersma WJA, Grootegoed JA, Brinkmann AO 1993 Localization and hormonal stimulation of phosphorylation sites in the LNCaP-cell androgen receptor. *Biochem J* 291:95-101

## References

- Kuiper GGJM, Brinkmann AO 1994a Steroid hormone receptor phosphorylation: is there a physiological role? *Mol Cell Endocrinol* 100:103-107
- Kuiper GGJM, Brinkmann AO 1994b Phosphotryptic peptide analysis of the human androgen receptor: detection of a hormone-induced phosphopeptide. *submitted*
- Kumar V, Green S, Staub A, Chambon P 1986 Localisation of the oestradiol-binding and putative DNA-binding domains of the oestrogen receptor. *EMBO J* 5:2231-2236
- Kumar V, Green S, Stack G, Berry M, Jin J-R, Chambon P 1987 Functional domains of the human estrogen receptor. *Cell* 51:941-951
- Kumar V, Chambon P 1988 The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. *Cell* 55:145-156
- Kumara-Siri MH, Shapiro LE, Surks MI 1986 Association of the 3,5,3'-triiodo-L-thyronine nuclear receptor with the nuclear matrix of cultured growth hormone-producing rat pituitary tumor cells (GC cells). *J Biol Chem* 261:680-685
- Kupfer SR, Marschke KB, Wilson EM, French FS 1993 Receptor accessory factor enhances specific DNA binding of androgen and glucocorticoid receptors. *J Biol Chem* 268:17519-17527
- Kurz EM, Cover AR, Sengelaub DR 1992 Testosterone fails to save androgen-sensitive rat motoneurons following early target removal. *Dev Brain Res* 70:181-189
- Kushner PJ, Hort E, Shine J, Baxter JD, Greene GL 1990 Construction of cell lines that express high levels of the human estrogen receptor and are killed by estrogens. *Mol Endocrinol* 4:1465-1473
- Lahooti H, White R, Danielian PS, Parker MG 1994 Characterization of ligand-dependent phosphorylation of the estrogen receptor. *Mol Endocrinol* 8:182-188
- Landschulz WH, Johnson PF, McKnight SL 1989 The DNA binding domain of the rat liver C/EBP is bipartite. *Science* 243:1681-1688
- Lanford RE, Butel JS 1984 Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. *Cell* 37:801-813
- Laskey RA, Dingwall C 1993 Nuclear shuttling: the default pathway for nuclear proteins. *Cell* 74:585-586
- La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischback KH 1991 Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352:77-79
- Latchman DS 1990 Eukaryotic transcription factors. *Biochem J* 270:281-289
- Laudet V, Hänni C, Coll J, Catzeflis F, Stéhelin D 1992 Evolution of the nuclear receptor gene superfamily. *EMBO J* 11:1003-1013
- Lebeau MC, Massol N, Herrick J, Faber LE, Renoir JM, Radanyi C, Baulieu EE 1992 P59, an hsp90-binding protein. *J Biol Chem* 267:4281-4284
- Lebkowski JS, Laemmli UK 1982 Non-histone proteins and long-range organization of HeLa interphase DNA. *J Mol Biol* 156:325-344
- Lee JW, Gulick T, Moore DD 1992 Thyroid hormone receptor dimerization function maps to a conserved subregion of the ligand binding domain. *Mol Endocrinol* 6:1867-1873
- Lee MS, Kliewer SA, Provencal J, Wright P, Evans RM 1993 Structure of the retinoid X receptor  $\alpha$  DNA binding domain: a helix required for homodimeric DNA binding. *Science* 260:1117-1121
- Lees JA, Fawell SE, Parker MG 1989 Identification of two transactivation domains in the mouse estrogen receptor. *Nucleic Acids Res* 17:5477-5488
- Lees JA, Fawell SE, White R, Parker MG 1990 A 22-amino-acid peptide restores DNA-binding activity to dimerization-defective mutants of the estrogen receptor. *Mol Cell Biol* 10:5529-5531
- Le Goff P, Montano MM, Schodin DJ, Katzenellenbogen BS 1994 Phosphorylation of the human estrogen receptor. *J Biol Chem* 269:4458-4466
- Lemesle-Varloot L, Ojasoo T, Morron JP, Raynaud JP 1992 A model for the determination of the 3D-spatial distribution of the functions of the hormone-binding domain of receptors that bind 3-keto-4-ene steroids. *J Steroid Biochem Mol Biol* 41:369-388
- Lewin B 1990 Commitment and activation at pol II promoters: a tail of protein-protein interactions. *Cell* 61:1161-1164
- Lillie JW, Green MR 1989 Transcription activation by the adenovirus E1a protein. *Nature* 338:39-44
- Lin YS, Ha I, Maldonado E, Reinberg D, Green MR 1991 Binding of general transcription factor TFIIB to an acidic activating region. *Nature* 353:569-571
- Lobaccaro J-M, Lumbroso S, Belon C, Galtier-Dereure F, Bringer J, Lesimple T, Heron J-F, Pujol H, Sultan C 1993a Male breast cancer and the androgen receptor gene. *Nature Genet* 5:109-110
- Lobaccaro J-M, Lumbroso S, Belon C, Galtier-Dereure F, Bringer J, Lesimple T, Namer M, Cutuli BF, Pujol H, Sultan C 1993b Androgen receptor gene mutation in male breast cancer. *Human Mol Genet* 2:1799-1802

- Lubahn DB, Joseph DR, Sullivan PM, Willard HF, French FS, Wilson EM 1988a Cloning of human androgen receptor complementary DNA and localization to the X chromosome. *Science* 240:327-330
- Lubahn DB, Joseph DR, Sar M, Tan J, Higgs HN, Larson RE, French FS, Wilson EM 1988b The human androgen receptor: Complementary deoxyribonucleic acid cloning, sequence analysis and gene expression in prostate. *Mol Endocrinol* 2:1265-1275
- Lubahn DB, Brown TR, Simental JA, Higgs HN, Migeon CJ, Wilson EM, French FS 1989 Sequence of the intron/exon junctions of the coding region of the human androgen receptor gene and identification of a point mutation in a family with complete androgen insensitivity. *Proc Natl Acad Sci USA* 86:9534-9538
- Lubahn DB, Moyer JS, Golding TS, Couse JF, Korach KS, Smithies O 1993 Alteration of reproductive function but not prenatal sexual development after insertional disruption of the mouse estrogen receptor gene. *Proc Natl Acad Sci USA* 90:11162-11166
- Luisi BF, Xu WX, Otwinowski Z, Freedman LP, Yamamoto KR, Sigler PB 1991 Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature* 352:497-505
- Madan AP, DeFranco DB 1993 Bidirectional transport of glucocorticoid receptors across the nuclear envelope. *Proc Natl Acad Sci USA* 90:3588-3592
- Mader S, Kumar V, de Verneuil H, Chambon P 1989 Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid-responsive element. *Nature* 338:271-274
- Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barceló J, O'Hoy K, Lemond S, Earle-Macdonald J, de Jong PJ, Wieringa B, Korneluk RG 1992 Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* 255:1253-1255
- Maity SN, Vuorio T, de Crombrughe B 1990 The B subunit of a rat heteromeric CCAAT-binding transcription factor shows a striking sequence identity with the yeast Hap2 transcription factor. *Proc Natl Acad Sci USA* 87:5378-5382
- Mansour AM, Nass S 1970 *In vivo* cortisol action on RNA synthesis in rat liver nuclei and mitochondria. *Nature* 228:665-667
- Marcelli M, Tilley WD, Wilson CM, Griffin JE, Wilson JD, McPhaul MJ 1990a Definition of the human androgen receptor gene structure permits the identification of mutants that cause androgen resistance: premature termination of the receptor protein at amino acid residue 588 causes complete androgen resistance. *Mol Endocrinol* 4:1105-1116
- Marcelli M, Tilley WD, Wilson CM, Wilson JD, Griffin JE, McPhaul MJ 1990b A single nucleotide substitution introduces a premature termination codon into the androgen receptor gene of a patient with receptor-negative androgen resistance. *J Clin Invest* 85:1522-1528
- Martinez E, Wahli W 1991 Characterization of hormone response elements. In: Parker MG (ed) *Nuclear hormone receptors*. Academic press, London, pp 125-153
- Mason SA, Housley PR 1993 Site-directed mutagenesis of the phosphorylation sites in the mouse glucocorticoid receptor. *J Biol Chem* 268:21501-21504
- Matsuura T, Ogata A, Demura T, Moriwaka F, Tashiro K, Koyanagi T, Nagashima K 1993 Identification of androgen receptor in the rat spinal motoneurons. *Neurosci Lett* 158:5-8
- McConnell M, Whalen AM, Smith DE, Fisher PA 1987 Heat shock-induced changes in the structural stability of proteinaceous karyoskeletal elements *in vitro* and morphological effects *in situ*. *J Cell Biol* 105:1087-1098
- McEwan IJ, Wright APH, Dahlman-Wright K, Carlstedt-Duke J, Gustafsson J-A 1993 Direct interaction of the taur1 transactivation domain of the human glucocorticoid receptor with the basal transcriptional machinery. *Mol Cell Biol* 13:399-407
- McGrath ME, Wagner RL, Apriletti JW, West BL, Ramalingam V, Baxter JD, Fletterick RJ 1994 Preliminary crystallographic studies of the ligand-binding domain of the thyroid hormone receptor complexed with triiodothyronine. *J Mol Biol* 237:236-239
- McPhaul MJ, Marcelli M, Tilley WD, Griffin JE, Wilson JD 1991 Androgen resistance caused by mutations in the androgen receptor gene. *FASEB J* 5:2910-2915
- McPhaul MJ, Marcelli M, Zoppi S, Wilson CM, Griffin JE, Wilson JD 1992 Mutations in the ligand-binding domain of the androgen receptor gene cluster in two regions of the gene. *J Clin Invest* 90:2097-2101
- Meek DW, Street AJ 1992 Nuclear protein phosphorylation and growth control. *Biochem J* 287:1-15
- Mermod N, O'Niell EA, Kelly TJ, Tjian R 1989 The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. *Cell* 58:741-753
- Metzger DA, Korach KS 1990 Cell-free interaction of the estrogen receptor with mouse uterine nuclear matrix: evidence of saturability, specificity, and resistance to KCl extraction. *Endocrinology* 126:2190-2195

