Functional characterisation of the CAG polymorphism in the androgen receptor- in vitro and in vivo

Nenonen, Hannah

2011

FUNCTIONAL CHARACTERISATION OF THE CAG POLYMORPHISM IN THE ANDROGEN RECEPTOR

-IN VITRO AND IN VIVO-

HANNAH NENONEN

LUND UNIVERSITY
Faculty of Medicine

MALMÖ

2011

FUNCTIONAL CHARACTERISATION OF THE CAG POLYMORPHISM IN THE ANDROGEN RECEPTOR

-IN VITRO AND IN VIVO-

HANNAH NENONEN

ACADEMIC DISSERTATION

By due permission from the Faculty of Medicine, Lund University, Sweden, to be publicly defended on April 1st, 2011, at 09:00 in Jubileumsaulan, entrance 59, Skåne University Hospital, Malmö

FACULTY OPPONENT

Professor Philippa Saunders

Human Reproductive Sciences Unit
Centre for Reproductive Biology
The Queen's Medical Research Institute
Edinburgh, Scotland, United Kingdom

ISBN 978-91-86671-75-4
ISSN 1652-8220

Cover art: “Hm-pipp-pi” Johan Petterson ©
Lund University
Faculty of Medicine Doctoral Dissertation Series 2011:27
Department of Clinical Sciences, Malmö
Molecular Genetic Reproductive Medicine

"DNA makes RNA,
RNA makes proteins,
and proteins make us"

Francis Crick.

Till min familj

Abstract

The androgen receptor (AR) is the mediator of androgen actions. In the AR coding region there is a polymorphic CAG repeat encoding a stretch of the amino acid glutamine (Q). The repeat length modulates receptor activity and is normally distributed between 10-30 CAG with a median length of 22 repeats in white men. At the start of this work, a negative linear association between AR function and the CAG repeat number was generally assumed. This assumption was supported by clinical findings in patients with Kennedy's disease, which is a neuromuscular disorder caused by an abnormally expanded CAG repeat (>40 CAG). However, *in vivo* data concerning the association between CAG numbers within normal length and androgenic effects were conflicting. As understanding the impact of CAG number on the AR activity is important for proper interpretation of this polymorphism and risk of pathological conditions other than Kennedy's disease, the purpose of this study was to examine the influence of CAG length, if any, on AR activity.

Firstly an *in vitro* study was performed. A reporter gene with a human androgen responsive promoter was used in a transactivation assay. The repeat lengths included were 16, 22 and 28 CAG, which represent a short, the median, and a long repeat within the normal human range. The study showed that the AR with median repeat length had the highest activity *in vitro*. Secondly, the effect of the CAG repeat in relation to two androgen dependent conditions, infertility and PCa, was analysed in two separate meta-analyses. When stratifying the CAG repeats into three groups, shorter than median, median and longer than the median CAG length, the meta-analysis on infertile men showed 20% increased risk of infertility in men harbouring other repeat lengths than the most common. On the other hand, CAG number did not have any effect on PCa risk.

The AR regulates the expression of prostate specific antigen (PSA). Thus the expression of PSA can be used as a marker of AR activity in tissue. We measured the AR protein and PSA amount in human prostate tissue from 19 men with known CAG length. Those who were carriers of 22 CAG had lower AR amount and higher PSA than counterparts with other CAG lengths, but this was not statistically significant, probably due to the small study size.

Taken together, these studies indicate that the median length of the androgen receptor CAG repeat is associated with optimal activity, *in vitro* and *in vivo.*

Contents

List of abbreviations

AF-1	activation function 1
AF-2	activation function 2
AF-5	activation function 5
AIS	androgen insensitivity syndrome
AncSR1	ancestor steroid receptor 1
AncSR2	ancestor steroid receptor 2
AR	androgen receptor
ARA24	androgen receptor associated protein 24
ARA70	androgen receptor associated protein 70
BMI	body mass index
bp	base pair
CAIS	complete androgen insensitivity syndrome
CRPC	castrate resistant prostate cancer
DBD	DNA binding domain
DHT	5 α-dihydrotestosterone
ER	estrogen receptor
FSH	follicle stimulating hormone
GnRH	gonadotropin releasing hormone
GR	glucocorticoid receptor
HAT	histone acetyltransferase
HDAC6	histone deacetylation complex 6
HRE	hormone responsive element
HSP	heat shock protein
LBD	ligand binding domain
LH	luteinizing hormone
MAIS	mild androgen insensitivity syndrome
MR	mineralocorticoid receptor
NLS	nuclear localisation signal
NR	nuclear receptor
NTD	N-terminal domain
PAIS	partial androgen insensitivity syndrome
PCa	prostate cancer
PCR	polymerase chain reaction
qPCR	real-time quantitative PCR
PR	progesterone receptor
PSA	prostate specific antigen
RCC1	regulation of chromosome condensation 1
SBMA	spinal bulbar muscular artrophy
SRD5A	5 α-reductase
TF	transcription factor
TRFI	time resolved fluorescence imaging
VDR	vitamin D receptor

List of papers

I. **CAG repeat number is not inversely associated with androgen receptor activity *in vitro***
Hannah Nenonen, Christel Björk, Paal-Andre Skjaerpe, Aleksander Giwercman, Lars Rylander, Johan Svartberg, and Yvonne L Giwercman. Molecular Human Reproduction, Vol.16, No.3 pp. 153-157, 2010.

II. **Non-linear association between androgen receptor CAG repeat length and risk of male subfertility- a meta-analysis**
Hannah A Nenonen, Aleksander Giwercman, Erik Hallengren, and Yvonne L Giwercman. International Journal of Andrology, Vol.33, pp. 1-6, 2010.

III. **No difference in prostate cancer risk in relation to androgen receptor CAG repeats- a stratified analysis**
Hannah A Nenonen and Yvonne L Giwercman.
Submitted manuscript.

IV. **Evaluation of androgen receptor amount in prostate tissues from patients with known androgen receptor CAG length**
Hannah A Nenonen, Giuseppe Lippolis, Azharuddin Sajid, Anders Bjartell, and Yvonne L Giwercman.
Manuscript.

Background

Transcription factors

Proteins that bind to DNA and regulate gene transcription of other genes are termed transcription factors. Generally they play a role in positioning the DNA polymerase II at the transcription initiation site that is required for transcription to take place. Specific transcription factors only bind to specific DNA sequences in certain genes. One of these transcription factors, the androgen receptor (AR) is the main topic of this thesis.

The nuclear receptor superfamily

Nuclear receptors (NRs) are one of the largest groups of transcriptional regulators in animals [1]. They constitute a superfamily of phylogenetically related proteins and to date 49 genes for NRs have been found in humans [2]. The NRs regulate homeostasis, reproduction, development and metabolism and they bind to sequences in the DNA known as hormone responsive elements (HRE). The most evident difference between NRs and other transcription factors is their capacity to bind small hydrophobic molecules both intracellularly and specifically, and to mediate a fast response by regulating gene expression directly [3]. Based on their mode of function NRs can be divided into three main groups:

1) Steroid receptors, NRs that are activated by specific hydrophobic ligands that bind to a deep internal cavity within the protein [4].

2) Thyroid hormone, vitamin D and retinoid receptors [4].

3) Orphan receptors [4].

All members of the NR superfamily have a common structure that is divided into 5 to 6 regions [5]. The N-terminal domain (NTD) is highly variable and contains at least one constitutionally active transactivation region (AF-1). The length of this region varies greatly between receptors, from 23 amino acids (aa) in the vitamin D receptor to 550 aa in the AR, mineralocorticoid receptor and the glucocorticoid receptor. The DNA-binding domain (DBD) is the most conserved region and contains the part of the receptor that confers specific DNA sequence recognition. The DBD-domain is also involved in receptor dimerisation [1].

The hinge region of the NRs is situated between the ligand binding domain (LBD) and the DBD. It is variable and contains a nuclear localization signal (NLS). This

domain may act as a flexible linker allowing the protein to adopt different conformations [3]. The LBD of the NRs is the largest domain and is moderately conserved. The 12 α-helices making up the structure of this domain are better conserved than the sequence itself. In the LBD there are many important areas, the ligand-dependent activation function 2 (AF-2), a strong dimerization interface, and sometimes a repression function [1]. Some NRs also contain a fifth domain. The role of this domain is still unknown but it may be involved in co-factor recruitment or antagonist action [3]. NRs form monomers, homodimers, or heterodimers before they enter the cell nucleus and bind to DNA.

NRs with ligand binding domains that are unable to interact with co-activators generally act as repressors of transcription. They can bind to NR response elements on the DNA or dimerize with other receptors which then are unable to activate their target genes [6].

Orphan receptors are constitutively active, they may be bound by an unknown ligand or be ligand independent. Several of these receptors do not have any ligand binding pocket, but activate gene expression even so, others can bind hydrophobic molecules that change the baseline activity of the receptor [7].

Steroid receptors

In vertebrates there are six related steroid receptors (SRs). Based on phylogeny they are divided into two groups; the estrogen receptor α and estrogen receptor β group, and the androgen, progesterone, mineralocorticoid and glucocorticoid receptors (AR, PR, MR and GR) forming the other group. They all descend from the same ancestral receptor through two large-scale genome expansions [8]. The first existing steroid receptor, ancestral steroid receptor was an estrogen-activated receptor (AncSR1) which in the first expansion (duplication) gave rise to an ER and a 3-ketosteroid receptor (AncSR2) [9]. In the second expansion the 3-ketosteroid receptor became a corticoid receptor and a receptor for 3-ketogonadal steroids (progestins). At a later stage these receptors duplicated again, and the six SRs evolved (Fig. 1) [8].

ERα and ERβ are transcribed from two different genes. ERα was the first known estrogen receptor whereas ERβ was discovered 10 years later [10; 11]. The ligand, 17β-estradiol is the same for both receptors, and is necessary for the reproductive function and menstrual cycle in women [3]. 17β-estradiol affects bone growth and the cessation of bone growth in both men and women. In men it has also been shown to affect the brain and the reproductive tissues [12]. The PR is activated by progesterone that plays a role in various parts of the body such as normal female reproductive function especially ovulation and uterine implantation, the brain, the cardiovascular system, bone and the central nervous system [13-17].

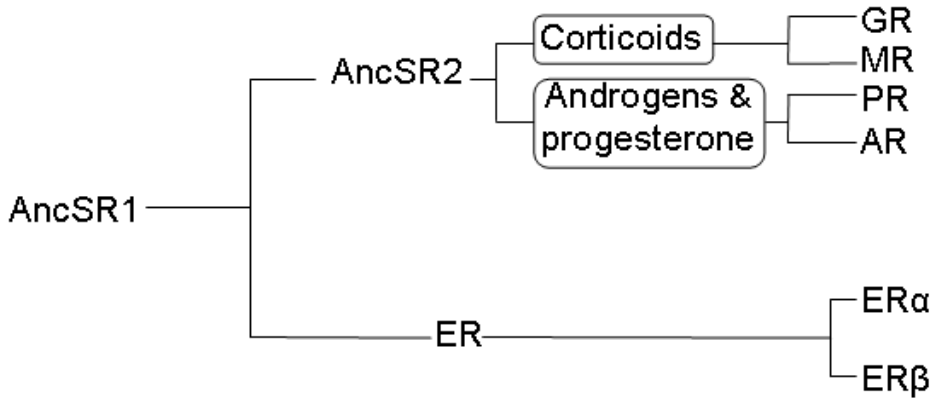

Figure 1. Schematic view of the evolution of the steroid receptors. The first receptor was ancestral steroid receptor 1 (AncSR1). The first duplication resulted in AncSR2 and ER. In the second duplication ER became ERα and ERβ, and AncSR2 became one receptor for corticoids and one for androgens and progesterones. Later the corticoid receptor developed into the GR and MR. The combined androgen and progesterone receptor became the AR and the PR.

There are two isoforms of PR, PR-A and PR-B, they overlap in their expression in female reproductive tissue but vary in their interaction with cofactors and can act on different promoters [14].

The MR and GR are closely related and partly share the same ligands, mineralocorticoids. The MR only binds mineralocorticoids, hormones that are involved in the control of salt and water homeostasis by regulating sodium, potassium and hydrogen ions across tight epithelia. The MR is not a strong transcriptional activator and can form dimers with GR [3]. The GR mainly binds glucocorticoids. Glucocorticoids are important for the metabolism of carbohydrates and lipids but also have an effect on stress response and inflammation.

In the ligand unbound state the ER and PR are located in the nucleus, the MR is present in the nucleus and cytoplasm in its unbound form but after ligand binding it becomes nuclear [18-20]. The AR and GR are predominantly cytoplasmic in their ligand unbound form [21-24]. In their inactive state SRs are bound to heat shock protein (HSP) complexes. All SRs are bound by a minimal complex which is essential for ligand responsive signalling, it consists of Hsp90, Hsp70, Hsp40, Hop and p23. The complex keeps the receptors in a conformation that allows the ligand binding pocket to be accessible [25-27]. When ligand dissociates into the cell the receptors are activated, the minimal complex dissociates, and the receptor is transported into the cell nucleus where it binds to specific DNA sequences and

regulates expression of target genes. The specific DNA sequences recognized by SRs are called hormone response elements (HREs). They are nucleotide palindromes with three base pair (bp) spacing. The ER can also bind to and transactivate from widely spaced direct repeats [3].

Androgens and androgen regulation

Androgens are steroid hormones essential for normal development of the male-specific phenotype during embryogenesis, in the establishment of sexual maturation at puberty, and in the maintenance of the male reproductive function, spermatogenesis, and sexual behaviour throughout life. Androgens also affect functions in non-reproductive tissue, such as bone, skeletal muscle and hair growth in both males and females [28]. The major circulating androgen is testosterone which to 90% is produced from cholesterol in the testis by the Leydig cells (Fig. 2). The remaining 10% is produced in the adrenal glands. Testosterone is also irreversibly converted into 5α-dihydrotestosterone (DHT) by the enzyme steroid 5α-reductase (SRD5A).

Both androgens act through the AR [29]. Testosterone has a two fold lower affinity than DHT for the AR and dissociates from the receptor five times faster than DHT [30]. The production of testosterone is controlled by the hormones in the hypothalamus-pituitary-gonadal axis (Fig. 3).

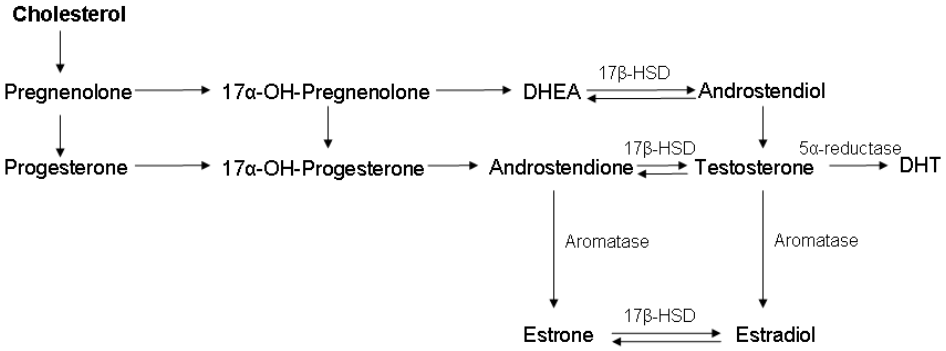

Figure 2. The androgen synthesis pathway in the Leydig cells in the testis. 17β-HSD: 17 β-hydroxysteroid dehydrogenase.

