2009-01-01

Generation of a Reporter Library of ERalpha Interacting DNA Sequences

Cormac Jennings
Technological University Dublin

Follow this and additional works at: https://arrow.tudublin.ie/scienmas

Part of the Biochemistry, Biophysics, and Structural Biology Commons

Recommended Citation

This Theses, Masters is brought to you for free and open access by the Science at ARROW@TU Dublin. It has been accepted for inclusion in Masters by an authorized administrator of ARROW@TU Dublin. For more information, please contact yvonne.desmond@tudublin.ie, arrow.admin@tudublin.ie, brian.widdis@tudublin.ie.

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License
Generation of a Reporter Library of ERα Interacting DNA Sequences

By
Cormac Jennings B.Sc

MPhil in Biomedical Science

School of Biological Sciences,
Dublin Institute of Technology,
Kevin Street,
Dublin 8

In conjunction with the FAS Science Challenge Internship programme at
Baylor College of Medicine,
Houston, Texas
Abstract

Oestrogens are a group of steroid hormones of which oestradiol is the major form. Oestrogens are present in men and women and some of their main physiological roles involve female and male reproduction, bone metabolism and homeostasis. Oestrogens carry out their actions by diffusing across the cell membrane and binding to their receptor called the oestrogen receptor (ER), of which two types exist, ERα and ERβ. ERs regulate gene expression through the binding of DNA, most notably the estrogen response element (ERE).

Breast cancer is the most common malignancy in women in industrialised countries and 65 percent of breast cancers are ER positive. Oestrogen is involved in the growth and differentiation of epithelial cells in the mammary glands, and the mitogenic effects of oestrogens on breast epithelial cells are in part due to increased expression of genes involved in the regulation of the cell cycle. Selective oestrogen receptor modulators (SERMs), including Tamoxifen and ICI 182,780, are compounds that are designed to be effective in breast cancer prevention and treatment as they inhibit the proliferative effects of oestrogen that are mediated through the oestrogen receptor. Tamoxifen binds the ER in breast cancer cells and has an antioestrogenic effect on growth regulatory genes, while ICI 182,780 ICI182, 780 induces ER degradation resulting in a marked reduction in the cellular concentration of ER.

The oestrogen receptor binds to EREs to mediate gene expression. However, there are many types of EREs including half site and multiple EREs. It is also known that ER binds co-operatively to other transcription factors also to regulate gene expression.
The aim of this project was to generate an ERE reporter library that would allow questions such as what adjacent transcription factor binding motifs are found co-associated with EREs and, secondly, whether the level of gene expression varied between ERα and ERβ.

To generate a reporter library two methods were used, Chromatin Immuno Precipitation (ChIP) and Dam ID. Using Dam ID ERα sequences were isolated and a reporter library was generated. This yielded information on the types of transcription factor binding motifs that were found in these sequences, for example, ERE half sites, SP1 and AP1. A comparison was made between the levels of expression via ERα and ERβ and it was found that levels were on average two fold higher with ERα.
Declaration

I certify that this thesis which I now submit for examination for the award of MPhil, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or part for another award in any Institute.

The work reported on in this thesis conforms to the principles and requirements of the Institute’s guidelines for ethics in research.

The Institute has permission to keep, lend or copy this thesis in whole or in part, on condition that any such use of the material of the thesis be duly acknowledged.

Signature:__________________________________
Cormac Jennings

Date: ____________________________________
31st July 2009
Acknowledgements

I would like to express profound gratitude to my supervisors; Dr. Bert O’Malley for giving me the opportunity to carry out research in the Molecular and Cellular Biology Department of the Baylor College of Medicine and providing great supervision and advice; Dr David Lonard for his non-stop encouragement and suggestions throughout this research work; Dr Fergus Ryan, Dr Alice McEvoy and Joe Vaughan for their invaluable support and supervision, not to mention incredible patience, in helping me to write up my Masters.

I am also deeply indebted to my laboratory colleagues, Dr Sean McGuire, Vladimir Stanisic, Nejla Yosef and Dr Amber Johnson for their valuable help throughout this study and I wish them all the best with their future research.

I would like to acknowledge the role FAS played in giving me the opportunity to study in the Baylor College of Medicine through the FAS Science Challenge programme. A special word of recognition for the work of John Cahill, Manager of the FAS Science Challenge programme, Grainne Timlin, Dr Denis Headon and Dr Austin Cooney and for all who helped set up this great programme for Irish students allowing us the chance to study in world class institutions.

I would like to say a very special thank you to Dr Pauline Ward for making us feel at home on arrival and throughout our stay in Houston and ensuring we were not wanting for anything. Your hospitality and dedication was greatly appreciated.

I would also like to thank the rest of the “Houston Twelve”, Kieran Harvey, Claire Walshe, Claire Wynne, Aisling Ni hAnnaidh, Nicholas Nolan, Alan Taylor, Richard Piggot, Aoife Gilmartin, Yvonne Smith, Colm Craven and Mark Boylan for making everyday we spent together in Houston so enjoyable and for providing so many cherished memories.

Lastly a special thank you to my family Maura, Michael, Dónal and Micheál for all their support throughout my study. A sincere thanks to my parents for visiting me no matter where my study has taken me. I hope you both enjoyed the trips.
List of Figures

Fig. 1.1 Illustration Outlining Oestradiol Synthesis (Ray et al, 2006)............................12
Fig. 1.2 Chemical Structure of Oestradiol (http://medchem.rutgers.edu).......................14
Fig. 1.3 Oestradiol Levels During the Menstrual Cycle (www.asus.net.pl)....................15
Fig. 1.4 Chemical Structure of Oestradiol (http://medchem.rutgers.edu/MedChemII/steroid) ..................................................................................................................17
Fig. 1.5 Structures of 17β-Oestradiol and its main A- and D-ring metabolites (Mueck et al, 2002) ..................................................................................................................18
Fig. 1.6 Functional Domains of Human ERα (Pavao et al, 2000).................................20
Fig. 1.7 The Amino Acid Residues Making the ‘P box’ and the ‘D Box’ in Oestrogen Receptor ..................................................................................................................21
Fig. 1.8 Structure of the Human ERβ gene, protein and mRNA isoforms (Zhao et al, 2008) .......................................................................................................................23
Fig. 1.9 Functional Domains of human ERβ (Pavao et al, 2000)..................................24
Fig. 1.10 Illustration Outlining the Homology between ERα and ERβ............................25
Fig. 1.11 Physiological Role of Oestrogen Receptors (Gruber et al, 2002) ....................29
Fig. 1.12 Chemical Structure of Tamoxifen (http://www.pharmacology2000.com) .......32
Fig. 1.13 Chemical Structure of Raloxifene (http://www.chemblink.com) ....................33
Fig. 1.14 Chemical Structure of ICI 182,780 (http://www.marquette.edu) ..................34
Fig. 1.15 Structural Representation of hsp90 (Dollins et al, 2007) ..................................35
Fig. 1.16 Structure of ERα DBD on an Oestrogen Response Element (ERE) (http://www.cbs.cnrs.) .................................................................................................38
Fig. 1.17 Schematic diagram of p300 (Kim et al, 2001) ..................................................44
Fig. 1.18 Some of the interactions involved in ERα mediated Transcription (Kim et al, 2001) ..........................................................45
Fig. 1.19 Cyclic Model of ER Transcription Complex Assembly (Shang et al, 2000) ....46
Fig. 1.20 DNA dependent ERα:SP1 Transactivation ....................................................49
Fig. 1.21 DNA Independent ERα:SP1 Transactivation ..................................................50
Fig. 1.22 Illustration Outlining the ChIP Process (http://www.activemotif.com) .........56
Fig. 1.23 Interaction of ERα:Dam Fusion Protein with DNA and methylation of GATC sequences ...........................................................................................................57
Fig. 2.1 pERE-Elb-Luc Luciferase Reporter Vector .........................................................63
Fig. 2.2 A Forward and Reverse Primers used In the Generation of the ERα PCR Product ....................................................................................................................73
Fig. 2.2 B Forward and Reverse Primers used in the Generation of the ERα PCR Product ....................................................................................................................74
Fig. 2.3 Sequence of Adaptors used for Ligation ............................................................85
Fig. 2.4 Sequence of PCR Primer, Annealing and Extension Step ................................87
Fig. 2.5 Map of pCR8/GW/TOPO Entry Vector (www.invitrogen.com) .........................90
Fig. 2.6 Newly Generated pGL3 reporter Destination Vector ......................................94
Fig. 2.7 LR Recombination Reaction .............................................................................96
Fig. 3.1 KpnI and Hind III restriction digestion analysis of plasmids resulting following ligation of Nhel, Bgl II digested pERE-Elb-Luc vector with ChIP isolated DNA sequences ...........................................................................................................101
Fig. 3.2 KpnI and Hind III restriction digestion analysis of plasmids resulting following ligation of Nhel, Bgl II digested pERE-Elb-Luc vector with ChIP isolated DNA sequences ...........................................................................................................101
Fig. 3.3: Luciferase Assay Results following co-transfection of WT ERα expression vector, aa 202/205 ERα mutant expression vector or no expression vector with the pERE-E1b-Luc reporter vector. Transfected cells were treated with Ethanol, Oestradiol or Tamoxifen.

Fig. 3.4: Luciferase Assay Results following co-transfection of WT ERα expression vector, aa 202/205 ERα mutant expression vector or no expression vector with a luciferase reporter vector containing DNA sequence 7 as pulled down using ChIP assay. Transfected cells were treated with Ethanol, Oestradiol or Tamoxifen.

Fig. 3.5: Luciferase Assay Results following co-transfection of WT ERα expression vector, aa 202/205 ERα mutant expression vector or no expression vector with a luciferase reporter vector containing DNA sequence 12 as pulled down using ChIP assay. Transfected cells were treated with Ethanol, Oestradiol or Tamoxifen.

Fig. 3.6: Luciferase Assay Results following co-transfection of WT ERα expression vector, aa 202/205 ERα mutant expression vector or no expression vector with a luciferase reporter vector containing DNA sequence 44 as pulled down using ChIP assay. Transfected cells were treated with Ethanol, Oestradiol or Tamoxifen.

Fig. 3.7: Overview of the DamID Mediated ERα gDNA Interacting Sequence Isolation Process.

Fig. 3.8 A: Forward and Reverse Primers used in the Generation of the ERα PCR Product.

Fig. 3.8 B: Forward and Reverse Primers used in the Generation of the ERα PCR Product.

Fig. 3.9: 2% Agarose gel containing Molecular Weight Ladder, Lane 1 and 2 ERα PCR products. Lane 3 and 4 Dam PCR Products.

Fig. 3.10: 2% Agarose Gel Containing Molecular Weight Marker, Lanes 1 and 2 ERα PCR Products. Lanes 3 and 4 Dam PCR Products and Lanes 6 and 7 ERα:Dam fusion PCR Product.

Fig. 3.11: Overview of the Processes involved in the Generation of Individual ERα and Dam PCR Products, followed by the Generation of the ERα:Dam Fusion PCR Product.

Fig. 3.12: Overview of the Process Involved in the Generation of the ERα:Dam fusion Expression Vector.

Fig. 3.13: Overview of the Process involved in the Generation of the ERα:Dam Fusion Expression Vector.

Fig. 3.14: Restriction Analysis of pCR3.1 ERα:Dam fusion vectors run on 2% Agarose Gel.

Fig. 3.15: Overview of Luciferase Assay for the Functional Testing of the pCR3.1 ERα:Dam Expression Vector.

Fig. 3.16: Luciferase Assay Results Following Co-Transfection of the pERE-E1b-Luc reporter vector and Increasing Concentrations of the ERα:Dam Expression Vector, or Standard Concentration of WT ERα Expression Vector as Indicated. Cells were treated with Ethanol, Oestradiol or Tamoxifen for 24 hours. RLU=Relative Light Units.

Fig. 3.17: Overview of the ERα:Dam Transfection Process and DNA Concentration Measurement.

Fig. 3.18: Overview of the LM-PCR Process.
Fig. 3.19: Agarose Gel Electrophoresis of LM-PCR Products from LM-PCR gDNA isolated from HeLa cells transfected with ERα:Dam and WT ERα expression vectors and Treated with EtOH or Oestradiol for 24 hrs. gDNA was digested with Dpn I where indicated prior to PCR amplification.

Fig. 3.20 Overview of the TOPO Cloning Process.

Fig. 3.21 Overview of the Process of Plasmid Isolation and Subsequent Plasmid Analysis.

Fig. 3.22 EcoR1 Restriction Digest analysis of pCR8 / GW /TOPO plasmids generated following ligation with Sequences Isolated from cells transfected with the ERα:Dam fusion Protein and treated with Ethanol.

Fig. 3.23 EcoR1 Restriction Digest analysis of pCR8 / GW /TOPO plasmids generated following ligation with Sequences Isolated from cells transfected with the ERα:Dam fusion Protein and treated with Oestradiol.

Fig. 3.24 EcoR1 Restriction Digest analysis of pCR8 / GW /TOPO plasmids generated following ligation with Sequences Isolated from cells transfected with the ERα:Dam fusion Protein and treated with Tamoxifen.

Fig. 3.25 Selection of DNA Sequences of Inserts of Human Origin.

Fig. 3.26 Diagram of the newly constructed pGl3 Basic Destination vector with Cassette insert.

Fig. 3.27 Luciferase Assay Results following co-transfection of newly generated pGl3 Basic Luciferase Reporter Constructs, pGL3 Basic and pERE-EIb-Luc with pCR3.1 hERα followed by treatment with Ethanol, Oestradiol and Raloxifene (Luciferase assay performed by David Lonard on constructs).

Fig. 3.28 Luciferase Assay Results following co-transfection of newly generated pGL3 Basic Luciferase Reporter Constructs, pGL3 Basic and pERE-EIb-Luc with the ERα DBD Mutant C202H / C205 expression vector followed by Ethanol, Oestradiol and Raloxifene hormone treatment. (Luciferase assay performed by David Lonard on constructs).

Fig. 3.29 Luciferase Assay Results following co-transfection of newly generated pGL3 Basic Luciferase Reporter Constructs, pGL3 Basic and pERE-EIb-Luc with the pCR3.1 flag ERβ expression vector followed by Ethanol, Oestradiol and Raloxifene hormone treatment. (Luciferase assay performed by David Lonard on constructs).
List of Tables

**Table 3.1** The DNA concentrations and A260/280 readings of isolated gDNA from hormone treated HeLa cells transfected with ERα:Dam and WT ERα.................123

**Table 3.2:** Table showing number of colonies following overnight growth of One Shot Competent E.coli cells transformed with Topo cloning reaction products...........131

**Table 3.3** List of Genes in Close Proximity to the Isolated Sequences, and whether or not these Genes are Regulated by Oestradiol...........................................................146

**Table 3.4** Isolated Sequences Containing Known ERα Receptor Binding Sites........147
Introduction

1.1 Oestrogens

Oestrogens are a group of steroid compounds so called due to their involvement in the oestrous cycle. Oestrogens, like all steroid hormones, diffuse across the cell membrane and when inside the cell they bind to their receptors termed oestrogen receptors (Kim et al, 2001). Oestradiol, oestriol and oestrone are the three types of oestrogens, with oestradiol representing the major oestrogen. Oestrogens are present in both men and women, and they are usually present at significantly higher levels in women of reproductive age.

They function as the primary female sex hormone promoting the development of female secondary sex characteristics such as breasts, the thickening of the endometrium, and other aspects of the menstrual cycle. Oestrogen regulates functions of the reproductive system important to the maturation of sperm in males and it may also be necessary for healthy libido. Oestrogens are also involved in bone growth and homeostasis (Migliaccio et al, 1996).

1.1.1 Oestrogen Synthesis

In females of reproductive age the ovary is the most important site of oestrogen biosynthesis, and this process occurs in a cyclic fashion. The hormones FSH and LH stimulate the production of oestrogen in the ovaries (Bulum et al, 2000). Oestrogen production occurs in the primarily developing follicles in the ovaries and the corpus leuteum. Synthesis of oestrogens starts in the interna cells of the ovaries, by the synthesis of androestenedione from cholesterol. Androstenedione has moderate androgenic activity and it crosses the basal membrane into the surrounding granulosa
cells where it is converted to oestrone or oestradiol, either immediately or through testosterone (Fig. 1.1). The enzyme that catalyses the conversion of testosterone to oestradiol (Fig. 1.1) and of androstenedione to oestrone is aromatase. The aromatase gene is expressed in human tissues and cells such as ovarian granulosa cells, placental syncytiotrophoblast, adipose tissue and skin fibroblasts, and the brain (Bulum et al, 2000).

Intracellular cyclic adenosine monophosphate (cAMP) regulates aromatase expression in the ovary. Upon binding of follicle-stimulating hormone (FSH) to its G-protein-coupled receptor in the granulosa cell membrane, intracellular cAMP levels rise and enhance binding of two critical transcription factors [i.e., steroidogenic factor-1 (SF-1)
and cAMP response element binding protein (CREB)], to the proximal promoter II of the aromatase gene. This results in aromatase expression and consequently oestrogen secretion from the preovulatory follicle (Bulum et al, 2000).

In postmenopausal women the major sites of oestrogen biosynthesis are extraglandular tissues such as skin and adipose tissue. Aromatase expression is controlled by the cytokines interleukin (IL)-6, IL-11, tumor necrosis factor alpha (TNFα) and also glucocorticoids in adipose tissue and skin fibroblasts (Simpson et al, 1994). The major precursor of oestrogen in adipose tissue and skin is androstenedione of adrenal origin (Bulum et al, 2000). Androstenedione is then converted to oestradiol by 17β hydroxysteroid dehydrogenase (17β-HSD) (reductase) activity in these peripheral tissues.

1.1.2 Oestradiol

Oestradiol (Fig. 1.2) is the most thoroughly studied oestrogen and represents the major oestrogen found in humans. Although it is often called the female hormone it is also present in males where it is involved in maturation of sperm amongst other functions. Oestradiol functions mainly in reproductive and sexual processes but it also affects other organs and bone structure. It is the most active of the naturally occurring oestrogens (Jager et al, 2000).

The primary sources of oestradiol in women are the theca and granulosa cells of the ovaries (Gruber et al, 2002). In the ovaries the granulosa cells produce oestradiol by the aromatisation of androstenedione to oestrone, followed by the conversion of oestrone to oestradiol by 17beta-hydroxysteroid reductase. In postmenopausal women then,
oestradiol is no longer solely an endocrine factor. Instead, it is produced in a number of extragonadal sites and acts locally at these sites as a paracrine or even intracrine factor. These sites include the mesenchymal cells of adipose tissue including that of the breast, osteoblasts and chondrocytes of bone, the vascular endothelium and aortic smooth muscle cells, and numerous sites in the brain (Simpson, 2002).

Fig. 1.2 Chemical Structure of Oestradiol [http://medchem.rutgers.edu]

1.1.2.1 Oestradiol Metabolism

In the serum, oestradiol reversibly binds to sex-hormone–binding globulin, a β-globulin, and binds with less affinity to albumin in while only about 2 to 3 percent is free and biologically active (Gruber et al, 2002). Deactivation of oestradiol includes its conversion to the less active oestrogens, oestrone and oestriol. Oestradiol is metabolized by sulfation or glucuronidation in the liver, and the conjugates are excreted via the kidneys. Some of the water-soluble conjugates are excreted via the bile duct, and hydrolysis of some of these conjugates by the intestinal flora and subsequent reabsorption of the estrogen results in an enterohepatic circulation. This enterohepatic circulation contributes to maintaining oestradiol levels (Gruber et al, 2002).
1.1.2.2 Oestradiol Effects

Oestradiol acts as a growth hormone for tissue of the reproductive organs in females, supporting lining of the vagina, the cervical glands, the endometrium and also the lining of the fallopian tubes. Oestradiol also enhances the growth of myometrium and appears necessary to maintain the oocytes in the ovary (Hewitt, 2003).

Oestradiol measures <50 μg/ml at menstruation, rises with follicular development, drops briefly at ovulation, and rises again during the leuteal phase for a second peak where, in conjunction with progesterone, oestradiol prepares the endometrium for implantation. At the end of the leuteal phase oestradiol levels drop to their menstrual levels (Fig. 1.3) unless pregnancy occurs as during pregnancy levels rise due to increased aromatisation by the placenta of prohormones produced by the fetal adrenal gland.

Oestradiol has an important role in the development of female secondary sex characteristics. These changes in secondary sex characteristics are initiated at puberty, are most enhanced during the reproductive years and become less pronounced with declining oestradiol levels after menopause. The main effects of oestradiol in these developments are the enhancement of breast development.

Fig. 1.3 Oestradiol Levels During the Menstrual Cycle (www.asus.net.pl)
1.1.2.3 Medical Applications

Contraceptive devices contain a synthetic form of oestradiol called ethinyloestradiol. Combined forms of hormonal contraception contain ethinyloestradiol and a progestin, act to inhibit LH, FSH and GnRH. Inhibition of these hormones is the mechanism by which these birth control methods work as it leads to prevention of ovulation and thus prevention of pregnancy.

1.1.3 Oestriol

Oestriol (Fig. 1.4) is the least potent member of the oestrogenic hormone family (Jager et al, 2000). The low estrogenic potency of oestriol in clinical use is due to the short half-life of the binding at the estrogen receptor (Mueck et al, 2002). Oestriol levels in non-pregnant women do not change much after menopause and also levels of oestriol are not significantly different from levels in men.