## References

- Meyer ME, Pornon A, Jingwei JI, Bocquel MT, Chambon P, Gronemeyer H 1990 Agonist and antagonist activities of RU486 on the functions of the human progesterone receptor. *EMBO J* 12:3923-3932
- Meyer M-E, Quirin-Stricker C, Lerouge T, Bocquel M-T, Gronemeyer H 1992 A limiting factor mediates the differential activation of promoters by the human progesterone receptor isoforms. *J Biol Chem* 267:10882-10887
- Mhatre AN, Trifiro MA, Kaufman M, Kazemi-Esfarjani P, Figlewicz D, Rouleau G, Pinsky L 1993 Reduced transcriptional regulatory competence of the androgen receptor in X-linked spinal and bulbar muscular atrophy. *Nature Genet* 5:184-188
- Michalak M, Milner RE, Burns K, Opas M 1992 Calreticulin. *Biochem J* 285:681-692
- Miesfeld R, Godowski PJ, Maler BA, Yamamoto KR 1987 Glucocorticoid receptor mutants that define a small region sufficient for enhancer activation. *Science* 236:423-427
- Migliaccio A, Rotondi A, Auricchio F 1986 Estradiol receptor: phosphorylation on tyrosine in uterus and interaction with antiphosphotyrosine antibody. *EMBO J* 5:2867-2872
- Mirkovitch J, Mirault M-E, Laemlli UK 1984 Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell* 39:223-232
- Mishari M, Atger M, d'Auriol L, Loosfelt H, Meriel C, Fridlansky F, Guiochon-Mantel A, Galibert F, Milgrom E 1987 Complete amino acid sequence of the human progesterone receptor deduced from cloned cDNA. *Biochem Biophys Res Commun* 143:740-748
- Mitchell PJ, Tjian R 1989 Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245:371-378
- Miyata Y, Yahara I 1991 Cytoplasmic 8S glucocorticoid receptor binds to actin filaments through the 90-kDa heat shock protein moiety. *J Biol Chem* 266:8779-8783
- Moudgil VK 1990 Phosphorylation of steroid hormone receptors. *Biochim Biophys Acta* 1055:243-258
- Mowszowicz I, Doukani A, Giacomini M 1988 Binding of the androgen receptor to the nuclear matrix of human foreskin. *J Steroid Biochem* 29:715-719
- Moyer ML, Borror KC, Bona BJ, DeFranco DB, Nordeen SK 1993 Modulation of cell signaling pathways can enhance or impair glucocorticoid-induced gene expression without altering the state of receptor phosphorylation. *J Biol Chem* 268:22933-22940
- Mulder E, Vrij L, Foekens JA 1981 Extraction of nuclear androgen receptors from rat prostate with different reagents. *Mol Cell Endocrinol* 23:283-296
- Muller M, Renkawitz R 1991 The glucocorticoid receptor. *Biochim Biophys Acta* 1088:171-182
- Muller M, Baniahmad C, Kaltschmidt C, Renkawitz R 1991a Multiple domains of the glucocorticoid receptor involved in synergism with the CACCC box factor(s). *Mol Endocrinol* 5:1498-1503
- Muller M, Baniahmad C, Kaltschmidt C, Schüle R, Renkawitz R 1991b Co-operative transactivation of steroid receptors. In: Parker MG (ed) *Nuclear hormone receptors*. Academic press, London, pp 155-174
- Myers RH, MacDonald ME, Koroshetz WJ, Duyao MP, Ambrose CM, Taylor SAM, Barnes G, Srinidhi J, Lin CS, Whaley WL, Lazzarini AM, Schwarz M, Wolff G, Bird ED, Vonsattel J-PG, Gusella JF 1993 De novo expansion of (CAG)<sub>n</sub> repeat in sporadic Huntington's disease. *Nature Genet* 5:168-173
- Nagafuchi S, Yanagisawa H, Sato K, Shirayama T, Ohsaki E, Bundo M, Takeda T, Tadokoro K, Kondo I, Murayama N, Tanaka Y, Kikushima H, Umino K, Kurosawa H, Furukawa T, Nihei K, Inoue T, Sano A, Komure O, Takahashi M, Yoshizawa T, Kanazawa I, Yamada M 1994 Dentatorubral and pallidolusian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. *Nature Genet* 6:14-18
- Nakamura M, Mita S, Murakami T, Uchino M, Watanabe S, Tokunaga M, Kumamoto T, Ando M 1994 Exonic trinucleotide repeats and expression of androgen receptor gene in spinal cord from X-linked spinal and bulbar muscular atrophy. *J Neurol Sci* 122:74-79
- Nardulli AM, Lew D, Erijman L, Shapiro DJ 1991 Purified estrogen receptor DNA binding domain expressed in *Escherichia coli* activates transcription of an estrogen-responsive promoter in cultured cells. *J Biol Chem* 266:24070-24076
- Nemoto T, Ohara-Nemoto Y, Denis M, Gustafsson JA 1990 The transformed glucocorticoid receptor has a lower steroid-binding affinity than the nontransformed receptor. *Biochemistry* 29:1880-1886
- Nemoto T, Ohara-Nemoto Y, Ota M 1992 Association of the 90-kDa heat shock protein does not affect the ligand-binding ability of androgen receptor. *J Steroid Biochem Mol Biol* 42:803-812
- Nervi C, Grippo JF, Sherman MI, George MD, Jetten AM 1989 Identification and characterization of nuclear retinoic acid-binding activity in human myeloblastic leukemia HL-60 cells. *Proc Natl Acad Sci USA* 86:5854-5858
- Newmark JR, Hardy DO, Tonb DC, Carter BS, Epstein JI, Isaacs WB, Brown TR, Barrack ER 1992 Androgen receptor gene mutations in human prostate cancer. *Proc Natl Acad Sci USA* 89:6319-6323

- Nigg EA, Baeuerle PA, Lührmann R 1991 Nuclear import-export: in search of signals and mechanisms. *Cell* 66:15-22
- Ohara-Nemoto Y, Nemoto T, Ota M 1991 The Mr 90,000 heat shock protein-free androgen receptor has a high affinity for steroid, in contrast to the glucocorticoid receptor. *J Biochem* 109:113-119
- Okret S, Wikström A-C, Wrangé O, Andersson B, Gustafsson J-A 1984 Monoclonal antibodies against the rat liver glucocorticoid receptor. *Proc Natl Acad Sci USA* 81:1609-1613
- O'Malley BW 1989 Did eucaryotic steroid receptors evolve from intracrine gene regulators? *Endocrinology* 125:1119-1120
- O'Malley BW, Conneely OM 1992 Orphan receptors: in search of a unifying hypothesis for activation. *Mol Endocrinol* 6:1359-1361
- Orr HT, Chung M-y, Banfi S, Kwaitkowski jr TJ, Servadio A, Beaudet AL, McCall AE, Duvick LA, Ranum LPW, Zoghbi HY 1993 Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type I. *Nature Genet* 4:221-226
- Ortí E, Mendel DB, Smith LI, Munck A 1989 Agonist-dependent phosphorylation and nuclear dephosphorylation of glucocorticoid receptors in intact cells. *J Biol Chem* 264:9728-9731
- Ortí E, Bodwell JE, Munck A 1992 Phosphorylation of steroid hormone receptors. *Endocr Rev* 13:105-128
- Pabo CO, Sauer RT 1992 Transcription factors: structural families and principles of DNA recognition. *Annu Rev Biochem* 61:1053-1095
- Pakdel F, Reese JC, Katzenellenbogen BS 1993a Identification of charged residues in an N-terminal portion of the hormone-binding domain of the human estrogen receptor important in transcriptional activity of the receptor. *Mol Endocrinol* 7:1408-1417
- Pakdel F, Le Goff P, Katzenellenbogen BS 1993b An assessment of the role of domain F and PEST sequences in estrogen receptor half-life and bioactivity. *J Steroid Biochem Mol Biol* 46:663-672
- Palvimo JJ, Kallio PJ, Ikonen T, Mehto M, Jänne OA 1993 Dominant negative regulation of trans-activation by the rat androgen receptor: roles of the N-terminal domain and heterodimer formation. *Mol Endocrinol* 7:1399-1407
- Pearce D, Yamamoto KR 1993 Mineralocorticoid and glucocorticoid receptor activities distinguished by nonreceptor factors at a composite response element. *Science* 259:1161-1165
- Perlmann T, Eriksson P, Wrangé O 1990 Quantitative analysis of the glucocorticoid receptor-DNA interaction at the mouse mammary tumor virus glucocorticoid response element. *J Biol Chem* 265:17222-17229
- Perrot-Appinat M, Logeat F, Griyer-Picard MT, Milgrom E 1985 Immunocytochemical study of mammalian progesterone receptor using monoclonal antibodies. *Endocrinology* 116:1473-1483
- Peters R 1986 Fluorescence microphotolysis to measure nucleocytoplasmic transport and intracellular mobility. *Biochim Biophys Acta* 864:305-359
- Peterson MG, Tanese N, Pugh BF, Tjian R 1990 Functional domains and upstream activation properties of cloned human TATA binding protein. *Science* 248:1625-1630
- Pfahl M 1993 Nuclear receptor/AP-1 interaction. *Endocr Rev* 14:651-658
- Pham TA, Hwung Y-P, Santiso-Mere D, McDonnell DP, O'Malley BW 1992 Ligand-dependent and -independent function of the transactivation regions of the human estrogen receptor in yeast. *Mol Endocrinol* 6:1043-1050
- Picard D, Yamamoto KR 1987 Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. *EMBO J* 6:3333-3340
- Picard D, Salsler SJ, Yamamoto KR 1988 A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor. *Cell* 54:1073-1080
- Picard D, Khursheed B, Garabedian MJ, Fortin MG, Lindquist S, Yamamoto KR 1990 Reduced levels of hsp90 comprise steroid receptor action *in vivo*. *Nature* 348:166-168
- Pinsky L, Trifiro M, Kaufman M, Beitel LK, Mhatre A, Kazemi-Esfarjani P, Sabbaghian N, Lumbroso R, Alvarado C, Vasiliou M, Gottlieb B 1992 Androgen resistance due to mutation of the androgen receptor. *Clin Invest Med* 15:456-472
- Poletti A, Weigel NL 1993 Identification of a hormone-dependent phosphorylation site adjacent to the DNA-binding domain of the chicken progesterone receptor. *Mol Endocrinol* 7:241-246
- Power RF, Mani SK, Codina J, Conneely OM, O'Malley BW 1991b Dopaminergic and ligand-independent activation of steroid hormone receptors. *Science* 254:1636-1639
- Power RF, Conneely OM, O'Malley BW 1992 New insights into activation of the steroid hormone receptor superfamily. *Trends Pharmacol Sci* 13:318-323
- Pratt WB, Jolly DJ, Pratt DV, Hollenberg SM, Giguere V, Cadepond FM, Schweizer-Groyer G, Catelli M-G, Evans RM, Baulieu E-E 1988 A region in the steroid binding domain determines formation of the non-DNA-binding, 9S glucocorticoid receptor complex. *J Biol Chem* 263:267-273