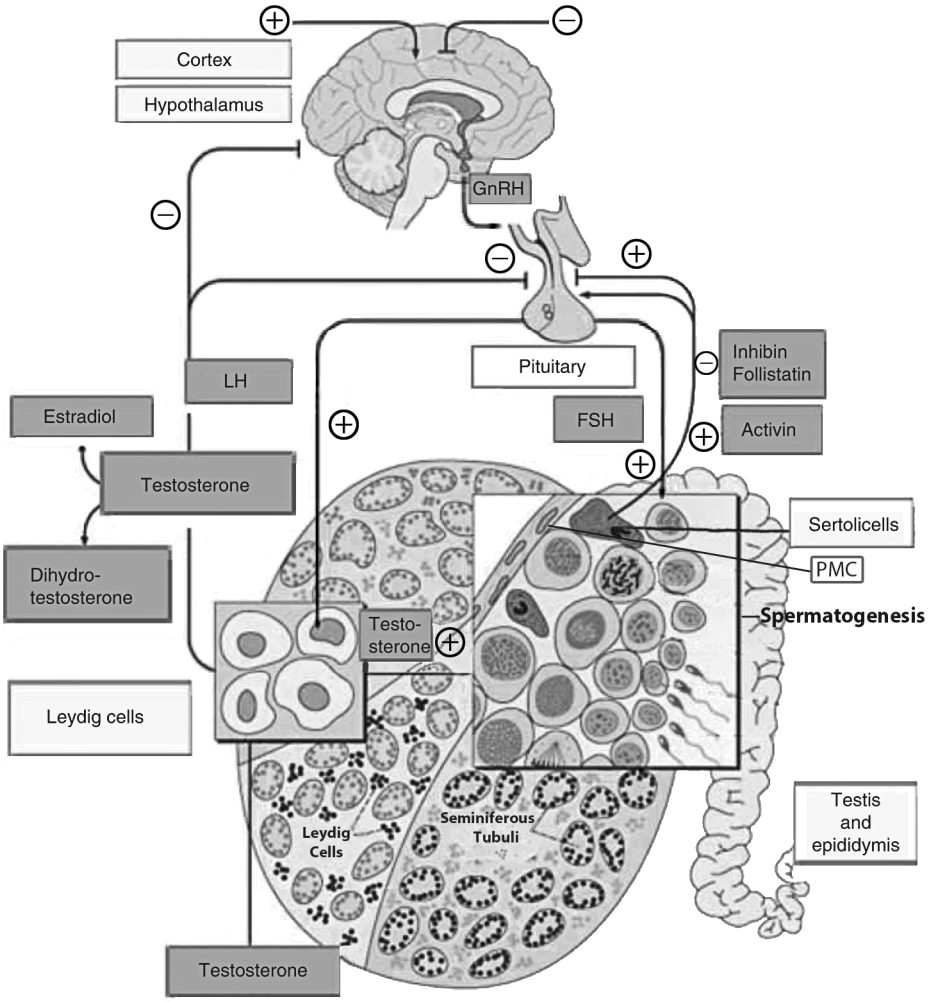

Figure 3. The hormones and cells in the hypothalamus-pituitary-gonadal axis. LH and FSH are released from the pituitary gland under the control of GnRH produced in the hypothalamus. LH acts on the Leydig cells in the testis. The Leydig cells respond by producing testosterone that stimulates spermatogenesis, acting via the Sertoli cells in the seminiferous tubulis. FSH acts directly on the Sertoli cells. Testosterone can be converted into DHT and Estradiol. Testosterone acts inhibitory on the release of GnRH. Sertoli cells produce inhibin B and follistatin that act inhibitory on the FSH release whereas activin acts stimulatory. Estradiol and dihydrotestosterone have an inhibitory effect on LH and FSH secretion. Testosterone also enters the circulation and affects other tissues. FSH: follicle stimulating hormone. LH: luteinizing hormone. PMC: peritubular myoid cells. Modified from Nieschlag et al. 3rd edition, Andrology 2010 [12]

Luteinizing hormone (LH) and follicle-stimulating hormone (FSH) are produced and released from the pituitary gland under the control of gonadotropin-releasing hormone (GnRH). Luteinizing hormone acts on the Leydig cells, found between the seminiferous tubules in the testes which produce and release testosterone that stimulates the maturation of germ cells through the Sertoli cells in the seminiferous tubules. Follicle-stimulating hormone acts directly on Sertoli cells. The action of testosterone and FSH in the testes leads to the maturation of spermatogonia into mature sperm. There is no expression of AR or the FSH receptor in the gametes. Therefore the effects of testosterone and FSH are most probably mediated by other cells in the testicle, such as the Sertoli cells.

The testosterone produced also acts inhibitory on the release of GnRH from the hypothalamus, affecting secretion of LH and FSH. In response to testosterone and FSH stimulation Sertoli cells produce inhibin B and follistatin that act inhibiting on release of FSH from the pituitary gland [12].

The androgen receptor

The androgen receptor was the last receptor in the SR family to be cloned [31; 32]. The AR mediates the androgen actions. The AR is not only expressed in reproductive organs; it also has a function and is expressed in muscles, brain, kidney, spleen, heart, liver and the salivary glands [12].

The AR gene is located on the short arm of chromosome X (Xq11-12) and includes more than 90 kilo base pairs of DNA containing eight exons and giving an approximately 2757 bp open reading frame (depending on the length of the polymorphic repeats) (Fig. 4) [33]. The AR genomic organization is conserved throughout mammalian evolution. The localization to the X-chromosome in mammals could be a sign of conservation of the particular loci including the AR and surrounding genes [34].

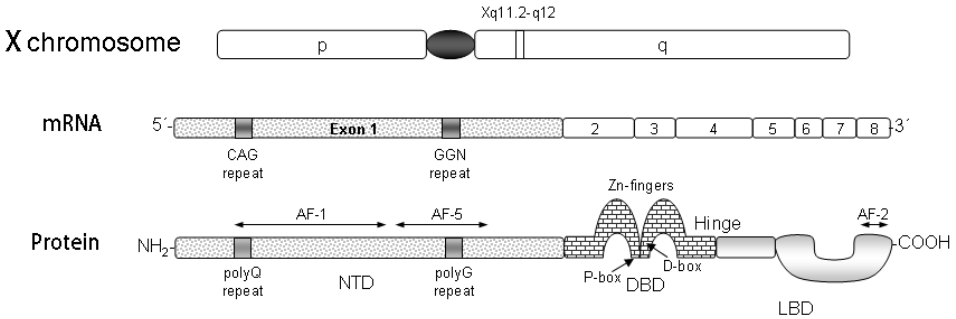

Figure 4. Schematic view of the chromosomal location, mRNA organisation, and protein domains of the AR.

The N-terminal domain

The N-terminal domain (NTD), also termed the transactivating domain, is encoded by exon 1 and is the largest part of the protein (bp 1-537 according to [33]). The other SRs do not have as large N-terminal domains as the AR and there is less than 50% homology with AR NTDs from other species [35]. To date there is no crystal structure available for the AR NTD. The NTD most probably is a largely globular and flexible domain in the ligand unbound state and may need to interact with other proteins to adopt its active conformation, which could contribute to the crystallisation difficulties [36]. The main role of the NTD is to recruit other proteins that influence the transcriptional activity of the receptor [35].

In the NTD there are two overlapping areas responsible for the transactivating function, activation function-1 (AF-1) (bp 100-370) and activation function-5 (AF-5) (bp 360-528) [37-39]. These areas contain various peptide elements such as microsatellite repeats, surfaces for protein-protein interaction, phosphorylation and sumoylation regulatory sites. The AF-1 has the strongest transactivation potential when the receptor is activated by ligand, whereas in a truncated AR, lacking the LBD, the AF-5 is responsible for the activity, which then becomes constitutive [39].

There is strong interaction between the NTD and the LBD (NTD/LBD interaction) in the activated AR [40], and it has been shown that these interactions are necessary for complete AR activity [41]. The strongest motif in the NTD is the ^{23}FQNLF^{27} motif which binds to the AF-2 in the LBD. This motif is conserved among the AR in different species [42-46]. The precise function of the NTD/LBD interaction in the AR is unknown but it is believed that it facilitates the activation of the receptor by making protein-protein and protein-DNA interacting areas available at the surface, thereby facilitating the receptor transactivation [35].

In the NTD there are several repetitive sequences, two of them are polymorphic and will be described more closely in the following paragraphs. The repetitive sequences in the NTD contribute to its flexibility and can also be a source of differences in for example receptor activity. Two of the repetitive sequences are six glutamine residues at approximately bp 614-631 and eight prolines at around bp 1478-1501. The polymorphic CAG and GGN stretches are located further downstream. A polymorphism is defined as a genetic variant that appears in at least 1% of a population. The polymorphisms in the AR are triplet repeats, the CAG stretch encoding the aa glutamine (Q), commonly named the CAG or the polyQ repeat, and the triplet repeat of GGN, where N can be any nucleotide, encoding the aa glycine (G), designated the GGN or the polyG repeat. The other SRs do not have long polymorphic stretches of triplet repeats but have single nucleotide polymorphisms that can affect receptor activity [47-49].

The CAG polymorphism

The CAG repeat consensus sequence is $(CAG)_nCAA$ [32]. This stretch normally varies between 11-31 CAG and also varies depending on ethnicity [50; 51]. In white men the most common length is 22 (Fig. 5) whereas in African men 18 repeats is most frequent and Asians most often have a slightly longer repeat length of 23 CAG [51]. The repeat is coded from bp 596 to approximately bp 630. It is included in the AF-1 region of the NTD. Similar CAG tracts are present in other transcription factors such as the TATA box binding protein that is a component of the transcription factor 2D and important for control of the RNA polymerase II function [52]. To date only one disease is known to be directly associated with the length of the AR CAG repeat, Kennedy´s disease, also known as spinal bulbar muscular artrophy (SBMA) [50]. SBMA is a late onset disease that causes muscular degeneration and androgen insensitivity due to degradation of motor neurons [50], described in more detail on page 27.

At the beginning of this work there was a common belief that the CAG length was inversely associated to AR activity. This was mainly based on *in vitro* studies on a few ARs with varying repeat lengths [53-55] and the clinical findings in patients with Kennedy´s disease.

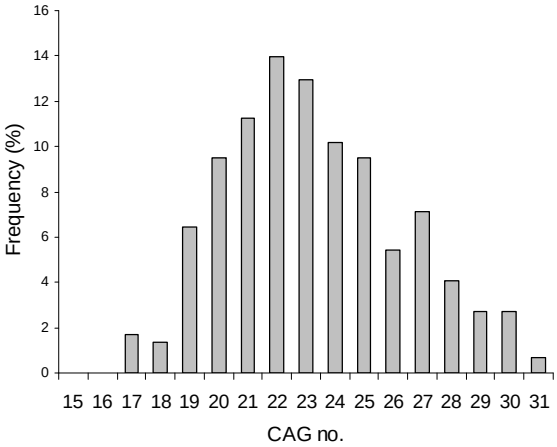

Figure 5. Distribution of the AR CAG repeat length in Swedish men. Modified from: Giwercman et al. 1998 Clinical Genetics 54:435-436.

However these previous studies did not clearly show a correlation between CAG length and receptor activity. In one of these [55], three AR genotypes were studied with 15, 20 or 31 CAG. When the activity of the receptor with 15 CAG was compared to the activity of the AR with 31 CAG a significant difference was found, but no difference was shown when the activity of any of the lengths was compared to the AR with 20 CAG [55]. In a study by Chamberlain [54], the activity of ARs with CAG lengths of 25, 35, 49 and 77 were compared. There was a significant difference in transactivation between the ARs with 25 CAG and the two longest repeats but no difference between those in the normal or upper normal range (25-35 CAG) [54]. In the third study, receptors with 15, 24 and 31 CAG were compared in various cell lines [53]. In monkey kidney cells (COS-1) and the androgen independent prostate cancer cell line PC-3, the AR with 31 repeats had lower activity than the receptor with 15 CAG. No difference in activity was shown compared to the 24 CAG genotype. In the breast cancer cell line MCF-7 there were no differences in activity between any of the AR genotypes [53].

The previously mentioned theory was widely accepted although Buchanan 2004 [56] showed that the NTD/LBD interaction in the AR was intact in receptors with CAG repeat lengths of 16-29, which also gave high transactivation. Repeat lengths of 9 or 35 residues gave a significantly lower transactivating capacity. This meant that the length of the glutamine tract was important for correct AR action but that both long and short repeats caused lower transactivation. In the same study these *in vitro* results were corroborated by the fact that 91-99% of all CAG lengths in African American, Asian, Mexican American and white men were shown to be 16-29 repeats [51; 56]. These results suggested a different view of the role of the CAG repeat in AR activity. The repeat acted as a modifier of AR activity, by keeping the repeat within certain lengths, optimal and specific activity could be retained. These specific lengths could have most advantageous and strict interaction with co-regulators, and be optimal for particular activity in various cell types [56]. These studies were the basis for the studies included in this thesis.

To date the AR CAG repeat length and its association to reproductive function has been analyzed with a linear approach. However, studies on patients and reproductive function in relation to CAG length, using a linear approach, have given inconclusive results [55; 57-67].

The GGN polymorphism

The GGN repeat is located downstream from the CAG repeat at approximately bp 1709-1780 and encodes a stretch of the aa glycine. The letter N in GGN represents any nucleotide of cytosine, thymine or guanine and the consensus sequence of the repeat is $(GGT)_3GGG(GGT)_2(GGC)_n$ [68].

Less is known about the GGN repeat than the CAG repeat. A total deletion of the GGN repeat decreases the AR activity by approximately 30%, suggesting that it is important for correct receptor function [69].

There are two dominant GGN alleles in the white population, 23 GGN is the most common, closely followed by 24 GGN [70]. The 23 GGN allele in combination with the median CAG length (22 CAG) has been shown to give higher AR activity compared to longer and shorter GGN alleles (27.24 and 10 GGN) *in vitro* [71]. The effect of the GGN repeat on human reproductive parameters has also been studied *in vivo* where it was shown that <23GGN repeats was associated with lower semen volume compared to subjects with ≥23GGN. Moreover, men with <23 GGN who were exposed to maternal smoking during pregnancy, had a higher BMI in comparison to men with any other GGN length, irrespective of whether their mother smoked or not during pregnancy [72].

Similar to the CAG repeat the GGN repeat length also varies between populations. African populations have the shortest repeat lengths and also the largest length variations whereas white and Asian men have longer repeat lengths, with a narrower span [73].

This in combination with CAG length data has been proposed to be a factor that partly explains why prostate cancer is more common in African-American men compared to white and Asian men. African-Americans have the highest incidence of prostate cancer in the world, whereas Asian men having the longest CAG and GGN repeats, have the lowest incidence. The incidence in white men is in between these two populations [74]. A few years ago the GGN polymorphism was suggested to be involved in hypospadias and chryptorchidism [75]. This has later been confirmed in other studies [76; 77].

The DNA binding domain

The DNA binding domain shows high evolutionary conservation and the human sequence is identical to that of the rat [78]. It is approximately 80 aa long and stretches over exon 2 and 3. The domain consists of two zinc fingers formed from three α-helices that interact with DNA and a C-terminal extension [79]. Each zinc finger contains one zinc atom that is coordinated by four cysteine amino acids. The first zinc finger includes a sequence element termed the proximal-box (P-box) which is identical in the AR, GR, MR and PR. The P-box is a five aa long sequence consisting of glycine, serine, cysteine, lysine and valine that co-ordinates specific interactions between the protein and the major groove of the DNA where the HREs are located. The residues involved in DNA recognition are the glycine, serine and the valine [40; 80; 81]. The second zinc finger contains the distal-box (D-box), a site for DBD/DBD homo-dimerisation of ARs [81; 82].

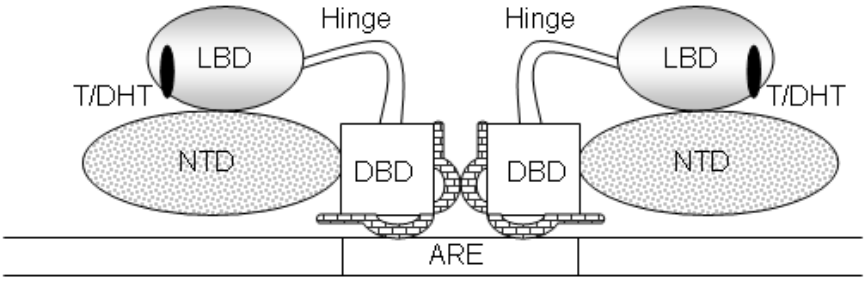

Figure 6. Schematic drawing of the AR as a homodimer bound to an androgen response element in the DNA. The D-box in the DBD is in contact with the DNA, the P-box connects the two receptors. ARE: androgen responsive element.

The dimerisation of the receptors fixes the aa in the P-box, making interaction with DNA possible (Fig. 6.). The HREs recognized by the AR, PR, GR and MR are inverted repeats of 5´-AGAACA-3´ or similar motifs with a three nucleotide spacer in between, resulting in a total length of 15 nucleotides [83]. There are two types of androgen response elements in the DNA, one type that is recognized by the AR, PR, GR and MR, called classical AREs and another type which more often has direct repeats of 5´-AGAACA-3´ or similar motifs that are not recognized by other receptors. These motifs are termed selective AREs (sARE) [81; 84].

In transiently transfected cells sAREs were recognized by AR and PR-B and luciferase transcription was induced, whereas GR and MR could not induce transcription. Transactivation by PR-A was only just detectable in the same experiment. To study the effect of chromatin on recognition, the activity of these receptors also was compared in a cell line stably transfected with a classical ARE or an sARE. All receptors showed high transcriptional activity on the classical ARE. On the sARE, the AR induced high transcription whereas the MR and GR showed no transactivation capability, PR-B showed lower transactivation with the sARE, and PR-A had very low or no transactivation on either element. The binding of the receptors to the different DNA sequences was investigated by electric mobility shift assay, showing that only the AR interacted with the sARE. It also revealed that the AR had higher affinity for the classical ARE than GR, PR or the MR had [85]. The most conserved sAREs are direct repeats of hexamers that have a G at nucleotide position -6, A at -4 and C at -3 in the motif with a half site spaced by 3 nucleotides, and a head-to-head or head-to-tail orientation of the repeats [84; 85].

The AR DBD needs a carboxyterminal extension of at least four aa for binding of classical AREs (^{625}TLGA^{628}) and 12 residues for binding to sAREs (^{625}TLGARKLKKLGN^{636}). These are located in the hinge region of the protein

[79; 81; 86; 87]. The ARE selectivity may also be promoted by dimerisation of the receptors that can affect binding at promoter sites. The AR, GR, PR and ER all prefer to dimerise in a head to head fashion [88-91].

The hinge region

The DBD and the LBD are separated by the hinge region located in the 3´ end of exon 3, and 5´ part of exon 4. It is a small region, from approximately aa 623 to 671, that is poorly conserved among the steroid receptors [40]. The nuclear localisation signal (NLS) is found in the hinge region. It binds to the nuclear import factor importin-α that mediates the transport of the AR through the nuclear pore complex into the cell nucleus. Mutations in the NLS that reduce the binding affinity to importin-α have been associated with the androgen insensitivity syndrome (AIS) where AR activity is impaired or totally lacking [92]. Cutress et al. [92] proposed that the role of the hinge region is more complex than just harbouring the NLS, as mutations in the region also affect transactivating capacity *in vitro.* Stronger evidence for this was found *in vivo* when the AR of a patient with partial androgen insensitivity syndrome (PAIS) who did not respond to androgen treatment was analysed. Sequencing of the AR revealed a mutation in the hinge region that did not affect nuclear transport particularly, but decreased receptor activity by interrupting the NTD and LBD interactions [93]. A serine residue (Ser^{650}) in the hinge region has been shown to be mutated to a glycine in an infertility patient with hypogonadism and scrotal hypoplasia [94]. This particular serine residue is normally phosphorylated and can be involved in regulation of receptor translocalisation [95; 96].

