MOLECULAR CHARACTERIZATION OF ESTROGEN RECEPTORS WITH FOCUS ON BREAST CANCER

Milica Putnik
MOLECULAR CHARACTERIZATION OF ESTROGEN RECEPTORS WITH FOCUS ON BREAST CANCER

Milica Putnik

Stockholm 2012
To my mother,
and other heroes,
who fight the Breast Beast
ABSTRACT

Estrogen signaling is mediated by estrogen receptors (ERs), ERα and ERβ. Aberrant estrogen signaling is involved in breast cancer development. ERα is one of the key biomarkers for diagnosis and treatment of breast cancer. Unlike ERα, ERβ is still not introduced as a marker for diagnosis and established as a target of therapy. Numerous studies suggest antiproliferative effects of ERβ, however its role remains to be fully explored. Albeit important, ERα is not a perfect marker, and some aspects of ERα function are still unclear. This thesis aims to characterize distinct molecular facets of ER action relevant for breast cancer and provide valuable information for ER-based diagnosis and treatment design.

In PAPER I, we analyzed the functionality of two common single nucleotide polymorphisms in the 3’ untranslated regions of ERβ, rs4986938 and rs928554, which have been extensively investigated for association with various diseases. A significant difference in allelic expression was observed for rs4986938 in breast tumor samples from heterozygous individuals. However, no difference in mRNA stability or translatability between the alleles was observed.

In PAPER II, we provided a more comprehensive understanding of ERβ function independent of ERα. A global gene expression analysis in a HEK293/ERβ cell model identified a set of ERβ-regulated genes. Gene Ontology (GO) analysis showed that they are involved in cell-cell signaling, morphogenesis and cell proliferation. Moreover, ERβ expression resulted in a significant decrease in cell proliferation.

In PAPER III, using the human breast cancer MCF-7/ERβ cell model, we demonstrated, for the first time, the binding of ERα/β heterodimers to various DNA-binding regions in intact chromatin.

In PAPER IV, we investigated a potential cross-talk between estrogen signaling and DNA methylation by identifying their common target genes in MCF-7 cells. Gene expression profiling identified around 150 genes regulated by both 17β-estradiol (E2) and a hypomethylating agent 5-aza-2'-deoxycytidine. Based on GO analysis, CpG island prediction analysis and previously reported ER binding regions, we selected six genes for further analysis. We identified BTG3 and FHL2 as direct target genes of both pathways. However, our data did not support a direct molecular interplay of mediators of estrogen and epigenetic signaling at promoters of regulated genes.

In PAPER V, we further explored the interactions between estrogen signaling and DNA methylation, with focus on DNA methyltransferases (DNMT1, DNMT3a and DNMT3b). E2, via ERα, up-regulated DNMT1 and down-regulated DNMT3a and DNMT3b mRNA expression. Furthermore, DNMT3b interacted with ERα. siRNA-mediated DNMT3b depletion increased the expression of two genes, CDKN1A and FHL2. We proposed that the molecular mechanism underlying regulation of FHL2 and CDKN1A gene expression involves interplay of DNMT3b and ERα.

In conclusion, the studies presented in this thesis contribute to the knowledge of ERβ function, and give additional insight into the cross-talk mechanisms underlying ERα signaling with ERβ and with DNA methylation pathways.
LIST OF PUBLICATIONS

I. **Putnik M**, Zhao C, Gustafsson J-Å, Dahlman-Wright K
   *Effects of two common polymorphisms in the 3’ untranslated regions of estrogen receptor beta on mRNA stability and translatability.*
   BMC Genetics, 2009, 10:55

II. Zhao C, **Putnik M**, Gustafsson J-Å, Dahlman-Wright K
    *Microarray analysis of altered gene expression in ERbeta-overexpressing HEK293 cells.*
    Endocrine, 2009, 36(2), 224-32

III. Papoutsi Z, Zhao C, **Putnik M**, Gustafsson J-Å, Dahlman-Wright K
    *Binding of estrogen receptor α/β heterodimers to chromatin in MCF-7 cells.*
    J Mol Endocrinol, 2009, 43(2), 65-72

IV. **Putnik M**, Zhao C, Gustafsson J-Å, Dahlman-Wright K
    *Gene regulation by estrogen signaling and DNA methylation in MCF-7 breast cancer cells.*
    Manuscript

V. **Putnik M**, Zhao C, Dahlman-Wright K
    *Interplay of estrogen receptor α and DNA methyltransferases with focus on DNMT3b.*
    Manuscript

*Publication not included in the thesis*

Lundholm L, **Putnik M**, Otsuki M, Andersson S, Ohlsson C, Gustafsson JA, Dahlman-Wright K
*Effects of estrogen on gene expression profiles in mouse hypothalamus and white adipose tissue: target genes include glutathione peroxidase 3 and cell death-inducing DNA fragmentation factor, alpha-subunit-like effector A.*
J Endocrinol. 2008, 196(3), 547-57
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF</td>
<td>Activation function</td>
</tr>
<tr>
<td>AIs</td>
<td>Aromatase inhibitors</td>
</tr>
<tr>
<td>AP-1</td>
<td>Activator protein 1</td>
</tr>
<tr>
<td>BRCA</td>
<td>Breast cancer susceptibility protein</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>Co-IP</td>
<td>Co-immunoprecipitation</td>
</tr>
<tr>
<td>DAC</td>
<td>5-aza-2’-deoxycytidine</td>
</tr>
<tr>
<td>DBD</td>
<td>DNA-binding domain</td>
</tr>
<tr>
<td>DNMT</td>
<td>DNA methyltransferase</td>
</tr>
<tr>
<td>E2</td>
<td>17β-estradiol</td>
</tr>
<tr>
<td>ERE</td>
<td>Estrogen response element</td>
</tr>
<tr>
<td>ERK</td>
<td>Extracellular signal-regulated kinases</td>
</tr>
<tr>
<td>ERα</td>
<td>Estrogen receptor α</td>
</tr>
<tr>
<td>ERβ</td>
<td>Estrogen receptor β</td>
</tr>
<tr>
<td>ERβ2</td>
<td>Estrogen receptor β 2 (ERβcx)</td>
</tr>
<tr>
<td>FHL2</td>
<td>Four and a half LIM domains 2</td>
</tr>
<tr>
<td>FOXA1</td>
<td>Forkhead box protein A1</td>
</tr>
<tr>
<td>GO</td>
<td>Gene ontology</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-wide association studies</td>
</tr>
<tr>
<td>HDACs</td>
<td>Histone deacetyltransferases</td>
</tr>
<tr>
<td>HER2</td>
<td>Human epidermal growth factor receptor 2</td>
</tr>
<tr>
<td>LBD</td>
<td>Ligand-binding domain</td>
</tr>
<tr>
<td>LD</td>
<td>Linkage disequilibrium</td>
</tr>
<tr>
<td>MAP kinase</td>
<td>Mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MBD</td>
<td>Methyl-CpG-binding domain protein</td>
</tr>
<tr>
<td>MeCP2</td>
<td>Methyl CpG binding protein 2</td>
</tr>
<tr>
<td>miRNA</td>
<td>microRNA</td>
</tr>
<tr>
<td>NCoR1</td>
<td>Nuclear receptor corepressor 1</td>
</tr>
<tr>
<td>NR</td>
<td>Nuclear receptor</td>
</tr>
<tr>
<td>PR</td>
<td>Progesterone receptor</td>
</tr>
<tr>
<td>pS2</td>
<td>TFF1, trefoil factor 1</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering RNA</td>
</tr>
<tr>
<td>SMRT</td>
<td>Silencing Mediator for Retinoid and Thyroid-hormone receptors</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>Sp1</td>
<td>Specificity protein 1</td>
</tr>
<tr>
<td>SRC</td>
<td>Steroid receptor coactivator</td>
</tr>
<tr>
<td>TF</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>Tet</td>
<td>Tetracycline</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

1.1 BREAST CANCER

Breast cancer is the form of cancer with the highest global incidence [1]. Of all cancer types diagnosed in women, it has the highest incidence (22%) and mortality rate (13%), and presents one of the biggest medical challenges of the modern world. According to estimates of lifetime risk, about 12% of women in the general population will develop breast cancer sometime during their lives. Established risk factors are age, high mammographic density, older age at first child, low number of children, hormone replacement therapy, ethnicity, high body mass index, exposure to ionizing radiation, increased alcohol consumption, low physical activity and genetic factors [2]. Only 5-10% of breast cancer cases are hereditary. Mutations in BRCA1 and BRCA2 tumor-suppressor genes are the prime examples of hereditary causes of breast cancer. These mutations are uncommon, but associated with high risk of early development of breast and ovarian cancer [3, 4].

Breast cancer is a complex disease, with heterogeneous molecular background. This, in combination with individual variations in origin, grade and stage at diagnosis are reasons for the lack of a universal cure for breast cancer. Breast cancer classifications based on histopathology include tumor stage (size, invasiveness, metastatic status), grade and origin. Breast cancer can originate in the lobular or ductal epithelium of the mammary gland, which represents the basis of the histopathological classification to lobular and ductal carcinoma, respectively.

Mammary ductal carcinoma is the most common type of breast cancer in women [5]. The mammary gland duct epithelium is comprised of two distinct cell types, basal-like (myoepithelial) and luminal cells. Breast cancer can originate from either cell type and it can be classified by cellular features and molecular features (Table 1) [6, 7]. Myoepithelial cells form an outer layer of the mature mammary duct, surrounding an inner layer of milk-secreting luminal cells. It is believed that mammary stem cells reside in a basal position between these two cell types and give rise to progenitor cells and both lineages of fully differentiated cells (Figure 1) [8-10]. Female mammary gland growth, development and function are regulated by hormones and growth factors. Elevated levels of estrogen, progesterone and growth factors, such as epidermal-, fibroblast- and insulin-like growth factors, induce proliferation and differentiation of mammary stem cells, suggesting possible mechanisms of cancer development [11-13].

The majority of hormones and growth factors exert their biological functions by binding to specific proteins - receptors, which became crucial biomarkers for diagnosis and treatment of breast cancer. Tumor biomarkers are usually proteins measured either in serum, plasma or tumor tissue and they are used to identify individuals with increased predisposition to develop a cancer, screen for early malignancies and/or assist
in cancer diagnosis, therapeutic strategy and prognosis. **Estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2)** are the most significant biomarkers in breast cancer diagnostics and treatment. Hence, breast cancers are classified by ER, PR and HER2 status (Table 1) [6, 7, 14]. About 70% of breast cancer types belong to the ER+ luminal subtype. The luminal A subtype shows good prognosis and it is less invasive than the luminal B and ER- subtypes.

Table 1. Classification of breast cancer.