## References

- Pratt WB, Sanchez ER, Bresnick EH, Meshinchi S, Scherrer LC, Dalman FC, Welsh MJ 1989 Interaction of the glucocorticoid receptor with the Mr 90,000 heat shock protein: an evolving model of ligand-mediated receptor transformation and translocation. *Cancer Res* 49:2222s-2229s
- Pratt WB 1990 Interaction of hsp90 with steroid receptors: organizing some diverse observations and presenting the newest concept. *Mol Cell Endocrinol* 74:c69-c76
- Pratt WB, Czar MJ, Stancato LF, Owens JK 1993 The hsp56 immunophilin component of steroid receptor heterocomplexes: could this be the elusive nuclear localization signal-binding protein? *J Steroid Biochem Mol Biol* 46:269-279
- Pratt WB 1993 The role of heat shock proteins in regulating the function, folding, and trafficking of the glucocorticoid receptor. *J Biol Chem* 268:21455-21458
- Ptashne M 1988 How eukaryotic transcriptional activators work. *Nature* 335:683-689
- Ptashne M 1989 How gene activators work. *Scientific Am* Jan:25-31
- Pugh BF, Tjian R 1991 Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes Dev* 5:1935-1945
- Ratajczak T, Wilkinson SP, Brockway MJ, Hähnel R, Moritz RL, Begg GS, Simpson RJ 1989 The interaction site for tamoxifen aziridine with the bovine estrogen receptor. *J Biol Chem* 264:13453-13459
- Razin SV, Yarovaya OV 1985 Initiated complexes of RNA polymerase II are concentrated in the nuclear skeleton associated DNA. *Exp Cell Res* 158:273-275
- Redmond T, Sanchez ER, Bresnick EH, Schlesinger MJ, Toft DO, Pratt WB, Welsh MJ 1989 Immunofluorescence colocalization of the 90-kDa heat-shock protein and microtubules in interphase and mitotic mammalian cells. *Eur J Cell Biol* 50:66-75
- Reese JC, Wooge CH, Katzenellenbogen BS 1992 Identification of two cysteines closely positioned in the ligand-binding pocket of the human estrogen receptor: roles in ligand binding and transcriptional activation. *Mol Endocrinol* 6:2160-2166
- Refetoff S, Weiss RE, Usala SJ 1993 The syndromes of resistance to thyroid hormone. *Endocr Rev* 14:348-399
- Richards RI, Sutherland GR 1992 Heritable unstable DNA sequences. *Nature Genet* 1:7-9
- Riegman PHJ, Vlietstra RJ, van der Korput JAGM, Brinkmann AO, Trapman J 1991 The promoter of the prostate-specific antigen gene contains a functional androgen response element. *Mol Endocrinol* 5:1921-1930
- Rigaud G, Roux J, Pictet R, Grange T 1991 *In vivo* footprinting of rat TAT gene: dynamic interplay between the glucocorticoid receptor and a liver-specific factor. *Cell* 67:977-986
- Rihs H-P, Peters R 1989 Nuclear transport kinetics depend on phosphorylation-site-containing sequences flanking the karyophilic signal of the simian virus 40 T-antigen. *EMBO J* 8:1479-1484
- Ris-Stalpers C, Kuiper GGJM, Faber PW, Schweikert HU, van Rooij HCJ, Zegers ND, Hodgins MB, Degenhart HJ, Trapman J, Brinkmann AO 1990 Aberrant splicing of androgen receptor mRNA results in synthesis of a nonfunctional receptor protein in a patient with androgen insensitivity. *Proc Natl Acad Sci USA* 87:7866-7870
- Ris-Stalpers C, Verleun-Mooijman MCT, Trapman J, Brinkmann AO 1993 Threonine on amino acid position 868 in the human androgen receptor is essential for androgen binding specificity and functional activity. *Biochem Biophys Res Commun* 196:173-180
- Robbins J, Dilworth SM, Laskey RA, Dingwall C 1991 Two interdependent basic domains in nucleoplasmic nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* 64:615-623
- Robinson SI, Nelkin BD, Vogelstein B 1982 The ovalbumin gene is associated with the nuclear matrix of chicken oviduct cells. *Cell* 28:99-106
- Roche PJ, Hoare SA, Parker MG 1992 A consensus DNA-binding site for the androgen receptor. *Mol Endocrinol* 6:2229-2235
- Rodriguez R, Weigel NL, O'Malley BW, Schrader WT 1990 Dimerization of the chicken progesterone receptor *in vitro* can occur in the absence of hormone and DNA. *Mol Endocrinol* 4:1782-1790
- Rojiani MV, Finlay BB, Gray V, Dedhar S 1991 *In vitro* interaction of a polypeptide homologous to human Ro/SS-A antigen (calreticulin) with a highly conserved amino acid sequence in the cytoplasmic domain of integrin alpha subunits. *Biochemistry* 30:9859-9866
- Ruff VA, Yem AW, Munns PL, Adams LD, Reardon IM, Deibel Jr MR, Leach KL 1992 Tissue distribution and cellular localization of hsp56, an FK506-binding protein. Characterization using a highly specific polyclonal antibody. *J Biol Chem* 267:21285-21288
- Rundlett SE, Wu X-P, Miesfeld RL 1990 Functional characterizations of the androgen receptor confirm that the molecular basis of androgen action is transcriptional regulation. *Mol Endocrinol* 4:708-714