1.1.3.1 Oestriol Metabolism

Oestriol is formed primarily by the breakdown of 16-hydroxyestrone (Fig. 1.5) and is regarded as the main terminal product of oestradiol metabolism (Mueck et al, 2002). Oestriol is only produced in significant amounts during pregnancy when it is synthesized in the placental syncytiotrophoblasts by aromatisation of 16α-Hydroxyandrostenedione. The latter compound is derived from 16α-Hydroxyepiandrosterone sulfate, which is produced by the fetal liver and desulfated in the placenta. 16α-Hydroxyepiandrosterone sulfate in turn is derived from dehydroepiandrosterone sulfate produced in the fetal adrenal gland. The combination of
the fetal adrenal gland and liver and the placenta has been referred to as the "fetoplacental unit of steroid biosynthesis" (Gruber et al, 2002).

![Chemical Structure of Oestradiol](http://medchem.rutgers.edu/MedChemII/steroid)

**1.1.4 Oestrone**

Oestrone is the third member of the oestrogenic hormones and is excreted by the ovary. Oestrone is the most prevalent oestrogen in postmenopausal women (Jager et al, 2000).

**1.1.4.1 Oestrone Metabolism**

Oestradiol is converted to oestrone by oxidation in the C17 position, a process that is reversible. The balance is generally in favor of oestrone formation, a situation also characterized by the fact that oestradiol breakdown to oestrone takes place rapidly, but the back reduction of oestrone to oestradiol much more slowly (Mueck et al, 2002).

From oestrone onwards, breakdown continues by two different pathways, namely, by hydroxylation of the A-ring on the one hand and the D-ring on the other. Once these products are formed, they usually can no longer be reduced back to oestrone. The main
metabolites formed by A-ring metabolism are 2-hydroxyestrone and 4-hydroxyestrone, and by D-ring metabolism 16α-hydroxyestrone and oestriol (Mueck et al, 2002).

16α-Hydroxyestrone (Fig. 1.5) was regarded mainly as a biologically unimportant intermediate for the formation of oestriol; however as the amounts of oestrone and oestriol excreted in the urine, are approximately of the same quantitative order, this suggests that 16α-Hydroxyestrone, like estriol, is a terminal product of human oestrogen metabolism (Mueck et al, 2002). Kinetic studies have shown that 2-Hydroxyestrone (Fig. 1.5) is metabolized very rapidly in the blood, more rapidly than other known natural steroids. 4-Hydroxyestrone (Fig. 1.5) is detected only in very small amounts in humans; however it still possesses estrogenic properties (Mueck et al, 2002).
1.2 Oestrogen Receptor

Oestrogen Receptor is a ligand activated receptor which is a member of the steroid / nuclear receptor superfamily that has 60 members (Lander et al, 2001). The vast majority of the members of the steroid hormone receptor superfamily are ligand dependent transactivators. There are two isoforms of Oestrogen Receptor, Oestrogen Receptor α (ERα) and Oestrogen Receptor β (ERβ). These are Class I nuclear receptors that bind to DNA as homodimers. In the absence of hormone, ER exists in an unactivated, untransformed state, as a monomer, complexed with heat shock proteins. In response to oestrogen hormone binding to ER, the receptor undergoes physico-chemical changes including phosphorylation at specific serine and tyrosine residues that are accompanied by conformational changes (Pavao et al, 2000).

These conformational changes result in the dissociation of heat shock proteins from the activated complex and the formation of a homodimer with high affinity for oestradiol and DNA. The activated ER homodimer binds to a specific hormone response element called oestrogen response element (ERE), which is located in the 5’ flanking regions of oestrogen responsive genes, regulating their transcriptional activity through various pathways, including the recruitment of coactivators and corepressors. ER is not dependent on DNA interactions to regulate transcriptional activity as it can do so by binding to other transcriptional factors such as AP-1 and Sp1 (Pavao et al, 2000).

1.2.1 Oestrogen Receptor alpha (ERα)

ERα was first identified in the 1960s (Jensen et al, 1962), shown to be a ligand activated transcription factor (O’Malley et al, 1968) and cloned in 1986 (Greene et al, 1986). Human ERα has a molecular weight of 66 – 70 kDa and is comprised of 595
amino acids (Pavao et al, 2001). The ER\(\alpha\) gene is located on the long arm of chromosome 6 (Enmark et al, 1999). ER\(\alpha\) is localised in breast, brain, cardiovascular system and bone, as well as the urigenital tract and it is the predominant ER subtype in the liver and uterus (Taylor et al, 2000).

The recognition of 6 different functional regions (A-F) was made possible through cloning and functional studies with ER\(\alpha\) cDNAs from different species. These different functional regions were found to be responsible for the following specific functions: DNA binding, ligand binding, dimerisation, protein binding and transcriptional activation (Fig. 1.6).

The hypervariable A/B domain (aa 1-184) located in the amino terminal region of ER\(\alpha\) exhibits very little conservation between species (Krust et al, 1986) and also among the nuclear receptor superfamily (Seagraves, 1991). The hypervariable A/B domain contains ligand independent activation function 1 (AF1), and is thought to be important for transactivation and also responsible for gene and cell specificity. The AF1 region of ER\(\alpha\) interacts with differential transcription regulators and coactivators that affect ligand independent transactivation. Kinase dependent phosphorylation also regulates
the activity of AF-1, and these individual pathways vary between promoter type (Tzukerman et al, 1994).

Region C (aa 185-263) contains the DNA binding domain (DBD) and this region is highly conserved across species and also members of the nuclear receptor superfamily (Kumar et al, 1987). The DBD consists of two functionally distinct zinc-finger motifs (CI and CII) spanning 60-70 amino acids and is responsible to specific binding of the receptor to EREs located in target genes (Gronemeyer, 1991). Each zinc finger contains four cysteine molecules that coordinate the binding of a zinc atom. At the base of the CI zinc finger is a region called the “P box” that contains three specific amino acids responsible for the DNA binding specificity (Nelson et al, 1995) (Fig. 1.7).

![Fig. 1.7 The Amino Acid Residues Making the ‘P box’ and the ‘D Box’ in Oestrogen Receptor](Ruff et al, 2000)

Located in the second zinc finger are residues that form the so-called “D box” and are involved in dimerisation. The DBD also contains two α helices, one of which binds the major groove of DNA, making specific contacts and is known as the recognition helix, while the other helix spans COOH terminus of the second zinc finger and forms a right angle with the recognition helix (Aranda et al, 2001).
Region D (aa 264-302) is the hinge region, which separates the DBD and the ligand binding domain (LBD) (Pavao et al, 2001). The flexibility of the secondary structure of region D is thought to allow the receptor to undergo conformational changes that occur during activation. It is also thought to be important in receptor dimerisation (Kumar et al, 1988) and to facilitate rotation of the DBD. Nuclear localization signals (NLS), responsible for nuclear localization, are contained in region D and the carboxy-terminal (C termial) portion of region C.

Region E (aa 303-553) is located in the C-terminal portion of the receptor and is conserved among species. The E region contains the LBD, which consists of 12 $\alpha$-helices, H1 to H12 arranged together in an antiparallel, three-layered sandwich and also includes $\beta$ strands. Helices H1 to H11 form a hydrophobic pocket responsible for ligand binding whose entrance is guarded by H12 (Bourguet et al, 1995). Region E also contains AF-2, which is a second activation functional domain, and is involved in transactivation in conjunction with the A/B domain. The AF-2 interaction surface is composed of amino acids in helix 3, 4, 5 and 12 and, upon ligand binding, the position of helix 12 is altered.

Also located in Region E is the receptor dimerisation domain and domains for binding of coactivators, co-repressors and heat shock proteins and another nuclear localization signal (Nilsson et al, 2001). It has been shown through deletion and mutation studies that ER$\alpha$ dimerisation is mediated through helices 7-10 (Fawell et al, 1990, Lees et al, 1999).
Region F (aa 554 – 595) is located at the C-terminal. Its exact function is unknown. But in some members of the steroid nuclear receptor family it is thought to play a role in distinguishing between agonist and antagonist binding to the receptor molecule (Pavao et al, 2001).

### 1.2.2 Oestrogen Receptor Beta (ERβ)

ERα was thought to be the only receptor for oestrogen. However, in 1996 a second genetically distinct receptor, ERβ, was identified and cloned from rat prostate and ovary (Kuiper et al, 1996) and subsequently the human form was cloned (Enmark et al, 1997) (Fig. 1.8). The full-length hERβ protein is 530 amino acids and is 58-62 kDa in size. Nevertheless, it has been shown to be produced in multiple isoforms (Peterson et al, 1998). The ERβ gene is located at the q22-24 band of chromosome 14. ERβ is localised in breast, brain, cardiovascular system, bone as well as the argental tract and it is the predominant of the two ER isoforms in the colon brain and prostate (Taylor et al, 2000).

*Fig. 1.8 Structure of the Human ERβ gene, protein and mRNA isoforms (Zhao et al, 2008)*
ERβ also contains 6 functional domains (Fig. 1.9), A-F. The A/B domain (aa 1-148) contains a repressor domain instead of an AF-1 activation domain. Upon deletion of this repressor domain the overall transcriptional activity of the receptor increases. The A/B domain is the least conserved of the domains between ERα and ERβ with only 30% similarity at the amino acid level. The C domain (DBD) is highly homologous (96%) between ERβ and ERα and contains a dimerisation domain and transactivation domain and the DBD (Pavao et al, 2001). The amino acid sequence of the “P box” is identical between ERα and ERβ. Also located between the C and D domain (hinge region) is a nuclear translocation domain. The hinge region is not well conserved between the two receptors.

![Fig. 1.9 Functional Domains of human ERβ (Pavao et al, 2000)](image)

The E region (aa 304-500) contains a dimerisation domain, a transactivation domain and a nuclear translocation domain as well as the LBD and has only 59% homology between ERα and ERβ (Fig. 1.10) (Pavao et al, 2001). The F region is found between amino acids 500 – 530.
Fig. 1.10 Illustration Outlining the Homology between ERα and ERβ

(http://www.cellscience.com/reviews4/Molecular_Basis_Steroid_Action_Prostate.html)
1.3 Physiological Roles of Oestrogen Receptors

Oestrogen receptors control a variety of different physiological processes (Fig. 1.11) and are also implicated in diseases. The following are the physiological processes in which oestrogen receptors are involved.

1.3.1 Female and Male Production

The lobular units of the terminal ducts of the breast tissue of young women are highly responsive to oestrogen. In breast tissue, oestrogens stimulate the growth and differentiation of the ductal epithelium, induce mitotic activity of ductal cylindric cells, and stimulate the growth of connective tissue (Gruber et al, 2002). ERα and ERβ have different roles and distributions in the mammary glands. ERα in mammary stroma cells is involved in the oestrogen proliferative response in the epithelium, which involves the stroma-derived hepatocyte growth factor (HGF). Integrin signaling is thought to interact with HGF and oestrogen to modulate their proliferative effects. Differentiation of the epithelium is mediated by the expression of progesterone by oestrogen bound ERα. The density of oestrogen receptors in breast tissue is highest in the follicular phase of the menstrual cycle and falls after ovulation (Gruber et al, 2002).

Oestrogens are essential for fertility, and ERs are involved in ovulation, implantation, pregnancy maintenance and childbirth. It is thought that ERα may mediate the proliferative effects, while ERβ mediates the differentiative effects of oestrogen within the follicles.

Oestrogen receptors are abundant throughout the male reproductive tract, but ERα is primarily localised in the efferent ductal epithelium, where its expression is even more
abundant than even the female reproductive tract. The primary function of oestrogen in the male tract is the regulation of fluid reabsorption in the efferent ductules via ERα. Oestrogen is also responsible for maintaining a differentiated epithelial morphology. Knockout of the receptor results in dilation of cauda epididymal sperm, disruption of sperm morphology, inhibition of sodium transport and subsequent water reabsorption and eventual decreased fertility (Hess, 2003).

1.3.2 Skeletal System

Oestrogens have an important role in bone metabolism and homeostasis, with effects on skeletal growth as well as bone maturation. In adolescence oestrogens are involved in the modeling of bone, initializing pubertal bone growth and limiting longitudinal bone growth in women (Migliaccio et al, 1996). In adults oestrogens are important in bone maintenance by promoting bone formation through osteoblasts, and repressing bone resorption via osteoclasts. ERα is involved in the growth promoting effects of oestrogens and ERβ is involved during puberty, limiting longitudinal and radial bone growth in females (Nilsson et al, 2001).

Oestradiol regulates several factors that are important in regulating differentiation and function of osteoclasts and osteoblasts. Oestradiol stimulates the synthesis and secretion of anabolic growth factor IGF-1 in osteoblasts, and inhibits the cytokines IL-1 (interleukin) and tumour necrosis factor and IL-6, which are involved in bone resorption (Roodman et al, 1996). Osteoclastic function and differentiation is regulated by osteoprotegrin (OPG), osteoclast differentiation factor (ODF) and receptor activator of NFκB (RANK). RANK interacts with ODF, and this interaction allows differentiation and activation of osteoclasts. Oestradiol stimulates the synthesis and expression of
OPG (Hofbauer et al, 1999), which is a protein that is involved in the inhibition of osteoclast function as OPG secreted by osteoblasts is an endogenous decoy receptor for ODF and it competes with RANK for binding of ODF and thus inhibits osteoclast formation.

The suppression of osteoclastic bone resorption and stimulation of osteoblastic bone formation form the basis for the bone-preserving effects of oestradiol. ER\(\alpha\) is thought to mediate the growth-promoting effects of E2 but is not thought to be involved in the maintenance of trabecular bone. ER\(\beta\) is thought to carry out its actions during pubertal growth by terminating the growth spurt in females, limiting longitudinal and radial bone growth (Nilsson et al, 2001). The action of oestradiol on bone maintenance and homeostasis is evident in women as oestrogen levels decrease following menopause and this decrease is associated with osteoporosis.

1.3.3 Central Nervous System

Oestrogens are also thought to have neuroprotective actions. In brain tissue from adult rats, oestrogens induce synaptic and dendritic remodeling and cause glial activation. In neurons of the hippocampus, an area involved in memory, oestrogens increase the density of \(N\)-methyl-D-aspartate receptors and increase neuronal sensitivity to input mediated by these receptors. In cultured human neuroblastoma cells, oestrogens have neuroprotective effects and reduce the generation of beta-amyloid peptides (Gruber et al, 2002).
1.3.4 Vascular Effects

Oestrogens are thought to be natural vasoprotective agents. Oestrogen receptors have been detected in smooth-muscle cells of coronary arteries and endothelial cells in various sites. Oestrogens cause short-term vasodilation by increasing the formation and release of nitric oxide and prostacyclin in endothelial cells. They also reduce vascular smooth-muscle tone by opening specific calcium channels through a mechanism that is dependent on cyclic guanosine monophosphate (Gruber et al, 2002).

Fig. 1.11 Physiological Role of Oestrogen Receptors (Gruber et al, 2002)
1.4 Selective Oestrogen Receptor Modulators

Oestrogen is important for many developmental processes in both men and women, as it affects growth, differentiation and function of tissues of the reproductive system. It also plays important roles in maintaining bone density and protecting against osteoporosis. However, its aberrant expression also has a role to play in breast cancer due to its effect on growth. Prolonged stimulation of oestrogen receptor on breast ductal epithelium contributes to the development and progression of breast cancer. In these cases treatments that are designed to block the effects of oestrogens are important options in halting cancer progression (Kent et al, 2000).

Selective oestrogen receptor modulators (SERMs) are a type of compound, which are designed to be effective in breast cancer prevention and treatment as they inhibit the proliferative effects of oestrogen that are mediated through the oestrogen receptor. SERMs bind the oestrogen receptor, alter its conformation and, thereby, facilitate the binding of coregulatory proteins that can activate or repress transcriptional activation of oestrogen target genes (Kent et al, 2000).

SERMs are divided into three major categories:

1. Triphenylethylene derivatives
2. Other nonsteroidal compounds
3. Steroidal compounds that have more complete antioestrogenic activity
1.4.1 Triphenylethylenes

Tamoxifen

Tamoxifen (Fig. 1.12) was synthesized in the 1960s and demonstrated antiproliferative effects in the breast (Jordan, 1994). Tamoxifen was shown to be an effective therapy in patients with metastatic breast cancer, whose tumors expressed ER (Osborne, 1998). The major metabolites of tamoxifen are N-desmethyltamoxifen and trans-4-hydroxytamoxifen, which has an affinity for ER similar to that of 17-β oestradiol (Buckley et al, 1989).

Tamoxifen has oestrogenic as well as anti-estrogenic effects. It has agonist effects on bone (Love et al, 1992), blood lipids and also the endometrium (Osborne, 1998) while having antioestrogenic effects on the CNS and vaginal mucosa. The mixed oestrogenic and antioestrogenic effects of Tamoxifen are species, tissue and cell dependent. Tamoxifen has also been shown to have oestrogenic effects on certain genes while having antioestrogenic effects on other genes in the same cell (Osborne et al, 2000).

Tamoxifen binds to the ER in breast cancer cells and antagonizes the effect of oestrogen on a variety of growth-regulatory genes (Dhingra, 1999). The predominant effect it has is cytostatic with the induction of a G1 cell cycle block thereby slowing cell proliferation, thus inhibiting cancer growth (Osborne, 1998). Tamoxifen’s dual activities provide some advantages for women receiving Tamoxifen treatment. Not only is there a beneficial effect through inhibition of proliferation of breast ductal epithelium and breast cancer, but due to Tamoxifen’s oestrogenic effects in certain tissues, women may also benefit from maintenance of bone density and a reduction in cholesterol. The associated reduction in cholesterol is due to its oestrogenic activity in
the liver, as serum concentrations of total cholesterol and low-density lipoprotein are reduced by tamoxifen, while the preservation of bone density is due to Tamoxifen’s oestrogenic effect in bone.

However, due to its varying activity in different tissues, there are side effects associated with Tamoxifen. The most serious of these side effects is associated with its oestrogenic activity in the endometrium, which can result in endometrium hyperplasia and also low-grade endometrial cancers (Wilking et al, 1997). Also due to its antioestrogenic effects on the vaginal mucosa, menopausal symptoms can result (Love et al, 1991).

![Chemical Structure of Tamoxifen](http://www.pharmacology2000.com)

**Fig. 1.12 Chemical Structure of Tamoxifen** (http://www.pharmacology2000.com)

### 1.4.2 Other Nonsteroidal Compounds

**Raloxifene**

Raloxifene (Fig. 1.13) is a benzothiophene derivative and it binds to the ER with an equal affinity to that of oestradiol (Black et al, 1983). Raloxifene has similar oestrogenic activity to that of Tamoxifen, with the exception of the endometrium where Raloxifene possesses less oestrogen agonist activity (Balfour et al, 1998). Raloxifene is
an inhibitor of cultured breast cancer cells and it possesses antitumor activity in rat mammary tumor models similar to that of Tamoxifen (Poulin et al, 1989).

Development of Raloxifene for its use in breast cancer was stopped. However, it was found to maintain bone density due to its oestrogenic effect on bone; it was developed for osteoporosis for which it now an approved drug.

Fig. 1.13 Chemical Structure of Raloxifene (http://www.chemblink.com)

1.4.3 Steroidal Antioestrogens

ICI182, 780

ICI182, 780 (Fig. 1.14) is a derivative of oestradiol with a long, hydrophobic side chain at the 7 α position. It has a pure antioestrogenic profile on all genes and in all tissues (Wakeling et al, 1991). ICI182, 780 antioestrogenic mechanism is based on it’s blocking of ER transactivation coming from both the AF-1 and AF-2 domains (Wakeling et al, 1995). ICI182, 780 induces ER degradation resulting in a marked
reduction in the cellular concentration of ER (Davouis et al, 1992), and has been shown to impair ER dimerisation (Fawell et al, 1990). Due to its antioestrogenic effects on the cellular concentration of ER, ICI182, 780 is a potent inhibitor of transcription of oestrogen-regulated genes.

ICI182, 780 has a binding affinity to ER similar to oestradiol, but 100 times greater than Tamoxifen and in preclinical models ICI182, 780 displays antitumor activity. It has been shown that certain Tamoxifen-resistant cell lines may remain sensitive to growth inhibition by ICI182, 780 (Lykkesfelt et al, 1994), and also when MCF-7 human breast cancer cells are grown in an athymic mouse model, ICI182, 780 is a much more potent inhibitor of tumorigenesis (Osborne et al, 1995), and it is displays more potency in inducing tumor regression in mice with established tumors.

ICI182, 780 may not cross the blood brain barrier and, therefore, might not cause the hot flushes that are associated with other SERMs (Wade et al, 1993). All in all due to its antioestrogenic profile ICI182, 780 may not be the most desirable SERM; however it may be a superior antitumor agent.

![Chemical Structure of ICI 182,780](http://www.marquette.edu)

*Fig. 1.14 Chemical Structure of ICI 182,780* (http://www.marquette.edu)
1.5 Oestrogen Receptor Activation

Oestrogen receptor is bound by the 90 kDa heat shock protein (hsp 90) in the absence of ligand, which is abundant in most cells and is further increased by heat or cellular stress (Fig. 1.15). ER is recovered from hormone free target cells extracted in low salt medium in the inactive 8-9S form, complexed with a set of proteins including a dimer of hsp90, a 59kDa immunophilin protein of the FK 506-binding protein class and a 23-kDa protein (Smith et al, 1993). Of these proteins that are extracted complexed to inactive ER, only hsp90 is reported to have an intrinsic capability of specific receptor binding. In hormone treated cells ER is tightly associated with nuclear components, and ER can be extracted as the sole constituent in a smaller 4-5S form, free of any detectable hsp90 (Sabbah et al, 1996). These findings have led to the opinion that ligand binding to the ER promotes the dissociation of hsp90, allowing the receptor to bind to the EREs of regulated genes (Baulieu, 1987).

