DISSERTATIONS IN
HEALTH
SCIENCES

HARRI MAKKONEN

Androgen Receptor-Mediated Gene Activation in Prostate Cancer Cells

PUBLICATIONS OF THE UNIVERSITY OF EASTERN FINLAND
Dissertations in Health Sciences

UNIVERSITY OF
EASTERN FINLAND

HARRI MAKKONEN

Androgen Receptor-Mediated Gene Activation in Prostate Cancer Cells

To be presented by permission of the Faculty of Health Sciences, University of Eastern Finland for public examination in the Auditorium L21, Snellmania building, University of Eastern Finland, on Saturday 18th September 2010, at 12 noon

Publications of the University of Eastern Finland
Dissertations in Health Sciences
19

Institute of Biomedicine
School of Medicine, Faculty of Health Sciences
University of Eastern Finland
Kuopio
2010

Kopijyvä Oy
Kuopio, 2010

Editors:
Professor Veli-Matti Kosma, M.D., Ph.D.
Department of Pathology, Institute of Clinical Medicine
School of Medicine, Faculty of Health Sciences

Professor Hannele Turunen, Ph.D.
Department of Nursing Science
Faculty of Health Sciences

Distribution:

University of Eastern Finland Library / Sales of Publications
P.O.Box 1627, FI-70211 Kuopio, Finland
http://www.uef.fi/kirjasto

ISBN: 978-952-61-0173-6 (print)
ISBN: 978-952-61-0174-3 (pdf)
ISSN: 1798-5706 (print)
ISSN: 1798-5714 (pdf)
ISSNL: 1798-5706

Author's address: Institute of Biomedicine, School of Medicine
University of Eastern Finland
P.O.Box 1627
FI-70211 Kuopio, Finland

Supervisors: Professor Jorma Palvimo, Ph.D.
Institute of Biomedicine
University of Eastern Finland
Kuopio, Finland

Docent Sami Väisänen, Ph.D.
Department of Biosciences
University of Eastern Finland
Kuopio, Finland

Reviewers: Professor Pirkko Härkönen, M.D., Ph.D.
Institute of Biomedicine
University of Turku
Turku, Finland

Professor Jukka Hakkola, M.D., Ph.D.
Department of Pharmacology and Toxicology
University of Oulu
Oulu, Finland

Opponent: Docent Pekka Kallio, Ph.D.
Orion Pharma
Turku, Finland

Makkonen, Harri Tapani. Androgen receptor-mediated gene activation in prostate cancer cells. Publications of the University of Eastern Finland. Dissertations in Health Sciences 19. 2010. 84 p.
ISBN: 978-952-61-0173-6 (print)
ISBN: 978-952-61-0174-3 (pdf)
ISSN: 1798-5706 (print)
ISSN: 1798-5714 (pdf)
ISSNL: 1798-5706

ABSTRACT

Androgens, testosterone and 5α-dihydrotestosterone are responsible for the male phenotype and sexual characteristics. The effects of androgens are mediated by a specific nuclear receptor referred to as the androgen receptor (AR). Androgen-bound AR is localized to the nucleus where it binds to its response elements (AREs) and activates target gene transcription. AR-mediated transcription is crucially involved in normal prostate development and maintenance as well as in the development of prostate cancer (PC) that is the most common cancer in western males. AR-mediated transcription has been studied in detail only at the level of a few model genes from which *prostate specific antigen* (*PSA*) has been the most studied. This study had two main objectives: firstly, to clarify the mechanisms of AR-mediated gene regulation at the chromatin level in PC cells and secondly, to study the molecular mechanisms of PC and drug resistance of PC cells. In the first part, the AR-dependent transcription of two different AR target genes was characterized; *ETS-like transcription factor 4* (*ELK4*) was a novel and *FK506-binding protein 51* (*FKBP51*) a rarely studied AR target gene. It was demonstrated that *ELK4* harbors two functional AREs in its proximal promoter. Instead, the AREs of *FKBP51* are located in distal intronic and upstream enhancers. Since the glucocorticoid receptor (GR) shares partly the same binding sites with AR, it was decided to compare the regulation of *FKBP51* between these two receptors. Interestingly, only minor differences in the regulatory mechanisms were found. Both receptors brought about similar changes in the chromatin structure and covalent histone modifications as well as RNA polymerase II occupancy. The main differences were observed in binding affinity and periodicity of AR binding on certain AREs/GREs. It was also found that AR, but not GR, could regulate *FKBP51*'s neighbor gene *Chromosome 6 open reading frame 81* (*C6orf81*) due to different receptor binding affinity to the closest enhancer region rather than due to differential binding of the insulator protein CTCF. The second part of the thesis examined the mechanisms involved in PC development and progression as well as the mechanisms of drug resistance. It was shown that ELK4 can promote PC cell growth and its expression is elevated during cancer progression. By comparing LNCaP and VCaP cells, it was also found that overexpression of AR in VCaP cells could affect the function of antiandrogen drugs. In conclusion, this dissertation provides new information on AR-mediated gene activation and molecular mechanisms of PC progression and drug resistance, which may be applied to development of new PC therapies.

National Library of Medicine Classification: QU 475, WJ 762, WJ 875

Medical Subject Headings (MeSH): Androgens; Cell Line, Tumor; Enhancer Elements, Genetic; Receptors, Androgen; Promoter Regions, Genetic; Prostatic Neoplasms; Transcription, Genetic

Makkonen, Harri Tapani. Androgeenireseptorivälitteinen geeniaktivaatio eturauhassyöpäsoluissa. Publications of the University of Eastern Finland. Dissertations in Health Sciences 19. 2010. 84 sivua.
ISBN: 978-952-61-0173-6 (print)
ISBN: 978-952-61-0174-3 (pdf)
ISSN: 1798-5706 (print)
ISSN: 1798-5714 (pdf)
ISSNL: 1798-5706

TIIVISTELMÄ

Androgeenit, testosteroni ja 5α-dihydrotestosteroni, vastaavat miehisten sukupuolisominaisuuksien kehittymisestä. Tumareseptoreihin kuuluva androgeenireseptori (AR) välittää androgeenien vaikutuksia elimistössä. Androgeenien sitoutuessa AR:iin reseptori-ligandikompleksi siirtyy tumaan, jossa se sitoutuu androgeenivaste-elementteihin (ARE) ja sitä kautta aktivoi kohdegeeniensä transkriptiota. Eturauhasen normaali kehitys ja toiminnan ylläpito ovat riippuvaisia AR-välitteisestä geeninsäätelystä. AR on mukana myös länsimaisten miesten yleisimmän syövän, eturauhassyövän (PC), kehittymisessä. AR:n säätelemää transkriptiota on tutkittu yksityiskohtaisesti ainoastaan muutaman geenin osalta, joista selvästi tutkituin on *eturauhaselle spesifinen antigeeni* (*PSA*). Väitöskirjan ensimmäisenä tavoitteena oli tutkia AR-välitteisen geeninsäätelyn mekanismeja kromatiinitasolla PC-soluissa. Niitä tutkittiin kahden kohdegeenin avulla, joista *ELK4* oli uusi ja *FKBP51* ennestään vähän tutkittu androgeeneilla säädelty geeni. *ELK4*:n lähisäätelyalueelta löydettiin kaksi ARE:ä, joiden välityksellä AR aktivoi geeniä. Sen sijaan *FKBP51*:n ARE:t sijaitsevat kaukana intronissa ja ylävirran geenien välisellä alueella olevissa lisääjäelementeissä. *FKBP51*:n AR-välitteistä säätelyä verrattiin myös glukokortikoidireseptori (GR) -välitteiseen säätelyyn, koska osan GR:n ja AR:n vaste-elementeistä tiedetään olevan samoja. Säätelymekanismien välillä löytyi ainoastaan pieniä eroja; lähinnä reseptorien sitoutumisaffiniteettien ja -aikojen väliltä. Molemmat reseptorit aiheuttivat samansuuntaisia muutoksia kromatiinin rakenteessa, histonien kovalenttisissa muokkauksissa ja RNA-polymeraasi II:n sijoittumisessa kromatiinille. Huomattiin myös, että toisin kuin GR, AR säätelee *FKBP51*:n viereistä geeniä *C6orf81*:tä johtuen ennemminkin reseptorien erilaisista sitoutumisaffiniteeteista läheisimmälle lisääjäelementille kuin eristäjäproteiini CTCF:n erilaisesta sitoutumisesta. Väitöskirjan toisena tavoitteena oli tutkia PC:n sekä sen lääkeresistenssin syntyyn liittyviä molekyylitason mekanismeja. ELK4:n määrän havaittiin lisääntyvän sitä mukaa, kun PC etenee pahempilaatuiseksi, ja että ELK4 edistää PC-solujen kasvua. PC-lääkkeinä käytettyjen antiandrogeenien ominaisuuksia verrattiin LNCaP- ja VCaP-soluissa. Tulosten perusteella VCaP-solujen AR:n ylituotanto vaikuttaa lääkeaineiden kykyyn estää AR-välitteistä geeninsäätelyä. Tämä väitöstutkimus toi täten uutta tietoa AR-välitteisestä geeninsäätelystä sekä PC:n ja lääkeresistenssin kehittymisen molekyylitason mekanismeista. Tuloksia voitaneen soveltaa uusien androgeenivaikutusten salpaamiskeinojen kehitystyössä.

Luokitus: QU 475, WJ 762, WJ 875

Yleinen suomalainen asiasanasto (YSA): androgeenit; eturauhassyöpä; geenit; reseptorit; syöpäsolut

Acknowledgements

This study was carried out in the Institute of Biomedicine, School of Medicine, and Faculty of Health Sciences at the University of Eastern Finland.

I would like to thank all directions and the people who have participated in this project. Especially, I wish to express my thanks to:

Professor Jorma Palvimo, my principal supervisor, for his invaluable support in all divisions of the work. Docent Sami Väisänen, my second supervisor, for his enormous background support during the project and careful reading of the manuscript.

Professors Pirkko Härkönen and Jukka Hakkola, the official reviewers, for their constructive criticism and comments. Doctor Ewen MacDonald for revising the language of this thesis. Docent Pekka Kallio for his agreement to be my official opponent in public defence.

All coauthors of the original articles: Tiina Jääskeläinen, Tiina Pitkänen-Arsiola, Miia Rytinki, Kati Waltering, Mikko Mättö, Tapio Visakorpi, Miia Kauhanen and Ville Paakinaho. Merja Räsänen and Eija Korhonen for skillful technical support.

All current and former members of prof. Palvimo's group and all the other people in medical biochemistry unit.

Katri for her patience during the project, especially after our first-born child, and for careful reading of all the initial manuscripts I have been written. Jalo and Nuutti for their being in my life.

All financial supporters: Academy of Finland, Association for International Cancer Research, Finnish Cancer Foundation, Sigrid Jusélius Foundation, University of Kuopio (travel grant), Graduate School of Molecular Medicine (travel grant), Faculty of Health Sciences/University of Eastern Finland, Kuopio Naturalists' Society, Helena Vuorenmies Foundation, Karjalan Sivistysseura, Finnish Union of Experts in Science, Kuopio University Foundation, Emil Aaltonen Foundation, Finnish Cultural Foundation North Savo Regional Fund, Ida Montini Foundation, North Savo Cancer Foundation, and The Finnish Medical Foundation.

Kuopio, September 2010

Harri Makkonen

List of original publications

This dissertation is based on the following original publications referred to in the text by their corresponding Roman numerals (I-IV). In addition, some unpublished results will be presented.

I **Makkonen H, Jääskeläinen T, Pitkänen-Arsiola T, Rytinki M, Waltering KK, Mättö M, Visakorpi T, Palvimo JJ.** Identification of ETS-like transcription factor 4 as a novel androgen receptor target in prostate cancer cells. Oncogene. 2008 27:4865-76.

II **Makkonen H, Kauhanen M, Paakinaho V, Jääskeläinen T, Palvimo JJ.** Long-range activation of FKBP51 transcription by the androgen receptor via distal intronic enhancers. Nucleic Acids Res. 2009 37:4135-48.

III **Paakinaho V, Makkonen H, Jääskeläinen T, Palvimo JJ.** Glucocorticoid Receptor Activates Poised FKBP51 Locus through Long-Distance Interactions. Mol Endocrinol. 2010 24:511-25.

IV **Makkonen H, Kauhanen M, Jääskeläinen T, Palvimo JJ.** Androgen receptor amplification is reflected in the transcriptional responses of Vertebral-Cancer of the Prostate cells. Mol Cell Endocrinol. In press.

The publishers of the original publications have kindly granted permission to reprint the articles in this dissertation.

Contents

Abbreviations

3C	chromatin conformation capture
AF	activation function
Ago	argonaute
AIS	androgen insensitivity syndrome
AP-1	activator protein 1
AR	androgen receptor
ARE	androgen response element
ARKO	ubiquitous knock-out of the AR
BIC	bicalutamide
BRE	TFIIB response element
BRG1	brahma-related gene 1
BRM	brahma
BTA	basal transcription apparatus
CARM1	coactivator-associated arginine methyltransferase 1
CBP	CREB-binding protein
CREB	cAMP response element-binding protein
CTA	cyproterone acetate
CTCF	CCCTC-binding factor
CTD	C-terminal domain
DBD	DNA-binding domain
DHT	5α-dihydrotestosterone
DNMT	DNA methyltransferase
DPE	downstream promoter element
DR3	direct repeat separated by three nucleotides
ELK4	ets-like TF 4
ER	estrogen receptor
ERG	v-ets erythroblastosis virus E26 oncogene homolog
ETS	E twenty-six or E26 transformation-specific
ETV	ets variant
FKBP51	FK506-binding protein 51
FOXA1	forkhead box A1
GATA1	GATA-binding protein 1
GnRH	gonadotropin-releasing hormone
GR	glucocorticoid receptor
HAT	histone acetyltransferase
HDAC	histone deacetylase
HDM	histone demethylases
HMT	histone methyltransferase
holo-AR	androgen-bound AR
Inr	initiator element
IR3	inverted repeat separated by three nucleotides
ISWI	imitation of SWI
JmjC	jumonji-C-terminal
LBD	ligand-binding domain
LH	luteinizing hormone
LNCaP	PC derived from lymph node metastasis
LSD	lysine-specific demethylase
Mi-2/NuRD	mi-2/nucleosome remodeling deacetylase complex
mRNA	messenger RNA
MTE	motif ten element
ncRNA	non-protein-coding RNA
NR	nuclear receptor
NTD	N-terminal domain
P/CAF	p300/CBP-associated factor
PC	prostate cancer
PML	promyelocytic leukaemia
PR	progesterone receptor
PSA	prostate specific antigen
PTGS	post-transcriptional gene silencing
RNAP	RNA polymerase
S/MAR	scaffold/matrix attachment regions
SAP-1	SRF accessory protein 1
siRNA	small interfering RNA
SLC45A3	solute carrier family 45 member 3
SR	steroid receptor
SRC	steroid receptor coactivator

SRF	serum response factor
SUMO	small ubiquitin-like modifier
SWI/SNF	switch mating type/sucrose non-fermenting
TAF	TBP-associated factor
TBP	TATA-binding protein
TCF	ternary complex factor
TF	transcription factor
TGS	transcriptional gene silencing
TMPRSS2	transmembrane serine protease 2
TRAMP	transgenic adenocarcinoma of the mouse prostate
TSS	transcription start site
UTR	untranslated region
VCaP	PC derived from vertebrae metastasis

1. Introduction

The human genome contains approximately 20,000 to 25,000 protein-coding genes. In addition, there are thousands of genes expressing non-coding RNAs (ncRNA), which are not coding for proteins. However, out of the 3.2 billion nucleotides, only 1–2% code for amino acid sequences of proteins (exons). The rest of the genome consists of introns, intergenic regions, regulatory regions, repetitive sequences, telomeres, non-coding genes, etc. In the post-genomic era, the future goal is to understand how genes are regulated, how they are connected together, and how the regulation changes for example during aging, illness, nutrition, and drug administration. Knowing the gene regulation *per se* is a prerequisite for developing advanced cures for many diseases.

Gene regulation is a complex event which consists of several simultaneous and sequential processes. The chromatin structure has a significant role in transcription initiation and therefore the first step in the gene regulation is usually local chromatin remodeling. Nuclear proteins that have the capability to modify chromatin structure appropriately in response to gene activation or repression signals are recruited to the chromatin by activated transcription factors (TFs) which regulate the corresponding gene. The final goal in transcription initiation is to activate an enzyme synthesizing RNA, an RNA polymerase (RNAP), *via* the recruited coactivators.

Androgen receptor (AR) is a hormone-inducible TF belonging to the nuclear receptor (NR) superfamily, which is the largest family of DNA-binding TFs in humans. The natural ligands for AR are the male sex hormones, testosterone and 5α-dihydrotestosterone (DHT). Testosterone is produced mainly by testes and the more potent DHT in the target tissues. Androgen-bound AR (holo-AR) is translocated from the cytosol to the nucleus where it binds to its response elements (AREs) or interacts with other sequence-specific transcription factors, resulting in activation or repression of gene transcription. Perturbations in AR-mediated gene regulation are strongly linked to the development of prostate cancer (PC). PC is the leading diagnosed cancer and the second leading cancer-caused death in western males. The search for an efficient therapy for PC will depend on careful examination of the mechanisms of AR-mediated gene regulation.

In this dissertation, the mechanisms of AR-mediated gene activation will be discussed at the example of two model genes, *ETS-like TF 4* (*ELK4*) and *FK506-binding protein 51* (*FKBP51*). In addition, the molecular mechanisms of PC development, especially the role of ELK4, will be discussed as well as the mechanisms of development of drug resistance in PC.

2. Review of literature

2.1 NUCLEAR ARCHITECTURE

The nucleus is a membrane-bounded cell organelle containing most of the genetic material of the eukaryotic cell. The nuclear envelope is composed of two concentric lipid bilayers, the inner and outer nuclear membranes, from which the outer membrane is continuous with the endoplasmic reticulum. The structure of the envelope is supported by an underlying, fibrous meshwork called the nuclear lamina, which is composed of fibrous lamin proteins. The nuclear lamina is directly attached to the lipids in the inner nuclear membrane as well as to proteins with in the membrane. Special structures called the nuclear pore complexes serve as the sole channels through the envelope. Macromolecules, such as proteins and RNAs, and small polar compounds can travel from the cytoplasm to the nucleus, or *vice versa*, through these channels. In humans, the genetic material, chromatin, is divided into 46 parts called chromosomes, which are attached to the nuclear matrix, the major non-chromatin component of the nucleus. The isolated character of the nucleus is responsible for the main differences between prokaryotic and eukaryotic gene regulation. Overall, the nuclear architecture is dynamic rather than a stable structure and it plays a remarkable role in gene regulation (Lanctôt *et al*. 2007). In the next chapters, the nuclear architecture will be discussed in the context of transcriptional regulation.

2.1.1 Nuclear matrix and subnuclear structures

The solid component of the nucleus can be divided into the nuclear matrix and chromatin. There are also numerous soluble components (nucleoplasm) in the nucleus, such as proteins, RNAs, electrolytes, nucleotides, etc. The nuclear matrix is a poorly studied structure and it was originally defined as the non-chromatin structures of the nucleus readily observed in unextracted cells under the electron microscope (Fawcett 1966). It is also called nuclear scaffold or nucleoskeleton, because it bears some similarities to the cytoskeleton. The matrix consists of two separate parts: the nuclear lamina and the internal nuclear matrix that are, however, connected. The internal matrix was first discovered as RNA containing protein structures called the fibrogranular

ribonucleoprotein (RNP) network. In later studies, the functions of the substructures of the RNP were resolved. For example, structures called perichromatin fibrils were found to be the sites of RNA transcription and the interchromatin granule clusters were observed to be involved in RNA splicing (Nickerson 2001). The core of the matrix network is formed by 10-nm branched filaments, which are composed of A- and B-type lamin and emerin proteins. These core filaments are then covered with proteins and RNAs, producing rough surface of the filaments and more evident granules (Elcock and Bridger 2008). The protein composition of the matrix is quite complex and more than 400 proteins are known to associate with the nuclear matrix (Mika and Rost 2005).

The concept of the nuclear matrix is not very clear, and sometimes certain special structures within the nucleus, such as nucleolus, nuclear speckles, PML-bodies, and Cajal bodies, are counted as parts of the nuclear matrix or at least the nuclear matrix is postulated to be involved in the formation of these structures (Fig. 1). Nucleoli are prominent structures where the ribosomal RNA is transcribed by RNAPI and coupled with the ribosomal proteins (Hernandez-Verdun 2006). One function of nuclear speckles is to act as RNA splicing factories, whereas Cajal bodies are involved in biogenesis of nuclear RNA (Handwerger and Gall 2006). PML-bodies are formed mainly by SP100 and promyelocytic leukemia (PML) proteins. The exact role of PML-bodies is not known, but they have been often linked to tumor suppression, apoptosis, transcriptional regulation, and DNA repair (Bernardi and Pandolfi 2007).

As mentioned above, the chromatin is attached to the nuclear matrix. Each chromosome is located in its own chromosomal territory, which is defined by the nuclear matrix. The regions of the chromatin bound to the matrix are called scaffold/matrix attachment regions (S/MARs). In the human genome, the average distance between two S/MARs is about 50 to 200 kb meaning that the chromatin forms loops of that length (Eivazova *et al*. 2009, Linnemann *et al*. 2007, Heng *et al*. 2004). The S/MARs, however, are not distributed evenly, but for example, in telomeres at the ends of each chromosome, an S/MAR can be found at one kb intervals (Luderus *et al*. 1996). The loop formation has a significant role in gene regulation. One loop usually contains genes that are coordinately regulated and thus might share the regulatory regions (Fig. 2). S/MAR forms a physical and regulatory boundary between two loops meaning that it could function as a kind of insulator. The best known MAR binding protein is special AT-rich sequence-binding protein 1 (SATB1), which acts as a link between the chromatin and the nuclear matrix (Galande *et al*. 2007).

Interestingly, SATB1 is also known to interact with PML-bodies and to regulate transcription (Kumar *et al*. 2007). In addition, other factors can form chromatin loops and function as insulators, but these are not necessarily bound to the S/MAR regions, but to specific binding sites. The master insulator protein is CCCTC-binding factor (CTCF), which has several other functions as well, such as the regulatory function in gene expression by mediating long-range chromatin interactions (Phillips and Corces 2009, Nunez *et al*. 2009). The relationship between S/MARs and other insulators is poorly understood. However, Antes *et al*. (2001) proposed that S/MARs function as structural boundaries and CTCF-binding sites as functional boundaries. Nevertheless, the exact relationship between these different regions remains unclear. Even though the role of the loops is to isolate the different gene regions from each other, the loop structures are dynamic rather than fixed (Galande *et al*. 2007). For example, induction of gene expression can require remodeling of the loop. Chromosomal regions containing active genes can even escape from the chromosomal territory of the corresponding chromosome. This loop remodeling is believed to be an important event in transcriptional regulation (Fig. 2) (Fraser and Bickmore 2007). Perturbations in the nuclear organization found in several diseases, especially in cancers, are further confirming the importance of appropriate internal organization of the nucleus for proper cellular function (Elcock and Bridger 2008). In conclusion, the nuclear matrix does not only attach the nuclear components together, but it is a dynamic structure that has a remarkable role in nuclear functions.

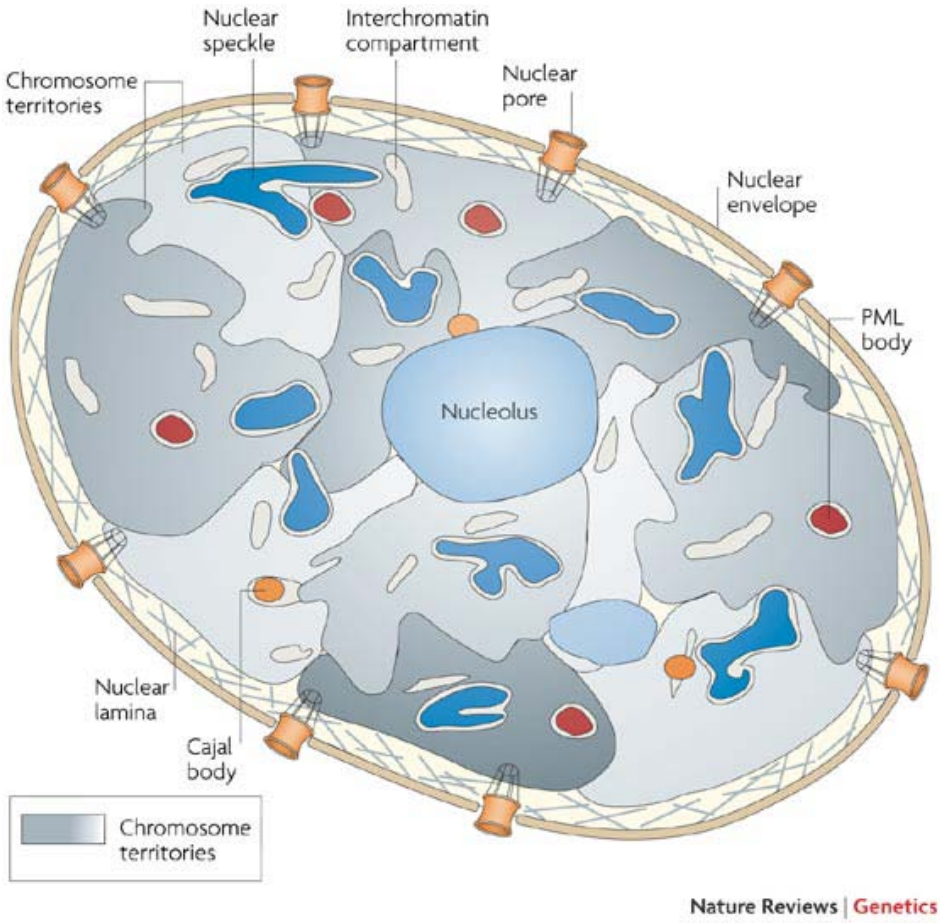

Figure 1. The special structures of the nucleus. (Reprinted from Lanctôt *et al*. 2007 with kind permission of Nature Publishing Group.)

