Estrogen Receptor Target Gene: An Evolving Concept

Jason S. Carroll and Myles Brown

Division of Molecular and Cellular Oncology, Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115

Estrogen receptor (ER) functions as a transcription factor to induce gene expression events sufficient for cell division and breast cancer progression. A significant body of work exists on the identification of ER gene targets and the cofactors that contribute to these transcription events, yet surprisingly little is known of the cis-regulatory elements involved. In this review, we investigate the advances in technology that contribute to a comprehensive understanding of ER target genes and explore recent work identifying cis-regulatory domains that augment transcription of these targets. Specifically, we find that ER association with gene targets results from an association with the pioneer factor FoxA1, responsible for recruitment of ER to the genome. Recruitment of ER to the genome does not occur at promoter proximal regions, but instead involves distal enhancer elements that function to tether the ER complex to the target gene promoters. These advances in technology permit a more detailed investigation of ER activity and may aid in the development of superior drug interventions. (Molecular Endocrinology 20: 1707–1714, 2006)

The identification of the estrogen receptor (ER) by Jensen in 1960 (1) shifted the paradigm of steroid hormone action from an enzymatic one to a model whereby steroids diffuse into cells and interact with a specific receptor to elicit defined biological responses. Pioneering work by O’Malley et al. (2) demonstrated that ER functions primarily as a transcription factor to regulate gene expression at the mRNA level. The cloning of ERα (3, 4) allowed the definition of specific functional domains within the receptor including separable DNA binding, ligand binding, and trans-activation domains, and cemented its characterization as a ligand-dependent transcription factor. In addition, the cloning of the receptor and the availability of various classes of ligands including selective ER modulators such as tamoxifen led to the conclusion that ER regulates gene expression in a cell type, promoter, and ligand-specific fashion.

The more recent identification of a second ER protein (ERβ) (5, 6) raised the possibility of ER subtype and cell-specific gene targets and models of action that involved cooperation between the two ER proteins as well as potential competition have been proposed. The presence and role of ERβ in the mammary gland remains controversial where the antiproliferative effects of the selective ER modulator, tamoxifen, in breast cancer patients relies primarily on the presence and function of ERα. Therefore, ERα independently appears to possess important transcriptional activity in breast cancer cells where the most advanced analyses have been performed, and thus this review will focus on ERα.

Identification of Estrogen-Responsive Elements (ERE)

Early work on the Xenopus vitellogenin gene identified a minimal ERE core sequence: GGTCANNTGTGACC (Ref. 7; reviewed in Ref. 8). This ERE sequence was shown to function in an orientation and distance-independent manner, both of which are properties of an enhancer (8). Controversy still exists concerning ER DNA binding via ERE half sites, although a number of examples exist (9–12). Since the identification of a canonical ERE, several computational approaches have been undertaken to identify target genes based on the presence of EREs within promoter proximal regions (13, 14). In one of the most comprehensive studies, Bourdeau and co-workers screened for all EREs in the human and mouse genomes and identified in excess of 70,000 EREs within the human genome, over 17,000 of which were within 15 kb of mRNA start sites. Elimination of EREs that were not conserved
between the human and mouse genomes reduced the number of gene proximal EREs to 660. A number of these sites were validated as genuine ER interaction sites, supporting to some degree the use of computational models to predict putative ER target genes.

**CLONING OF ESTROGEN TARGET GENES**

A small number of estrogen target genes such as ovalbumin have been known from the earliest recognition of the function of ER as a transcription factor. A second wave of target gene identification involved the use of differential cloning techniques. The cloning of cDNAs induced by estrogen in breast cancer cells led to the identification of a number of gene targets including TFF-1/PS2 (15, 16), a gene expressed primarily in gastric mucosa and breast epithelia; cathepsin D, a lysosomal proteinase (17), and more recently, EBAG9 (18), a gene of unknown function with similarity to a surface antigen gene called RCAS1 (19). The identification of c-Myc (20, 21) and cyclin D1 (22) as estrogen targets provided direct evidence that ER could induce specific gene events that on their own were sufficient to cause cell cycle progression and cell division (23). The subsequent cloning of the approximately 1-kb promoter proximal regions of a number of these target genes into reporter vectors and mutagenesis of various DNA binding elements within the promoter regions led to the hypothesis that nonclassical mechanisms of ER transcription exist that do not involve canonical EREs. Instead, ER-mediated transcription in reporter assays involved other transcriptional elements such as AP-1, NF-κB, cAMP-like elements, and Sp-1 DNA binding motifs (24–26). As such, two independent mechanisms for ER-mediated gene activation exist, namely direct ER-mediated transcription via EREs and indirect or nonclassical mechanisms via other DNA binding elements and potentially involving other transcription factors. These methods required a gene-by-gene approach and focused primarily on proximal promoter regions that were thought to be the mediators of transcriptional regulation.