The ligand binding domain

The LBD constitutes the 3´ portion of exon 4 and exons 5 to 8. It is approximately 50% identical to the LBDs of the GR, MR and PR [97]. The crystal structure of the LBD with bound ligand has been solved, and it is a 12 α-helix sandwich with a central ligand binding cavity. The structure of the LBD is more conserved among SRs than the DNA sequence [98; 99]. There are 18 residues in direct contact with the ligand but the whole LBD is important for proper function. Upon ligand binding to the LBD a conformational change takes place. The α-helix number 12 in the LBD repositions so that the ligand binding cavity is closed and a hydrophobic cleft is exposed on the LBD surface. This hydrophobic cleft is the activation function 2 (AF-2) [100].

The AF-2 motif in the LBD mainly binds the NTD at a specific aa sequence coding for phenylalanine (F), glutamine (Q), asparagine (N), leucine (L) and phenylalanine (F), called in brief the ^{23}FQNLF^{27} motif where 23 is the number of

the first aa, and 27 the number of the last. In other NRs the AF-2 mainly interacts with another aa motif found on other interacting proteins the LxxLL motif coding for leucine (L), two residues of any other aa (xx) ending with two more leucines (LL) [39]. The AR AF-2 is surrounded by a charged clamp containing lysine and glutamate that can interact with FxxLF motifs but not with LxxLL motifs [101; 102].

Regulation of AR expression and activity

The regulation of AR expression and activity is possible at all levels, from control of gene transcription, mRNA half-life to protein turnover and activity regulation.

The human AR gene contains a large 5´-untranslated region (UTR) and at least two transcription initiation sites over 1000 nt upstream from the ATG translation initiation codon [103-105]. The promoter region of the AR gene lacks TATA and CCAAT boxes, motifs generally recognised by the transcriptional machinery. Instead the AR gene has a GC-rich region where the transcription factor specificity protein 1 binds that could be important for transcription initiation [104; 106; 107].

The expression of AR is regulated by androgens in both negative and positive directions. In the prostate cancer cell line LNCaP it has been shown that testosterone, DHT and the synthetic testosterone R1881 down regulate AR mRNA levels, as the transcription of AR was suppressed at the same time as the half-life of AR mRNA was increased [105; 108-110].

Up-regulation of AR protein has been shown in the transiently transfected monkey kidney COS-1 cell line after androgen treatment. The AR half-life increased in relation to the androgen concentration. Untreated AR had a half-life of approximately 1.5 h. After treatment AR half-life increased to 11-14 h depending on the androgen concentration used [111; 112]. The importance of androgen binding for AR protein stabilisation was shown when an AR with a mutation in the LBD (V889M) had lower binding affinity than the wild-type receptor, and also a higher degradation rate [112]. Androgens have also been shown to decrease the AR mRNA level in rat ventral prostate [109; 113]. These varying results point towards cell type specific mechanisms in regulation of AR expression.

Posttranscriptional regulation of AR mRNA is partly due to the heterogeneous nuclear ribonucleoprotein K (HnRNP-K) which binds to the 5´ untranslated region of the AR mRNA. This binding inhibits RNA translation and thereby reduces AR protein amount [114]. Regulation also takes place at the posttranslational level. Interactions with heat shock proteins (HSPs) are crucial for stabilization of the protein [115].

Directly after AR production the protein is phosphorylated and the majority of the phosphorylation sites (serines) are found in the NTD [35]. Most phosphorylations are carried out by mitogen-activated protein kinase, Akt and protein kinase C signalling [116]. Phosphorylations have been shown to affect activity, recruitment to DNA, enhancers and hormone binding [117]. The phosporylation of Ser^{308} in the NTD inhibits AR activity [118] as well as Akt mediated phosphorylation of Ser^{213} in the NTD and Ser^{791} in the LBD [119-121]. The NTD phosphorylation sites may also be important for phosphorylation of other parts of the receptor because phosphorylations can mediate conformational changes and make other phosphorylation sites, such as Ser^{650} more or less available for phosphorylation or dephosphorylation [117]. Some genes regulated by the AR require the binding of various factors simultaneously at several sites for full transcription. An example of this is the prostate specific antigen gene, where full activation requires the binding of AR, co-activators and RNA polymerase II at the enhancer and promoter region simultaneously. In contrast, repression of the gene can be carried out only when factors are bound to the promoter and not the enhancer. This phenomenon is called chromatin looping. The DNA forms a loop that brings two distant sites into close contact, which is essential for gene activation [122]. The activation and action of the AR is described graphically in Fig. 7.

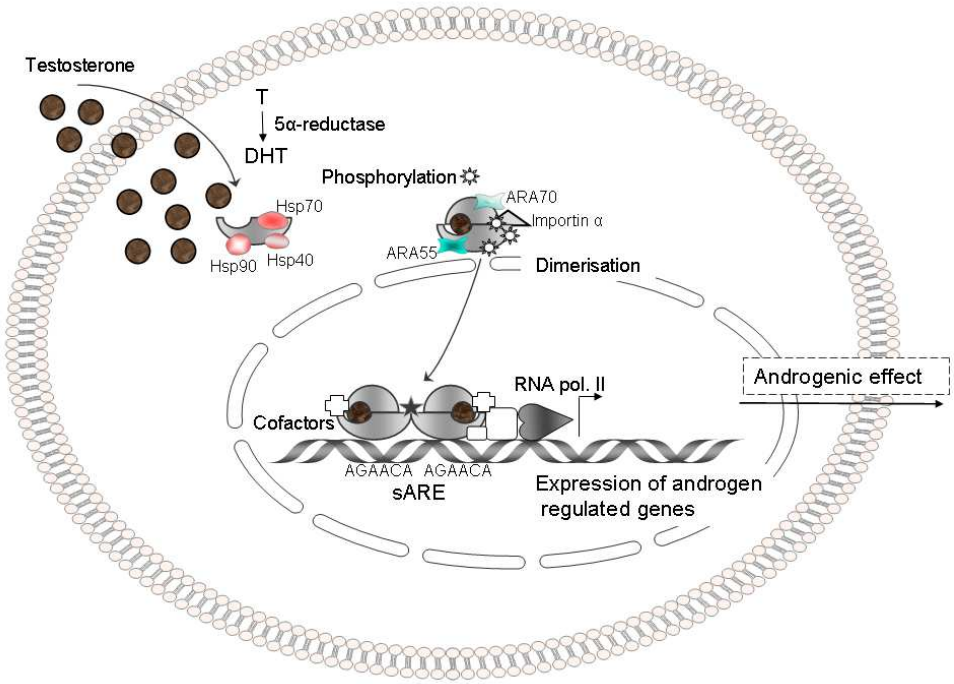

Figure 7. The mechanism of androgen action in the cell. The unactivated AR is bound by heat shock proteins, Hsp40, Hsp70 and Hsp90. When androgen dissociates into the cell testosterone or DHT bind to the ligand binding domain. Conformational changes take place, co-factors bind (ARA55 and ARA70) and the receptor is phosporylated. Importin α binds to the NLS in the hinge region and mediates transport into the nucleus. The AR dimer binds to the specific ARE (sARE). Transcriptional co-regulators and the transcription complex with RNA polymerase II bind in to the DNA. The androgen regulated gene is transcribed and the androgenic effect takes place. ARA55: androgen receptor associated protein 55. ARA70: androgen receptor associated protein 70.

Co-regulators

Androgens are essential for AR activity but small interacting proteins, termed co-regulators, can also influence in both an activating and repressing manner. Most co-regulators are unspecific and act upon many receptors and genes regulating various effects. They can act by binding to the AR protein directly or by binding to the DNA, in complex with other proteins, for example in the transcription initiation complex, or separately.

Among the first AR co-regulators to be discovered was the androgen receptor associated protein 70 (ARA70) that at the time was claimed to be an AR specific co-activator [123]. The primary interaction site for ARA70 is the LBD and it enables the antiandrogens hydroxyflutamide and bicalutamide to become AR ligands and increase transcriptional activity [124]. Increased expression of ARA70 in prostate cancer cell lines has been shown to activate the AR when stimulated with very low concentrations of androgens or oestradiol [124].

Later, ARA70 was shown to interact with other NRs and the increased AR activity seen with ARA70 was not different from the increase caused by other co-activators [125]. Also, mutations in the LBD that disrupted interaction with ARA70 only partly decreased the AR activity [126].

There are also co-activators that interact with the CAG repeat. One example is ARA24/Ras related nuclear protein (Ran). It is a small GTPase protein first identified as a complex with the protein regulator of chromosome condensation 1 [127]. ARA24/Ran is involved in many processes such as nuclear transport of protein and RNA, cell cycle progression, RNA and DNA synthesis [128]. ARA24/Ran increases AR sensitivity to DHT *in vitro* and interacts with the CAG region in a CAG length dependent manner. The association between ARA24/Ran and the NTD as well as the co-activation function of ARA24/RAN is decreased when the CAG repeat length increases. This was demonstrated by using repeat lengths with 25 and 49 CAG. The AR activity was enhanced by ARA24/RAN at both physiological and very low androgen levels, which could imply a role in castrate-resistant prostate cancer. A weaker interaction between ARA24/RAN and AR with longer glutamine lengths could be a cause of the aggregation of AR in the cytoplasm seen in patients with Kennedy´s disease [129]. The best described co-activators for the AR and other transcription factors are the three members of the p160 family, steroid receptor co-activator 1 (SRC-1), SRC-2/Transcription initiation factor 2 (TIF-2), and SRC-3/receptor associated co-activator 3/amplified in breast cancer 1. All family members have multiple LxxLL motifs that interact with the AF-2 region in the LBD of other NRs but in the AR the AF-5 region in the NTD is a stronger interaction surface [130]. The p160 co-activators also affect

transcription factors that regulate expression of other genes, for example they influence cell proliferation and cell survival [131].

Whereas SRC-1 and SRC-2 appear to be needed for optimal transcription activity SRC-3/RAC3/AIB1 has been shown to be over-expressed in various cancers, both hormone dependent and non dependent [132-134].

The AR activity is also regulated by co-repressors. An example of an AR co-repressor is glycogen synthase kinase 3 β (GSK3β). It was highly expressed in prostate cells and has been shown to interact with the AR both *in vitro* and *in vivo*. It acts by phosphorylating the NTD and suppresses the transactivational activity by reducing the interaction between the NTD and the LBD, thereby affecting dimerisation [135]. Another co-repressor is HDAC1, a histone deacetylase that binds directly to the AR and down-regulates transcription without affecting AR protein levels. HDAC1 binds to the AF-2 motif and most probably deacetylates the receptor causing inactivation [136].

Human disorders related to the AR

Androgen insensitivity syndrome

Mutations in the AR gene result in various grades of androgen insensitivity syndrome (AIS). There are three main types of androgen insensitivity syndrome (AIS). In complete androgen insensitivity syndrome (CAIS) the AR is inactive. This creates a female phenotype at birth although the child´s karyotype is XY. Partial androgen insensitivity syndrome (PAIS) is a heterogenous form of AIS and presents as various degrees of female virilization or male feminisation due to diverging degrees of AR activity [137]. The mildest form of AIS is mild androgen insensitivity syndrome (MAIS) which gives a phenotypic and genetic male but with AR activity defects that can be manifested as oligospermia, gynecomastia or minor hypospadias [138; 139]. There are at least 300 different mutations causing AIS in exons 2-8 but only 23 in exon 1 of the AR. Approximately 70% of these mutations are inherited and 30% are de novo mutations, not found in the maternal AR gene (http://androgendb.mcgill.ca/).

Kennedy's disease

Kennedy's disease, also known as SBMA was described in 1968 [140]. It is a late-onset progressive motor-neuron disease affecting muscles in the hips, shoulders and later also bulbar muscles resulting in difficulty with walking, swallowing and talking. These patients also may have gynecomastia and reduced fertility, showing

phenotypic traits similar to PAIS. The CAG repeat length is inversely correlated to age at onset and disease severity [50]. Longer repeats give an earlier onset and a more severe disease [141]. The elongated CAG tracts cause aggregation of AR and other proteins in neural and non-neural tissues, eventually leading to cell and tissue death. The neuro-pathological symptoms are not due to the loss of AR function but to a gain of neurotoxic function of the AR [142]. It has been proposed that the elongated CAG tract leads to misfolded proteins that fail to be degraded due to impairments in the proteolytic pathway and therefore form intermolecular complexes that accumulate in the cells [143]. Mouse models have shown that the pathology of CAG repeats is androgen dependent. In clinical trials reduced androgen production induced by leuprorelin decreased inclusion formations in scrotal skin, but muscular function was not significantly improved [144]. The fact that patients with Kennedy's disease have low AR activity has motivated studies on CAG repeat length and AR activity. Also, it has led to the common dogma that longer CAG tracts result in lower AR activity.

Other repeat expansion diseases

There are two categories of repeat expansion diseases, those such as SBMA where the repeat is located in an exon and those where the repeat is located in an intron. Both types of repeats can have significant impact, for example by causing chromosome fragility, silencing the genes in which they are located, modulating transcription and translation, and by inhibiting proteins involved in processes such as splicing and cell architecture [145]. Other diseases caused by an extended glutamine repeat are Huntington disease, and the spino-cerebellar ataxias that are due to an expanded glutamine repeat in the huntingtin, and the ataxin proteins, respectively.

Triplet repeats have been found in 15 neurological disorders of which eight result from an expanded CAG repeat and are neurodegenerative. All these diseases are progressive, and have an earlier onset, and are more severe the longer the CAG stretch. Another similarity is that only specific neurons in each disease are affected, although the mutated protein is found throughout the body [146].

Prostate Cancer

Circulating androgens and the AR are essential for prostate development and also to some extent for prostate cancer (PCa) development. In the prostate, androgen effects are mediated by DHT which is metabolised from testosterone by SRD5A2.

Recently it was discovered that the AR has varying functions in specific prostate cells and also that its effects differ in normal and malignant tissue. In prostate

stromal cells it enhances proliferation, in epithelial luminal cells it is a survival factor that promotes cancer progression, and in basal cells it inhibits metastasis [147]. These results are contradictory to the accepted idea that AR activity always promotes cancer progression. However, high expression of the AR in prostate does not automatically cause PCa, since mice with over-expression of wild-type AR developed prostate intraepithelial neoplasia (PIN) [148] whereas over-expression of an AR with a missense mutation caused PCa [149]. The AR in PCa has been found to have a range of mutations but they are very seldom found in untreated PCa. Moreover, although these mutations are not always found in the primary tumour, they are present in the metastases [150; 151]. In the Finnish population the substitution of leucine for arginine at aa position 726 (Arg726Leu) was found to increase the risk of PCa [152] but these findings were not confirmed in a subsequent study of a North American population [153].

Since the first suggestion that the AR CAG repeat length could be associated with PCa risk in 1994 [154], there have been numerous publications on this topic. Today this matter is still debated and some studies show associations with PCa risk [155-161] whereas others do not [162-168]. PCa aggressiveness and age at onset has also been studied with inconclusive results. Stage and grade of the disease have been associated with short CAG repeats in a number of studies [156; 165; 169-171] and younger age at onset has been related to shorter CAG repeats in some reports [156; 160; 168; 172; 173]. Other studies have not found any association between CAG repeat length and age at onset or disease aggressiveness [174; 175]. The largest study to date on PCa risk and CAG repeat length has been carried out by Lindström et al. and includes more than 6000 cases and controls [163]. No association between CAG length and PCa was reported but a relation between CAG length and circulating levels of testosterone and estradiol was observed, as has been found earlier in another large elderly cohort [176]. High androgen levels have been thought to increase PCa risk [177-182] but in a large meta-analysis no association was found between circulating levels of testosterone, DHT, or E and PCa [183]. This was corroborated by Lindström et al. who did not find higher PCa incidence in men with higher testosterone levels [163].

Infertility

Infertility affects 15% of couples worldwide. In 20% of these cases infertility is caused by a male factor and in 60-75% of male infertility cases the aetiology is unknown [184]. Seven mutations in the AR primarily causing infertility have been reported, five in exon 1, two in exon 8 and one in exon 5 [185]. Men with these mutations develop normally, have a male phenotype but a defect in spermatogenesis. In mouse models where the AR was selectively not expressed in Sertoli cells, a phenotypic normal male was developed but spermatogenesis was

disregulated [186; 187]. Male infertility may therefore be caused by specific functional limitations of AR action in Sertoli cells.

Since the discovery of an association between an expanded CAG repeat and Kennedy's disease, there has been a general belief that longer CAG stretches give a less active AR and therefore could cause infertility. In a large meta-analysis by Davis-Dao on mainly white men [188], infertile men were shown to have 0.19 repeat longer CAG stretches than fertile men. In a similar study on Asian men, the mean difference in CAG length was 1.34 repeats [67]. The data in these studies were analysed in a linear model based on the early *in vitro* studies mentioned and on the theory of an inverse relation between CAG length and AR activity. Although both studies found a slightly longer CAG stretch in infertile men the biological effect of less than one or 1.5 repeats could be questioned.