2.1.2 Chromatin organization and structure

As discussed above, the chromosomes are located in their own nuclear territory and are attached to the nuclear matrix, but how are the genes localized in the territories? One general concept is that the active, gene-rich chromatin (euchromatin) is located toward the nuclear center, whereas the inactive and gene-poor chromatin, such as heterochromatin, is mostly located near the nuclear lamina and is tightly bound to the nuclear matrix (Kumaran *et al*. 2008). However, that is not an absolute rule as Finlan *et al*. (2008) noted that the location of a gene at the nuclear periphery is not incompatible with active transcription. Moreover, Gilbert *et al*. (2004) reported that some active genes can be found within the large heterochromatin fibers and conversely, inactive genes in the euchromatin fibers, suggesting that the predominant chromatin

conformation of the fiber does not, however, directly define the activity of all genes within the fiber. In the chromosomal territory, genes are also organized in a nonrandom fashion. Active genes seem to locate close to the boundary of the territory while inactive genes are located in interior regions of the territory (Cremer *et al*. 2001). However, also this rule is something of a generalization, since some activated genes are still located in the center of the territory (Mahy *et al*. 2002). Many gene-rich, constantly active gene clusters, such as the major histocompatibility complex, are located in the chromatin loops, which have escaped from their chromosomal territories to the interchromosomal space (Volpi *et al*. 2000). Enhancer elements, such as β-globin locus control region, can promote the escape from the territory (Noodermeer *et al*. 2008).

The traditional thinking has been that TFs are attracted to the chromatin during transcription. However, a novel concept suggests that TF complexes called transcription factories, rather than activated genes, are stationary structures that recruit transcribable chromatin (Fig. 2). These factories are located in the boundary of the territory and in the interchromosomal space and are bound to the nuclear matrix. These factories are rich in RNAP and certain TFs and are thus regulating a cluster of genes that are controlled by the same stimuli. The driving force that actually moves the chromatin fibers towards the factories remains elusive although some explanations have been proposed. For example, it has been postulated that RNAP itself would be responsible of chromatin retraction (Schneider and Grosschedl 2007). Interestingly, the genes regulated by a given transcription factory do not have to be located in the same chromosome. Thus, a certain TF binding site from one chromosome can also regulate genes which are located in different chromosomes. This type of interchromosomal regulation is called *in trans* regulation, while intrachromosomal regulation is called *in cis* regulation. In addition to transcription, also other events, such as DNA replication, occur similarly by specific, fixed protein factories through which the chromatin fiber is retracted (Göndör and Ohlsson 2009).

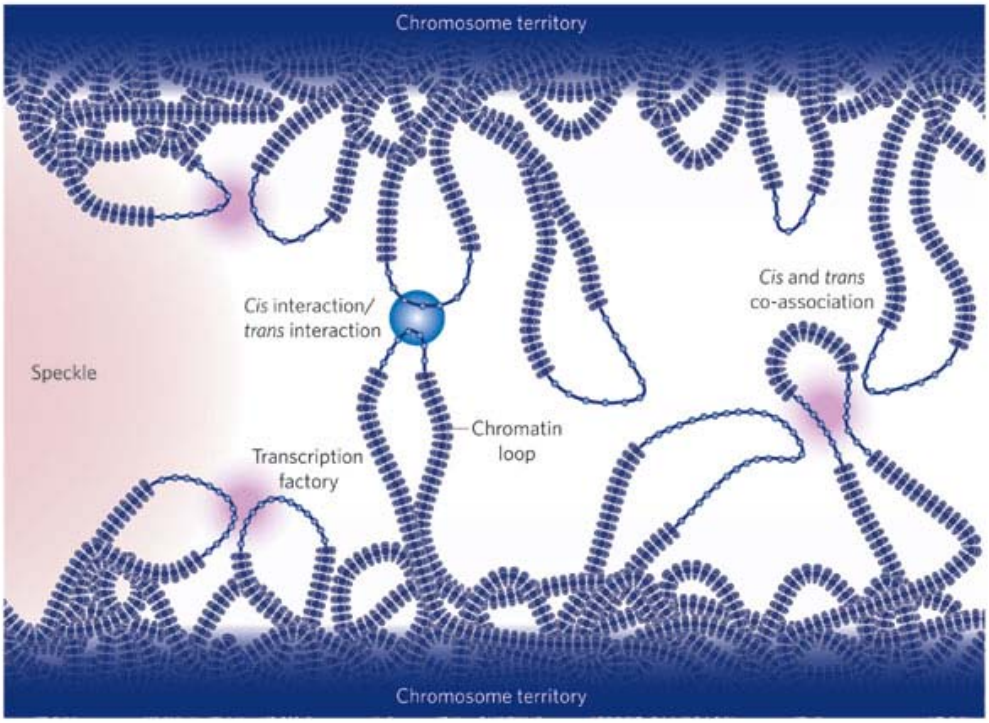

Figure 2. Chromatin looping and transcription factories. (Reprinted from Fraser and Bickmore 2007 with kind permission of Nature Publishing Group.)

Two special structures can be found in the chromosomes. The telomeres are regions/structures found at both ends of each chromosome. They are formed by several kb of repetitive sequence TTAGGG and specific proteins associated with this sequence. The last few hundred bases of the telomeres consist of single stranded DNA, which forms a structure called the t-loop. The role of the telomeres is to protect the chromosomal ends from degradation. One interesting feature of the telomeres is that they become shortened in every mitosis cycle. Ultimately, they have completely disappeared, which prevents further cell divisions. Certain types of cells, such as stem cells and cancer cells, express the enzyme called telomerase, which extends the telomeres and thus enables unlimited number of cell divisions (Artandi and DePinho 2010). The second special structure is the centromere found at the center of each chromosome. It is a region in the chromosome on which a complex directing the chromosome segregation is assembled during cell division (Morris and Moazed 2007). The complex is called the kinetochore and it connects the chromosome and the spindle microtubules (Santaguida and Musacchio 2009).

The DNA sequence of the centromere is not conserved, but it contains hierarchical arrays of simple sequence, such as the 171-bp repeats of alphoid DNA in mammalian cells and the chromatin at the centromere region is epigenetically modified, which causes the recruitment and assembly of the kinetochore proteins (Bloom and Joglekar 2010).

The total length of DNA molecules in a single human cell is about two meters. One could ask the question, how does something of that length fit into the spherical structure whose diameter is in micrometer scale? The answer is efficient packing. At the first level of packing, 146 bp of the negatively charged DNA is wrapped 1.65 turns around the octameric positively charged globular protein complex called the nucleosome. The core of the nucleosome consists of two copies of each of the histone proteins called H2A, H2B, H3, and H4. Each histone consists of a globular part and N- and C-terminal tails, which are often subjected to post-translational modifications (Luger *et al*. 1997). The average density of the nucleosomes is one in every 200 bp meaning that there is nucleosome free linker DNA between two nucleosomes which has a length of around 60 bp. This level of packing produces a chromatin fiber whose diameter is around 10 nm. This type of fiber is usually called beads-on-a-string or euchromatin and it has been usually perceived as transcriptionally active chromatin. At the next level of the packing, linker histones, such as histone 1 (H1), bind to the adjacent nucleosomes bringing the nucleosomes nearer to each other. This process produces a fiber of diameter of 30 nm. The chromatin is then further condensed to produce finally over 10,000-fold compaction in comparison to naked DNA. This type of chromatin is called heterochromatin and it is usually transcriptionally inactive (Horn and Peterson 2002). In addition to histones, the chromatin contains a huge number of nonhistone proteins, which are responsible for transcriptional regulation or performance. These proteins are called TFs. In transcriptional regulation, the structure of the chromatin is usually modified by specific TFs. This type of regulation can also be epigenetic and will be discussed below.

2.1.3 Structure of human genome

In addition to the 20,000 to 25,000 genes coding for proteins, the human genome consists of several other types of DNA sequences. The content of the genome can be divided into nonrepetitive and repetitive sequences. For example, exons are usually located within nonrepetitive sequences, meaning that there is only one copy of the sequence in the haploid genome. In human, more than half of the genome consists of repetitive sequence meaning that the

sequence is found more than once in the genome. The nonrepetitive portion of the genome consists of exons (~1%), introns (~24%) and other intergenic DNA (~22%) for example containing regulatory regions. According to these percentages, the genes cover only 25% of the genome. The average human gene contains 7 exons, which are 145 bp in length and the total length of the gene is 27 kb. The repetitive portion of the genome contains transposons (45%), large duplications (5%), simple repeats (3%), and pseudogenes (0.1%, ~3000 copies). The main origin of the repetitive sequences is probably retroviruses, but their exact role is not well known. Previously it was thought that the repetitive sequence is only "junk DNA", but the present understanding is that it has some functional role for example in gene regulation (International Human Genome Sequencing Consortium 2001). Even though the classical genes cover only one quarter of human genome, also other regions of the genome are transcribed. These ncRNAs could function as epigenetic regulators of transcription and will be discussed more extensively in the next chapter (The ENCODE Project Consortium 2007).

2.2 REGULATION OF TRANSCRIPTION

Gene expression consists of multiple sequential and simultaneous processes. First, a gene is transcribed to heterogeneous nuclear RNA (hnRNA), then spliced and modified to messenger RNA (mRNA) and finally translated into the amino acid sequence of a protein in cytosol or rough endoplasmatic reticulum by ribosomes (O'Malley *et al*. 1977). All of these steps can be regulated, but in this chapter the focus will be on regulation of transcription. Transcription is regulated by two main mechanisms: by TF binding to specific DNA elements and by modification of the chromatin structure. These two mechanisms are strictly linked together and actual order of which one is "egg" and which one is "chicken" is not clearly known. Some of the factors needed for regulation of transcription are heritable and they can be inherited either by genetically or epigenetically. The genetic heritability can be defined as inherited DNA sequences of the regulatory regions and the epigenetic heritability as heritable phenotype changes that do not involve alterations in DNA sequence. For example, the epigenetic changes involve DNA methylation, histone modifications and ncRNA-based silencing (Bernstein *et al*. 2007). In the subsequent section, the TFs and their DNA-binding sites will be discussed together as well as the chromatin modifications and other epigenetic

factors. Since the role of the chromatin organization and nuclear matrix in gene regulation has been described above, this will not be discussed further.

2.2.1 Transcription factors and DNA-binding sites

TFs are proteins involved in the initiation of transcription. In humans, over 2,000 genes code for TFs and they are generally divided into two categories: general TFs and other TFs. RNAPs are not counted as TFs, but they are DNA-dependent RNAP enzymes activated by TFs and are thus discussed in this chapter. Many TFs recognize a specific DNA binding sequence; these factors are called sequence-specific DNA-binding TFs (Pan *et al*. 2010). Some factors do not bind directly to DNA, but are bound to other TFs. These are usually called cofactors or coregulators, or corepressors or coactivators. TFs have several functions, such as chromatin modification and RNAP activation. Moreover, some of them have enzymatic activity and some can function as receptors for internal or external signals (Brivanlou and Darnell 2002).

2.2.1.1 RNA polymerases

There are three RNAPs in eukaryotes: I, II, and III. They have typically ~12 subunits from which three are common for all RNAPs (Young 1991). The genes coding for these subunits are referred with letters and numbers to, for example like POLR1A, which stands for Polymerase (RNA) I polypeptide A. From the gene bank, subunits for RNAPI from A to E, for RNAPII from A to L, and for RNAPIII from A to H, can be found (www.ncbi.nml.nih.gov). Each RNAP has specific genes for which it transcribes. RNAPI transcribes genes coding for ribosomal RNAs 18S and 28S in nucleolus and has actually the most prominent RNA synthesis capability in terms of quantity. RNAPIII is responsible for transcribing genes coding for ribosomal 5S RNA, transfer RNAs, and other small RNAs in nucleoplasm. RNAPII is the most complex RNAP and is responsible for transcribing most of the hnRNAs, which are precursors for mRNAs coding for proteins (Archambault and Friesen 1993). The largest subunit of RNAPII (POLR2A) has a unique C-terminal domain (CTD), which has been linked to several functions of RNAPII, such as an interaction with DNA and histone displacement during elongation. The CTD consist of multiple repeats (~50) of heptameric amino acid sequence (YSPTSPS), whose serine and threonine residues can be phosphorylated. The phosphorylation of the CTD is the final activating signal for RNAPII to initiate and continue transcription. Phosphorylation at serine 5 is needed for transcription initiation, while phosphorylation at serine 2 is needed for RNAPII elongation

(Buratowski 2009). The role of the phosphorylation at serine 7 is not very well characterized, but it has been linked to small nuclear RNA gene expression (Chapman *et al*. 2007, Egloff *et al*. 2007).

2.2.1.2 General transcription factors and the core promoter

General TFs are factors/protein complexes that are needed for RNAP recruitment to the core promoter of a gene and for transcription initiation by RNAP. In this section, only the factors involved in transcription of RNAPII regulated genes will be discussed. The core promoter is defined as a region having all the binding sites needed for RNAPII to bind and function (Fig. 3). The size of the core promoter is approximately 100 bp and the transcription start site (TSS) lies at the center of the core. General TFs and RNAPII together constitute the basal transcription apparatus (BTA) needed for every promoter to initiate transcription. In transcriptional activation, distal DNA-bound TFs interact with and activate the BTA leading to initiation of transcription, so the actual regulation of the transcription is mainly done by other TFs rather than by general TFs. The core promoter contains few conserved elements from which the one at the center is the initiator element (Inr), whose DNA sequence can be described as YYANWYY, where Y is either T or C, N is any nucleotide, and W is either A or T. The adenosine of the sequence is the actual TSS (+1). Approximately 10%–15% of all promoters contain sequence TATAWAAR, where R is either A or G, at ~25 bp upstream of the TSS and which is called the TATA-box. Usually the promoters, which do not have a TATA-box, contain a downstream promoter element (DPE), whose consensus sequence is RGWYVT, where V is either A, C, or G, and this being located at ~30 bp downstream from the TSS. Other less studied elements have also been found, such as motif ten element (MTE) just upstream from DPE. The presence of these elements defines the type of transcription initiation. Two-thirds of human promoters have a characteristic of disperse initiation and the rest display the characteristics of focused initiation. In disperse initiation, the transcription starts from many weak TSSs, while in focused initiation, the TSS is strictly defined. Promoters having a TATA-box tend to be of the focused type as well as highly regulated genes, while constitutively expressed genes are typically of the disperse type (Juven-Gershon *et al*. 2010).

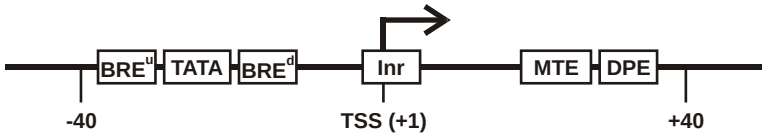

Figure 3. The structure of core promoter. Abbreviations are found in the abbreviation list.

Before polymerase can bind to a promoter and start transcription, a protein complex called positioning factor needs to bind first on the core promoter. In humans, the factor is called TFIID (II subscript stands for RNAPII), which is composed of TATA-binding protein (TBP) and several (up to 14) TBP-associated factors (TAFs). Although the name of the TBP refers to its ability to bind TATA-box, TBP or related factors (TRFs) are still needed for promoters that are TATA-box-less. The composition of TAFs can vary for example, depending on cell type and thus some variant TFIIDs can recognize alternative promoters of genes. TBP is essential, especially for RNAPII, to locate promoters to bind, because RNAPII does not have any intrinsic promoter recognizing property. In TATA-box-less promoters, TAFs recognize elements other than a TATA-box, such as Inr or DPE, for positioning the promoters. Next, TFIIA joins the complex and activates TBP's DNA-binding ability by removing TAF1 from DNA-binding surface of TBP (Cler *et al*. 2009). Then TFIIB binds both upstream and downstream (elements called BRE^u and BRE^d, Fig. 3) to TBP determining the polarity of the promoter, *i.e*. which strand is template and which way RNAPII faces. TFIIB actually forms the surface that is recognized by RNAPII (Deng and Roberts 2007). For example, TFIIF is responsible for recruiting RNAPII to the assembling complex, since it binds tightly to RNAPII. It has also helicase activity and it is actually responsible for DNA melting in transcription initiation together with TFIIE and TFIIH (Eichner *et al*. 2010). Next these latter two factors join the complex (Tanaka *et al*. 2009). TFIIH has multiple enzymatic activities *e.g*. it can achieve the phosphorylation of serine 5 and 7 of CTD. The serine 5 phosphorylation is needed for the release of RNAPII from the core promoter (Achtar *et al*. 2009). The timing of the phosphorylation on serine 7 by CDK7 (a subunit of TFIIH) is not known, but it has been suggested that this occurs before transcription initiation (Boeing *et al*. 2010). The initial transcription of many genes is stopped rapidly and the RNA formed is degraded. Subsequently, a complex called positive transcription elongation factor b (P-TEFb) is recruited which can phosphorylate serine 2 of CTD (Lenasi and Barboric 2010). Finally, the actual transcription can start and most of the initiation complex factors are

dissociated from the promoter, except that RNAPII together with the factors needed for elongation, *i.e.* TFIIS complex (Kim *et al*. 2007).

2.2.1.3 Coregulators

As mentioned above, other TFs rather than general TFs regulate the rate of transcription. The effects of TFs on transcription need to be mediated in someway to the BTA. In some cases, DNA-binding TFs can interact directly with the apparatus, but in most cases there are other factors, called coregulators, between them. A coregulator can function either as coactivator or corepressor depending on its effect on transcription. Coregulators are usually recruited to the chromatin by TFs or by other coregulators. However, all of them do not interact with the BTA, but instead modify the chromatin structure locally. Coregulators can be divided into three categories: covalent modifiers of the chromatin, chromatin remodeling complexes, and mediator complexes. The best well characterized coregulators are recruited by NRs and thus the focus of this section will be on those factors (Rosenfeld *et al*. 2006, O'Malley 2007).

Covalent modifiers of the chromatin possess an enzymatic activity to either add or remove small molecules or proteins to or from the bases of the DNA or amino acid residues of the histone proteins. In some cases, a coregulator neither adds nor removes molecules, but instead changes (isomerizes) the structure of its substituent. The inserted or removed compound is either an acetyl group (ac), a methyl group (me), ubiquitin (ub), a small ubiquitin-like modifier (SUMO), ADP-ribose (ADPr), or phosphate residue (p). The DNA can be modified only by methyl group, but the histones can be modified by all the substituents (Kouzarides 2007a). In addition to TFs, these cofactors can be recruited at the chromatin by other coregulators, such as p160-family coactivators that have no or at best only modest, intrinsic histone acetyltransferase activity, or by some corepressors, such as nuclear receptor corepressor 1 and 2 (NCoR1 and NCoR2) that function as linkers between TF and repressive chromatin modifying enzyme (Privalsky 2004). The specific effects of the modifications on transcription as well as the specific coregulators will be discussed in chapter 2.2.2.

Chromatin remodeling complexes are usually ATP-dependent enzymes that modify the structure, position, or existence of a certain nucleosome. These complexes have an important role in regulation of transcription; especially in TF binding to its binding element on DNA. Nucleosomes normally inhibit the binding, but dissociation (eviction) or moving (sliding) of the nucleosome from

its initial position uncovers the binding site and thus enables the binding of a TF (Becker and Hörz 2002, Workman 2006, Gutiérrez *et al*. 2007). The complexes are classified into four classes depending on the central ATPase. The central ATPase of the switch mating type/sucrose non-fermenting (SWI/SNF) complex is either brahma (BRM) or brahma-related gene 1 (BRG1), that of the imitation of SWI (ISWI) complex is ISWI-ATPase, that of the mi-2/nucleosome remodeling deacetylase (Mi-2/NuRD) complex is chromodomain-helicase-DNA binding protein (CHD), and that of the INO80 complex is INO80 (Hogan and Varga-Weisz 2007). In addition to the transcription, the above complexes have specific roles also in other functions in the nucleus. For example, ISWI plays a role in replication and INO80 in DNA repair and chromosome segregation (Farrants 2008, Hur *et al*. 2010). In contrast to the other complexes, the Mi-2/NuRD complex is involved in transcription repression and can be defined as corepressor complex, since it has also histone deacetylase activity (Gao *et al*. 2009). In addition, SWI/SNF has a crossactivity to other chromatin modifications, since it has a role in DNA demethylation or at least its loss causes DNA methylation (Banine *et al*. 2005).

As the name suggests, the mediator complex mediates the activation signal from a TF to the BTA. It is a large complex composed of ~20 proteins (called MEDs or TRAPs) and its total mass is over 1 MDa. It interacts directly with the RNAPII, especially with the hypophosphorylated CTD (Chadick and Asturias 2005). The mediator can recruit $TF_{II}H$, $TF_{II}E$, and $TF_{II}S$ to the core promoter by interacting directly with these factors. Due to recruitment of $TF_{II}H$, it enhances the phosphorylation of the CTD and thus activates the transcription initiation (Guglielmi *et al*. 2007, Esnault *et al*. 2008, Boeing *et al*. 2010). The phosphorylation of the CTD causes dissociation of the mediator, enabling the reinitiation of the transcription (Casamassimi and Napoli 2007). The other part of the complex is interacting with TFs or other coregulators and the mediator can even recruit them to the chromatin. Thus, the mediator is a key factor between TFs and the BTA in transcription initiation (Huang *et al*. 2003).

2.2.1.4 Sequence-specific DNA-binding transcription factors

Regulatory DNA can be defined as a DNA sequence involved in the regulation of transcription. These regions can be located either in the nonrepetitive or repetitive part of the genome, in intergenic or intragenic regions, in exons, introns or untranslated regions (UTRs) (Carroll *et al*. 2006, Wang *et al*. 2007, Bolton *et al*. 2007). These regions contain specific DNA sequences recognized by sequence-specific DNA-binding TFs. These factors, activators or repressors,

activate or repress, respectively, transcription of genes by binding to the specific binding sites. Depending on the TF, it can either interact with and activate directly the BTA or recruit coregulators which can activate or repress the BTA or modify chromatin. Nevertheless, most TFs can interact directly and *via* the coregulators with the BTA (Brivanlou and Darnell 2002). Some factors, such as CTCF, interact neither directly nor indirectly with the BTA, but are rather structural TFs. These factors regulate the chromatin bending or loop formation and thus enable the long range interactions between other TFs and the BTA (Phillips and Corces 2009). The binding site or a cluster of binding sites that recruits activating or repressing TFs is called either an enhancer or a silencer, respectively. The location of the enhancer or silencer of a certain gene can be from tens of nucleotides to several kbs in either direction from the TSS and the orientation of the element does not matter (Carroll *et al*. 2006, Wang *et al*. 2007, Bolton *et al*. 2007).

Sequence-specific DNA-binding TFs can be categorized into several classes depending on their properties or activating mechanisms. Brivanlou and Darnell (2002) have proposed the following categories for activating TFs: The main groups are constitutive and regulatory TFs. Then the regulatory TFs can be divided into cell-specific and signal-dependent, and then signal-dependent TFs further into steroid receptors (SR), TFs for internal signals, and cell surface receptor activated TFs, which can be further divided into two subcategories. The constitutive TFs, such as specificity protein 1 (SP1), are expressed constantly in all the cell types and they are responsible for the regulation of transcription of constitutively active genes, such as tubulin, and the basal activity of other genes. The cell-specific TFs, such as GATA-binding protein 1 (GATA1), are expressed in tissue-specific manner and are involved mostly in developmental processes. These factors are mainly regulated by their expression rather than external or internal activation signals. As the name suggests, the signal dependent TFs, such as AR, are activated by external or internal signals. The cell surface receptor dependent TFs are triggered by the activation cascade starting from ligand binding to its cell surface receptor. Internal signal-dependent TFs, such as p53, are activated by internal signals, such as DNA damage in the case of p53.

SRs is a subclass of endocrine receptors which is one of the subclasses of NRs. The superfamily of NRs consists of 48 members and it is the largest family of ligand activated transcription factors in humans. One common property to all NRs is the lipophilicity of their ligands. The other two subclasses of NR are orphans and adopted orphans. The orphans are NRs

whose natural ligand does not exist or at least has not been found. The adopted orphans are a group of orphans for which the natural ligand was found after the receptor had been cloned. The ligand specificity and binding affinity is very different between the endocrine and adopted orphan receptors. Endocrine receptors are very specific for their ligands, whereas most adopted orphans have large ligand binding pockets that decrease the ligand binding affinity and specificity (Chawla *et al*. 2001). In addition, a more sophisticated categorizing system, based on the sequence homology of the NRs, has been developed, where the receptors are divided into seven groups (0 to VI). In fact, the systematic names of the NR genes are based on this group numbering, for example *AR*'s systematic name is *NR3C4* indicating that *AR* belongs into class 3C of the NRs (Aranda and Pascual 2001).

2.2.1.5 Androgen and glucocorticoid nuclear receptors

The SR subfamily includes AR, glucocorticoid receptor (GR), progesterone receptor (PR), mineralocorticoid receptor (MR) and two estrogen receptors (ERα, ERβ), from which all but the ERs share the same DNA-binding sites, albeit specific sites are also found (Huang *et al*. 2010). The general protein structure of SRs is common for all and several structurally distinct domains have been identified: the N-terminal domain (NTD), the DNA-binding domain (DBD), the ligand-binding domain (LBD), and a hinge region between DBD and LBD, which contains the nuclear localization signal (NLS). NTD is not well conserved in the SR family, for example the sequence similarity between AR-NTD and PR-NTD is only 20%. Conversely, the conservation of DBD is very high (~80%) between all SRs except for ER-DBD, whose sequence is only 59% similar to that of AR. The differences in ER-DBD compared to other SRs can thus explain the difference in DNA sequence recognition (Gao *et al*. 2005, Huang *et al*. 2010). The transactivation function of SRs is mediated mainly by two functional regions: activation function 1 and 2 (AF-1, AF-2), which are located in the NTD and LBD, respectively. There is a general mechanism to explain how SRs function for all SR's. First, a hydrophobic ligand diffuses through the plasma membrane and binds to the SR monomer in the cytosol. The ligand binding causes a conformational change in the receptor's LBD, leading to dissociation of the associated chaperone complex, and phosphorylation, dimerization and nuclear translocation of the receptor. In the nucleus, the receptor dimer binds to its response elements, recruits coregulators and then it can activate transcription (Biddie *et al*. 2010).