Other proposed mechanisms of estrogen-regulated transcription involve either an indirect nongenomic function of ER or the action of other membrane-bound nonnuclear receptor family receptors for estrogen. In one model, cytoplasmic or membrane-bound ER can initiate rapid signaling events that ultimately lead to changes in estrogen-regulated target genes. In support of this model are the reported interactions of ERα with a number of signaling proteins including Src, Shc, ras, PI3-kinase, and G proteins (27–31). Recently, however, work has suggested that estrogen can function through the G protein-coupled receptor, GPR30 (32–34). These studies suggest that signaling through GPR30 may play a role in the cellular response to estrogen. Further studies are needed to determine the relative contribution of either or both of these pathways to estrogen signaling.

**SIMULTANEOUS IDENTIFICATION OF MULTIPLE ESTROGEN TARGETS**

The use of differential sequence display proved to be an informative technique to identify multiple ligand-specific gene targets simultaneously within a cell line, or for the comparison of differentially regulated genes between different cell lines or tumors. A novel estrogen-regulated gene with similarities to the human megakaryocyte CD63 antigen mRNA was identified in MCF-7 cells (35), which was subsequently validated as an estrogen-modulated gene in T-47D cells. Also, a homolog of the Na⁺ H⁺ exchange regulatory factor (NHE-RF) was found to be an estrogen-regulated gene in breast cancer cell lines (36). NHE-RF functions to regulate protein kinase A activity and therefore likely contributes to estrogen-mediated signal transduction modulation. The ICERE-1 gene was found to be underrepresented in two ER-positive breast cancer cell lines by differential display (37), which was subsequently confirmed in a larger panel of cell lines. The ICERE-1 gene product was shown to lack similarities to any known proteins but was later shown to partially rescue the drug-resistant phenotype in a melanoma cell line (38), supporting a role for growth control of tumor cells. The XBP-1 transcription factor was one of a number of novel estrogen-regulated genes identified in a separate investigation (39). Interestingly, XBP-1 is known to augment ER-mediated transcription itself, thereby initiating a feedforward pathway (40). A larger study employing differential display identified 127 cDNAs with specific expression in either ER⁺/PR⁺ or ER⁻/PR⁻ tumors, after which eight were shown to be novel transcripts with no similarity to characterized genes (41).

Suppressive subtractive hybridization was also used to identify estrogen-regulated genes in MCF-7 cells. All of these were shown to be negatively regulated by the antiestrogen tamoxifen (42) supporting their identity as direct ER-regulated targets. Within this list was GREB1, which has not yet been assigned a function due to the lack of homology to any known proteins, but was subsequently shown to be a direct ER binding target (43). An independent investigation by Yang et al. (44) revealed a total of 10 differentially expressed transcripts when comparing ER-negative and -positive cell lines, which included a number of known genes including GATA-3, as well as several uncharacterized expressed sequence tags. A similar approach was used to identify 29 cDNAs that were differentially expressed between ER-positive and -negative cell lines (45), including a number of previously identified estrogen targets as well as the Cdk inhibitor p21*Waf1/Cip1*, a gene directly implicated in regulating cell cycle progression (46).

**EXPRESSION MICROARRAY EXPERIMENTS**

The advent of expression microarrays afforded the ability to investigate global gene changes after ligand
treatment. A significant number of studies have been published detailing microarray-based gene changes after nuclear receptor activation, although we will only focus on a handful of studies as examples of the technique. The use of cDNA arrays led to identification of GATA3 as an estrogen-regulated gene target in MCF-7 and T-47D breast cancer cell lines (47), confirming previous data from differential display experiments. Oligonucleotide microarrays have been used in combination with RNA from ZR75-1 cell lines to identify a number of previously identified estrogen targets including pS2/TFF-1, Cathepsin D, RIP140/NRIP-1, and c-myc (48), although a significant number of new targets were identified, including E16, a cationic amino acid transporter. Katzenellenbogen and co-workers (49) conducted a detailed set of microarrays experiments over a time course of estrogen treatment. Of the genes regulated, the highest proportion of estrogen-induced genes were those involved in transcriptional regulation and cell proliferation. An interesting conclusion was that approximately 70% of the changes after estrogen treatment were down-regulated genes, including a number of proapoptotic genes (49), fitting with a model of estrogen-induced cell survival. The mechanism of negative regulation by estrogen was not revealed by these studies and did not distinguish between direct transcriptional inhibition, physiologic squelching by sequestration of limiting factors away from these genes, or induction of inhibitory factors. It is possible that all of these mechanisms may play a role.