<table>
<thead>
<tr>
<th>Classification of breast cancer</th>
<th>ER+</th>
<th>ER-</th>
<th>Normal-like/ unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Luminal A</strong></td>
<td>ER+PR+HER2–</td>
<td>ER+PR+HER2±</td>
<td>ER–PR–HER2– (triple negative)</td>
</tr>
<tr>
<td><strong>Luminal B</strong></td>
<td>ER+PR+HER2–</td>
<td>ER+PR+HER2±</td>
<td>ER–PR–HER2– (triple negative)</td>
</tr>
<tr>
<td><strong>Basal-like</strong></td>
<td>ER–PR–HER2– (triple negative)</td>
<td>ER–PR–HER2+</td>
<td>?</td>
</tr>
<tr>
<td><strong>HER2+</strong></td>
<td>ER–PR–HER2– (triple negative)</td>
<td>ER–PR–HER2+</td>
<td>?</td>
</tr>
<tr>
<td><strong>Normal-like/ unclassified</strong></td>
<td>ER–PR–HER2– (triple negative)</td>
<td>ER–PR–HER2+</td>
<td>?</td>
</tr>
<tr>
<td><strong>Incidence rate</strong></td>
<td>56-61%</td>
<td>9-16%</td>
<td>8-20%</td>
</tr>
<tr>
<td><strong>5-year survival rate</strong></td>
<td>95%</td>
<td>50%</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Therapy</strong></td>
<td>Hormonal</td>
<td>Hormonal, Chemotherapy, HER2 blockers</td>
<td>Chemotherapy</td>
</tr>
</tbody>
</table>

In general, early detection and diagnosis, and surgery remain the first-line and most successful strategies in breast cancer management. However, adjuvant treatments are required in most cases, in order to increase the survival rate, eliminate (micro)metastases, or where surgery cannot be performed. These include radiotherapy and systemic therapies - chemotherapy and targeted treatments.
Chemotherapy of breast cancer includes treatment with cytotoxic drugs, which target inhibition of cell mitosis or promoting apoptosis. Chemotherapy decreases mortality rate by 20-30% in 10 years [15]. The advantage of chemotherapy is low incidence of resistance, due to rather general biological effects, which can however lead to a wide range of severe side-effects [16].

ER and HER2, the most significant biomarkers, were essential for development of targeted treatments of breast cancer - hormonal therapy and HER2 blockers, respectively. Hence, ER and HER2 status is crucial for choice of therapy. PR status is usually positively associated with that of ER, and is considered as a marker of functional ER.

Amplification and/or overexpression of the HER2 gene are associated with the pathogenesis and progression of the aggressive HER2+ types of breast cancer [17]. Thus, the HER2 protein, a transmembrane tyrosine kinase exposed on the cell surface, became an important biomarker and target of therapy. HER2 blockers, such as trastuzumab (trade name Herceptin) and pertuzumab (trade name Omnitarg), are monoclonal antibodies, binding selectively to HER2 and blocking the dimerization, required for HER2 function [18, 19]. It has been shown that one year of treatment with trastuzumab after adjuvant chemotherapy has a significant overall survival benefit after a median follow-up of 2 years [20]. However, the majority of patients with metastatic breast cancer who initially respond to trastuzumab develop resistance within one year of treatment [21, 22]. Furthermore, HER2-blockers are considered expensive and can cause cardiac dysfunction [23].

ER status is a most crucial marker for breast cancer classification and treatment. Most breast cancers are ER+ and those patients typically receive hormonal (endocrine) therapy after completion of chemotherapy. Hormonal therapy is used to block the estrogen production or estrogen receptor signaling.

1.2 ESTROGEN RECEPTORS

The ERs belong to the nuclear receptor (NR) superfamily of ligand-regulated transcription factors. ERα was the first ER to be characterized, and it is used as a marker for diagnosis and treatment of breast cancer [24-26]. ERβ, discovered in the mid 1990’s [27], remains less characterized including its relation to breast cancer development.

Both ERs are co-expressed in a number of tissues including the mammary gland, epididymis, thyroid, adrenal, bone, and certain regions of the brain [28]. However, within some of these tissues, such as the endometrium and the prostate, they are expressed in different cell types [29, 30].

ERα is a dominant ER in the uterus, liver, kidney, and heart, whereas ERβ is a dominant ER in the ovaries, prostate, lung, gastrointestinal tract, bladder, and
hematopoietic and central nervous system [31]. Estrogen signaling plays a critical role in many physiological processes, including regulation of development, growth and function of many organ systems in the body. Therefore, aberrations in estrogen signaling are associated not only with different types of cancer, including breast, endometrial and ovarian cancers, but also with diseases such as osteoporosis, depression and eating disorders [32].

1.2.1 Structural and functional organization

ERα and ERβ encoding genes are located on different chromosomes (6 and 14, respectively) [27, 33], having a typical NR domain organization and sharing relatively high protein domain homology. ER proteins include six functional domains referred to by letters A-F [34-36]. The A and B protein domains contain the ligand-independent transcription activation function-1 (AF-1), and sites for phosphorylation. The C domain contains the DNA-binding domain (DBD). The D domain contains nuclear localization sequences. The E domain is the ligand-binding domain (LBD) and it contains the ligand-dependent transcription activation function-2 (AF-2). The F domain is involved in coregulator recruitment. ERα and ERβ share 96% amino acid identity in the DBD, approximately 53% amino acid identity in the LBD and 30% or less in other domains, involved in transactivation and localization (Figure 2) [35]. The latter may explain some of the observed differences in transcriptional activation potential of the two receptors.

Activated estrogen receptors form dimers. It is assumed that in cell types where the two receptor subtypes are co-expressed, the formation of α/β heterodimers plays an important role in estrogen signaling, affecting patterns of gene regulation distinct from those regulated by the ER homodimers [28].

Both ER encoding genes have complex promoter structures. The ERα gene is transcribed from at least nine promoters (A, B, C, D, T2, T, E1, F and E), into multiple transcripts that can vary in their 5’ untranslated regions (5’UTRs) [37]. The significance of the multiple promoters in the ERα gene is still unclear. However, there are many examples of a tissue-specific usage of particular promoters and with associated production of different ERα mRNA variants. For instance, ERα transcripts derived from promoter B showed the highest expression in human breast cancer MCF-7 cells [38, 39], and promoters E1 and E2 are used predominantly in the liver [40, 41]. The ERβ gene is transcribed from at least two promoters, named 0K and 0N [42], giving rise to two different messenger RNAs, which display distinct tissue distribution [43].

Both ER full-length mRNAs are encoded by eight exons. Additionally, both ERs are expressed in a variety of isoforms, due to alternative splicing [34]. The most relevant ERα splice variants, with regard to functionality and occurrence, are ERα46 and ERαΔ3. ERα46 was named after its predicted molecular weight of 46 kDa and corresponds to a deletion of the first coding exon [44]. It inhibits the function of the wild-type ERα and forms dimers with ERα and β [45]. ERαΔ3 is a result of a deletion
of exon 3, hence it lacks part of the DNA-binding domain. It inhibits the function of ERα [46].

ERβ2, also known as ERβcx, is the best characterized ERβ isoform [47]. It utilizes an alternative last exon, and consequently encodes a variant receptor with an altered C-terminus. The amino acids corresponding to exon 8 are replaced with 26 unique amino acids, giving rise to an altered F domain and truncated protein [48]. ERβ2 has undetectable affinity for ligands and cannot activate transcription of an estrogen response element-driven reporter [47, 49]. ERβ2 has been found to bind ERα and inhibit ligand-induced ERα transcriptional activity, most likely by mediating ERα protein degradation [50, 51]. This suggests that ERβ2 has an important role in neutralizing the function of ERα, hence ERβ2 may be significant for diagnosis and treatment of breast cancer [52].

Figure 2. Structural organization of ERα and ERβ genes and proteins. For the genes, exons, including UTRs, are indicated with boxes, and introns as lines in between. For the protein, the numbers on the right indicate the total size of the protein in amino acids (aa) and kilodaltons (kDa). Functional protein domains are marked by letters A-F. Homology between ERα and ERβ domains is marked in %. The position of two SNPs studied in this thesis (PAPER I), rs4986938 and rs928554, is indicated.

1.2.2 Ligands

The main natural endogenous ER ligands (agonists) are estrogen hormones. The predominant estrogen in the body is 17β-estradiol (E2), secreted by the ovaries during the female reproductive period. E2 metabolites, estrone (E1) and estriol (E3) are also ER ligands. E1 is the main estrogen before puberty and after menopause and is synthesized by the ovaries and the adipose tissue. E3 is produced mainly by the placenta during pregnancy. Of all three estrogens, E2 has the highest affinity for ERα, and equal binding affinity for ERα and ERβ. E1 has preferential binding affinity for ERα over ERβ, whereas E3 has preferential binding affinity for ERβ over ERα [53].
Antiestrogens (ER antagonists) bind ERs in a manner similar to estrogens, but induce a different conformation of the ligand-binding domain [54, 55]. This results in a lack of recruitment of coactivators by the AF-2 domain. Most ER antagonists act by targeting the ER competitively, by binding and blocking access of other possible ligands. Some ER antagonists possess partial agonist activity, dependent on the cell type and tissue, and they are referred to as selective ER modulators (SERMs). The most common SERMs are tamoxifen and raloxifene. Fulvestrant/ICI 182,780, is a complete ER antagonist, and a SERD - selective estrogen receptor down-regulator. It binds to ER and inhibits its activity by nuclear export and degradation [56].

Relatively low ligand-binding domain homology between the two ERs allowed the development of ER subtype-selective ligands. PPT (propylpyrazole triol) is a commonly used synthetic ERα selective agonist, with a 410-fold relative binding affinity for ERα versus ERβ [57]. It shows E2-like properties in many different tissues [58]. DPN (diarylpropionitrile) is the most commonly used synthetic ERβ selective agonist, with 70-fold binding higher binding affinity for ERβ compared to ERα [59]. Some phytoestrogens, plant-derived compounds with steroid structure and estrogen-like properties, such as genistein and coumestrol, have higher affinities for ERβ than for ERα [60]. Selected ER ligands are presented in Figure 3.

![Figure 3. Chemical structures of representative ER agonists and antagonists. Chemical formulas are obtained from METLIN, Metabolite and Tandem MS Database [61].]
1.2.3 Signaling pathways

ER activation can either be ligand-dependent or independent. The classical model of ER action involves activation of the ER by a ligand, upon which the receptors form dimers and to bind specific DNA sequences, such as estrogen response elements (EREs). By transcription factor cross-talk, ligand-activated ERs can regulate transcription indirectly, through interaction with other transcription factors (TFs), such as members of the activating protein-1 (AP-1) and specificity protein-1 (Sp1) families [62-64]. By non-genomic mechanisms, ligand binds to ERs localized in the cell membrane, which leads to activation of signal transduction pathways in the cytoplasm [65, 66]. Ligand-independent pathways include receptor phosphorylation by growth factor signaling via activation of kinases, such as HER2-regulated mitogen-activated protein (MAP) kinases ERK1 and ERK2 [67] (Figure 4).

Figure 4. Simplified models of ER signaling pathways. 1. Classical model. Liganded nuclear ERs bind directly to EREs in target gene promoters. 2. Transcription factor cross-talk. Liganded nuclear ER complexes are tethered through protein-protein interactions to a TF-complex that contacts the target gene promoter containing TF-response element (TFRE). 3. Non-genomic mechanisms. Liganded extranuclear ER complexes activate protein-kinase cascades. 4. Ligand-independent pathways. Growth factors (GF) activate protein-kinase cascades via growth factor receptors in the membrane, leading to phosphorylation (P) and activation of nuclear ERs at EREs.

ER-mediated gene transcription also requires the recruitment of coregulatory proteins, which form complexes with ER through protein-protein interactions. Coregulators can be either coactivators, required for transcriptional activation, or corepressors, which are involved in decrease of the transcriptional activity.
The best characterized coactivators belong to the steroid receptor coactivator family (SRC), which binds to the AF-2 domain of the ER protein. SRC coactivators were found to recruit enzymes that mediate histone acetylation and methylation, associated with an open chromatin structure and increased transcriptional output [68, 69]. Antagonist-bound receptors interact with corepressors, such as nuclear receptor corepressor 1 (NCoR1) and silencing mediator of retinoid and thyroid hormone receptors (SMRT). NCoR1 and SMRT in turn recruit large repressor complexes including histone deacetylases (HDACs) that repress gene activity by maintaining or reinforcing a repressive chromatin state [70, 71].