Fig. 1.15 Structural Representation of hsp90 (Dollins et al, 2007)
ERs DNA binding activity is modulated by the concentration of hsp90 as at low concentrations the receptor is capable of forming complexes with the ERE while at higher concentrations of hsp90 the binding activity of ER is inhibited. This shows that the binding of ER to ERE is inversely dependent to the concentration of hsp90, illustrating that hsp90 does not suppress ER function merely by steric hindrance as hsp90 is capable of dissociating ER from its cognate ERE by a dynamic and specific process rather than by simple trapping. In order for hsp90 to form a complex with ER it needs to bind to both the LBD and the highly positively charged region situated at the C-terminus of the DBD, between amino acids 251 and 271 (Sabbah et al, 1996).

It is thought hsp90 may play multiple roles in the binding of ER. At physiological temperatures hsp90 might have an important role in signal transduction by interacting with and inducing and / or stabilizing an active form of the receptor and also preventing unproductive interactions of the receptor with other proteins (Sabbah et al, 1996).

At elevated temperatures and under conditions of cellular stress, the reversibility of ER binding to DNA raises the possibility that the increased nuclear concentration of hsp90 might interfere negatively in the regulation of transcription as it has been shown that an increasing concentration of nuclear hsp90 is correlated with repression of transcription by nuclear steroid receptors (Sabbah et al, 1996).
1.6 Oestrogen Receptor DNA interactions

ER mediated gene expression is a multistep process involving the conversion of ER from an inactive state to a transcriptionally active state and, followed by this active form, binding to its response element located in the promoter of oestrogen target genes, termed the oestrogen response element (ERE), and recruitment of cofactors leading to gene transactivation (Martinez et al, 1991). This process is initiated by a conformational change in ER upon oestrogen binding, its dissociation from chaperone proteins and its dimerisation.

The derived minimal consensual ERE sequence is a 13 bp pallindromic inverted repeat 5’GGTCAnnnTGACC-3’ (Cowley et al, 1999) and differs in only two bp in each half site from the GRE (Klock, 1987). The ERE was first identified by aligning sequences with shared homologies in the 5’flanking sequences of the oestrogen-regulated vitellogenin genes A1, A2, B1 and B2 from Xenopus laevis and the chicken apo-VLDLI I gene (Walker et al, 1984).

Specific contacts between the ER dimer and the sugar-phosphate backbone of the ERE are important in sequence recognition and high affinity binding (Koszewski et al, 1991). Each ER monomer is bound to DNA in the major groove with the dimer located predominantly on one face of the double helix (Fig. 1.16).
The specific amino acids within the “P Box” of zinc finger C I (Fig. 1.16) interact in the major groove in a sequence specific manner (Mader et al, 1989). The fourth base pair of the ERE half site (AGG\text{TCA}) provides a positive contact for the P-box, whereas the third base pair (AG\text{GTC}) provides binding energy (Schwabe et al, 1995). The C II zinc finger is involved in half-site ERE spacing recognition and ER dimerisation (Martinez et al, 1989).

The interaction of ER\(\alpha\) with the ERE has been studied using phosphate methylation, ethylation and thymine interference assays. Phosphate methylation assays have shown that ER\(\alpha\) forms the strongest interaction with the underlined nucleotides 5’-GGTCAGCGTGACC-3’ (Koszewski et al, 1991), whereas, the ethylation and thymine interference assays indicate that ER\(\alpha\) contacts the underlined nucleotides in the chicken vitellogenin II ERE 5’-CTGGTCAGCGCTGACCGG-3’ (Obourn et al, 1993). Thus nucleotide recognition by ER\(\alpha\) differs with varying techniques.

Maximal binding of ER\(\alpha\) to an ERE requires ligand binding as this is thought to stabilize the ER\(\alpha\) - ERE interactions (Gronemeyer et al, 1991). The HMG proteins,
HMG-1 and HMG-2, which are architectural proteins involved in chromatin function, have been shown to stabilize ER\(\alpha\)-ERE binding. The way they are thought to facilitate this increased stability is by inducing structural changes in the target DNA that enhance ER-ERE binding (Bustin et al, 1996).

Most oestrogen-regulated genes contain single or multiple copies of EREs that deviate from the consensus ERE by one or more nucleotides (Anolik et al, 1995). These EREs still enable oestrogen regulated target gene expression through ER; however they are less potent enhancers of transcription than the consensus ERE. The more nucleotide changes there are from the consensus within a half-site ERE palindrome, the lower the ER\(\alpha\) binding affinity and transcriptional activity. Furthermore, EREs in which nucleotides are altered in each arm of the palindrome show lower transcriptional activity than EREs containing alterations in only one half of the palindrome (Driscoll et al, 1998). Also the amount of transcriptional activity detected from the same ERE varies between cell types, indicating that cell specific factors, including the presence or absence of coactivators, corepressors and also the concentration of these proteins, plays an important role in the level of transcriptional activity (Klinge, 2001).

Each type of ligand and ERE induces different conformational changes in ERs. However, ligand binding to ER does not affect the binding specificity or affinity of ERs to EREs. This suggests that ERE and ligand induced ER conformational changes are not integrated to provide ERs with distinct functional features. It is thought that the binding affinity of ER to an ERE effects the extent of cofactor interaction, while the ligand induced structural changes on ER effects the affinity of ER to a cofactor. Thus ER ligand and ERE-induced conformational changes are not interrelated as the structure
of the ligand dictates the affinity of ER for cofactors independently from the ERE sequence (Yi et al, 2002).

Oestradiol induced conformational changes in the LBD of the ER are important for the efficient recruitment of cofactors such as SRC-1, TIF-1 and TIF-2 which interact with the AF-2 domain. Tamoxifen and ICI182, 780 impede the interactions of both subtypes of receptors with cofactors by inducing conformations in the LBD that are distinct from those of the unliganded or oestradiol bound receptor (Yi et al, 2002).

The effects of oestradiol, Tamoxifen and ICI182, 780 on the pattern of AF-2 dependent cofactor interactions with the ER are independent from the identity of ERE sequences. The amount of cofactor recruitment is ultimately determined by the ERE sequence. This is because the sequence is critical for the affinity of ER – ERE interaction, thus the relative amount of receptor binding, therefore the amount of cofactor recruitment, which correlates with transcription. ERE induced changes in the receptor conformation do not alter interaction of the oestradiol-ERα complex (Yi et al, 2002).

The extent of ligand or AF-2-dependent cofactor recruitment by both receptor subtypes is altered by the two independent factors: ER ligands and ERE sequences. Following ligand binding, interaction of the ERs with cofactors is affected by the ERE sequence, which determines the binding affinity and consequently the relative amount of ERs that use the same contacts in DNA. The interference of cofactor interaction with ERs by an antagonist occurs independently of ERE sequence. This is due to the fact that the ERE sequence influences the amount of receptor binding to the DNA, and it also indirectly influences the amount of coactivator that is recruited to the ER bound to an ERE.
sequence. The ability of ER to differentially recruit a cofactor could contribute to ER subtype-specific gene responses (Yi et al, 2002).

ERα shows higher transcriptional activation than ERβ (Tremblay et al, 1997). ERα binds to the Xenopus vitellogenin A2 ERE with higher affinity than ERβ (Klinge, 2001). Although both receptors show a similar binding specificity to EREs, ERα binds to an ERE with an approximately 2-fold higher affinity than ERβ. ERα and ERβ utilise the same nucleotides to bind to EREs, despite their differences in affinity (Klinge, 2001). The ERα:ERβ heterodimer also binds with an affinity similar to that of ERα rather than ERβ. ERα binds the Xenopus vitellogenin A2 ERE with higher affinity than any other natural ERE (Klinge, 2001).

Flanking sequences also impact ER binding to an ERE and transcriptional activation. A study carried out on genes that are highly upregulated by oestrogen, showed that the ERE in the promoter sequences of these genes is flanked by an AT rich region (Anoilk et al, 1993). These flanking AT rich regions enhance ER binding to the ERE and one possible way that the AT rich regions may alter the binding affinity is by altering the DNA conformation. AT enriched regions of DNA are more easily deformed than random DNA and, as ERα binding to an ERE results in the bending of DNA toward the major groove, these AT rich regions could enhance deformation (Wang et al, 1987). This DNA bending is thought to facilitate transcription as it facilitates interactions with components of the transcription complex bound at a different site (Kerpolla et al, 1991).
1.7 Mechanism of Transcriptional Activation by ER

Target gene expression following the activation of ER in response to oestradiol or other agonists is thought to occur by one of two mechanisms. In the first mechanism the ligand bound homodimer binds directly to its response element termed the oestrogen response element (ERE). Once bound to this DNA response element, the ER homodimer is able to recruit and interact directly both with coactivator proteins and with components of the RNA polymerase II transcriptional initiation complex resulting in enhanced transcription.

In the second mechanism ER interacts with another DNA bound transcription factor. This binding of ER may stabilize the binding of the transcription factor to the DNA, and also may act by recruiting additional coactivators that have a role in transcription initiation. In this mechanism the ER homodimer does not interact directly with DNA and instead may bind additional transcription factors and targets such as Sp1, RARα, insulin-like growth factor-binding protein 4, transforming growth factor α, bcl-2 and also interactions with AP-1.

1.7.1 Gene activation through ligand Dependent ER action

In vivo ERs bind to EREs in the promoters or regulatory regions of oestrogen responsive genes assembled into chromatin in the nuclear environment of the cell. As chromatin acts as a general repressor of RNA polymerase II mediated transcription it has an important functional consequence on ERα and other transcription factors mediating transcription of target genes (Wolffe et al, 1998). Complexes such as chromatin remodeling complexes and coactivators are recruited in order for transcription factors to overcome this chromatin-mediated repression of transcription.
Using energy stored as ATP chromatin remodeling complexes are able to mobilise or structurally alter nucleosomes, allowing greater access of transcriptional machinery to the DNA template (Kingston et al, 1999).

Coactivators play multiple roles in nuclear receptor mediated gene transcription as they function as bridging factors to recruit other cofactors to chromatin bound receptors (Mckenna et al, 1999); they acetylate nucleosomal histones and protein factors at the promoters of target genes which loosens chromatin structure and facilitates remodeling; they recruit RNA pol II and other components of the basal transcriptional machinery to the hormone regulated promoters (Leo et al, 2000).

The SRC family of proteins contains three structurally and functionally related members unified under the nomenclature of SRC1, SRC2 and SRC3 (Li et al, 1998) that function primarily but not exclusively as coactivators for nuclear receptors (Mckenna et al, 1999). The SRC proteins bind directly to liganded nuclear receptors via $\alpha$ helical motifs related to the sequence Leu–x–x–Leu–Leu (referred to as LXXLL motifs or NR boxes) (Heery et al, 1997). The NR boxes are located in the receptor interaction domain (RID) of the SRC proteins and interact with a hydrophobic groove on the surface of the receptor LBD (Dairmont et al, 1998). The SRC proteins, via distinct activation domains, contribute to gene transactivation. One such activation domain functions as a p300 and CBP interaction domain (PID). The SRC family may also possess a weak histone acetyltransferase (HAT) activity (Spencer, 1997) although such an activity has not been detected universally for the SRC family (Sheppard et al, 2001).
p300 and CBP are large highly related multifunctional coactivators sharing many structural and functional attributes, referred to collectively as p300/CBP, that act as coactivators for DNA binding transcription factors including nuclear receptors (Vo et al, 2001). The conserved motifs and functional domains in p300/CBP include a bromodomain, three cystine-histidine (C/H) rich regions (C/H1, C/H2, C/H3), a glutamine rich region, an intrinsic acetyltransferase activity and an SRC interaction domain (Fig. 1.17).

The bromodomain is a histone interacting module required for the direct interaction of p300/CBP with chromatin (Manning et al, 2001), and is found in many chromatin and transcriptional related factors (Winston et al, 1999). The C/H3 region is the site of interaction of a number of different transcription-related factors including RNA Pol II complexes, TFIIB (Nakajima et al, 1997). The glutamine rich region contains the SRC interaction domain (Kamei, 1996). The intrinsic acetyltransferase activity is capable of acetylating free or nucleosomal histones as well as SRC family members and some transcriptional activator proteins (Vo et al, 2001). All these p300/CBP domains are required to allow gene transactivation to occur via the oestrogen receptor (Kraus et al, 1999).
ER transcription complex assembly is thought to occur in a sequential fashion, leading to the activation of gene expression. The p160 coactivators, CBP, p300, pCAF and PBP are recruited in a specific order to the ER transcription complex after oestrogen stimulation in MCF-7 breast cancer cells (Shang et al, 2000). Also the ER transcription complex is thought to repeatedly cycle on and off target promoters in the presence of continuous stimulation by oestrogen (Shang et al, 2000). This cycling of the oestrogen transcription complex may be due to the covalent modification of coactivators (Fig. 1.18).

The following is thought to occur in the sequential build up of the ER transcription complex leading to target gene expression. Upon the binding of agonist to ERα, the liganded ERα binds to it’s ERE on DNA. Recruitment of the HAT containing p160-p300 complex and the CBP complex occurs following binding of the ERα homodimer to an ERE (Shang et al, 2000). Through subsequent histone acetylation the p300 complex modifies local chromatin structure, thus facilitating RNA Pol II recruitment. p300 only acts in the initial cycle of transcription initiation and not in the subsequent cycles thus showing that perhaps histone acetylation by p300 is long lived and not
required for the subsequent cycles. CBP replaces p300 in the transcription complex bringing in pCAF at the start of transcription when the C-terminal domain of Pol II is phosphorylated (Shang et al, 2000).

CBP acetylates p160 leading to the release of p160 with ERα. Following this CBP and pCAF disassemble and the cycle is repeated (Fig. 1.19) (Shang et al, 2000).
1.7.2 ER mediated gene expression via binding to Adjacent Transcription Factors

It had been thought that the only method of ERα mediated gene expression was through its direct binding to DNA via its response element the ERE. However, it is now known that ER can stimulate the expression of genes without directly binding to DNA. This is known because several genes not containing an ERE have been found to be regulated by oestradiol. ERα and ERβ have been shown to regulate these genes without binding directly to the promoter DNA. In these cases ERs bind indirectly to DNA via interactions with other transcription factors such as AP1, SP1, NFκB and GATA-1.

1.7.2.1 ER Interactions with AP1

One method of oestrogen regulation of genes that do not contain EREs is through interactions of ER with members of the fos, jun family of transcription factors that bind to AP-1 sites. Examples of genes that are activated by ER:AP1 interactions include IGF-1, ovalbumin and collagenase. It has been shown that the ERα DBD is not required for hormonal activation through AP1, but that AP1 proteins are required for activation to occur (Web et al, 1995). Interactions of ERα and c-Jun have been shown to occur in vitro and the amino acids 259-302 located in the hinge region of ERα are responsible for these interactions.

ERα and ERβ have been shown to differentially express/repress gene activation through their binding to AP1, following treatment by agonists and antagonists. The agonist oestradiol and antagonists Tamoxifen and ICI 164384 have been shown to activate gene transcription at AP1 sites through interactions with ERα, while oestradiol and DES inhibit transcription at these AP1 sites in the presence of ERβ. Tamoxifen and ICI 164384 act as transcriptional activators at AP1 sites in the presence of ERβ (Paech et al,
Molecular analysis has shown that the domains of ER responsible for AP1 interactions vary with the cell type and also the ligand involved (Weatherman and Scanlon, 2001).

It is thought that both AF dependent and independent pathways are involved in activation of gene expression through ER:AP1 interactions (Web et al, 1999). Oestrogen gene activation through ERα:AP1 interactions is thought to require AF-1 and AF-2 and their interaction with the p160 family of coactivators while tamoxifen is thought to activate ERα:AP1 mediated gene expression through AF-1 and also an AF independent pathway. ICI 164384 regulates gene expression through ERβ:AP1 interactions via an AF independent mechanism involving interactions with the corepressor N-CoR (Bjornstrom et al, 2002).

1.7.2.2 ER Interactions with SP1

Another method of ER target gene transcription through indirect binding with DNA is via interactions with Specificity protein 1 (SP1). SP1 was the first transcription factor identified in the early 1980s (Dylan et al, 1985). It is a member of the Sp/krupple like family (KLF), which consists of at least 25 components. The Sp1 consensus binding site is 5’-(G/T)GGGCGG(G/A)(G/A)(G/T)-3’. The ER interaction with SP1 involves a DNA dependent mechanism (Fig. 1.20) in which an ERE half site and GC rich site are required for transactivation and also a DNA independent mechanism which requires only a GC rich region.

The transactivation of genes by ER following interaction with the ER half site and SP1 bound to its GC rich region binding site has been shown for gene promoters of creatine
kinase B (Wu Peng et al, 1992), Cathepsin D (Krishnan et al, 1995), hsp 27 (Porter et al, 1996), progesterone receptor genes (Petx et al 2000), TGFα (Vyhlidal et al, 2000) and initially in the c-myc promoter (Dubik et al, 1987). It has been established that these promoters do not to contain a consensus full site ERE but instead a ERE half site and also a GC rich region. The ERα generally requires its DBD to allow transactivation to occur in these instances.

It is also thought that ER:SP1 interactions leading to gene activation may occur without ER binding directly to DNA. It was found that mutation of the ERE half site in the hsp27 gene promoter still allowed gene expression, which led to the thought that in some cases GC rich regions may be all that is required for gene activation despite the presence of an ERE half site. ERα and ERβ have been shown to interact with the C-terminal DBD of SP1 (Porter et al, 1997), while SP1 has also been shown to interact with multiple regions on ERα (Saville et al, 2000). In this mode of ERα transactivation through non-DNA contacts it is thought that ERα interacts with the GC rich bound SP1 and thus mediates transactivation of genes containing GC rich regions (Fig. 1.21).
ER:SP1 gene transactivation has been shown both in normal breast tissue as well as in breast cancer cells. Non breast cancer cell line genes regulated by ERα:SP1 interactions include the small conductance Ca\(^+\) - activated potassium channel rat SK3 (rSK3) (Jackobson et al, 2003), receptor for advanced glycalation end products (RAGE) (Tanaka et al, 2000), low density lipoprotein receptor (LDLR) (Li et al, 2001) and epidermal growth factor receptor (EGFR) (Salvatori et al, 2000). The genes regulated by ERα:SP1 genes in breast cancer cell lines include DNA polymerase α (Samudio et al, 2001), adenosine deaminase (ADA), thymidylate synthase (Xie et al, 2000), cyclin D1 (Castro-Rivera et al, 2001), retinoic acid receptor α (RARα) (Sun et al, 1998), VEGF (Stoner et al, 2004), bcl2 (Dong et al, 1999) and pS2 (Sun et al, 2005).
1.7.3 ER activation by Kinases

ER activation can also occur in the absence of ligand by a variety of different factors and pathways. Examples include activation of ER by activators of the PKA and PKC pathways (Cho et al, 1993), cyclins A (Neuman et al, 1997) and D (Zwijsen et al, 1997), peptide growth factors (Pietras et al, 1995), 8-bromo-cyclic adenosine monophosphate (cAMP) (Ince et al, 1994), caveolin (Schlegel et al, 1999) and insulin (Patrone et al, 1996). In addition, growth factors, which are involved in cell proliferation and differentiation, activate ER: examples are EGF, IGI-1 and herregulin-2 (Pietras et al, 1995) and transforming growth factor (TGFα) (Ignar-Trowbridge et al, 1996).

EGF is known to act as a promoter of cell growth, to induce oestrogen responsive genes (Nelson et al, 1991) and it has been shown in the uteri of ERKO mice that EGF is unable to induce DNA synthesis and transcription, which suggests that EGF may act through ER in the absence of oestrogen. IGF-1 activates the tyrosine receptor IGFR-1, regulates uterine cell proliferation in vivo and, like EGF, is dependent on ERα for its regulation of cell proliferation (Klotz et al, 2002). Activation of ERα by IGF-1 and EGF is through MAPK-dependent phosphorylation at Ser 118. The PI3-K-dependent phosphorylation of ERα at Ser 167, which enhances ERα-dependent transcriptional activity and cell proliferation, is also activated by IGF-1.
1.8 Oestrogen Receptor and Breast Cancer

In industrialised countries breast cancer is the most common malignancy in women, and it is the most common important cause of cancer-associated morbidity and mortality (Giacinti et al, 2006). Breast cancer was known to be associated with oestrogen more than 100 years ago, when Beatson used ovariectomy to prevent tumor recurrence and actually induce regression of the primary tumor (Sommer et al, 2001). Sixty five percent of primary breast cancers are ER positive (Giacinti et al, 2006).

In normal mammary glands oestrogen is involved in the growth and differentiation of epithelial cells. Thus the mitogenic effects of oestrogens on breast epithelial cells are due to, at least in part, the increased expression of genes involved in the regulation of the cell cycle (Vendrell et al, 2004). Although the mitogenic effect of oestrogen has been known for a while, little is known about the actual mechanism of oestrogen mediated proliferation.

Some studies on oestrogen regulated gene expression in breast cancer cells have identified certain target genes that may be involved in cell proliferation and, in particular, breast cell proliferation. NOV, which encodes for a negative regulator of cell proliferation, is down regulated by oestrogen in breast cancer cells (Vendrell et al, 2004). WNT2 expression, which is associated with abnormal proliferation in human breast tissue (Huguet et al, 1994), has been demonstrated to increase following oestrogen exposure (Vendrell et al, 2004). Oestrogen has also been shown to induce the down regulation of the TACC gene family members, TACC1, EFNA1 and ZNF217 (Vendrell et al, 2004). Members of this family of proteins are defined by the presence of the so-called TACC domain, a predicted coiled-coiled region in their carboxy
terminus (Lappin et al, 2002). The TACC gene family has been reported to contribute to the development of cancer (Raff, 2002), and TACC1 mRNA down regulations have been observed in breast tumors (Conte, 2002).