Androgens are steroid hormones that function *via* AR and are responsible for the development of male sexual characteristics during embryogenesis and puberty as well as maintaining them after puberty. Testosterone and DHT are the two most potent natural androgens, with testosterone being produced mainly by Leydig cells of the testes and DHT locally in target tissues by 5α-reductase enzyme from testosterone (Gao *et al*. 2005). In androgen free conditions, AR is inactive and is incorporated into the chaperone/immunophilin complex in the cytosol. The complex consists of heat shock protein 90 (HSP90) as the main chaperone and at least two co-chaperones: p23 and either immunophilin protein (a protein that binds immunosuppressive drugs) FKBP51, FKBP52, or Cyp40, or non-immunophilin protein PP5 (Pratt and Toft 1997, Heitzer *et al*. 2007).

In response to androgen exposure, by the mechanisms discussed above, the receptor homodimer is translocated to the nucleus where it binds to ARE that is a prerequisite for AR-mediated transactivation, but not necessarily for AR-mediated transrepression (Gao *et al*. 2005). The mechanisms of AR-mediated gene repression are less studied than those of gene activation. However, it appears that the transcriptional repression does not require interaction of the receptor with specific DNA elements but interference with other sequence-specific TFs. For example, AR can repress the activity of activator protein 1 (AP-1) by interfering with its DNA binding (Kallio *et al*. 1995). Moreover, AR can form a complex with RelA (an activating-subunit of nuclear factor κB, NFκB), which leads to their mutual inhibition (Palvimo *et al*. 1996). AR and androgens can also have non-genomic actions, for example AR can activate mitogen-activated protein kinases (MAPKs) by transcription-independent mechanisms and DHT can bind to membrane-associated AR, which leads to rapid increase of intracellular calcium concentration (Foradori *et al*. 2008). The consensus sequence of the ARE is AGAACAnnnTGTTCT that is a type of inverted repeat separated by three nucleotides (IR3)-element. If that is also the consensus sequence of GR, PR, and MR binding elements, how can there be genes that are activated only by androgens? Claessens *et al*. (2001) proposed that in addition to palindromic AREs there are also so-called AR specific binding elements that are a type of direct repeat separated by three nucleotides (DR3). However, the AR specificity is not absolute, since it was recently shown that also PR can bind to DR3-type AREs (Denayer *et al*. 2010). Moreover, recent studies have suggested that only 10% of all the human AREs are canonical and the rest are more or less noncanonical *i.e*. non-IR3-type (Wang *et al*. 2007, Verrijdt *et al*. 2006, Bolton *et al*. 2007). Irrespective of the type of the element,

AR recognizes and binds to the ARE through two tandem zinc fingers (that contain regions called P-box and D-box, respectively) formed by eight cysteine residues in DBD and by two central Zn^{2+}. The first zinc finger is responsible for specific DNA-recognition, whereas the latter one is needed in AR homodimerization. The orientation of the AR monomers depends on the type of ARE: IR3 element prefers a head-to-head orientation, whereas DR3 prefers a head-to-tail orientation (Verrijdt *et al*. 2003, Gelmann 2002). The binding affinity of AR to an ARE certainly depends on the DNA sequence of the half sites, but recent studies have indicated that in addition to ARE itself, proximal surrounding binding sites for cell-specific TFs play an important role in AR binding efficiency. These factors include at least forkhead box A1 (FOXA1), GATA2 and OCT1 (Wang *et al*. 2007, Gao *et al*. 2003, Jia *et al*. 2008).

After DNA binding, AR interacts directly with the components of BTA and recruits coregulators, such as p160-family coactivators (steroid receptor coactivator 1, 2, or 3; SRC-1,-2, or -3) that facilitate the recruitment of histone modification enzymes, such as p300, cAMP response element-binding protein (CREB)-binding protein (CBP), p300/CBP-associated factor (P/CAF), coactivator-associated arginine methyltransferase 1 (CARM1), chromatin remodeling complexes, such as SWI/SNF, the mediator complex, and many other proteins (over 150 coregulators are known for AR) that are involved in a wide variety of functions, such as in proteasome-mediated protein degradation and SUMO modifications (Fig. 4) (Heemers and Tindall 2007). In contrast to the other NRs, ligand-independent AF-1 rather than ligand-dependent AF-2 plays a major role in AR mediated transactivation and thus also the coactivator recruitment differs from the others. AR-LBD interacts poorly with the LXXLL-motif found in many coactivators; instead it interacts with FXXLF-motifs found for example in its NTD and in some AR specific coactivators, such as ARA70. The coactivators that have the LXXLL-motif interact with the NTD and DBD instead of LBD, and in the situations when the coactivator is overexpressed, it can interact also with the LBD (He *et al*. 2002).

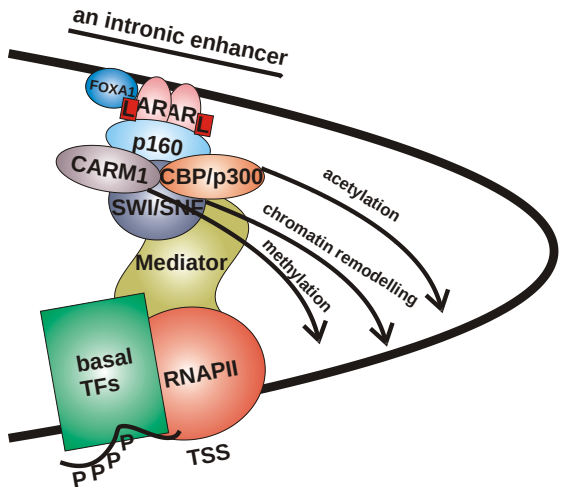

Figure 4. Simplified model of androgen receptor-mediated transcription activation. L, ligand; P, phosphate residue; other abbreviations are found in the abbreviation list.

Glucocorticoids are steroid hormones that can function *via* both GR and MR due to the high similarity of the receptors' LBD (Sorrells and Sapolsky 2007). The most potent natural glucocorticoid is cortisol that is produced by adrenal cortex and its production is regulated by hypothalamus and hypophysis hormones. Glucocorticoids regulate many genes involved in gluconeogenesis as well as in lipid and amino acid metabolism (Heitzer *et al*. 2007). They also negatively regulate immunoreactions and are thus widely used as immunosuppressive drugs, for example in astma (De Bosscher and Haegeman 2009). In addition, glucocorticoids are mediating stress reactions in the body and thus they can also regulate mental functions of the brain (Spijker and van Rossum 2009). The mechanisms of GR action are very similar to that of AR and are thus not discussed further.

2.2.2 Histone and DNA modifications and epigenetics

Histones and DNA are naturally strongly bound together due to their opposite net charge: DNA is negatively whereas histones are positively charged. However, a too strong interaction between these two partners leads to chromatin condensation that inhibits TF binding to their binding elements on DNA and thus prevents transcription. Hence, in order to obtain the correct transcription level of a gene, the chromatin has to be modified so that it will decondensate and enable TF binding. These modifications include chromatin remodeling as well as covalent modifications of histones, especially the histone tails, and DNA (Kouzarides 2007a, Li *et al*. 2007a). Many of these modifications

are linked together, meaning that one modification may be a prerequisite for the next. Together these modifications are suggested to form an epigenetic code, *i.e*. a pattern of certain modifications that poses specific effects on transcription. For example, the epigenetic marks can form a specific binding surface for TF or coregulator (Fuchs *et al*. 2006, Turner 2007). Since chromatin remodeling is discussed above, the focus will be on the covalent modifications. The hereditary nature of these modifications will be also discussed as will be the ncRNAs as regulators of transcription.

2.2.2.1 DNA methylation

DNA methylation is a key chromatin modification linked to gene regulation, imprinting, heterochromatin assembly and X-chromosome inactivation. In eukaryotes, the DNA methylation occurs only on carbon five of the cytosine base that, at least in mouse brain and embryonic cells, can be further converted to 5-hydroxymethyl cytosine (Klose and Bird 2006, Tahiliani *et al*. 2009). In most cases, the DNA methylation is linked to transcriptional repression (Klose and Bird 2006). The methylation occurs mostly in CpG islands found in promoter region of most genes, but there is also some non-CpG methylation. The non-CpG methylation is common only in undifferentiated cells, such as stem cells and other embryonic cells, and thus seems to be linked to pluripotency and differentiation together with 5-hydroxymethyl cytosine (Lister *et al*. 2009, Tahiliani *et al*. 2009). This modification is strongly linked to other chromatin modifications such as histone methylation. The repressive property of the modification is, at least in part, mediated by recruitment of histone modification enzymes by methyl-CpG-binding proteins (Klose and Bird 2006). Recently, dynamic and cyclical methylation of CpG islands of transcriptionally active genes have been reported, suggesting that the DNA methylation can be relatively dynamic and involved also in active transcription (Métivier *et al*. 2008, Kangaspeska *et al*. 2008). The process, where one allele of the same gene is closed during embryogenesis, is called imprinting. DNA methylation is the key modification that mediates this process (Weaver *et al*. 2009). Interestingly, also some RNAs can be methylated at the same carbon, but the functional importance of that modification remains elusive (Motorin *et al*. 2010).

The DNA is methylated by the enzymes called DNA methyltransferases (DNMTs). They can be classified into two main groups: *de novo* (DNMT3a and DNMT3b) and maintenance DNMTs (DNMT1). The substrate of the *de novo* DNMTs is the CpG islands that are not methylated on either strand, whereas

maintenance DNMTs copy pre-existing methylation patterns onto the new DNA strand during DNA replication. The enzymes that can reverse the methylation, *i.e*. DNA demethylases, have not been found, albeit some candidates have been proposed, such as methyl-CpG binding domain protein 2 (MBD2) (Patra *et al*. 2008). However, the demethylation does occur, and a few indirect mechanisms have been postulated, such as DNA repair, where the whole methyl-cytosine is removed and replaced with non-methyl cytosine (Gehring *et al*. 2009). A massive active demethylation occurs in early embryogenesis, where paternal DNA is demethylated before the first replication, suggesting that there must be also direct mechanisms for demethylation of DNA (Abdalla *et al*. 2009). The 5-methylcytosine is converted to 5-hydroxymethylcytosine by enzymes called TET1, -2 and -3 (Tahiliani *et al*. 2009).

2.2.2.2 Histone acetylation

Histones can be acetylated at their lysine residues in a very dynamic fashion during transcription. One of the main mechanisms by which acetylation can activate transcription is to decrease of the positive charge of the histones and thus histone-DNA interactions. The other mechanism for transcriptional activation is that the acetylated lysines form a binding surface for many coactivator proteins, such as that have a specific chromatin-binding region called bromodomain. H2A can be acetylated at lysine 5 (K5) and K9; H2B at K5, K12, K15, K20, and K120; H3 at K4, K9, K14, K18, K23, K36, and K56; and H4 at K5, K8, K12, K16, and K91 (Wang *et al*. 2008b, Verdone *et al*. 2006, Kouzarides 2007b). In contrast to the other acetylated lysines, acetylated H3K4 (H3K4ac) has been linked to heterochromatin assembly and thus to transcriptional repression. However, this modification differs from the other lysine acetylations in that it is linked to replication rather than to the dynamic transcription regulation in interphase cells (Xhemalce and Kouzarides 2010). H3K56 is not located in the histone tail, but in the globular part of the H3 and it has been reported that acetylation of the residue does not have any effect on histone-DNA interaction, but instead it influences chromatin remodeling by SWI/SNF complexes (Neumann *et al*. 2009). The histone acetylation is not distributed evenly in the genome, but it is enriched around the TSSs of the active or poised, *i.e*. epigenetically preprogrammed, genes, and thus it can be used as an epigenetic marker for locating active genes (Fig. 5) (Li *et al*. 2007a).

The histone acetylation occurs by the function of enzymes called histone acetyltransferases (HATs) that are usually found in coactivator complexes. In

humans, at least 17 HATs have been found of which CBP, p300, and P/CAF are best characterized (Allis *et al*. 2007). The acetylation is reversed by histone deacetylases (HDACs) that can be divided into four groups (I-IV): Class I consists of HDACs 1, 2, 3, and 8; Class II of HDACs 4, 5, 6, 7, 9, and 10; Class III of silent information regulators 1-7 (sirtuin 1-7 or SIRT1-7); and Class IV of HDAC11 (de Ruijter *et al*. 2003, Brandl *et al*. 2009). HDACs are usually found in corepressor complexes such as in Mi2/NuRD complex and mammalian switch-independent 3 protein (mSin3) complex (Privalsky 2004).

2.2.2.3 Histone methylation

Lysine and arginine (R) residues of the histone tails or globular parts of H2B, H3, and H4 can be either mono- (K and R), di- (K and R (symmetric or asymmetric)), or trimethylated (only K). The specific residues that can be modified are H2B's K5, H3's R2, K4, R8, K9, R17, R26, K27, K36, and K79, and H4's R3 and K20 (Barski *et al*. 2007, Kouzarides 2007b). The effect of a given modification on transcription depends strongly on the residue being modified, the number of methyl groups added, and the location of the modification with regard to the gene. The methylated residues that are linked to transcription activation are H3K4, H3K9, H3K20, H3K27, H3K36, H3K79, H3R2, H3R17, H3R26, and H4R3, and those linked to repression are H3K9, H3K27, H3K36, H3K79, and H4K20 (Li *et al*. 2007a, Barski *et al*. 2007). Most of the studied modifications have been methylations of H3K4, H3K9, H3K27, and H3K36, and thus their role in transcription is best known. Trimethylation of H3K4 and, to some extent, its dimethylation are enriched around TSSs of the active genes together with acetylation of H3 and H4, and thus these modifications are very good markers for an active gene (Fig. 5) (Santos-Rosa *et al*. 2002, Barski *et al*. 2007, Wang *et al*. 2008b, Guenther *et al*. 2007). In contrast, monomethylation of H3K4 is enriched mainly in the enhancer regions (Heintzman *et al*. 2007). Interestingly, monomethylation of H3K9 is enriched also around TSSs of active genes, whereas di- and trimethylation of H3K9 are linked to transcriptional repression or heterochromatin together with trimethylation of H3K27. In heterochromatin formation, heterochromatin protein 1 (HP1) recognizes and binds specifically to di- and trimethylated H3K9. The modification is also linked to DNA methylation, since a high level of H3K9 trimethylation and the absence of H3K4 trimethylation correlate with high levels of DNA methylation (Cheng and Blumenthal 2010). The location of H3K9 methylation relative to a TSS determines also the role of the modification; methylation of upstream region is linked to transcriptional repression and that of downstream region to

activation (Li *et al*. 2007a). Trimethylation of H3K36 can be enriched in the body of active genes being even more enriched in exons and 3′ end of the genes and thus it can be also involved in RNA splicing. The modification is needed for RNAP elongation and it inhibits aberrant transcription initiation by recruiting the Rpd3S HDAC complex that deacetylases histones. Rpd3S recognize the modification by its chromodomain, which is a common protein region for recognizing methylated histones (Kolasinska-Zwierz *et al*. 2009, Li *et al*. 2007b). Not only is histone methylation very complicated *per se*, but the fact that it is also linked to other modifications of the histones, makes it and the whole histone code even more complex to understand. For example phosphorylation of H3 threonine 6 (T6) controls demethylation of H3K4 and monoubiquitination of H2BK120 is needed for H3K4 and H3K79 methylation (Metzger *et al*. 2010, Lee *et al*. 2007). Moreover, histone methylation can even impose the ligand dependency in NR-mediated gene activation (Garcia-Bassets *et al*. 2007).

Enzymes that methylate histones are called histone methyltransferases (HMTs). In humans, 48 "suppressor of variegation, enhancer of zeste, trithorax" (SET)-domain containing HMTs and DOT1L that does not contain the SET-domain, have been found. HMTs are specified to methylate certain residues. For example, ASH1L catalyzes only the reaction, where H3K4me2 is transformed to H3K4me3. HMTs can form complexes either with other coactivators or corepressors and thus the role of methylation in transcription is very variable. HTMs are also linked to many diseases. For example, overexpression of AR coactivator CARM1, which only methylates arginine residues, is linked to PC development and a polycomb group (PcG) protein EZH2, that maintains undifferentiated state of embryonic stem cells *via* catalyzing trimethylation of H3K27, is linked to many cancers, including lymphoma, melanoma, bladder, breast, colorectal, gastric, and PCs (Albert and Helin 2010). The histone methylation can be reversed indirectly by active histone exchange, by proteolytic removal of the histone amino-termini or by conversion of methylarginine to citrulline by peptidylarginine deiminase, or directly by enzymes called histone demethylases (HDMs). HDMs can be classified into two classes depending on their structure and activation mechanisms: lysine-specific demethylases (LSDs) and jumonji-C-terminal (JmjC)-domain containing demethylases. The main functional difference between these two groups is that only JmjC-domain containing HDMs can demethylate trimethylated lysine residues due to the distinct chemistry of the enzymatic reaction (Mosammaparast and Shi 2010). LSD1 was the first

demethylase identified and it was not found until six years ago (from the time of writing this thesis) and it specifically demethylates H3K4me2 and H3K4me1, but not H3K4me3 (Shi *et al*. 2004). JmjC-domain containing retinoblastoma binding protein 2 (RBP2) can demethylate H3K4me3 (Klose *et al*. 2007). However, when LSD1 and JMJD2C interact with holo-AR, they promote AR-mediated transcription by demethylating di- and trimethylated H3K9, respectively (Wissmann *et al*. 2007). Interestingly, HDAC inhibitors, a putative class of anticancer agents, reduce also H3K4 demethylation by LSD1, suggesting that HDACs and HDMs may have some functional interplay (Lee *et al*. 2006). This is further confirmation that epigenetic markers and the enzymes involved in the marking can function together in an organized fashion rather than independently of each other.

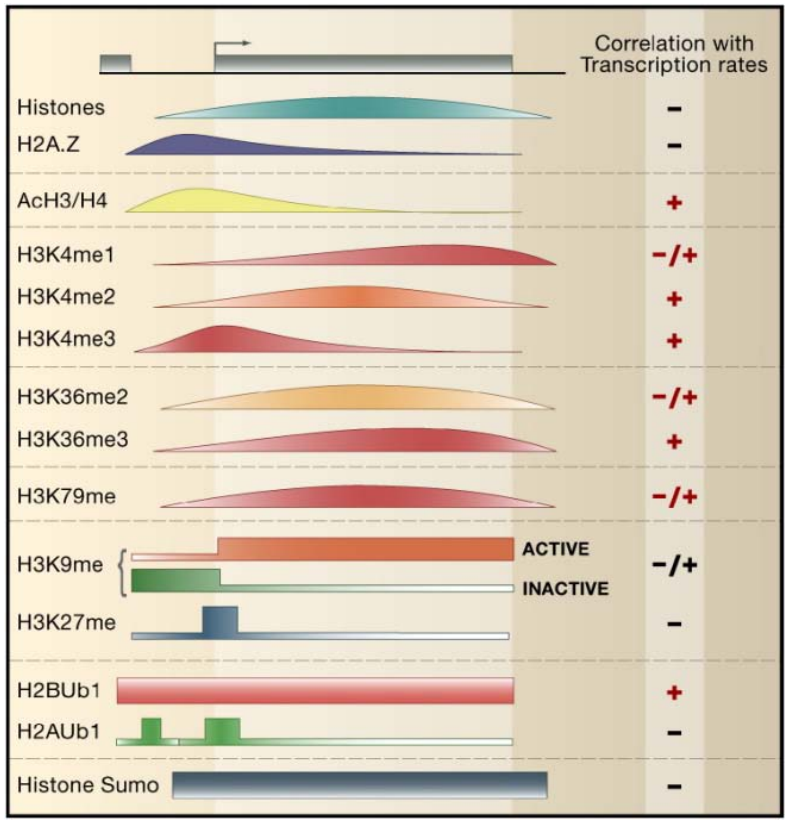

Figure 5. A summary of the effects of histone modifications and histone variants on transcription. (Reprinted from Li *et al*. 2007a with kind permission of Elsevier.)

2.2.2.4 Other chromatin modifications and histone variants

In addition to the above DNA and histone modifications, histone phosphorylation, sumoylation, ubiquitination, ADP-ribosylation, proline isomerization, arginine deimination, histone variants, and DNA breaks are chromatin modifications involved in the regulation of transcription and these processes are mediated by specific enzymes that function as coregulators (Kouzarides 2007a).

Serine (S) and threonine residues of the histones can be phosphorylated by specific kinases. H3S10p is linked to chromatin condensation as well as to transcriptional activation, but the precise mechanisms are not known (Johansen and Johansen 2006, Hayashi-Takanaka *et al*. 2009). H3T11p is mediated by protein-kinase-C-related kinase 1 (PRK1) and this is one of the essential modifications for AR-mediated transcription, since it accelerates demethylation of H3K9 by JMJD2C (Metzger *et al*. 2008).

Histone sumoylation is linked to transcriptional repression (Fig. 5) (Shiio and Eisenman 2003). At least in yeasts, all of the core histones can be sumoylated on specific lysine residues. The sumoylation represses transcription by competing with the activating histone modifications at the same lysine residues (Nathan *et al*. 2006).

Histone monoubiquitination is linked to both transcriptional repression and activation (Fig. 5). Only two residues, H2AK119 and H2BK120, have been shown to be modified by ubiquitin (Weake and Workman 2008). H2AK119ub can repress transcription by inhibiting RNAPII elongation by preventing recruitment of elongation factor FACT (Zhou *et al*. 2008). Deubiquitination of H2AK119ub by 2A-DUB deubiquitinase enhances histone acetylation and H1 dissociation and thus activates transcription (Zhu *et al*. 2007). H2BK120ub is located at transcribed regions of active genes, but it can also be a repressive marker for certain genes (Shema *et al*. 2008, Minsky *et al*. 2008).

The function of ADP-ribosylation on transcription is not fully understood. The modification can be either of mono or poly type, but only mono-ADP-ribosylation is linked to transcriptional regulation. Only one residue of all the core histones, H2B glutamate 2 (E2), has been found to be modified by mono-ADP-ribosylation. Instead, many glutamate and arginine residues of linker histone H1 and its variants can also be modified. The modification may also undergo some interplay with other histone modifications, such as acetylation, but its main function is perhaps to be a marker for DNA damage (Hossa *et al*. 2006).

Proline (P) can exist either in a *trans* or in *cis* conformation, with the *trans* being the predominant form (~90%). The structure of a protein can change dramatically due to isomerization of proline residues. Proline isomerases are enzymes that can change the proline's conformation in either direction. H3P30 and H3P38 are known targets for Fpr4 isomerase. The conformation of H3P38 is important for H3K36 methylation *i.e*. the methylation can occur only when H3P38 is in the *trans* conformation. Since the *trans* conformation is the predominant form, Fpr4 activity moves the equilibrium towards the *cis*-conformation, and thus Frp4 activity inhibits H3K36 methylation and transcriptional activity (Nelson *et al*. 2006).

In the arginine deimination reaction, arginine is transformed to citrulline by a specific enzyme called PADI4. The deimination takes place with non- or monomethylated H3R2 and H4R3, but not on dimethylated versions. Thus deimination is one of the indirect mechanisms of demethylation and its effect on transcription is mostly due to demethylation, with the effect usually being repressive (Wang *et al*. 2004).

In addition to the canonic histones, depending on the genomic location and the activity of the genes, the histones can be replaced with histone variants. All of the histones possess variant versions, albeit variant H4 (H4V) exists only in trypanosome (a primitive eukaryote), and they all have a specific role in chromatin functions (Talbert and Henikoff 2010). In humans, five different H3 variants have been found: H3.1, H3.2, H3.3, H3t, and CenH3. H3.1 and H3.2 are canonical replication specific variants and they are commonly found in S-phase cells, especially in heterochromatin. H3.3 is replication independent and is located at the promoters, enhancers and body of active genes that contain active histone marks. H3.1, H3.2, and H3.3 can be involved also in epigenetic memory. H3t is testis specific and CenH3 centromere specific variant of H3 (Hake and Allis 2006). H2B has only two variants and both, TH2B and H2BFWT, are testis specific. H2A can have six variants: H2A.X, H2A.Z, macroH2A, H2A.Bbd, H2AL1,L2, and H2Abd. The function of macroH2A and H2A.Bbd is not known and H2AL1,L2 is a testis specific variant. H2A.X and H2Abd are known to be linked to the repair of dsDNA breaks. The most important H2A variant for transcription is H2A.Z which possesses both repressive and active effects on transcription. It is enriched at the boundaries of nucleosome free regions, such as TSSs, and facilitates RNAPII recruitment. However, the variant is also enriched in heterochromatin, pointing to a role in transcriptional repression (Fig. 5) (Talbert and Henikoff 2010).

Topoisomerase IIβ (TopoIIβ) is an enzyme that can generate breaks in dsDNA. Interestingly, the TopoIIβ mediated dsDNA break is needed for transcription initiation together with subsequent poly-ADP-ribose polymerase 1 (PARP1) DNA repair complex activity. The mechanism to explain how the DNA breakage and repair can enhance transcription is based on H1 exchange and thus on chromatin conformation changes (Ju *et al*. 2006, Ju and Rosenfeld 2006).

2.2.2.5 Non-coding RNAs

ncRNAs are transcribed RNAs that are not translated to proteins. Until recently, their important roles in transcriptional regulation and epigenetic silencing were not appreciated. The gene silencing mechanisms can be categorized into two groups: small RNA-directed transcriptional gene silencing (TGS) and siRNA-targeted post-transcriptional gene silencing (PTGS). In TGS, the gene expression is inhibited before transcription, whereas in PTGS, transcription is intact, but translation is inhibited due to degradation of the synthesized mRNA. The other fundamental difference between TGS and PTGS is that TGS can be heritable, whereas PTGS is mostly transient (Morris 2009).