1.3 ESTROGEN RECEPTOR SIGNALING AND BREAST CANCER

Normally, ERα is found to be expressed in only 7-10% of luminal cells in the mammary gland, although its levels vary during the menstrual cycle [72-74]. ERα levels are often found to be increased in breast cancer and it is regarded as a positive regulator of cell proliferation [75]. However, high ERα expression is considered a marker of good prognosis, since such patients (ER+) have higher chance to benefit from antiestrogen therapies and have an increased survival rate. DNA methylation of the ERα promoter has been proposed as a mechanism for the development of ER– cell lines as well as primary tumors [76-78]. Patients diagnosed with ER– breast cancer have generally poor survival rate, increased metastasis and relapse occurrence [79].

ERβ is found to be expressed in 80-85% of the cells in the normal breast tissue, and hence it is considered to be a predominant ER in breast [34]. Yet the role of ERβ in the breast remains largely unclear, and it is still not introduced as a marker and a target in diagnostics and treatment. Reported antiproliferative effects, interplay with ERα and other mechanistic features of ERβ have a potential to contribute to our understanding of breast cancer and improvement of current treatments. Protein assays generally suggest that the ERβ protein expression is a favorable prognostic factor, correlating with known biomarkers such as low histological grade, PR expression, longer disease-free survival, and response to antiestrogen therapy [80, 81]. Decreased expression of ERβ in preinvasive in situ carcinoma, and its antiproliferative and anti-invasive properties in vitro suggest that ERβ has a role in maintaining the benign phenotype, perhaps as a tumor suppressor [82]. It has also been shown that the ERβ promoter is methylated MCF-7 cells, as well as in some breast cancer tumors [43, 83], leading to loss of ERβ expression.

1.3.1 Cistrome

Cistrome is a term used to define a set of DNA-binding sites (cis-regulatory elements) recruiting a certain transcription factor. The ER cistrome includes the ER binding sites associated with the classical model of ER action, as well as those associated with transcription factor cross-talk.
The genome-wide maps for ER binding sites have recently been described using chromatin-immunoprecipitation assays (ChIP) combined with high-throughput genomic technologies. The number of ERα binding regions, primarily in MCF-7 breast cancer cell lines, ranges from approximately 10,000 to 15,000 [84, 85]. Additionally, these studies show that the ERE is the most predominant motif in ERα-binding regions, but also identify other motifs, such as AP-1, Sp-1 and Forkhead motifs as enriched in ERα-binding regions. Recently, ERα binding events were assayed in different subtypes of ER+ breast tumors for the first time [86]. Interestingly, more than 70% of ERα binding sites enriched in the samples with poor outcome (ER+PR-HER2-, ER+PR+HER2–) and metastasis overlapped with ER-binding events in MCF-7 cells. In addition, it was shown that breast tumors with different clinical outcome have distinct ERα-binding profiles.

Several studies showed that 50-60% of the FOXA1 (Forkhead box protein A1) binding regions overlap with ERα-binding regions [86-88]. It is suggested that FOXA1 acts as a pioneering factor of ER signaling, facilitating ERα binding [89].

Using the ChIP-chip approach, Krum et al. compared the ERα-binding site profile in MCF-7 cells with that in the U2OS osteosarcoma cell line. Strikingly, less than 15% of the ERα-binding regions were common between the two cell lines. Importantly, FOXA1 is not expressed in U2OS cells and its DNA motif was not enriched within ERα binding regions in these cells [90].

ERβ binding regions in breast cancer cells have so far been studied only in cell lines with overexpressed ERβ, due to the lack of an appropriate breast cell line that expresses ERβ endogenously. Several studies examined ERβ-binding sites in MCF-7 cells engineered to express ERβ. Charn et al. [91] examined the location of ERα and ERβ DNA-binding regions in MCF-7 cells engineered to express one or both ERs upon E2 treatment. They identified a higher number of sites bound by ERα than by ERβ. More than 70% of ERβ binding regions overlapped with ERα binding regions when either ER subtype was expressed alone. However, only about 30% of the binding regions were shared when both ERs were expressed together, suggesting a competition between the ER subtypes with regard to selection of DNA-binding regions.

Our group identified ERβ-binding regions on a genome-wide scale in MCF-7 cells using the ChIP-chip approach [92]. Interestingly, 60% of the genomic regions bound by ERβ contained AP-1-like binding sites together with ERE-like sites. Co-occupancy of ERβ and AP-1 on chromatin was demonstrated and siRNA-mediated knockdown of expression of AP-1 family members (c-Fos and c-Jun) decreased ERβ recruitment to chromatin. These results suggest that the transcription factor AP-1 collaborates with ERβ in mediating estrogen responses in breast cancer cells.

Most of the studies investigated ER cistromes in MCF-7 breast cancer cells under similar conditions of E2 treatment. However, the number of regions detected differs between the studies and the binding profiles show a limited overlap. These differences could be due to biological variation between the MCF-7 (sub)lines, different cell handling protocols used in different laboratories, choice of ERα antibody or different platforms for detection of ChIP DNA sequences [93]. Additionally, MCF-7 cells with
exogenous ERβ vary in the employed ERβ-inducible systems and the ratio of ERα versus ERβ levels.

Despite the differences, several observations are consistent between the studies: (1) both ERs bind across the genome; (2) many regulated genes are associated with multiple ER binding regions; (3) ERα binding regions are more significantly associated with estrogen up-regulated, than down-regulated genes and 4) ERα and ERβ share many common binding regions.

These findings suggest that in estrogen-responsive breast cancer the final cellular response to estrogen is likely to depend upon the relative concentration of the two ERs in the cell, their activation status, DNA binding kinetics and the presence of other factors, such as FOXA1 and AP-1, influencing their respective functions.

1.3.2 Transcriptome

The transcriptome defines the set of RNA molecules, including mRNA, rRNA, tRNA, and additional non-coding RNAs, present in one or a population of cells at a given time.

Several reports have described global gene expression profiles in ERα-expressing breast cancer cell lines in response to E2 treatment [94-96]. These studies have reported different numbers of E2-regulated genes via ERα in MCF-7 breast cancer cells, ranging from 200 to 1500. Gene expression profiling have confirmed the regulation of several well known ER target genes in breast cancer cells such as pS2 (trefoil factor 1; TFF1), ADORA1 (adenosine A1 receptor), GREB1 (growth regulation by estrogen in breast cancer 1), MYC (v-myc myelocytomatosis viral oncogene homolog), CCND1 (cyclin D1) and IGFBP4 (insulin-like growth factor binding protein 4). ERα-regulated genes can be categorized into those that modulate the cell cycle, transcriptional regulation, morphogenesis, and apoptosis, compatible with a role of estrogen, via ERα, in inducing breast cancer cell proliferation and survival [97].

ERβ signaling is studied mainly in cell lines stably expressing ERβ either with or without ERα, due to the lack of cell lines expressing endogenous ERβ. Studies examining ERβ gene expression profiles in ERα-positive breast cancer cell lines stably expressing ERβ have provided insights into the interplay between ERα and ERβ in gene regulation. It is established that ERα and ERβ share some target genes, although each receptor also appears to have distinct sets of downstream target [98-101]. In these studies, co-expression of ERβ with ERα was found to significantly impact the E2-induced transcriptional response by ERα.

Monroe et al. attempted to identify targets unique for each ER in the ER– osteosarcoma U2OS cell line that was stably transfected with either ERα or ERβ [102]. Only 17 common genes were identified, suggesting that the transcriptional effects of E2 via ERα and ERβ, are largely distinct in these cells.
In other ER– cells, such as HEK293 (human embrional kidney, PAPER II) and Hs578T (human ductal breast carcinoma, [101]), ERβ alone has been overexpressed in order to identify its specific targets. Only three genes (PTGER4, ENPP2, and DKK1) were found to be commonly regulated in these cell lines, suggesting that ERβ evokes distinct gene responses in different types of target cells. Despite the differences, both studies reported inhibition of cell proliferation by ERβ expression independently of ERα, suggesting a similar function of ERβ in different cell types. Further studies are needed to clarify the molecular mechanisms by which ERβ elicits inhibitory effects on cell proliferation.

1.4 DNA METHYLATION

DNA methylation, an epigenetic modification, is a chemical change of the DNA sequence catalyzed by the enzymes DNA methyltransferases (DNMTs), which most commonly occurs at CpG dinucleotides in mammals [103].

Three DNMTs, encoded by distinct genes, have been identified in mammals, DNMT1, DNMT3a and DNMT3b. DNMT1 maintains the CpG methylation pattern during DNA replication and repair, whereas DNMT3a and DNMT3b establish the initial CpG methylation pattern de novo [103]. Furthermore, DNMT3b has been shown to be expressed as several splice variants. Most of them have altered catalytic activity, adding to the functional complexity of DNMTs. Methyl-CpG-binding domain protein family (MBD) members (MeCP2, MBD1, MBD2) which share a methyl-CpG-binding domain, have a specific affinity for methylated CpG sites, are shown to regulate gene expression by interacting with other epigenetic modulators, HDACs in particular [104].

DNA methylation is a post-replication modification, almost exclusively found on cytosines within CpG dinucleotides [105]. Genomic regions rich in CpG dinucleotides are called CpG islands. A CpG island is a region at least 200 bp long, with a GC percentage that is greater than 50% and with an observed/expected CpG ratio that is greater than 60%. CpG islands are located in the 5’-regions of 70% of all genes [106].

In cancer cells, CpG islands that are normally unmethylated can become methylated (hypermethylation) [107], which may result in repression of tumor-suppressor genes and genomic instability, through silencing of DNA repair genes, and chromatin condensation [108]. Additionally, CpG dinucleotides in other regions can become unmethylated (hypomethylation), leading to gene reactivation and ultimately to the up-regulation or overexpression of proto-oncogenes, increased recombination and mutation rates [109].

Gene silencing by hypermethylation of promoter genes is an important mechanism of carcinogenesis that offers opportunities for novel diagnostic and therapeutic strategies. More than 100 genes have been reported to be hypermethylated in breast tumors or breast cancer cell lines. Many of them play important roles in cell-cycle regulation, apoptosis, tissue invasion and metastasis, angiogenesis and hormone signaling [110,
Two established inhibitors of DNA methylation, 5-azacytidine (trade name Vidaza) and 5-aza-2'-deoxycytidine (trade name Decitabine), used in treatment of myelodysplastic syndromes, are in phase I/II of clinical trials for breast cancer treatment [76, 112]. Vidaza and Decitabine are cytosine analogues and act by their incorporation into DNA in the place of the natural base, cytosine, during DNA replication leading to covalent trapping of DNMTs. This causes the depletion of active DNMTs and demethylation of genomic DNA through cell division [113]. One disadvantage of these for treatment regiments is that the compounds are highly unstable in neutral aqueous solutions, and more stable derivatives and compounds are under development, such as zebularine and procainamide [112, 114].