- Rupprecht R, Arriza JL, Spengler D, Reul JM, Evans RM, Holsboer F, Damm K 1993 Transactivation and synergistic properties of the mineralocorticoid receptor: relationship to the glucocorticoid receptor. *Mol Endocrinol* 7:597-603
- Rusconi S, Yamamoto KR 1987 Functional dissection of the hormone and DNA binding activities of the glucocorticoid receptor. *EMBO J* 6:1309-1315
- Sadowski I, Ma J, Triezenberg S, Ptashne M 1988 Gal4-VP16 is an unusually potent transcriptional activator. *Nature* 335:563-564
- Sadowski I, Ptashne M 1989 A vector for expressing GAL4(1-147) fusions in mammalian cells. *Nucleic Acids Res* 17:7539
- Sai T, Seino S, Chang C, Trifiro M, Pinsky L, Mhatre A, Kaufman M, Lambert B, Trapman J, Brinkmann AO, Rosenfield RL, Liao S 1990 An exonic point mutation of the androgen receptor gene in a family with complete androgen insensitivity. *Am J Hum Genet* 46:1095-1100
- Saito H, Uchida H 1977 Initiation of the DNA replication of bacteriophage lambda in *Escherichia coli* K12. *Mol Biol* 113:1-5
- Samuels HH, Forman BM, Horowitz ZD, Ye ZS 1988 Regulation of gene expression by thyroid hormone. *J Clin Invest* 81:957-967
- Sanchez ER, Hirst M, Scherrer LC, Tang H-Y, Welsh MJ, Harmon JM, Simons Jr SS, Ringold RM, Pratt WB 1990 Hormone-free mouse glucocorticoid receptors overexpressed in the chinese hamster ovary cells are localized to the nucleus and are associated with both hsp70 and hsp90. *J Biol Chem* 265:20123-20130
- Santen RJ 1992 Clinical review 37: endocrine treatment of prostate cancer. *J Clin Endocrinol Metab* 75:685-689
- Sar M, Lubahn DB, French FS, Wilson EM 1990 Immunohistochemical localization of the androgen receptor in rat and human tissues. *Endocrinology* 127:3180-3186
- Sartorius CA, Tung L, Takimoto GS, Horwitz KB 1993 Antagonist-occupied human progesterone receptors bound to DNA are functionally switched to transcriptional agonists by cAMP. *J Biol Chem* 268:9262-9266
- Schatt MD, Rusconi S, Schaffner W 1990 A single DNA-binding transcription factor is sufficient for activation from a distant enhancer and/or from a promoter position. *EMBO J* 9:481-487
- Schena M, Freedman LP, Yamamoto KR 1989 Mutations in the glucocorticoid receptor zinc finger region that distinguish interdigitated DNA binding and transcriptional enhancement activities. *Genes Dev* 3:1590-1601
- Scherrer LC, Dalman FC, Massa E, Meshinchi S, Pratt WB 1990 Structural and functional reconstitution of the glucocorticoid receptor-hsp90 complex. *J Biol Chem* 265:21397-21400
- Schlesinger MJ 1990 Heat shock proteins. *J Biol Chem* 265:12111-12114
- Schoenberg MP, Hakimi JM, Wang S, Bova GS, Epstein JI, Fischbeck KH, Isaacs WB, Walsh PC, Barrack ER 1994 Microsatellite mutation (CAG<sub>24-18</sub>) in the androgen receptor gene in human prostate cancer. *Biochem Biophys Res Commun* 198:74-80
- Schowalter DB, Sullivan WP, Mailhe NJ, Dobson ADW, Conneely OM, O'Malley BW, Toft DO 1991 Characterization of progesterone receptor binding to the 90- and 70-kDa heat shock proteins. *J Biol Chem* 266:21165-21173
- Schrader WT, O'Malley BW 1972 Progesterone-binding components of chick oviduct: characterization of purified subunits. *J Biol Chem* 247:51-59
- Schrader WT 1993 Insight: Subunit functions of the steroid/thyroid receptor family. *Mol Endocrinol* 7:1241-1243
- Schreiber SL, 1991 Chemistry and biology of the immunophilins and their immunosuppressive ligands. *Science* 251:283-287
- Schreiber V, Molinette M, Boeuf H, de Murcia G, Ménéssier-de Murcia J 1992 The human poly(ADP-ribose) polymerase nuclear localization signal in a bipartite element functionally separate DNA binding and catalytic activity. *EMBO J* 11:3263-3269
- Schröder FH 1991 Endocrine therapy: where do we stand and where are we going? *Cancer Surveys* 11:177-194
- Schuchard M, Subramaniam M, Ruesink T, Spelsberg TC 1991 Nuclear matrix localization and specific matrix DNA binding by receptor binding factor 1 of the avian oviduct progesterone receptor. *Biochemistry* 30:9516-9522
- Schüle R, Müller M, Kaltschmidt C, Renkawitz R 1988 Many transcription factors interact synergistically with steroid receptors. *Science* 242:1418-1420
- Schwabe JWR, Neuhaus D, Rhodes D 1990 Solution structure of the DNA-binding domain of the oestrogen receptor. *Nature* 348:458-461
- Schwabe JWR, Chapman L, Finch JT, Rhodes D 1993 The crystal structure of the estrogen receptor DNA-

## References

- binding domain bound to DNA: how receptors discriminate between their response elements. *Cell* 75:567-578
- Schweikert H-U 1993 The androgen resistance syndromes: clinical and biochemical aspects. *Eur J Pediatr* 152:s50-s57
- Scott MP, Carrol SB 1987 The segmentation and homeotic gene network in early *Drosophila* development. *Cell* 51:689-698
- Seed B, Sheen J-Y 1988 A simple phase-extraction assay for chloramphenicol acetyltransferase activity. *Gene* 67:271-277
- Severne Y, Wieland S, Schaffner W, Rusconi S 1988 Metal binding 'finger' structures in the glucocorticoid receptor defined by site-directed mutagenesis. *EMBO J* 7:2503-2508
- Shemshedini L, Ji J, Brou C, Chambon P, Gronemeyer H 1992 *In vitro* activity of the transcription activation functions of the progesterone receptor. *J Biol Chem* 267:1834-1839
- Sheridan PL, Krett NL, Gordon JA, Horwitz KB 1988 Human progesterone receptor transformation and nuclear down-regulation are independent of phosphorylation. *Mol Endocrinol* 2:1329-1342
- Sheridan PL, Francis MD, Horwitz KB 1989 Synthesis of human progesterone receptors in T47D cells. *J Biol Chem* 264:7054-7058
- Shibuya H, Nonneman DJ, Huang TH-M, Ganjam VK, Mann FA, Johnson GS 1993 Two polymorphic microsatellites in a coding segment of the canine androgen receptor gene. *Animal Genet* 24:345-348
- Silva CM, Cidlowski JA 1989 Direct evidence for intra- and intermolecular disulfide bond formation in the human glucocorticoid receptor. *J Biol Chem* 264:6638-6647
- Silva CM, Cidlowski JA 1992 The effect of oxidation/reduction on the charge heterogeneity of the human glucocorticoid receptor. *J Steroid Biochem Mol Biol* 41:1-10
- Silver PA 1991 How proteins enter the nucleus. *Cell* 64:489-497
- Simental JA, Sar M, Lane MV, French FS, Wilson EM 1991 Transcriptional activation and nuclear targeting signals of the human androgen receptor. *J Biol Chem* 266:510-518
- Simons Jr SS, Pumphrey JG, Rudikoff S, Eisen HJ 1987 Identification of cysteine 656 as the amino acid of hepatoma tissue culture cell glucocorticoid receptors that is covalently labeled by dexamethasone. *J Biol Chem* 262:9676-9680
- Simons Jr SS, Sistare FD, Chakraborti PK 1989 Steroid binding activity is retained in a 16 kDa fragment of the steroid binding domain of rat glucocorticoid receptors. *J Biol Chem* 264:14493-14497
- Sleddens HFBM, Oostra BA, Brinkmann AO, Trapman J 1992 Trinucleotide repeat polymorphism in the androgen receptor gene (AR). *Nucleic Acids Res* 20:1427
- Smith DF, Toft DO 1993 Steroid receptors and their associated proteins. *Mol Endocrinol* 7:4-11
- Smith CL, Conneely OM, O'Malley BW 1993 Modulation of the ligand-independent activation of the human estrogen receptor by hormone and antihormone. *Proc Natl Acad Sci USA* 90:6120-6124
- Snell RG, MacMillan JC, Cheadle JP, Fenton I, Lazarou LP, Davies P, MacDonald ME, Gusella JF, Harper PS, Shaw DJ 1993 Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. *Nature Genet* 4:393-397
- Spanjaard RA, Darling DS, Chin WW 1991 Ligand-binding and heterodimerization activities of a conserved region in the ligand-binding domain of the thyroid hormone receptor. *Proc Natl Acad Sci USA* 88:8587-8591
- Stein GS, Lian JB, Dworetzky SI, Owen TA, Bortell R, Bidwell JP, van Wijnen AJ 1991 Regulation of transcription-factor activity during growth and differentiation: involvement of the nuclear matrix in concentration and localization of promoter binding proteins. *J Cell Biochem* 47:300-305
- Strähle U, Boshart M, Klock G, Stewart F, Schütz G 1989 Glucocorticoid- and progesterone-specific effects are determined by differential expression of the respective hormone receptors. *Nature* 339:629-632
- Stringer KF, Ingles CJ, Greenblatt J 1990 Direct and selective binding of an acidic activation domain to the TATA-box factor TFIID. *Nature* 345:783-786
- Strömstedt P-E, Berkenstam A, Jörnvall H, Gustafsson J-A, Carlstedt-Duke J 1990 Radiosequence analysis of the human progestin receptor charged with [<sup>3</sup>H]promegestone. A comparison with the glucocorticoid receptor. *J Biol Chem* 265:12973-12977
- Stuurman N, de Jong L, van Driel R 1992a Nuclear frameworks: concepts and operational definitions. *Cell Biol Int Rep* 16:837-852
- Stuurman N, de Graaf A, Floore A, Josso A, Humbel B, de Jong L, van Driel R 1992b A monoclonal antibody recognizing nuclear matrix-associated nuclear bodies. *J Cell Sci* 101:773-784
- Stuurman N, Floore A, Colen A, de Jong L, van Driel R 1992c Stabilization of the nuclear matrix by disulfide bridges: identification of matrix-associated polypeptides that form disulfides. *Exp Cell Res* 200:285-294
- Sultan C, Lumbroso S, Poujol N, Belon C, Boudon C, Lobaccaro J-M 1993 Mutations of androgen receptor