1.8.1 ER negative and Tamoxifen Resistant Breast Cancer

As sixty five percent of primary breast cancers are ER positive, therapy with antioestrogens represents an attractive option, as reducing oestrogen receptor levels or altering oestrogen receptor activity can induce cancer regression. However, up to one third of breast cancers lack ER\(_\alpha\) at the time of diagnosis and a fraction of the breast cancers that are initially ER\(_\alpha\) positive, lose ER expression during progression of the tumor (Yang et al, 2001). This absence of ER is correlated with a more malignant disease, as these types of cancer deprive us of an important possibility of patient care by endocrine therapy.

1.8.1.1 Methylation of the ER gene Promoter

A significant fraction of breast cancers lack ER as a result of aberrant methylation of GpC islands, cytosine–guanine rich areas that are located in the 5’ regulatory regions of the ER\(_\alpha\) gene (Lapidus et al, 1998). Methylation is the covalent addition of a methyl group to the 5’ position of cytosine, predominantly within the CpG dinucleotide (Robertson, 1998). As the ER\(_\alpha\) promoter is hypermethylated, its expression is silenced.

This CpG island methylation is not random and genes that regulate cell growth are more likely to be susceptible to hypermethylation, so cells that lack them have a growth advantage (Costello et al, 2000). Methylation of the ER\(_\alpha\) promoter is an example of growth advantage because breast cancer is normally a hormone dependent tumor, which
means oestrogens can regulate their growth through binding to ERs. This means that breast cancer incidence and proliferation in ER positive breast cancers would be increased on exposure to oestrogen; however, antioestrogenic drugs can control this. The growth of ER negative breast cancers is not under the control of oestrogen, so they cannot be stopped by endocrine therapy. These tumors are more aggressive resulting in poorer prognosis (Giacinti et al, 2006).

1.8.1.2 ER mutations

ER missense mutations have been reported to be present in 1% of primary breast tumors (Roodi et al, 1995) and this value may be even higher for metastatic lesions (Zhang et al, 1997). These mutations in the ER gene have been shown to result in changes in the receptor activation by conferring constitutive activity, thought to be due to conformational changes (Zhang et al, 1997), and also interference with cofactor binding (Wies et al, 1996).

One such ER mutation, Tyr 537 Asn, affects a major site of ligand-induced phosphorylation in the LBD, resulting in a constitutively active receptor, which may bind to coactivators even in the absence of ligand (Wies et al, 1996). Another ER mutant Lys 303 Arg displays enhanced binding to the coactivator SRC-2 and results in the stimulation of proliferation even in the presence of low levels of oestrogen, suggesting a possible role in early carcinogenesis (Fuqua, 2000).
1.9 Methods used for Identifying Transcription Factor Binding Sites

The methods that are used to allow the isolation of transcription factor binding sequences are:

1) Chromatin immunoprecipitation which involves the determination of an endogenous proteins DNA binding location and can be used with direct cloning

2) DamID which involves the generation of a fusion protein and is based on the detection of protein binding using targeted DNA adenine methylation.

Because both these techniques are based on different principles they can be used together to increase the likelihood of gaining an insight into specific proteins DNA interactions.

1.9.1 ChIP Assay

ChIP is a powerful tool for studying protein DNA interactions. Using formaldehyde intact cells are fixed, which cross-links and preserves protein/DNA interactions. The gDNA is then sheared into small uniform fragments using sonication or enzymatic digestion. Using an antibody directed against the specific DNA binding protein of interest, the specific protein/DNA interactions are immunoprecipitated. The protein-G coated magnetic beads facilitate immunoprecipitation. The beads have a high capacity binding for IgG and also low non-specific binding. Following immunoprecipitation of the protein/DNA complexes cross-linking is reversed, and the proteins are removed by Proteinase K treatment and the DNA is recovered (Fig. 1.22). The DNA can then be cloned into a vector or PCR analysis can be carried out.

ChIP assay is in theory an ideal method for the isolation of ERα interacting sequences as protein/DNA interactions are fixed while in an endogenous chromosomal context.
Thus ChIP results reflect the influence that chromatin has on transcription factor binding and also reflects the effects of cellular regulatory proteins.

1.9.2 DamID

DamID is second technique that will be used for the isolation of ERα genomic interacting sequences. DamID was originally developed in Drosophila (Van Steensel et al, 2000) and is based on targeted adenine methylation. DamID relies on the low level in vivo expression of a fusion protein consisting of *Escherichia coli* DNA adenine methyltransferase (Dam) and the transcription factor of interest, in this case ERα. Dam methylates adenines in the sequence GATC (Fig. 1.23), which is a modification that is not present in most eukaryotes (Vogel et al, 2007). Upon in vivo expression of
the fusion protein in cultured cells, Dam is targeted to its fusion partners’ DNA binding sites, resulting in the methylation of nearby GATC sequences.

To detect and isolate the sequences that have been methylated by Dam, the gDNA is first isolated and then treated with the restriction enzyme Dpn I, which selectively cuts methylated GATC sequences. This generates a pool of blunt stranded DNA, to which a double stranded adaptor oligonucleotide is ligated. DNA containing unmethylated GATC sequences are destroyed following digestion using the restriction enzyme Dpn II, that specifically cuts unmethylated GATC sequences. A PCR primer, identical to the 3’end of the double stranded adaptor, is used to amplify adaptor-ligated sequences (Vogel et al, 2007). These PCR amplified sequences can then be cloned into a vector to allow sequencing and analysis of their functional activity. In this way DamID can be used for isolation of ERα DNA binding sequences and facilitate the generation of a reporter library of genomic ER interacting cis elements.

![Fig. 1.23 Interaction of ERα:Dam Fusion Protein with DNA and methylation of GATC sequences](image)

There are advantages and disadvantages to both ChIP and DamID and if both can be used in combination these disadvantages can be minimized. The advantages of ChIP
are that it can be performed on the endogenous protein whereas DamID relies on an exogenously prepared fusion protein. As ChIP is used on the endogenous protein it may allow the determination of a truer picture of what exactly is going on in the cell and it also means that the protein of interest does not have to be altered in any way. DamID is based on the fusion of the protein of interest with DNA adenine methyltransferase.

This fusion of Dam to the protein of interest could reduce the protein’s ability to bind DNA and/or could cause the protein to lose its genomic binding specificity. Unlike ChIP, DamID is not dependent on the availability of high quality antibodies against the protein of choice, which can be a limitation in some ChIP assays. To carry out a ChIP assay, the quantity of cells typically required is 10 to 100 fold greater than DamID, which can be performed on ~10^6 cells. ChIP is dependent on the use of cross-linking reagents which can result in cross-linking artifacts. DamID is not dependent on these cross-linking reagents as it involves the identification of DNA protein interactions using adenine methylation and subsequent PCR specific amplification, thus ruling out the possibility of cross-linking artifacts.
1.10 Aims of the Project

Many important processes in the cell rely on protein DNA interactions for their initiation and subsequent regulation including transcription, DNA repair and replication. To gain a greater insight into the steps involved in these cell processes knowledge of where the regulatory proteins bind can be of great benefit. One is able to gain a greater knowledge of their separate and combined functions by ascertaining where these individual proteins bind.

Oestrogen as well as various other ligands, mediates its effects through the activation of the ER. The activated receptor is able to regulate gene expression through binding DNA. Trans factors involved in ER\(\alpha\) mediated transcription include the direct or indirect binding of the receptor to DNA, coactivators or corepressors binding the receptor and also receptor post translational modification coding such as methylation and acetylation. Cis regulatory elements that can be involved in ER\(\alpha\) mediated transcription include ERE full sites, ERE half sites and a full spectrum of other transcription factor binding sites (i.e. Fox A1, AP-1, SF-1, MAF, PAX2, PAX3 and CDX) in cis with EREs.

Most of our knowledge of oestrogen regulated gene expression is based on a few model genes such as Xenopus Vitellogenin A2 enhancer and the human pS2 (TFF1) genes. Both of these genes are upregulated by oestrogen and have relative weak partial agonist response to 4-hydroxytamoxifen. However, gene expression profiling has revealed many genes that depart from this classical model of ER mediated regulation; some are negatively regulated by oestradiol while 4-hydroxytamoxifen positively regulates others.
Not a lot is known about exactly how the ER interacting DNA sequence regulates its binding and its subsequent modification of gene expression and, also, if ER binds to DNA sequences that do not contain an ERE. Therefore, the aim of this project is to try and answer some of the following questions:

1. Which method of isolating transcription factor binding, ChIP or DamID, is most efficient for isolation of ERα interacting sites?
2. Which type of ERE is most commonly isolated, ERE full site, half site or full and half site?
3. What adjacent transcription factor binding motifs are found co-associated with the isolated EREs?
4. Do the levels of ERα and ERβ mediated transcription vary depending on ERE type?
5. Does treatment with agonists or partial agonist/antagonists cause a variation in ERα mediated transcription following binding to isolated EREs?
6. What is the difference in gene expression levels following ERE mediated binding of ERα and ERβ?
Materials and Methods

2.1 ChiP Mediated Isolation of ERα Interacting Sequences

2.1.1 ChiP Assay

Culture and Treatment

MCF-7 cells were grown to a confluency of 70-80% in a 15cm plate with DMEM media containing 10% steroid stripped foetal bovine serum. The cells were treated with 0.5 μl ethanol, Oestradiol $10^{-6}$ M dissolved in ethanol and Tamoxifen $10^{-4}$ M dissolved in ethanol for 24 hours.

Fixation

Following 24 hours of ligand treatment the media was aspirated off the wells. 20ml of fixation solution was added to each plate and incubated on a shaking platform for 10 minutes at room temperature. Following incubation the fixation solution was removed and a wash was performed by adding 10 ml of cold PBS to each plate. The plate was rocked for 5 seconds and the PBS poured off. 10 ml of Glycine-Stop Fix solution was added to each plate to stop the fixing reaction. An incubation was carried out for 5 minutes at room temperature, followed by removal of the Glycine-Stop Fix solution and a wash with PBS. 2ml of cell scraping solution was then added to each well and the cells were scraped using a rubber scraper. The solution containing the cells was stored in a tube on ice. The cells were then pelleted by centrifugation at 14,000 rpm for 15 minutes @ 4°C, and the resulting supernatant was removed.
Preparation of Cell Extracts for Immunoprecipitation

Cells were resuspended in 1ml of ice-cold lysis buffer containing 5 μl Protease inhibitor cocktail (PIC) and 5 μl PMSF. These were then incubated on ice for 30 minutes. Following incubation the cells were dounced using an ice-cold dounce homogeniser, which aids in nuclei release. The cells were transferred to centrifuge tubes and centrifuged at 5,000 rpm for 10 minutes at 4°C. The supernatant was discarded and the pellet was resuspended in 350μl of digestion buffer. This was incubated at 37°C for 5 minutes. 17 μl of the enzymatic shearing cocktail, diluted 1:100 in 50% glycerol and dH2O, was added. This was incubated for 1 hour at 37°C. The shearing reaction was stopped by adding 7μl of ice-cold 0.5 M EDTA and incubated on ice for 10 minutes. The samples were centrifuged at 15,000 rpm for 10 minutes at 4°C. The supernatant, which contains the sheared chromatin, was removed and put into a new tube ready for immunoprecipitation.

Immunoprecipitation and Proteinase K Treatment

The immunoprecipitation reaction was performed in one tube by adding 50 μl of sheared chromatin sample, 25 μl of Protein G Magnetic Beads, 10 μl ChIP 1 Buffer, 1 μl PIC, 9 μl H2O and 5 μl of ERα (HCD20):sc 534. The reactions were incubated on an end-to-end rotator overnight at 4°C. Using a magnet, the beads were clumped allowing disposal of the supernatant. The beads were washed once with 800 μl of ChIP Buffer 1 and twice with 800 μl of ChIP Buffer 2. The beads were then resuspended in 50 μl of Resuspension buffer and incubated for 15 minutes at room temperature in an end on end incubator. 50 μl of the Reverse Cross-Linking Buffer was added and the beads were immediately pelleted at the side of the tube using a magnet. The supernatant was
transferred to a fresh tube. Next the tubes were incubated at 65°C for two and a half hours. Following incubation, 2 μl of Proteinase K was added to each tube. These were then incubated for 1 hour at 37°C. 2μl of Proteinase K stop solution was added followed by centrifugation. The isolated gDNA was then digested with 1 μl each of the restriction enzymes Bgl II and Nhe I.

2.1.2 Vector Digestion

The reporter vector pERE-EIb-LUC (Fig. 2.1) was used as the vector for cloning the isolated DNA sequences into. This Luciferase reporter vector contains an ERE derived from the Xenopus vitellogenin A2 enhancer. This ERE is first removed from the pERE-EIb-LUC reporter by digestion with the restriction enzymes Bgl II and Nhe I. This generates an open Luciferase expression vector into which the ChIP isolated DNA sequences could be ligated and subsequently functionally tested following transfection and luciferase assay.

![Diagram of pERE-EIb-LUC Luciferase Reporter Vector](image-url)
The following digestion reactions were set up: 5 μl pERE-EIb-LUC, 1 μl NEB 2 Buffer, 1 μl Bgl II, 1 μl Nhe I and 1 μl H2O. The digestion reaction was carried out at 37°C for 1 hour.

2.1.3 Digestion Reaction Analysis on Agarose Gel

A 2% agarose gel containing 0.5 μg/ml ethidium bromide was prepared by weighing out 2g of agarose and adding 100 ml of TAE buffer followed by heat treatment. The gel was then placed in an electrophoresis chamber and wells created using a comb. Once set, the gel was covered with TAE buffer. 3μl of molecular weight marker and the 10 μl digestion products were pipetted into the wells. The gel was then run at 100 volts for 35 mins. The gel was examined under a UV lumimeter.

2.1.4 Gel Clean up

The QIAquick system utilises the selective binding properties of a uniquely designed silica membrane. The membrane is uniquely adapted to purify DNA from both aqueous solutions and agarose gels, and up to 10 μg DNA can bind to each QIAquick column. DNA adsorbs to the silica membrane in the presence of high concentrations of salt while contaminants pass through the column.

The binding buffers in QIAquick Spin Kits provide the correct salt concentration and pH for adsorption of DNA to the QIAquick membrane. The adsorption of nucleic acids to silica surfaces occurs only in the presence of a high concentration of chaotropic salts, which modify the structure of salt. Adsorption of DNA to silica also depends on pH. Adsorption is typically 95% if the pH is < 7.5 and is reduced drastically at higher pH. The binding buffer PBI and solubilisation buffer QG contain a pH indicator, which
appears as yellow below a pH of 7.5, allowing easy determination of the optimal pH for DNA binding.

During the DNA adsorption step, unwanted primers and impurities, such as salts, enzymes, unincorporated nucleotides, agarose and ethidium bromide do not bind to the silica membrane but flow through the column. Salts are quantitatively washed away by the ethanol containing Buffer PE.

Elution efficiency is strongly dependent on the salt concentration and pH of the elution buffer. Elution is most efficient under basic conditions and low salt concentrations. The DNA is eluted with the supplied Buffer EB. The maximum elution efficiency is achieved between pH of 7.0 and 8.5.

**Procedure**

The digested pERE-EIb-Luc DNA fragment was excised from the agarose gel under the UV lumimeter with a clean sharp scalpel. A QIAquick Gel Extraction was then carried out. Three volumes of Buffer QG were added to one volume of gel. This was incubated at 50 °C until the gel slice had completely dissolved. After the gel slice had dissolved completely, 1 gel volume of isopropanol was added to the samples followed by mixing. This step increases the yield of DNA fragments < 500 bp and > 4 kb. The solutions were then applied to the QIAquick column and centrifuged for 1 minute. The flow-through was discarded and the QIAquick column was placed back in the same collection tube. 0.5 ml of Buffer QG was added to QIAquick column and centrifuged for 1 minute and the flow through was again discarded. For the wash step, 0.75 ml of Buffer PE was added to the QIAquick column and centrifuged for 1 minute, followed
by discarding the flow-through. The QIAquick column was centrifuged for an additional 1 minute at max speed and then placed into a clean 1.5 ml tube. The DNA was then eluted by adding 50 μl of Buffer EB to the center of the QIAquick membrane followed by centrifugation for 1 min.

2.1.5 Ligation of ChIP Isolated DNA Sequences and Linearised Vector

The following ligation reactions were set up: 1 μl Bgl II and Nhe I digested pERE-Elb-LUC, 5 μl ChIP isolated DNA, 1.5 μl 10mM ATP, 1.5 μl 10 x Fast-link Buffer, 1 μl Fast-link DNA Ligase and 5 μl H2O. A 15 μl negative control reaction was also set up as follows: 1.5 μl 10 x Fast-link ligation Buffer, 1.5 μl 10mM ATP, 1 μl Apa I BglII digested pCR3.1 hERα, 10μl H2O, 1μl Fast-link Ligase. Both reactions were incubated for 45 minutes at room temperature. The reactions were then incubated at 70°C for 15 minutes to inactivate the Fast-Link DNA ligase.

2.1.6 Transformation of Ligation Reaction

A vial of One Shot TOP10 chemically competent E.coli was thawed for each transformation. To each vial of One Shot cells, 5 μl of the ligation reactions were added followed by gentle mixing. A positive transformation was also set up by adding 1 μl the pUC19 control into a separate vial of One Shot cells. The vials were then incubated on ice for 30 minutes. A heat shock process was carried out on the cells for 30 seconds at 42 °C. After this the vials were placed on ice for 2 minutes. 250 μl of S.O.C Medium, prewarmed to 37 °C, was added to each vial. They were then shaken horizontally at 37°C for 1 hour at 225 rpm in a shaking incubator.
Following incubation 70μl from each transformation was spread on a pre-warmed ampicillin plates and incubated overnight at 37 °C. The remaining transformation mix was stored at 4°C so additional cells could be plated out the next day, if needed.

2.1.7 Growth of Resulting Colonies and Plasmid Isolation
All resulting bacterial colonies on the ampicillin plate containing competent cells, transformed with vector ligated with ChIP isolated sequences, were picked and grown overnight. First 3ml of LB and 15 μl of Ampicillin (10 mg/ml) were added to test tubes. Each individual colony was picked using a pipette tip and placed in the test tube. All tubes were then incubated overnight in a 37 °C shaker.

2.1.8 Plasmid Isolation
The ChargeSwitch-Pro Plasmid Miniprep Kit contains all components for the rapid and efficient isolation of highly pure plasmid DNA from E. coli cells. The purification columns contain a novel ChargeSwitch membrane that is positively charged at low pH and neutral at pH 8.5, to bind and elute plasmid DNA without the use of harsh reagents.

The kit allows preparation of cell lysates with a modified alkaline lysis procedure and subsequent purification of plasmid DNA using a simple centrifugation protocol. In low pH conditions, the ChargeSwitch membrane binds the negatively charged nucleic acid backbone. Proteins and other contaminants are not bound and are washed away using wash buffers.

Elution of the bound DNA is carried out by neutralizing the charge of the membrane from raising the pH to 8.5 using a low-salt buffer. The purified plasmid DNA is then
ready for use in mammalian transfections, automated fluorescent DNA sequencing and analysis using restriction digestion.

Procedure

The plasmids were isolated from the overnight bacterial suspensions using the ChargeSwitch-Pro Plasmid MiniPrep Kit. 1.5 ml of the overnight bacterial culture was harvested by centrifugation. The cell pellet was resuspended in 250 μl of Resuspension Buffer premixed with RNase A. 250 μl of lysis buffer was then added and the solution was mixed by inverting the capped tube until the solution became viscous. The solutions were incubated at room temperature for 2 – 5 mins. 250 μl of Precipitation Buffer was added, with immediate mixing by inversion until a white precipitate was formed.

The solutions were then centrifuged at 1500 rpm for 10 minutes to pellet the debris. The supernatant was transferred onto the ChargeSwitch-Pro Miniprep Column, which had been inserted in a clean collection tube, and the column centrifuged at 15000 rpm for 30 – 60 seconds. The flow-through was discarded and the column re-inserted into the same collection tube. 750 μl of Wash Buffer 1 was added to the column, which was centrifuged at 15000 rpm for 30 – 60 seconds. The flow-through was discarded and 250 μl of Wash Buffer 2 was added to the column, which was centrifuged at 15000 rpm for 30 – 60 seconds. The flow-through was discarded, the column was inserted into an Elution Tube and 25 μl – 100 μl of elution buffer was added onto the column. This was centrifuged at 15000 rpm for 30 – 60 seconds.
2.1.9 Plasmid Analysis

The isolated plasmids were screened for the presence of inserts using restriction digestion analysis. The following digestion was carried out: 7.5 μl isolated Plasmid DNA, 1 μl NEB 2 Buffer, 0.5 μl Kpn I and 0.5 μl Hind III. Restriction digestion was carried out for 1 hour at 37°C.

All digested plasmids were run on a 2% agarose gel as outlined in step 2.1.3 above. This allowed the visual detection of the presence of any plasmids containing ChIP isolated DNA inserts.

2.1.10 Transfection and Luciferase Assays of Plasmids Containing DNA Inserts

Lipofectamine is a specially designed cationic lipid that facilitates the delivery of DNA into cells. The basis structure of cationic lipids consists of a positively charged head group and one or two hydrocarbon chains. The charged head group governs the interaction between the lipid and the phosphate backbone of the nucleic acid, and facilitates DNA condensation. The positive surface charge of the liposomes also mediates the interaction of the nucleic acid and the cell membrane, allowing the fusion of the liposome/nucleic acid (transfection complex) with the negatively charged cell membrane. Sufficient cationic liposome has to be mixed with the negatively charged DNA to ensure a net positive charge, facilitating this interaction with the membrane.

The transfection complex is then thought to enter the cell through endocytosis. Once inside, the complex must escape the endosomal pathway, diffuse through the cytoplasm and enter the nucleus. Plasmid DNA must be delivered to the nucleus for transcription to occur and this is thought to occur via the nuclear pore complex. Only a fraction of
the DNA delivered to the cells makes it to the nucleus for transcription, with eventual export of the mRNA to the cytoplasm for protein production.