In TGS, ncRNAs are transcribed from the anti-sense strand of the silenced gene or its promoter. The RNAs can be either short or long strands. The transcription is usually bidirectional, *i.e*. both the sense (the mRNA) and the antisense (the ncRNA) RNAs are transcribed simultaneously. The main function of these RNAs is to regulate the expression of the corresponding gene. The RNAs targeted to the promoter region hybridized with the target DNA sequence, where they recruit a protein called argonaute 1 (Ago-1), which is able to recruit histone and DNA modification enzymes, such as HDAC1, DNMT3a, and DNMT1. These then evoke increased methylation of H3K9, H3K27, and DNA that, for one, silences the gene. If the ncRNA is expressed for a long time (3 to 4 days), the epigenetic changes become stable and inheritable. If the target sequence of the ncRNA is located at the TSS or at the boundary of an exon and intron, Ago-2 instead of Ago-1 mediates the effects such as inhibition of RNAPII recruitment or changes in RNA splicing. In PTGS, small antisense RNA of the mRNA of gene to be silenced is hybridized with the mRNA leading to dicer-mediated mRNA degradation. As mentioned above, PTGS is transient, but easy to execute in the laboratory by transfecting the cells with small interfering RNAs (siRNAs) (Morris 2009).

ncRNAs can also be involved in transcriptional activation. The effect is not direct activation, but instead more resembling derepression. Activating ncRNAs are targeted against long antisense RNAs that are inhibitory. These anti-antisense RNAs decrease epigenetic effects of anti-sense RNAs by some unknown mechanism and more mRNAs are translated into proteins (Morris 2009). Recently, a new class of ncRNAs called enhancer RNAs (eRNAs) has been found. The exact role of these RNAs is not clear, but they are thought to be necessary for promoter activation of the corresponding gene (Kim *et al*. 2010).

2.2.2.6 Epigenetic inheritance

After the publication of Charles Darwin's theory of evolution, according to which inherited properties of an individual are not caused by environmental factors, but only by natural selection, Jean-Baptiste Lamarck's theory of inheritance of acquired characters was forgotten for many years. Today, it looks that both scientists were right. Inheritance of acquired characters is called epigenetic inheritance and this consists of all the changes in usage of genetic information that can be passed on to the next generation. The epigenetic information is represented by the chemical chromatin modifications discussed above.

How are these modification patterns then copied to daughter cells during mitosis or meiosis? This question has remained rather an enigma, albeit some mechanisms have been proposed. One clear mechanism involves DNA methylation by DNMT1, where the enzyme copies the methylation patterns to the newly synthesized DNA strand during mitosis. The mechanism on how histone modification patterns are copied to the daughter cells is poorly understood. However, the mechanism to explain how heterochromatin is copied to the daughter chromatin strand has been partly resolved. In the mechanism, trimethylated H3K9 recruits HP1 that can recruit histone methyltransferases that methylate newly synthesized chromatin strands during mitosis (Bernstein *et al*. 2007, Portin 2009). In addition, histone replacement with the variant species by nucleosome remodeling factors during replication has been proposed to be one mechanism in epigenetic inheritance (Henikoff 2008).

Altought epigenetic inheritance during simple mitosis is hard to understand, it is even more difficult to comprehend the mechanism how the epigenetic information is transferred from parents to offspring in multicellular organisms. For example, starvation of the father in his childhood may affect

also the daughter's epigenetic profile in metabolic genes, but the mechanisms have not been clarified. Possible candidate factors that carry the epigenetic memory from one individual to another are small ncRNAs. In fact, these ncRNAs may be the main factors mediating all the epigenetic information and the chromatin modifications are simply the end points of the action (Kouzarides 2007a). However, more studies are needed to clarify the nature of ncRNAs and the whole nature of epigenetic inheritance *per se*.

2.3 PROSTATE CANCER

The prostate gland is a male-specific organ that is located around the urethra just under the urinary bladder in the pelvic cavity. It produces most of the seminal plasma (the cell-less part of the sperm). Androgens, especially DHT, and AR are essential for prostate cell proliferation and differentiation during embryogenesis and also after birth in puberty, when the prostate differentiates to its final form. During adulthood, androgens are responsible for maintaining the prostate function. Histologically, the prostate tissue can be divided into stroma or connective tissue and glandular ducts. The wall of the glandular ducts is formed by simple epithelium that excretes the seminal plasma into the lumen of the ducts. The ducts ultimately open into the urethra, through which the sperm is transported outside the body during ejaculation. Prostate hyperplasia is a very common age-related benign process that causes urination problems due to physical pressure on the urethra. However, benign hyperplasia can transform into malign prostate adenocarcinoma that is the most prevalent form of PC, but the cancer development does not require preceding benign hyperplasia. In fact, it is not known for sure whether benign hyperplasia is a susceptibility factor for PC. PC is almost invariably androgen-dependent at its onset, but the androgen-dependency may diminish while the cancer progresses (Heinlein and Chang 2004). PC is the most diagnosed cancer and second common cancer-related cause of death in western males (Jemal *et al*. 2008). For example in Finland, over 4,000 men (0.16% of male population, 30.7% of all the male cancers) receive a PC diagnosis every year (www.cancer.fi, www.cancerregistry.fi). In addition, many PCs remain undiagnosed, because for example from 70–80% of over 80-year old men are found to have asymptomatic PC post-mortem, even though the cause of death was some other illness (Editors of the physician's database 2009). Since the risk to contract the disease increases with aging and people are living longer, PC

has become more and more common on a global level. In addition, life-style factors, such as increased intake of saturated fats and red meat and decreased intake of vegetables and fruits, are linked to a higher risk for development of symptomatic PC (Tomlins *et al*. 2006b). Thus, the mechanisms contributing to this disease are under extensive investigations within efforts to develop novel treatments for the disease.

2.3.1 Diagnostics and therapy

The symptoms in PC are very similar to that of benign prostate hyperplasia, *i.e*. urination problems including for example weak spout and frequent need for urination. In addition to these symptoms, the back surface of the prostate can become nodular when palpated through the rectum that is undetectable in prostate hyperplasia. The most widely used biomarker for PC is an increased level of prostate specific antigen (PSA, also called kallikrein 3 or KLK3) in the serum. The degree of the increase correlates well to the disease stage; the levels increase while the cancer progresses. *PSA* is an androgen-regulated gene, whose protein product is a serine protease secreted by prostate epithelial cells into the glandular ducts, where it degrades the large proteins produced by seminal vesicles to prevent coagulation of the sperm. Normally, very low levels of PSA can seep into the circulation, but in PC, due to abnormalities of the glandular ducts, the passage becomes easier and thus the PSA levels are increased in serum (Heinlein and Chang 2004). The diagnosis can be further ensured by prostate biopsy and subsequent histological determination, *i.e*. histopathological diagnosis. If the cancer has developed to the metastatic state, bone pains can be felt, because the most prevalent metastatic tissue of PC is bone, especially the ribs and vertebrae (Editors of the physician's database 2009). The grade of PC is determined by a PC-specific scoring system called the Gleason score that was originally established in 1966 (Gleason 1966). In the technique, a histological sample is graded from 1 to 5 depending on the cell differentiation and other growth properties of the tumor. The most common and second common growth pattern of the same sample is graded separately and the sum of the scores is the final Gleason score (2–10). The higher the score, the more aggressive is the cancer (Lotan and Epstein 2010).

The type of the therapy depends on the disease state, *i.e*. whether the cancer has not spread (is it inside the surrounding capsule of the prostate), locally spread, or spread (metastatic). In fact, most PCs are local and asymptomatic and that progress very slowly and never spread outside the capsule. Advanced cancer, however, spread usually first to the closest lymph nodes *via* lymphatic

vessels and later to the para-aortic nodes. The cancer cells can also spread *via* blood vessels to axial skeleton, especially to lumbar spine (Clarke *et al*. 2009). The strategy of therapies varies depending on the country, and the Finnish system is focused below. The cancer that has not spread can be cured by radical prostatectomy, where the prostate is removed surgically, or by radiation therapy either traditionally through the skin (external) or intraprostatically by small radiating granules. The locally spread PC can be treated by prostatectomy, external radiation, or androgen deprivation therapy that includes chemical or surgical castration, and/or antiandrogen therapy (Fig. 6). Metastases are treated usually endocrinologically by chemical or surgical castration and antiandrogen therapy. Though prostatectomy is powerful therapy for PC, it causes many regrettable side-effects, since almost every patient becomes impotent and half of them suffer urinary incontinence after the operation. In medical castration, gonadotropin-releasing hormone (GnRH)-analogs are used to lower the cyclical release of GnRH needed for luteinizing hormone (LH) excretion from the pituitary gland. LH is a hormone that stimulates androgen production by Leydig cells and thus a decline in LH decreases also androgen production (Editors of the physician's database 2009).

Antiandrogens are either steroidal or non-steroidal AR antagonists. They compete with the androgens for binding to the ligand binding pocket of AR (Fig. 6). Antagonist-bound AR is unable to recruit coactivators and/or nuclear translocation is inhibited, and thus the AR-mediated transcription becomes depressed. That leads to inhibition of AR-mediated cell proliferation and even apoptosis of the prostate epithelial cells. Unfortunately, the androgen deprivation therapy usually fails and the disease becomes reactivated, leading to death (Powell *et al*. 2006). The mechanism of the reactivation will be discussed below. Three antiandrogens are used as medicines in Finland and of these, only cyproterone acetate (CTA) is a steroidal compound, whereas bicalutamide (BIC) (for structures, *cf*. Fig. 13) and flutamide are non-steroidal drugs (Kariaho *et al*. 2010). CTA is actually a partial agonist, *i.e*. it has some agonistic properties. In addition to competitive inhibition of AR, the second effect of CTA on androgen action is inhibition of LH excretion that finally decreases androgen production by the testes. BIC and flutamide are pure antagonists and of these BIC is more used due to worse side-effects of flutamide (Powell *et al*. 2006). However, the overexpression of AR can convert BIC to act as a partial agonist that is one of the possible mechanisms of the PC reactivation (Chen *et al*. 2004). Recently, novel antiandrogens RD162 (Fig. 13) and MDV3100 have been developed that can function as an antagonist also in

AR-overexpressing cells (Tran *et al*. 2009, Scher *et al*. 2010). Non-competitive AR inhibitors have also been developed, but none of them is presently in clinical use (Jones *et al*. 2009). In addition to antiandrogens, 5α-reductase inhibitors can be used to reduce DHT synthesis in peripheral tissues (Walsh 2010).

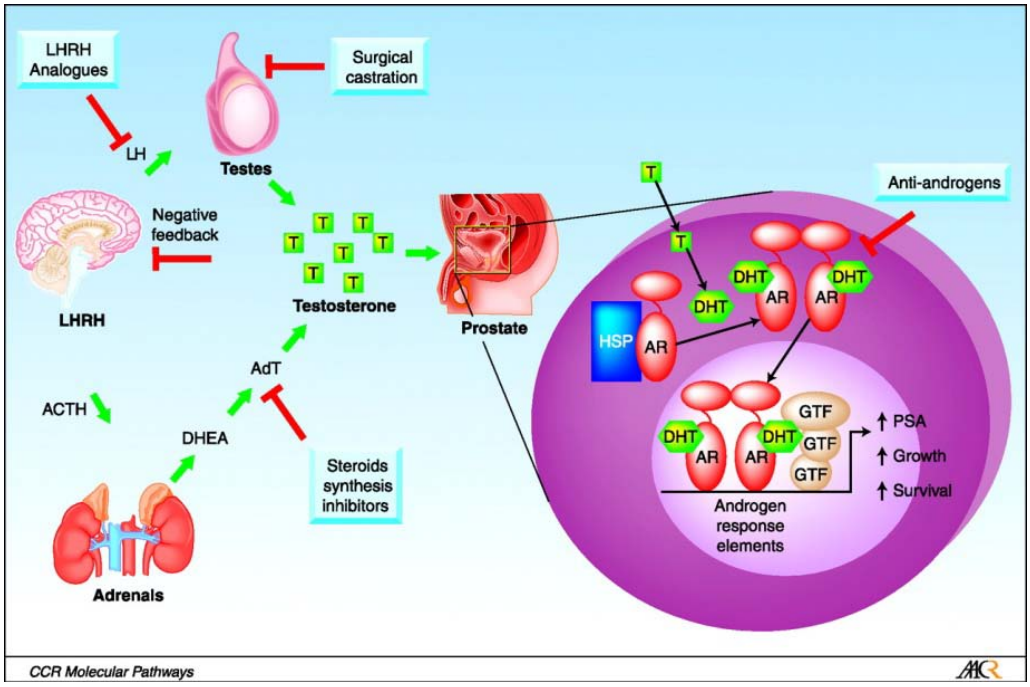

Figure 6. The strategies in androgen deprivation therapy. (Reprinted from Attar *et al*. 2009 with kind permission of AACR Publications.)

2.3.2 Molecular mechanisms

Molecular mechanisms involved in PC have been investigated extensively for many decades. Initially, Charles Huggins and Clarence Hodges (1941) showed that the PC growth was dependent on androgens. Still today, it is generally believed that the dysfunction of AR-mediated gene regulation is a major factor in PC development (Heinlein and Chang 2004, Culig *et al*. 2002). However, an excess of androgens does not cause PC in the healthy man, suggesting that androgens are not involved in the initial carcinogenesis, but in its late progression (Hsing 2001). However, androgens may promote chromosomal translocations that lead to overexpression of growth promoting genes (Mani *et al*. 2009). These translocations are believed to be the main promoting factors for PC together with concurrent loss of phosphatase and tensin homolog (PTEN).

PTEN is a tumor suppressor inactivating v-akt murine thymoma viral oncogene homolog 1 (AKT1), *i.e.* one of the kinases activating cell growth (Kumar-Sinha *et al*. 2008, Sircar *et al*. 2009, Squire 2009, Xin *et al*. 2006). The role of androgens and AR in PC progression as well as the role of chromosomal translocations will be discussed below more extensively.

2.3.2.1 Chromosomal translocations

Transmembrane serine protease 2 (*TMPRSS2*) is an androgen regulated gene that is expressed in human prostate, colon, stomach, and salivary gland, and whose function is unknown (Lin *et al*. 1999, Vaarala *et al*. 2001, Kim *et al*. 2006). Interestingly, five years ago it was found that *TMPRSS2* forms recurrent fusions with E twenty-six or E26 transformation-specific (ETS) TF genes (Tomlins *et al*. 2005). ETS TFs regulate genes involved in several important cellular functions such as growth, apoptosis, development, and differentiation, and they can be thus called oncogenes (Oikawa and Yamada 2003). *TMPRSS2* can be fused at least with *v-ets erythroblastosis virus E26 oncogene homolog* (*ERG*), *ets variant 1* (*ETV1*), *ETV4*, and *ETV5* from which the fusion with *ERG* is clearly the most prevalent (*TMPRSS2-ERG* is found in over 50% of all the PCs and the rest constitute less than 10%) (Tomlins *et al*. 2005, Tomlins *et al*. 2006a, Helgeson *et al*. 2008, Kumar-Sinha *et al*. 2008). In fact, the high prevalence makes the *TMPRSS2-ERG* fusion very unique and an important marker for PC, since the fusion is never found in benign prostate hyperplasia (Kumar-Sinha *et al*. 2008). The *TMPRSS2-ERG* fusion is formed by ~3 Mb (or less) deletion between the genes, occasionally by translocation (Yoshimoto *et al*. 2006, Tu *et al*. 2007, Iljin *et al*. 2006). The mechanisms of the deletions or translocations are not well known, but holo-AR may promote the fusion formation (Mani *et al*. 2009). The fusion mRNA contains usually 5'UTR (exon 1) of the *TMPRSS2* and the whole coding region (exons 4-11) of *ERG*, resulting in full length ERG protein production by translation (Tomlins *et al*. 2005, Tu *et al*. 2007). Since *TMPRSS2* is an androgen-regulated gene and it has AREs in its promoter region, the fusion with the ETS genes subjugates them under AR regulation. This subjugation leads to overexpression of ETS genes in prostate cells that express AR and have high DHT concentration. Unexpectedly, when examined, the TMPRSS2-ERG fusion did not enhance PC growth in mice, but instead enhanced the invasiveness of cancer cells (Tomlins *et al*. 2008). Neither vitamin D induced TMPRSS2-ERG fusion mRNA nor ERG protein overexpression can promote PC growth (Washington and Weigel 2010). In contrary, positive effect of the fusion on cell proliferation has been also reported (Wang *et al*. 2008a).

Overexpression of ERG was associated with high expression of HDAC1 and low expression of its target genes, suggesting that the fusion can lead to epigenetic reprogramming (Iljin *et al*. 2006). In fact, recently it has been reported that overexpression of *TMPRSS2-ERG* fusion can increase H3K27 trimethylation of tumor suppressor genes, leading to their downregulation and dedifferentiation of the cancer cells (Yu *et al*. 2010).

In addition to *TMPRSS2*, also other genes have been found as 5'-partners in ETS-fusions in PCs. For example, these include androgen regulated gene *solute carrier family 45 member 3* (*SLC45A3*) and androgen insensitive gene *DEAD (Asp-Glu-Ala-Asp) box polypeptide 5* (*DDX5*) that can form fusions with *ERG* (*SLC45A3*), *ETV1* (*SLC45A3*), *ETV4* (*DDX5*), and *ETV5* (*SLC45A3*) (Han *et al*. 2008, Helgeson *et al*. 2008, Tomlins *et al*. 2007). The prevalence of these fusions is very minimal (1% to 2%) compared to the major fusion *TMPRSS2-ERG* (over 50%) (Esgueva *et al*. 2010, Han *et al*. 2008). Interestingly, *SLC45A3* can form also fusion transcripts with *ELK4*, a member of the ternary complex factor (TCF) subfamily of ETS TFs (Maher *et al*. 2009, Rickman *et al*. 2009). The mechanism of the fusion mRNA formation does not involve any DNA rearrangement, but rather other mechanisms such as intergenic trans-splicing. Maher *et al*. (2009) showed that the chance to have any other known fusion transcript that needs DNA rearrangement together with SLC45A3-ELK4 expression is low, suggesting that they can exclude each other. Rickman *et al*. (2009) reported, however, that there was no true mutually exclusive expression. A detailed discussion about SLC45A3-ELK4 chimaeric transcript can be found in the "Results and discussion" section.

2.3.2.2 The role of AR and coregulators in drug resistance

Most PCs are initially sensitive to androgens. However, androgen deprivation does not kill all the cells, since some of them may have developed mechanisms that confer resistance to apoptosis. These mechanisms develop in the early stage of the carcinogenesis and they include for example overexpression of an antiapoptotic protein B-cell CLL/lymphoma 2 (BCL-2) and other mutations in cell growth genes (Dong 2006). During the androgen deprivation therapy, these surviving cells develop new, circulating androgen-independent mechanisms for enhancing their proliferation. These mechanisms lead to drug resistance and runaway growth of PC cells and finally to death of the patient (Heinlein and Chang 2004). Though the mechanisms are not dependent on circulating androgens, the role of AR is still pivotal and AR-mediated transcription can be activated in several ways, *e.g*. mutations in *AR* including

point mutations and trinucleotide repeats, androgen-independent activation of AR, intratumoral androgen production, and overexpression of AR coregulators and AR itself. In addition, totally AR-independent mechanisms have been suggested for drug resistance (Fig. 7) (Bonkhoff and Berges 2010).

AR contains two trinucleotide repeats, poly-CAG (polyglutamine) and poly-GGN (polyglycine), in its NTD that vary in the number of repeats. The number of the repeats does not usually change *via* somatic mutations, but it is an inherited property. Thus, the number of repeats is a susceptibility factor for drug resistance rather than an actual mechanism for its development during deprivation therapy. However, the number of polyglutamine repeats negatively correlates to AR expression levels and AR-mediated transcription, and thus it may be linked to PC. The decrease in AR-mediated transcription by long glutamine repeats (the genotype in Kennedy's disease) is caused, at least in part, by worse recruitment of coregulators. In humans, the number of repeats varies from 7 to 36 with 22 being the most prevalent in Caucasian males. The association of polyglycine with PC has not been extensively studied and the results are inconsistent (Choong and Wilson 1998, Palazzolo *et al*. 2008, Heinlein and Chang 2004). Several point mutations of AR are associated with PC and drug resistance. They are usually targeted to the LBD of the AR, leading to a widened ligand binding pocket that decreases the ligand specificity. Thus, other ligands, such as estrogens, progestins and even antiandrogens, can bind and activate AR (Gottlieb *et al*. 2004, Heinlein and Chang 2004). For example, over one quarter of hormone-refractory metastatic PCs have the point mutation T877A (alanine), which is also found in a lymph node PC (LNCaP) cell line (Gaddipati *et al*. 1994, Veldscholte *et al*. 1990). Interestingly, E231G (glycine) mutation can even initiate the carcinogenesis in prostate tissue (Han *et al*. 2005). In addition to ligand unspecificity, the mutations can evoke different recruitment of coregulators by AR and thus different gene expression profiles (Brooke *et al*. 2008).

AR can also be activated ligand-independently by other pathways, such as by several growth factors and inflammation signals (Zhu and Kyprianou 2008, Kaarbo *et al*. 2007). For example, overexpression of HER-2/neu tyrosine kinase can modulate the AR-signaling to function under low androgen conditions and interleukin 6 (IL-6) can activate AR in a ligand-independent manner (Craft *et al*. 1999, Malinowska *et al*. 2009). Moreover, several studies have claimed that there is an overexpression of AR coregulators as the key factor in the development of drug resistance (Golias *et al*. 2009, Culig *et al*. 2004, Chmelar *et al*. 2006). For example, overexpression of SRC-1 and SRC-2 in recurrent PCs

increased the AR-mediated transcription with physiological concentrations of adrenal androgens (Gregory *et al*. 2001). However, in another study, overexpression of AR coregulators could not be observed in clinical samples (Linja *et al*. 2004). Overexpression of the coactivators may promote a prolonged AR interaction with AREs that leads to androgen-independent activation of AR (Shi *et al*. 2008). Even though AR is usually the main regulator of PC cell growth, AR can be also bypassed by several mechanisms. For example, overexpression of ERα can subjugate the growth control of the cells under estrogen or progesterone regulation instead of androgens (Bonkhoff and Berges 2009).

Even though androgen depletion therapy decreases blood androgen levels almost completely (95%), the levels are decreased only by 50-80% in the prostate tissue. Nonetheless, that level of decrease can kill most of the cells, but some can still survive. The surviving cells are in some way hypersensitized to the low level of androgens, a process that can occur by the mechanisms discussed above and/or by AR overexpression (Heinlein and Chang 2004). The overexpression can be caused for example by *AR* gene amplification (Koivisto *et al*. 1997, Gregory *et al*. 2001, Linja *et al*. 2001). For example, in a vertebrae PC (VCaP) cell line, at least five copies of *AR* were found leading to overexpression of AR (Liu *et al*. 2008). Interestingly, it has been reported that the overexpression can transform the pure antiandrogen BIC to act as an agonist, suggesting that the overexpression can directly lead to the development of drug resistance (Chen *et al*. 2004). Moreover, overexpression of AR can directly sensitize hormone-refractory cells to low levels of androgens that lead to increased AR-mediated transcription (Waltering *et al*. 2009). The sensitizing is facilitated by intratumoral synthesis of androgens that cannot be blocked by chemical castration and this further depresses the efficiency of the treatments (Mostaghel *et al*. 2007, Montgomery *et al*. 2008).

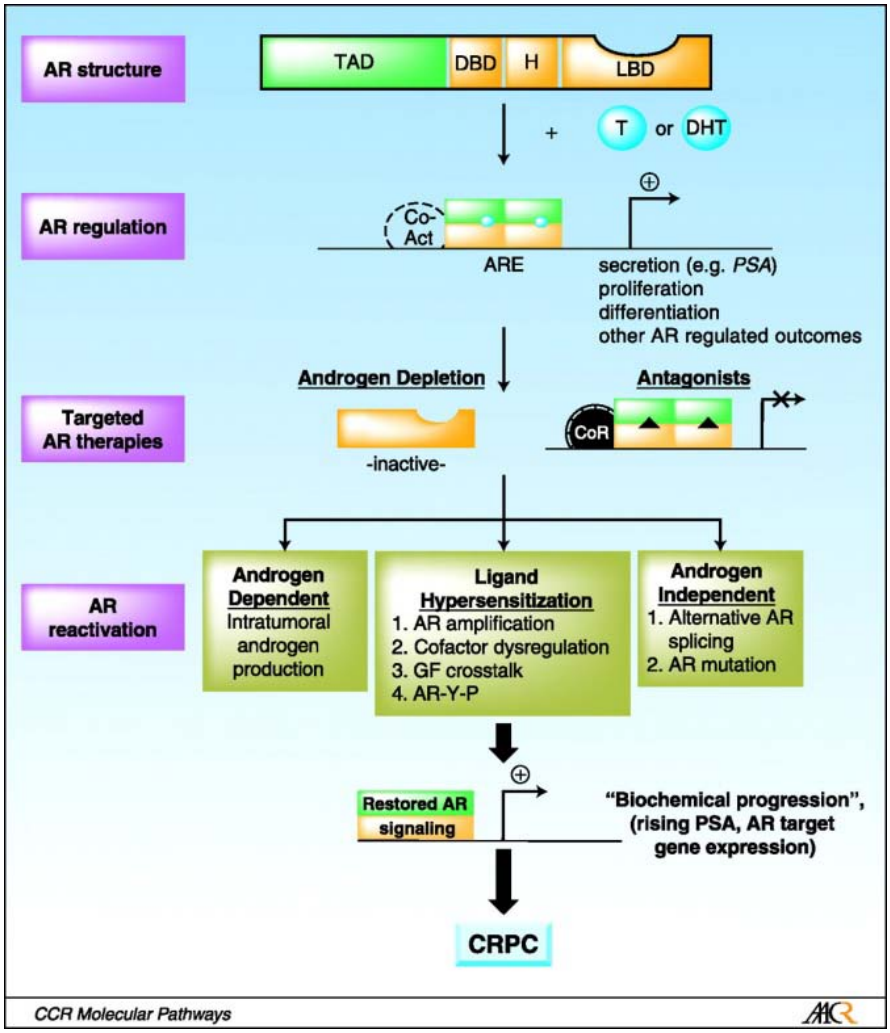

Figure 7. The mechanisms on developing castration-resistant prostate cancer (CRPC). (Reprinted from Knudsen and Scher 2009 with kind permission of AACR Publications.)