SERIAL ANALYSIS OF GENE EXPRESSION (SAGE)

SAGE libraries to identify differentially regulated genes on a scale comparable to microarray analysis, has also been used. These studies have illuminated the role of WISP-2 as a differentially regulated estrogen gene, as well as validation of previously identified ER targets (50). However, these few genes were the fruits of more than 30,000 sequencing reactions in both nontreated and estrogen-treated MCF-7 cells. In support of this study, work from Polyak and co-workers (51) used a similar approach to also validate known estrogen targets as transcripts increased by estrogen treatment of ZR75-1 cells. A number of the other targets identified included the pro-proliferative gene cyclin D1, the antiapoptotic factor TIT-5, and EIT-6. Interestingly, EIT-6 was estrogen induced in more than one breast cancer cell line and was shown to promote colony growth in vitro, supporting its role as a mediator of cell division. A total of 61 tags were observed to change after estrogen treatment, including 22 that were down-regulated. However, approximately 45,000 sequencing events from each library (untreated and estrogen treated) were required to identify these 61 tags, highlighting the large-scale sequencing required to adequately cover transcript changes on a genome-wide scale.

CHROMATIN IMMUNOPRECIPITATION (ChIP) ANALYSIS OF TARGET SITES

The application of ChIP to clarify protein-DNA binding dynamics has provided significant information about the order of protein association with endogenous promoter regions. Previous work from our laboratory and others used ChIP to map protein recruitment to the promoters of TFF-1 and Cathepsin D. A cyclic pattern of ER association was observed on these promoter regions, with maximal recruitment at 45 min after estrogen stimulation. A number of additional proteins subsequently associate with the promoter regions including p300, p160 cofactors, CBP, pCAF, CARM1, and RNA PolII, all of which then cycle off the promoter (52–54). Coincident changes in histone acetylation occur to promote a localized euchromatic environment permissive for transcription. Subsequent rounds of transcription involve a majority of the same proteins, although some differences in protein usage exist. In contrast to the cycling kinetics observed by ChIP, the use of fluorescence recovery after photobleaching to assess the rate of ER association and dissociation from DNA suggests a model of mobility that is measured in seconds rather than almost an hour (55). However, the use of fluorescence recovery after photobleaching does not distinguish the bulk of ER protein from the ER that is associated with gene targets in chromatin and, therefore, may not accurately represent the mobility of transcriptionally functional ER complexes. That said, however, the two models of ER kinetics are not necessarily mutually exclusive. It is possible that ER cycles on and off of chromatin with two kinetic profiles, namely a rapid one that is measured in seconds and an average of the longer changes in ER-chromatin association that can be measured in minutes to hours.

ChIP-BASED METHODS

A more recent approach to identify ER binding sites to define new target genes as well as cis-regulatory regions, used ChIP combined with sequencing. Work by Laganiere et al. (56, 57) using ChIP-cloning identified DNA sequences associated with ER under in vivo conditions, by sequencing cloned DNA fragments that coprecipitate with ER after estrogen stimulation of cell lines. The benefit of this method is that it allows for identification of cis-regulatory regions without bias toward promoter regions or known gene targets. This method has successfully been applied to identify the known target, TFF-1, as well as 11 other targets, including RARA (57). Most of these new ER binding sites
were shown to be bona fide ER binding sites as well as interacting domains for the cofactors, SRC-1 and SRC-3, validating the technique as a method for identifying previously undocumented ER-chromatin interaction sites. However, similar to SAGE-based approaches, large-scale sequencing is essential and all sequences identified need to be validated by directed ChIP and PCR because the presence of enriched DNA sequences are embedded within large amounts of background DNA that will also be precipitated.