Following DNA transfection, detection of transient expression of the gene lasts for about 1 to 7 days. Within a few days most of the foreign DNA is degraded by nucleases or diluted by cell division and after a week its presence is normally no longer detected.

Firefly luciferase is a 61kDa monomeric protein that does not require posttranslational processing for enzymatic activity. Thus it serves as a genetic reporter immediately upon translation. Photon emission is achieved through oxidation of beetle luciferin in a reaction that requires ATP, Mg$^{2+}$, and O$_2$. Luciferase agents incorporate coenzyme A (CoA), which provides favorable reaction kinetics.

**Procedure**

Transfection of HeLa cells and subsequent luciferase assay can be used to test the functionality of the plasmids containing ChIP isolated ER$\alpha$ sequences in conjunction with the ER$\alpha$ WT expression vector or the ER$\alpha$ mutant expression vector C202H / C205 H. This mutant possesses two mutations to two of the Zinc co-ordinating cystines in the first Zinc finger, more specifically the P-Box. This form of the receptor is transcriptionally inactive when assayed on an ERE-containing luciferase reporter, as it cannot bind to an ERE.

The HeLa cells were plated at a density of 50,000 cells per well the day before the transfection. A master mix of Lipofectamine and DMEM containing no serum (1µl and
25 μl per well respectively) according to the amount of wells was made up initially. A master mix of reporter vectors and receptor expression vectors was made up by adding 250 ng of each individual plasmid containing isolated sequences and 2 ng of WT ERα or C202H / C205 H expression vectors. 200 μl of Lipofectamine/DMEM mix was added to the tubes containing the plasmids. These were incubated for 45 minutes at room temperature.

The medium was removed from each well containing HeLa cells and 200 μl of serum free DMEM was added, followed by 50 μl of Vector / Lipofectamine / DMEM without serum mix. The cells were incubated in a CO2 incubator. After 3 hours the DMEM was removed from the individual wells using a sterile automated pipette technique and 500 μl of DMEM containing serum was added. The cells were incubated overnight in a CO2 incubator.

After 24 hours 0.5 μl of ethanol or the appropriate hormones, oestradiol x 10⁻⁶ M or Tamoxifen 10⁻⁴ M were added to each of the wells. This was then incubated overnight for 24 hours. The cells were now ready for luciferase assay.

A passive lysis technique was used to measure luciferase gene expression. The growth medium was removed from the wells using a sterile automated pipette technique. Into each well 50 μl of 1 x passive lysis buffer (PLB) was added. This promotes the rapid lysis of cultured mammalian cells without the need for cell scraping. The plates were placed onto a rocking orbital shaker for 15 minutes to ensure complete coverage of the cell monolayer with 1 x PLB. The cell lysates were added to new tubes. Luciferase
Assay reagent was added to each tube. The luciferase assays were then carried out on the plate reading luminometer.
2.2 Generation of ERα:Dam Fusion PCR Product

2.2.1 Generation of ERα PCR Product

PCR amplification was performed on a 50μl reaction containing 0.5μl ERα forward primer (Fig. 2.2 A), 0.5μl ERα reverse primer (Fig. 2.2B), 2μl pCR3.1 hERα, 41.2μl of H2O, 5μl of 10 x Accuprime Rx Mix and 0.8μl Accuprime Pfix DNA Polymerase. The PCR cycle was 5 mins at 94°C, 1 min at 94 °C, 1 min at 58 °C, 1 min at 64 °C, a repeat of 25 cycles of the 94°C to 64°C stages followed by 5 mins at 72 °C.

ERα Forward Primer

GAC ATG GTG GAG ATC TTC GAC ATG CTG CTG

ERα Reverse Primer

GCA GAG GGT TTC CCT GCC ACA GTC AAG AAA AAT CGC GCT TTT TTG AAG

Fig. 2.2 A Forward and Reverse Primers used In the Generation of the ERα PCR Product

2.2.2 PCR Generation of Dam PCR Product

E.coli gDNA isolated by centrifuging 1.5 ml of E.coli bacterial culture. The pellet was resuspended in 500 μl TE. A 250 μl phenol-chloroform extraction was followed by a 250 μl extraction using chloroform. PCR amplification was performed on a 50μl reaction containing 0.5μl Dam forward primer (Fig. 2.2 B), 0.5μl Dam reverse primer (Fig. 2.2 B), 2μl E.coli gDNA, 41.2μl of H2O, 5μl of 10 x Accuprime Rx Mix and 0.8μl Accuprime Pfix DNA Polymerase. The PCR cycle was 5 mins at 94°C, 1 min at 94 °C,
1 min at 58 °C, 1 min at 64 °C, a repeat of 25 cycles of the 94°C to 64°C stages followed by 5 mins at 72 °C.

Dam Forward Primer

```
ER\(\alpha\) 3'End

* * * * * * * * * * * * * * * * * * * Dam 5'End
CTT CAA AAA AGC GCG ATT TTT CTT GAC TGT GGC AGG GAA ACC CTC TGC
```

Dam Reverse Primer

```
Apa I Restriction Site Stop Dam 3'End

* * * * * * * * * * * * * * * * * * * *
ATC GCA GGG CCC GAA TTA TTT TTT CGC GGG TGA AAC GAC
```

**Fig. 2.2 B Forward and Reverse Primers used in the Generation of the ER\(\alpha\) PCR Product**

### 2.2.3 ER\(\alpha\) and Dam PCR Product Gel Separation

A 2 % agarose gel was made by weighing out 2g of Agarose and adding 100 ml of TAE buffer followed by heat treatment. The gel was then placed in an electrophoresis chamber and wells created using a comb. Once set, the gel was covered with TAE buffer. 3\(\mu\)l of molecular weight marker and 15 \(\mu\)l ER\(\alpha\) and Dam PCR products were pipetted into the wells. The gel was then run at 100 volts for 35 mins. The gel was examined under a UV lumimeter.

### 2.2.4 Gel Excision of PCR Products and their Subsequent Clean Up

The ER\(\alpha\) and Dam DNA fragments were excised from the agarose gel under the UV lumimeter with a clean sharp scalpel. A QIAquick Gel Extraction was then carried as per 2.1.4.
2.2.5 Preparation of the ERα:Dam Fusion Protein PCR Product

PCR amplification was performed on a 50μl reaction containing 0.5μl ERα forward primer, 0.5μl Dam reverse primer, 1μl Gel purified ERα PCR product, 1μl Gel purified Dam PCR product, 41.2μl of H2O, 5μl of 10 x Accuprime Rx Mix and 0.8μl Accuprime Pfix DNA Polymerase. The PCR cycle was 5 mins at 94°C, 1 min at 94°C, 1 min at 58°C, 1 min at 64°C, a repeat of 25 cycles of the 94°C to 64°C stages followed by 5 mins at 72°C.

2.2.6 ERα:Dam Product Clean up, Enzyme Digestion, Separation and Gel Extraction

The ERα:Dam PCR product was run on an agarose gel as per step 2.1.3 above. The DNA fragment was then excised from the agarose gel, and a Gel clean up performed using QIAquick Gel purification kit as per step 2.1.4 above.

Following Gel purification, the PCR product was digested with the restriction enzymes Bgl II and Apa I. The following restriction digest was set up: PCR 25 μl Gel purified ERα:Dam PCR product, 3 μl 10 x NEB Buffer 2, 1 μl Bgl II, 1 μl Apa I. The reaction was carried out at 37°C for 1 hour.

Use QIAquick PCR Purification Kit as explained above to remove Accuprime Pfix DNA Polymerase and dNTPs from the reaction.

Following enzyme digestion, 15μl of the reaction was run on a 2% agarose gel as outlined in step 2.1.3 above. The DNA fragment was then excised from the agarose gel.
and gel purification was performed using a QIAquick Gel Purification Kit as outlined in step 2.1.4 above.
2.3 Generation of the ERα:Dam Expression Vector

2.3.1 Restriction Digestion of pCR 3.1 hERα

To facilitate ligation of the newly generated ERα:Dam fusion PCR product the pCR3.1 hERα vector was digested with the restriction enzymes Bgl II and Apa I. The following restriction digestion was set up: 1 µl pCR3.1 hERα (1 µg/µl), 1 µl 10 x NEB 2 buffer, 1 µl Bgl II, 1 µl Apa I and 6 µl H2O. This restriction digestion reaction was carried out for 2 hours at 37°C.

The pCR3.1 hERα digestion reaction was then run on a 2 % agarose gel at 100 volts as outlined in step 2.1.3 above. The Bgl II, Apa I digested pCR3.1 hERα DNA fragment was isolated from the gel and, subsequently, the gel purified using QIAquick gel purification kit as outlined in step 2.1.4 above.

2.3.2 Ligation of ApaI, Bgl II digested ERα:Dam and pCR3.1 hERα

The fast-link DNA ligation kits provide reagents optimized for the construction of recombinant vectors, or the ligation of ds DNA to ds DNA in a short time. Uses of the kit include cloning of DNA into prokaryotic or eukaryotic vectors and bacteriophage lambda vectors.

Procedure:

To allow ligation of the ERα:Dam fusion PCR product to the digested pCR3.1 hERα:Dam vector a ligation reaction was carried out using fast-link ligase. The following 15 µl ligation reaction was set up: 1.5 µl 10 x Fast-link ligation Buffer, 1.5 µl
10mM ATP, 1 µl Apa I BglII digested pCR3.1 hERα, 3µl Apa I Bgl II digested ERα:Dam PCR product, 7µl H₂O, 1µl Fast-link Ligase.

A 15 µl negative control reaction was also set up as follows: 1.5 µl 10 x Fast-link ligation Buffer, 1.5 µl 10mM ATP, 1 µl Apa I BglII digested pCR3.1 hERα, 10µl H₂O, 1µl Fast-Link Ligase.

Both reactions were incubated for 30 mins at room temperature. The reactions were then incubated at 70°C for 15 minutes to inactivate the Fast-Link DNA ligase.

2.3.4 Transformation of E.coli with Ligation Reactions

Transformation was carried out as per section 2.1.6. Following incubation 70µl from each transformation was spread on pre-warmed ampicillin plates and incubated overnight at 37°C. The remaining transformation mix was stored at 4°C so additional cells could be plated out the next day, if needed.

2.3.5 Growth of Resulting Colonies and Plasmid Isolation

All resulting bacterial colonies on the ampicillin plate containing competent cells transformed with pCR3.1 ERα:Dam Vector were picked and grown overnight. First 3ml of LB and 15 µ of Ampicillin were added to test tubes. Each individual Colony was picked using a pipette tip and placed in the test tube. All tubes were then incubated overnight in a 37°C shaker.
2.3.6 Plasmid Isolation

The plasmids were isolated as in section 2.1.8.

2.3.7 Plasmid Analysis for Presence of ERα:Dam insert

The plasmids were analysed for the presence of an ERα:Dam amplicon by enzyme
digestion and subsequently analysed on a 2% agarose gel. The following 10 μl
restriction digestion reaction was set up: 8 μl of Purified Vector DNA, 0.5 μl of Bgl II,
0.5 μl of Apa I and 1 μl of 10 x NEB 2 buffer. The restriction digestion reactions were
carried out at 37°C for 1 hour.

The reactions containing the digested vectors were then run on a 2 % agarose gel at 100
volts, as explained in step 2.1.3, to allow the visual detection of ERα:Dam amplicon
inserts.
2.4 Functional Analysis of ERα:DAM Vector

2.4.1 Luciferase assays on newly generated ERα:Dam fusion vectors

The transfection of HeLa cells with the newly generated ERα:Dam fusion vector followed by luciferase assay allow functionality testing of the expression vector. Initially it can be used to test the functionality of the newly generated ERα:Dam expression vector in conjunction with the luciferase reporter vector pERE-E1b-LUC, which contains an ERE derived from the Xenopus vitellogenin A2 enhancer, hormone treatment and a luciferase assay. This allows the determination of not only whether the newly generated expression vector ERα:Dam is able to express the fusion protein, but also, if this fusion protein can bind to the ERE sequence in the pERE-E1b-LUC and cause expression of the Luciferase protein.

This ERα:Dam driven expression of luciferase protein can be determined using the subsequent luciferase assay which allows the detection of any expressed luciferase protein. As HeLa cells do not contain any endogenous ERα, the detection of luciferase protein would show that the ERα:Dam expression vector is able to express the ERα:Dam fusion protein and that this protein is able to bind to an ERE and drive the expression of luciferase, thus proving its expression and binding capabilities.

Procedure:

This transfection protocol can be used

1) to test the functionality of the generated ERα:DAM vector in conjunction with pERE-E1b-LUC using a luciferase assay and

2) in the subsequent PCE amplification of methylated sequences following the transformation.
The newly generated ERα:Dam vector was made up to concentrations of 5ng, 10ng, 20ng, 100ng and 250ng. A master mix of Lipofectamine and DMEM containing no serum (3μl and 250 μl per well respectively) was made up initially according to the amount of wells. The appropriate concentration of ERα:DAM Vector was added to the individual aliquots of Lipofectamine / DMEM without serum mix. The reporter vector pERE-E1b-LUC was also added to each tube.

To each well containing HeLa cells 1 ml of DMEM was added, followed by 250 μl of Vector / Lipofectamine / DMEM without serum mix. After 3 hours the DMEM was removed from the individual wells using a sterile automated pipette technique. 2 ml of DMEM containing serum was added to each individual well and these were left to incubate overnight.

After 24 hours either 0.5 μl ethanol, Oestradiol 10^-6 dissolved in ethanol or Tamoxifen 10^-4 dissolved in ethanol was added to each of the wells. This was then incubated overnight for 24 hours. The cells were ready for luciferase assay. A passive lysis technique was used to measure luciferase gene expression. The growth medium was removed from the wells using a sterile automated pipette technique. 500 μl of 1 x passive lysis buffer (PLB) was added into each well. This promotes the rapid lysis of cultured mammalian cells without the need for cell scraping. The plates were placed onto a rocking orbital shaker for 15 minutes to ensure complete coverage of the cell monolayer with 1 x PLB. The cell lysates are added to new tubes. 100 μl of Luciferase Assay reagent was added to each tube. The luciferase assays were then carried out on the plate reading luminometer.
2.5 Isolation of ERα DNA Interacting Sequences

2.5.1 In Vivo Methylation of ERα DNA interacting sequences

Transfection was also used to methylate the gDNA of the HeLa cells using the ERα:Dam expression vector. As the vector had been determined to be functionally active following the luciferase assay, the transfection reaction could be used to methylate in vivo the GATC sequences, following treatment with hormone and binding of the ERα:Dam to ERE sequences in the DNA of the HeLa cells. This allows the DAM protein to methylate GATC sequences in the region of the ERE to which it has bound. Following isolation of the gDNA from the wells the methylated sequences can be specifically amplified using PCR.

Procedure:

The transfection assay was carried out as per step 2.1.10 outlined above. Instead of a luciferase being carried out following transfection, the DNA is isolated and the methylated sequences are subsequently amplified using methylation specific PCR amplification.

2.5.2 Precipitation of gDNA from Transfected Cells

The gDNA of the HeLa cells has been methylated following transfection with the ERα:Dam expression vector, as the ERα:DAM fusion protein will have bound to ERE sequences and methylated the nearby GATC sequences. The DNA is now isolated to allow subsequent PCR amplification of these methylate GATC sequences in order to isolate the nearby EREs. This was done using an alcohol based precipitation technique.
Procedure:

The transfected HeLa cells were harvested using a cell scraper and placed in 5ml of lysis buffer (10 mM Tris pH 8.0, 10 mM EDTA, 0.5% SDS, Proteinase K 200 μg / ml). This solution was incubated at 55°C for 2 hours. 200 μl of 5 M NaCl was then added to the lysate. Extraction was carried out by adding 2 ml of phenol:chloroform:isoamyl alcohol. The solution was incubated for 5 minutes and centrifuged at 500 rpm and the top layer was saved. This process was repeated. Next another extraction was carried out by adding 2 ml of chloroform, followed by centrifugation for 5 minutes at 5000 rpm and the saving of the top layer.

The tube containing the top layer was incubated at 55°C for 1 hour to evaporate all the chloroform. 100 μg of RNAse A was added and the solution was incubated at 37°C for 1 hour. Another extraction using phenol:chloroform:isoamyl and again saving the top layer, was carried out. Next 10 ml of ethanol was added precipitating the DNA, which could then be spooled using a pipette tip. The spooled DNA was washed in 70% ethanol for 5 minutes and the pipette tip was transferred into a 1.5 ml eppendorf tube with 500 μl of TE buffer. This was allowed sit overnight at 4°C for the DNA to dissolve.

2.5.3 Measurement of Isolated gDNA Concentration

A 1 in 50 dilution of all the DNA solutions was prepared by adding 2 μl of DNA to 98 μl of H2O. The spectrophotometer was blanked using 100 μl of H2O. The dilution factor was set and the absorbance at 260 nm (A260) of the isolated gDNA was
measured. The \( A_{260} / A_{280} \) ratio was also measured to determine the purity of the samples.

2.5.4 DpnI restriction digestion of the isolated HeLa gDNA

The isolated HeLa cell gDNA is digested with the restriction enzyme Dpn I, which specifically recognizes methylated GATC sequences. The isolated gDNA from the HeLa cells contains methylated GATC sequences following transfection with the ER\( \alpha \):Dam expression vector. These methylated GATC sequences are digested by Dpn I thus generating a pool of blunt-ended DNA fragments with 5’TC and 3’GA\textsuperscript{met}, allowing their subsequent amplification by ligation mediated methylation specific PCR.

Procedure:

A 10 \( \mu l \) DpnI digestion was set up on all HeLa isolated gDNA samples. This restriction digestion consisted of 2.5 \( \mu l \) of gDNA (2.5 \( \mu g \)), 1 \( \mu l \) of 10 x NEB 4 Buffer, 0.5 \( \mu l \) Dpn I (10 units) and 6 \( \mu l \) of H\textsubscript{2}O. A 10 \( \mu l \) negative control reaction consisting of all reactants except Dpn I was also set up as follows: 2.5 \( \mu l \) HeLa gDNA, 1 \( \mu l \) 10 x NEB 4 Buffer, and 6.5 \( \mu l \) H\textsubscript{2}O. The digestion and negative control reactions were carried out overnight at 37°C. Following completion of digestion reaction, the DpnI was inactivated by incubating for 20 minutes at 80°C.

2.5.5 Ligation of Adaptors

To allow PCR amplification of the Dpn I digested sequences, a double stranded adaptor nucleotide is ligated to the blunt ends. The double stranded adaptor (Fig. 2.3) is
generated from the ligation of two separate single strand oligonucleotides, with one strand having a 5’ overhang.

The 5’ overhang of the double strand adaptor ensures directional ligation to the blunt end DNA sequences. A double stranded adaptor is used as double stranded DNA is ligated more efficiently than single stranded DNA.

Right Adaptor, Top Strand
AdRT 5’CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGA 3’

Right Adaptor Bottom Strand
AdRb 5’TCCTCGGCCG 3’

1st Step : Formation of Adaptor
AdRT 5’CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGA 3’
AdRb 3’GCCGGCTCTC 5’

2nd Step Adaptor Ligation
5’CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGA TCNNNN
3’GCCGGCTCTCAGNNNN

Fig. 2.3 Sequence of Adaptors used for Ligation
Procedure:

The first step involved the generation of a stock of double stranded adaptor by pipetting 50 μl of AdRt (100 μM) and 50 μl of AdRb (100 μM) into a tube together. This was placed in a beaker with boiled water, and the beaker was let cool to room temperature, allowing the adaptors to anneal slowly.

Next a ligation reaction was set up to allow ligation of the double stranded adaptors to the Dpn I digested HeLa cell material. A 20 μl ligation reactions was set up as follows: 10 μl DpnI digested gDNA, 0.75 μl newly ligated double stranded adaptor, 1.5μl 10 x Fast-link ligation Buffer, 0.75 μl 10 mM ATP, 1 μl Fast-Link Ligase and 6 μl H2O. Ligation reactions were carried out overnight at 16°C and subsequently the Fast-link ligase was inactivated by heating the reaction to 70°C for 15 minutes.

2.5.6 Dpn II Digestion

DpnII is a restriction enzyme that digests DNA fragments containing unmethylated GATC sequences. The adaptor ligated Dpn I digested sequences are treated with Dpn II. This ensures only fragments in which all GATC sequences are methylated, making them resistant to degradation, are subsequentially amplified.

Procedure:

DpnII digestion reactions were set up as follows: 20 μl of the ligation reaction, 5 μl 10 x Dpn II buffer, 1 μl Dpn II and 24 μl H2O. The Dpn II digestion reactions were carried out for 1 hour at 37 °C.
2.5.7 PCR Amplification of Methylated Sequences

PCR amplification of the methylated sequences with ligated double stranded adaptor is carried out using an individual PCR primer (Fig. 2.4) that is identical to the 15 most 3', nucleotides of the AdRt oligonucleotide sequence and the 5'TC nucleotides. This primer design prevents the amplification of non specific DNA products which may be derived from random DNA breaks, as the PCR primer overlaps with the GA dinucleotide located at the ends of the DpnI digested fragments, ensuring only these fragments are amplified in the PCR thereby suppressing amplification of non specific ligation products, which generally will not have a GA dinucleotide at their ends.