2.3.3 Animal and cell models

In addition to prostate development, androgens are also responsible for the development of many other male and female organs and after development they remain crucial for organ function. Malfunction of AR can cause several diseases in addition to PC, *i.e.* androgen insensitivity syndrome (AIS), Kennedy's disease, and infertility. To study the physiological role of AR and androgens, the development of AR knock-out and knock-in animal models is crucial and a number of AR knock-out mice have been developed (Kerkhofs *et al*. 2009). Men having AIS have female external sexual characteristics, such as

female genitalia including clitoris and blind end vagina, but the internal female organs, such as the uterus, are absent. Moreover, the testes are undescended, the prostate, vas deferens, epididymis, and seminal vesicle are absent, female-type breasts are developed and the axillary and pubic hair are absent (Galani *et al*. 2008). Forty years ago, a mouse model having the AIS phenotype was described (Lyon and Hawkes 1970). The mice were infertile and their testes were small and undescended, suggesting that androgens regulate the inguinoscrotal migration. Also several transgenic mice models with a ubiquitous knock-out of the AR (ARKO) mirrors the findings about AIS phenotype (Yeh *et al*. 2002, Kerkhofs *et al*. 2009). Female ARKO mice have also been developed. The studies with these mice have been shown that AR plays an important role in the normal development and functions of ovaries, uterus and breasts in females (Zhou 2009). Interestingly, male ARKO mice have impaired skeletal muscle development, whereas female ARKO mice have normal muscle function, suggesting that AR and androgens regulate muscle growth and strength in males, but not in females (MacLean *et al*. 2008). Cell-specific ARKO mice models have also been developed. Sertoli cell-specific ARKO mice males show a normal male phenotype, except that the testes are smaller and spermatogenesis is impaired, suggesting that androgens are essential for Sertoli cell function in spermatogenesis. In addition, Leydig cell and peritubular myoid cell-specific ARKO mice exhibited smaller testes, but only Leydig cell-specific ARKO mice were infertile. Together these data suggest that AR has a pivotal role in male spermatogenesis (Wang *et al*. 2009b). In addition to knock-out models, some interesting knock-in mice have been developed. For example, a mouse model for Kennedy's disease, where the number of CAG repeats was increased, displayed spinal and bulbar muscular atrophy and loss of fertility (Yu *et al*. 2006). By changing the second zinc finger of AR's DBD with that of GR's (specificity affecting AR knock-in (SPARKI)), some effect on fertility and reproductive organs could be observed, suggesting that AR-specific AREs are important for male fertility, but not for the anabolic effects of AR (Schauwaers *et al*. 2007).

Transgenic adenocarcinoma of the mouse prostate (TRAMP) model has been developed as a way of studying PC. The TRAMP mice have rat probasin promoter driving prostate-specific epithelial expression of the simian virus 40 (SV4O) large T antigen (Skalnik *et al*. 1991, Greenberg *et al*. 1995). Large T antigen can promote tumorigenesis of many organs by affecting retinoblastoma protein (Rb) and p53, the main tumor suppressors of the human cells (Sáenz Robles and Pipas 2009). The TRAMP mice develop prostate

tumors as early as on age of 10 weeks and represent a very useful animal model for studying PC development and progression (Greenberg *et al*. 1995). For example, TRAMP mice were used to study the effect of SRC-3 on PC progression and it was found that SRC-3-knock-out TRAMP-mice did not develop PC, suggesting that AR coregulators are essential for PC progression (Chung *et al*. 2007). In addition, PC cell lines derived from TRAMP mice have been established (Foster *et al*. 1997).

Several prostate cell lines have been established (van Bokhoven *et al*. 2003). Many of them are derived from PC metastasis, such as DU145 (brain), PC-3 (bone), LNCaP (lymph node), VCaP (vertebral body), and DuCaP (dura), but some of them are from primary PC, such as 22Rv1 and CWR-R1. The most widely used cell line is LNCaP with the T877A mutation in the LBD of AR that leads to unspecificity of the ligands (Veldscholte *et al*. 1990, Kuil and Mulder 1996, Bohl *et al*. 2007). VCaP cells overexpress wild-type AR due to *AR* amplification (Korenchuk *et al*. 2001, Liu *et al*. 2008). Since VCaP cells have the main PC-specific *TMPRSS2-ERG* gene fusion and AR is overexpressed, VCaP cells are a very useful model cell line for PC (Tomlins *et al*. 2005).

2.3.4 AR target genes

All the effects on cellular function by androgens and AR are a consequence of altered gene expression of primary and secondary AR target genes. The primary target genes are those that AR regulates directly by binding to the promoter or the enhancer regions of the gene, whereas secondary target genes are not directly regulated by AR binding, but by other TFs activated by AR. The secondary target genes need *de novo* protein synthesis and are thus induced later than the primary genes. AR regulates genes involved in several cellular functions, such as in stimulating proliferation, inhibition of apoptosis, and production of secreted proteases like PSA (Lamont and Tindall 2010). The estimated number of protein-coding mRNAs expressed by AR lies somewhere around 1,000, but in addition a huge amount of ncRNAs are expressed (Dehm and Tindall 2006). Recently, a microarray experiment with LNCaP cells treated with R1881, a synthetic AR agonist, showed that most of the affected genes were involved in transforming growth factor β (TGFβ) signaling pathway, fatty acid metabolism, steroid biosynthesis, or genes with unknown function (Fig. 8) (Ngan *et al*. 2009). Even though several AR target genes have been identified with genome-wide technologies, the molecular mechanisms of AR-regulated transcription have been traditionally studied only in one model gene, *PSA* that contains two AREs in its proximal promoter and a more distal upstream

enhancer region (Cleutjens *et al*. 1996, Cleutjens *et al*. 1997, Schuur *et al*. 1996, Jia *et al*. 2006). Since the mechanisms can vary in a gene- and cell-specific way, it is important to study the mechanisms of the AR-mediated transcription in a detailed fashion using several AR target genes and cell models. Moreover, the disease state of PC, *i.e*. whether it is androgen-dependent or -independent, can alter the pattern of the genes regulated by AR, further pointing out the importance of using a repertuare of models (Wang *et al*. 2009a). A better understanding of the complex mechanisms of AR function will hopefully facilitate the development of improved treatments for PC.

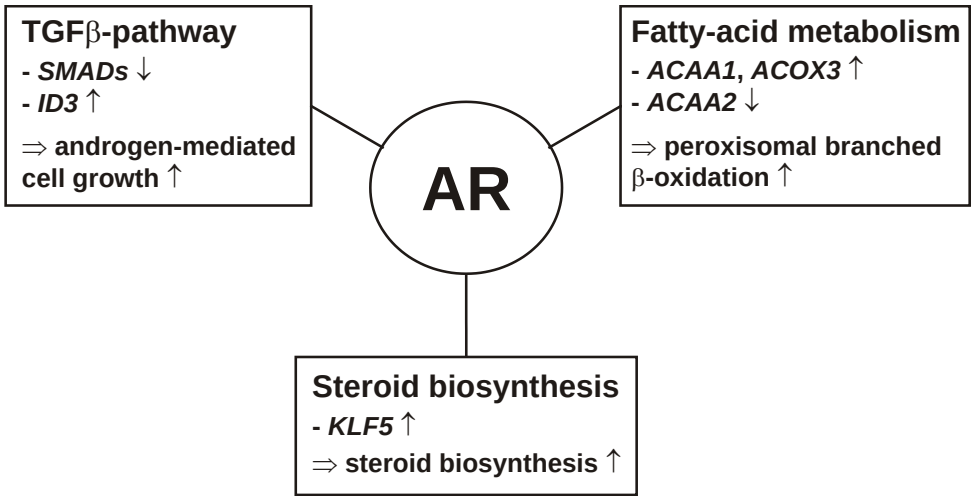

Figure 8. The main pathways and examples of specific genes regulated by androgen receptor according to data from Ngan *et al*. 2009.

3. Aims of the study

At the beginning of this study, only a few AR target genes had been characterized in detail at the chromatin level. As most knowledge about AR-mediated transcription was based only on studies about *PSA* regulation, it was decided to characterize novel AR target genes and to clarify in more detail the mechanisms of AR-mediated transcription. In addition, the molecular mechanisms of AR-mediated PC progression were aimed to study. The more detailed aims of this study were:

1. To characterize the AR-mediated regulation of *ELK4* at chromatin level.

2. To study the role of ELK4 in PC development and progression.

3. To characterize the AR-mediated regulation of *FKBP51* at the chromatin level.

4. To compare the AR- and GR-mediated regulation of *FKBP51*.

5. To study the effects of elevated AR expression on target gene responses to different ligands by comparing two different cell models.

4. Materials and methods

In this thesis, a wide range of molecular biology methods were used (Table 1). Detailed descriptions of the materials and methods can be found in the referred original articles and in Table 2.

Table 1. Materials and methods used in this thesis.

Material or Method	Original article
Cell culture	I, II, III, IV
PCR-primers, EMSA-probes and siRNAs	I, II, III, IV, Table 2
Antibodies	I, II, III, IV
DNA constructs	I, II, III, IV
Quantitative reverse-transcriptase PCR (qRT-PCR)	I, II, III, IV
Reporter gene analysis (RGA)	I, II, III
Electrophoretic mobility shift assay (EMSA)	I, II
Western blotting	I, II, III, IV
Chromatin immunoprecipitation (ChIP)	I, II, III, IV
Cell proliferation assay	I
Tissue microarrays	I

Table 2. Primers used in Fig. 9.

Transcript	Forvard 5'-3'	Reverse 5'-3'
ELK4-E1-E2	AGGCTGAGGGCGGAGAGG	CCATCATTAGAGGTCCAACAG
ELK4-E4-E5	CTGTTGCTCCCCTAAGTCCA	CCAGCCCAGACAGAGTGAAT
SLC45A3-E1-E2	GGCGGAACCAGCCTGCAC	CTGCTTCGTCTCGGCTCTG
SLC45A3-E4-E5	CGCCATCTCCCTGGTCTTC	CAGTGTCCCCTCGGTATTTG
SLC45A3-ELK4-E1-E2	GGCGGAACCAGCCTGCAC	CCATCATTAGAGGTCCAACAG
SLC45A3-ELK4-E4-E2	CGCCATCTCCCTGGTCTTC	CCATCATTAGAGGTCCAACAG

5. Results and discussion

5.1 ELK4 IS A NOVEL AR TARGET IN PC (I)

ELK4, also called serum response factor (SRF) accessory protein 1 (SAP-1), is a member of the TCF subfamily of ETS domain TFs. The TCF subfamily contains also ELK1 and ELK3 (also called NET, SAP-1, ERP). TCFs are activated by MAPKs and they regulate immediate early genes, such as *c-fos* that encodes a subunit of AP-1, by binding to a specific DNA element called serum response element (SRE) *via* their ETS-domain (Dalton and Treisman 1992, Buchwalter *et al*. 2004, Shaulian and Karin 2002). They usually form a ternary complex together with SRF (Mo *et al*. 2001), but they can also regulate target gene expression without SRF (Yamazaki *et al*. 2003). TCFs, like all the ETS domain TFs, are regulators of cell life and death and thus potential oncogenes, since for example blockade of TCF-mediated transcription leads to growth arrest and triggers apoptosis (Vickers *et al*. 2004). TCFs share common features and functional redundancy between them is possible, at least between ELK1 and ELK4, since knock-out mice had relatively subtle defects (Cesari *et al*. 2004a, Cesari *et al*. 2004b, Ayadi *et al*. 2001, Costello *et al*. 2004).

Aberrant expressions of ETS genes are strongly linked to PC progression (Kumar-Sinha *et al*. 2008). Together with *ERG*, *ETV1*, *ETV4*, and *ETV5* (Kumar-Sinha *et al*. 2008), *ELK4* has also been shown to be overexpressed in PC according to microarray data by Edwards *et al*. (2005). Due to the potential role of ELK4 in PC progression, it was decided to study its role in PC in more detail as well as the regulation mechanisms of its gene expression. It was hypothesized that *ELK4* might be regulated by AR and this formed the starting point for this study, *i.e*. explore its androgen-sensitivity. It was found that *ELK4* is an androgen regulated gene, whereas other TCFs were unaffected or slightly downregulated by R1881. In fact, the downregulation of *ELK3* has also been shown by other groups (Bolton *et al*. 2007) and that may be important in AR-mediated cell growth, since ELK3 normally represses the expression of *c-fos* and thus restricts cell proliferation (van Riggelen *et al*. 2005). Two putative AREs were in the proximal promoter of *ELK4* according to *in silico* analysis. The functionality and AR binding capability of putative AREs were confirmed by using a number of different methods including electrophoretic mobility

assays (EMSA), reporter gene assays (RGA), and chromatin immunoprecipitation (ChIP). The more distal ARE2 was shown to be important in prostate derived cells, but not in non-prostate cells, whereas more proximal ARE1 showed an opposite dependency, suggesting that functionality of the AREs is tissue specific. Moreover, a tissue specific TF FOXA1, but not GATA2, was shown to be important for the correct function of ARE2 in PC cells and the enrichment of FOXA1 to its binding element adjacent to ARE2 was shown by ChIP analysis. Others have also emphasized the importance of FOXA1 for DNA binding of AR in PC cells (Wang *et al*. 2007, Gao *et al*. 2003, Jia *et al*. 2008, Mirosevich *et al*. 2005, Mirosevich *et al*. 2006), suggesting that tissue-specific TFs, such as FOXA1, GATA2, and OCT1, define the tissue specific expression of AR target genes.

Transcription variant A or 1 of *ELK4* is encoded by five exons, of which the first one contains only 5′ UTR. Instead, transcription variant B or 2 is encoded by three exons, from which the first two and the beginning of the exon 3 are the same as those of variant A (Fig. 9 and 10A). Therefore, to distinguish between the two variants in study I, it was necessary to design the qRT-PCR primers to the 3′-end of the gene, accurately to exons 4 and 5, and to 2 and 3 of variant A and B, respectively. Recently, Rickman *et al*. (2009) reported that, instead of endogenous *ELK4*, a fusion transcript SLC45A3-ELK4 was upregulated by androgens in LNCaP cells. Since the intergenic space between *ELK4* and *SLC45A3* is only ~25 kb and the direction of the genes is the same, the possibility of chimaeric transcript formation is increased. The fusion transcripts are reported to contain 5′-region of *SLC45A3* (typically only exon 1 or 4) fused to exon 2 of *ELK4*. The exon 1 of *ELK4* and the exon 5 of *SLC45A3* were always absent in the fusion transcripts, suggesting that these exons represent normal expression of *ELK4* and *SLC45A3*, respectively. Moreover, the exon 1 of *ELK4* was also absent in the fusion transcript reported in another study (Maher *et al*. 2009). To re-evaluate and confirm the expression of *ELK4* and *SLC45A3-ELK4* in our experimental system, new primers were designed to test the expression of the endogenous and fusion transcripts in LNCaP and VCaP cells (Table 2). The cells were treated with 10 nM R1881 for 12 h and the mRNAs were analyzed by qRT-PCR. In agreement with Rickman *et al*. (2009), androgen induction of the endogenous *ELK4* (ELK4-E1-E2) was poor in LNCaP cells (1.3-fold) and absent in VCaP cells (1.0-fold) (Fig. 10B,C). Instead, the androgen induction of the endogenous *SLC45A3* (E1-E2 and E4-E5) was detected in both cell lines, but interestingly the inductions were evidently higher in LNCaP cells than VCaP cells (E1-E2, 64-fold vs. 2.8-fold; E4-E5, 17-

fold vs. 2.7-fold). The higher inductions in LNCaP cells than in VCaP cells were probably due to their lower basal expression levels. Similarly, Maher *et al*. (2009) and Rickman *et al*. (2009) observed higher expression of SLC45A3-ELK4 fusion transcripts in LNCaP cells than in VCaP cells. The androgen inductions were also higher in LNCaP cells than VCaP cells (SLC45A3-E1-ELK4-E2, 9.7-fold vs. 3.4-fold; SLC45A3-E4-ELK4-E2, 47-fold vs. 5.1-fold), suggesting that the fusion has a prior oncogenic role in LNCaP cells. According to these results and those of Maher *et al*. (2009), the fusion transcript containing exon 4, but not exons 1-3, of *SLC45A3* fused to exon 2 of *ELK4* is evidently more abundant in both cell lines compared to the SLC45A3-E1-ELK4-E2 fusion. Despite the apparently androgen-insensitive expression of the normal *ELK4*, the AREs found in the proximal promoter of *ELK4* are likely to be involved in the AR-promoted transcription of the fusion products. However, further studies are needed to clarify the precise mechanism how the AR contributes to the formation of the fusion transcripts. In fact, the androgen regulation of endogenous *SLC45A3* has not been investigated at whole locus and chromatin level and the AREs mediating the regulation remain unknown. The AREs found in *ELK4* promoter can also regulate the expression of other transcription variants of *ELK4* that do not contain the exon 1. Few alternative variants have been reported that do not contain the same exon 1 found in variants A and B (Fig. 9). In addition, an alternative promoter is found around exon 2, pointing to an alternative TSS at the beginning of exon 2 (www.genomatix.de). Together these results suggest that the expression of the coding region of *ELK4* is regulated by androgens. This may occur either directly *via* the two AREs found in the proximal promoter of *ELK4*, or through other currently unknown AREs regulating the *SLC45A3* or by a combination of both mechanisms. ChIP deep sequencing analyses of AR-binding sites in LNCaP and VCaP cells may at least in part help to resolve these questions.

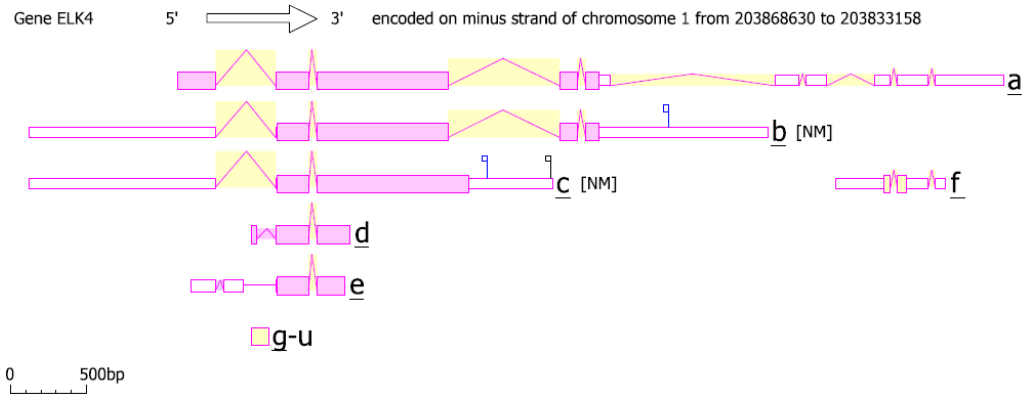

Figure 9. Alternative transcription variants of ELK4. In this diagram, variant b corresponds to variant A in the main text. This is the main variant of ELK4. Moreover, variant c corresponds to variant B in the main text. This variant shares the first two exons with the main variant A, but the exon 3 is longer and exons 4 and 5 are absent. Other variants are less abundant and are not mentioned in the main text. Open boxes indicate UTRs, closed boxes coding sequence and angled lines depict introns. The picture was captured from www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html.

In addition to androgen regulation of *ELK4*, its expression was evaluated in different PCs and its effect on PC cell proliferation. Using siRNAs against ELK4 it was found that its downregulation attenuated the growth of LNCaP cells, suggesting that ELK4 can promote the proliferation of PC cells. Interestingly, in tissue microarray experiments as well as in RNA microarrays conducted by others (Yu *et al*. 2004), ELK4 were overexpressed in hormone-refractory PCs, suggesting that it may have a role in PC progression. In fact, the overexpression of *c-fos*, a target gene of ELK4, can promote the growth of androgen-independent PC (Edwards *et al*. 2004). Another target gene, early growth response 1 (EGR1), can enhance the invasiveness of the aggressive hormone-refractory PC cells by upregulating human protease-activated receptor 1 (HPAR1) (Clarkson *et al*. 1999, Salah *et al*. 2007). Recent findings that SLC45A3-ELK4 fusion transcripts, which encode the whole coding region of *ELK4* and thus full length ELK4 protein, are recurrently overexpressed in PCs, together with the functional evidence in this thesis, strongly suggest that ELK4 has a major role in PC progression (Maher *et al*. 2009, Rickman *et al*. 2009). Moreover, the expression of the fusion transcript can be measured by noninvasive assays from the urine, suggesting that the ELK4 can also be used as a diagnostic marker for PC (Rickman *et al*. 2009). In conclusion, these data suggest that *ELK4* is a novel androgen regulated gene overexpressed in advanced PC that promotes PC proliferation and is thus a potential target for PC therapy together with other ETS-domain TFs overexpressed in PC.

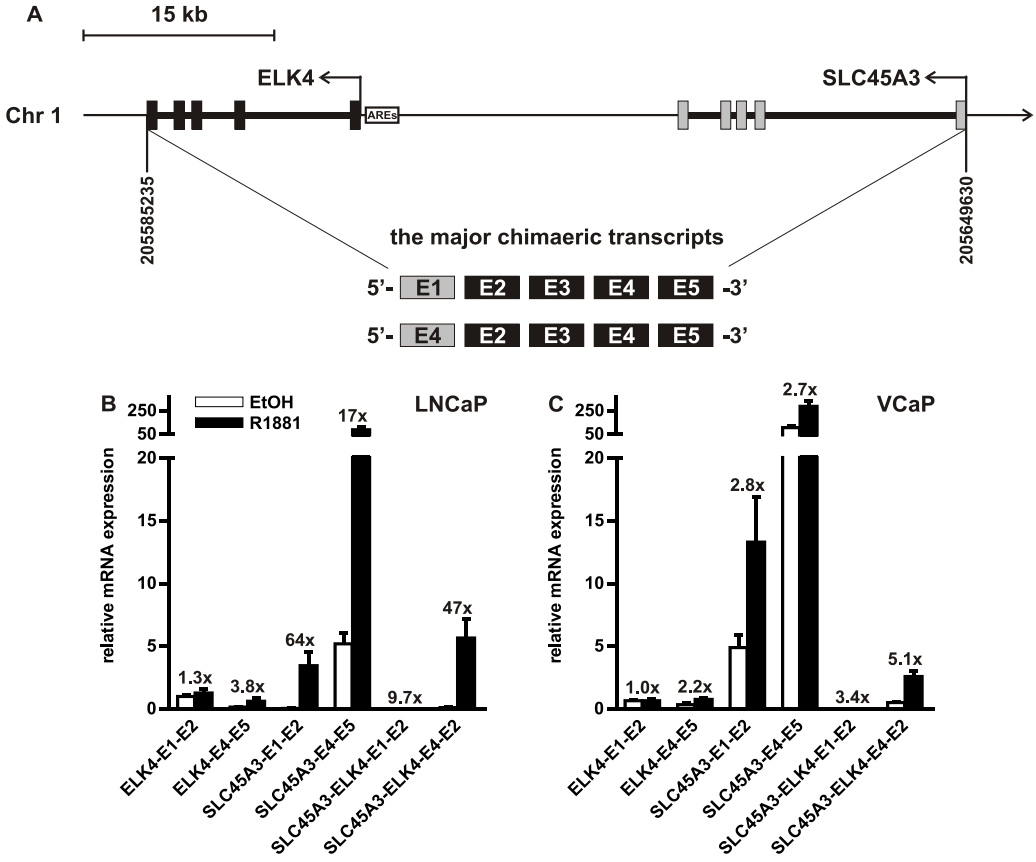

Figure 10. Chimaeric transcription of *SLC45A3* and *ELK4*. (A) Schematic view of the *SLC45A3-ELK4* locus and the major chimaeric fusion transcripts reported by Maher *et al*. (2009) and Rickman *et al*. (2009). The arrow indicates the direction to the end of the long arm (q) of chromosome 1 and the angled arrows depict the TSSs of genes. Black vertical boxes indicate the exons of *ELK4*, grey the exons of *SLC45A3*, and thickened line the gene body. LNCaP (B) or VCaP (C) cells were treated either with vehicle (ethanol) or 10 nM R1881 for 12 h and the mRNAs of the indicated transcripts (primers in Table 2) were analyzed by qRT-PCR likewise in II. Analyzed GAPDH mRNA levels were used to normalize the amounts of total RNA between the samples. The results were calculated using the formula $2^{-(\Delta\Delta Ct)}$, where ΔΔCt is ΔCt(R1881)–ΔCt(EtOH), ΔCt is Ct(gene X)–Ct(GAPDH) and Ct is the cycle at which the threshold is crossed, and finally the value of ethanol treated ELK4-E1-E2 LNCaP-sample was set as one. The number above the bars indicates the ligand induction. Columns represent the mean ± SD of three independent experiments.

5.2 TRANSCRIPTION OF FKBP51 IS REGULATED BY DISTAL ANDROGEN AND GLUCOCORTICOID RECEPTOR-BOUND ENHANCERS (II, III)

Many genes are known to be regulated by androgens and AREs have been characterized genome-widely, but the detailed molecular mechanisms on the

holo-AR-mediated gene regulation of most genes remain elusive. The genome-wide NR-binding studies have shown that most of the binding sites are not located in upstream promoter, but in introns, exons and downstream regions (Bolton *et al*. 2007, Carrol *et al*. 2006, Yu *et al*. 2010). For example, only 4% of ER-binding sites are located in 1-kb proximal promoter region of ER target genes, suggesting that the traditional textbook model of gene regulation needs to be modified (Carrol *et al*. 2006). However, the genome-wide ChIP technologies (ChIP-seq, ChIP-on-chip) and genome-wide expression assays (RNA-seq, RNA-microarrays) do not give the final answer to clarifying the relationship between individual binding sites and target genes, since those assays only assume that the closest binding site is the major regulatory region of a certain gene (Barski and Zhao 2009). The best way to study the relations in a genome-wide manner is to use assays that determine direct chromatin interactions. These include chromatin conformation capture (3C)-based assays, such as chromosome conformation capture carbon copy (5C) and chromatin interaction analysis using paired-end tag sequencing (ChIA-PET) (Fullwood and Ruan 2009).