A ChIP-SAGE approach was recently published where ChIP of the nuclear receptor coactivator, SRC-3 after estrogen treatment was used to purify pieces of associating chromatin, which were subsequently analyzed by SAGE-based approaches (58). Approximately 7000 tags were sequenced, almost half of which corresponded to repetitive DNA and could not be assigned unique loci. The remaining tags were grouped into regions of DNA containing multiple 21-bp tags, as would be expected based on the average 0.5-

ChIP-CHIP

More recently, ChIP-microarray (ChIP-chip) approaches have been undertaken to identify ER binding sites in an unbiased manner. ER ChIP of MCF-7 cells followed by hybridization of associated chromatin to microarrays that contain probes representing approximately 9000 GC-rich regions, identified 70 putative ER binding sites. GC-rich regions are known to be biased toward the transcriptional start sites of genes (59) and therefore the microarrays likely overrepresent promoter sequences. Within the list of 70 putative sites, a number of enriched DNA binding motifs were identified including ERE half sites and Sp-1 sites.

Recent work from our laboratory used a ChIP-chip approach to identify ER binding sites on a chromosome-wide scale, using ChIP in combination with microarrays that cover the entire nonrepetitive sequence of chromosomes 21 and 22 at 35-bp resolution (60). In this study, we identified 57 ER binding sites across the 35 million bp represented on the microarrays, almost all of which were not in promoter proximal regions, but instead existed up to 150 kb from putative gene targets. These binding sites represent a small fraction of predicted EREs, confirming that the presence of an ERE within DNA is insufficient to determine an ER binding site. We further showed that the ER binding sites function as binding sites for RNA PolII and the p160 cofactor, AIB1, in an estrogen-dependent manner. Although RNA PolII is found associated with the distal ER binding sites, it has been difficult to determine whether these are direct RNA PolII binding sites, or whether RNA PolII is tethered to these regions via a protein complex that involves promoter sequences, or both. However, in at least two tested examples, we could show that the distal enhancer and the promoter of target genes physically interact upon estrogen addition, confirming that the distal cis-regulatory regions function by bringing the proteins at the enhancers in contact with the promoter. Much work has focused on elucidating the mechanisms by which proteins at enhancers interact with and regulate promoter sequences in other biological systems (reviewed in Refs. 61 and 62). The capturing of enhancer-promoter interactions in our system suggests a looping model whereby intervening chromatin is looped out as the proteins at the enhancer interact with the promoters to initiate transcription. Recent work in our laboratory has focused on androgen receptor regulation of the promoter and enhancer (~4 kb upstream) of the PSA target gene. Cloning of the entire enhancer-promoter region into a vector and the introduction of an insulator in between these domains inhibits transcription of a reporter gene (63). This suggests that proteins track from the enhancer to the promoter and that looping between the enhancer and the promoter is not the sole method of communication between the proteins at the enhancer and the promoter region. It is likely that a combination of these mechanisms exists such that chromatin looping can facilitate interaction of the proteins at the enhancers within the promoter proximal region after which RNA PolII tracking occurs on a more localized level to correctly position the transcription machinery at the transcription start site. However, it is possible that the proteins at the enhancer function only to modulate the chromatin structure and nucleosome positions at the promoter region and to promote a transcriptionally permissive state. This is supported by work on the HNF4-α enhancer and promoter region, where the proteins at the enhancer function primarily to modify acetylation at the promoter and do not appear to transfer proteins from the enhancer to the gene proximal region (64).

Using the 57 ER binding sites from our ChIP-chip data, a search for enriched motifs identified two motifs, namely an ERE and a Forkhead motif. Previous work has identified a role for the Forkhead protein, FoxA1/HNF-3α, in glucocorticoid receptor- and androgen receptor-mediated transcription (65, 66). Moreover, a correlation between FoxA1 and ER expression has been previously documented in breast cancer cell lines (67). We identified common recruitment of FoxA1 to almost half of the ER binding sites in chromosomes 21 and 22, although unlike ER, FoxA1 generally was associated with chromatin in the absence of estrogen and dissociated from the DNA after estrogen addition. FoxA1 is known to bind to condensed heterochromatin via its winged helix DNA binding domains and can mimic histone H1 and H5 proteins, thereby functioning as a pioneer factor to
facilitate subsequent protein binding events, including ER binding (68, 69). Coupled with this, FoxA1 has been shown to induce localized chromatin remodeling on its own, suggesting that it not only promotes association of other proteins but also can independently enhance euchromatic conditions. The specific targeted silencing of FoxA1 in breast cancer cells inhibited ER association with the chromatin and abrogated estrogen-mediated transcription on a chromosome-wide level (60), supporting a model whereby FoxA1 is often essential for ER association with cis-regulatory regions.