Aims

Understanding the impact of CAG number on AR activity is important for proper interpretation of receptor function and the role it has in pathological, androgen related diseases, other than Kennedy's disease.
The purpose of this thesis was to examine the influence of CAG length, if any, on AR activity and expression by studying receptors with CAG lengths in the normal range.

The specific aims were to:

- Determine the effect of CAG lengths within normal range on AR activity *in vitro*

- Re-analyse data on AR CAG repeat length in relation to risk of male infertility and prostate cancer

- Measure AR protein amount and activity *in situ* in human prostate tissue from men with known AR CAG repeat length.

Materials and methods

Studies included

This thesis is based on four studies. The first was an *in vitro* study where mammalian cells were transfected with different AR genotypes with varying numbers of CAG length and constant GGN length. The second and third studies were meta-analyses where the association between CAG length in relation to infertility and PCa was examined. The majority of the data for these studies was obtained from published scientific articles. In both studies the CAG length was stratified into three groups. In the fourth study, prostate tissue samples from men with known CAG length were analysed for AR and PSA content.

Subjects

The material for the meta-analyses was obtained through PubMed (www.ncbi.nih.gov). In the cases where no free full-text version of the article of interest was available, the author was contacted directly, and a request made for a copy of the data. Only publications on white men were included to avoid genetic heterogeneity.

In the meta-analysis on the association between infertility risk and CAG length a total of 3915 men were included, of whom 1831 were fertile and 2084 infertile. The data was extracted from 15 articles, and also data on 172 men from the outpatient clinic at the Reproductive Medicine Centre (RMC) Malmö, Sweden was included. All data was divided into three groups depending on CAG length. CAG 22 and 23 was used as the reference and the other groups were CAG<22 and CAG>23.

The other meta-analysis included PCa cases and controls. In total 7483 white men were included, 4067 controls and 3416 with diagnosed PCa. The data was originally published in 12 separate publications. A similar procedure was chosen for analysis of PCa risk in relation to CAG repeat length. The same CAG length groups were used in this analysis as in the infertility analysis, CAG 22 and 23 was the reference group which was compared to CAG<22 and CAG>23.

In the study on patient material, AR expression and activity in prostate tissue needle biopsies from 19 men with normal CAG lengths (14-28 CAG) were included. The biopsies were taken for PCa diagnosis purposes before any eventual treatment was started. The samples were collected during the period 1997-2002

and had been paraffin embedded. Two of these samples only contained benign tissue and two other samples only had malignant tissue.

Molecular methods

Gene amplification and plasmid construction

The polymerase chain reaction (PCR) is a highly sensitive detection method that is used on a daily basis in the field of molecular biology. PCR amplifies a specific target DNA sequence by copying the DNA. The specificity of the PCR is defined by the choice of primers. Primers are small oligonucleotides that are designed to be complementary to the sequences flanking the genomic area of interest. The PCR reaction has three general steps. First, the DNA is heated up to approximately 96 °C and separates into its two complementary strands (denaturation). At a lower temperature, the specific oligo-primers bind to their complementary single DNA strands (annealing). The thermostable DNA polymerase enzyme included in the reaction requires this oligo-primer binding prior to initiating copying of the single-stranded DNA adjacent to the primer-binding site in a 5´-3´direction (extension). Two new double-stranded templates of the region of interest are produced through incorporation of deoxynucleotides provided in the reaction mix, at the optimum processing temperature for the polymerase enzyme, 68-72 °C. These three steps are repeated, and the target DNA is amplified in an exponential fashion [189]. There are several variants of PCR, real-time PCR also termed quantitative PCR (qPCR), used for quantification of a specific sequence or gene product, and reverse transcriptase PCR (RT-PCR) which uses RNA as starting material transcribing it to cDNA prior to quantification by qPCR or amplification by PCR.

In this thesis PCR was used for amplification of the human AR 23 GGN region, that was selected to generate an expression plasmid with the combination 16 or 28 CAG and 23 GGN. A plasmid expressing an androgen receptor with 22 CAG and 23 GGN repeats was already available in the research laboratory. These genotypes were chosen to compare the median CAG length in combination with the most common GGN length in white men (22 CAG 23 GGN) to longer and shorter CAG repeats within the normal range.

The DNA template for the PCR was human DNA, and the reaction was run under the following conditions: 1 min denaturation at 96 ºC, followed by 1 min annealing at 56 ºC and 3 min extension at 72 ºC. The sequence of the forward primer was 5´-CCAGAGTCGCGACTACTACAACTTTCC-3´ and the reverse primer sequence was 5´-CCAGAACACAGAGTGACTCTGCC-3´.

The PCR products were digested with Kpn1 and BstE11/Eco911, resulting in a 241 bp product, only containing the required GGN stretch.

Two expression vectors (pCMV4) containing full length AR cDNA with the required CAG repeat length in combination with an unwanted GGN repeat were digested with the same enzymes as the PCR product to cut out the GGN repeat. The opened plasmids, without GGN repeat, were then purified.

The digested and purified PCR product (23 GGN) was ligated into the opened pCMV4 expression vectors, resulting in two vectors with 23 GGN in combination with 16 or 28 CAG repeats. These plasmids were transformed into *Escherichia. coli DH5α* by electroporation, amplified and purified. Finally correct incorporation of the GGN fragment and the CAG sequence was verified by direct sequencing on an ABI Prism3730 GeneticAnalyzer (Applied Biosystems).

Transfection of mammalian cells

Mammalian cells can be either transiently transfected or stably transfected. Transiently transfected cells do not integrate the foreign DNA into the chromosomal DNA, and only express the genes in the transfected DNA for a short period of time until the foreign DNA is degraded. In stable transfection the foreign DNA is integrated into the cell genome and is replicated and transferred to the daughter cells during each cell division. Stable transfection often requires that the transfected DNA contains a selective gene, such as antibiotic resistance, to keep the DNA integrated into the genome. In the current work african green monkey kidney cells (COS-1) were used for transient transfection of AR expressing plasmids. COS-1 cells were chosen as they do not express an endogenous AR, are easy to transfect, and have been used in similar studies previously [53; 55]. They express the SV40 large tumour antigen that enables transcription initiation at the SV40 origin site in the pCMV4 expression vector [190]. This results in a high copy number of the vector with the AR cDNA in each transfected cell [191].

The methods used for transfection of plasmids into mammalian cells are based on various ways of promoting DNA uptake. In methods using calcium phosphate small precipitates are formed between the DNA and calcium that then are adsorbed by the cell through a poorly understood mechanism. Electroporation, when cells are given a short electric shock, is an efficient method to transfer DNA directly into the cell nucleus. In liposome based methods a complex is formed between the DNA and lipids. The complex formed fuses with the cell membrane or is taken up into the cell by endocytosis. We used a transfection reagent containing a blend of lipids (Fugene 6, Roche Diagnostics).

The plasmids transfected into the cells were 600 ng pCMV4 plasmid containing the AR genotype of interest, 200 ng of the reporter plasmid pGL3hPSALuc2, and

one of either of the transfection efficiency controls, 1 ng of pCH110β-gal or 5 ng pRL-SV40renilla. The pCMV4 plasmid has a strong CMV promoter and as described earlier can be replicated in COS-1 cells. The reporter plasmid had a human prostate specific promoter in front of the luciferase gene. The two transfection efficiency controls were used to investigate if the results could be influenced by any common transfection efficiency plasmid.

Cells were grown and kept in Dulbecco´s modified eagle media containing 10 % heat inactivated fetal bovine serum supplemented with 2 mM glutamine and 0.02% gentamicin.

For transfection 150 000 cells were seeded 24h before transfection in 12-well plates.

After transfection, cells were left for 24 h before they were washed and new media with 10 or 100 nM of DHT or R1881, or media with no hormone was added. Methyltrienolone (R1881) is a synthetic testosterone that can not be converted into DHT by SRD5A. Since the COS-1 cells are known to metabolise testosterone, R1881 was used to study the effect of testosterone. The serum used in the media was stripped from hormones.

Luciferase and renilla activities were assessed by the Dual Luciferase Reporter assay, whereas the combination of β-gal and luciferase was assessed by the Dual Light Luciferase assay. Total protein amount in each sample was measured by the method of Bradford [192]. The experiment was repeated up to 31 times with DHT and six times with R1881.

To rule out the possibility that differences in activity depended on varying AR transcription amount, mRNA was prepared from transfected cells. The mRNA was then reverse transcribed to complementary DNA (cDNA) and quantified by quantitative real-time PCR.

Quantitative real-time PCR

In quantitative real-time PCR (qPCR) the outcome of each PCR cycle is determined in “real” time. The number of target DNA molecules can be determined by measuring at what cycle a certain threshold level is reached (the Ct value). This enables accurate determination at an unparalleled dynamic range. There are two main methods of measuring the amount of PCR product based on the use of labelled hydrolysis probes, or the incorporation of fluorescent dyes.

One probe based system uses labelled probes i.e. small single stranded DNA molecules (oligonucleotides) of the chosen sequence conjugated with a fluorescent dye at the 5´ end, the reporter, and, in close spatial proximity, a quencher at the 3´ end of the intact molecule that reduces the fluorescence emitted by the reporter

fluorophore. As long as the probe is intact, it can bind (hybridise) to the single stranded DNA sequence of interest. If the primers bind to the DNA, and a thermostable DNA polymerase with 5´ exonuclease activity is used in the reaction mix, primer extension will take place, the hybridized probe will be cleaved, and the quencher will separate from the fluorophore. Continuous excitation allows online monitoring of the fluorescence signal in "real time". In presence of target DNA the fluorescence signal produced by the reporter dye increases, and the number of target molecules can be determined by measuring at what cycle a certain threshold level is reached, the Ct value.

The second method is less precise as both specific and non-specific products give a signal. It is based on the use of a fluorescent dye that binds directly to the newly synthesized double stranded DNA. With every cycle, the amount of double stranded DNA and fluorescence increases. When using this method, no probe is needed and it is reliable as long as the product size is verified by generating a melting curve, that indicates the melting temperature of the products amplified during the reaction. By comparing the amount of the gene of interest to the amount of a known endogenous control gene, run in a parallel reaction, a relative amount of the gene of interest can be obtained. The endogenous control gene is a gene that is constantly similarly expressed in the starting material [193].

For the measurement of AR expression in COS-1 cells, qPCR using a fluorescent dye (SYBRGreen, Stratagene, CA, USA) was performed. During the process of RNA extraction the RNA was treated with DNase to ensure that no residual plasmid DNA remained which could be used as false template in the subsequent qPCR reaction. Primers that specifically bind to the polyA tail of mRNA molecules were used in the reverse transcriptase reaction to produce cDNA.

The AR specific primers in the qPCR reaction were designed to be intron/exon spanning to facilitate the discrimination of products from chromosomal DNA and cDNA. Primers used for AR detection were, forward 5´-AGCCTATTGCGAG AGAGCTG-3´ and reverse 5´-GCTTCACTGGGTGTGGAAAT-3´.

The chosen endogenous gene, glyceraldehyde 3-phosphate dehydrogenase (GAPDH), was detected with the forward primer 5´-CGACCACTTTGT CAAGCTCA-3´ and reverse primer 5´-AGGGGTCTACATGGCAACTG-3´. The AR mRNA was quantified by using the comparative C_T method. The method is based on normalization of fluorescence values for the gene of interest to the values for the endogenous control. Three samples of each genotype were analysed in triplicates and the data were compared with the expression of 22 CAG treated with 10 nM DHT, which was set as a reference value and included in each qPCR run.

Enzyme linked immunosorbent sandwich assay

Methods for estimating or determining specific protein amount most often are antibody-based techniques. Western blot is commonly used for detection of proteins in cell lysates. It is a combination of fractionating the proteins in the sample based on molecular weight by polyacrylamide gel electrophoresis, transferring all separated proteins to a solid membrane and then detecting specific proteins by using primary antibodies and secondary antibodies conjugated to a detectable reagent. The technique is valuable for detecting specific proteins, but is a crude method for determining the specific protein amount. Instead of using western blotting for determining the amount of a specific protein, enzyme-linked immunosorbent sandwich assay (ELISA) can be used. In this work a sandwich ELISA was the chosen method to rule out that differences in measured activity depended on varying AR protein amount.

The sandwich ELISA allows accurate determination of protein amount. It is an antibody-based method where specific antibodies are used to coat a microtiter plate. The sample is added, the protein of interest (antigen) is bound to the antibody and the unbound residual material in the sample can be washed away. The amount of bound protein can be determined by using a second antibody conjugated to a colorimetric, enzyme linked or fluorescent/chemiluminescent antibody, which develops a measurable color/fluorescence when substrate is added. To increase the sensitivity of the assay, an unconjugated secondary antibody can be used before a third conjugated antibody is added. Apart from the sandwich method there are several variations of ELISA. The indirect ELISA is not a sensitive method but can be used for qualitative analysis. In indirect ELISA the sample is incubated in an empty well. Antigens in the sample adhere to the surface. The surface is then blocked to inhibit binding of the next antibody to the empty surface. Primary antibody against the antigen is added and a secondary antibody linked to an enzyme is then used for detection. Another variation of ELISA is the competitive ELISA, where the sample first is incubated with specific antibodies. Then the mix is added to wells coated with antigen. The secondary antibody is linked with an enzyme and after addition of substrate a measurable colour is developed. Standard samples, with known antigen concentration, can be included in the reaction to give a standard curve. This curve can then be used to determine the exact amount of antigen in each sample. There is no purified AR protein available on the market, so for this work the exact amount of AR could not be detected, only relative amounts in different samples.

In all ELISA reactions lysates from the prostate cancer cell line LnCaP were included as positive controls. Untransfected COS-1 cells were used as negative controls. The appropriate amount of total protein to be used was decided by

running a dilution curve on one sample before the analysis of all samples. The dilution curve showed which dilution to use to obtain optical density values within the linear range.

Quantification of AR and AR activity in tissue

Antibody staining and TRFI

Serum PSA measurement is commonly used to screen for prostate disease and in the follow up of patients with PCa. A high PSA value could be a sign of aberrant prostate growth but does not directly indicate PCa. Expression of the prostate specific antigen (PSA) gene is regulated by the androgen receptor. In other words, the PSA gene is a downstream target of the AR. Therefore AR and PSA protein amount was measured directly in human prostate tissue to determine if AR expression and activity varied depending on CAG length. Common immunohistochemistry can be used for specific protein quantification, but the signal obtained is not always linear to protein amount, and the background is often high. The chosen method for direct quantification was time resolved fluorescence imaging (TRFI). First the method was verified on AR transfected and paraffin embedded COS-1 cells treated with 10 nM DHT. The cells were transfected as mentioned earlier with ARs harbouring 16, 22 or 28 CAG in combination with 23 GGN. After harvesting and fixation the cells were paraffin embedded. Sections were then prepared for TRFI for AR protein measurement. TRFI is a method where specific antibodies directly or indirectly linked to lanthanides are used in combination with image acquisition in an epifluorescence microscope. The lanthanides can be excited at specific wavelengths and have a long decay time. The long decay time makes it possible to obtain images with low background noise and autofluorescence. In comparison to immunohistochemistry, TRFI gives a linear relation between signal intensity and the specific protein expression; it also allows an automated and improved quantification and evaluation of cellular parameters.

The primary antibodies were directed against AR and PSA. They were conjugated to Europium (Eu) and Terbium (Tb) respectively making it possible to measure both AR and PSA amount from the same sample. One sample from each patient was analysed once. The primary areas of interest were prostate glands, where AR expressed in the nucleus of epithelial cells and PSA found in the stromal cells could be detected. The Eu signal was obtained with an emission filter set at 615 nm and for Terbium a filter at 545 nm was used. This gave separate images with

specific Eu and Tb signals from exactly the same area. These images were then analysed, resulting in specific density values for each region of interest in every image. The mean intensity of the background was subtracted from each region of interest making it possible to compare the images to each other.

Statistical analysis

The choice of statistical method was based on the distribution pattern and number of data points. The non-parametric tests were chosen when data was not normally distributed or when the number of observations was small. In the analysis of the transactivation assay data, the groupwise comparisons were performed with a test where pairwise comparisons were possible. All statistical tests were two-tailed with statistical significance defined as $p<0,05$.

In each experiment the mean value of the AR with 22 CAG stimulated with 10 nM DHT or R1881 was set to 100% or 1. The values of the other genotypes were expressed relative to it.

To compare the activities in the transactivation assay, the results of all three genotypes were compared to each other with the non-parametric Friedman´s test. If the test showed a significant difference ($p<0,05$) between genotypes they were pairwise compared with the non-parametric Wilcoxon signed ranks test. The ELISA and qPCR results were compared using the same statistical tests.

In the meta-analysis the three CAG repeat length groups within each study were of similar size. Both data sets were analysed by binary logistic regression with fertility or PCa as the dependent variable. The risk of infertility or PCa was also assessed with the independent sample t-test, where the mean CAG length of all included men in the respective study was compared.

In the study on AR and PSA protein amount in prostate tissue the samples were grouped into similar groups as in the meta-analysis. The four samples with 22 CAG represented the median repeat length in white men and were one group. Samples with CAG lengths of 14 to 18 were in the second group (n=9) representing shorter normal repeat lengths, and patients with 26 to 28 CAG were in the third group representing longer normal CAG lengths (n=6). Non-parametric tests were used for analysis. The Kruskal Wallis test was used for the over all comparison of the three groups. Groupwise comparisons were made with the Mann Whitney test. The same statistical methods were chosen for comparison of the AR amount obtained by TRF in COS-1 cells. The mean density value of cells treated with empty vector was considered background and subtracted from every data point obtained in AR transfected cells.