- gene in androgen receptor insensitivity syndromes. *J Steroid Biochem Mol Biol* 46:519-530
- Superti-Furga G, Bergers G, Picard D, Busslinger M 1991 Hormone-dependent transcriptional regulation and cellular transformation by *fos*-steroid receptor fusion proteins. *Proc Natl Acad Sci USA* 88:5114-5118
- Suzuki H, Sato N, Watabe Y, Masai M, Seino S, Shimazaki J 1993 Androgen receptor gene mutations in human prostate cancer. *J Steroid Biochem Mol Biol* 46:759-765
- Takimoto GS, Tasset DM, Eppert AC, Horwitz KB 1992 Hormone-induced progesterone receptor phosphorylation consists of sequential DNA-independent and DNA-dependent stages: analysis with zinc finger mutants and the progesterone antagonist ZK98299. *Proc Natl Acad Sci USA* 89:3050-3054
- Takimoto GS, Horwitz KB 1993 Progesterone receptor phosphorylation. Complexities in defining a functional role. *Trends Endocrinol Metab* 4:1-7
- Tan J-A, Joseph DR, Quarmby VE, Lubahn DB, Sar M, French FS, Wilson EM 1988 The rat androgen receptor: primary structure, autoregulation of its messenger ribonucleic acid, and immunocytochemical localization of the receptor protein. *Mol Endocrinol* 2:1276-1285
- Tanaka M, Herr W 1990 Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. *Cell* 60:375-386
- Tanese N, Pugh BF, Tjian R 1991 Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. *Genes Dev* 5:2212-2224
- Tasset D, Tora L, Fromental C, Scheer E, Chambon P 1990 Distinct classes of transcriptional activating domains function by different mechanisms. *Cell* 62:1177-1187
- The Huntington's disease collaborative research group 1993 A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72:971-983
- Thorburn A, Moore R, Knowland J 1988 Attachment of transcriptionally active DNA sequences to the nucleoskeleton under isotonic conditions. *Nucleic Acids Res* 16:7183-7183
- Tilley WD, Marcelli M, Wilson JD, McPhaul MJ 1989 Characterization and expression of a cDNA encoding the human androgen receptor. *Proc Natl Acad Sci USA* 86:327-331
- Timmers M, Sharp PA 1991 The mammalian TFIID protein is present in two functionally distinct complexes. *Genes Dev* 5:1946-1956
- Tjian R, Maniatis T 1994 Transcriptional activation: a complex puzzle with few easy pieces. *Cell* 77:5-8
- Tora L, Mullick A, Metzger D, Ponglikitmongkol M, Park I, Chambon P 1989 The cloned human oestrogen receptor contains a mutation which alters its hormone binding properties. *EMBO J* 8:1981-1986
- Towbin H, Staehelin T, Gordon J 1979 Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedures and some applications. *Proc Natl Acad Sci USA* 76:4350-4354
- Trapman J, Klaassen P, Kuiper GGJM, van der Korput JAGM, Faber PW, van Rooij H CJ, Geurts van Kessel A, Voorhorst MM, Mulder E, Brinkmann AO 1988 Cloning, structure and expression of a cDNA encoding the human androgen receptor. *Biochem Biophys Res Commun* 153:241-248
- Traut TW 1988 Do exons code for structural or functional units in proteins? *Proc Natl Acad Sci USA* 85:2944-2948
- Truss M, Chalepakis G, Beato M 1992a Interplay of steroid hormone receptors and transcription factors on the mouse mammary tumor virus promoter. *J Steroid Biochem Mol Biol* 43:365-378
- Truss M, Beato M 1993 Steroid hormone receptors: interaction with deoxyribonucleic acid and transcription factors. *Endocr Rev* 14:459-479
- Tsai SY, Carlstedt-Duke J, Weigel NL, Dahlman K, Gustafsson J-A, Tsai M-J, O'Malley BW 1988 Molecular interactions of steroid hormone receptor with its enhancer element: evidence for receptor dimer formation. *Cell* 55:361-369
- Tsai SY, Tsai M-J, O'Malley BW 1989 Cooperative binding of steroid hormone receptors contributes to transcriptional synergism at target enhancer elements. *Cell* 57:443-448
- Tung L, Mohamed MK, Hoeffler JP, Takimoto GS, Horwitz KB 1993 Antagonist-occupied human progesterone B-receptors activate transcription without binding to progesterone elements and are dominantly inhibited by A-receptors. *Mol Endocrinol* 7:1256-1266
- Tzukerman MT, Esty A, Santiso-Mere D, Danielian P, Parker MG, Stein RB, Pike JW, McDonnell DP 1994 Human estrogen receptor transactivational capacity is determined by both cellular and promoter context and mediated by two functionally distinct intramolecular regions. *Mol Endocrinol* 8:21-30
- Umesono K, Evans RM 1989 Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* 57:1139-1146
- Vallee BL, Coleman JE, Auld DS 1991 Zinc fingers, zinc clusters, and zinc twists in DNA-binding protein domains. *Proc Natl Acad Sci USA* 88:999-1003
- van der Kwast ThH, Schalken J, Ruizeveld de Winter JA, van Vroonhoven CCJ, Mulder E, Boersma W, Trapman J 1991 Androgen receptors in endocrine-therapy-resistant human prostate cancer. *Int J Cancer*

## References

- 48:189-193
- van Driel R, Humbel B, de Jong L 1991 The nucleus: a black box being opened. *J Cell Biochem* 47:311-316
- van Laar JH, Voorhorst-Ogink MM, Zegers ND, Boersma WJA, Claassen E, van der Korput JAGM, Ruizeveld de Winter JA, van der Kwast ThH, Mulder E, Trapman J, Brinkmann AO 1989 Characterization of polyclonal antibodies against the N-terminal domain of the human androgen receptor. *Mol Cell Endocrinol* 67:29-38
- van Laar JH, Bolt-de Vries J, Zegers ND, Trapman J, Brinkmann AO 1990 Androgen receptor heterogeneity and phosphorylation in human LNCaP cells. *Biochem Biophys Res Commun* 166:193-200
- van Laar JH, Berrevoets CA, Trapman J, Zegers ND, Brinkmann AO 1991 Hormone-dependent androgen receptor phosphorylation is accompanied by receptor transformation in human lymph node carcinoma of the prostate cells. *J Biol Chem* 266:3734-3738
- van Steensel B, van Haarst AD, de Kloet ER, van Driel R 1991 Binding of corticosteroid receptors to rat hippocampus nuclear matrix. *FEBS Lett* 292:229-231
- van Wijnen AJ, Bidwell JP, Fey EG, Penman S, Lian JB, Stein JL, Stein GS 1993 Nuclear matrix association of multiple sequence-specific DNA binding activities related to SP-1, ATF, CCAAT, C/EBP, OCT-1, and AP-1. *Biochemistry* 32:8397-8402
- Vegeto E, Shahbaz MM, Wen DX, Goldman ME, O'Malley BW, McDonnell DP 1993 Human progesterone receptor A form is a cell- and promoter-specific repressor of human progesterone receptor B function. *Mol Endocrinol* 7:1244-1255
- Veldscholte J, Ris-Stalpers C, Kuiper GGJM, Jenster G, Berrevoets C, Claassen E, van Rooij HCJ, Trapman J, Brinkmann AO, Mulder E 1990 A mutation in the ligand binding domain of the androgen receptor of human LNCaP cells affects steroid binding characteristics and response to anti-androgens. *Biochem Biophys Res Commun* 173:534-540
- Veldscholte J, Berrevoets CA, Zegers ND, van der Kwast ThH, Grootegoed JA, Mulder E 1992a Hormone-induced dissociation of the androgen receptor-heat-shock protein complex: use of a new monoclonal antibody to distinguish transformed from nontransformed receptors. *Biochemistry* 31:7422-7430
- Veldscholte J, Berrevoets CA, Brinkmann AO, Grootegoed JA, Mulder E 1992b Anti-androgens and the mutated androgen receptor of LNCaP cells: differential effects on binding affinity, heat shock protein interaction, and transcription activation. *Biochemistry* 31:2393-2399
- Verheijen R, van Venrooij W, Ramaekers F 1988 The nuclear matrix: structure and composition. *J Cell Sci* 90:11-36
- Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu Y-H, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, Eussen BE, van Ommen G-JB, Blonden LAJ, Riggins GJ, Chastain JL, Kunst CB, Galjaard H, Caskey CT, Nelson DL, Oostra BA, Warren ST 1991 Identification of a gene (FMR-1) containing a CCG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65:905-914
- Wagner P, Kunz J, Koller A, Hall MN 1990 Active transport of proteins into the nucleus. *FEBS Lett* 275:1-5
- Wahli W, Martinez E 1991 Superfamily of steroid nuclear receptors: positive and negative regulators of gene expression. *FASEB J* 5:2243-2249
- Walsh CT, Zydowski LD, McKoen FD 1992 Cyclosporin A, the cyclophilin class of peptidylprolyl isomerases, and blockage of T cell signal transduction. *J Biol Chem* 267:13115-13118
- Walter P, Green S, Greene G, Krust A, Bornert J-M, Jeltsch JM, Staub A, Jensen E, Scrcce G, Waterfield M, Chambon P 1985 Cloning of the human estrogen receptor cDNA. *Proc Natl Acad Sci USA* 82:7889-7893
- Wansink DG, Schul W, van der Kraan I, van Steensel B, van Driel R, de Jong L 1993 Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J Cell Biol* 122:283-293
- Washburn T, Hocutt A, Brautigan DL, Korach KS 1991 Uterine estrogen receptor *in vivo*: phosphorylation of nuclear specific forms on serine residues. *Mol Endocrinol* 5:235-242
- Webster NJG, Green S, Jin JR, Chambon P 1988 The hormone-binding domains of the estrogen and glucocorticoid receptors contain an inducible transcription activation function. *Cell* 54:199-207
- Webster NJG, Green S, Tasset D, Ponglikitmongkol M, Chambon P 1989 The transcriptional activation function located in the hormone-binding domain of the human estrogen receptor is not encoded in a single exon. *EMBO J* 8:1441-1446
- Wei LL, Gonzalez-Aller C, Wood WM, Miller LA, Horwitz KB 1990 5'-Heterogeneity in human progesterone receptor transcripts predicts a new amino-terminal truncated "C"-receptor and unique A-receptor messages. *Mol Endocrinol* 4:1833-1840