A subsequent ER ChIP-chip publication has confirmed the finding of a requirement for FoxA1 during ER-mediated transcription (70). However, this study focused primarily on promoter regions of approximately 18,000 genes, but still identified an enrichment of EREs within the bound promoter regions.

Although all the ChIP-chip experiments to date focus on one specific cell line of interest, two pieces of evidence suggest that the ER binding sites we identified in MCF-7 cells play a broader biological role. The first is that directed ChIP of ER in another breast cancer cell line (T47D cells) followed by PCR of a number of newly identified sites confirmed a very high degree of concordance in ligand-induced ER binding sites. Second, there is a significant degree of sequence conservation at the ER binding sites between the human and mouse genomes, with very little sequence identity in the immediate surrounding sequence, suggesting evolutionary conservation at

---

**Fig. 1. A Revised Model of ER-Mediated Transcription**

In breast cancer cells, the Forkhead protein, FoxA1, interacts with cis-regulatory regions in heterochromatin and in combination with adjacent DNA binding elements, such as EREs, facilitates the interaction of ER with chromatin. Subsequent to ER association, recruitment of p160 cofactors, other chromatin remodeling proteins and RNA PolII occur at these distal enhancer sites. Histone modification occurs to promote a permissive chromatin environment. In a number of tested cases, the distal transcription machinery can interact with the promoters of target genes to initiate gene transcription.
these discrete binding regions within the chromatin. However, ER ChIP-chip on a genome-wide scale in different cell lines is required to identify whether the binding sites observed in MCF-7 breast cancer cells are the same binding sites in other cell types, or whether cell type-specific ER binding sites exist, possibly as a result of differences in pioneer factors.

NEW MODEL OF AN ER TARGET GENE

Our recent ChIP-chip analysis leads to a revised model of ER action. This model for ER-mediated transcription (Fig. 1) involves the presence of a pioneer factor such as FoxA1 on chromatin, that in combination with adjacent EREs, facilitates ER association with discrete regions within the genome. These sites are often in regions far from the transcription start site of target genes. Binding of ER to these distal enhancers is followed by the formation of a chromatin loop that promotes the physical contact between the enhancer and the proximal promoter. The juxtaposition of the proteins involved in transcriptional activation with the promoter allows for the initiation of gene transcription.

CONCLUSIONS

The importance of ER as a target of therapy in breast cancer has been the stimulus for understanding both the factors involved in assisting ER in regulating transcription and on identifying the specific gene targets and the DNA elements responsible for activation or inhibition. Surprisingly little is known of the actual cis-regulatory elements involved. The completion of the human genome sequence and the advent of technologies such as tiling arrays for the whole human genome for the first time makes a comprehensive analysis of the genomic targets of ER action possible. The full understanding of the trans-acting factors and cis-regulatory targets of ER action in various estrogen-responsive cell types will support the development of improved selective ER modulators useful for the prevention and treatment of breast cancer and other diseases.

Acknowledgments

We thank Danielle Lynch for her help with graphics and Mitch Lazar, Timothy Geistlinger, and Danielle Lynch for critical reading of the manuscript.

Received August 15, 2005. Accepted December 28, 2005.

Address all correspondence and requests for reprints to: Myles Brown, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115. E-mail: myles_brown@dcri.harvard.edu.

J.S.C. has nothing to declare. M.B. is a consultant to Novartis Pharmaceuticals and has received lecture fees from Wyeth Pharmaceuticals.

REFERENCES

presentation of the EBAG9 gene at 8q23 associated with early-stage breast cancers. Clin Cancer Res 7:3526–3532


27. Prall OW, Rogan EM, Musgrove EA, Watts CK, Sutherland RL 1998 c-Myc or cyclin D1 mimics estrogen effects on cyclin E-Cdk2 activation and cell cycle reentry. Mol Cell Biol 18:4499–4508


17. Prall OW, Rogan EM, Musgrove EA, Watts CK, Sutherland RL 1998 c-Myc or cyclin D1 mimics estrogen effects on cyclin E-Cdk2 activation and cell cycle reentry. Mol Cell Biol 18:4499–4508