1.4.1 Cross-talk between DNA methylation and estrogen signaling in breast cancer

Many studies have investigated the cross-talk between epigenetic modifications and estrogen-mediated gene regulation, mostly in breast cancer cell lines. However, the majority of studies focus on mediators of histone modification. Histone H3K9 acetylation and H3K4 methylation, associated with an open chromatin structure and increased transcriptional output, were observed at the promoters and enhancers of the active ERα target genes pS2 and GREB1 [115, 116]. Enzymes that mediate these modifications, such as arginine methyltransferase CARM1 and H3K4 methyltransferase SMYD3, were found recruited to the promoters of these genes upon treatment with E2 [68, 69]. Conversely, marks of repression (HDACs, H3K9 methylation) were observed in the promoters of E2-repressed genes [70, 71]. It was suggested that histone H3 and H4 acetylation and methylation, as well as DNA methylation, of the pS2 promoter are added and removed in a cyclical fashion [117]. The recruitment of cofactors, ERα, and RNA Polymerase II also occurred in a cyclical manner, producing transcriptional ‘waves’. The authors observed DNA methylation at the end of each productive transcription cycle. DNA methylation correlated with the occurrence of the MBD proteins, DNMT1, DNMT3a/b and the chromatin remodeler SWI/SNF. Furthermore, the authors suggest that DNMT3a/b is involved in both methylation and demethylation of the pS2 promoter. These data suggest that both histone modifications and DNA methylation may be intricate parts of the regular ERα transcriptional cycle.

Some studies show indirectly a relationship between DNA methylation and estrogen signaling, e.g. PR promoter methylation after ERα loss [118], or global methylation pattern changes in antiestrogen-resistant breast cancer cells [119]. The promoter of the CXCR4 gene, involved in the induction of proliferation, was found to be demethylated by tamoxifen in MCF-7 cells [120]. A genome-wide study, which combined DNA methylation and ERα binding assays in mammosphere-derived epithelial cells and MCF-7 cells, identified 11 large chromosomal zones, including a total of 108 genes that might undergo estrogen-mediated epigenetic repression [121].
Interestingly, ERα gene expression can be reactivated by inhibition of methylation in ER– cells, using demethylating agents or reducing DNMT expression [122, 123]. However, clinical data remains contradictory. For example, Lapidus et al. found hypermethylation of the ERα promoter region in ER– tumors [124], whilst other reports show no correlation between the gene methylation pattern and ERα gene expression in breast tumors [125].

Modified DNA methylation patterns in the ERβ promoters have been described for several forms of cancer [43]. Our group has previously shown that ERβ expression in tumors was inversely correlated with promoter methylation and that treatment with 5-aza-2′-deoxycytidine led to increased ERβ expression [126], which was confirmed by others [83, 127].

1.5 SINGLE NUCLEOTIDE POLYMORPHISMS

Single nucleotide polymorphisms (SNPs) are defined as a single base change in the DNA sequence that occurs more frequently than in 1% of individuals in a population. The corresponding nucleotides that differ within individuals are referred to as alleles, and chromosomal position of an allele is called locus. Alleles of loci in close vicinity tend to be inherited together. A set of SNPs that are statistically associated and therefore transmitted together is defined as a haplotype. Linkage disequilibrium (LD) is used to describe the non-random association of alleles at two or more loci. Unlike SNPs, genetic mutations that are recognized to increase e.g. breast cancer risk within families, such as those in BRCA genes, are much less frequent in the population. However, mutations confer higher risk [3].

**Candidate gene association studies** assay the effects of genetic variants in a gene potentially contributing to disease in a case-control material. These studies can be performed relatively quickly and inexpensively. However, the candidate gene approach is limited by the still incomplete knowledge on the genetic background of the investigated diseases. An example of a gene variant identified by this approach and confirmed in a large Breast Cancer Consortium study is a SNP in the caspase 8 gene (CASP8) coding region, which results in a substitution of aspartic acid for histidine [128]. The functional implications of this SNP on CASP8 protein function have not yet been identified.

**Genome-wide association studies (GWAS)** use high-throughput technologies to assay alleles of a large number of SNPs, typically 1 million in a case-control material. In one such study, the strongest association with breast cancer susceptibility was found for a SNP in the fibroblast growth factor receptor 2, FGFR2, positioned in intron 2, suggesting no direct effect on FGFR2 protein [129]. However, SNPs within introns may affect alternative splicing. The FGF signaling pathway has been shown to be important in mammary tumorogenesis and FGFR2 encodes a transmembrane tyrosine kinase involved in mammary gland development and breast carcinogenesis. Interestingly, FGFR2 expression, as well as the identified FGFR2 SNP, is associated
with ER+ tumors, suggesting interplay between these pathways. The possible function of the identified SNP in relation to FGFR and ER signaling remains to be determined.

1.5.1 SNPs in estrogen receptor genes

Approximately 2800 SNPs have been identified in the ERα gene [130]. Polymorphisms in ERα are associated with breast cancer, endometrial cancer, lupus nephritis, menstrual disorder, Alzheimer’s disease, osteoporosis and coronary artery disease [131, 132].

Several potentially functional SNPs in ERα have been evaluated for their association with breast cancer as well as other estrogen-related diseases. The most commonly studied are rs2234693 (T↔C) and rs9340799 (A↔G), both located in the first intron [133-135]. A potentially functional ERα SNP, rs2747648, was analyzed for association with breast cancer risk using a large familial study population [136]. It is located in the last exon and in silico studies show that it affects the binding of microRNA miR-453, which is stronger when the C allele is present, which can be associated with relative increased miRNA-mediated ERα repression, and decreased breast cancer risk. Another ERα breast cancer associated SNP (rs851987) is located in the promoter CpG island [137], in a predicted binding site for MeCP2. The T allele was associated with a protective effect regarding breast cancer risk. However, it remains to be shown that the different alleles of the SNP actually display different binding of MeCP2 with associated effects on DNA methylation and potentially ERα gene expression. Of 14 ERα SNPs associated with disease in GWAS, 4 are associated with breast cancer [138].

Around 720 SNPs have been identified in the ERβ gene [130]. Polymorphisms in the ERβ gene have been correlated to pathological states such as ovulatory dysfunctions, hypertension, bone mineral density, androgen levels and breast cancer. rs4986938, rs928554 and rs1256049 are frequent ERβ polymorphisms that have been associated with diseases, including breast cancer [132, 139-141]. None of these polymorphisms change the amino acid sequence of the ERβ protein. rs4986938 is a G↔A exchange in the 3′-UTR of exon 8. rs928554 is a G↔A exchange in the 3′UTR of exon 9. These SNPs display strong LD. rs1256049 is a G↔A exchange in the coding region of exon 6. A recent meta-analysis suggested that rs4986938 is related to breast cancer risk, acting as a modifier of the relationship between breast cancer risk and environmental factors, while the rs1256049 association is only due to high LD with rs4986938 [139]. Furthermore, rs4986938 has been found to be associated with psychiatric disorders, such as bulimia, Parkinson disease and Alzheimer disease [142].

Several groups have studied SNPs located in the ERβ promoter region, including rs2987983 and rs35036378, for association with breast cancer, proposing their effect on transcriptional regulation of gene expression [143, 144]. To date, no GWAS have revealed association between ERβ SNPs and disease [145].

Overall, the functional significance of many disease-associated SNPs has not been clarified. There is evidence that intronic SNPs may contribute to alternative splicing and 5′UTR SNPs to gene promoter activity. SNPs in coding regions and regulatory
regions are non-synonymous and likely to affect gene function [146]. Finally, SNPs in the 3’UTR may contribute to transcript stability and translatability [147].

1.6 HORMONAL THERAPY OF BREAST CANCER

Drugs used in hormonal therapy target estrogen signaling, and they include antiestrogens and aromatase inhibitors (AIs).

Antiestrogens disrupt estrogen signaling, including its down-stream proliferative effects. According to the producer’s (AstraZeneca) market analysis, the antiestrogen tamoxifen (trade name Nolvadex) is today’s best-selling hormonal anti-cancer drug. It is used in treatment of ER+ positive cancers, both in pre- and postmenopausal women. Tamoxifen itself has relatively low affinity for the ER, and whilst its metabolites, 4-hydroxytamoxifen and endoxifen, bind the ER with 30-100 times higher affinity [148].

The antagonistic activity of tamoxifen is crucial for breast cancer treatment, whereas in the bone, the endometrium and the cardiovascular system tamoxifen displays agonistic properties [149]. In the bone, tamoxifen was found to prevent osteoporosis [150, 151]. However, tamoxifen has a proliferative effect on the endometrium, increasing the risk for uterine cancer [152]. It is believed that the reason for tamoxifen being only a partial antagonist lies in the differential response of the two ERs. It has been demonstrated that tamoxifen can act as an agonist through ERα, and as an antagonist through ERβ [98], which can be due to their structural differences in the N-terminal domain that regulates interactions with coregulators. This suggests that the co-expression of the two ERs may affect the response to tamoxifen therapy.

Raloxifene (trade name Evista) is a SERM used in treatment of ER+ breast cancer. Side-effects of raloxifene are reportedly less profound than tamoxifen, particularly decreased risk of endometrial cancer and thrombosis. However their overall therapeutic effects are similar [153].

Fulvestrant (trade name Faslodex) is used as second-line therapy in postmenopausal women who had relapsed or progressed after previous hormonal therapy [154]. In ER+ breast cancer postmenopausal patients, fulvestrant was shown to have similar efficacy to tamoxifen [155, 156].

Aromatase inhibitors target estrogen production, by blocking aromatases, enzymes that catalyze estrogen synthesis. AIs are used in treatment of almost exclusively postmenopausal women, where the estrogen production is “outsourced” from ovaries to other tissues, such as liver, adrenal glands and fat. AIs are ineffective in premenopausal women, since the ovarian estrogen production is still active and under control of the hypothalamus and pituitary axis, via positive feedback. The AI-induced decrease in estrogen would activate the axis to stimulate estrogen production in the ovary, counteracting the AI effect. Acquired resistance to AIs eventually occurs, as well as some of the side effects: hot flushes, sweating, joint and muscle pain, and osteoporosis.
However, it has been suggested that AIs are highly effective in those ER+ tumors where tamoxifen, the most commonly used ER antagonistic drug, “switched sides” - displaying agonistic activities and inducing tumor growth [157]. It is proposed that optimizing the combinational AI+tamoxifen therapy may lead to improved treatments [158, 159].

1.6.1 Resistance to antiestrogen therapy

About 40% of ER+ tumors fail to initially respond to tamoxifen therapy and breast tumors initially responding to tamoxifen will in many cases develop resistance to this treatment. De novo and acquired resistance to antiestrogen therapy has been extensively studied, particularly for tamoxifen, however the underlying molecular mechanisms remain elusive.

About 10% of the patients who possess variant forms of the gene CYP2D6, whose cytochrome product is involved in the tamoxifen metabolism, may not receive full benefit from tamoxifen due to impaired CYP2D6 activity [160]. That is described as de novo (intrinsic) resistance. Additionally, it has been suggested that overexpression of HER2 can be another mechanism of intrinsic tamoxifen resistance [161], which is consistent with the lower efficacy of tamoxifen in ER+HER2+ patients.

Many different mechanisms can contribute to acquired resistance. There is evidence that high levels of coactivators, such as SRC-1 and SRC-3, may enhance the agonistic activity of tamoxifen and contribute to resistance [162, 163]. On the other hand, progressive reductions in corepressor activity during tamoxifen therapy may enhance the agonist effects of tamoxifen on the ER contributing to resistance [164]. For example, NCoR1 only weakly associated with ER in the absence of ligand, but did so avidly in the presence of hydroxytamoxifen [165]. When NCoR activity was blocked using a purified specific antibody, hydroxytamoxifen was converted into an agonist in MCF-7 cells.