FKBP51 is an immunophilin protein that functions as co-chaperone in SR-chaperone complexes (Pratt and Toft 1997). Even though AR can directly interact and form a complex with FKBP51, FKBP52 is more important for AR-mediated developmental processes during embryogenesis (Yong *et al*. 2007). However, the expression of *FKBP51* is increased in PC and it promotes AR-mediated transcription and PC cell growth, pointing to a role in PC progression (Amler *et al*. 2000, Velasco *et al*. 2004, Febbo *et al*. 2005, Periyasamy *et al*. 2010, Ni *et al*. 2010). Interestingly, in contrast to androgen-mediated transcription, the overexpression of *FKBP51* negatively correlates to GR and PR activity and thus it has been implicated in glucocorticoid and progestin resistance (Reynolds *et al*. 1999, Hubler *et al*. 2003). These findings suggest that the co-chaperone function is SR-dependent. *FKBP51* is regulated by androgens, glucocorticoids, and progestins (Amler *et al*. 2000, Vermeer *et al*. 2003, Hubler *et al*. 2003, Magklara and Smith 2008). Since the expression of *FKBP51* is more sensitive to depletion of intraprostatic androgens than any other AR target gene (Mostaghel *et al*. 2007), it was hypothesized that *FKBP51* would be mainly under AR regulation and influence of other DNA-binding TFs plays a secondary role in PC cells. Moreover, *FKBP51* has been shown to be a very sensitive glucocorticoid target gene in lung epithelia cells, suggesting that the expression of *FKBP51* is, for one, mainly regulated by GR in lung cells

(Woodruff *et al*. 2007). Thus *FKBP51* was chosen as a model gene for studying AR and GR-mediated transcriptional regulation.

It was found that *FKBP51* is directly, rapidly and strongly induced by R1881 and dexamethasone (a synthetic GR agonist) in VCaP as well as in LNCaP cells and in A549 cells (lung epithelial cancer cells), respectively. The main expressed transcript was variant 1 that differs from its 5′-end compared to the longer variant 2 (Fig. 11 and Fig. 1 in III). The expression was increased very rapidly compared to for example the situation with *PSA*, whose expression increased very slowly in response to androgen treatment and the maximum ligand induction was poor compared to that obtained with *FKBP51*. Cycloheximide treatment did not diminish androgen-dependent expression of *FKBP51*, but instead it slightly increased its overall expression probably due to some indirect mechanisms. The androgen induction of a classical AR target gene *PSA* was diminished by cycloheximide treatment, which when considered together with its slow androgen induction indicates that *de novo* synthesis is needed for full androgen induction of *PSA*, but not for that of *FKBP51*. These data further supports the concept that *FKBP51* is regulated mainly by AR or GR and other TFs do not have any major role in its regulation.

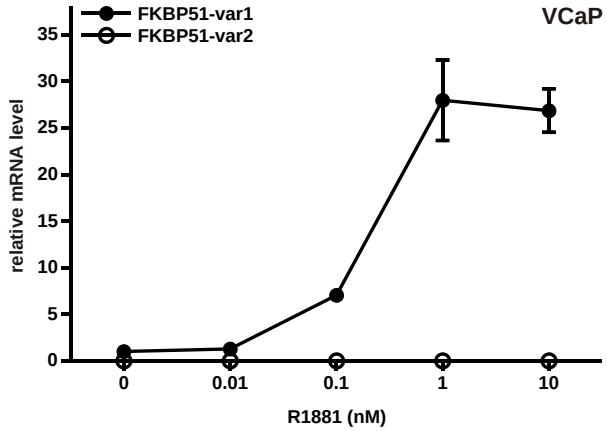

Figure 11. Androgen upregulates FKBP51 transcription variant 1 expression, but not variant 2. The cells were treated with the indicated concentrations of synthetic androgen R1881 for 12 h and mRNAs of indicated transcription variant were analyzed by RT-qPCR analysis likewise in II and III. Total RNA levels between samples were normalized using mRNA levels of GAPDH. The results were calculated using the formula $2^{-(\Delta\Delta Ct)}$, where ΔΔCt is ΔCt(R1881)–ΔCt(EtOH), ΔCt is Ct(gene X)–Ct(GAPDH) and Ct is the cycle at which the threshold is crossed, and finally the value of ethanol treated variant 1-sample is set as one. Columns represent the mean ± SD of three independent experiments.

The rapid and strong androgen-dependency indicates that several AREs/GREs are needed for *FKBP51* regulation. To that end, the whole *FKBP51*

locus was scanned *in silico* to find putative AREs. Initially, 13 AREs were found, from which all but one was located in the intronic regions of the gene. Later, the scan was expanded to more distal 5′ intergenic region and four additional AREs were found. The functionality of the AREs was studied by cloning ~300-bp regions containing AREs with luciferase reporter gene plasmid and their ability to function as AR-regulated enhancers was assessed. The region containing two AREs located at ~34 kb (-3) upstream from the TSS was shown to have the best functionality in VCaP cells by androgen treatment, whereas the region located at ~87 kb downstream that contained also two AREs functioned best in A549 cells in response to glucocorticoid treatment (Fig. 12A, II, III, data not shown). The GR-responsiveness of the region located at ~87 kb downstream has been previously reported as well as its responsiveness also to PR and AR (Hubler and Scammell 2004, Magee *et al*. 2006). Interestingly, in COS-1 cells that do not express endogenous AR or GR, the differences between AR and GR were absent, suggesting that the regulatory differences between the receptors are rather caused by cell line specific accessory TFs than DNA-binding ability of the receptors *per se* (Fig. 12B). Thus, these data suggest that the most functional AREs of *FKBP51* are IR3-type/non-AR-specific AREs.

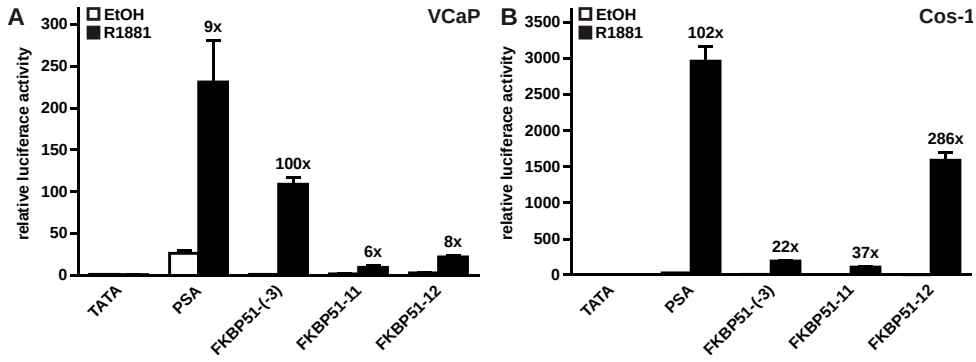

Figure 12. The upstream enhancer of *FKBP51* functions in prostate cell-specific way. (A) VCaP or (B) COS-1 were transfected with reporter construct driven by 5.8-kb PSA promoter (pPSA5.8-LUC), reporter construct containing only TATA-box (pTATA-LUC) or different LUC constructs driven by 0.3-kb FKBP51 fragments harboring AR enhancers likewise in II and III. For COS-1 (A) analyses, pSG5-hAR was cotransfected with the reporter constructs. Cotransfection of pCMVβ and β-galactosidase activity was used for normalization of transfection efficiency. The cells were treated with vehicle (EtOH) or 10 nM R1881 for 24 h before harvesting the cells for reporter analyses. Results are shown as relative LUC activity, with the activity of pTATA-LUC in the absence of R1881 set as 1, and fold inductions of androgen-treated samples in the relation to the activity of ethanol-treated samples are shown above the columns. Columns represent the mean ± SD of three independent experiments.

The *in vivo* binding of AR and GR was studied in ChIP assays. The binding patterns were rather similar between the receptors, but the relative binding efficiency between the different regions was different as was the kinetics of the binding. The best AR and GR binding was found in the region located at ~87 kb downstream which together with the previous findings emphasizing the importance of the region in SR-mediated regulation of *FKBP51* (Hubler and Scammell 2004, Magee *et al*. 2006). The region located at ~34 kb upstream was shown to represent the main difference in binding efficiency between the receptors, since AR binding to it was equal to that of the region located at ~87 kb downstream, but the GR binding was only less than half of that value. These data further suggest that the region located at ~34 kb upstream can function as a prostate-specific enhancer. Interestingly, the kinetics of the receptor binding was different: AR-binding generally peaked at the 60 min time point, whereas GR bound in two waves at the 40 min time point and at 100 min time point. The two-wave-kinetics was seen only at regions ~34 kb upstream and ~87 kb downstream, indicative of a role in transcription initiation, whereas the other two main binding sites at ~78 kb and ~98 kb downstream appeared to be involved in later transcription enhancement. A similar delay in binding was seen with AR, but the binding to the initiator AREs was of the one-wave type. The importance of the difference in kinetics between the receptors remains unclear, but at least the final outcome, *i.e*. the mRNA expression, follows the equal kinetics between the receptors. The receptor binding was similarly blocked by the antiandrogen, BIC and the antiglucocorticoid, RU486 in LNCaP and A549 cells, respectively, but not in VCaP cells overexpressing AR, suggesting that elevated AR levels affect the efficiency of the antiandrogen treatment. However, BIC was almost unable to recruit RNAPII to the chromatin.

The chromatin markers were very similar between the cell lines: the total H3 levels were decreased, acetylation of H3 and methylation of H3K4 marked the TSS and enhancer regions (albeit slightly decreased after androgen/glucocorticoid treatment) and H3K36me3 marked the gene body. In fact, the histone modifications patterns were very similar to that of genome-wide data from other cell lines (*e.g*. B-lymphocyte derived cell line GM12878) produced by ENCODE project (genome.ucsc.edu), suggesting that the modification patterns of *FKBP51* are not restricted only to prostate or lung derived cells. RNAPII occupancy mirrored the trimethylation of H3K36 and interestingly both of them as well as acetylation of H3 were absent in the region unique to the transcription variant 2, further suggesting that variant 1 is

the major transcript of the gene. The decrease in H3 density was shown to be due to recruitment of SWI/SNF complex to the chromatin in the both cell lines, highlighting the importance of ATP-dependent chromatin remodeling in *FKBP51* regulation. Low ligand effects on H3K4me3 and H3ac together with rapid mRNA expression of *FKBP51* suggest that the locus is already poised for transcription and only minor modifications need to be done after hormone treatment.

Due to close proximity of the *C6orf81* to upstream enhancer region of *FKBP51*, the androgen and glucocorticoid responsiveness of the gene as well as the responsiveness of 3'-neighboring gene tubby like protein 1 (TULP1) were tested. *TULP1* did not show any response in either cell line. Interestingly, *C6orf81* was highly induced by androgens in VCaP cells, but not by glucocorticoids in A549 cells even though it has been reported as the GR target gene (Reineke *et al*. 2007). The difference can be explained by poorer GR binding and glucocorticoid responsiveness of the upstream enhancer in ChIP and RGA, respectively. Surprisingly, in osteosarcoma U2OS-GR cells that have been stably-transfected with GR-expression vector, *C6orf81* expression was induced by glucocorticoids and GR could efficiently bind to the upstream enhancer region. This further suggests that the function of the enhancer is cell line-specific and the enhancer is important for AR- and GR-mediated regulation of *C6orf81* (Paakinaho 2010). Another explanation for the unresponsiveness of *C6orf81* in A549 cells could have been a cell-specific insulator between the upstream enhancer and TSS of *C6orf81*. The binding of master insulator protein CTCF and cohesin, an insulator specific binding-partner of CTCF, were however similar in the two cell lines that was not very unexpected, because the CTCF-binding patterns have been shown to be very similar between different cell lines (Phillips and Corces 2009, Cuddapah *et al*. 2009, Heintzman *et al*. 2009). The role of CTCF binding on *FKBP51* locus may be to mediate 3D chromatin interactions instead of insulating the activity of enhancers. A model was proposed where CTCF mediates the formation of a loop that brings all the AR/GR-bound enhancers into one complex around TSS, where they activate RNAPII.

The old textbook version about gene regulation, where the TFs bind two-dimensionally to close proximity of the TSSs of genes and activate their transcription, have been discredited during last few years. These results further support the importance of the 3D-structure of the chromatin in gene regulation, especially the chromatin loop formation that is a prerequisite for long-range chromatin interactions. The absolute relationship between the

found enhancer regions and androgen or glucocorticoid regulation of *FKBP51* or *C6orf81* cannot be determined from these present results. Instead, studies that clarify the direct interactions between the enhancers and TSS of the gene would perhaps provide some hint about the relation, though only disruption of the enhancer *in vivo* would give the final answer. Due to complexity of the regulatory regions of *FKBP51*, the manual 3C would be daunting, but instead high-throughput 3C-based methods would be reasonable for that purpose (Fullwood and Ruan 2009). In conclusion, these results highlight the emerging biological importance of very long-range regulation *via* distal enhancers by AR and GR that should be also taken into account when studying the role of AR in PC development.

5.3 ELEVATED AR LEVELS INFLUENCE THE RESPONSE OF PC CELLS TO ANTIANDROGENS (IV)

The development of drug resistance is a major problem in current PC therapy. The role of AR in the resistance is undisputable and several different defects in AR action have been reported, such as mutations in AR, ligand-independent activation of AR, and overexpression of AR and its coregulators (Heinlein and Chang 2004). The overexpression of AR can occur in several ways, but the amplification of *AR* is the most probable alternative (Koivisto *et al*. 1997, Gregory *et al*. 2001, Linja *et al*. 2001, Liu *et al*. 2008). The overexpression of AR is thought to sensitize the cells to lower amounts of androgens and this is achieved by intratumoral androgen synthesis (Waltering *et al*. 2009, Mostaghel *et al*. 2007, Montgomery *et al*. 2008). Furthermore, the antagonist BIC can also acquire agonistic properties due to overexpression of AR (Chen *et al*. 2004). Recently established antiandrogens RD162 and MDV3100 are, however, resistant to overexpression of AR, probably due to their inhibitory ability for AR nuclear translocation (Tran *et al*. 2009, Scher *et al*. 2010) (Fig. 13). A limited number of cell models for studying the effects of overexpression on AR-mediated transcription are available. The most widely used PC cell model LNCaP has AR expression level equal to normal prostate cells and the receptor has a mutation that affects its sensitivity to antiandrogens (Gaddipati *et al*. 1994, Veldscholte *et al*. 1990, Kuil and Mulder 1996, Bohl *et al*. 2007). The recently established PC cell line VCaP possesses most of the common PC properties, such as overexpression of AR, derivation from vertebrae metastasis, and it has the major gene fusion *TMPRSS2-ERG* that is found in

half of all PCs (Korenchuk *et al*. 2001, Liu *et al*. 2008, Tomlins *et al*. 2005). The AR expressed by VCaP cells is not mutated, which is an advantage in studying the effects of the overexpression in isolation from the other faults in AR-mediated transcription. To that end, LNCaP cells expressing normal levels were chosen with VCaP cells expressing elevated levels of AR as comparison partners in order to study the effects of overexpression on AR-mediated transcription by androgens and antiandrogens.

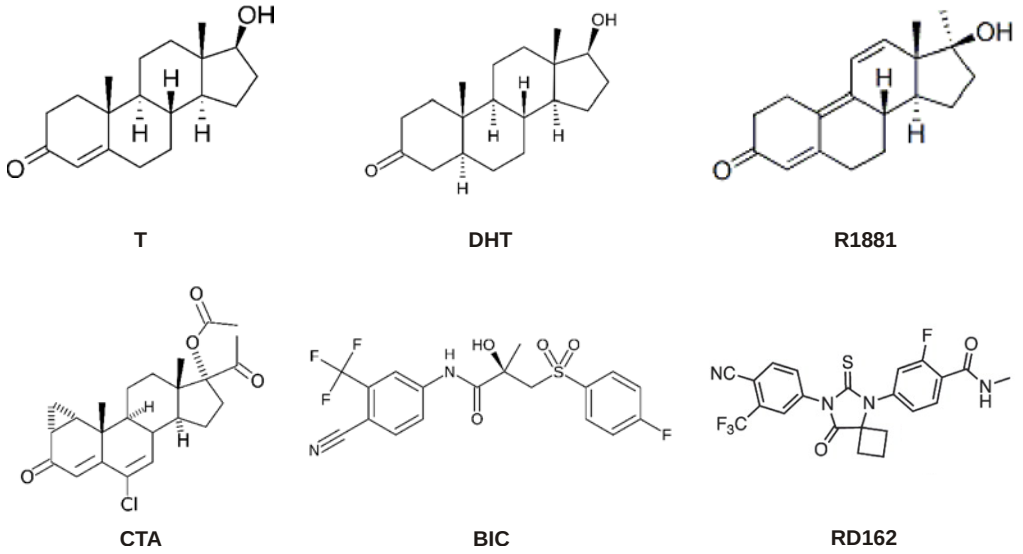

Figure 13. Chemical structures of selected androgens and antiandrogens. Androgens: T, testosterone; DHT, 5α-dihydrotestosterone; R1881, methyltrienolone. Antiandrogens: CTA, cyproterone acetate; BIC, bicalutamide; RD162.

Initially, the effect of androgens and antiandrogens were examined on *AR* expression. The AR levels were approximately ten times higher in VCaP cells than in LNCaP cells both at mRNA and protein levels. After twelve-hours of androgen treatment, however the difference had almost completely disappeared, though it was still 1.5-fold. The decrease in the difference was caused by strong downregulation of AR in VCaP cells and slight upregulation in LNCaP cells, suggesting that the autoregulation of AR is not the same in these two cell lines, perhaps due to overexpression of TMPRSS2-ERG fusion in VCaP cells (Yu *et al*. 2010). Comparing the accumulation of mRNA by increasing R1881 and CTA concentrations revealed that already the basal expression level of seven of the nine AR target genes examined was higher in VCaP cells, suggesting that overexpressed but apparently unliganded AR still possess some residual activity. In fact, abolishing AR by siRNA treatment

decreased the basal expression level of six of the target genes, further supporting the notion that unliganded AR has residual, androgen-independent activity in VCaP cells. Interestingly, the expression level of the main PC diagnostic marker and the most studied AR target gene, *PSA* was significantly higher in LNCaP cells expressing lower levels of AR, whereas expression of a highly androgen sensitive gene *FKBP51* (II) was significantly higher in VCaP cells expressing high levels of AR, further demonstrating that the *PSA* is a rather insensitive AR target gene despite its stature as a model gene.

Due to T877A mutation in LNCaP cells, a partial AR agonist CTA can function as a full agonist (Veldscholte *et al*. 1990, Kuil and Mulder 1996, Bohl *et al*. 2007). In order to determine whether the overexpressed AR could also increase the agonistic properties of CTA, the mRNA accumulation and AR-binding to the main AREs regulating their cognate genes by R1881 and CTA treatment were determined. Interestingly, in both cells, AR target gene expression levels were lower when treated with CTA compared to that of R1881, but the chromatin binding efficiency of AR was at the same level. These results suggest that CTA neither inhibits the nuclear translocation nor DNA binding, but the transactivation capability of AR is diminished. Probably due to the difference in the AR level, the overall AR binding was higher in VCaP cells than in LNCaP cells when comparing the values with samples precipitated by a non-specific antibody. The binding efficiency was also well correlated to the androgen induction of the target gene mRNAs by R1881 treatment, but not by that of CTA. Together these data indicate that the AR-binding affinity can estimate the extent of ligand induction when the cells are treated with a full agonist R1881, but not necessarily when treated with a partial agonist CTA, even though the AR levels would be overexpressed.

In contrast to CTA, the T877A mutation in LNCaP's AR does not have any effect on the antiandrogenic properties of BIC (Berrevoets *et al*. 1993), unlike the overexpression of AR (Chen *et al*. 2004). Opposite to BIC, novel antiandrogens RD162 and MDV3100 have been shown to be resistant to losing efficacy due to overexpression (Tran *et al*. 2009, Scher *et al*. 2010). In fact, it was found that RD162 could compete better with R1881 than BIC and it did not have any agonistic properties when applied on its own. Instead, it decreased the basal activity of some target genes in a similar fashion as siRNA treatment against AR, suggesting that it may actively diminish the residual activity of overexpressed AR, *i.e*. it may function as an inverse agonist. The DNA binding of AR was also inhibited more efficiently by RD162 than BIC and the BIC-induced DNA binding was also reflected to the mRNA expression of target

genes to some extent. In addition, larger differences between the cell lines were revealed by BIC treatments than those of RD162, suggesting that BIC is more sensitive to AR overexpression. The differences in DNA binding between the cell lines were significantly higher than the differences in mRNA accumulation achieved by both antagonists. Even though RD162 was more efficient than BIC in decreasing DNA binding of holo-AR, the decrease was only approximately 50% from the maximum binding, which suggests that the ability of RD162 to inhibit DNA binding of AR is sensitive to AR overexpression, but not its effect on mRNA expression.

In conclusion, the development of new therapies for advanced PC is problematic because the molecular mechanisms of the drug resistance are not fully understood. According to these results and work done with the novel antiandrogens, RD162 and MDV3100, one can say that drug development has moved one step closer to finding efficacious treatment for PC (Tran *et al*. 2009, Scher *et al*. 2010). The main advantages of RD162 compared to BIC are that RD162 does not induce either target gene expression or DNA binding of AR and its competitive inhibition of the receptor is more efficient than that of BIC's. Even though it have been reported that BIC clearly induces AR target gene expression in cells engineered to overexpress AR (Chen *et al*. 2004), the present results with more natural VCaP cells overexpressing AR suggest that the overexpression *per se* is not sufficient to convert the antagonist to act as an agonist, but other factors are also needed.

6. Summary and conclusions

Prostate cancer is the most common malignancy in western males. The development of efficacious therapies is thus very important for public health and the national economy. Knowledge about the molecular mechanisms underlying the disease is a prerequisite for the development of an efficient cure. In this study, these mechanisms were explored taking as an example of a few AR target genes and two PC cell lines.

The role of AR is pivotal in PC development and progression. Thus studying the mechanisms of AR action may provide the answers and further help to clarify the progression of this disease. The AR-mediated transcription has been traditionally studied using only *PSA* as a model gene. However, it was found that *PSA* is actually a rather atypical AR target gene, since its androgen dependency is low and the regulatory regions driving its expression are located unusually close to the TSSs. *FKBP51* was found to be a highly androgen and glucocorticoid-sensitive gene, whose AR/GR-bound enhancers are located very distal from the TSS resembling the average SR target gene. Moreover, the standard model for PC has previously been the LNCaP cell line that does not represent the most common PC type. In this study, the properties of the prototype PC cell line VCaP were characterized. These cells overexpress AR and have the most common gene fusion, TMPRSS2-ERG, in their genome. The AR target genes that mediate the oncogenic properties of AR activity are potential drug targets in future PC therapies. The functional implications about the role of ELK4 in PC progression and its androgen-dependency have inspired others to develop a potential diagnostic method for PC. Encouraging information on the superiority of a novel antiandrogen RD162 over the older antiandrogens in AR target gene responses was provided.

In conclusion, novel information about AR-mediated long-range transcriptional regulation was discovered. This may have implications about which are the important AR target genes in PC progression. In the future, the molecular mechanisms involved in the drug resistance should be explored intensively, so that the development of novel therapies would be accelerated. The use of novel high-throughput genome-wide sequencing technologies, such as ChIP-seq and ChIA-PET, may be one way to gain a deeper understanding about AR-mediated transcriptional regulation.

7. References

Abdalla H, Yoshizawa Y, Hochi S. Active demethylation of paternal genome in mammalian zygotes. J Reprod Dev. 2009 55:356-60.

Akhtar MS, Heidemann M, Tietjen JR, Zhang DW, Chapman RD, Eick D, Ansari AZ. TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. Mol Cell. 2009 34:387-93.

Albert M, Helin K. Histone methyltransferases in cancer. Semin Cell Dev Biol. 2010 21:209-20.

Allis CD, Berger SL, Cote J, Dent S, Jenuwien T, Kouzarides T, Pillus L, Reinberg D, Shi Y, Shiekhattar R, Shilatifard A, Workman J, Zhang Y. New nomenclature for chromatin-modifying enzymes. Cell. 2007 131:633-6.

Amler LC, Agus DB, LeDuc C, Sapinoso ML, Fox WD, Kern S, Lee D, Wang V, Leysens M, Higgins B, Martin J, Gerald W, Dracopoli N, Cordon-Cardo C, Scher HI, Hampton GM. Dysregulated expression of androgen-responsive and nonresponsive genes in the androgen-independent prostate cancer xenograft model CWR22-R1. Cancer Res. 2000 60:6134-41.

Antes TJ, Namciu SJ, Fournier RE, Levy-Wilson B. The 5' boundary of the human apolipoprotein B chromatin domain in intestinal cells. Biochemistry. 2001 40:6731-42.

Aranda A, Pascual A. Nuclear hormone receptors and gene expression. Physiol Rev. 2001 81:1269-304.

Archambault J, Friesen JD. Genetics of eukaryotic RNA polymerases I, II, and III. Microbiol Rev. 1993 57:703-24.

Artandi SE, DePinho RA. Telomeres and telomerase in cancer. Carcinogenesis. 2010 31:9-18.

Attar RM, Takimoto CH, Gottardis MM. Castration-resistant prostate cancer: locking up the molecular escape routes. Clin Cancer Res. 2009 15:3251-5.

Ayadi A, Zheng H, Sobieszczuk P, Buchwalter G, Moerman P, Alitalo K, Wasylyk B. Net-targeted mutant mice develop a vascular phenotype and up-regulate egr-1. EMBO J. 2001 20:5139-52.