Results and discussion

At the start of this work, a negative linear association between AR function and the CAG repeat number was generally assumed. This was based on clinical findings in men with the neuromuscular disorder Kennedy's disease who present with an abnormal expansion of the CAG stretch to more than 40 repeats, and on a few *in vitro* studies from the 1990s [53-55]. However, *in vivo* data concerning the association between CAG number and androgenic effects were conflicting.

As most previous *in vitro* studies were based on extreme CAG lengths, and on reporter-systems containing viral promoters, the objective of the first study was to investigate ARs with CAG lengths within normal range (16, 22 and 28) in a reporter-assay with the human PSA promoter as the target (Paper I). We also wished to elucidate whether the interpretation of the results depended on the methods used for adjustment of transfection efficiency.

With ß-galactosidase as transfection control, 22 CAG length had the highest activity (100%) compared to 16 CAG (mean 78% [range 41- 132], p=0.005) and 28 CAG (68% [26-162], p=0.006), whereas adjusting for renilla-luciferase resulted in 16 CAG behaving similarly to 22 CAG (104% [56- 165], p=0.7) and 28 CAG having lower activity (59% [33-101], p=0.004). When using the co-transfected transfection controls the negative control samples transfected with empty AR vector displayed considerable background activity.

Adjusting for AR protein gave another activity distribution, the 22 CAG genotype showed the highest activity; 16 CAG and 28 CAG displaying 20% (10-47, $p<0.0001$) and 12% (5-21, $p<0.0001$) thereof. Similar results were obtained with adjustment for total protein. The negative controls showed no or very little activity. By normalising for AR protein-content, the highest AR activity was confined to the 22 CAG and not 16 CAG, contrary to the findings with transfected control vectors. Thus, the relation between CAG repeat length and AR activity may not be strictly linear, as had been assumed in the majority of previously published association studies. This may, at least partly, explain the discrepancy in data aiming to link physiological conditions to CAG repeat length. Re-analysing data in a stratified manner may give more important information on the relation between AR CAG repeat length and disease predisposition.

In general, the glutamine length variance is influenced by both specific sequence characteristics and the specific role of the glutamine tract within the protein structure and function [194]. This is illustrated by the fact that some proteins are polymorphic whereas others are not [195], indicating critical reasons for

maintaining a particular CAG repeat length in some proteins. With respect to the AR, there seems to be a length of approximately 10-30 repeats that is tolerated for proper AR function and that can be balanced by both AR protein amount, as has been shown previously [196; 197], and also by the secretion of sex hormones. In line with this hypothesis, it has been proposed that as 91–99% of the CAG alleles across different ethnic groups are between 16 and 29 residues, this could be a range that would maintain maximum interaction between the transactivating domain and the hormone binding domain of the AR [56]. Shorter or longer repeats than the critical range could be a more important mediator of disease phenotype than a stepwise reduction in activity with increasing CAG length across the entire range [56].

The CAG stretch of the AR protein is situated in the transactivating domain that interacts with the hormone binding domain. The fact that mutations in the CAG stretch can disrupt this interaction [56] provides evidence that the polyglutamine tract plays a crucial role in ensuring proper function of the human AR. The interplay between the transactivating domain and the hormone binding domain has previously been shown to be significantly reduced by shorter or longer CAG repeats than the normal range [56]. This finding generated the hypothesis that the polyglutamine tracts serve as flexible spacers to separate regions of biological activity while maintaining the capacity to interact with co-regulators and the transcription machinery. Such a mechanism could explain how both increased and decreased CAG lengths can influence AR function. To test this hypothesis, data from two recent meta-analyses, -one on infertile men and one on PCa cases, - was collected and re-analysed in a stratified manner (Paper II and III). In both studies <1 CAG repeat differed between cases and controls when analysed in a linear regression model, assuming that AR function diminishes with increasing CAG length. Such a small difference is unlikely to explain why PCa develops in some men. Neither does it give any improvement in methods for disease prediction. On the other hand, one single repeat difference in the GGN repeat can affect receptor activity. An AR with 23 GGN had significantly higher activity than a receptor with 24 GGN *in vitro* [71]. Also phenotypic differences were found when comparing men with 23 and 24 GGN [72].

Thus, by analysing data sets with a stratified analysis, more information on the role of the AR in these diseases could possibly be obtained.

The infertility study included 3915 men, of whom 1831 were fertile and 2084 infertile (Paper II). Data was divided into three categories: CAG<22, CAG 22-23 (reference group) and CAG>23 and then re-analysed in a binary logistic regression model. When comparing the groups regarding differences in mean length, no significant difference was found between the fertile and infertile men (equal variance, p= 0.615). However, when the CAG lengths were divided into three

groups (<22, 22-23 and >23 CAG) for the stratified analysis, men with CAG<22 and CAG>23 had 20% increased odds ratio of infertility compared to carriers of the median lengths (p=0.03, 95%CI 1.02-1.39 for CAG<22 and p=0.02, 1.03-1.44 for CAG>23). These results indicate that data from previous studies should be re-analysed regarding CAG number and androgenic endpoints.

The PCa-study (Paper III) comprised genetic data on 3416 white men with PCa and 4067 controls stratified into three groups in the same manner as in the previous infertility-study. The mean CAG length for all men was 21.8 with a range of 6-42 CAG repeats. For the controls, the mean repeat length was 21.9, and for cases 21.6 CAG. In an un-stratified comparison of CAG length in cases vs. controls, a significant difference between all cases and controls was observed (p=0.001 unequal variance). However, in the stratified analysis, no statistically significant difference in prostate cancer risk between the groups was noted (CAG<22, p=0.071, OR=1.12 CI 0.99-1.26; CAG>23, p=0.780, OR=1.02 CI=0.89-1.16). Also, when the cases with extremely long (>30) or short (<10) CAG repeats were excluded from the stratified analysis (n= 61) no significant differences were found (CAG<22, p=0.072, OR=1.12, CI 0.99-1.26; CAG>23, p=0.823, OR=1.02, CI=0.89-1.16). This result does not necessarily mean that there is no link, but the current study had not enough statistical power to show any association between PCa and CAG length although more than 7400 men were included.

It is generally believed that androgens are necessary for prostate carcinogenesis, and androgen ablation is a corner stone in the treatment of the disease. Because of the study design, we did not have access to hormonal data. However, in a recent report, including 18 prospective studies on in total more than 3800 prostate cancer cases and 6400 controls, no association between risk for subsequent disease and serum concentrations of androgens or estrogens were found [183]. Only one hormone measurement was carried out in the studies included in the Roddam et al. analysis [183]. The androgen levels at that time-point may not necessarily reflect the hormone exposure relevant to target cells in the prostate at the relevant time in life. Moreover, even if androgens would not be causative of prostate cancer, the disease could well progress driven by androgen action on a basic level once the disease is manifest. The amount of androgens needed could be so small that the differences in AR efficiency, to which the CAG repeats contribute, do not matter.

As the in vitro experiment indicated that AR protein amount was dependent on the number of CAG repeats (Paper I), we wanted to examine whether this also would be the case in vivo. Higher AR protein expression for CAG repeats associated with lower receptor activity would point towards cellular compensatory mechanisms adjusting the receptor amount to give efficient activity in the presence of ligand.

To study the AR and PSA protein amount in vivo prostate biopsies from 19 men with median length (22 CAG), short CAG length (14-18 CAG) and long CAG length (26-28 CAG) were included (Paper IV). The tissue contained both malignant and benign areas. The method used for AR and PSA measurement was time resolved fluorescent imaging (TRFI). Because PSA is synthesized and secreted by normal and malignant epithelial cells of the human prostate in response to androgens it can be used as a measure of AR activity.

For validation of the TRFI method, COS-1 cells were transfected with different AR genotypes (16, 22 or 28 CAG with 23 GGN) and the AR protein amount was measured by TRFI. The transfected 22 CAG genotype showed lower AR amount in comparison to the transfected 16 or 28 CAG genotypes, when measured by TRFI. These results were comparable with the ELISA studies on transfected cells (16vs22 $p<0,0001$, 16vs28 $p<0,0001$, 22vs28 $p<0,0001$).

The median AR amount in benign samples was for short CAG length 1003 ± 339 count per pixel (cpp) (147%), median CAG 682 ± 252 cpp (100%), and long CAG tract 835 ± 557 cpp (122%). The median PSA amount in benign samples and corresponding CAG lengths was 315 ± 109 (85%), 371 ± 200 (100%), and for long l312 ± 93 (84%). Thus, these data indicated higher AR protein expression for the CAG numbers associated with lower receptor activity. These results should, however, be taken with some caution. None of the differences were statistically significant (p interval: 0.088-1.00), perhaps due to the relatively low number of samples included and high sample-to-sample variation.

In the malignant tissue, the AR median amount was for short CAG 1038 ± 343 cpp (121%), median 859 ± 671 cpp (100%), long CAG 730 ± 733 cpp (122%). The median PSA amount in malignant samples was for short CAG length 374 ± 136 (127%), median length 295 ± 116 (100%), long CAG 325 ± 200 (110%). Thus, the results were similar to the findings for the benign tissue, however, with p interval: 0.201.-0.914.

The TRFI study shows that there may be CAG length dependent differences in AR amount and androgen activity in benign prostate. This result is similar to previous *in vitro* results [197], where lowest protein amount (Fig. 8) and highest activity (Fig. 9) was seen in the AR with 22 repeats compared to a CAG length of 16 or 28, and is also similar to the study on infertile men (Paper II).

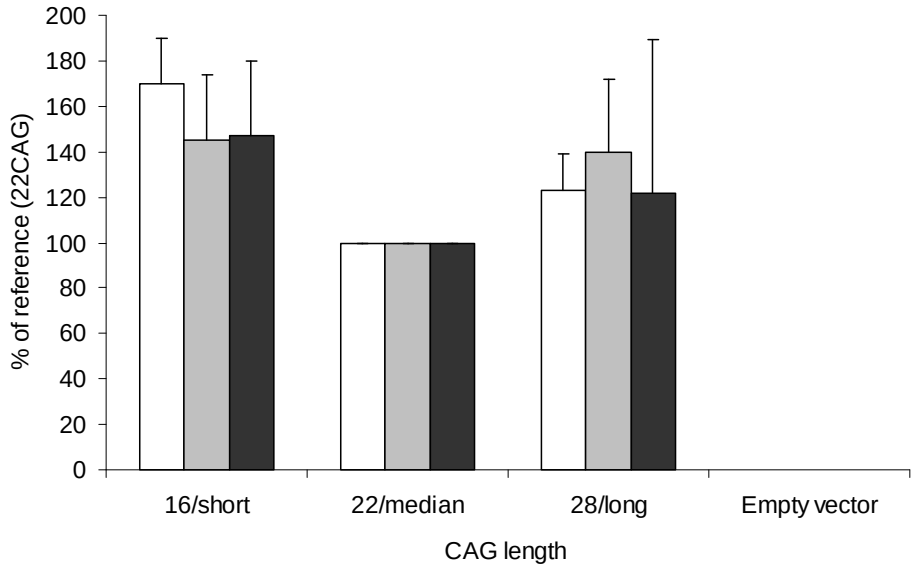

Figure 8. Androgen receptor mRNA (white bars) and protein amount (grey bars) in COS-1 cells and AR amount in benign prostate tissue (black bars).

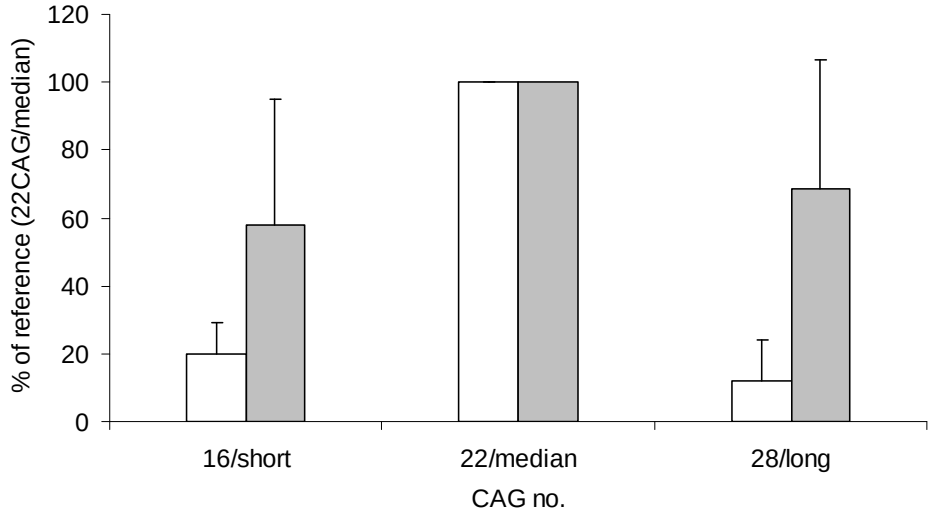

Figure 9. Androgen receptor activity in COS-1 cells (white bars) and benign prostate tissue (grey bars), adjusted for AR protein amount.

In a previous study where prostate tissues from patients were examined by immunohistochemistry, samples with CAG repeats <21 in combination with <23 GGN had a more intense AR staining and generally a higher Gleason score than samples from patients with CAG repeats >21 [199]. This result is difficult to compare with our study because of the differences in methods used. In the previous study, the genotyping was carried out by estimation of size using a ladder created with known allele sizes, not by direct sequencing, which is a precise method for genetic analyses. The TRFI technique used in the current work is considered to be a more sensitive method than immunohistochemistry. Moreover, the TRFI study results were validated with ELISA. Data was also analysed differently in our study. Rodriguez-Gonzalez et al. [199] used linear regression models, assuming a linear relationship between CAG number and intensity, in contrast to the stratified method we have been using.

In the current work, no CAG dependent differences in AR or PSA expression were detected in malignant tissue. This could be due to loss of normal androgen regulation in tumours, or to the known genetic changes that occur in tumour tissue for example amplifications, mutations and deletions. Animal studies show similar results. When human AR was expressed in healthy mice, those with 21 CAG had testosterone concentrations very similar to normal for mice, whereas mice carrying CAG 12 or 48 had higher concentrations [200], possibly compensating for lower AR function. This also resembles the observations made in infertile men, who may have a well functioning AR, although at a suboptimal level [201], but who in a considerable number of cases fail to compensate with higher androgen output. Such failure to compensate with higher androgen output may be due to testicular dysgenesis, or disruption of the feed-back mechanism, or a combination of both.

Because of the small study size as well as large intra- and inter-variability between samples, our work should be considered as a pilot study, particularly as we did not have access to more prostate biopsies from men with CAG lengths in the outer regions of the normal range. A larger study would be of importance to confirm these results.

The trend towards lower AR levels in patients with 22 CAG could however be an important finding and should be investigated further. In an unpublished study from our group (Lundberg Giwercman Y *et al.* unpublished data) it was found that men with median long CAG repeats in the AR naturally have higher PSA concentration in the circulation compared to men with longer and shorter repeats. Our results from the PSA measurements in prostate tissue and also our *in vitro* study supports these unpublished results, indicating that men with median CAG number may, on average, have higher PSA levels, due to a more efficient AR.

The data could also be studied in relation to GGN length, disease progression and hormone levels.

In the last decade the role of the AR and androgens in PCa development has been revised [181-183; 202-205]. For a long time the AR and androgen levels were considered to be causative of PCa. Because shorter CAG repeats were believed to give a more active receptor, they were also seen as a source of the malignant growth, although serum androgen levels have been difficult to correlate to initiation of disease, possibly because of lack of knowledge of the time period when hormones play an important role, and difficulties in assessing hormonal exposures retrospectively. Sperm output and the ability to father children can be used as an indicator of the long-term androgen status. In the largest study of PCa risk and number of children fathered, 48 850 PCa cases were identified through the Cancer Registry [206]. Being childless or having fathered only one child was associated with 20% reduced risk for PCa compared to having fathered ≥2 children. This finding was later confirmed in a Danish study [207]. However, neither the Swedish nor the Danish study register data allowed discrimination between different causes of male childlessness. To study the association between male fertility and PCa risk, a nested case-control study on more than 11 000 male participants was performed [208]. Childless men had a 50% lower risk of PCa compared to those who had fathered children. This finding supports the theory that normal testicular function, and hence most probably steroidogenesis is an important contributing factor in the development of PCa, although it may not be causative. This also indicates that a man with PCa most probably has been fertile with good AR function when young.

In summary, the results of this work show that the most common CAG repeat length gives optimal AR activity both *in vitro* and *in vivo*, and may protect against infertility. These findings should be used as a base for future studies within reproductive medicine and also for re-analysis of previous results.

The relation between AR activity and CAG repeat length and association with disorders in the male reproductive organs can be summarised as follows (Fig. 10).

- AR activity is highest within the most common CAG repeat range whereas AR protein amount is lower because the receptor has optimal activity. AR genotypes with other CAG repeat lengths compensate the lower activity with increased protein expression.
- A shorter or longer CAG repeat increases infertility risk that at least partly could be caused by low AR activity.
- CAG length does not affect PCa risk, indicating that the AR is not a cause of that disease.