Cross-talk between ER signaling and the growth factor receptor pathways, such as HER2, epidermal growth factor receptor (EGFR) and insulin-like growth factor receptor (IGFR) pathways, can affect cell growth and patterns of acquired resistance [166, 167]. For example, ER can be phosphorylated by the downstream components of the HER2 signaling pathway, the MAP kinases ERK1 and ERK2, leading to ligand-independent ER activation [67]. Therefore, increased ERK activity could potentially contribute to resistance to endocrine therapy, shown in several breast cancer cell-line models of endocrine resistance and suggested by clinical observations [168-170]. In addition to activating ER directly, kinase-mediated growth factor signaling may also modulate ER activity indirectly by phosphorylation of coregulators, enhancing the activity of coactivators and attenuating corepressor activity [171]. Experimental findings show that HER2 and EGFR levels rise after administration of tamoxifen and fulvestrant to breast cancer cells, suggesting a possible mechanism of acquired resistance [172]. Additionally, ER can activate HER2 and EGFR, and their downstream kinases, which was also implicated in resistance to tamoxifen [65, 66].
Acquired tamoxifen resistance can occur due to loss of ERα expression, but this is observed in only about 15% of breast cancer patients [173]. In fact, up to 20% of patients who have relapsed on tamoxifen, respond to AIs or fulvestrant, suggesting that ER continues to regulate growth even after the acquired resistance [174, 175]. DNA methylation of the ERα promoter has been proposed as a mechanism for the development of ER− cell lines as well as primary tumors [76-78]. MCF-7 cell lines resistant to either tamoxifen or fulvestrant were shown to have distinct epigenetic profiles compared to sensitive MCF-7 cell lines, suggesting that epigenetic mechanisms might be involved in antiestrogen resistance [119].

Loss of ERβ expression, commonly observed in breast cancer, has also been implicated in tamoxifen resistance. In line with this, it has been shown that high levels of ERβ may improve disease-free and overall survival in patients treated with tamoxifen [176]. Interestingly, another hypomethylating agent, procainamide, was found to decrease tamoxifen resistance by inducing ERβ overexpression in breast cancer patients [177].
2 AIMS OF THE THESIS

Clarifying the molecular pathways of estrogen signaling is important for our understanding of breast cancer development, diagnosis and treatment. These molecular pathways display significant complexity, including the existence of two distinct ERs, ERα and ERβ, their numerous splice variants and polymorphisms, as well as their interplay, both mutual and with other signaling pathways. ER status is not a perfect marker for responsiveness to antiestrogen therapy, and additional markers are needed to more accurately define patients who will benefit from hormone therapy. Additionally, intrinsic and acquired resistance to this treatment represents a significant clinical challenge. This work attempts to address some molecular details and mechanisms relating to the role of ERs in breast cancer with the ultimate hope that this knowledge will in the future contribute to stratify diagnosis and improve therapy for breast cancer patients.

The general aim of this thesis was to characterize molecular mechanisms of ER action, focusing on breast cancer. It addresses the interplay of ERα with ERβ and DNA methylation pathways, the role of ERβ in global gene expression, and the functionality of two common ERβ SNPs. Specifically, the aims were:

PAPER I Functional characterization of two common SNPs positioned in the ERβ 3’UTRs that have been shown to be associated with several diseases.

PAPER II Description of ERβ regulatory functions independent of ERα.

PAPER III Demonstration of the recruitment of ERα/β heterodimers to various DNA-binding regions in intact chromatin.

PAPER IV Investigation of global gene regulation by estrogen signaling and DNA methylation.

PAPER V Investigation of the interplay between ERα and main regulators of DNA methylation, DNMTs.
3 METHODOLOGICAL CONSIDERATIONS

3.1 CELL LINES

The majority of the constituent studies in this thesis are based on experiments in immortalized human cell lines. Cell lines offer attractive systems for molecular mechanistic studies due to their ease of manipulation and propagation.

The MCF-7 (Michigan Cancer Foundation 7) cell line is the most commonly used cell line in ER signaling studies related to breast cancer. The receptor status of MCF-7 is ER+ and the cell line is of luminal origin. There are different MCF-7 (sub)lines used in different laboratories, varying in E2-response, proliferation rate and chromosomal structure [178]. In PAPER I, PAPER IV and PAPER V, we used the MCF-7S strain, originally generated by Dr. S. Shafie at the NIH in 1981 [179]. MCF-7 cells with an inducible Tet-system used in PAPER III were obtained from Clontech.

The HEK293 (Human Embryonic Kidney 293) cell line is of epithelial origin. It was selected for the studies in PAPER II due to the lack of endogenous ERs. In this paper we used HEK293 cells with an inducible Tet-system obtained from Clontech. Due to high transfection efficiency [180], the HEK293 cell line from American Type Culture Collection (ATCC) was selected for the majority of the experiments in PAPER I due to its convenience in transfection assays.

HeLa (Henrietta Lacks) is the first generated human cell line, derived from cervical cancer of the patient that the line was named after [181]. Similar to HEK293, it lacks endogenous ERs and can be efficiently transfected. It was used in PAPER I, to confirm data obtained in HEK293 cell line.

COS-7 (CV-1 Origin SV40 7) is a monkey kidney cell line, obtained by immortalizing a CV-1 cell line derived from kidney cells of the African green monkey [182]. Similar to HEK293 and HeLa cell lines, this cell line is often used for in vitro studies, especially transient transfection-based assays including co-immunoprecipitation, which we employed in PAPER V.

3.1.1 Tet gene expression systems

Due to lack of a breast cancer cell line that expresses significant amounts of ERβ, we have generated stable cell lines expressing ERβ (PAPER II and PAPER III). The expression of ERβ in these cell lines were controlled by Tet gene expression systems.

Tet gene expression systems are commercially available as Tet-Off and Tet-On cell lines, providing regulated, high-level gene expression [183]. In the Tet-Off system,
gene expression is turned on when the antibiotics tetracycline (Tet) or doxycycline (Dox; a Tet derivative) are removed from the culture medium. In contrast, expression is turned on in the Tet-On system by the addition of Dox. The first critical component of the Tet Systems is the tetracycline-controlled transactivator (tTA), a modified E.coli TetR protein. In the Tet-Off system, tTA is encoded by the pTet-Off regulator plasmid. In the Tet-On system, tTA is modified into rtTA (“reversed” tTA) and it is encoded by the pTet-On regulator plasmid. rtTA protein is capable of binding the operator only when bound by doxycycline, hence the Tet-On system is not responsive to Tet (Figure 5). In both Tet-On and Tet-Off systems, transcription is turned on or off in response to the antibiotic in a precise and dose-dependent manner [184].

Figure 5. Schematic outline of gene regulation in the Tet-Off and Tet-On Systems. Tet-Off: tTA binds the TRE and activates transcription in the absence of Tet or Dox. The TRE is located upstream of the minimal immediate early promoter of cytomegalovirus (PminCMV), which is silent in the absence of activation. tTA binds the TRE - and thereby activates transcription of Gene X - in the absence of Tet or Dox. Tet-On: rtTA binds the TRE and activates transcription in the presence of Dox. From Clontech Laboratories, Inc.® Tet-Off and Tet-On Gene Expression Systems User Manual.
The second critical component is the response plasmid which expresses the gene of interest (in our case ERβ) under control of a tetracycline-response element (TRE). A response plasmid without the gene of interest is used as a negative control (Mock). The response plasmid is usually pBI-EGFP, where TRE is bidirectional (Figure 6), i.e. designed to co-express both the gene of interest and enhanced green fluorescent protein (EGFP). Hence, when TRE is activated, both the gene of interest and EGFP are expressed, and the cells are visible as green, which facilitates clone selection.

![Figure 6. Restriction map of pBI-EGFP (left) and detailed organization of the Pbi-1 bidirectional promoter (right). MCS is acronym for Multiple Cloning Site into which the ERβ cDNA is cloned. Pbi-1 bidirectional promoter contains the TRE. From Clontech Laboratories, Inc.® pBI-EGFP Vector Information.](image)

### 3.2 GENE EXPRESSION ASSAYS

Gene expression assays can target either a single gene or have a genome-wide approach. Both approaches were extensively used in the studies described in this thesis.

#### 3.2.1 Quantitative polymerase chain reaction

Quantitative polymerase chain reaction (qPCR, real-time PCR) is a widely used approach for expression analysis at low throughput. It is based on detection of produced PCR products in real time by fluorescence produced by reporter molecules, the most common being the SYBR Green dye and TaqMan probes. The SYBR Green dye functions by non-specifically binding to double-stranded DNA. TaqMan is a fluorophore- and quencher-containing probe that specifically hybridizes to the gene of interest.

To quantify the PCR product, we used the standard $2^{-\Delta\Delta C_{t}}$ method, which assumes 100% efficiency for the PCR reaction. The Ct (Cycle threshold) value represents the number of cycles required for the fluorescent signal to reach a set threshold in the linear phase of the qPCR production. $\Delta C_{t}$ is the difference between the Ct values of two
samples. The exponential value, $2^{\Delta C_t}$, represents the relative fold change between two samples. To normalize for individual sample variations, Ct values from a housekeeping gene are used, which is then subtracted from the first one, giving the $\Delta \Delta C_t$ value. The final exponential value, $2^{\Delta \Delta C_t}$, represents the relative fold change between two samples normalized by the Ct values from a housekeeping gene. Three of the most commonly used human housekeeping genes, Gus, 18S and 36B4 (RPLP0), were used in our studies, due to high and stable expression levels in the used cell lines.

### 3.2.2 Expression microarrays

Microarray technology is used to measure gene expression on a genome-wide scale. It is based on high-density arrays that contain thousands to millions of oligonucleotide probes to detect cDNA targets. Differences in expression levels determined by microarrays are usually confirmed by qPCR. In most cases, qPCR confirms the general trend of a change obtained from a microarray, but the absolute changes can vary, due to different sensitivities of the two assays and different designs of probes and primers. While qPCR employs two primers and optionally a probe amplifying an approximately 100bp long sequence, microarrays used in this study employ multiple probes that cover the longer parts, and sometimes the entire mRNA transcripts.

Two Affymetrix® expression array types were used in our studies. In PAPER II, we employed the Gene 1.0 ST Array, which contains approximately 26 probes spread across the transcripts of 28,869 annotated genes. GeneChip HT HG-U133+ PM 96-Array Plate was used in PAPER IV. This array contains probe sets with 9 or 10 probe pairs per transcript of more than 47,000 transcripts and variants of more than 33,000 well-characterized genes and UniGene clusters. It is a 3’-based array, which means that all probes are positioned at the 3’ end of mRNAs. Sample processing and data analysis was performed at the Bioinformatics and Expression Analysis core facility at the Karolinska Institutet (www.bea.ki.se).

### 3.3 IMMUNOPRECIPITATION ASSAYS

Immunoprecipitation assays employ antibodies in order to isolate proteins and screen for their interactions, including with other proteins (Co-IP) or chromatin (ChIP).

#### 3.3.1 Protein complex immunoprecipitation (Co-IP)

We used Co-IP in PAPER V, to identify possible protein complexes between ERα and DNMTs. The method is based on precipitation of a protein complex, using an antibody specific for one of the complex members, in our case ERα. Precipitated proteins are then separated by SDS PAGE and proteins present in the complex are detected by a Western blot assay, using an antibody specific for other possible member of the complex, in our case a DNMT. When an interaction is detected, it can be confirmed by
switching the antibodies at the respective steps. However this is not always successful as an antibody may work for protein precipitation but not in Western blot assays, and vice versa.

3.3.2 Chromatin immunoprecipitation (ChIP)

ChIP was critical central assay in three of our studies: PAPER II, PAPER III and PAPER IV. It was used to determine the location of ERα and/or ERβ DNA binding sites.