Banine F, Bartlett C, Gunawardena R, Muchardt C, Yaniv M, Knudsen ES, Weissman BE, Sherman LS. SWI/SNF chromatin-remodeling factors induce changes in DNA methylation to promote transcriptional activation. Cancer Res. 2005 65:3542-7.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. Cell. 2007 129:823-37.

Barski A, Zhao K. Genomic location analysis by ChIP-Seq. J Cell Biochem. 2009 107:11-8.

Becker PB, Hörz W. ATP-dependent nucleosome remodeling. Annu Rev Biochem. 2002 71:247-73.

Bernardi R, Pandolfi PP. Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. Nat Rev Mol Cell Biol. 2007 8:1006-16.

Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. Cell. 2007 128:669-81.

Berrevoets CA, Veldscholte J, Mulder E. Effects of antiandrogens on transformation and transcription activation of wild-type and mutated (LNCaP) androgen receptors. J Steroid Biochem Mol Biol. 1993 46:731-6.

Biddie SC, John S, Hager GL. Genome-wide mechanisms of nuclear receptor action. Trends Endocrinol Metab. 2010 21:3-9.

Bloom K, Joglekar A. Towards building a chromosome segregation machine. Nature. 2010 463:446-56.

Boeing S, Rigault C, Heidemann M, Eick D, Meisterernst M. RNA polymerase II C-terminal heptarepeat domain Ser-7 phosphorylation is established in a mediator-dependent fashion. J Biol Chem. 2010 285:188-96.

Bohl CE, Wu Z, Miller DD, Bell CE, Dalton JT. Crystal structure of the T877A human androgen receptor ligand-binding domain complexed to cyproterone acetate provides insight for ligand-induced conformational changes and structure-based drug design. J Biol Chem. 2007 282:13648-55.

Bolton EC, So AY, Chaivorapol C, Haqq CM, Li H, Yamamoto KR. Cell- and gene-specific regulation of primary target genes by the androgen receptor. Genes Dev. 2007 21:2005-17.

Bonkhoff H, Berges R. From pathogenesis to prevention of castration resistant prostate cancer. Prostate. 2010 70:100-12.

Bonkhoff H, Berges R. The evolving role of oestrogens and their receptors in the development and progression of prostate cancer. Eur Urol. 2009 55:533-42.

Brandl A, Heinzel T, Krämer OH. Histone deacetylases: salesmen and customers in the post-translational modification market. Biol Cell. 2009 101:193-205.

Brivanlou AH, Darnell JE Jr. Signal transduction and the control of gene expression. Science. 2002 295:813-8.

Brooke GN, Parker MG, Bevan CL. Mechanisms of androgen receptor activation in advanced prostate cancer: differential co-activator recruitment and gene expression. Oncogene. 2008 27:2941-50.

Buchwalter G, Gross C, Wasylyk B. Ets ternary complex transcription factors. Gene. 2004 324:1-14.

Buratowski S. Progression through the RNA polymerase II CTD cycle. Mol Cell. 2009 36:541-6.

Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, Brodsky AS, Keeton EK, Fertuck KC, Hall GF, Wang Q, Bekiranov S, Sementchenko V, Fox EA, Silver PA, Gingeras TR, Liu XS, Brown M. Genome-wide analysis of estrogen receptor binding sites. Nat Genet. 2006 38:1289-97.

Casamassimi A, Napoli C. Mediator complexes and eukaryotic transcription regulation: an overview. Biochimie. 2007 89:1439-46.

Cesari F, Brecht S, Vintersten K, Vuong LG, Hofmann M, Klingel K, Schnorr JJ, Arsenian S, Schild H, Herdegen T, Wiebel FF, Nordheim A. Mice deficient for the ets transcription factor elk-1 show normal immune responses and mildly impaired neuronal gene activation. Mol Cell Biol. 2004a 24:294-305.

Cesari F, Rennekampff V, Vintersten K, Vuong LG, Seibler J, Bode J, Wiebel FF, Nordheim A. Elk-1 knock-out mice engineered by Flp recombinase-mediated cassette exchange. Genesis. 2004b 38:87-92.

Chadick JZ, Asturias FJ. Structure of eukaryotic Mediator complexes. Trends Biochem Sci. 2005 30:264-71.

Chapman RD, Heidemann M, Albert TK, Mailhammer R, Flatley A, Meisterernst M, Kremmer E, Eick D. Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. Science. 2007 318:1780-2.

Chawla A, Repa JJ, Evans RM, Mangelsdorf DJ. Nuclear receptors and lipid physiology: opening the X-files. Science. 2001 294:1866-70.

Chen CD, Welsbie DS, Tran C, Baek SH, Chen R, Vessella R, Rosenfeld MG, Sawyers CL. Molecular determinants of resistance to antiandrogen therapy. Nat Med. 2004 10:33-9.

Cheng X, Blumenthal RM. Coordinated chromatin control: structural and functional linkage of DNA and histone methylation. Biochemistry. 2010 49:2999-3008.

Chmelar R, Buchanan G, Need EF, Tilley W, Greenberg NM. Androgen receptor coregulators and their involvement in the development and progression of prostate cancer. Int J Cancer. 2007 120:719-33.

Choong CS, Wilson EM. Trinucleotide repeats in the human androgen receptor: a molecular basis for disease. J Mol Endocrinol. 1998 21:235-57.

Chung AC, Zhou S, Liao L, Tien JC, Greenberg NM, Xu J. Genetic ablation of the amplified-in-breast cancer 1 inhibits spontaneous prostate cancer progression in mice. Cancer Res. 2007 67:5965-75.

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

Clarke NW, Hart CA, Brown MD. Molecular mechanisms of metastasis in prostate cancer. Asian J Androl. 2009 11:57-67.

Clarkson RW, Shang CA, Levitt LK, Howard T, Waters MJ. Ternary complex factors Elk-1 and Sap-1a mediate growth hormone-induced transcription of egr-1 (early growth response factor-1) in 3T3-F442A preadipocytes. Mol Endocrinol. 1999 13:619-31.

Cler E, Papai G, Schultz P, Davidson I. Recent advances in understanding the structure and function of general transcription factor TFIID. Cell Mol Life Sci. 2009 66:2123-34.

Cleutjens KB, van der Korput HA, van Eekelen CC, van Rooij HC, Faber PW, Trapman J. An androgen response element in a far upstream enhancer region is essential for high, androgen-regulated activity of the prostate-specific antigen promoter. Mol Endocrinol. 1997 11:148-61.

Cleutjens KB, van Eekelen CC, van der Korput HA, Brinkmann AO, Trapman J. Two androgen response regions cooperate in steroid hormone regulated activity of the prostate-specific antigen promoter. J Biol Chem. 1996 271:6379-88.

Costello PS, Nicolas RH, Watanabe Y, Rosewell I, Treisman R. Ternary complex factor SAP-1 is required for Erk-mediated thymocyte positive selection. Nat Immunol. 2004 5:289-98.

Craft N, Shostak Y, Carey M, Sawyers CL. A mechanism for hormone-independent prostate cancer through modulation of androgen receptor signaling by the HER-2/neu tyrosine kinase. Nat Med. 1999 5:280-5.

Cremer M, von Hase J, Volm T, Brero A, Kreth G, Walter J, Fischer C, Solovei I, Cremer C, Cremer T. Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. Chromosome Res. 2001 9:541-67.

Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome Res. 2009 19:24-32.

Culig Z, Comuzzi B, Steiner H, Bartsch G, Hobisch A. Expression and function of androgen receptor coactivators in prostate cancer. J Steroid Biochem Mol Biol. 2004 92:265-71.

Culig Z, Klocker H, Bartsch G, Hobisch A. Androgen receptors in prostate cancer. Endocr Relat Cancer. 2002 9:155-70.

Dalton S, Treisman R. Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. Cell. 1992 68:597-612.

De Bosscher K, Haegeman G. Minireview: latest perspectives on anti-inflammatory actions of glucocorticoids. Mol Endocrinol. 2009 23:281-91.

de Ruijter AJ, van Gennip AH, Caron HN, Kemp S, van Kuilenburg AB. Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem J. 2003 370:737-49.

Dehm SM, Tindall DJ. Molecular regulation of androgen action in prostate cancer. J Cell Biochem. 2006 99:333-44.

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

Deng W, Roberts SG. TFIIB and the regulation of transcription by RNA polymerase II. Chromosoma. 2007 116:417-29.

Dong JT. Prevalent mutations in prostate cancer. J Cell Biochem. 2006 97:433-47.

Edwards J, Krishna NS, Mukherjee R, Bartlett JM. The role of c-Jun and c-Fos expression in androgen-independent prostate cancer. J Pathol. 2004 204:153-8.

Edwards S, Campbell C, Flohr P, Shipley J, Giddings I, Te-Poele R, Dodson A, Foster C, Clark J, Jhavar S, Kovacs G, Cooper CS. Expression analysis onto microarrays of randomly selected cDNA clones highlights HOXB13 as a marker of human prostate cancer. Br J Cancer. 2005 92:376-81.

Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S. Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. Science. 2007 318:1777-9.

Eichner J, Chen HT, Warfield L, Hahn S. Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. EMBO J. 2010 29:706-16.

Eivazova ER, Gavrilov A, Pirozhkova I, Petrov A, Iarovaia OV, Razin SV, Lipinski M, Vassetzky YS. Interaction in vivo between the two matrix attachment regions flanking a single chromatin loop. J Mol Biol. 2009 386:929-37.

Elcock LS, Bridger JM. Exploring the effects of a dysfunctional nuclear matrix. Biochem Soc Trans. 2008 36:1378-83.

ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007 447:799-816.

Esgueva R, Perner S, J LaFargue C, Scheble V, Stephan C, Lein M, Fritzsche FR, Dietel M, Kristiansen G, Rubin MA. Prevalence of TMPRSS2-ERG and SLC45A3-ERG gene fusions in a large prostatectomy cohort. Mod Pathol. 2010 23:539-46.

Esnault C, Ghavi-Helm Y, Brun S, Soutourina J, Van Berkum N, Boschiero C, Holstege F, Werner M. Mediator-dependent recruitment of TFIIH modules in preinitiation complex. Mol Cell. 2008 31:337-46.

Farrants AK. Chromatin remodelling and actin organisation. FEBS Lett. 2008 582:2041-50.

Fawcett DW. On the occurrence of a fibrous lamina on the inner aspect of the nuclear envelope in certain cells of vertebrates. Am J Anat. 1966 119:129-45.

Febbo PG, Lowenberg M, Thorner AR, Brown M, Loda M, Golub TR. Androgen mediated regulation and functional implications of fkbp51 expression in prostate cancer. J Urol. 2005 173:1772-7.

Finlan LE, Sproul D, Thomson I, Boyle S, Kerr E, Perry P, Ylstra B, Chubb JR, Bickmore WA. Recruitment to the nuclear periphery can alter expression of genes in human cells. PLoS Genet. 2008 4:e1000039.

Foradori CD, Weiser MJ, Handa RJ. Non-genomic actions of androgens. Front Neuroendocrinol. 2008 29:169-81.

Foster BA, Gingrich JR, Kwon ED, Madias C, Greenberg NM. Characterization of prostatic epithelial cell lines derived from transgenic adenocarcinoma of the mouse prostate (TRAMP) model. Cancer Res. 1997 57:3325-30.

Fraser P, Bickmore W. Nuclear organization of the genome and the potential for gene regulation. Nature. 2007 447:413-7.

Fuchs J, Demidov D, Houben A, Schubert I. Chromosomal histone modification patterns--from conservation to diversity. Trends Plant Sci. 2006 11:199-208.

Fullwood MJ, Ruan Y. ChIP-based methods for the identification of long-range chromatin interactions. J Cell Biochem. 2009 107:30-9.

Gaddipati JP, McLeod DG, Heidenberg HB, Sesterhenn IA, Finger MJ, Moul JW, Srivastava S. Frequent detection of codon 877 mutation in the androgen receptor gene in advanced prostate cancers. Cancer Res. 1994 54:2861-4.

Galande S, Purbey PK, Notani D, Kumar PP. The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. Curr Opin Genet Dev. 2007 17:408-14.

Galani A, Kitsiou-Tzeli S, Sofokleous C, Kanavakis E, Kalpini-Mavrou A. Androgen insensitivity syndrome: clinical features and molecular defects. Hormones (Athens). 2008 7:217-29.

Gao H, Lukin K, Ramírez J, Fields S, Lopez D, Hagman J. Opposing effects of SWI/SNF and Mi-2/NuRD chromatin remodeling complexes on epigenetic reprogramming by EBF and Pax5. Proc Natl Acad Sci U S A. 2009 106:11258-63.

Gao N, Zhang J, Rao MA, Case TC, Mirosevich J, Wang Y, Jin R, Gupta A, Rennie PS, Matusik RJ. The role of hepatocyte nuclear factor-3 alpha (Forkhead Box A1) and androgen receptor in transcriptional regulation of prostatic genes. Mol Endocrinol. 2003 17:1484-507.

Gao W, Bohl CE, Dalton JT. Chemistry and structural biology of androgen receptor. Chem Rev. 2005 105:3352-70.

Garcia-Bassets I, Kwon YS, Telese F, Prefontaine GG, Hutt KR, Cheng CS, Ju BG, Ohgi KA, Wang J, Escoubet-Lozach L, Rose DW, Glass CK, Fu XD, Rosenfeld MG. Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. Cell. 2007 128:505-18.

Gehring M, Reik W, Henikoff S. DNA demethylation by DNA repair. Trends Genet. 2009 25:82-90.

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

Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP, Bickmore WA. Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. Cell. 2004 118:555-66.

Gleason DF. Classification of prostatic carcinomas. Cancer Chemother Rep. 1966 50:125-8.

Golias Ch, Iliadis I, Peschos D, Charalabopoulos K. Amplification and co-regulators of androgen receptor gene in prostate cancer. Exp Oncol. 2009 31:3-8.

Gottlieb B, Beitel LK, Wu JH, Trifiro M. The androgen receptor gene mutations database (ARDB): 2004 update. Hum Mutat. 2004 23:527-33.

Greenberg NM, DeMayo F, Finegold MJ, Medina D, Tilley WD, Aspinall JO, Cunha GR, Donjacour AA, Matusik RJ, Rosen JM. Prostate cancer in a transgenic mouse. Proc Natl Acad Sci U S A. 1995 92:3439-43.

Gregory CW, He B, Johnson RT, Ford OH, Mohler JL, French FS, Wilson EM. A mechanism for androgen receptor-mediated prostate cancer recurrence after androgen deprivation therapy. Cancer Res. 2001 61:4315-9.

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. Cell. 2007 130:77-88.

Guglielmi B, Soutourina J, Esnault C, Werner M. TFIIS elongation factor and Mediator act in conjunction during transcription initiation in vivo. Proc Natl Acad Sci U S A. 2007 104:16062-7.

Gutiérrez JL, Chandy M, Carrozza MJ, Workman JL. Activation domains drive nucleosome eviction by SWI/SNF. EMBO J. 2007 26:730-40.

Göndör A, Ohlsson R. Chromosome crosstalk in three dimensions. Nature. 2009 461:212-7.

Hake SB, Allis CD. Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis". Proc Natl Acad Sci U S A. 2006 103:6428-35.

Han B, Mehra R, Dhanasekaran SM, Yu J, Menon A, Lonigro RJ, Wang X, Gong Y, Wang L, Shankar S, Laxman B, Shah RB, Varambally S, Palanisamy N, Tomlins SA, Kumar-Sinha C, Chinnaiyan AM. A fluorescence in situ hybridization screen for E26 transformation-specific aberrations: identification of DDX5-ETV4 fusion protein in prostate cancer. Cancer Res. 2008 68:7629-37.

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

Handwerger KE, Gall JG. Subnuclear organelles: new insights into form and function. Trends Cell Biol. 2006 16:19-26.

Hassa PO, Haenni SS, Elser M, Hottiger MO. Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? Microbiol Mol Biol Rev. 2006 70:789-829.

Hayashi-Takanaka Y, Yamagata K, Nozaki N, Kimura H. Visualizing histone modifications in living cells: spatiotemporal dynamics of H3 phosphorylation during interphase. J Cell Biol. 2009 187:781-90.

He B, Minges JT, Lee LW, Wilson EM. The FXXLF motif mediates androgen receptor-specific interactions with coregulators. J Biol Chem. 2002 277:10226-35.

Heemers HV, Tindall DJ. Androgen receptor (AR) coregulators: a diversity of functions converging on and regulating the AR transcriptional complex. Endocr Rev. 2007 28:778-808.

Heinlein CA, Chang C. Androgen receptor in prostate cancer. Endocr Rev. 2004 25:276-308.

Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenkov VV, Stewart R, Thomson JA, Crawford GE, Kellis M, Ren B. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature. 2009 459:108-12.

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007 39:311-8.

Heitzer MD, Wolf IM, Sanchez ER, Witchel SF, DeFranco DB. Glucocorticoid receptor physiology. Rev Endocr Metab Disord. 2007 8:321-30.

Helgeson BE, Tomlins SA, Shah N, Laxman B, Cao Q, Prensner JR, Cao X, Singla N, Montie JE, Varambally S, Mehra R, Chinnaiyan AM. Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. Cancer Res. 2008 68:73-80.

Heng HH, Goetze S, Ye CJ, Liu G, Stevens JB, Bremer SW, Wykes SM, Bode J, Krawetz SA. Chromatin loops are selectively anchored using scaffold/matrix-attachment regions. J Cell Sci. 2004 117:999-1008.

Henikoff S. Nucleosome destabilization in the epigenetic regulation of gene expression. Nat Rev Genet. 2008 9:15-26.

Hernandez-Verdun D. Nucleolus: from structure to dynamics. Histochem Cell Biol. 2006 125:127-37.

Hogan C, Varga-Weisz P. The regulation of ATP-dependent nucleosome remodeling factors. Mutat Res. 2007 618:41-51.

Horn PJ, Peterson CL. Molecular biology. Chromatin higher order folding--wrapping up transcription. Science. 2002 297:1824-7.

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

Huang P, Chandra V, Rastinejad F. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. Annu Rev Physiol. 2010 72:247-72.

Huang ZQ, Li J, Sachs LM, Cole PA, Wong J. A role for cofactor-cofactor and cofactor-histone interactions in targeting p300, SWI/SNF and Mediator for transcription. EMBO J. 2003 22:2146-55.

Hubler TR, Denny WB, Valentine DL, Cheung-Flynn J, Smith DF, Scammell JG. The FK506-binding immunophilin FKBP51 is transcriptionally regulated by progestin and attenuates progestin responsiveness. Endocrinology. 2003 144:2380-7.

Hubler TR, Scammell JG. Intronic hormone response elements mediate regulation of FKBP5 by progestins and glucocorticoids. Cell Stress Chaperones. 2004 9:243-52.

Huggins C, Hodges CV. Studies on prostatic cancer. I. The effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate. Cancer res. 1941 1:293-297.

Hur SK, Park EJ, Han JE, Kim YA, Kim JD, Kang D, Kwon J. Roles of human INO80 chromatin remodeling enzyme in DNA replication and chromosome segregation suppress genome instability. Cell Mol Life Sci. 2010 Mar 17 (epub)

Iljin K, Wolf M, Edgren H, Gupta S, Kilpinen S, Skotheim RI, Peltola M, Smit F, Verhaegh G, Schalken J, Nees M, Kallioniemi O. TMPRSS2 fusions with oncogenic ETS factors in prostate cancer involve unbalanced genomic rearrangements and are associated with HDAC1 and epigenetic reprogramming. Cancer Res. 2006 66:10242-6.

International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature. 2001 409:860-921.

Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun MJ. Cancer statistics, 2008. CA Cancer J Clin. 2008 58:71-96.

Jia L, Berman BP, Jariwala U, Yan X, Cogan JP, Walters A, Chen T, Buchanan G, Frenkel B, Coetzee GA. Genomic androgen receptor-occupied regions with different functions, defined by histone acetylation, coregulators and transcriptional capacity. PLoS One. 2008 3:e3645.

Jia L, Shen HC, Wantroba M, Khalid O, Liang G, Wang Q, Gentzschein E, Pinski JK, Stanczyk FZ, Jones PA, Coetzee GA. Locus-wide chromatin remodeling and enhanced androgen receptor-mediated transcription in recurrent prostate tumor cells. Mol Cell Biol. 2006 26:7331-41.

Johansen KM, Johansen J. Regulation of chromatin structure by histone H3S10 phosphorylation. Chromosome Res. 2006 14:393-404.

Jones JO, Bolton EC, Huang Y, Feau C, Guy RK, Yamamoto KR, Hann B, Diamond MI. Non-competitive androgen receptor inhibition in vitro and in vivo. Proc Natl Acad Sci U S A. 2009 106:7233-8.

Ju BG, Lunyak VV, Perissi V, Garcia-Bassets I, Rose DW, Glass CK, Rosenfeld MG. A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. Science. 2006 312:1798-802.

Ju BG, Rosenfeld MG. A breaking strategy for topoisomerase IIbeta/PARP-1-dependent regulated transcription. Cell Cycle. 2006 5:2557-60.

Juven-Gershon T, Kadonaga JT. Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev Biol. 2010 339:225-9.

Kaarbø M, Klokk TI, Saatcioglu F. Androgen signaling and its interactions with other signaling pathways in prostate cancer. Bioessays. 2007 29:1227-38.

Kallio PJ, Poukka H, Moilanen A, Jänne OA, Palvimo JJ. Androgen receptor-mediated transcriptional regulation in the absence of direct interaction with a specific DNA element. Mol Endocrinol. 1995 9:1017-28.

Kangaspeska S, Stride B, Métivier R, Polycarpou-Schwarz M, Ibberson D, Carmouche RP, Benes V, Gannon F, Reid G. Transient cyclical methylation of promoter DNA. Nature. 2008 452:112-5.

Kariaho E, Juuti H, Hannula K, Hednäs P, Ruponen M, Tunderman P. Pharmaca Fennica 2010. Lääketietokeskus Oy. ISSN 0355-7472.

Kerkhofs S, Denayer S, Haelens A, Claessens F. Androgen receptor knockout and knock-in mouse models. J Mol Endocrinol. 2009 42:11-7.

Kim B, Nesvizhskii AI, Rani PG, Hahn S, Aebersold R, Ranish JA. The transcription elongation factor TFIIS is a component of RNA polymerase II preinitiation complexes. Proc Natl Acad Sci U S A. 2007 104:16068-73.

Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010 465:182-7.

Kim TS, Heinlein C, Hackman RC, Nelson PS. Phenotypic analysis of mice lacking the Tmprss2-encoded protease. Mol Cell Biol. 2006 26:965-75.

Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci. 2006 31:89-97.

Klose RJ, Yan Q, Tothova Z, Yamane K, Erdjument-Bromage H, Tempst P, Gilliland DG, Zhang Y, Kaelin WG Jr. The retinoblastoma binding protein RBP2 is an H3K4 demethylase. Cell. 2007 128:889-900.

Knudsen KE, Scher HI. Starving the addiction: new opportunities for durable suppression of AR signaling in prostate cancer. Clin Cancer Res. 2009 15:4792-8.

Koivisto P, Kononen J, Palmberg C, Tammela T, Hyytinen E, Isola J, Trapman J, Cleutjens K, Noordzij A, Visakorpi T, Kallioniemi OP. Androgen receptor gene amplification: a possible molecular mechanism for androgen deprivation therapy failure in prostate cancer. Cancer Res. 1997 57:314-9.

Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J. Differential chromatin marking of introns and expressed exons by H3K36me3. Nat Genet. 2009 41:376-81.

Korenchuk S, Lehr JE, MClean L, Lee YG, Whitney S, Vessella R, Lin DL, Pienta KJ. VCaP, a cell-based model system of human prostate cancer. In Vivo. 2001 15:163-8.

Kouzarides T. Chromatin modifications and their function. Cell. 2007a 128:693-705.

Kouzarides T. Snapshot: Histone-modifying enzymes. Cell. 2007b 131:822.

Kuil CW, Mulder E. Deoxyribonucleic acid-binding ability of androgen receptors in whole cells: implications for the actions of androgens and antiandrogens. Endocrinology. 1996 137:1870-7.

Kumar PP, Bischof O, Purbey PK, Notani D, Urlaub H, Dejean A, Galande S. Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. Nat Cell Biol. 2007 9:45-56.

Kumaran RI, Thakar R, Spector DL. Chromatin dynamics and gene positioning. Cell. 2008 132:929-34.

Kumar-Sinha C, Tomlins SA, Chinnaiyan AM. Recurrent gene fusions in prostate cancer. Nat Rev Cancer. 2008 8:497-511.

Lamont KR, Tindall DJ. Androgen regulation of gene expression. Adv Cancer Res. 2010 107:137-62.

Lanctôt C, Cheutin T, Cremer M, Cavalli G, Cremer T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat Rev Genet. 2007 8:104-15.

Lee JS, Shukla A, Schneider J, Swanson SK, Washburn MP, Florens L, Bhaumik SR, Shilatifard A. Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. Cell. 2007 131:1084-96.

Lee MG, Wynder C, Bochar DA, Hakimi MA, Cooch N, Shiekhattar R. Functional interplay between histone demethylase and deacetylase enzymes. Mol Cell Biol. 2006 26:6395-402.

Lenasi T, Barboric M. P-TEFb stimulates transcription elongation and pre-mRNA splicing through multilateral mechanisms. RNA Biol. 2010 7:(epub).

Li B, Carey M, Workman JL. The role of chromatin during transcription. Cell. 2007a 128:707-19.

Li B, Gogol M, Carey M, Lee D, Seidel C, Workman JL. Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. Science. 2007b 316:1050-4.

Lin B, Ferguson C, White JT, Wang S, Vessella R, True LD, Hood L, Nelson PS. Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. Cancer Res. 1999 59:4180-4.

Linja MJ, Porkka KP, Kang Z, Savinainen KJ, Jänne OA, Tammela TL, Vessella RL, Palvimo JJ, Visakorpi T. Expression of androgen receptor coregulators in prostate cancer. Clin Cancer Res. 2004 10:1032-40.

Linja MJ, Savinainen KJ, Saramäki OR, Tammela TL, Vessella RL, Visakorpi T. Amplification and overexpression of androgen receptor gene in hormone-refractory prostate cancer. Cancer Res. 2001 61:3550-5.