- Westphal U 1978 In: O'Malley BW, Birnbaumer L (eds) Receptors and hormone action II. Acad Press, New York, pp 443-473
- Wharton KA, Yedvobnick B, Finnerty VG, Artavanis-Tsakonas S 1985 Opa: a novel family of transcribed repeats shared by the notch locus and other developmentally regulated loci in *D. melanogaster*. *Cell* 40:55-62
- White J, Brou C, Wu J, Lutz Y, Moncollin V, Chambon P 1992 The acidic transcriptional activator GAL-VP16 acts on preformed template-committed complexes. *EMBO J* 11:2229-2240
- Wieland S, Döbbeling U, Rusconi S 1991 Interference and synergism of glucocorticoid receptor and octamer factors *EMBO J* 10:2513-2521
- Wikström A-C, Bakke O, Okret S, Brönnegard M, Gustafsson J-A 1987 Intracellular localization of the glucocorticoid receptor: evidence for cytoplasmic and nuclear localization. *Endocrinology* 120:1232-1242
- Williams T, Tjian R 1991 Analysis of the DNA-binding and activation properties of the human transcription factor AP-2. *Genes Dev* 5:670-682
- Wilson EM, Colvard DS 1984 Factors that influence the interaction of androgen receptors with nuclei and nuclear matrix. *Ann NY Acad Sci* 438:85-100
- Wilson CM, McPhaul MJ 1994 A and B forms of the androgen receptor are present in human genital skin fibroblasts. *Proc Natl Acad Sci USA* 91:1234-1238
- Wolffe AP 1994 Transcription: in tune with the histones. *Cell* 77:13-16
- Wong C-i, Zhou Z-x, Sar M, Wilson EM 1993 Steroid requirement for androgen receptor dimerization and DNA binding. *J Biol Chem* 268:19004-19012
- Wooster R, Mangion J, Eeles R, Smith S, Dowsett M, Averill D, Barrett-Lee P, Easton DF, Ponder BAJ, Stratton MR 1992 A germline mutation in the androgen receptor gene in two brothers with breast cancer and Reifenstein syndrome. *Nature Genet* 2:132-134
- Wrange O, Eriksson P, Perlmann T 1989 The purified activated glucocorticoid receptor is a homodimer. *J Biol Chem* 264:5253-5259
- Wrenn CK, Katzenellenbogen BS 1993 Structure-function analysis of the hormone binding domain of the human estrogen receptor by region-specific mutagenesis and phenotypic screening in yeast. *J Biol Chem* 268:24089-24098
- Wright PE 1989 What can two-dimensional NMR tell us about proteins? *Trends Biochem Sci* 14:255-260
- Wright APH, Gustafsson J-A 1991 Mechanism of synergistic transcriptional activation by the human glucocorticoid receptor. *Proc Natl Acad Sci USA* 88:8283-8287
- Wright APH, McEwan IJ, Dahlman-Wright K, Gustafsson J-A 1991 High level expression of the major transactivation domain of the human glucocorticoid receptor in yeast cells inhibits endogenous gene expression and cell growth. *Mol Endocrinol* 5:1366-1372
- Wright APH, Zilliacus J, McEwan IJ, Dahlman-Wright K, Almlöf T, Carlstedt-Duke J, Gustafsson J-A 1993 Structure and function of the glucocorticoid receptor. *J Steroid Biochem Mol Biol* 47:11-19
- Yamamoto KK, Gonzalez GA, Biggs WH, Montminy MR 1988 Phosphorylation-induced binding and transcriptional efficacy of nuclear factor CREB. *Nature* 334:494-498
- Ylikomi T, Bocquel MT, Berry M, Gronemeyer H, Chambon P 1992 Cooperation of proto-signals for nuclear accumulation of estrogen and progesterone receptors. *EMBO J* 11:3681-3694
- Zawel L, Reinberg D 1992 Advances in RNA polymerase II transcription. *Curr Opin Cell Biol* 4:488-495
- Zawel L, Reinberg D 1993 Initiation of transcription by RNA polymerase II: a multi-step process. *Prog Nucleic Acid Res* 44:67-108
- Zechel C, Shen X-Q, Chambon P, Gronemeyer H 1994a Dimerization interfaces formed between the DNA binding domains determine the cooperative binding of RXR/RAR and RXR/TR heterodimers to DR5 and DR4 elements. *EMBO J* 13:1414-1424
- Zechel C, Shen X-Q, Chen J-Y, Chen Z-P, Chambon P, Gronemeyer H 1994b The dimerization interfaces formed between the DNA binding domains of RXR, RAR and TR determine the binding specificity and polarity of the full-length receptors to direct repeats. *EMBO J* 13:1425-1433
- Zegers ND, Claassen E, Neelen C, Mulder E, van Laar JH, Voorhorst MM, Berrevoets CA, Brinkmann AO, van der Kwast ThH, de Winter JAR, Trapman J, Boersma WJA 1991 Epitope prediction and confirmation for the human androgen receptor: generation of monoclonal antibodies for multi-assay performance following the synthetic peptide strategy. *Biochim Biophys Acta* 1073:23-32
- Zhou Z-x, Sar M, French FS, Wilson EM 1993 Molecular biological aspects of the human androgen receptor relating to disease. In: Moudgil VK (ed) Steroid hormone receptors: basic and clinical aspects. Birkhäuser, Boston, pp 407-426
- Zhou Z-x, Sar M, Simental JA, Lane MV, Wilson EM 1994a A ligand-dependent bipartite nuclear targeting

## References

- signal in the human androgen receptor. *J Biol Chem* 269:13115-13123
- Zhou Z-x, Wong C-i, Sar M, Wilson EM 1994b The androgen receptor: an overview. In: Bardin CW (ed) *Recent progress in hormone research*, Academic Press, Inc, San Diego, vol 49:249-274
- Zilliacus J, Dahlman-Wright K, Wright A, Gustafsson J-A, Carlstedt-Duke J 1991 DNA binding specificity of mutant glucocorticoid receptor DNA-binding domains. *J Biol Chem* 266:3101-3106
- Zlatanova JS, van Holde KE 1992 Chromatin loops and transcriptional regulation. *Crit Rev Eukaryot Gene Expr* 2:211-224
- Zoppi S, Wilson CM, Harbison MD, Griffin JE, Wilson JD, McPhaul MJ, Marcelli M 1993 Complete testicular feminization caused by an amino-terminal truncation of the androgen receptor with downstream initiation. *J Clin Invest* 91:1105-1112

## SUMMARY

The steroid hormones testosterone (T) and dihydrotestosterone (DHT) exert their functions concerning male sex differentiation and fertility, through the intracellular androgen receptor (AR). In the absence of ligand, the AR is phosphorylated and associated with heat shock proteins (hsps). Upon binding of T or DHT, the hormone-AR complex accumulates in the nucleus and associates with the nuclear matrix (NM). Subsequently, the AR interacts as a dimer with specific DNA sequences, located in the flanking regions of target genes, resulting in the up- or down-regulation of the expression of these genes. The AR belongs to the superfamily of ligand-dependent transcription regulators, of which all members typically contain at least four functional domains: 1) an N-terminal modulating domain, 2) a DNA binding domain (DBD), 3) a hinge region, and 4) a C-terminal ligand binding domain (LBD) (Chapter 1). Series of AR deletion- and substitution mutants were constructed, and transfected into cultured cells to express the various AR proteins. The ability of these receptor proteins to execute the different AR features was examined.

### *Post-translational modification (Chapter 2)*

The unliganded AR migrates as a doublet of 110 kDa and 112 kDa isotypes, when analyzed by SDS-PAGE. The upshift of the 110 kDa form into the 112 kDa isotype reflects phosphorylation, which is dependent on the Ser-80 and Ser-93 Ser-Pro directed kinase consensus sites. The (non-metabolizable) androgen R1881 induces a second migration upshift resulting in a 110-112-114 kDa AR triplet. This ligand-induced 114 kDa upshift is dependent on the Ser-80 consensus site, and the amount of the 114 kDa isotype is decreased for AR mutants unable to bind DNA or unable to activate transcription. These observations suggest a multi-step process of AR phosphorylation: a basal phosphorylation of sites Ser-80 and Ser-93 and two hormone-dependent phosphorylation steps, of which one step occurs during or following transcription activation. Although the ligand-induced AR upshift implies a role for phosphorylation in AR transactivation, substitution of the proposed Ser-80 or Ser-93 phosphorylation sites had no effect on the AR's capacity to activate transcription.

### *Heat shock protein association (Chapter 3)*

After translation, the AR binds different heat shock proteins (hsps) with apparent molecular masses of 90 (hsp90), 70 (hsp70), and 56 (hsp56) kDa. Hsp90 binds to the AR LBD and dissociates from the AR upon ligand binding. Although the detailed function of these associated hsps is at present unknown, it has been suggested that they are important for ligand binding, receptor trafficking and nuclear import.

### *Subcellular localization and nuclear import (Chapter 4)*

The large AR protein needs specific nuclear localization signals (NLSs) to enter the nucleus. The NLS essential for proper AR nuclear import, is bipartite and consists of two (basic) lysine- and arginine-rich motifs, spaced by 10 amino acids. The signal is located directly C-terminal of the DBD, and is highly conserved among steroid receptors and similar to NLSs in many other nuclear proteins.

### *Steroid binding (Chapter 5)*

Androgens bind with high affinity and specificity to the ligand binding domain (LBD) which is located in the C-terminal part of the AR. The LBD has a length of approximately 250 amino acids and is conserved among the nuclear receptor family members. Binding of ligand transforms the AR to an active state. In general, the LBD is considered to be an inhibiting

## Summary

domain because its deletion results in a constitutively active transcription factor. Almost any mutation in the LBD affects the ability to bind androgens, which is exemplified by the vast number of amino acid substitutions identified in patients with androgen insensitivity.

### *Dimerization (Chapter 6)*

Specific and cooperative DNA binding is facilitated by receptor dimerization. Two different dimerization domains have been identified in steroid receptors: 1) the D-box located in the second zinc-finger of the DBD, and 2) a hydrophobic heptad repeat in the LBD.

### *Nuclear matrix association (Chapter 7)*

Upon ligand binding, the AR strongly associates with the nuclear matrix (NM), which is facilitated by the AR LBD. Deletion of the LBD revealed a second and weaker NM binding which involves the N-terminal domain and/or DBD. AR-NM association might position and anchor the heteromeric transcription machinery to specific sites and might ensure efficient gene transcription.

### *DNA binding (Chapter 8)*

The DBD consists of two zinc-finger structures in which in each module a zinc ion is liganded by four cysteines. The N-terminal zinc-finger specifically interacts with the base pairs of the DNA element to which nuclear receptors bind (termed hormone response element [HRE]). The second zinc-finger is mainly involved in the protein-protein interactions of the two receptor molecules that bind to the palindromic HRE.