This assay is designed to isolate DNA sequences that DNA-associated proteins bind to. An antibody, or in some cases a pool of antibodies, specific for a putative DNA-binding protein is used to immunoprecipitate the protein-DNA complex from cellular lysates after cross-linking, usually by formaldehyde. The cells are then lysed and the protein-bound DNA is broken into pieces of 0.2-1.0 kb in length using sonication. Protein-DNA complexes are then precipitated with a specific antibody, followed by reversing cross-links using high temperature (65°C), which allows the DNA to be separated from the protein. The identity and quantity of the DNA fragments isolated are then determined by classical PCR, qPCR or using high-throughput assays. The latter include ChIP-chip and ChIP-DSL (DNA selection and ligation), based on high-density oligonucleotide arrays, and ChIP-PET and ChIP-Seq, based on high-throughput sequencing [84]. Re-ChIP is used to investigate the simultaneous binding of two proteins to the same DNA region(s), employing two, or more, sequential precipitation steps as exemplified in PAPER III.
4 RESULTS AND DISCUSSION

4.1 PAPER I

EFFECTS OF TWO COMMON POLYMORPHISMS IN THE 3' UNTRANSLATED REGIONS OF ESTROGEN RECEPTOR BETA ON mRNA STABILITY AND TRANSLATABILITY

This study was the first report to functionally characterize two common SNPs in the 3'UTRs of ERβ, focusing on the differences between alleles with regard to mRNA stability and translatability. The goal was to define the molecular mechanisms by which SNPs in ERβ are associated with disease.

ERβ1 and ERβ2 use different last exons, exons 8 and 9, respectively, and thus have different 3'UTRs. The SNPs in focus of this study are referred to as rs4986938 and rs928554, positioned in the 3'UTR of exon 8 and 9, respectively. Both are G→A transitions. A large number of studies report associations of these two SNPs with diseases, including bulimia, cardiovascular diseases, osteoporosis and breast cancer [140, 141, 185-187]. The SNPs are found to be associated either individually or within a haplotype block.

We examined allelic expression for the two SNPs in breast tumor samples from heterozygous individuals. For each sample and SNP, the relative levels of the cDNA (as a measure of the mRNA levels) versus the genomic DNA were determined for both alleles. A significant difference in allelic expression level was observed for rs4986938, but not for rs928554.

To further investigate potential molecular effects of the two SNPs, a cell model system was employed. We used a modified PGL3 basic vector, where the ERβ promoter and the ERβ 3'UTRs are flanking the luciferase gene including the different alleles of the investigated SNPs. mRNA stability was assayed by determining mRNA levels after inhibition of transcription by actinomycin D. A luciferase assay was used to determine protein levels as a measure of mRNA translatability. We did not observe any differences in mRNA stability or translatability between the two alleles for the investigated SNPs. Although the allelic expression assay suggested that rs4986938 alleles had an impact on mRNA levels the sample number was small (n=5) and the observed result could not be accounted for by molecular mechanisms investigated in cell models.

These results indicate that the observed associations between ERβ 3'UTR SNPs and disease susceptibility are most likely due to linkage disequilibrium with another gene variant, rather than the variant itself being the susceptibility factor.
4.2 PAPER II

MICROARRAY ANALYSIS OF ALTERED GENE EXPRESSION IN ERβ-OVEREXPRESSING HEK293 CELLS

The aim of this study was to gain a global understanding of ERβ-dependent gene regulation independent of ERα. The experimental approach was a global gene expression profiling analysis for HEK293 Tet-On cell lines with inducible version of ERβ1 and ERβ2. The analysis showed that ERβ2 had no effect on mRNA expression under the applied experimental conditions and with the filters employed to identify differentially expressed mRNAs (unpublished data).

A total of 332 genes and 210 genes were found to be up-regulated and down-regulated by ERβ, respectively. The Gene Ontology analysis revealed that ERβ-induced and ERβ-repressed genes were involved in cell-cell signaling, morphogenesis, and cell proliferation. The ERβ repressive effect on genes related to proliferation was further studied by functional proliferation assays, where ERβ expression resulted in a significant decrease in cell proliferation.

To identify primary ERβ target genes, we examined 20 ERβ-regulated genes selected from the global gene expression profiling experiment, using ChIP assays for regions bound by ERβ. Our results showed that ERβ recruitment was significant to regions associated with 8 ERβ up-regulated genes and 5 ERβ down-regulated genes. Six binding regions were located within genes, four downstream and two upstream of genes.

Our findings on ERβ regulatory functions independent of ERα were consistent with previous reports in two ER- cell lines, the human breast cancer Hs578T cell line and the osteoblastic U2OS cell line, in which stably expressed ERs were investigated for their global gene expression profiles [101, 102]. However, of 95 genes found to be regulated 2-fold or more, by ERβ in Hs578T/ERβ cells, only 3 (PTGER4, ENPP2, and DKK1) were identified in our study. One of them (ENPP2) was identified as a primary ERβ target gene in our study. For another gene (PTGS2), we show down-regulation by ERβ, whereas the study in Hs578T cells reports up-regulation by ERβ. The observed discrepancies may be the result of different origin of the cell systems (HEK293 versus Hs578T breast cancer cells), achieved levels of stably expressed ERβ, number of genes on the array (28,869 genes in our study versus 8,700 genes in the Hs578T study), and applied filters for identifying regulated genes (1.5-fold vs. 2-fold). However, both studies report inhibition of cell proliferation by ERβ expression independently of ERα, suggesting a similar function of ERβ in different cell types.

This study provides novel information on the gene regulatory function of ERβ independent of ERα and identified a number of primary ERβ target genes. The results of GO analysis and proliferation assays are consistent with an antiproliferative role of ERβ independent of ERα, which supports ERβ as a potential marker of good prognosis in breast cancer treatment.
4.3 PAPER III

BINDING OF ESTROGEN RECEPTOR α/β HETERO DIMERS TO CHROMATIN IN MCF-7 CELLS

This study was the first demonstration of ERα/β heterodimers recruitment to various DNA-binding regions in intact chromatin. Although the formation of ERα/β heterodimers has been previously demonstrated using GST-pull down and gel-shift assays [35, 188], their exact role in estrogen signaling remains unclear.

We applied the Re-ChIP assay to study the simultaneous presence of ERα and ERβ on various DNA-binding regions in intact chromatin. Following E2-treatment for 45 min, ERα/β heterodimers were isolated by precipitation with an anti-ERβ antibody followed by an anti-ERα antibody in an MCF-7 Tet-Off cell line that stably expresses an inducible version of ERβ (MCF-7/ERβ) together with endogenous ERα. MCF-7/Mock was used as a negative control.

We initially applied the Re-ChIP method to assay ERα/β heterodimers to a promoter region of the pS2 gene, known to bind both ERα and ERβ. Recruitment to the pS2 promoter was only detected in ERβ-ERα immunoprecipitated fractions from the MCF-7/ERβ cells, but not from the MCF-7/Mock cells. The fold-enrichment for heterodimers was significantly lower than when the interaction was assayed for each ER alone, which could be explained by loss of material during the Re-ChIP assay, low fraction of heterodimers compared to homodimers or lower DNA binding affinity of heterodimers compared to homodimers, which has been previously suggested [176].

We further examined the binding of ERα/β heterodimers to 12 binding regions, identified as regions recruiting both ERα and ERβ by a ChIP-chip assay (from [92] and unpublished data). Among the 12, 8 binding sites were located within genes, and 4 were located in the promoter regions. The Re-ChIP assay demonstrated recruitment of ERα/β heterodimers to all selected regions. Additionally, all tested binding regions exhibited significant recruitment of ERα and ERβ individually. However, similarly to what was observed for the pS2 binding region, significantly lower recruitment of heterodimers was observed in comparison to when the interaction was assayed for each ER alone. Furthermore, it is notable that different binding regions exhibited differences in fold enrichment of ERα/β heterodimers. This could be attributed to variations in affinity of heterodimers to various ER-binding regions depending upon the sequence of the binding regions.

Using the CONSITE program, we found that all of the selected binding sites contained half-ERE motifs, 70% contained AP1-motifs, and only 23% contained full ERE motifs. Furthermore, our data showed that the ratio of enrichment in the ERα/β Re-ChIP assay versus the enrichment in the ERα and ERβ ChIP assays varied among the selected sites. For example, for genes such as NBPF4, NOTCH2, NBPF15, and PRUNE, this ratio is high, suggesting that heterodimer recruitment may play a more significant role in ER-mediated transcriptional regulation of these genes.
To verify the effect of ERβ on transcriptional activity of ERα, we evaluated the mRNA levels of some selected genes after E2 treatment. Our data showed that mRNA levels of genes that are implicated in cell proliferation, like pS2, and ADORA1 were down-regulated in the MCF-7/ERβ cell line compared with the MCF-7/Mock cell line. The mRNA levels of NBPF1 and NBPF4, whose function is not fully explored, were also decreased in the presence of ERβ. mRNA levels of BCL9, which is participating in cell proliferation and development, were up-regulated when ERβ is expressed. Finally, the mRNA levels of NOTCH2 and PRUNE were unaffected by the presence of ERβ. Interestingly, both these genes have been implicated in the aggressiveness of breast neoplasm.

Considering that all examined binding regions exhibited significant ERα/β heterodimer recruitment, our data suggested that heterodimerization is a frequent mechanism by which ERα and ERβ interact in estrogen signaling.

**4.4 PAPER IV**

**GENE REGULATION BY ESTROGEN SIGNALING AND DNA METHYLATION IN MCF-7 BREAST CANCER CELLS**

In this study we investigated a potential regulatory cross-talk between estrogen signaling and DNA methylation by identifying their common target genes and exploring potential underlying molecular mechanisms in human MCF-7 cells.

We compared effects on global gene expression profiles in response to E2 and the hypomethylating agent 5-aza-2'-deoxycytidine (DAC). We focused on the up-regulated genes in the DAC-regulated group, since DAC-induced hypomethylation leads to up-regulation of genes that are normally directly silenced by DNA methylation. Changes in gene expression profiles in response to E2 and DAC co-treatment were not further explored, as DAC down-regulated both mRNA and protein levels of ERα in our experiments (data not shown), in accordance with previous findings for MCF-7 cells [189].

A total of 88 genes were up-regulated by both E2 and DAC (E↑D↑ group) and 58 genes were down-regulated by E2 and up-regulated by DAC (E↓D↑ group). Candidate common target genes were selected for further analysis using GO Analysis, previously reported association with breast cancer, estrogen signaling and/or DNA methylation, reported ER recruitment, and predicted CpG islands. Based on these selection criteria, six genes were selected for a detailed investigation in relation to regulation by E2 and DAC: Three from the E↑D↑ group (BTG3, FHL2 and PMAIP1) and three from the E↓D↑ group (BTG2, CDKN1A and TGFB2). Real-time PCR analysis confirmed changes in gene expression derived from microarray data for selected genes. We further confirmed that the selected genes were regulated through ER by E2 and tamoxifen co-
treatment. Finally, a dose-response effect of DAC on the induction of the expression of these genes was observed.

As the regions of the selected genes previously shown to recruit ERα do not contain predicted CpG islands, we assessed ERα recruitment to the CpG island containing promoter regions of these genes. ERα showed significant recruitment to all investigated promoters, however the observed recruitment was much lower to the CpG island containing promoter regions than to the previously reported ER binding regions of these genes. The lack of an effect of E2 in this assay suggests ligand-independent ERα recruitment. Additionally, DAC treatment did not affect ERα recruitment to the promoter regions, suggesting that the promoter methylation status has no effect on ERα recruitment.