Linnemann AK, Platts AE, Doggett N, Gluch A, Bode J, Krawetz SA. Genomewide identification of nuclear matrix attachment regions: an analysis of methods. Biochem Soc Trans. 2007 35:612-7.

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009 462:315-22.

Liu W, Xie CC, Zhu Y, Li T, Sun J, Cheng Y, Ewing CM, Dalrymple S, Turner AR, Sun J, Isaacs JT, Chang BL, Zheng SL, Isaacs WB, Xu J. Homozygous deletions and recurrent amplifications implicate new genes involved in prostate cancer. Neoplasia. 2008 10:897-907.

Lotan TL, Epstein JI. Clinical implications of changing definitions within the Gleason grading system. Nat Rev Urol. 2010 7:136-42.

Ludérus ME, van Steensel B, Chong L, Sibon OC, Cremers FF, de Lange T. Structure, subnuclear distribution, and nuclear matrix association of the mammalian telomeric complex. J Cell Biol. 1996 135:867-81.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997 389:251-60.

Lyon MF, Hawkes SG. X-linked gene for testicular feminization in the mouse. Nature. 1970 227:1217-9.

Editors of the physician's database. Eturauhasen syöpä. Duodecim. 2009 125:2455-8.

MacLean HE, Chiu WS, Notini AJ, Axell AM, Davey RA, McManus JF, Ma C, Plant DR, Lynch GS, Zajac JD. Impaired skeletal muscle development and function in male, but not female, genomic androgen receptor knockout mice. FASEB J. 2008 22:2676-89.

Magee JA, Chang LW, Stormo GD, Milbrandt J. Direct, androgen receptor-mediated regulation of the FKBP5 gene via a distal enhancer element. Endocrinology. 2006 147:590-8.

Magklara A, Smith CL. A composite intronic element directs dynamic binding of the progesterone receptor and GATA-2. Mol Endocrinol. 2009 23:61-73.

Maher CA, Kumar-Sinha C, Cao X, Kalyana-Sundaram S, Han B, Jing X, Sam L, Barrette T, Palanisamy N, Chinnaiyan AM. Transcriptome sequencing to detect gene fusions in cancer. Nature. 2009 458:97-101.

Mahy NL, Perry PE, Gilchrist S, Baldock RA, Bickmore WA. Spatial organization of active and inactive genes and noncoding DNA within chromosome territories. J Cell Biol. 2002 157:579-89.

Malinowska K, Neuwirt H, Cavarretta IT, Bektic J, Steiner H, Dietrich H, Moser PL, Fuchs D, Hobisch A, Culig Z. Interleukin-6 stimulation of growth of prostate cancer in vitro and in vivo through activation of the androgen receptor. Endocr Relat Cancer. 2009 16:155-69.

Mani RS, Tomlins SA, Callahan K, Ghosh A, Nyati MK, Varambally S, Palanisamy N, Chinnaiyan AM. Induced chromosomal proximity and gene fusions in prostate cancer. Science. 2009 326:1230.

Métivier R, Gallais R, Tiffoche C, Le Péron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G, Benes V, Jeltsch A, Gannon F, Salbert G. Cyclical DNA methylation of a transcriptionally active promoter. Nature. 2008 452:45-50.

Metzger E, Imhof A, Patel D, Kahl P, Hoffmeyer K, Friedrichs N, Müller JM, Greschik H, Kirfel J, Ji S, Kunowska N, Beisenherz-Huss C, Günther T, Buettner R, Schüle R. Phosphorylation of histone H3T6 by PKCbeta(I) controls demethylation at histone H3K4. Nature. 2010 464:792-6.

Metzger E, Yin N, Wissmann M, Kunowska N, Fischer K, Friedrichs N, Patnaik D, Higgins JM, Potier N, Scheidtmann KH, Buettner R, Schüle R. Phosphorylation of histone H3 at threonine 11 establishes a novel chromatin mark for transcriptional regulation. Nat Cell Biol. 2008 10:53-60.

Mika S, Rost B. NMPdb: Database of Nuclear Matrix Proteins. Nucleic Acids Res. 2005 33(Database issue):D160-3.

Minsky N, Shema E, Field Y, Schuster M, Segal E, Oren M. Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. Nat Cell Biol. 2008 10:483-8.

Mirosevich J, Gao N, Gupta A, Shappell SB, Jove R, Matusik RJ. Expression and role of Foxa proteins in prostate cancer. Prostate. 2006 66:1013-28.

Mirosevich J, Gao N, Matusik RJ. Expression of Foxa transcription factors in the developing and adult murine prostate. Prostate. 2005 62:339-52.

Mo Y, Ho W, Johnston K, Marmorstein R. Crystal structure of a ternary SAP-1/SRF/c-fos SRE DNA complex. J Mol Biol. 2001 314:495-506.

Montgomery RB, Mostaghel EA, Vessella R, Hess DL, Kalhorn TF, Higano CS, True LD, Nelson PS. Maintenance of intratumoral androgens in metastatic prostate cancer: a mechanism for castration-resistant tumor growth. Cancer Res. 2008 68:4447-54.

Morris CA, Moazed D. Centromere assembly and propagation. Cell. 2007 128:647-50.

Morris KV. RNA-directed transcriptional gene silencing and activation in human cells. Oligonucleotides. 2009 19:299-306.

Mosammaparast N, Shi Y. Reversal of Histone Methylation: Biochemical and Molecular Mechanisms of Histone Demethylases. Annu Rev Biochem. 2010 Apr 7 (epub).

Mostaghel EA, Page ST, Lin DW, Fazli L, Coleman IM, True LD, Knudsen B, Hess DL, Nelson CC, Matsumoto AM, Bremner WJ, Gleave ME, Nelson PS. Intraprostatic androgens and androgen-regulated gene expression persist after testosterone suppression: therapeutic implications for castration-resistant prostate cancer. Cancer Res. 2007 67:5033-41.

Motorin Y, Lyko F, Helm M. 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. Nucleic Acids Res. 2010 38:1415-30.

Nathan D, Ingvarsdottir K, Sterner DE, Bylebyl GR, Dokmanovic M, Dorsey JA, Whelan KA, Krsmanovic M, Lane WS, Meluh PB, Johnson ES, Berger SL. Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. Genes Dev. 2006 20:966-76.

Nelson CJ, Santos-Rosa H, Kouzarides T. Proline isomerization of histone H3 regulates lysine methylation and gene expression. Cell. 2006 126:905-16.

Neumann H, Hancock SM, Buning R, Routh A, Chapman L, Somers J, Owen-Hughes T, van Noort J, Rhodes D, Chin JW. A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. Mol Cell. 2009 36:153-63.

Ngan S, Stronach EA, Photiou A, Waxman J, Ali S, Buluwela L. Microarray coupled to quantitative RT-PCR analysis of androgen-regulated genes in human LNCaP prostate cancer cells. Oncogene. 2009 28:2051-63.

Ni L, Yang CS, Gioeli D, Frierson H, Toft DO, Paschal BM. FKBP51 promotes assembly of the Hsp90 chaperone complex and regulates androgen receptor signaling in prostate cancer cells. Mol Cell Biol. 2010 30:1243-53.

Nickerson J. Experimental observations of a nuclear matrix. J Cell Sci. 2001 114:463-74.

Noordermeer D, Branco MR, Splinter E, Klous P, van Ijcken W, Swagemakers S, Koutsourakis M, van der Spek P, Pombo A, de Laat W. Transcription and chromatin organization of a housekeeping gene cluster containing an integrated beta-globin locus control region. PLoS Genet. 2008 4:e1000016.

Nunez E, Fu XD, Rosenfeld MG. Nuclear organization in the 3D space of the nucleus - cause or consequence? Curr Opin Genet Dev. 2009 19:424-36.

Oikawa T, Yamada T. Molecular biology of the Ets family of transcription factors. Gene. 2003 303:11-34.

O'Malley BW, Towle HC, Schwartz RJ. Regulation of gene expression in eucaryotes. Annu Rev Genet. 1977 11:239-75.

O'Malley BW. Coregulators: from whence came these "master genes". Mol Endocrinol. 2007 21:1009-13.

Paakinaho V. Activation of FKBP51 locus by glucocorticoid receptor in human lung cancer cells. Master's thesis. University of Eastern Finland.

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

Palvimo JJ, Reinikainen P, Ikonen T, Kallio PJ, Moilanen A, Jänne OA. Mutual transcriptional interference between RelA and androgen receptor. J Biol Chem. 1996 271:24151-6.

Pan Y, Tsai CJ, Ma B, Nussinov R. Mechanisms of transcription factor selectivity. Trends Genet. 2010 26:75-83.

Patra SK, Patra A, Rizzi F, Ghosh TC, Bettuzzi S. Demethylation of (Cytosine-5-C-methyl) DNA and regulation of transcription in the epigenetic pathways of cancer development. Cancer Metastasis Rev. 2008 27:315-34.

Periyasamy S, Hinds T Jr, Shemshedini L, Shou W, Sanchez ER. FKBP51 and Cyp40 are positive regulators of androgen-dependent prostate cancer cell growth and the targets of FK506 and cyclosporin A. Oncogene. 2010 29:1691-701.

Phillips JE, Corces VG. CTCF: master weaver of the genome. Cell. 2009 137:1194-211.

Portin P. The elusive concept of the gene. Hereditas. 2009 146:112-7.

Powell SM, Brooke GN, Whitaker HC, Reebye V, Gamble SC, Chotai D, Dart DA, Belandia B, Bevan CL. Mechanisms of androgen receptor repression in prostate cancer. Biochem Soc Trans. 2006 34:1124-7.

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

Privalsky ML. The role of corepressors in transcriptional regulation by nuclear hormone receptors. Annu Rev Physiol. 2004 66:315-60.

Reineke EL, Liu H, Lam M, Liu Y, Kao HY. Aberrant association of promyelocytic leukemia protein-retinoic acid receptor-alpha with coactivators contributes to its ability to regulate gene expression. J Biol Chem. 2007 282:18584-96.

Reynolds PD, Ruan Y, Smith DF, Scammell JG. Glucocorticoid resistance in the squirrel monkey is associated with overexpression of the immunophilin FKBP51. J Clin Endocrinol Metab. 1999 84:663-9.

Rickman DS, Pflueger D, Moss B, VanDoren VE, Chen CX, de la Taille A, Kuefer R, Tewari AK, Setlur SR, Demichelis F, Rubin MA. SLC45A3-ELK4 is a novel and frequent erythroblast transformation-specific fusion transcript in prostate cancer. Cancer Res. 2009 69:2734-8.

Rosenfeld MG, Lunyak VV, Glass CK. Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. Genes Dev. 2006 20:1405-28.

Sáenz Robles MT, Pipas JM. T antigen transgenic mouse models. Semin Cancer Biol. 2009 19:229-35.

Salah Z, Maoz M, Pizov G, Bar-Shavit R. Transcriptional regulation of human protease-activated receptor 1: a role for the early growth response-1 protein in prostate cancer. Cancer Res. 2007 67:9835-43.

Santaguida S, Musacchio A. The life and miracles of kinetochores. EMBO J. 2009 28:2511-31.

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T. Active genes are tri-methylated at K4 of histone H3. Nature. 2002 419:407-11.

Schauwaers K, De Gendt K, Saunders PT, Atanassova N, Haelens A, Callewaert L, Moehren U, Swinnen JV, Verhoeven G, Verrijdt G, Claessens F. Loss of androgen receptor binding to selective androgen response elements causes a reproductive phenotype in a knockin mouse model. Proc Natl Acad Sci U S A. 2007 104:4961-6.

Scher HI, Beer TM, Higano CS, Anand A, Taplin ME, Efstathiou E, Rathkopf D, Shelkey J, Yu EY, Alumkal J, Hung D, Hirmand M, Seely L, Morris MJ, Danila DC, Humm J, Larson S, Fleisher M, Sawyers CL; Prostate Cancer Foundation/Department of Defense Prostate Cancer Clinical Trials Consortium. Antitumour activity of MDV3100 in castration-resistant prostate cancer: a phase 1-2 study. Lancet. 2010 375:1437-46.

Schneider R, Grosschedl R. Dynamics and interplay of nuclear architecture, genome organization, and gene expression. Genes Dev. 2007 21:3027-43.

Schuur ER, Henderson GA, Kmetec LA, Miller JD, Lamparski HG, Henderson DR. Prostate-specific antigen expression is regulated by an upstream enhancer. J Biol Chem. 1996 271:7043-51.

Shaulian E, Karin M. AP-1 as a regulator of cell life and death. Nat Cell Biol. 2002 4:E131-6.

Shema E, Tirosh I, Aylon Y, Huang J, Ye C, Moskovits N, Raver-Shapira N, Minsky N, Pirngruber J, Tarcic G, Hublarova P, Moyal L, Gana-Weisz M, Shiloh Y, Yarden Y, Johnsen SA, Vojtesek B, Berger SL, Oren M. The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. Genes Dev. 2008 22:2664-76.

Shi XB, Xue L, Zou JX, Gandour-Edwards R, Chen H, deVere White RW. Prolonged androgen receptor loading onto chromatin and the efficient recruitment of p160 coactivators contribute to androgen-independent growth of prostate cancer cells. Prostate. 2008 68:1816-26.

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell. 2004 119:941-53.

Shiio Y, Eisenman RN. Histone sumoylation is associated with transcriptional repression. Proc Natl Acad Sci U S A. 2003 100:13225-30.

Sircar K, Yoshimoto M, Monzon FA, Koumakpayi IH, Katz RL, Khanna A, Alvarez K, Chen G, Darnel AD, Aprikian AG, Saad F, Bismar TA, Squire JA. PTEN genomic deletion is associated with p-Akt and AR signalling in poorer outcome, hormone refractory prostate cancer. J Pathol. 2009 218:505-13.

Skalnik DG, Dorfman DM, Williams DA, Orkin SH. Restriction of neuroblastoma to the prostate gland in transgenic mice. Mol Cell Biol. 1991 11:4518-27.

Sorrells SF, Sapolsky RM. An inflammatory review of glucocorticoid actions in the CNS. Brain Behav Immun. 2007 21:259-72.

Spijker AT, van Rossum EF. Glucocorticoid receptor polymorphisms in major depression. Focus on glucocorticoid sensitivity and neurocognitive functioning. Ann N Y Acad Sci. 2009 1179:199-215.

Squire JA. TMPRSS2-ERG and PTEN loss in prostate cancer. Nat Genet. 2009 41:509-10.

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009 324:930-5.

Talbert PB, Henikoff S. Histone variants--ancient wrap artists of the epigenome. Nat Rev Mol Cell Biol. 2010 11:264-75.

Tanaka A, Watanabe T, Iida Y, Hanaoka F, Ohkuma Y. Central forkhead domain of human TFIIE beta plays a primary role in binding double-stranded DNA at transcription initiation. Genes Cells. 2009 14:395-405.

Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, Menon A, Jing X, Cao Q, Han B, Yu J, Wang L, Montie JE, Rubin MA, Pienta KJ, Roulston D, Shah RB, Varambally S, Mehra R, Chinnaiyan AM. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature. 2007 448:595-9.

Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE, Cao Q, Prensner JR, Rubin MA, Shah RB, Mehra R, Chinnaiyan AM. Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia. 2008 10:177-88.

Tomlins SA, Mehra R, Rhodes DR, Smith LR, Roulston D, Helgeson BE, Cao X, Wei JT, Rubin MA, Shah RB, Chinnaiyan AM. TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. Cancer Res. 2006 66:3396-400.

Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, Shah RB, Pienta KJ, Rubin MA, Chinnaiyan AM. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science. 2005 310:644-8.

Tomlins SA, Rubin MA, Chinnaiyan AM. Integrative biology of prostate cancer progression. Annu Rev Pathol. 2006 1:243-71.

Tran C, Ouk S, Clegg NJ, Chen Y, Watson PA, Arora V, Wongvipat J, Smith-Jones PM, Yoo D, Kwon A, Wasielewska T, Welsbie D, Chen CD, Higano CS, Beer TM, Hung DT, Scher HI, Jung ME, Sawyers CL. Development of a second-generation antiandrogen for treatment of advanced prostate cancer. Science. 2009 324:787-90.

Tu JJ, Rohan S, Kao J, Kitabayashi N, Mathew S, Chen YT. Gene fusions between TMPRSS2 and ETS family genes in prostate cancer: frequency and transcript variant analysis by RT-PCR and FISH on paraffin-embedded tissues. Mod Pathol. 2007 20:921-8.

Turner BM. Defining an epigenetic code. Nat Cell Biol. 2007 9:2-6.

Vaarala MH, Porvari KS, Kellokumpu S, Kyllönen AP, Vihko PT. Expression of transmembrane serine protease TMPRSS2 in mouse and human tissues. J Pathol. 2001 193:134-40.

Walsh PC. Chemoprevention of prostate cancer. N Engl J Med. 2010 362:1237-8.

Waltering KK, Helenius MA, Sahu B, Manni V, Linja MJ, Jänne OA, Visakorpi T. Increased expression of androgen receptor sensitizes prostate cancer cells to low levels of androgens. Cancer Res. 2009 69:8141-9.

van Bokhoven A, Varella-Garcia M, Korch C, Johannes WU, Smith EE, Miller HL, Nordeen SK, Miller GJ, Lucia MS. Molecular characterization of human prostate carcinoma cell lines. Prostate. 2003 57:205-25.

van Riggelen J, Buchwalter G, Soto U, De-Castro Arce J, Hausen HZ, Wasylyk B, Rösl F. Loss of net as repressor leads to constitutive increased c-fos transcription in cervical cancer cells. J Biol Chem. 2005 280:3286-94.

Wang J, Cai Y, Yu W, Ren C, Spencer DM, Ittmann M. Pleiotropic biological activities of alternatively spliced TMPRSS2/ERG fusion gene transcripts. Cancer Res. 2008a 68:8516-24.

Wang Q, Li W, Liu XS, Carroll JS, Jänne OA, Keeton EK, Chinnaiyan AM, Pienta KJ, Brown M. A hierarchical network of transcription factors governs androgen receptor-dependent prostate cancer growth. Mol Cell. 2007 27:380-92.

Wang Q, Li W, Zhang Y, Yuan X, Xu K, Yu J, Chen Z, Beroukhim R, Wang H, Lupien M, Wu T, Regan MM, Meyer CA, Carroll JS, Manrai AK, Jänne OA, Balk SP, Mehra R, Han B, Chinnaiyan AM, Rubin MA, True L, Fiorentino M, Fiore C, Loda M, Kantoff PW, Liu XS, Brown M. Androgen receptor regulates a distinct transcription program in androgen-independent prostate cancer. Cell. 2009a 138:245-56.

Wang RS, Yeh S, Tzeng CR, Chang C. Androgen receptor roles in spermatogenesis and fertility: lessons from testicular cell-specific androgen receptor knockout mice. Endocr Rev. 2009b 30:119-32.

Wang Y, Wysocka J, Sayegh J, Lee YH, Perlin JR, Leonelli L, Sonbuchner LS, McDonald CH, Cook RG, Dou Y, Roeder RG, Clarke S, Stallcup MR, Allis CD, Coonrod SA. Human PAD4 regulates histone arginine methylation levels via demethylimination. Science. 2004 306:279-83.

Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ, Zhao K. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet. 2008b 40:897-903.

Washington MN, Weigel NL. 1{alpha},25-Dihydroxyvitamin D3 inhibits growth of VCaP prostate cancer cells despite inducing the growth-promoting TMPRSS2:ERG gene fusion. Endocrinology. 2010 151:1409-17.

Weake VM, Workman JL. Histone ubiquitination: triggering gene activity. Mol Cell. 2008 29:653-63.

Weaver JR, Susiarjo M, Bartolomei MS. Imprinting and epigenetic changes in the early embryo. Mamm Genome. 2009 20:532-43.

Velasco AM, Gillis KA, Li Y, Brown EL, Sadler TM, Achilleos M, Greenberger LM, Frost P, Bai W, Zhang Y. Identification and validation of novel androgen-regulated genes in prostate cancer. Endocrinology. 2004 145:3913-24.

Veldscholte J, Ris-Stalpers C, Kuiper GG, Jenster G, Berrevoets C, Claassen E, van Rooij HC, Trapman J, Brinkmann AO, Mulder E. A mutation in the ligand binding domain of the androgen receptor of human LNCaP cells affects steroid binding characteristics and response to anti-androgens. Biochem Biophys Res Commun. 1990 173:534-40.

Verdone L, Agricola E, Caserta M, Di Mauro E. Histone acetylation in gene regulation. Brief Funct Genomic Proteomic. 2006 5:209-21.

Vermeer H, Hendriks-Stegeman BI, van der Burg B, van Buul-Offers SC, Jansen M. Glucocorticoid-induced increase in lymphocytic FKBP51 messenger ribonucleic acid expression: a potential marker for glucocorticoid sensitivity, potency, and bioavailability. J Clin Endocrinol Metab. 2003 88:277-84.

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

Verrijdt G, Tanner T, Moehren U, Callewaert L, Haelens A, Claessens F. The androgen receptor DNA-binding domain determines androgen selectivity of transcriptional response. Biochem Soc Trans. 2006 34:1089-94.

Vickers ER, Kasza A, Kurnaz IA, Seifert A, Zeef LA, O'donnell A, Hayes A, Sharrocks AD. Ternary complex factor-serum response factor complex-regulated gene activity is required for cellular proliferation and inhibition of apoptotic cell death. Mol Cell Biol. 2004 24:10340-51.

Wissmann M, Yin N, Müller JM, Greschik H, Fodor BD, Jenuwein T, Vogler C, Schneider R, Günther T, Buettner R, Metzger E, Schüle R. Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. Nat Cell Biol. 2007 9:347-53.

Volpi EV, Chevret E, Jones T, Vatcheva R, Williamson J, Beck S, Campbell RD, Goldsworthy M, Powis SH, Ragoussis J, Trowsdale J, Sheer D. Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. J Cell Sci. 2000 113:1565-76.

Woodruff PG, Boushey HA, Dolganov GM, Barker CS, Yang YH, Donnelly S, Ellwanger A, Sidhu SS, Dao-Pick TP, Pantoja C, Erle DJ, Yamamoto KR, Fahy JV. Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. Proc Natl Acad Sci U S A. 2007 104:15858-63.

Workman JL. Nucleosome displacement in transcription. Genes Dev. 2006 20:2009-17.

Xhemalce B, Kouzarides T. A chromodomain switch mediated by histone H3 Lys 4 acetylation regulates heterochromatin assembly. Genes Dev. 2010 24:647-52.

Xin L, Teitell MA, Lawson DA, Kwon A, Mellinghoff IK, Witte ON. Progression of prostate cancer by synergy of AKT with genotropic and nongenotropic actions of the androgen receptor. Proc Natl Acad Sci U S A. 2006 103:7789-94.

Yamazaki Y, Kubota H, Nozaki M, Nagata K. Transcriptional regulation of the cytosolic chaperonin theta subunit gene, Cctq, by Ets domain transcription factors Elk-1, Sap-1a, and Net in the absence of serum response factor. J Biol Chem. 2003 278:30642-51.

Yeh S, Tsai MY, Xu Q, Mu XM, Lardy H, Huang KE, Lin H, Yeh SD, Altuwaijri S, Zhou X, Xing L, Boyce BF, Hung MC, Zhang S, Gan L, Chang C. Generation and characterization of androgen receptor knockout (ARKO) mice: an in vivo model for the study of androgen functions in selective tissues. Proc Natl Acad Sci U S A. 2002 99:13498-503.

Yong W, Yang Z, Periyasamy S, Chen H, Yucel S, Li W, Lin LY, Wolf IM, Cohn MJ, Baskin LS, Sánchez ER, Shou W. Essential role for Co-chaperone Fkbp52 but not Fkbp51 in androgen receptor-mediated signaling and physiology. J Biol Chem. 2007 282:5026-36.

Yoshimoto M, Joshua AM, Chilton-Macneill S, Bayani J, Selvarajah S, Evans AJ, Zielenska M, Squire JA. Three-color FISH analysis of TMPRSS2/ERG fusions in prostate cancer indicates that genomic microdeletion of chromosome 21 is associated with rearrangement. Neoplasia. 2006 8:465-9.

Young RA. RNA polymerase II. Annu Rev Biochem. 1991 60:689-715.

Yu J, Yu J, Mani RS, Cao Q, Brenner CJ, Cao X, Wang X, Wu L, Li J, Hu M, Gong Y, Cheng H, Laxman B, Vellaichamy A, Shankar S, Li Y, Dhanasekaran SM, Morey R, Barrette T, Lonigro RJ, Tomlins SA, Varambally S, Qin ZS, Chinnaiyan AM. An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. Cancer Cell. 2010 17:443-54.

Yu YP, Landsittel D, Jing L, Nelson J, Ren B, Liu L, McDonald C, Thomas R, Dhir R, Finkelstein S, Michalopoulos G, Becich M, Luo JH. Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy. J Clin Oncol. 2004 22:2790-9.

Yu Z, Dadgar N, Albertelli M, Gruis K, Jordan C, Robins DM, Lieberman AP. Androgen-dependent pathology demonstrates myopathic contribution to the Kennedy disease phenotype in a mouse knock-in model. J Clin Invest. 2006 116:2663-72.

Zhou W, Zhu P, Wang J, Pascual G, Ohgi KA, Lozach J, Glass CK, Rosenfeld MG. Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. Mol Cell. 2008 29:69-80.

Zhou X. Roles of Androgen Receptor in Male and Female Reproduction: Lessons from Global and Cell Specific Androgen Receptor Knockout (ARKO) Mice. J Androl. 2009 Dec 3 (epub).

Zhu ML, Kyprianou N. Androgen receptor and growth factor signaling cross-talk in prostate cancer cells. Endocr Relat Cancer. 2008 15:841-9.

Zhu P, Zhou W, Wang J, Puc J, Ohgi KA, Erdjument-Bromage H, Tempst P, Glass CK, Rosenfeld MG. A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation. Mol Cell. 2007 27:609-21.

www.ncbi.nml.nih.gov, www.cancer.fi, www.cancerregistry.fi, www.genomatix.de, genome.ucsc.edu