PCR Primer
AdR-PCR 5’GGTCGCGGCCGAGGA 3’

3rd Step Primer Annealing and Extension
5’CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGA

Fig. 2.4 Sequence of PCR Primer, Annealing and Extension Step

Procedure:
A 50 μl PCR reaction was set up as follows: 4 μl Dpn II digested DNA, 5μl 10 x Accuprime II Buffer, 39.5 μl H₂O, 1.25 μl PCR Primer (50 μm) and 0.25 μl Accuprime Taq Polymerase. The PCR cycler was set up as follows: 3 minutes at 98°C, 1 minute at 94 °C, 1 minute at 65 °C, 15 minutes at 72°C, 1 minute at 94°C, 1 minute at 94°C, 1
minute at 65°C, 10 minutes at 72°C, repeat last three cycles 3 times, 1 minute at 94°C, 1 minute at 65°C, 2 minutes at 72°C and finally a repeat of the last three steps 25 times.

2.5.8 Analysis of PCR Products on Agarose Gel

5µl of the PCR generated products are analysed on a 2% agarose gel as outlined in step 2.1.3 above.

2.5.9 Purification of PCR Products

The remaining contents of the PCR generated products, which are going to be used for cloning, were purified using the Qiagen PCR purification kit. 5 volumes of Buffer PBI was added to 1 volume of the PCR sample and mix. The QIAquick spin column was then placed in the 2 ml collection tube provided. The samples were applied to the QIAquick column and then centrifuged at 10,000 rpm for 60 s.

The flow through was discarded and the QIAquick column was placed back into the same tube. 0.75 ml of wash Buffer PE was added to the QIAquick column and centrifuged at 10,000 rpm for 60 s. The flow through was discarded and the QIAquick column was placed back into the same tube and centrifuged for an additional 1 minute. The QIAquick column was placed into a clean 1.5 ml centrifuge tube and the DNA was eluted by adding 50 µl of Buffer EB to the centre of the QIAquick membrane, followed by centrifugation.
2.6 Cloning of PCR Products using pCR8/GW/TOPO TA Cloning Kit

2.6.1 Cloning reaction using pCR8/GW/Topo Vector

The pCR8/GW/TOPO TA cloning Kit combines Invitrogen’s TOPO cloning and gateway technologies to facilitate 5-minute, one-step cloning of Taq Polymerase-amplified PCR products into plasmid vector with > 95% efficiency. Clones may then be easily sequenced and characterized. The clones may also be transferred from the pCR8/GW/TOPO entry vector (Fig. 2.5) to a Gateway destination vector of choice for expression of the gene of interest in many systems.

The pCR8/GW/TOPO vector is supplied with single 3’-thymidine (T) overhangs for TA Cloning and also with Topoisomerase I covalently bound to the vector. Taq polymerase has a non-template-dependent terminal transferase activity that adds a single deoxadenosine (A) to the 3’ends of PCR products. The linearized vector supplied in this kit has single, overhanging deoxythymidine (T) residues. This allows PCR inserts to ligate efficiently with the vector.

Topoisomerase I from *Vaccinia* virus binds to duplex DNA at specific sites (CCCTT) and cleaves the phosphodiester backbone in one strand (Shuman, 1991). The energy from the broken phosphodiester backbone is converted by formation of a covalent bond between the 3’ phosphate of the cleaved strand and a tyrosyl residue (Tyr-274) of topoisomerase I. The phospho-tyrosol bond between the DNA and enzyme can subsequently be attacked by the 5’hydroxyl of the original cleaved strand, reversing the reaction and releasing topoisomerase. TOPO Cloning exploits this reaction to efficiently clone PCR products.
Procedure:

TOPO cloning reactions were set up as follows: 4 μl of PCR product, 1 μl Salt Solution and 1 μl of Topo Vector. A negative control reaction was set up as follows: 4 μl H2O, 1 μl Salt Solution and 1 μl Topo Vector. The reactions were mixed gently and incubated at room temperature for 30 minutes.

2.6.2 Transformation of Cloning Reactions into Chemically Competent Cells

2 μl of the TOPO cloning reactions were added to a vial of One Shot Chemically Competent E.coli and mixed gently. A positive transformation reaction was set up by transforming 1 μl of the pUC19 control plasmid. The transformations were incubated on ice for 30 minutes followed by heat shock for 30 seconds at 42 °C. The tubes were incubated in a 37 °C horizontal incubator at 200 rpm for 1 hour. 50 μl from each transformation reaction was spread on a spectinomycin plate and incubated overnight at 37°C.
2.6.3 Growth of Resulting Colonies and Plasmid Isolation

All colonies on the spectinomycin plate containing competent cells transformed with PCR product were picked using a pipette and placed into a 3 ml LB culture containing 30 μl of spectinomycin. All tubes were incubated overnight in a 37 °C shaker.

Plasmid DNA was purified from 1.5 ml of culture using the ChargeSwitch-Pro Plasmid MiniPrep Kit as outlined in step 2.1.8 above.

2.6.4 Plasmid Analysis using Restriction Digestion

The plasmids were analyzed for the presence of an insert by enzyme digestion using the restriction enzyme EcoRI. The following digestion reactions were set up: 8 μl Purified Topo Plasmid DNA, 1 μl EcoRI and 1 μl 10 x NEB 3 Buffer. The restriction digestion reactions were carried out at 37°C for 1 hour.

2.6.5 Visualization of Digested Plasmid Products

The reactions containing the digested vectors were run on a 2 % agarose gel at 100 volts, as outlined in step 2.1.3 above, to allow the visual detection of any DNA inserts.
2.7 DNA Sequencing and Subsequent Analysis

2.7.1 DNA Sequencing

All DNA samples for sequencing were made up to a concentration of 200 ng/μl using TE buffer. The primer used for sequencing was the T7 primer.

2.7.2 DNA Sequence Analysis

All subsequent sequences were blasted against the human genome using IMGT blast. This allowed the determination of whether the isolated sequences were of human or plasmid origin.

After blasting of the sequences the transcription factor binding site search engine Promo was used to determine whether the isolated sequences contained ERα binding sites. The transcription factor binding sites, which were searched for, included EREs, SP-1, and AP-1 interacting sites.
2.8 Creation of Expression Vector Containing Isolated Sequences

2.8.1 Generation of Destination Vector

To functionally test the isolated sequences they had to be transferred from the pCR8/GW/Topo cloning vector into a destination vector with reporter capabilities. The Gateway conversion system easily allows the conversion of any vector of choice into a gateway vector. This is achieved by ligating a blunt ended cassette containing attR sites flanking the ccdB gene and the chloramphenicol resistance gene into the multiple cloning site of the chosen destination vector.

The destination vector of choice is the pGL3-Basic vector (Fig. 2.6), which is 4118 bp in length and is a luciferase expression vector. The vector is first linearized by digestion with the restriction enzymes NheI and Hind III. Then the vector is treated with alkaline Phosphatase to remove the 5’phosphates, thus decreasing the chance of vector religation. Following a ligation reaction using fast-link ligase, the reading frame shown below is ligated to the vector.

Procedure:

To allow ligation of the cassettes the pGL3-Basic vector was digested with the restriction enzymes Nhe I and Hind III as follows: 1 μl pGL3 Basic, 1μl Hind III, 1 μl Nhe I and 7 μl Neb 4 Buffer.

The digested vector was then treated with APex Heat-Labile Alkaline Phosphatase, which is a novel Phosphatase that dephosphorylates 5’-phosphates. To the completed restriction digest, 1 μl of APex Heat-Labile Phosphatase was added. This was
incubated at 37°C for 10 minutes followed by heat inactivation of the Phosphatase at 70°C for 5 minutes.

The reading cassette was ligated to the digested pGL3-Basic Vector using Fast-link ligase. The following ligation reaction was set up: 1μl Dephosphorylated pGL3 Basic, 2 μl Gateway Reading Cassette, 1 μl Fast-link Ligation Buffer, 0.75 μl 10 mM ATP, 1 μl Fast-Link Ligase and 10.25 μl H2O. The ligation reaction was incubated for 1 hour at room temperature.

Fig. 2.6 Newly Generated pGL3 reporter Destination Vector

2.8.2 Transformation of One Shot ccdB Survival Competent Cells

Next the ligation reaction was transformed into One Shot ccdB Survival Competent Cells. 1 μl of the ligation reaction was added to 1 vial of the One Shot ccdB Competent Cells and then mixed gently. The vials were incubated on ice for 30 minutes followed by heat shock for 30 seconds at 42°C.
The vials were then placed back into the ice for another two minutes. 250 μl of SOC medium was added to the vial which was incubated for 1 hour at 37°C. 25 μl of the transformation reaction was spread on an ampicillin plate and incubated overnight at 37°C. Resulting colonies were picked and grown overnight in LB medium containing 30 μg / ml chloramphenicol. The plasmid DNA was isolated using the Qiagen plasmid isolation kit as outlined on step 2.1.8.
2.9 Transfer of Isolated Sequences Into Destination Vector

2.9.1 LR Recombination Reaction

The isolated sequences of human origin had to be transferred from the pCR8/GW/TOPO entry vector into the newly generated destination vector using the LR recombination reaction (Fig. 2.7). The gateway system takes advantage of the site-specific recombination reactions enabling the bacteriophage λ to integrate and excise itself out of a bacterial chromosome. The LR clonase mix that consists of integrase, integration host factor and the phage excisionase, catalyses the LR reaction, in which the DNA insert is transferred from the entry clone to the destination clone.

In this case the LR clonase mix transfers the isolated sequence which is flanked by attL1 and attL2 sites in the pCR8/GW/TOPO entry vector into the newly generated pGL3 Basic destination vector with ligated cassette containing two attR sites. Following recombination of the matching attR and the attL sites the DNA fragment of interest is flanked by attB sites. A LR recombination reaction is set up for all pCR8/GW/TOPO vectors that had been determined to contain inserts of human origin after sequencing. The LR recombination reactions were transformed into competent...
bacteria and grown overnight on ampicillin selective plates. All individual colonies were grown overnight in selective media and a plasmid preparation was carried out on individual cultures to allow isolation of the newly generated pGL3 Basic luciferase reporter vectors containing DNA inserts of human origin.

**Procedure:**

To transfer all isolated DNA sequences of human origin into the reporter destination vector to allow function analysis a LR Recombination reaction was carried out. The following outlines the reactions: 5 μl of the pCR8/GW/Topo entry vectors with isolated sequences, 1 μl pGL3 Basic modified Destination Vector, 2 μl TE buffer and 2 μl of LR Clonase Enzyme mix. A positive reaction was also set up as follows: 1 μl pENTR-gus, 1 μl pGL3 Basic modified destination vector 6 μl TE buffer and 2 μl of LR Clonase Enzyme mix. The reactions were incubated at 25°C for 1 hour. 1 μl of proteinase K was added to each reaction and incubated for 10 minutes at 37°C. 1 μl of each reaction was transformed into One Shot Competent E.coli as outlined in step 2.1.6.
2.10 Determination of Functional Activity using Luciferase Assays

2.10.1 Functional Analysis of Isolated Sequence using ERα

To determine if the sequences were able to bind ERα and allow subsequent gene transactivation a luciferase assay was carried out following co-transfection of HeLa cells with 2.5 ng pCR3.1 hERα and each individual newly generated reporter destination vector containing isolated DNA sequences of human origin as outlined in step 2.1.10. Raloxifene hormone treatment was used instead of Tamoxifen.

2.10.2 Determination of Indirect or Direct DNA binding

To determine if ERα was binding to the isolated sequences directly or indirectly to cause luciferase gene expression a luciferase assay was carried out following co-transfection of HeLa cells with 2.5 ng of the ERα DNA binding mutant C202H / C205H expression vector and each individual newly generated reporter destination vector containing isolated DNA sequences of human origin as outlined in step 2.1.10. Raloxifene hormone treatment was used instead of Tamoxifen

2.10.3 Functional Analysis of Isolated Sequences using ERβ

To determine if ERβ was able to bind the isolated sequences of human origin and regulate gene expression a luciferase assay was carried out following co-transfection of HeLa cells with 2.5 ng pCR3.1 hERβ and each individual newly generated reporter destination vector containing isolated DNA sequences of human origin as outlined in step 2.1.10. Raloxifene hormone treatment was used instead of Tamoxifen.
Results

This study aimed to use ChIP and DamID to generate a reporter library of genomic ERE cis elements that could be functionally evaluated for their ability to regulate ERα mediated gene expression, thereby gaining a greater understanding of ERα mediated transcription.

3.1 ChIP

ChIP was the first method employed to generate a reporter library of genomic ERE cis elements, which could then be functionally evaluated. ChIP assays allow the detection of Protein: DNA interactions using the principle of immunoprecipitation. Using ChIP, cells are fixed using formaldehyde, which is a cross-linking agent and thus preserves protein: DNA interactions at that certain moment. The DNA is the enzymatically digested and the protein: DNA complexes can be isolated using a specific antibody directed against the protein of interest. The ChIP assay in theory would allow the isolation of EREs with prebound ERα in the absence of ligand induction, EREs with bound ERα following agonist treatment and also EREs with bound ERα following antagonist treatment. These precipitated ERα:DNA complexes are isolated using antibodies directed against ERα, in this case ERα (HC20): sc 534 as per Materials and Methods section 2.1.1.

HeLa cells were treated with 0.5 μl ethanol, Oestradiol 10^-6 and Tamoxifen 10^-4 for 24 hours. The genomic DNA was then isolated using a dounce homogeniser as per Materials and Methods section 2.1.1. Next an interaction reaction containing digested DNA with bound ERα and anti-ERα antibodies was set up as per Materials and
Methods section 2.1.1. The anti ERα : ERα : DNA complexes were now bound to the magnetic beads following the interaction reaction. They were then eluted using elution and reverse cross-linking reagents as per Materials and Methods section 2.1.1. The ERα bound gDNA sequences were now isolated and ready for cloning into a reporter vector.

The reporter vector used as the cloning vector was pERE-E1b-LUC as per Materials and Methods section 2.1.2. This contained a consensus ERE sequence, which was excised using the restriction enzymes Nhe I and Bgl II. The isolated gDNA ERα bound sequences were then ligated to the digested reporter vector using the fast-link ligase kit as per Materials and Methods section 2.1.5. The ligation reactions were then transformed into DH10B competent cells and, following incubation, were grown overnight on ampicillin selective plates. Colonies were picked into LB and incubated overnight at 37°C in a shaking incubator as per Materials and Methods section 2.1.7.

The bacterial cultures were treated using a plasmid mini prep as per Materials and Methods section 2.1.8, allowing isolation of transformed plasmids. Following purification, the plasmids were analysed by restriction enzyme digestion using KpnI and Hind III as per Materials and Methods section 2.1.9. The digested material was run on a gel to allow the detection of any plasmids containing inserts. Sample results of this are presented in Figures 3.1 and 3.2.
Fig. 3.1 KpnI and Hind III restriction digestion analysis of plasmids resulting following ligation of NheI, Bgl II digested pERE-E1b-Luc vector with ChIP isolated DNA sequences

Fig. 3.2 KpnI and Hind III restriction digestion analysis of plasmids resulting following ligation of NheI, Bgl II digested pERE-E1b-Luc vector with ChIP isolated DNA sequences
All 3 plasmids containing inserts were functionally assayed by carrying out a luciferase assay using the reporter constructs alone, or in conjunction with the pCR3.1 hERα expression vector, which encodes WT ERα, or the ERα C202H / C205H expression vector which encodes a hERα mutant containing the double mutation C202H / C205H within the DBD. The luciferase assays were carried out on cells treated with ethanol, the agonist oestradiol and the partial agonist / antagonist Tamoxifen. A control co-transfection using pERE-E1b-Luc reporter vector was initially performed to demonstrate that activation of the pERE-luc construct could be detected with WT ERα in the presence of oestradiol and Tamoxifen. These results are presented in Fig. 3.3. This process was repeated using constructs containing the ChIP isolated sequences, shown in Figs 3.4, 3.5 and 3.6.

![Fig. 3.3 : Luciferase Assay Results following co transfection of WT ERα expression vector, aa 202/205 ERα mutant expression vector or no expression vector with the pERE-E1b-Luc reporter vector. Transfected cells were treated with Ethanol, Oestradiol or Tamoxifen](image-url)
Fig. 3.4 : Luciferase Assay Results following co transfection of WT ERα expression vector, aa 202/205 ERα mutant expression vector or no expression vector with a luciferase reporter vector containing DNA sequence 7 as pulled down using ChIP assay. Transfected cells were treated with Ethanol, Oestradiol or Tamoxifen.

Fig. 3.5 : Luciferase Assay Results following co transfection of WT ERα expression vector,, aa 202/205 ERα mutant expression vector or no expression vector with a luciferase reporter vector containing DNA sequence 12 as pulled down using ChIP assay. Transfected cells were treated with Ethanol, Oestradiol or Tamoxifen.
Fig. 3.6: Luciferase Assay Results following co-transfection of WT ERα expression vector, aa 202/205 ERα mutant expression vector or no expression vector with a luciferase reporter vector containing DNA sequence 44 as pulled down using ChIP assay. Transfected cells were treated with Ethanol, Oestradiol or Tamoxifen.

When carrying out functional analysis on these ChIP isolated sequences it was found that sequence 12 (Fig. 3.5) and sequences 44 (Fig. 3.6) may contain DNA sequences that are ERα binding sites. The basal level of expression is the level of luciferase gene expression when the cells were treated with ethanol. On treatment with oestradiol the level of gene expression increased 2 fold for sequence 12 (Fig. 3.5) and 12 fold for sequence 44 (Fig. 3.6). On treatment with Tamoxifen the level of gene expression increased 2 fold for sequence 12 (Fig. 3.5), and 2 fold for sequence 44 (Fig. 3.6) These results indicate that these sequences may facilitate ERα gene transactivation through its binding and subsequent recruitment of coactivators and the Pol II complex.
Sequence 7 did not seem to adhere to the classical model of ERα mediated gene transactivation as, upon oestradiol treatment, no fold increase in gene expression occurred (Fig. 3.4).

The ChIP method had not proved efficient for the isolation of genomic ERE cis elements. This method did not show a good efficiency even though there were colonies following transformation of the ligation reactions. But on analysis it was found that many of these plasmids did not contain inserts and, instead, were religated vectors as is demonstrated in Fig. 3.1. Even though the vectors were treated with alkaline Phosphatase, this religation problem persisted throughout my research.

Taken together this data demonstrates that ChIP was not efficient in isolating ERα binding sequences so the alternative DamID method was employed.
3.2 Isolation of ERE cis Elements using DamID

The ChIP assay had failed to efficiently generate a reporter library of genomic ERE cis elements despite much modification and repetition so another technique had to be employed to try and achieve our goal. This technique is Dam ID and is based on the generation of a fusion protein consisting of DNA adenine methyltransferase (Dam) and the DNA interacting protein of interest, in this case ERα. Following in vivo expression of this chimaeric protein using transfection assays, the dam protein is targeted to the ERα binding site, and this in turn leads to the methylation of adenines in GATC sequences that are close to the ERα binding sites (Fig. 1.23). Selective amplification of these methylated sequences is then carried out using LM-PCR.

An overview of the experimental approach used for the DamID ERα interacting sequence isolation process is illustrated in Fig. 3.7. The results of each step will be presented.
Fig. 3.7 Overview of the DamID Mediated ERα gDNA Interacting Sequence Isolation Process

Step 1
Generation of ERα : DAM Fusion Product

Step 2
Generation of ERα : DAM Expression Vector

Step 3
Functional Testing of the ERα : DAM Expression Vector Using Luciferase Assays

Step 4
Transfections using ERα : DAM Expression Vector, DNA Isolation and Concentration Determination

Step 5
Specific Amplification of methylated ERE sequences using LM-PCR

Step 6
Topo Cloning of the Generated PCR Products

Step 7
Plasmid Isolation and Subsequent Screening For Presence of Inserts

Step 8
Sequencing of Plasmids containing Inserts

Step 9
Generation of Reporter Vectors Containing the DamID Isolated Sequences

Step 10
Function Analysis of the Isolated Sequences Contained in the Newly Generated Reporter Vectors
3.2.1 Step 1: Generation of ERα: DAM PCR Product

The first steps involved in the generation of an ERα: Dam Fusion PCR product was the production of the individual PCR products for both ERα and DAM using specific primers, as per Materials and Methods sections 2.2.1 and 2.2.2. pCR3.1 hERα, which is an expression vector for ERα, was used as the PCR template for ERα while bacterial genomic DNA isolated from E.coli was used as the PCR DNA template for the amplification of Dam.

The ERα forward primer contains a Bgl II restriction recognition site at its 5’end and a sequence specific for the internal part of ERα at its 3’end. The ERα reverse primer contains a 3’ ERα binding sequence at its 5’end and a 5’Dam binding sequence at its 3’end (Fig. 3.8 A). Thus the ERα generated following PCR will contain a Bgl II recognition sequence at its 5’end and an overlapping 5’Dam sequence at its 3’end. The 5’Bgl II restriction recognition site is to facilitate ligation of this ERα:Dam into a Bgl II and Apa I digested pCR3.1 hERα vector while the Dam sequence at its 3’end is to facilitate the generation of the ERα:Dam fusion protein through overlapping with the Dam PCR product (Fig. 3.8 B). The ERα PCR product will contain just the 3’ region of the receptor, and this is done to avoid Taq Polymerase generated mistakes which may occur in amplification of a full length ERα product.

The Dam forward primer is a reverse complement of the ERα reverse primer and it contains a 5’ERα sequence and a 3’Dam sequence (Fig. 3.8 B). The Dam reverse primer contains at its 5’ end a sequence complementary to the 3’end of the Dam protein and at its 3’end an Apa I recognition sequence (Fig. 3.8 B). The Dam forward primer
contains the complementary sequence to the ER$\alpha$ 3’end to facilitate the generation of the ER$\alpha$:Dam fusion protein through overlapping with the ER$\alpha$ PCR product while the reverse primer contains an Apa I site to allow subsequent ligation of the ER$\alpha$:Dam fusion protein into the digested pCR3.1 hER$\alpha$ expression vector. In this way the fusion protein is generated with the Dam protein attached to the C-Terminal of ER$\alpha$.