Although DAC treatment activated the expression of all selected genes, the bisulfate sequencing assay showed that only the promoters of the BTG3 and FHL2 genes were methylated. However, E2 treatment had no effect on the methylation status of these promoters. PMAIP1, BTG2, CDKN1A and TGFB2 are also found to be up-regulated by DAC. However, as their promoters were not methylated even in the absence of DAC, this suggests that DAC regulated the expression of these genes either via hypomethylation of other methylated DNA regions, such as CpG shores, shelves and open seas, or indirectly, through hypomethylation of other genes.

Our results support the previous report that BTG3 (B-cell translocation gene 3) is directly regulated by DNA methylation in MCF-7 cells [211]. Furthermore, we show that E2 affects the expression of this gene. However, our data does not support that this effect involves changes in DNA methylation status. Genistein, an ER ligand, has been shown to have the same hypomethylating effect as DAC on the BTG3 promoter in renal and prostate cancer cells, hence estrogen effects on DNA methylation might display cell type selective mechanisms [190, 191]. Interestingly, the BTG gene family encodes proteins that appear to have antiproliferative properties [192]. Apart from BTG3, another member of the family, BTG2, was identified in our study.

We confirm the previous data that FHL2 (four and a half LIM domains 2) gene is regulated by E2 in MCF-7 cells [119]. Fan et al. observed that long term disruption of estrogen signaling using fulvestrant can lead to hypermethylation of the FHL2 promoter with the associated loss of E2 responsiveness. However, they showed no changes in FHL2 methylation upon long term tamoxifen treatment, suggesting that diverse antiestrogens can exert different effects on the DNA methylation status of the FHL2 promoter. Our study, focusing on regulation of promoter methylation upon short term E2 treatment, does not support a connection between short term E2 treatment and FHL2 promoter methylation.

In summary, we identified a set of genes regulated by both estrogen signaling and DNA methylation. However, our data does not support a direct molecular interplay of mediators of estrogen and epigenetic signaling at promoters of regulated genes.
4.5 PAPER V

INTERPLAY OF ESTROGEN RECEPTOR α AND DNA METHYLTRANSFERASES WITH FOCUS ON DNMT3B

In this study, we explored the interplay of estrogen signaling and DNMTs.

Reports on the effect of estrogen signaling on DNMT expression are limited, and most of them study the endometrium, since its lining undergoes cyclic regeneration and it is hypothesized that this process is regulated epigenetically [193]. Additionally, prolonged exposure to estrogen is a risk factor for endometrial cancer [194]. We observed that in MCF-7 cells estrogen signaling affects the mRNA expression of the two DNMT classes, the maintenance DNMT, DNMT1 and the de novo DNMTs, DNMT3a and DNMT3b, differently. E2 induces the expression of DNMT1, and decreases the expression of DNMT3a and DNMT3b. E2 increased DNMT1 mRNA expression only at 24h, but not at 48 and 72h, which could be due to additional regulatory mechanisms contributing at later time points. It is well established that estrogen signaling controls the cell cycle in MCF-7 cells by mediating the G1-S phase transition [195], which can be related to the increase in DNMT1 expression observed after 24h E2 treatment, and due to a requirement for maintenance of the DNA methylation pattern during replication. ER cistrome assays and gene expression profiling assays of potential E2-regulated genes did not detect DNMTs as primary ER targets in breast cancer cells.

Using a Co-IP assay, we identified protein-protein interactions between ERα and DNMT3b1. We could not detect protein-protein interactions between endogenous ERα and transfected c-myc tagged DNMT3b in MCF-7 cells, possibly be due to low transfection efficiency in these cells. The interaction between ERα and endogenous DNMT3b could not be studied due to the lack of an antibody against endogenous DNMT3b suitable for the Co-IP assay.

Interestingly, in comparison to DNMT1 and DNMT3a, DNMT3b has been more clearly implicated in breast cancer. It was shown that about 30% of breast cancer patients had increased DNMT3b expression in tumor tissue compared to normal breast tissue, while DNMT1 and DNMT3a were overexpressed in only 5% and 3% of breast carcinomas, respectively [196]. Elevated expression of DNMT3b was shown to be significantly associated with hypermethylation and subsequent reduced ERα expression and higher histological grade, pointing to a potential involvement of DNMT3b in breast tumor progression and aggressiveness [197]. A strong correlation between total DNMT activity and overexpression of DNMT3b was reported in a subset of breast cancer cells correlating with hypermethylation of methylation-sensitive genes, including ERα. No correlation between total DNMT activity with the expression of DNMT3a or DNMT1 was observed. It was recently demonstrated by the same group that inhibition of DNMT3b by siRNA-mediated knockdown can increase the chemotherapeutic efficacy in breast cancer cells [198]. The authors suggested that DNMT3b has a predominant role over DNMT3a and DNMT1 in breast tumorigenesis.
Since DNA methylation is involved in the process of gene silencing, we investigated whether depletion of DNMT3b can lead to up-regulation of expression of a set of genes we identified as regulated by both estrogen signaling and DNA methylation in PAPER IV. We found that the expression of FHL2 and CDKN1A was increased after siRNA-mediated DNMT3b knockdown.

In line with our results, the mRNA levels of CDKN1A were previously found to be up-regulated in MCF-7 cells after DAC-treatment or siRNA-mediated depletion of DNMT3b [199]. In PAPER IV, the CDKN1A promoter was found to be unmethylated, suggesting that DNMT3b regulate CDKN1A expression indirectly, which has also been proposed by others [200]. Alternatively, regulation of DNA methylation might be mediated via DNA-methylation of non-CpG island motifs such as CpG shores, shelves and open seas [201, 202].

Our results suggest that E2 treatment and DNMT3b depletion did not display synergistic effects on FHL2 and CDKN1A expression. This is line with the results of PAPER IV, where we demonstrated that the E2-mediated effect of ERα recruitment is selective to the enhancers of these genes, which contain no predicted CpG islands.

The FHL2 promoter was found to be methylated in MCF-7 cells in PAPER IV, as suggested by others [119]. However, to our knowledge, the molecular mechanisms that regulate FHL2 promoter methylation have not been described. Interestingly, FHL2 was found to be overexpressed in patients with ICF (Immunodeficiency, Centromere Instability, Facial abnormalities) syndrome, caused by mutations in the DNMT3b gene that lead to impaired catalytic activity of the enzyme [203]. This supports our findings and suggests that FHL2 promoter methylation, and subsequent gene expression, is regulated by DNMT3b. Considering that DNMT3b and ERα may be in the same complex, we can hypothesize that ERα can regulate FHL2 expression together with DNMT3b on a same regulatory element, possibly a CpG island in the promoter region. Whether promoter methylation pattern and the protein expression of FHL2 are directly regulated by DNMT3b, remains to be confirmed. In line with our hypothesis, additional approach could be to assay FHL2 promoter methylation status after ERα depletion.
5 GENERAL CONCLUSIONS AND FUTURE PERSPECTIVES

5.1 ESTROGEN RECEPTOR B SNPS AND DISEASE SUSCEPTIBILITY

Several studies have associated SNPs in ERβ 3’UTRs with various diseases including breast cancer. However the function of these SNPs in disease etiology remains unknown. We have addressed the function of two SNPs in the ERβ 3’UTRs that have been frequently assayed and associated with disease in candidate gene association studies. Our studies do not support a function of these SNPs in disease etiology. Future studies should address additional potential functions of these SNPs as well as the potential function of SNPs in LD with these SNPs. In general, development in our understanding of SNP function has been rather slow in comparison with the generation of data about associations between SNPs and disease. Novel approaches might be needed to address the function of SNPs, ideally in a high throughput format.

5.2 ESTROGEN RECEPTOR B FUNCTIONALITY

Antiproliferative properties of ERβ have been shown in a number of in vitro studies and ERβ has been correlated with markers of good prognosis in many clinical studies. However, the general clinical significance of ERβ is not yet established. Unlike ERα, ERβ is not generally introduced as a biomarker in breast cancer diagnostics. Reasons include the lack of well validated ERβ antibodies, and still rather limited knowledge about the molecular mechanism of action of ERβ and its function in breast cancer. In general, ERβ levels are decreased in breast cancer. It is believed that ERβ is the predominant ER in the mammary epithelium, but that its expression is reduced and possibly ultimately lost during cancer progression, which would restrict its exploration as a drug target.

In our studies we have addressed the function of ERβ in the absence of ERα as well as the cross-talk between ERα and ERβ at the level of heterodimer binding to DNA in the context of the intact chromatin. We used HEK293/ERβ cells to confirm the antiproliferative function of ERβ as well as to identify a palette of ERβ-target genes in the absence of ERα. We used MCF-7/ERβ cells to provide further evidence for the existence of ERα/β heterodimers.

These findings, as well as many others that study the function of ERβ in the context of cell model systems that stably overexpress ERβ, remain to be further investigated, in appropriate cell line models that express endogenous ERβ and more importantly in vivo. There is a constant influx of data on ERα genome-wide expression profiles and DNA binding events in various cell lines and breast tumors, but such information is still
lacking for ERβ. Future effort should focus on the identification of appropriate model systems to study the endogenous ERβ, where re-activation of ERβ by hypomethylating drugs could constitute one approach. The identification of antibodies compatible with the ChIP assay on breast cancer samples should also be prioritized.

5.3 ESTROGEN SIGNALING AND DNA METHYLATION

Interplay between the pathways of estrogen signaling and DNA methylation has been explored, but it is not well understood. Aiming to identify common targets of estrogen signaling and DNA methylation on a genome-wide scale in breast cancer cells, we used gene expression profiling and identified about 150 genes that were regulated by both pathways. After selecting and analyzing a subset of six genes, we concluded that there was no direct molecular interplay of mediators of estrogen and epigenetic signaling at these promoters. We cannot exclude the possibility that some of the remaining ~140 genes could be targets for a direct molecular interplay of these two pathways. Additionally, our studies were limited to a single time point of E2 exposure and ER ligands, including E2, could have direct effects on methylation patterns at other time points.

Most of the methylation screening assays used so far assessed the methylation changes in the promoter regions, which was also the focus of our study. It is possible that estrogen signaling affects methylation of other regions in the genome reported for differential methylation patterns, such as shores, shelves and open sea. Platforms covering wider DNA regions, such as Infinium Human Methylation 450 BeadChip, or the MeDIP assay, which assays methylation across the complete genome, could be used to test this hypothesis.

Instead of genomically aberrant breast cancer cells, another relevant approach would be assessing the estrogen signaling-related changes in global DNA methylation in mammary epithelial cells.

Nevertheless, we have identified one significant gene, FHL2, regulated by both pathways, but via distinct regulatory elements. E2-activated ERα is recruited to its distant enhancer and we hypothesize that this is the mechanism behind E2-induction of FHL2 expression. Unliganded ERα is recruited to the CpG island within the FHL2 promoter. Furthermore, we found that DNMT3b depletion increases FHL2 mRNA expression, suggesting that this DNA methyltransferase regulates FHL2 promoter methylation, which remains to be confirmed. Interestingly, we show that ERα and DNMT3b can be in the same complex, implying that the unliganded ERα could also be involved in methylation of FHL2 promoter. FHL2 is a transcription factor found to interact with many other factors, including estrogen receptor [204]. It has been suggested to repress ERα and ERβ transcriptional activity in breast cancer cells, together with the corepressor Smad4 [205]. ERs and FHL2 seem to regulate each other through a complex feedback mechanism that involves DNA methylation, and their intriguing relationship is worth exploring further.
We also show that ERα signaling affects the mRNA expression of DNMTs, suggesting global interplay between estrogen signaling and regulation of methylation. Future studies should focus on assaying DNMT protein expression and enzymatic activity in response to ER ligands in breast cancer cells.