### *Transcription regulation (Chapter 9)*

The N-terminal domain is essential for AR transcription activation and constitutes more than half of the entire receptor. This domain is absolutely not conserved among the nuclear receptor family members. Delimitation of the transcription activation unit (TAU), responsible for full transactivity of the wild type AR, revealed that almost the entire N-terminal domain was necessary (residues 1 to 485). In contrast, a small C-terminal part of the N-terminal domain (residues 360 to 528) was sufficient for the constitutively active AR (lacking its LBD) to induce full transcription. This demonstrates the capacity of the AR to use different N-terminal regions to induce transactivation. It is believed that TAUs represent interfaces that facilitate protein-protein interactions. Through association with other transcription factors, the AR can initiate the assembly of the basal transcription machinery and thereby regulate the expression of target genes.

### *AR mutations and disease (Chapter 10)*

A large number of mutations have been identified in the AR, which result in impairment or loss of AR functioning in individuals with androgen insensitivity syndrome (AIS). The complete form of AIS is characterized by 46, XY, external female phenotype. Interestingly, mutations are restricted to the DBD and LBD, although few N-terminal mutations (all resulting in premature translational termination) have been identified.

Kennedy's disease (a motor-neuron degenerative disorder) is correlated with an extension of the N-terminal polymorphic glutamine stretch from an average of 21 residues to more than 40 glutamines. AR mutants containing an extended Gln-repeat were less capable of activating transcription, which corresponds to the clinical features of gynaecomastia and reduced fertility in some Kennedy's disease patients.

Although some AR mutations have been identified in prostate cancer and male breast cancer, the total group of tumors screened is large and it can be concluded that AR mutations are not a common inducer of these cancers.

## SAMENVATTING

De steroïdhormonen testosteron (T) en dihydrotestosteron (DHT) zijn verantwoordelijk voor de ontwikkeling en het functioneren van de mannelijke geslachtsorganen en kenmerken. T en DHT oefenen hun werking uit door specifiek te binden aan de androgeenreceptor (AR), een eiwit dat in androgeen-doelwitcellen tot expressie komt. Na synthese wordt de AR gefosforyleerd en associeert zich tevens met "heat-shock" eiwitten (hsp's). Binding van T of DHT aan de AR resulteert in een ophoping van het AR-hormoon complex in de celkern en in de associatie van het complex met de kernmatrix (NM). Het AR-hormoon complex is in staat aan specifieke DNA-elementen te binden die gelokaliseerd zijn in de promotorgebieden van doelwitgenen. Deze DNA-binding resulteert in de modulatie van de expressie van deze genen. De AR is een lid van de grote familie van ligand-afhankelijke transcriptie-regulatoren. De leden van deze kernreceptorfamilie zijn opgebouwd uit tenminste vier functionele domeinen: 1) het N-terminale transactivatie domein, 2) het DNA-bindend domein, 3) het "hinge"-gebied en 4) het ligand-bindend domein (Hoofdstuk 1). Verschillende constructen werden gemaakt die coderen voor AR mutanten waarin een deel van het receptoreiwit ontbreekt of waarvoor één of meerdere aminozuren vervangen zijn. Deze AR mutanten zijn vervolgens tot expressie gebracht in cellen en zijn getest op hun capaciteit om de bovengenoemde AR functies uit te voeren.

Met behulp van eiwitelectroforese- en "blotting"-technieken kan de wild-type AR (geïsoleerd uit cellen die gekweekt zijn in de afwezigheid van ligand) zichtbaar gemaakt worden als een doublet van 110-112 kDa (Hoofdstuk 2). Het 112 kDa AR isotype ontstaat uit het 110 kDa isotype door fosforylering. Deze overgang is afhankelijk van de Ser-80 en Ser-93 aminozuren die mogelijk gefosforyleerd worden door het Ser-Pro-specifiek kinase. Het (synthetische) ligand R1881 induceert een tweede overgang hetgeen resulteert in een 110-112-114 kDa AR triplet. De vorming van het ligand-geïnduceerde 114 kDa isotype vindt niet plaats in een AR mutant waarin Ser-80 vervangen is. De hoeveelheid van het hormoon-geïnduceerde isotype is minder in AR mutanten die niet in staat zijn DNA te binden of niet in staat zijn transcriptie van een reportergen te activeren. Deze bevindingen suggereren dat AR fosforylering in meerdere stappen verloopt: een basale fosforylering op aminozuren Ser-80 en Ser-93 en twee ligand-gestimuleerde fosforyleringsstappen waarvan één stap pas tijdens of na transactivatie optreedt. Ofschoon deze theorie een rol voor fosforylering in AR transactivatie impliceert, zijn de AR mutanten waarin de aminozuren Ser-80 of Ser-93 zijn vervangen, normaal in staat transcriptie te activeren (Hoofdstuk 2).

Na synthese bindt de AR aan heat-shock eiwitten (hsp's). Het hsp met een molecuulmassa van 90 kDa associeert met de AR via het ligand bindend domein. Na hormoonbinding dissocieert het AR-hsp complex. De rol van de geassocieerde hsp's in het functioneren van de AR is tot op heden onbekend. Mogelijk spelen heat-shock eiwitten een rol bij het transport van de AR naar de celkern en bij import in de celkern (Hoofdstuk 3).

De AR heeft een specifiek kernlokalisatiesignaal nodig om de celkern binnen te kunnen. Dit signaal is opgebouwd uit twee motieven (bestaande uit vooral basische aminozuren) die tien aminozuren van elkaar af liggen. Het kernlokalisatiesignaal is sterk geconserveerd tussen de verschillende steroïdreceptoren en vergelijkbare kernlokalisatiesignalen komen voor in veel andere kerneiwitten. Het AR-kernlokalisatiesignaal ligt op de grens van het DNA-bindend domein en het "hinge"-gebied (Hoofdstuk 4).

T en DHT binden met hoge affiniteit en specificiteit aan het ligand-bindend domein. Dit

## Samenvatting

domein betreft ongeveer de laatste 250 aminozuren van de AR, waarvan de aminozuurvolgorde redelijk geconserveerd is tussen de verschillende leden van de kernreceptorfamilie. Het ligand-bindend domein wordt beschouwd als een remmend domein, omdat deletie van het ligand-bindend domein resulteert in een constant actieve transcriptiefactor. In de intacte AR wordt deze remming van het ligand-bindend domein opgeheven door hormoonbinding. Bijna elke aminozuurverandering in dit domein resulteert in verlies of een afname van het vermogen om ligand te binden. Dit komt goed tot uitdrukking in het grote aantal mutaties in het AR ligand-bindend domein die zijn geïdentificeerd in patiënten met het androgeen-ongevoelighedsyndroom (Hoofdstukken 5, 10).

Specifieke en coöperatieve DNA-binding wordt mede bepaald door AR dimerisatie. Twee verschillende dimerisatiedomeinen zijn beschreven in steroidreceptoren: 1) De zogenaamde D-box, gesitueerd in de tweede zink-vinger van het DNA-bindend domein, en 2) een domein in het ligand-bindend domein dat bestaat uit een herhaling van hydrofobe aminozuren (Hoofdstuk 6). Het is niet bekend of deze gebieden ook als dimerisatiedomeinen fungeren in de AR.

Na ligand-binding associeert de AR sterk met de kernmatrix via het ligand-bindend domein. Een tweede, veel zwakkere kernmatrix-binding, wordt duidelijk na de analyse van een mutant waarvan het ligand-bindend domein is gedeleteerd. De associatie van de AR met de kernmatrix kan een rol spelen in het plaatsen en verankeren van het transcriptie-initiatiecomplex en kan bijdragen aan de efficiëntie van genexpressie (Hoofdstuk 7).

Het DNA-bindend domein bestaat uit twee zogenaamde zink-vingers. Elke zink-vinger is opgebouwd uit ongeveer 30 aminozuren waarvan vier cysteïneresiduen een zink ion complexeren. De N-terminale zink-vinger bindt specifiek aan de baseparen van het DNA element (hormoon responsief element). De tweede zink-vinger is betrokken bij eiwit-eiwit interacties van de AR homodimeer op het palindromische hormoon responsief element (Hoofdstuk 8).

Het N-terminale domein van de AR is essentieel voor het vermogen om transcriptie te activeren. Dit domein is opgebouwd uit meer dan 500 aminozuren waarbij de aminozuurvolgorde niet geconserveerd is tussen de leden van de kernreceptorfamilie. Vrijwel het gehele N-terminale domein (aminozuren 1-485) is betrokken bij de transactivatie van de wild-type AR. Dit in tegenstelling tot de "transcriptie-activatie-unit" (TAU) die verantwoordelijk is voor de transactiviteit van de constitutief actieve AR mutant (die het ligand-bindend domein mist). In deze situatie is het gebied tussen de aminozuren 360 en 528 voldoende voor de volledige activiteit van de AR mutant. Deze bevindingen tonen aan dat de AR in staat is verschillende gebieden van het N-terminale domein te gebruiken als TAU. Het wordt aangenomen dat TAU's gebieden zijn waar eiwit-eiwit interacties plaatsvinden. Door associatie met andere transcriptiefactoren initieert de AR de opbouw van het pre-initiatiecomplex en reguleert daarmee de expressie van doelwitgenen (Hoofdstuk 9).