**ER$\alpha$ Forward Primer**

Bgl II Restriction Site
- - - - - - - - - - - - - Internal ER$\alpha$ Sequence
GGC ATG GTG GAG ATC TTC GAC ATG CTG CTG

**ER$\alpha$ Reverse Primer**

GCA GAG GGT TTC CCT GCC ACA GTC AAG AAA AAT CGC GCT TTT TTG AAG Dam 5’End

**Fig. 3.8 A Forward and Reverse Primers used in the Generation of the ER$\alpha$ PCR Product**

**Dam Forward Primer**

ER$\alpha$ 3’End

CTT CAA AAA AGC GCG ATT TTT CTT GAC TGT GGC AGG GAA ACC CTC TGC

**Dam Reverse Primer**

Apa I Restriction Site Stop Dam 3’End

ATC GCA GGG CCC GAA TTA TTT TTT CGC GGG TGA AAC GAC
Following the PCR generation of the ERα and Dam individual PCR products, the products were run on a 2% agarose gel to allow detection of product and to correct size.

![Agarose gel showing molecular weight ladder and PCR products](image)

Fig. 3.9: 2% Agarose gel containing Molecular Weight Ladder, Lane 1 and 2 ERα PCR products

Lane 3 and 4 Dam PCR Products

The correct PCR product sizes were observed so a second PCR reaction was carried out with both the individual ERα and Dam PCR products and the ERα forward primer and the Dam reverse primer. Due to the overlapping regions of each product this second PCR reaction is able to generate the fusion protein following annealing and extension of the ERα and Dam reverse primers.
Using PCR templates of an ERα expression vector and gDNA isolated from E.coli, it was possible to PCR amplify individual PCR products for DAM and ERα. These products were then used in a further PCR reaction to generate the ERα:Dam fusion product shown in the agarose gel above.

The ERα:Dam fusion product generated is ligated back into the pCR3.1 hERα vector which has been digested with Bgl II, and thus generates a full length ERα with the Dam DNA sequence attached at its 3’end. When expressed this protein will consist of full length ERα with its fusion partner attached at the C-Terminal (Fig. 3.11).
**Fig. 3.11:** Overview of the Processes involved in the Generation of Individual ERα and Dam PCR Products, followed by the Generation of the ERα:Dam Fusion PCR Product

- **pCR 3.1 hERα Expression**
- **PCR Amplification using Specific Primers to ERα and DAM**
- **E.coli Genomic DNA isolated**

**PCR Reaction incorporating Individual ERα and Dam PCR Products and ERα forward Primer and Dam Reverse Primer**

**ERα : DAM fusion PCR Product**
3.2.2 Step 2: Generation of the ERα:Dam Expression Vector

Following the successful generation of the ERα : DAM fusion product, the next step was to incorporate this product into the pCR3.1 hERα expression vector. This was digested with Bgl II and Apa I restriction enzymes, as per Materials and Methods 2.3.1, which linearized the vector. The ERα:Dam fusion PCR product was also digested with Bgl II and Apa I as per Materials and Methods section 2.2.6 to allow successful ligation. Ligation of ERα:Dam to the linearized pCR3.1 hERα expression vector was carried out using the fast-link ligase kit.

Fig. 3.12 Overview of the Process Involved in the Generation of the ERα:Dam fusion Expression Vector
The products from the ligation reaction were then transformed into competent bacteria and allowed to grow overnight. Overnight transformed colonies were grown for a second night in a LB suspension. Overnight bacterial suspensions were treated with plasmid mini-prep to isolate any ERα:Dam containing vector. The vector was cut with Bgl II and Apa I and run on an agarose gel (Fig. 3.13) to allow the identification of the presence of a vector with the ERα:Dam insert as per Materials and Methods section 2.3.7.

Fig. 3.1 : Overview of the Process involved in the Generation of the ERα:Dam Fusion Expression Vector

Transformation of competent bacteria, subsequent growth of colony and plasmid preparation

Restriction Digestion of Plasmids and analysis on Agarose gel to detect plasmid with ERα : DAM insert
Newly generated pCR3.1 ERα:Dam expression vector following restriction digestion with the enzymes Bgl II and Apa I yielded a band at ~ 850 bp which represents the ERα:Dam fusion insert. This was the predicted size.
3.2.3 Step 3 Functional Testing of ERα : DAM expression Vector

The newly generated ERα:Dam fusion proteins ability to bind DNA and cause target gene expression must be determined. One potential disadvantage with the Dam ID method is that ERα may lose some or its entire DNA binding abilities due to its fusion with Dam. This would render the fusion protein useless for the isolation of ERα binding sites.

To test the ERα:Dam fusion proteins ability to bind DNA a luciferase assay is carried out (Fig. 3.15). This involves the cotransfection of HeLa cells with the fusion protein and an ERE containing luciferase reporter. The reporter vector used was pERE-E1b-LUC (Fig. 2.1 Materials and Methods) which is an ERE containing luciferase reporter vector. This reporter vector contains an ERE from the Xenopus vitellogenin A2 enhancer. Cotransfection of the expression vector containing the ERα:Dam fusion and the ERE luciferase reporter vector would test the ability of the ERα:Dam to bind a conventional ERE and initiate luciferase gene expression. The level of luciferase gene expression can then be measured using a luciferase assay.

The ability of the ERα:Dam to differentially regulate gene expression following the binding of activators and antagonists could also be tested. If the EREs ligand-binding domain was not affected by its fusion to the Dam protein, it can cause differential expression of the luciferase gene following treatment with various ligands. A hormone treatment step is carried out following co-transfection of HeLa cells with the ERα:Dam expression vector and the ERE reporter vector. The hormones used were oestradiol and tamoxifen. Ethanol was also administered to determine basal levels of luciferase gene expression. If ERαs LBD has not been affected following its fusion to Dam it will be
possible to see different levels of luciferase gene expression following treatment with ethanol, oestradiol and tamoxifen.

Fig. 3.15 Overview of Luciferase Assay for the Functional Testing of the pCR3.1 ERα:Dam Expression Vector
The activity of Dam is high when it is expressed in cultured cells. If moderate or high levels of Dam are expressed in the cell then this would cause saturating amounts of non-targeted methylation thus rendering this technique for the isolation of DNA ERα binding sites useless. Therefore, it is important to express the ERα:Dam fusion protein at very low levels. Using varying concentrations of the ERα:Dam reporter vector in the transfection and subsequent luciferase assays, it is possible to determine the lowest levels at which the ERα:Dam is still able to bind an ERE and cause gene trans activation. Establishing this ideal concentration of the ERα:Dam reporter vector minimizes the incidence of non-targeted methylation occurring.

A co-transfection of the pCR3.1 hERα expression vector with the pERE-E1b-Luc reporter vector is also carried out in duplicate so as to determine if the transfection and luciferase assays have worked. This also allows a comparison of gene expression levels between a wild type ERα and the ERα:Dam fusion protein.
Fig. 3.16  Luciferase Assay Results Following Co-Transfection of the pERE-E1b-Luc reporter vector and Increasing Concentrations of the ERα:Dam Expression Vector, or Standard Concentration of WT ERα Expression Vector as Indicated. Cells were treated with Ethanol, Oestradiol or Tamoxifen for 24 hours  RLU=Relative Light Units
From Fig. 3.16 we can see that there is clear evidence that the fusion protein is in fact able to bind to DNA and initiate luciferase gene expression. This proves that the ER\(\alpha\):Dam fusion protein when overexpressed in the HeLa cells was able to bind the ERE and transactivate the cotransfected pERE-E1b-Luc reporter vector. As HeLa cells do not contain endogenous ER\(\alpha\), no luciferase expression would have occurred had the ER\(\alpha\):Dam fusion protein been unable to bind the reporter vector ERE.

The second question related to the ability of ER\(\alpha\):Dam to cause differential gene expression levels following hormone treatment by agonists, antagonists or no hormone treatment at all. The hormones used for treatment were the agonist oestradiol, the partial agonist/antagonist tamoxifen and ethanol control, which simulated no hormone treatment. It is clear from Fig. 3.16 that the ER\(\alpha\):Dam fusion protein is able to cause differential levels of gene expression following various hormone treatments. The highest levels of gene expression are associated with oestradiol hormone treatment, followed by the gene expression levels associated with ethanol and the lowest levels are associated with tamoxifen hormone treatment. Gene expression levels following treatment with ethanol represents basal levels of luciferase expression.

Expression levels following treatment with oestradiol are 3-4 fold higher than the basal levels of expression (Fig. 3.16). This is due not only to activation of ER\(\alpha\):Dam through dimerisation and resulting increased interactions with the ERE but also as the activated receptor is able to recruit coactivators and transcriptional machinery enabling high levels of luciferase gene expression. Tamoxifen hormone treatment is associated with levels of gene expression that are half as high as the basal levels. Even though
tamoxifen bound ERα:Dam may bind the ERE it may be the recruitment of corepressors by tamoxifen bound ERα:Dam which acts to lower even the basal levels of luciferase gene transcription. This data shows that our ERα:Dam fusion protein is able to differentially regulate gene expression levels through the binding of various ligand.

A disadvantage of DamID is that, unless the fusion protein is expressed at low levels, saturated Dam protein levels may cause no-targeted methylation of GATC sequences. To reduce the possibility of this occurring a concentration of ERα:Dam must be chosen that is high enough to allow DNA binding and subsequent gene activation but also low enough to avoid saturating levels of the Dam protein. To determine the ideal concentration varying concentrations of the ERα:Dam fusion proteins were co-transfected with the pERE-E1b-Luc reporter vector. As can be seen from Fig. 3.16 all concentrations of ERα:Dam were able to bind the ERE and subsequently able to promote luciferase gene expression.

Maximum expression was observed at a concentration of 20 ng while minimum expression was observed at a concentration of 250 ng. The lowest concentration of transfected ERα:Dam fusion protein had the second highest expression values. Thus, due to its moderately high expression values and its lowest concentration, it was decided that a concentration of 5ng of ERα:Dam fusion protein would be used in subsequent steps to minimise the risk of non targeted methylation while still allowing moderate DNA binding and transactivation.
3.2.4 Step 4: Transfection of HeLa cells with ERα:Dam Expression Vector

The ERα:Dam expression vector had been shown not only to bind to an ERE and cause target gene expression but also to regulate luciferase expression differentially following binding of different hormones. The principle of DamID is based on the targeting of the Dam to its fusion partners genomic binding site where methylation of the GATC sequences in close proximity to the binding site occur. The ERα:Dam therefore had to be shown to bind the ERα DNA response element, namely, the ERE which had been established using luciferase assays.

The next step is to express the ERα:Dam expression vector in HeLa cells which, via binding of ERα to its endogenous DNA binding sites, will allow Dam methylation of nearby GATC sequences. 5ng of ERα:Dam was transfected into HeLa cells and hormone treatment was also carried out with oestradiol for 24 hours. Some transfected cells were also treated with ethanol, which would allow the isolation of ERα:Dam that binds DNA in absence of hormone treatment. A transfection reaction using wild type ERα was also set up which would act as a control in the following LM-PCR mediated specific amplification of methylated GATC sequences, as no GATC specific methylation should occur when cells are transfected with WT ERα alone. The WT ERα transfected cells were also treated with oestradiol so as to increase the occupancy of EREs by the expressed ERα protein.

Following transfection and subsequent hormone treatment as per Materials and Methods section 2.5.1, the gDNA of the cells was isolated using an ethanol precipitation method as per Materials and Methods section 2.5.2. The DNA was then quantified using a spectrophotometer as per Materials and Methods 2.5.3.
**Fig. 3.17 Overview of the ER\(\alpha\):Dam Transfection Process and DNA Concentration Measurement**

<table>
<thead>
<tr>
<th>Transfection Vector</th>
<th>Hormone Treatment</th>
<th>Isolated DNA Conc. (\mu g / \mu l)</th>
<th>(A_{260/280})</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER(\alpha):Dam</td>
<td>Ethanol</td>
<td>0.342</td>
<td>1.42</td>
</tr>
<tr>
<td>ER(\alpha) : Dam</td>
<td>Oestradiol</td>
<td>0.667</td>
<td>1.46</td>
</tr>
<tr>
<td>WT ER(\alpha)</td>
<td>Ethanol</td>
<td>0.405</td>
<td>1.42</td>
</tr>
</tbody>
</table>

**Table 3.1** The DNA concentrations and A260/280 readings of isolated gDNA from hormone treated HeLa cells transfected with ER\(\alpha\):Dam and WT ER\(\alpha\)

A reasonable yield of DNA was achieved (Table 3.1). The \(A_{260/280}\) indicates the quality of DNA extracted. The value is < 1.8 which may indicate protein contamination.
3.2.5 Step 5: Specific Amplification of methylated ERE sequences using LM-PCR

Following isolation of the partially methylated gDNA the next step was to specifically amplify these methylated sequences. This is achieved by using a methylation specific LM-PCR method. This LM-PCR procedure (as per Materials and Methods section 2.5.7) ensures that only methylated DNA fragments are amplified as the sequential use of DpnI and DpnII restriction digestion creates a double selection of methylated sequences. This is because only methylated GATC sequences are cut by DpnI and therefore ligated to the adaptors, while DpnII ensures that only fragments in which all GATC sequences are methylated, making them resistant to degradation, are subsequentially amplified.

Also any non specific DNA products which may be derived from random DNA breaks are excluded from the PCR by the specific Primer design, as the PCR primer overlaps with the GA dinucleotide located at the ends of the DpnI digested fragments; this ensures only these fragments are amplified in the PCR thus suppressing amplification of non specific ligation products, which generally will not have a GA dinucleotide at their ends.

Also included in the PCR reaction is a control for DpnI digestion. This is gDNA isolated from HeLa cells that are transfected with ERα:Dam, but do not undergo subsequent DpnI treatment, thus no adaptor mediated amplification of these methylated sequences should take place. Following the PCR reaction (as per Materials and Methods section 2.5.7) the samples are run on an agarose gel to allow visualisation of any resulting PCR products.
Fig. 3.18  Overview of the LM-PCR Process
Fig. 3.19: Agarose Gel Electrophoresis of LM-PCR Products from LM-PCR gDNA isolated from HeLa cells transfected with ERα:Dam and WT ERα expression vectors and Treated with EtOH or Oestradiol for 24 hrs. gDNA was digested with Dpn I where indicated prior to PCR amplification.

Lane 1 (Fig. 3.19) contains LM-PCR generated PCR products resulting from the amplification of GATC methylated sequences following the transfection of HeLa cells with the ERα:Dam expression vector and treatment with ethanol. The lane shows LM-PCR products ranging in size from ~400 bp to 850 bp. These resulting PCR products show that not only is the fusion product able to bind DNA but it is also causing methylation. Even though the ERα:Dam transfected HeLa cells were treated with
ethanol there are still PCR products as it is known that ERα is able to bind DNA even in the absence of hormone, though to a much lesser extent.

Lane 2 (Fig. 3.19) contains LM-PCR generated products resulting from the amplification of DNA containing methylated GATC sequences following the transfection of HeLa cells with the ERα:Dam fusion protein and their subsequent treatment with oestradiol. PCR products in this lane range in size from ~350 bp to > 1000bp. A greater spread in size of these PCR products is due to treatment of the HeLa cells with the agonist oestradiol, which enhances ERα binding to genomic DNA through its activation. In this way a greater quantity of ERα:Dam fusion proteins can bind DNA, enabling increased methylation of the genome by Dam, which results in a greater quantity of methylation specific PCR products.

Lane 3 (Fig. 3.19) contains LM-PCR generated products resulting from the amplification of DNA following the transfection of HeLa cells with the WT ERα protein and their subsequent treatment with oestradiol. Genomic methylation is not expected in this case as Dam is absent. However, it can be seem from the gel that PCR products ranging in size from ~750 to 850 bp are present. These amplified products are probably due to PCR amplification of bacterial plasmid. As the ERα expression vector is of bacterial origin it has methylated GATC sequences. Thus the resulting PCR products are generated from methylation specific amplification from DpnI digestion of methylated GATC sequences in the transfected WT ERα expression vector sequences following DpnI digestion. Smaller bands at 100bp in size are PCR primer dimers. These are the result of Primers annealing to each other forming a dimer, due to the low concentration of target amplicons.
Lane 4 (Fig. 3.19) only contains primer dimers. This is due to the fact that the DNA template in this case is gDNA from HeLa cells that have been transfected with the ERα:Dam expression vector, and treated with oestradiol; however no subsequent Dpn 1 digestion has taken place. As no Dpn 1 digestion has taken place, the methylated GATC sequences are not isolated and thus no methylation specific amplification takes place. This is the case for methylated GATC sequences in the vector. Lane 5 is a control on the Dpn 1 digestion step. No products are detected on the agarose gel as the primers have not been used up as indicated by the presence of primer dimers (Fig. 3.19).
3.3 Cloning and Sequencing of Isolated Sequences

3.3.1 Step 6 : Topo Cloning of the Generated PCR Products

As the ERα DNA binding sequences have now been isolated and amplified, the next step is to sub-clone them into a vector. The vector used to clone the PCR products is pCR8/GW/TOPO (Fig. 2.5 Materials and Methods). This vector contains a TOPO cloning site for efficient cloning of Taq amplified PCR products. Taq polymerase has a non-template dependent terminal transferase activity that adds a single deoxyadenosine (A) to the 3’ ends of PCR products. The pCR8/GW/TOPO vector has overhanging deoxythymidine (T) residues and this allows the PCR inserts to ligate efficiently with the vector (Fig. 3.20).

The TOPO cloning reactions were set up as per Materials and Methods section 2.6.1. A negative control was also included in which no PCR product was included to test for vector religation. The products of the cloning reactions were transformed into One Shot Competent E.coli as per Materials and Methods section 2.6.2 and plated onto spectinomycin plates for overnight growth. The plates were examined for the presence of colonies.
Fig. 3.20 Overview of the TOPO Cloning Process
The following table shows that the TOPO cloning reaction was very successful. ~700 colonies resulted following overnight growth of bacteria transformed with Topo cloning reactions of pCR8/GW/TOPO with PCR products generated from transfection of HeLa cells with ERα:Dam and subsequent treatment with ethanol, oestradiol or tamoxifen. The negative control in the TOPO cloning reaction consisted of all constituents apart from DNA. The transformation of bacteria with products of this negative control resulted in the growth of 6 colonies. These colonies represent re-ligated vector. This very low number for religated vector indicates that the vast majority of colonies that grew following transformation of bacteria with TOPO reactions containing DNA, may contain a religated plasmid with a DNA insert. This is investigated in the next step.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>No of Colonies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topo reaction with PCR Generated Products From HeLa cells transfected with ERα:Dam and treated with Ethanol</td>
<td>~ 670</td>
</tr>
<tr>
<td>Topo reaction with PCR Generated Products From HeLa cells transfected with ERα:Dam and treated with Oestradiol</td>
<td>~ 750</td>
</tr>
<tr>
<td>Topo reaction with PCR Generated Products From HeLa cells transfected with ERα:Dam and treated with Tamoxifen</td>
<td>~ 720</td>
</tr>
<tr>
<td>Negative Control containing TOPO only Reaction</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 3.2: Table showing number of colonies following overnight growth of One Shot Competent E.coli cells transformed with Topo cloning reaction products
3.3.2 Step 7: Plasmid Isolation and Subsequent Screening For Presence of Inserts

Individual colonies were picked and grown overnight, as per Materials and Methods section 2.6.3, in 3ml of media so as to increase plasmid containing bacterial numbers. To isolate the plasmids from the bacteria the individual overnight cultures were treated using a Plasmid Miniprep Kit per Materials and Methods section 2.6.3 (Fig. 3.21).

Isolated plasmids suspensions were then screened for the presence of inserts by restriction digestion. The pCR8/GW/TOPO cloning vector contains EcoRI sites flanking the TOPO cloning site. Restriction digestion of the plasmid suspensions using EcoRI enables the detection of inserts. This was carried out as per Materials and Methods section 2.6.4. The digested plasmids are then run on a 2% agarose gel, as per Materials and Methods section 2.6.5, to allow the visualization of the presence of inserts. The identification of two bands on a gel, one at ~2800 bp and another between 300 and 1300 bp, would indicate the presence of the linearized plasmid and a DNA insert, while an individual band on the gel at ~2800 bp would represent linearized plasmid only (Fig. 3.21).
Fig. 3.21 Overview of the Process of Plasmid Isolation and Subsequent Plasmid Analysis

1. Pick Individual Colonies and grow overnight in Culture Media
2. Overnight Bacterial Culture
3. Isolation of Bacterial Plasmids using Mini-prep
4. Screening for the presence of Inserts using Restriction Digestion
5. Plasmid DNA Binds To Membrane
6. Analysis of Digested TOPO plasmids using Gel Electrophoresis
7. Plasmid Without an Insert
8. Molecular Weight Ladder
9. Plasmids Containing Insert

ACGAATTCTCCCCCTT
TGCTTAAGCCGGG

LAGGCCGAATTC
TTCCCCCTTAAG

ERE

EcoR I EcoR I EcoR I

AAGGGCGAATTC
TTCCCGCTTAAG

133
EcoRI Restriction digested plasmid DNA was electrophoresed on a 2% agarose gel. Restriction digestion of Topo plasmids containing inserts results in the presence of two bands (Fig. 3.22 lane V6), while single bands on the gel represent plasmids without ligated sequences (Fig. 3.22 lane V7). The efficiency of Topo cloning is evident from analysis of the gels as the vast majority of the plasmids contain inserts, while only a small quantity represent religated plasmids.