In summary, this thesis addresses various mechanistic aspects of estrogen signaling in breast cancer cells. Hopefully, knowledge about estrogen and ER signaling, including the cross-talk between the two ERs and their respective cross-talk with other signaling pathway will suggest hypothesis for improved diagnostic criteria as well as therapeutic strategies that can be tested in an appropriate clinical setting.
6 ACKNOWLEDGEMENTS

It is widely considered that a **phenotype** - visible characteristics of every living being, such as physical appearance and behavior - is a result of both genetic background and environmental factors. I like to believe that I became a scientist more due to the environmental circumstances, than due to genetic predisposition (since there are other things I’m more talented in… although that’s again on the environment to evaluate). Therefore, I would like to thank all the people (and factors), without whom this would never come to be…

My supervisor **Karin Dahlman-Wright**. To start with, thank you for answering that, now probably silly, e-mail I sent back in 2006, looking for a PhD, when I was still addressing you as “Dear professor Dahlman-Wright…” You once called it “a lucky coincidence”; indeed I contacted the right place at the right time. Thank you for the perpetually positive attitude, which appears you also recognized in me, along with quality and capacity. Your effectiveness, pragmatism and wisdom, made everything, always, under control, including of course The Department of BioNuts! More than once I said that you and Chunyan are the best supervisors, and now it’s written black on white!

And that leads me to **Chunyan Zhao**, my co-supervisor. Thank you for your patience, expertise, constructive critique, ideas and solutions, inspiration and endless help with literally everything, from lab bench to office desk, and back. You made all this happen.

**Jan-Åke Gustafsson**, who initially considered the aforementioned e-mail, and gave me this wonderful chance to become a part of the Nuclear Receptor/Orphan Group. We all owe everything to you and ERβ, and I am proud to be a part of your team.

My former supervisors at the Institute of Nuclear Sciences “Vinča” in Belgrade, Serbia. **Vesna Vranić-Mandusić**, for introducing me to the world of ERβ and splice variants. **Bogomir Dimitrijević**, for fantastic lectures and mentorship, which convinced me to follow the mysterious ways of cancer.

To the current members of Karin D-W’s **“Estrogen Signaling Research”** group. This is the perfect opportunity to finally baptize the group and introduce the smart name that Marko coined, conveniently generating the acronym “ESR”. Thank you for that **Marko**, and for being the socializing booster, as well as the designated English lector. **Nina G**, my fellow PhD-fighter through thick and thin, but the best awaits us (unless, I can hear you pointing out, the world really ends this year)! Hui, thank you for posing as my opponent at the group meetings and all your input, help and knowledge. **Min**, good luck with hunting the candidate genes and proteins! **Malin H**, thank you for keeping everything in order. To the “New Blood” - **Jian, Amirhossein** and **Yichun** - I wish the odds of Western blots, Real-time PCRs and knock-downs are always in your favor! **Laura**, have I ever told you that, apart from languages, I love ER splice variants too? **Peik**, thank you for all distractions and completely science-unrelated chats.
To the past members of the group - Zoi, you could always make me smile! Yawen, I will never forget your kindness and laughter. Sadaf, I miss forging big life plans over late lunch! Maria, my Lab Police predecessor, I hope I was worthy of the task in all these years to follow. Lovisa, I wish you stayed a bit longer…

To the members of the big Nuclear Receptor Group, a.k.a. “Orphans”. Lotta “LiLo”, the symbol of friendship and a heart of our Novum clique; this whole experience would not have any substance without you! Tassos, for teaching the ways of tequila in combination with orange and cinnamon and for that refined taste in American cinema and TV shows. Patricia, I miss sharing the office with you, talking about books and travels. Tomas, for mirroring my attitude, sense of humor and love for watches - I suspect that being born on a same day makes such individuals somewhat alike! Amenä, I must mention the fun we had organizing Nov2k07, despite the heated discussions on e.g. menus and curtains! Saioa, for bringing the sunny spirit up here to the North. Eckardt, for the spot-on comments and great jokes during lunches and travels. Knut, for your great Norgen achievement that became synonymous with “fun times”. Agneta and Lars-Arne, for the energy and expertise. To Per, Marion and Kirsten, for all the chats.

To the “Orphans” of the past - Anna E, I knew we’d become buddies right from the start; geek girls come together sooner or later! Karin E, for partying, in the lab, Belgrade and beyond. Karolina, for the aura of positivity. Rezin, för vänskap och innebandy-teorin! Joëlle, for the methylation tutorship, and Ivan - you guys are the coolest. Anders Ström, one of the most helpful and kindest persons I had the opportunity to work with; for tips on the Tet systems in particular. Margaret Warner, for sharing the ERβcx passion. To Nina H, Pia, Marcela, Michel, Cecilia, Sebastian, Christophoros and many, many more…

To the members of the Bioinformatics and Expression Analysis Core Facility (“BEA”): Fredrik and David, for the immense help with the microarrays, as well as with the student courses. Patrick, you were always one of those good corridor spirits, thanks for defeating Baal and never ceasing to amuse. Susann, I miss chatting about the equal love for cats and dogs! Malin N and Marika, for all our Lab Fikas.

To Sam Okret and his group, I’m looking forward to reading more on ERβ from you! Anna I for the relaxed atmosphere in the lab, and Krishan, for all the hilarity and the 30 Rock “antibody”.

To Juha Kere and his group, for a smart tip that greatly helped our SNP study, and for creating a great work environment. Pauliina, for sharing my joie de vivre & vampirisme. To Elo, Kristiina, Isabel, Ghazaleh, Tiina and Myriam, for the great company, both in the cell and lunch rooms!

To my friends throughout Novum: To Johan H “Stålis”, for the nerding and naturally, the Peanut. To Johan D “Deffiz”, for being both a friend and a nemesis! Lois “The Storyteller”, for all the anecdotes, advice and support. Yongtao, for being that glue that brings everyone together. Helmi, for all the interesting topics we discussed. Peter Swoboda, for accomplishing the impossible - being a group leader, a joker, and
pronouncing my name just right. To many, many other friendly colleagues: Kristina, Karin F, Debora, Amir, Ammad, Michelle, Andreas, Tania... To Peter Zaphiropoulos and Lennart Nilsson, for the help and contribution to the Department.

To the Administration staff, for the endless help throughout the years. To Lena, Marie and Monica, my Three Fairy Godmothers. To Patricia, Jessica and Linda, for assisting with tedious poster printers, Visa applications and Macrogen orders. To the almighty Computer guys - Anders, Erik, Jonas and Rikard - for all my battles with PCs, Macs and e-mails. To Inger and Christina, for impeccably dealing with my orders and assuring that no thief lays a hand on them. Till Daniela, för alla våra projekter. Special thanks to Eva, Lars, Gunnel and Inger&Inger.

To my friends outside work. My two Swedish bros: Johan “Joey”, my confidant and brother I never had, Stockholm and I gravely miss you; and David “The Dawg”, my wingman and analyst, for teaching me how to be awesome. To the “sverbengelska” club: Sandra T, my rebound & fallout girl, always there to remind me that home is where the heart is; ljubim te za sav muffing & treligovanje we had tillsammans. Stefan, thanks for supporting the muffing with your trademark dry humor. Ljilja, one day we’ll travel to adventure again! En stor TACK till Kerstin Larsson, Sveriges P3 Radio och Metro Morgontidning, för språkträning! 😊

To John, thank you for being a Dancer, och Wolfie till min Hon Solo; there are no men like you, there’s only you. 🌐

To my beloved ones far away, yet so close. My friends in Belgrade: Mirna “Moo” and Branko “MaX” - for the chain e-mails, chats and online conferences about camera choices, Will Ferrell and HIMYM. I love you guys, and may the Schwarz always be with you! To Mila, my figuratively oldest and literally closest friend. To Ivana, for the traditional coffee walks on Kalemegdan and for actually being the only one curious about the details of my research. To Vlada, my Mulder. To Ivan, for sharing my dream that life can be just like in the movies. Dragana, I am finally earning that “doctressa” title! To my friends scattered around the world: Sandra L in Germany, for reviving my spirit each time we melt Skype until 3am on a workday; and Igor in Italy, for keeping me close to your heart. To my ”A Song of Ice and Fire“ forum lobby, especially Marina “Nemini“ and Jadranka “Queen“, for creatively distracting me when most needed. 🎖

To my family. My aunts Ana and Vera, for being symbols of fun and ingenuity. My cousin Jelena-Katarina, for believing in my scientific mind and for loving Sweden even more than I do! My grandparents, Mara and Ljuba, for the wonderful spirit.

Finally, my parents. My father Miša, the voice of reason, for your hardnosed genes and for believing that I will succeed in everything I plan. And my mother, Zorica. This all is for you and because of you. 🌹

This has been a wonderful journey. For those who don’t know, my last name means “traveler” in Serbian. Some say that happiness is not about the destination, but about the journey, however science seems to be about both…
7 REFERENCES

3. Ripperger T, Gadzieki D, Meindl A, Schlegelberger B: Breast cancer 
   susceptibility: current knowledge and implications for genetic counselling. Eur J Hum 
4. Mavaddat N, Antoniou AC, Easton DF, Garcia-Closas M: Genetic susceptibility 
   incidence of in situ and invasive ductal and lobular breast carcinomas: United States, 
6. Raica M, Jung I, Cimpean AM, Suciu C, Muresan AM: From conventional 
   pathologic diagnosis to the molecular classification of breast carcinoma: are we ready 
7. Schnitt SJ: Classification and prognosis of invasive breast cancer: from morphology 
8. Dontu G, Al-Hajj M, Abdallah WM, Clarke MF, Wicha MS: Stem cells in 
9. Tiede B, Kang Y: From milk to malignancy: the role of mammary stem cells in 
10. Sternlicht MD: Key stages in mammary gland development: the cues that regulate 
11. Stingl J: Estrogen and progesterone in normal mammary gland development and in 
12. LaMarca HL, Rosen JM: Minireview: hormones and mammary cell fate--what 
14. Nakshatri H, Srour EF, Badve S: Breast cancer stem cells and intrinsic subtypes: 
    different polychemotherapy regimens for early breast cancer: meta-analyses of long- 
    term outcome among 100,000 women in 123 randomised trials. Lancet 2011 379:432- 
    44.
16. Hassan MS, Ansari J, Spooner D, Hussain SA: Chemotherapy for breast cancer 
17. Harari D, Yarden Y: Molecular mechanisms underlying ErbB2/HER2 action in 
18. Le XF, Pruefer F, Bast RC, Jr.: HER2-targeting antibodies modulate the cyclin- 
    dependent kinase inhibitor p27Kip1 via multiple signaling pathways. Cell Cycle 2005 
    4:87-95.
    HER-2 receptor and breast cancer: ten years of targeted anti-HER-2 therapy and 


58. Harris HA, Katzenellenbogen JA, Katzenellenbogen BS: Characterization of the biological roles of the estrogen receptors, ERalpha and ERbeta, in estrogen target


84. Welboren WJ, Sweep FC, Span PN, Stunnenberg HG: Genomic actions of estrogen receptor alpha: what are the targets and how are they regulated? *Endocr Relat Cancer* 2009 16:1073-89.