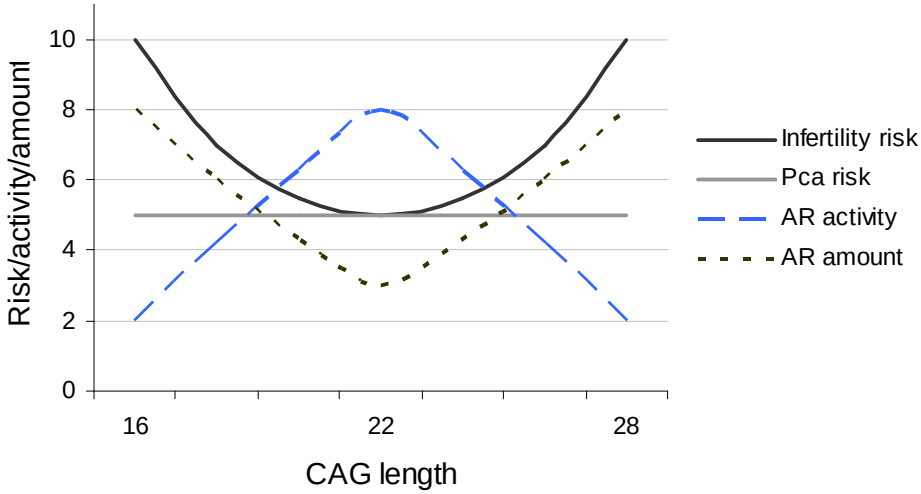

Figure 10. Schematic diagram representing the relation of the factors studied in this thesis. The Y-axis represents the risk, activity or amount, where 0 is low and 10 is high.

Future perspectives

Because of the finding of a non-linear relation between AR CAG length and receptor activity, this work can be seen as a new beginning in the study of the AR CAG repeat and its role in AR activity.

To improve understanding of AR function and the molecular mechanisms regulating it, *in vitro* studies similar to those described using human reporter promoters should be performed in various cell lines representing different tissues such as prostate, PCa and testis. Advances in TRFI methodology enable examination of larger data sets and this application could be used for protein measurement in other tissues, such as the testis. The TRFI could also be adapted to measure co-factor amount in tissue. This approach would give even more information in combination with AR protein amount and activity.

Interactions between cofactors and the AR could also be examined *in vitro*, in activity and interaction assays. In relation to the CAG repeat length, the ARA24 co-factor would be particularly interesting to study as it has been shown to interact directly with the NTD and the CAG repeat. To elucidate the effect of the CAG repeat on AR activity a mammalian two-hybrid system could be developed where the NTD and the LBD of the AR are expressed separately in combination with specific co-factors. In that way the effect of the co-factor on NTD/LBD interaction and AR activity could be studied.

The observed variations in expression of AR protein could be studied by blocking the proteosome in the cells. In that way any differences in protein degradation rate would be revealed. Also hormone binding assays could be performed to find out if differences in ligand binding could cause the varying activities. The sensitivity to low androgen levels could also be studied, as another activity pattern may be observed at very low or high androgen concentrations.

In the field of PCa the role of the AR has been constantly debated. In the future, gene webs or gene networks studies may provide a new approach to further our understanding of how various factors affect each other in the development of PCa.

Acknowledgements

This work was supported by grants from the Swedish Research Council, Swedish Cancer Society, the Research fund and Cancer Research fund of SUS, Malmö and the Gunnar Nilsson Cancer foundation.

I wish to express my gratitude to everybody who in any way has contributed to the work and completion of this thesis. A special thanks to:

Yvonne Giwercman for good supervision and advice that has helped me to become a better molecular biologist, for your contagious affection to your research, and for your aim to improve our manuscripts to perfection.

Aleksander Giwercman for inspiring scientific ideas, constructive advice, and putting the results of my work into a larger context.

Lars Rylander for your help, support, and encouragement.

All co-authors for good team-work, sharing of experiences, and expert advice.

The researchers who contributed their data for the meta-analyses in this thesis.

Elise Nilsson for your help, advice and patience regarding the IHC and TRFI work

My friends at the two closely connected research groups at CRC/RMC, you make lab. life even more fun and rewarding. Keep up the good work, and keep the fire burning!

Friends at CRC, RMC and CMP, your smiles make a difference ☺.

Anette Saltin, for your kind and efficient administrative help and for your never ending patience with my never ending questions.

My friends at the Virology department at SUS, Malmö, Anders, Ann-Sofie and Vilma.

Edward, Magnus, Erik and Tomas, thank you for introducing me to molecular biology and for giving me the opportunity to be a part of your research groups at Virology, GU, Gothenburg.

HNFF the best football club with the best team players.

My family, thank you for taking care of us, making me laugh and embellishing life.

Mum, Dad and Sofie, (Äiti, Isä ja isosisko) for encouragement, inspiration and your end-less love.

Mikael.

References

1. Robinson-Rechavi M, Escriva Garcia H, Laudet V (2003). The nuclear receptor superfamily. *J Cell Sci* 116**:** 585-6.

2. Robinson-Rechavi M, Carpentier AS, Duffraisse M, Laudet V (2001). How many nuclear hormone receptors are there in the human genome? *Trends Genet* 17**:** 554-6.

3. Gronemeyer H, Laudet V (ed.) (2002) *The Nuclear Receptor FactsBook*. Elsevier San Diego, 462pp.

4. Gronemeyer H, Gustafsson JA, Laudet V (2004). Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* 3**:** 950-64.

5. Green S, Kumar V, Krust A, Walter P, Chambon P (1986). Structural and functional domains of the estrogen receptor. *Cold Spring Harb Symp Quant Biol* 51 Pt 2**:** 751-8.

6. Santos GM, Fairall L, Schwabe JW (2010). Negative regulation by nuclear receptors: a plethora of mechanisms. *Trends Endocrinol Metab*.

7. Benoit G, Malewicz M, Perlmann T (2004). Digging deep into the pockets of orphan nuclear receptors: insights from structural studies. *Trends Cell Biol* 14**:** 369-76.

8. Thornton JW (2001). Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. *Proc Natl Acad Sci U S A* 98**:** 5671-6.

9. Eick GN, Thornton JW (2011). Evolution of steroid receptors from an estrogen-sensitive ancestral receptor. *Mol Cell Endocrinol*. 334(1-2):31-8.

10. Kuiper GG, Enmark E, Pelto-Huikko M, Nilsson S, Gustafsson JA (1996). Cloning of a novel receptor expressed in rat prostate and ovary. *Proc Natl Acad Sci U S A* 93**:** 5925-30.

11. Mosselman S, Polman J, Dijkema R (1996). ER beta: identification and characterization of a novel human estrogen receptor. *FEBS Lett* 392**:** 49-53.

12. Nieschlag Eberhard, Herman Behre, Susan Nieschlag (ed.) (2010) *Andrology Male reproductive Health and Dysfunction*. Springer: London.

13. Graham JD, Clarke CL (1997). Physiological action of progesterone in target tissues. *Endocr Rev* 18**:** 502-19.

14. Conneely OM, Mulac-Jericevic B, DeMayo F, Lydon JP, O'Malley BW (2002). Reproductive functions of progesterone receptors. *Recent Prog Horm Res* 57**:** 339-55.

15. Graham JD, Clarke CL (2002). Expression and transcriptional activity of progesterone receptor A and progesterone receptor B in mammalian cells. *Breast Cancer Res* 4**:** 187-90.

16. Li X, Lonard DM, O'Malley BW (2004). A contemporary understanding of progesterone receptor function. *Mech Ageing Dev* 125**:** 669-78.

17. Mote PA, Graham JD, Clarke CL (2007). Progesterone receptor isoforms in normal and malignant breast. *Ernst Schering Found Symp Proc***:** 77-107.

18. Htun H, Holth LT, Walker D, Davie JR, Hager GL (1999). Direct visualization of the human estrogen receptor alpha reveals a role for ligand in the nuclear distribution of the receptor. *Mol Biol Cell* 10**:** 471-86.

19. Lim CS, Baumann CT, Htun H, Xian W, Irie M, Smith CL, Hager GL (1999). Differential localization and activity of the A- and B-forms of the human progesterone receptor using green fluorescent protein chimeras. *Mol Endocrinol* 13**:** 366-75.

20. Stenoien DL, Mancini MG, Patel K, Allegretto EA, Smith CL, Mancini MA (2000). Subnuclear trafficking of estrogen receptor-alpha and steroid receptor coactivator-1. *Mol Endocrinol* 14**:** 518-34.

21. Baumann CT, Lim CS, Hager GL (1999). Intracellular localization and trafficking of steroid receptors. *Cell Biochem Biophys* 31**:** 119-27.

22. Htun H, Barsony J, Renyi I, Gould DL, Hager GL (1996). Visualization of glucocorticoid receptor translocation and intranuclear organization in living cells with a green fluorescent protein chimera. *Proc Natl Acad Sci U S A* 93**:** 4845-50.

23. Picard D, Yamamoto KR (1987). Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. *EMBO J* 6**:** 3333-40.

24. Tyagi RK, Lavrovsky Y, Ahn SC, Song CS, Chatterjee B, Roy AK (2000). Dynamics of intracellular movement and nucleocytoplasmic recycling of the ligand-activated androgen receptor in living cells. *Mol Endocrinol* 14**:** 1162-74.

25. Dittmar KD, Banach M, Galigniana MD, Pratt WB (1998). The role of DnaJ-like proteins in glucocorticoid receptor.hsp90 heterocomplex assembly by the reconstituted hsp90.p60.hsp70 foldosome complex. *J Biol Chem* 273**:** 7358-66.

26. Kosano H, Stensgard B, Charlesworth MC, McMahon N, Toft D (1998). The assembly of progesterone receptor-hsp90 complexes using purified proteins. *J Biol Chem* 273**:** 32973-9.

27. Pratt WB, Toft DO (1997). Steroid receptor interactions with heat shock protein and immunophilin chaperones. *Endocr Rev* 18**:** 306-60.

28. Matsumoto T, Shiina H, Kawano H, Sato T, Kato S (2008). Androgen receptor functions in male and female physiology. *J Steroid Biochem Mol Biol* 109**:** 236-41.

29. Russell DW, Wilson JD (1994). Steroid 5 alpha-reductase: two genes/two enzymes. *Annu Rev Biochem* 63**:** 25-61.

30. Grino PB, Griffin JE, Wilson JD (1990). Testosterone at high concentrations interacts with the human androgen receptor similarly to dihydrotestosterone. *Endocrinology* 126**:** 1165-72.

31. Chang CS, Kokontis J, Liao ST (1988). Molecular cloning of human and rat complementary DNA encoding androgen receptors. *Science* 240**:** 324-6.

32. Lubahn DB, Joseph DR, Sullivan PM, Willard HF, French FS, Wilson EM (1988). Cloning of human androgen receptor complementary DNA and localization to the X chromosome. *Science* 240**:** 327-30.

33. 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 U S A* 86**:** 9534-8.

34. Spencer JA, Watson JM, Lubahn DB, Joseph DR, French FS, Wilson EM, Graves JA (1991). The androgen receptor gene is located on a highly conserved region of the X chromosomes of marsupial and monotreme as well as eutherian mammals. *J Hered* 82**:** 134-9.

35. Shen HC, Coetzee GA (2005). The androgen receptor: unlocking the secrets of its unique transactivation domain. *Vitam Horm* 71**:** 301-19.

36. Reid J, Murray I, Watt K, Betney R, McEwan IJ (2002). The androgen receptor interacts with multiple regions of the large subunit of general transcription factor TFIIF. *J Biol Chem* 277**:** 41247-53.

37. Callewaert L, Van Tilborgh N, Claessens F (2006). Interplay between two hormone-independent activation domains in the androgen receptor. *Cancer Res* 66**:** 543-53.

38. Chamberlain NL, Whitacre DC, Miesfeld RL (1996). Delineation of two distinct type 1 activation functions in the androgen receptor amino-terminal domain. *J Biol Chem* 271**:** 26772-8.

39. Jenster G, van der Korput HA, Trapman J, Brinkmann AO (1995). Identification of two transcription activation units in the N-terminal domain of the human androgen receptor. *J Biol Chem* 270**:** 7341-6.

40. Claessens F, Denayer S, Van Tilborgh N, Kerkhofs S, Helsen C, Haelens A (2008). Diverse roles of androgen receptor (AR) domains in AR-mediated signaling. *Nucl Recept Signal* 6**:** e008.

41. Li J, Fu J, Toumazou C, Yoon HG, Wong J (2006). A role of the amino-terminal (N) and carboxyl-terminal (C) interaction in binding of androgen receptor to chromatin. *Mol Endocrinol* 20**:** 776-85.

42. Alen P, Claessens F, Verhoeven G, Rombauts W, Peeters B (1999). The androgen receptor amino-terminal domain plays a key role in p160 coactivator-stimulated gene transcription. *Mol Cell Biol* 19**:** 6085-97.

43. He B, Gampe RT, Jr., Kole AJ, Hnat AT, Stanley TB, An G, Stewart EL, Kalman RI, Minges JT, Wilson EM (2004). Structural basis for androgen receptor interdomain and coactivator interactions suggests a transition in nuclear receptor activation function dominance. *Mol Cell* 16**:** 425-38.

44. Langley E, Kemppainen JA, Wilson EM (1998). Intermolecular NH2-/carboxyl-terminal interactions in androgen receptor dimerization revealed by mutations that cause androgen insensitivity. *J Biol Chem* 273**:** 92-101.

45. Steketee K, Berrevoets CA, Dubbink HJ, Doesburg P, Hersmus R, Brinkmann AO, Trapman J (2002). Amino acids 3-13 and amino acids in and flanking the 23FxxLF27 motif modulate the interaction between the N-terminal and ligand-binding domain of the androgen receptor. *Eur J Biochem* 269**:** 5780-91.

46. He B, Wilson EM (2002). The NH(2)-terminal and carboxyl-terminal interaction in the human androgen receptor. *Mol Genet Metab* 75**:** 293-8.

47. Derijk RH, de Kloet ER (2008). Corticosteroid receptor polymorphisms: determinants of vulnerability and resilience. *Eur J Pharmacol* 583**:** 303-11.

48. Dumas I, Diorio C (2010). Polymorphisms in genes involved in the estrogen pathway and mammographic density. *BMC Cancer* 10**:** 636.

49. Thijssen JH (2009). Gene polymorphisms that may influence the biological effects of progestins. *Maturitas* 62**:** 366-70.

50. La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH (1991). Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352**:** 77-9.

51. 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**:** 241-53.

52. Nakamura K, Jeong SY, Uchihara T, Anno M, Nagashima K, Nagashima T, Ikeda S, Tsuji S, Kanazawa I (2001). SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum Mol Genet* 10**:** 1441-8.

53. Beilin J, Ball EM, Favaloro JM, Zajac JD (2000). Effect of the androgen receptor CAG repeat polymorphism on transcriptional activity: specificity in prostate and non-prostate cell lines. *J Mol Endocrinol* 25**:** 85-96.

54. Chamberlain NL, Driver ED, Miesfeld RL (1994). The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucleic Acids Res* 22**:** 3181-6.

55. Tut TG, Ghadessy FJ, Trifiro MA, Pinsky L, Yong EL (1997). Long polyglutamine tracts in the androgen receptor are associated with reduced trans-activation, impaired sperm production, and male infertility. *J Clin Endocrinol Metab* 82**:** 3777-82.

56. Buchanan G, Yang M, Cheong A, Harris JM, Irvine RA, Lambert PF, Moore NL, Raynor M, Neufing PJ, Coetzee GA, Tilley WD (2004). Structural and functional consequences of glutamine tract variation in the androgen receptor. *Hum Mol Genet* 13**:** 1677-92.

57. Badran WA, Fahmy I, Abdel-Megid WM, Elder K, Mansour R, Kent-First M (2009). Length of androgen receptor-CAG repeats in fertile and infertile Egyptian men. *J Androl* 30**:** 416-25.

58. Akinloye O, Gromoll J, Nieschlag E, Simoni M (2009). Androgen receptor gene CAG and GGN polymorphisms in infertile Nigerian men. *J Endocrinol Invest* 32**:** 797-804.

59. Casella R, Maduro MR, Misfud A, Lipshultz LI, Yong EL, Lamb DJ (2003). Androgen receptor gene polyglutamine length is associated with testicular histology in infertile patients. *J Urol* 169**:** 224-7.

60. Giwercman YL, Xu C, Arver S, Pousette A, Reneland R (1998). No association between the androgen receptor gene CAG repeat and impaired sperm production in Swedish men. *Clin Genet* 54**:** 435-6.

61. Martinez-Garza SG, Gallegos-Rivas MC, Vargas-Maciel M, Rubio-Rubio JM, de Los Monteros-Rodriguez ME, Gonzalez-Ortega C, Cancino-Villarreal P, de Lara LG, Gutierrez-Gutierrez AM (2008). Genetic screening in infertile Mexican men: chromosomal abnormalities, Y chromosome deletions, and androgen receptor CAG repeat length. *J Androl* 29**:** 654-60.

62. Mengual L, Oriola J, Ascaso C, Ballesca JL, Oliva R (2003). An increased CAG repeat length in the androgen receptor gene in azoospermic ICSI candidates. *J Androl* 24**:** 279-84.

63. Patrizio P, Leonard DG, Chen KL, Hernandez-Ayup S, Trounson AO (2001). Larger trinucleotide repeat size in the androgen receptor gene of infertile men with extremely severe oligozoospermia. *J Androl* 22**:** 444-8.

64. Rajpert-De Meyts E, Leffers H, Petersen JH, Andersen AG, Carlsen E, Jorgensen N, Skakkebaek NE (2002). CAG repeat length in androgen-receptor gene and reproductive variables in fertile and infertile men. *Lancet* 359**:** 44-6.

65. Westerveld H, Visser L, Tanck M, van der Veen F, Repping S (2008). CAG repeat length variation in the androgen receptor gene is not associated with spermatogenic failure. *Fertil Steril* 89**:** 253-9.