Vrijwel alle mutaties geïdentificeerd in individuen met het androgeen-ongevoelighedsyndroom betreffen veranderingen in het ligand-bindend domein en in het DNA-bindend domein. De enkele geïdentificeerde AR mutaties in het N-terminale domein leiden allemaal tot een vroegtijdige translatie-stop (Hoofdstuk 10). "Kennedy's disease" (een stoornis die gepaard gaat met een degeneratie van motor-neuronen) is gecorreleerd met een verlenging van een polyglutamine-keten in het N-terminale domein van de AR. De normale lengte van deze keten is gemiddeld 21 glutamines. Echter, in patiënten met "Kennedy's disease" is de lengte van de glutamine-keten toegenomen tot meer dan 40 residuen. AR mutanten met deze verlengde



glutamine-keten zijn minder goed in staat transcriptie te activeren. Deze verminderde AR activiteit komt overeen met de mogelijke klinische symptomen gynecomastie en verminderde fertiliteit (Hoofdstuk 10). Enkele AR mutaties zijn geïdentificeerd in prostaatkanker en in mannelijke borstkanker. Echter, de totale groep van tumoren getest op AR mutaties is groot, en aangenomen mag worden dat mutaties van de AR niet het algemene mechanisme is dat leidt tot het ontstaan, het uitzaaien, of het hormoon-onafhankelijk worden van dit soort tumoren.

LIST OF PUBLICATIONS

- Veldscholte J, Ris-Stalpers C, Kuiper GGJM, Jenster G, Berrevoets C, Claassen E, van Rooij HCJ, Trapman J, Brinkmann AO, Mulder E 1990 A mutation in the ligand binding domain of the androgen receptor of human LNCaP cells affects steroid binding characteristics and response to anti-androgens. *Biochem Biophys Res Commun* 173:534-540
- Ris-Stalpers C, Trifiro MA, Kuiper GGJM, Jenster G, Romalo G, Sai T, van Rooij HCJ, Kaufman M, Rosenfield RL, Liao S, Schweikert H-U, Trapman J, Pinsky L, Brinkmann AO 1991 Substitution of aspartic acid-686 by histidine or asparagine in the human androgen receptor leads to a functionally inactive protein with altered hormone binding characteristics. *Mol Endocrinol* 5:1562-1569
- Jenster G, van der Korput JAGM, van Vroonhoven C, van der Kwast ThH, Trapman J, Brinkmann AO 1991 Domains of the human androgen receptor involved in steroid binding, transcriptional activation, and subcellular localization. *Mol Endocrinol* 5:1396-1404
- Jenster G, van der Korput JAGM, Trapman J, Brinkmann AO 1992 Functional domains of the human androgen receptor. *J Steroid Biochem Mol Biol* 41:671-675
- Brinkmann AO, Jenster G, Kuiper GGJM, Ris C, van Laar JH, van der Korput JAGM, Degenhart HJ, Trifiro MA, Pinsky L, Romalo G, Schweikert H-U, Veldscholte J, Mulder E, Trapman J 1992 The human androgen receptor: structure/function relationship in normal and pathological situations. *J Steroid Biochem Mol Biol* 41:361-368
- Veldscholte J, Berrevoets CA, Ris-Stalpers C, Kuiper GGJM, Jenster G, Trapman J, Brinkmann AO, Mulder E 1992 The androgen receptor in LNCaP cells contains a mutation in the steroid binding domain which affects steroid binding characteristics and response to antiandrogens. *J Steroid Biochem Mol Biol* 41:665-669
- van der Kwast ThH, Zegers ND, Jenster G, Boersma W, Brinkmann AO, Trapman J 1992 Multi-assay performance of a monoclonal antibody directed against the human androgen receptor. *Prog Histo- Cytochemistry* 26:61-67
- Jenster G, Trapman J, Brinkmann AO 1993 Nuclear import of the human androgen receptor. *Biochem J* 293:761-768
- Kuiper GGJM, de Ruiter PE, Trapman J, Jenster G, Brinkmann AO 1993 In vitro translation of androgen receptor cRNA results in an activated receptor protein. *Biochem J* 296:161-167
- Jenster G, van der Korput JAGM, Trapman J, Brinkmann AO 1994 Identification of two transcription activation units in the N-terminal domain of the human androgen receptor. submitted
- Jenster G, de Ruiter PE, van der Korput JAGM, Kuiper GGJM, Trapman J, Brinkmann AO 1994 Changes in abundance of androgen receptor isotypes: effects of ligand treatment, glutamine-stretch variation, and mutation of putative phosphorylation sites. submitted
- van Steensel B, Jenster G, Damm K, Brinkmann AO, van Driel R 1994 Domains of the human androgen receptor and glucocorticoid receptor involved in binding to the nuclear matrix. *J Cell Biochem*, in press

## CURRICULUM VITAE

Guido Jenster was born on the 5<sup>th</sup> of May 1965 in the town of Boxmeer in The Netherlands. His secondary education (VWO-B) was taken at the Elshof College in Nijmegen. In 1983 he started the Biology study at the Agricultural University in Wageningen. His three doctoral stages included:

- Expression of oncogenes, growth factors and growth factor receptors in prostate carcinoma cell lines. Project leader: Dr. J. Trapman, Department of Pathology, Erasmus University Rotterdam, Rotterdam, The Netherlands
- Characterization of monoclonal antibodies against the putative growth factor receptor *neu*. Project leaders: Dr. M. van der Vijver and Dr. R. Nusse, The Netherlands Cancer Institute, Amsterdam, The Netherlands
- Expression of *bcr*, *abl* and the chimeric *bcr/abl* genes in transgenic mice. Project leaders: Dr. J. Groffen and Dr. N. Heisterkamp, Department of Pathology Research, Childrens Hospital of Los Angeles, Los Angeles, USA

After graduation in Cell-Biology in August 1989, a start was made with the PhD research in February 1990 at the Department of Endocrinology & Reproduction at the Erasmus University Rotterdam under supervision of Dr. A.O. Brinkmann, Prof.Dr. J.A. Grootegoed (Department of Endocrinology & Reproduction), and Dr.Ir. J. Trapman (Department of Pathology).

The writer of this thesis has been awarded a TALENT-fellowship by the Netherlands Organisation for Scientific Research to perform research over an one year period with Prof. B.W. O'Malley at the Department of Cell Biology, Baylor College of Medicine, Houston, Texas, USA, starting August 1994.

## DANKWOORD

Het mag duidelijk zijn dat dit proefschrift niet tot stand zou zijn gekomen zonder de hulp van anderen. Een woord van dank is daarom op zijn plaats.

Allereerst wil ik Hetty bedanken voor het uitstekend uitvoeren van vele (vaak dezelfde, saaie) experimenten en voor de zeer prettige samenwerking die de leuke publikaties en dit proefschrift tot resultaat hebben gehad.

Mijn twee begeleiders en co-promotoren Albert en Jan hebben me gezamenlijk het tweezijdige karakter van de wetenschap laten zien: het subjectieve en het objectieve, het pipeteren en het schrijven, het bedenken en het uitvoeren, het opkomen en het afgaan, Mol Endo en de niet reproduceerbaarheid (verzwakte versie van "Cell or hell"). De combinatie van een "vaderlijke" begeleiding en een realistische, nuchtere kijk op wetenschap hebben mijn wat jeugdige overenthousiasme tot een (hopelijk) acceptabel niveau ingetoomd.

Mijn dank gaat zeker ook uit naar het gezellige lab "Androdroom". De echte experimentatoren Cor B., Marja, Theo en Petra bedank ik voor het opzetten en in stand houden van de vele basisfaciliteiten op het lab. Petra neemt hier een speciale plaats in omdat ik het genoeg heb gehad om samen met haar de misschien wel 5 of 6 (en nog meer) androgeenreceptor isotypen te ontdekken. Eppo, Coos, Erna, Carrie, Jos, George, Cor K., Hennie en Annemie wil ik bedanken voor de spontane en geplande, interessante discussies en werkreflecties. George met zijn oninterrupteerbare eigen uit-frame gedachtengangen en Jos z'n onnavolgbare, niet te volgen, maar hard genoeg gesproken redeneringen hebben mij op heel wat nieuwe hypothesen gebracht. Jammer dat ze nooit waar bleken te zijn. Dankzij jullie allemaal heb ik ontdekt dat er naast "Cell or hell" ook nog zoiets is als MCE en Mol Endo.

Naast "Androdroom" ben ik alle medewerkers van de afdeling Endocrinologie & Voortplanting dank verschuldigd. De gezellige sfeer en verschillende faciliteiten op de afdeling vormen een goede basis voor al de lopende onderzoeken. Focko wil ik bedanken voor de voorspelbare, vaak gestelde, maar essentiële vragen na de colloquia en tijdens de borrels over waar je nou eigenlijk mee bezig bent en hoe reproduceerbaar die artefacten zijn. Hopelijk zijn mijn linea recta antwoorden wat dimensionaler overgekomen.

Anton, bedankt voor het kritisch meedenken met het onderzoek, het lezen en corrigeren van de publikaties en het becommentariëren van dit proefschrift.

Marc, Bonnie, Anna, Guido en Metin bedank ik voor hun inzet, gezelligheid en voor het inzicht dat ik vroeger als onervaren student (in tegenstelling tot jullie) behoorlijk lastig moet zijn geweest.

De diverse samenwerkingen met de afdelingen Pathologie, Urologie, Kindergeneeskunde, Genetica (Erasmus Universiteit Rotterdam) en Biochemie (EC Slater Instituut Universiteit van Amsterdam) hebben duidelijk hun vruchten afgeworpen. Met name wil ik Theo van der Kwast, Bas van Steensel en Roel van Driel bedanken voor de succesvolle en prettige samenwerking. Tenslotte verdient de hele groep van Jan Trapman een vermelding voor de intensieve, positieve en leuke samenwerking.

Het werk is gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO; Gebied Medische Wetenschappen); hiervoor mijn dank.

De drukkosten van dit proefschrift zijn medegefinancierd door Amersham B.V. Nederland en de Ernst Schering Research Foundation.