66. von Eckardstein S, Syska A, Gromoll J, Kamischke A, Simoni M, Nieschlag E (2001). Inverse correlation between sperm concentration and number of androgen receptor CAG repeats in normal men. *J Clin Endocrinol Metab* 86**:** 2585-90.

67. Yong EL, Loy CJ, Sim KS (2003). Androgen receptor gene and male infertility. *Hum Reprod Update* 9**:** 1-7.

68. Lubahn DB, Joseph DR, Sar M, Tan J, Higgs HN, Larson RE, French FS, Wilson EM (1988). The human androgen receptor: complementary deoxyribonucleic acid cloning, sequence analysis and gene expression in prostate. *Mol Endocrinol* 2**:** 1265-75.

69. Gao T, Marcelli M, McPhaul MJ (1996). Transcriptional activation and transient expression of the human androgen receptor. *J Steroid Biochem Mol Biol* 59**:** 9-20.

70. Lundin KB, Giwercman A, Richthoff J, Abrahamsson PA, Giwercman YL (2003). No association between mutations in the human androgen receptor GGN repeat and inter-sex conditions. *Mol Hum Reprod* 9**:** 375-9.

71. Lundin KB, Giwercman A, Dizeyi N, Giwercman YL (2007). Functional in vitro characterisation of the androgen receptor GGN polymorphism. *Mol Cell Endocrinol* 264**:** 184-7.

72. Lundin KB, Giwercman YL, Rylander L, Hagmar L, Giwercman A (2006). Androgen receptor gene GGN repeat length and reproductive characteristics in young Swedish men. *Eur J Endocrinol* 155**:** 347-54.

73. Kittles RA, Young D, Weinrich S, Hudson J, Argyropoulos G, Ukoli F, Adams-Campbell L, Dunston GM (2001). Extent of linkage disequilibrium between the androgen receptor gene CAG and GGC repeats in human populations: implications for prostate cancer risk. *Hum Genet* 109**:** 253-61.

74. Ross RK, Pike MC, Coetzee GA, Reichardt JK, Yu MC, Feigelson H, Stanczyk FZ, Kolonel LN, Henderson BE (1998). Androgen metabolism and prostate cancer: establishing a model of genetic susceptibility. *Cancer Res* 58**:** 4497-504.

75. Aschim EL, Nordenskjold A, Giwercman A, Lundin KB, Ruhayel Y, Haugen TB, Grotmol T, Giwercman YL (2004). Linkage between cryptorchidism, hypospadias, and GGN repeat length in the androgen receptor gene. *J Clin Endocrinol Metab* 89**:** 5105-9.

76. Castro-Nallar E, Bacallao K, Parada-Bustamante A, Lardone MC, Lopez PV, Madariaga M, Valdevenito R, Piottante A, Ebensperger M, Castro A (2010). Androgen receptor gene CAG and GGN repeat polymorphisms in Chilean men with primary severe spermatogenic failure. *J Androl* 31**:** 552-9.

77. Radpour R, Rezaee M, Tavasoly A, Solati S, Saleki A (2007). Association of long polyglycine tracts (GGN repeats) in exon 1 of the androgen receptor gene with cryptorchidism and penile hypospadias in Iranian patients. *J Androl* 28**:** 164-9.

78. Gelmann EP (2002). Molecular biology of the androgen receptor. *J Clin Oncol* 20**:** 3001-15.

79. Verrijdt G, Haelens A, Claessens F (2003). Selective DNA recognition by the androgen receptor as a mechanism for hormone-specific regulation of gene expression. *Mol Genet Metab* 78**:** 175-85.

80. Glass CK (1994). Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. *Endocr Rev* 15**:** 391-407.

81. Schoenmakers E, Alen P, Verrijdt G, Peeters B, Verhoeven G, Rombauts W, Claessens F (1999). Differential DNA binding by the androgen and glucocorticoid receptors involves the second Zn-finger and a C-terminal extension of the DNA-binding domains. *Biochem J* 341 (Pt 3)**:** 515-21.

82. Umesono K, Evans RM (1989). Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* 57**:** 1139-46.

83. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, Evans RM (1995). The nuclear receptor superfamily: the second decade. *Cell* 83**:** 835-9.

84. Claessens F, Alen P, Devos A, Peeters B, Verhoeven G, Rombauts W (1996). The androgen-specific probasin response element 2 interacts differentially with androgen and glucocorticoid receptors. *J Biol Chem* 271**:** 19013-6.

85. Denayer S, Helsen C, Thorrez L, Haelens A, Claessens F (2010). The rules of DNA recognition by the androgen receptor. *Mol Endocrinol* 24**:** 898-913.

86. Claessens F, Verrijdt G, Schoenmakers E, Haelens A, Peeters B, Verhoeven G, Rombauts W (2001). Selective DNA binding by the androgen receptor as a mechanism for hormone-specific gene regulation. *J Steroid Biochem Mol Biol* 76**:** 23-30.

87. Haelens A, Verrijdt G, Callewaert L, Christiaens V, Schauwaers K, Peeters B, Rombauts W, Claessens F (2003). DNA recognition by the androgen receptor: evidence for an alternative DNA-dependent dimerization, and an active role of sequences flanking the response element on transactivation. *Biochem J* 369**:** 141-51.

88. 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.

89. Roemer SC, Donham DC, Sherman L, Pon VH, Edwards DP, Churchill ME (2006). Structure of the progesterone receptor-deoxyribonucleic acid complex: novel interactions required for binding to half-site response elements. *Mol Endocrinol* 20**:** 3042-52.

90. Schwabe JW, Neuhaus D, Rhodes D (1990). Solution structure of the DNA-binding domain of the oestrogen receptor. *Nature* 348**:** 458-61.

91. Shaffer PL, Jivan A, Dollins DE, Claessens F, Gewirth DT (2004). Structural basis of androgen receptor binding to selective androgen response elements. *Proc Natl Acad Sci U S A* 101**:** 4758-63.

92. Cutress ML, Whitaker HC, Mills IG, Stewart M, Neal DE (2008). Structural basis for the nuclear import of the human androgen receptor. *J Cell Sci* 121**:** 957-68.

93. Deeb A, Jaaskelainen J, Dattani M, Whitaker HC, Costigan C, Hughes IA (2008). A novel mutation in the human androgen receptor suggests a regulatory role for the hinge region in amino-terminal and carboxy-terminal interactions. *J Clin Endocrinol Metab* 93**:** 3691-6.

94. Zuccarello D, Ferlin A, Vinanzi C, Prana E, Garolla A, Callewaert L, Claessens F, Brinkmann AO, Foresta C (2008). Detailed functional studies on androgen receptor mild mutations demonstrate their association with male infertility. *Clin Endocrinol (Oxf)* 68**:** 580-8.

95. Gioeli D, Black BE, Gordon V, Spencer A, Kesler CT, Eblen ST, Paschal BM, Weber MJ (2006). Stress kinase signaling regulates androgen receptor phosphorylation, transcription, and localization. *Mol Endocrinol* 20**:** 503-15.

96. Kesler CT, Gioeli D, Conaway MR, Weber MJ, Paschal BM (2007). Subcellular localization modulates activation function 1 domain phosphorylation in the androgen receptor. *Mol Endocrinol* 21**:** 2071-84.

97. Quigley CA, De Bellis A, Marschke KB, el-Awady MK, Wilson EM, French FS (1995). Androgen receptor defects: historical, clinical, and molecular perspectives. *Endocr Rev* 16**:** 271-321.

98. Matias PM, Donner P, Coelho R, Thomaz M, Peixoto C, Macedo S, Otto N, Joschko S, Scholz P, Wegg A, Basler S, Schafer M, Egner U, Carrondo MA (2000). Structural evidence for ligand specificity in the binding domain of the human androgen receptor. Implications for pathogenic gene mutations. *J Biol Chem* 275**:** 26164-71.

99. Sack JS, Kish KF, Wang C, Attar RM, Kiefer SE, An Y, Wu GY, Scheffler JE, Salvati ME, Krystek SR, Jr., Weinmann R, Einspahr HM (2001). Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone. *Proc Natl Acad Sci U S A* 98**:** 4904-9.

100. Moras D, Gronemeyer H (1998). The nuclear receptor ligand-binding domain: structure and function. *Curr Opin Cell Biol* 10**:** 384-91.

101. Dubbink HJ, Hersmus R, Verma CS, van der Korput HA, Berrevoets CA, van Tol J, Ziel-van der Made AC, Brinkmann AO, Pike AC, Trapman J (2004). Distinct recognition modes of FXXLF and LXXLL motifs by the androgen receptor. *Mol Endocrinol* 18**:** 2132-50.

102. Hur E, Pfaff SJ, Payne ES, Gron H, Buehrer BM, Fletterick RJ (2004). Recognition and accommodation at the androgen receptor coactivator binding interface. *PLoS Biol* 2**:** E274.

103. Faber PW, van Rooij HC, Schipper HJ, Brinkmann AO, Trapman J (1993). Two different, overlapping pathways of transcription initiation are active on the TATA-less human androgen receptor promoter. The role of Sp1. *J Biol Chem* 268**:** 9296-301.

104. Tilley WD, Marcelli M, McPhaul MJ (1990). Expression of the human androgen receptor gene utilizes a common promoter in diverse human tissues and cell lines. *J Biol Chem* 265**:** 13776-81.

105. Wolf DA, Herzinger T, Hermeking H, Blaschke D, Horz W (1993). Transcriptional and posttranscriptional regulation of human androgen receptor expression by androgen. *Mol Endocrinol* 7**:** 924-36.

106. Song CS, Her S, Slomczynska M, Choi SJ, Jung MH, Roy AK, Chatterjee B (1993). A distal activation domain is critical in the regulation of the rat androgen receptor gene promoter. *Biochem J* 294 (Pt 3)**:** 779-84.

107. Faber PW, van Rooij HC, van der Korput HA, Baarends WM, Brinkmann AO, Grootegoed JA, Trapman J (1991). Characterization of the human androgen receptor transcription unit. *J Biol Chem* 266**:** 10743-9.

108. 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.

109. Quarmby VE, Yarbrough WG, Lubahn DB, French FS, Wilson EM (1990). Autologous down-regulation of androgen receptor messenger ribonucleic acid. *Mol Endocrinol* 4**:** 22-8.

110. Yeap BB, Krueger RG, Leedman PJ (1999). Differential posttranscriptional regulation of androgen receptor gene expression by androgen in prostate and breast cancer cells. *Endocrinology* 140**:** 3282-91.

111. Kemppainen JA, Lane MV, Sar M, Wilson EM (1992). Androgen receptor phosphorylation, turnover, nuclear transport, and transcriptional activation. Specificity for steroids and antihormones. *J Biol Chem* 267**:** 968-74.

112. Zhou ZX, Lane MV, Kemppainen JA, French FS, Wilson EM (1995). Specificity of ligand-dependent androgen receptor stabilization: receptor domain interactions influence ligand dissociation and receptor stability. *Mol Endocrinol* 9**:** 208-18.

113. Shan LX, Rodriguez MC, Janne OA (1990). Regulation of androgen receptor protein and mRNA concentrations by androgens in rat ventral prostate and seminal vesicles and in human hepatoma cells. *Mol Endocrinol* 4**:** 1636-46.

114. Mukhopadhyay NK, Kim J, Cinar B, Ramachandran A, Hager MH, Di Vizio D, Adam RM, Rubin MA, Raychaudhuri P, De Benedetti A, Freeman MR (2009). Heterogeneous nuclear ribonucleoprotein K is a novel regulator of androgen receptor translation. *Cancer Res* 69**:** 2210-8.

115. Gibbs A, Schwartzman J, Deng V, Alumkal J (2009). Sulforaphane destabilizes the androgen receptor in prostate cancer cells by inactivating histone deacetylase 6. *Proc Natl Acad Sci U S A* 106**:** 16663-8.

116. Edwards J, Bartlett JM (2005). The androgen receptor and signal-transduction pathways in hormone-refractory prostate cancer. Part 1: Modifications to the androgen receptor. *BJU Int* 95**:** 1320-6.

117. Ward RD, Weigel NL (2009). Steroid receptor phosphorylation: Assigning function to site-specific phosphorylation. *Biofactors* 35**:** 528-36.

118. Zong H, Chi Y, Wang Y, Yang Y, Zhang L, Chen H, Jiang J, Li Z, Hong Y, Wang H, Yun X, Gu J (2007). Cyclin D3/CDK11p58 complex is involved in the repression of androgen receptor. *Mol Cell Biol* 27**:** 7125-42.

119. Lin HK, Yeh S, Kang HY, Chang C (2001). Akt suppresses androgen-induced apoptosis by phosphorylating and inhibiting androgen receptor. *Proc Natl Acad Sci U S A* 98**:** 7200-5.

120. Taneja SS, Ha S, Swenson NK, Huang HY, Lee P, Melamed J, Shapiro E, Garabedian MJ, Logan SK (2005). Cell-specific regulation of androgen receptor phosphorylation in vivo. *J Biol Chem* 280**:** 40916-24.

121. Wen Y, Hu MC, Makino K, Spohn B, Bartholomeusz G, Yan DH, Hung MC (2000). HER-2/neu promotes androgen-independent survival and growth of prostate cancer cells through the Akt pathway. *Cancer Res* 60**:** 6841-5.

122. Shang Y, Myers M, Brown M (2002). Formation of the androgen receptor transcription complex. *Mol Cell* 9**:** 601-10.

123. Yeh S, Chang C (1996). Cloning and characterization of a specific coactivator, ARA70, for the androgen receptor in human prostate cells. *Proc Natl Acad Sci U S A* 93**:** 5517-21.

124. Miyamoto H, Yeh S, Lardy H, Messing E, Chang C (1998). Delta5-androstenediol is a natural hormone with androgenic activity in human prostate cancer cells. *Proc Natl Acad Sci U S A* 95**:** 11083-8.

125. Gao T, Brantley K, Bolu E, McPhaul MJ (1999). RFG (ARA70, ELE1) interacts with the human androgen receptor in a ligand-dependent fashion, but functions only weakly as a coactivator in cotransfection assays. *Mol Endocrinol* 13**:** 1645-56.

126. Alen P, Claessens F, Schoenmakers E, Swinnen JV, Verhoeven G, Rombauts W, Peeters B (1999). Interaction of the putative androgen receptor-specific coactivator ARA70/ELE1alpha with multiple steroid receptors and identification of an internally deleted ELE1beta isoform. *Mol Endocrinol* 13**:** 117-28.

127. Bischoff FR, Ponstingl H (1991). Mitotic regulator protein RCC1 is complexed with a nuclear ras-related polypeptide. *Proc Natl Acad Sci U S A* 88**:** 10830-4.

128. Rush MG, Drivas G, D'Eustachio P (1996). The small nuclear GTPase Ran: how much does it run? *Bioessays* 18**:** 103-12.

129. Hsiao PW, Lin DL, Nakao R, Chang C (1999). The linkage of Kennedy's neuron disease to ARA24, the first identified androgen receptor polyglutamine region-associated coactivator. *J Biol Chem* 274**:** 20229-34.

130. Irvine RA, Ma H, Yu MC, Ross RK, Stallcup MR, Coetzee GA (2000). Inhibition of p160-mediated coactivation with increasing androgen receptor polyglutamine length. *Hum Mol Genet* 9**:** 267-74.

131. Yan J, Yu CT, Ozen M, Ittmann M, Tsai SY, Tsai MJ (2006). Steroid receptor coactivator-3 and activator protein-1 coordinately regulate the transcription of components of the insulin-like growth factor/AKT signaling pathway. *Cancer Res* 66**:** 11039-46.

132. Zhou HJ, Yan J, Luo W, Ayala G, Lin SH, Erdem H, Ittmann M, Tsai SY, Tsai MJ (2005). SRC-3 is required for prostate cancer cell proliferation and survival. *Cancer Res* 65**:** 7976-83.

133. Agoulnik IU, Vaid A, Bingman WE, 3rd, Erdeme H, Frolov A, Smith CL, Ayala G, Ittmann MM, Weigel NL (2005). Role of SRC-1 in the promotion of prostate cancer cell growth and tumor progression. *Cancer Res* 65**:** 7959-67.

134. Agoulnik IU, Vaid A, Nakka M, Alvarado M, Bingman WE, 3rd, Erdem H, Frolov A, Smith CL, Ayala GE, Ittmann MM, Weigel NL (2006). Androgens modulate expression of transcription intermediary factor 2, an androgen receptor

coactivator whose expression level correlates with early biochemical recurrence in prostate cancer. *Cancer Res* 66**:** 10594-602.

135. Wang L, Lin HK, Hu YC, Xie S, Yang L, Chang C (2004). Suppression of androgen receptor-mediated transactivation and cell growth by the glycogen synthase kinase 3 beta in prostate cells. *J Biol Chem* 279**:** 32444-52.

136. Gaughan L, Logan IR, Cook S, Neal DE, Robson CN (2002). Tip60 and histone deacetylase 1 regulate androgen receptor activity through changes to the acetylation status of the receptor. *J Biol Chem* 277**:** 25904-13.

137. Oakes MB, Eyvazzadeh AD, Quint E, Smith YR (2008). Complete androgen insensitivity syndrome--a review. *J Pediatr Adolesc Gynecol* 21**:** 305-10.

138. Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HW, Behre HM, Haugen TB, Kruger T, Wang C, Mbizvo MT, Vogelsong KM (2010). World Health Organization reference values for human semen characteristics. *Hum Reprod Update* 16**:** 231-45.

139. Hughes IA, Deeb A (2006). Androgen resistance. *Best Pract Res Clin Endocrinol Metab* 20**:** 577-98.

140. Kennedy WR, Alter M, Sung JH (1968). Progressive proximal spinal and bulbar muscular atrophy of late onset. A sex-linked recessive trait. *Neurology* 18**:** 671-80.

141. La Spada AR, Roling DB, Harding AE, Warner CL, Spiegel R, Hausmanowa-Petrusewicz I, Yee WC, Fischbeck KH (1992). Meiotic stability and genotype-phenotype correlation of the trinucleotide repeat in X-linked spinal and bulbar muscular atrophy. *Nat Genet* 2**:** 301-4.

142. Palazzolo I, Gliozzi A, Rusmini P, Sau D, Crippa V, Simonini F, Onesto E, Bolzoni E, Poletti A (2008). The role of the polyglutamine tract in androgen receptor. *J Steroid Biochem Mol Biol* 108**:** 245-53.

143. Paulson HL, Perez MK, Trottier Y, Trojanowski JQ, Subramony SH, Das SS, Vig P, Mandel JL, Fischbeck KH, Pittman RN (1997). Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3. *Neuron* 19**:** 333-44.

144. Banno H, Adachi H, Katsuno M, Suzuki K, Atsuta N, Watanabe H, Tanaka F, Doyu M, Sobue G (2006). Mutant androgen receptor accumulation in spinal and bulbar muscular atrophy scrotal skin: a pathogenic marker. *Ann Neurol* 59**:** 520-6.

145. Usdin K (2008). The biological effects of simple tandem repeats: lessons from the repeat expansion diseases. *Genome Res* 18**:** 1011-9.

146. Zoghbi HY, Orr HT (2000). Glutamine repeats and neurodegeneration. *Annu Rev Neurosci* 23**:** 217-47.

147. Niu Y, Altuwaijri S, Lai KP, Wu CT, Ricke WA, Messing EM, Yao J, Yeh S, Chang C (2008). Androgen receptor is a tumor suppressor and proliferator in prostate cancer. *Proc Natl Acad Sci U S A* 105**:** 12182-7.

148. Stanbrough M, Leav I, Kwan PW, Bubley GJ, Balk SP (2001). Prostatic intraepithelial neoplasia in mice expressing an androgen receptor transgene in prostate epithelium. *Proc Natl Acad Sci U S A* 98**:** 10823-8.

149. Han G, Buchanan G, Ittmann M, Harris JM, Yu X, Demayo FJ, Tilley W, Greenberg NM (2005). Mutation of the androgen receptor causes oncogenic transformation of the prostate. *Proc Natl Acad Sci U S A* 102**:** 1151-6.

150. Linja MJ, Visakorpi T (2004). Alterations of androgen receptor in prostate cancer. *J Steroid Biochem Mol Biol* 92**:** 255-64.

151. Shen MM, Abate-Shen C (2010). Molecular genetics of prostate cancer: new prospects for old challenges. *Genes Dev* 24**:** 1967-2000.

152. Mononen N, Syrjakoski K, Matikainen M, Tammela TL, Schleutker J, Kallioniemi OP, Trapman J, Koivisto PA (2000). Two percent of Finnish prostate cancer patients have a germ-line mutation in the hormone-binding domain of the androgen receptor gene. *Cancer Res* 60**:** 6479-81.

153. Gruber SB, Chen H, Tomsho LP, Lee N, Perrone EE, Cooney KA (2003). R726L androgen receptor mutation is uncommon in prostate cancer families in the united states. *Prostate* 54**:** 306-9.

154. Coetzee GA, Ross RK (1994). Re: Prostate cancer and the androgen receptor. *J Natl Cancer Inst* 86**:** 872-3.

155. Ashtiani ZO, Hasheminasab SM, Ayati M, Goulian BS, Modarressi MH (2010). Are GSTM1, GSTT1 and CAG Repeat Length of Androgen Receptor Gene Polymorphisms Associated with Risk of Prostate Cancer in Iranian Patients? *Pathol Oncol Res*.

156. Bratt O, Borg A, Kristoffersson U, Lundgren R, Zhang QX, Olsson H (1999). CAG repeat length in the androgen receptor gene is related to age at diagnosis of prostate cancer and response to endocrine therapy, but not to prostate cancer risk. *Br J Cancer* 81**:** 672-6.

157. Chen C, Lamharzi N, Weiss NS, Etzioni R, Dightman DA, Barnett M, DiTommaso D, Goodman G (2002). Androgen receptor polymorphisms and the incidence of prostate cancer. *Cancer Epidemiol Biomarkers Prev* 11**:** 1033-40.

158. Gsur A, Preyer M, Haidinger G, Zidek T, Madersbacher S, Schatzl G, Marberger M, Vutuc C, Micksche M (2002). Polymorphic CAG repeats in the androgen receptor gene, prostate-specific antigen polymorphism and prostate cancer risk. *Carcinogenesis* 23**:** 1647-51.

159. Hakimi JM, Schoenberg MP, Rondinelli RH, Piantadosi S, Barrack ER (1997). Androgen receptor variants with short glutamine or glycine repeats may identify unique subpopulations of men with prostate cancer. *Clin Cancer Res* 3**:** 1599-608.

160. Latil AG, Azzouzi R, Cancel GS, Guillaume EC, Cochan-Priollet B, Berthon PL, Cussenot O (2001). Prostate carcinoma risk and allelic variants of genes involved in androgen biosynthesis and metabolism pathways. *Cancer* 92**:** 1130-7.

161. Santos ML, Sarkis AS, Nishimoto IN, Nagai MA (2003). Androgen receptor CAG repeat polymorphism in prostate cancer from a Brazilian population. *Cancer Detect Prev* 27**:** 321-6.

162. Ekman P, Gronberg H, Matsuyama H, Kivineva M, Bergerheim US, Li C (1999). Links between genetic and environmental factors and prostate cancer risk. *Prostate* 39**:** 262-8.

163. Lindstrom S, Ma J, Altshuler D, Giovannucci E, Riboli E, Albanes D, Allen NE, Berndt SI, Boeing H, Bueno-de-Mesquita HB, Chanock SJ, Dunning AM, Feigelson HS, Gaziano JM, Haiman CA, Hayes RB, Henderson BE, Hunter DJ, Kaaks R, Kolonel LN, Le Marchand L, Martinez C, Overvad K, Siddiq A, Stampfer M, Stattin P, Stram DO, Thun MJ, Trichopoulos D, Tumino R, Virtamo J, Weinstein SJ, Yeager M, Kraft P, Freedman ML (2010). A large study of androgen receptor germline variants and their relation to sex hormone levels and prostate cancer risk. Results from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium. *J Clin Endocrinol Metab* 95**:** E121-7.

164. Mononen N, Ikonen T, Autio V, Rokman A, Matikainen MP, Tammela TL, Kallioniemi OP, Koivisto PA, Schleutker J (2002). Androgen receptor CAG polymorphism and prostate cancer risk. *Hum Genet* 111**:** 166-71.

165. Panz VR, Joffe BI, Spitz I, Lindenberg T, Farkas A, Haffejee M (2001). Tandem CAG repeats of the androgen receptor gene and prostate cancer risk in black and white men. *Endocrine* 15**:** 213-6.

166. Platz EA, Giovannucci E, Dahl DM, Krithivas K, Hennekens CH, Brown M, Stampfer MJ, Kantoff PW (1998). The androgen receptor gene GGN

microsatellite and prostate cancer risk. *Cancer Epidemiol Biomarkers Prev* 7**:** 379-84.

167. Price DK, Chau CH, Till C, Goodman PJ, Baum CE, Ockers SB, English BC, Minasian L, Parnes HL, Hsing AW, Reichardt JK, Hoque A, Tangen CM, Kristal AR, Thompson IM, Figg WD (2010). Androgen receptor CAG repeat length and association with prostate cancer risk: results from the prostate cancer prevention trial. *J Urol* 184**:** 2297-302.

168. Stanford JL, Just JJ, Gibbs M, Wicklund KG, Neal CL, Blumenstein BA, Ostrander EA (1997). Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer Res* 57**:** 1194-8.

169. Giovannucci E, Stampfer MJ, Krithivas K, Brown M, Dahl D, Brufsky A, Talcott J, Hennekens CH, Kantoff PW (1997). The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc Natl Acad Sci U S A* 94**:** 3320-3.

170. Ingles SA, Ross RK, Yu MC, Irvine RA, La Pera G, Haile RW, Coetzee GA (1997). Association of prostate cancer risk with genetic polymorphisms in vitamin D receptor and androgen receptor. *J Natl Cancer Inst* 89**:** 166-70.

171. Nam RK, Elhaji Y, Krahn MD, Hakimi J, Ho M, Chu W, Sweet J, Trachtenberg J, Jewett MA, Narod SA (2000). Significance of the CAG repeat polymorphism of the androgen receptor gene in prostate cancer progression. *J Urol* 164**:** 567-72.

172. Beilin J, Harewood L, Frydenberg M, Mameghan H, Martyres RF, Farish SJ, Yue C, Deam DR, Byron KA, Zajac JD (2001). A case-control study of the androgen receptor gene CAG repeat polymorphism in Australian prostate carcinoma subjects. *Cancer* 92**:** 941-9.

173. Hardy DO, Scher HI, Bogenreider T, Sabbatini P, Zhang ZF, Nanus DM, Catterall JF (1996). Androgen receptor CAG repeat lengths in prostate cancer: correlation with age of onset. *J Clin Endocrinol Metab* 81**:** 4400-5.

174. Correa-Cerro L, Wohr G, Haussler J, Berthon P, Drelon E, Mangin P, Fournier G, Cussenot O, Kraus P, Just W, Paiss T, Cantu JM, Vogel W (1999). (CAG)nCAA and GGN repeats in the human androgen receptor gene are not associated with prostate cancer in a French-German population. *Eur J Hum Genet* 7**:** 357-62.

175. Hsing AW, Gao YT, Wu G, Wang X, Deng J, Chen YL, Sesterhenn IA, Mostofi FK, Benichou J, Chang C (2000). Polymorphic CAG and GGN repeat lengths in the androgen receptor gene and prostate cancer risk: a population-based case-control study in China. *Cancer Res* 60**:** 5111-6.

176. Huhtaniemi IT, Pye SR, Limer KL, Thomson W, O'Neill TW, Platt H, Payne D, John SL, Jiang M, Boonen S, Borghs H, Vanderschueren D, Adams JE, Ward KA, Bartfai G, Casanueva F, Finn JD, Forti G, Giwercman A, Han TS, Kula K, Lean ME, Pendleton N, Punab M, Silman AJ, Wu FC (2009). Increased estrogen rather than decreased androgen action is associated with longer androgen receptor CAG repeats. *J Clin Endocrinol Metab* 94**:** 277-84.

177. Bosland MC (2000). The role of steroid hormones in prostate carcinogenesis. *J Natl Cancer Inst Monogr***:** 39-66.

178. Eaton NE, Reeves GK, Appleby PN, Key TJ (1999). Endogenous sex hormones and prostate cancer: a quantitative review of prospective studies. *Br J Cancer* 80**:** 930-4.

179. Hsing AW (1996). Hormones and prostate cancer: where do we go from here? *J Natl Cancer Inst* 88**:** 1093-5.

180. Hsing AW (2001). Hormones and prostate cancer: what's next? *Epidemiol Rev* 23**:** 42-58.

181. Platz EA, Giovannucci E (2004). The epidemiology of sex steroid hormones and their signaling and metabolic pathways in the etiology of prostate cancer. *J Steroid Biochem Mol Biol* 92**:** 237-53.

182. Shaneyfelt T, Husein R, Bubley G, Mantzoros CS (2000). Hormonal predictors of prostate cancer: a meta-analysis. *J Clin Oncol* 18**:** 847-53.

183. Roddam AW, Allen NE, Appleby P, Key TJ (2008). Endogenous sex hormones and prostate cancer: a collaborative analysis of 18 prospective studies. *J Natl Cancer Inst* 100**:** 170-83.

184. WHO (2000). *WHO Manual for the Standardised Investigation, Diagnosis and Management of the Infertile Male*. Cambridge University Press: Cambridge.

185. Gottlieb B, Lombroso R, Beitel LK, Trifiro MA (2005). Molecular pathology of the androgen receptor in male (in)fertility. *Reprod Biomed Online* 10**:** 42-8.

186. De Gendt K, Swinnen JV, Saunders PT, Schoonjans L, Dewerchin M, Devos A, Tan K, Atanassova N, Claessens F, Lecureuil C, Heyns W, Carmeliet P, Guillou F, Sharpe RM, Verhoeven G (2004). A Sertoli cell-selective knockout of the androgen receptor causes spermatogenic arrest in meiosis. *Proc Natl Acad Sci U S A* 101**:** 1327-32.

187. Denolet E, De Gendt K, Allemeersch J, Engelen K, Marchal K, Van Hummelen P, Tan KA, Sharpe RM, Saunders PT, Swinnen JV, Verhoeven G (2006). The

effect of a sertoli cell-selective knockout of the androgen receptor on testicular gene expression in prepubertal mice. *Mol Endocrinol* 20**:** 321-34.

188. Davis-Dao CA, Tuazon ED, Sokol RZ, Cortessis VK (2007). Male infertility and variation in CAG repeat length in the androgen receptor gene: a meta-analysis. *J Clin Endocrinol Metab* 92**:** 4319-26.

189. Joseph Sambrook DWR (2001). *Molecular Cloning: A Laboratory Manual (3-Volume Set)* 3rd edn. Cold Spring Harbor Laboratory; 3rd edition (January 15, 2001).

190. Gluzman Y (1981). SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* 23**:** 175-82.

191. Aruffo A (2001). Transient expression of proteins using COS cells. *Curr Protoc Neurosci* Chapter 4**:** Unit 4 7.

192. Bradford MM (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72**:** 248-54.

193. Primrose Sandy B, Twyman Richard M, Old Robert W (2001). *Principles of Gene Manipulation*, 6th edn. Blackwell publishing: Oxford.

194. Butland SL, Devon RS, Huang Y, Mead CL, Meynert AM, Neal SJ, Lee SS, Wilkinson A, Yang GS, Yuen MM, Hayden MR, Holt RA, Leavitt BR, Ouellette BF (2007). CAG-encoded polyglutamine length polymorphism in the human genome. *BMC Genomics* 8**:** 126.

195. Hayashi Y, Yamamoto M, Ohmori S, Kikumori T, Imai T, Funahashi H, Seo H (1999). Polymorphism of homopolymeric glutamines in coactivators for nuclear hormone receptors. *Endocr J* 46**:** 279-84.

196. Lieberman AP, Robins DM (2008). The androgen receptor's CAG/glutamine tract in mouse models of neurological disease and cancer. *J Alzheimers Dis* 14**:** 247-55.

197. Nenonen H, Bjork C, Skjaerpe PA, Giwercman A, Rylander L, Svartberg J, Giwercman YL (2010). CAG repeat number is not inversely associated with androgen receptor activity in vitro. *Mol Hum Reprod* 16**:** 153-7.

198. Nenonen HA, Giwercman A, Hallengren E, Giwercman YL (2010) Non-linear association between androgen receptor CAG repeat length and risk of male subfertility - a meta-analysis. *Int J Androl*.33

199. Rodriguez-Gonzalez G, Cabrera S, Ramirez-Moreno R, Bilbao C, Diaz-Chico JC, Serra L, Chesa N, Cabrera JJ, Diaz-Chico BN (2009). Short alleles of both GGN and CAG repeats at the exon-1 of the androgen receptor gene are associated to increased PSA staining and a higher Gleason score in human prostatic cancer. *J Steroid Biochem Mol Biol* 113**:** 85-91.

200. Albertelli MA, Scheller A, Brogley M, Robins DM (2006). Replacing the mouse androgen receptor with human alleles demonstrates glutamine tract length-dependent effects on physiology and tumorigenesis in mice. *Mol Endocrinol* 20**:** 1248-60.

201. Nenonen HA, Giwercman A, Hallengren E, Giwercman YL (2010). Non-linear association between androgen receptor CAG repeat length and risk of male subfertility - a meta-analysis. *Int J Androl*.33

202. Calistro Alvarado L (2010). Population differences in the testosterone levels of young men are associated with prostate cancer disparities in older men. *Am J Hum Biol* 22**:** 449-55.

203. Carpenter WR, Robinson WR, Godley PA (2008). Getting over testosterone: postulating a fresh start for etiologic studies of prostate cancer. *J Natl Cancer Inst* 100**:** 158-9.

204. Hsing AW, Chu LW, Stanczyk FZ (2008). Androgen and prostate cancer: is the hypothesis dead? *Cancer Epidemiol Biomarkers Prev* 17**:** 2525-30.

205. Sofikerim M, Eskicorapci S, Oruc O, Ozen H (2007). Hormonal predictors of prostate cancer. *Urol Int* 79**:** 13-8.

206. Giwercman A, Richiardi L, Kaijser M, Ekbom A, Akre O (2005). Reduced risk of prostate cancer in men who are childless as compared to those who have fathered a child: a population based case-control study. *Int J Cancer* 115**:** 994-7.

207. Jorgensen KT, Pedersen BV, Johansen C, Frisch M (2008). Fatherhood status and prostate cancer risk. *Cancer* 112**:** 919-23.

208. Ruhayel Y, Giwercman A, Ulmert D, Rylander L, Bjartell A, Manjer J, Berglund G, Giwercman YL (2010). Male infertility and prostate cancer risk: a nested case-control study. *Cancer Causes Control* 21**:** 1635-43.