The isolated products ranged in size from ~200 bp to ~700 bp. This variation in sequence size is a good indication that the majority of the sequences isolated are unique and it is not just one clone that has been incorporated.
Fig. 3.22  EcoRI Restriction Digest analysis of pCR8 / GW /TOPO plasmids generated following ligation with Sequences Isolated from cells transfected with the ERα:Dam fusion Protein and treated with Ethanol

Fig. 3.23  EcoRI Restriction Digest analysis of pCR8 / GW /TOPO plasmids generated following ligation with Sequences Isolated from cells transfected with the ERα:Dam fusion Protein and treated with Oestradiol
Fig. 3.24 EcoRI Restriction Digest analysis of pCR8/GW/TOPO plasmids generated following ligation with Sequences Isolated from cells transfected with the ERα:Dam fusion Protein and treated with Tamoxifen.
3.3.3 Step 8 : Sequencing of Plasmids containing Inserts

Following screening using restriction digestion with EcoRI and subsequent digest visualization using a 2% agarose gel, it was known which plasmids contained inserts and which contained only religated vector by the presence or absence of an insert. All vectors containing an insert were sequenced to allow possible identification of the presence of any EREs. The primers used for sequencing were T7 primers, as the pCR8/GW/TOPO vector contains a T7 priming site (Fig. 2.5 Materials and Methods).

The resulting DNA sequences (Fig. 3.25) were blasted against the human genome to determine firstly whether the sequences were human or whether they were partial plasmid sequences. Partial plasmids sequences would result due to the fact that the ERα:Dam expression vector contains methylated GATC sequences as it is of bacterial origin, so following DpnI digestion these digested vector sequences could be amplified using the methylation specific based PCR method. All identified human sequences were then screened for the presence of ERα interacting sites and any other transcription factor binding sites that ERα is known to bind co-operatively. Genes located in the vicinity of the blasted human sequences were identified and a literature review of these genes was carried out to determine if they were known to be regulated by oestrogen.
The following is a selective sample of some of the sequences isolated and also the hormone treatment used in their isolation.

Ethanol 2 (V2)
GATCCTGGAGATTTGGCTGCTTGGTTGAAAAGTATGTATGGGTCGCCTCTGTTTTTTAAGACACCTCAGCCCTGGGC
ATGAGCTGCCAAACACAGTGCATGCCCCTCACCTGCTCAGGATAGGAGACTCCGCTGGACTGGCTTCCCTGTCTT
GTGAAATAGGAAGTTGCTGGAACAGTCAAGATCTACAGAATACCTGGGATACCTTACTCAAGACCTCCTGCTCTT
AGTGTGATTTTTAATATCACTGGTTTTTCTCTCCCTCCCTTCTCTCTCTGACCATGGATGTGATGGAGGC
TCACTACTGCTTTTCTTGGCCAGGACTACTACATCTATAGCAAAGACATGGGAAATACTGGGCAAGAAGGAGGAGG
AGCCCTTGGGAGAGCTGCTGGACTGGCTTCCCTGTCTT

Ethanol 17 (V17)
GATCGTTATCTGTGAAATGGGAAAATGCCTGCAACATCAAGCTTGCACTCTACGTGAGCAGTGACCCTACTATC
TTAGCTATGATAACTGGCTAGGTGAGTTCTGCACACTTGAGGCAGTCCTGGGAGAGCGGGTCCTGACCAGCACACTGGCCGCTGGCCGG
AAATGTCCAGGCACAGGCCGGGTCTGAGCTGAGGCTCGAAT

Oestradiol 12 (E12)
GATCGGACACCAGTGTTTCACTAGCCTGACACTTGGCAGGGTACCCTGCAACATCAAGCTTGCACTCTACGTGAGCAGTGACCCTACTATC
TTAGCTATGATAACTGGCTAGGTGAGTTCTGCACACTTGAGGCAGTCCTGGGAGAGCGGGTCCTGACCAGCACACTGGCCGCTGGCCGG
AAATGTCCAGGCACAGGCCGGGTCTGAGCTGAGGCTCGAAT

Ethanol 30 (V30)
GATCCTCAGCCTCAGGCCGCCCGCCCTCTCCCTACATCCAGACCGCCCCCCGCACAAATTACTCGCCTGGGATCTGCGCC
ACACCCCCCTCTTGCGAGGGCCGCCCGCCCGCCCGCCCGCTTTGATGTTGGCTTCTGCTCGCAACTGCCCCTTCTCTTTAT
GGAGATGTGGTGGAAAGCCCTGCGGCTTTCTGCTCGCAACTGCCCCTTTTAT

Oestradiol 30 (E30)
GATCGGCCCTCTCTTCAAGCCGGGGCCCGCTTTCCACATCCAGACCGCCCCGCAACAAATTACTCGCCTGGGATCTGCGCC
ACACCCCCCTCTTGCGAGGGCCGCCCGCCCGCCCGCCCGCTTTGATGTTGGCTTCTGCTCGCAACTGCCCCTTCTCTTTAT
GGAGATGTGGTGGAAAGCCCTGCGGCTTTCTGCTCGCAACTGCCCCTTTTAT

Tamoxifen 14 (T14)
GATCTGAGCCAGGAGGGCCCGCCCTGACATCTGGGTTGGCTGCTGAGGGCCCGCCCGCCCGCTTTGATGTTGGCTTCTGCTCGCAACTGCCCCTTCTCTTTAT
GGAGATGTGGTGGAAAGCCCTGCGGCTTTCTGCTCGCAACTGCCCCTTTTAT

Tamoxifen 32 (T32)
GATCTAGAAATCAACCTGTTGTTGGATATATCCCTTCTTATGATATGAGATGGAAG
TCTAGCTCAATACCTACGCTCTGTTGCTGGGCTCTGCTGACACTTGGGAAAGCCCTGCGGCTTTCTGCTCGCAACTGCCCCTTCTCTTTAT
GGAGATGTGGTGGAAAGCCCTGCGGCTTTCTGCTCGCAACTGCCCCTTTTAT

Fig. 3.25 Selection of DNA Sequences of Inserts of Human Origin
Homo sapiens chromosome 17 genomic contig, alternate assembly
(based on HuRef SCAF_1103279188349)
Length=1133206
Features in this part of subject sequence:
ubiquitin specific protease 22
Ethanol 17 chr6:125,138,507-125,138,822

Homo sapiens chromosome 6 genomic contig, reference assembly
Length=61645385

Features in this part of subject sequence:

T-cell lymphoma breakpoint-associated target 1

Score = 579 bits (313), Expect = 5e-162
Identities = 315/316 (99%), Gaps = 0/316 (0%)
Strand=Plus/Plus

Query 132
GATCGTTATCTGTGAAATGGGAAAATGCCTGCAACATCAAGCTTGCACTCTACGTGAGCA 191
Sbjct 29201243
GATCGTTATCTGTGAAATGGGAAAATGCCTGCAACATCAAGCTTGCACTCTACGTGAGCA 29201302

Query 192
GTGACCCTACTATCTTAGCTATGATAACACTGCAAGGAGCTAAACCATCAGTCAGGGTTA 251
Sbjct 29201303
GTGACCCTACTATCTTAGCTATGATAACACTGCAAGGAGCTAAACCATCAGTCAGGGTTA 29201362

Query 252
CTGCACTGGACAGCTTCAGAACCCTCTTGACACTGATTTAATTGAAATAATGATTTTTTT 311
Sbjct 29201363
CTGCACTGGACAGCTTCAGAACCCTCTTGACACTGATTTAATTGAAATAATGATTTTTTT 29201422

Query 312
ttCCTTTTACTATCCTGAACATTTCTGTTCTTTTAGAGAGGTCAAGCTACTGAAATACCA 371
Sbjct 29201423
TTCTTTTACTATCCTGAACATTTCTGTTCTTTTAGAGAGGTCAAGCTACTGAAATACCA 29201482

Query 372
ATAGTATCCAAAGGTTCCAAAGGGTCAAGGCCAAGGGCAGCAACAAAAGTTTGTCATTGT 431
Sbjct 29201483
ATAGTATCCAAAGGTTCCAAAGGGTCAAGGCCAAGGGCAGCAACAAAAGTTTGTCATTGT 29201542

Query 432
ACAGCATAAGAGGATC 447
Sbjct 29201543
ACAGCATAAGAGGATC 29201558
**Ethanol 30 chr22:47,585,403-47,585,591**

ref|NT_011525.7|Hs22_11682 Homo sapiens chromosome 22 genomic contig, reference assembly
Length=1384186

Features flanking this part of subject sequence:
- 53738 bp at 5' side: family with sequence similarity 19 (chemokine (C-C motif)...
- 91014 bp at 3' side: hypothetical protein

Score = 350 bits (189), Expect = 1e-93
Identities = 189/189 (100%), Gaps = 0/189 (0%)
Strand=Plus/Plus

<table>
<thead>
<tr>
<th>Query</th>
<th>Sbjct</th>
</tr>
</thead>
<tbody>
<tr>
<td>138</td>
<td>GATCCTCAGTCAGGGCCAGGGCTCTTCAGGTTGGAGGCTGTTTGAGGAGGGCAGC 197</td>
</tr>
<tr>
<td></td>
<td>GATCCTCAGTCAGGGCCAGGGCTCTTCAGGTTGGAGGCTGTTTGAGGAGGGCAGC 219212</td>
</tr>
<tr>
<td>198</td>
<td>TGGGAAGATGGTGGAGGGAGGTGATTCTGCACACTTGAGGCAGTCCTGGGAGAGCGGGTC 257</td>
</tr>
<tr>
<td></td>
<td>TGGGAAGATGGTGGAGGGAGGTGATTCTGCACACTTGAGGCAGTCCTGGGAGAGCGGGTC 219272</td>
</tr>
<tr>
<td>258</td>
<td>CTGACCAGCACTGGCCGCTGGCCGGAATGTCCAGGCACAGGCGGCTGACTGTA 317</td>
</tr>
<tr>
<td></td>
<td>CTGACCAGCACTGGCCGCTGGCCGGAATGTCCAGGCACAGGCGGCTGACTGTA 219332</td>
</tr>
<tr>
<td>318</td>
<td>GGCTCGAAT 326</td>
</tr>
<tr>
<td></td>
<td>GGCTCGAAT 219341</td>
</tr>
</tbody>
</table>
Oestradiol 12 chr7:98,647,450-98,647,781

Homo sapiens chromosome 7 genomic contig, reference assembly
Length=64426257

Features flanking this part of subject sequence:
68111 bp at 5' side: Smad ubiquitination regulatory factor 1 isoform 1
121131 bp at 3' side: actin related protein 2/3 complex subunit 1A

Score = 616 bits (333), Expect = 4e-173
Identities = 333/333 (100%), Gaps = 0/333 (0%)
Strand=Plus/Minus

| Query | 135 | GGATCGGACACCAGTGTGTTTCCTAGCCTGATCTCCCCAGTGCTGAGGCTGTCTTGCAATGTT | 194 |
| Sbjct | 24044122 | GGATCGGACACCAGTGTGTTTCCTAGCCTGATCTCCCCAGTGCTGAGGCTGTCTTGCAATGTT | 24044063 |
| Query | 195 | GTTCTCTTTCCCTCCCAGGTATACTGAGTTCTGATGTGTTGCTAGGAGTAAAGCTACCTC | 254 |
| Sbjct | 24044062 | GTTCTCTTTCCCTCCCAGGTATACTGAGTTCTGATGTGTTGCTAGGAGTAAAGCTACCTC | 24044003 |
| Query | 255 | GAGCTTGAAAGTTGTATTTGAGAAATCTAAAGGCTGAGATGCTTTAAGCCAACCTTGTTG | 314 |
| Sbjct | 24044002 | GAGCTTGAAAGTTGTATTTGAGAAATCTAAAGGCTGAGATGCTTTAAGCCAACCTTGTTG | 24043943 |
| Query | 315 | AGGGGCTCCCCCTGGGCTGGAACCAACAAAGGACTCTGAGTTCTGACCTGGCTCCCTTGCCACATT | 374 |
| Sbjct | 24043942 | AGGGGCTCCCCCTGGGCTGGAACCAACAAAGGACTCTGAGTTCTGACCTGGCTCCCTTGCCACATT | 24043883 |
| Query | 375 | GGTGGGGAGGAGGCTCAACATGACTCAACAPTGGATTTGAGAGCTGTCTGGAATGTTTAAACCA | 434 |
| Sbjct | 24043882 | GGTGGGGAGGAGGCTCAACATGACTCAACAPTGGATTTGAGAGCTGTCTGGAATGTTTAAACCA | 24043823 |
| Query | 435 | GGTGTCACCCTCCACCTCTCTCGCTGGCTTTGAGATC | 467 |
| Sbjct | 24043822 | GGTGTCACCCTCCACCTCTCTCGCTGGCTTTGAGATC | 24043790 |
Oestradiol 30

ref|NT_033899.7|Hs11_34054 Homo sapiens chromosome 11 genomic contig, reference assembly
Length=38509590

Features in this part of subject sequence:
roundabout, axon guidance receptor, homolog 3

Score = 713 bits (386), Expect = 0.0
Identities = 388/389 (99%), Gaps = 0/389 (0%)
Strand=Plus/Minus

<table>
<thead>
<tr>
<th>Query 1</th>
<th>GATCGCCTCTTTTCCAGGGGCCCAGCTTTCCATCCAGACCCCAGCAGAATTACTGC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sbjct 28309187</td>
<td>GATCGCCTCTTTTCCAGGGGCCCAGCTTTCCATCCAGACCCCAGCAGAATTACTGC</td>
</tr>
<tr>
<td>Score 60</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query 61</th>
<th>CTGGGATCCTGGCCACACCCCCTCTTGCCAGGCCGCCGCCGCCCTTTGAGTGGGGCTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sbjct 28309127</td>
<td>CTGGGATCCTGGCCACACCCCCTCTTGCCAGGCCGCCGCCGCCCTTTGAGTGGGGCTT</td>
</tr>
<tr>
<td>Score 120</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query 121</th>
<th>TCCGTCCTCGAAGTCACCTCCCTTTTCCATGAATTCAGGATGTTGCAAGGCTTCCTCGTGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sbjct 28309067</td>
<td>TCCGTCCTCGAAGTCACCTCCCTTTTCCATGAATTCAGGATGTTGCAAGGCTTCCTCGTGG</td>
</tr>
<tr>
<td>Score 180</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query 181</th>
<th>CGTCTTTAGAGGCAAGTATGCATCCAGATCCCGGCCCCTTTGCCGCGAGGCTTCTCCCTCTAGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sbjct 28309007</td>
<td>CGTCTTTAGAGGCAAGTATGCATCCAGATCCCGGCCCCTTTGCCGCGAGGCTTCTCCCTCTAGT</td>
</tr>
<tr>
<td>Score 240</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query 241</th>
<th>CTCGGCAGCTGTCCTGAGTCCGAGCTCCGAGGCTCCAGCTTTGCTTCCGAGGCTTCTCCCTCTAGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sbjct 28308947</td>
<td>CTCGGCAGCTGTCCTGAGTCCGAGCTCCGAGGCTCCAGCTTTGCTTCCGAGGCTTCTCCCTCTAGT</td>
</tr>
<tr>
<td>Score 300</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query 301</th>
<th>GTGCGACCCCTTCTCTCCGAGGTCCTGAGTCCGAGGCTCCAGCTTTGCTTCCGAGGCTTCTCCCTCTCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sbjct 28308887</td>
<td>GTGCGACCCCTTCTCTCCGAGGTCCTGAGTCCGAGGCTCCAGCTTTGCTTCCGAGGCTTCTCCCTCTCC</td>
</tr>
<tr>
<td>Score 360</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query 361</th>
<th>CCATGGGCCGCTGGTCCCTGCCGAGATC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sbjct 28308827</td>
<td>CCATGGGCCGCTGGTCCCTGCCGAGATC</td>
</tr>
<tr>
<td>Score 389</td>
<td></td>
</tr>
</tbody>
</table>
Tamoxifen 14 chr1:11,754,255-11,754,675

Homo sapiens chromosome 1 genomic contig, reference assembly
Length=7590365
Features in this part of subject sequence:
  hypothetical protein

Query  131  GAGGATCTGAGCCAGGAGGGCCCTGACATGGCGCTGCTCGAGGGGCACCCCCCA  190
  Sbjct  6369458 GAGGATCTGAGCCAGGAGGGCCCTGACATGGCGCTGCTCGAGGGGCACCCCCCA  6369399
Query  191  TCTCTGAGTAAGTGAAGTCTAGCTCCAATACTCCAGCTCTGGTGTGGGCCTGCAATCATGT  250
  Sbjct  6369398 TCTCTGAGTAAGTGAAGTCTAGCTCCAATACTCCAGCTCTGGTGTGGGCCTGCAATCATGT  6369339
Query  251  GTATCAGCCCCAGAatgcaggccgtgctctctctatgctgtgtctctctctaaacaagtcac  310
  Sbjct  6369338 GTATCAGCCCCAGAatgcaggccgtgctctctctatgctgtgtctctctctaaacaagtcac  6369279
Query  311  ttaacttatgtctgctcctagtctcccccatcgtaaaaatgggcacagttcctctctctggtgggtc  370
  Sbjct  6369278 TTAACCTTATCTGTGCCTCAGTTCCCCATCCGTAATAATGGGCACAGTTCCCTCTGGGGTC  6369219
Query  371  actgtgaggctcaaatgagATTGTCTCTGGGAATGTGCCAAAGCAGATGCAAAATAAACA  430
  Sbjct  6369218 ACTGTGAGGCTCAAATGAGATTGTCTCTGGGAATGTGCCAAAGCAGATGCAAAATAAACA  6369159
Query  431  TTCATTCGCCCAGGGTCCCCCGCCCCTGCCAGTGTGGAGCTCAGGCCCTGCGTGGCCAG  490
  Sbjct  6369158 TTCATTCGCCCAGGGTCCCCCGCCCCTGCCAGTGTGGAGCTCAGGCCCTGCGTGGCCAG  6369099
Query  491  GCTCCATCTCCCCGCTTTCCAGCACCCCCACGCAGTCACCCAGCAGTCCAGGCCCTGCGTGGCCAG  550
  Sbjct  6369098 GCTCCATCTCCCCGCTTTCCAGCACCCCCACGCAGTCACCCAGCAGTCCAGGCCCTGCGTGGCCAG  6369039
Query  551  GATC  554
  Sbjct  6369038 GATC  6369035
Tamoxifen 32 ref|NT_007592.14|Hs6 7749 Homo sapiens chromosome 6 genomic contig, reference assembly
Length=48945890

Features flanking this part of subject sequence:
973519 bp at 5' side: IBR domain containing 2
398745 bp at 3' side: inhibitor of DNA binding 4, dominant negative helix-loop...

Score = 780 bits (422), Expect = 0.0
Identities = 424/425 (99%), Gaps = 0/425 (0%)
Strand=Plus/Plus

Query 1  GATCTAGAATCAACCTGTGTTTGGATATATCCTTGTATATGAGTAAAAAGGCCACTAAGT  60
|    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 10297068 GATCTAGAATCAACCTGTGTTTGGATATATCCTTGTATATGAGTAAAAAGGCCACTAAGT  10297127

Query 61  GACCAACTGCTTGGCTCCAGTTAAGATAGGCTCTATAATGCCAGCTCTCAGCATCCACCTGCA  120
|    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 10297128 GACCAACTGCTTGGCTCCAGTTAAGATAGGCTCTATAATGCCAGCTCTCAGCATCCACCTGCA  10297187

Query 121  TAGCACACTCACGCTGGAGGCCCTGAGGTCCTAGTGCTGCTCAGCATCCACCTGCA  180
|    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 10297188 TAGCACACTCACGCTGGAGGCCCTGAGGTCCTAGTGCTGCTCAGCATCCACCTGCA  10297247

Query 181  CCTACACTCATGATAGTAGCATGCAGAGCTGCGCTGGGCCTCAGCCACAGCCAGAGG  240
|    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 10297248 CCTACACTCATGATAGTAGCATGCAGAGCTGCGCTGGGCCTCAGCCACAGCCAGAGG  10297307

Query 241  AATATTGGTTGCATGGTGAAATCGGAGGGCTTGAGAATATTGCCCTGCTTCCTTTGCTT  300
|    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 10297308 AATATTGGTTGCATGGTGAAATCGGAGGGCTTGAGAATATTGCCCTGCTTCCTTTGCTT  10297367

Query 301  GTTTTTCTGCAATTATTTTATCTGAGAGAGCATCCTCTGCTCAGCATCCACCTGCA  360
|    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 10297368 GTTTTTCTGCAATTATTTTATCTGAGAGAGCATCCTCTGCTCAGCATCCACCTGCA  10297427

Query 361  AGGAGGGGCCGCCACACACAGGCAGATGCTGGAGGCTGCTTTAAAAACCTCTGGTGTTGCAGCTG  420
|    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 10297428 AGGAGGGGCCGCCACACACAGGCAGATGCTGGAGGCTGCTTTAAAAACCTCTGGTGTTGCAGCTG  10297487

Query 421  TGATC  425
|    ||||
Sbjct 1029748  TGATC  10297492
The following is a table of genes adjacent to DamID identified ERα binding sites. A literature review was carried out to determine if these adjacent genes have been shown to be regulated by oestradiol.

<table>
<thead>
<tr>
<th>Nearest Gene</th>
<th>Oestradiol Regulated</th>
<th>Nearest Gene</th>
<th>Oestradiol Regulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC113828</td>
<td>?</td>
<td>HEAT Repeat Containing 4</td>
<td>?</td>
</tr>
<tr>
<td>Ubiquitin specific peptidase 22</td>
<td>?</td>
<td>UDP-Gal: betaGal beta 1,3-galactosyltransferase polypeptide 6</td>
<td>?</td>
</tr>
<tr>
<td>Signal Reconstitution Particle 14 kDa</td>
<td>?</td>
<td>Family With Sequence Similarity 132, Member A</td>
<td>?</td>
</tr>
<tr>
<td>Hormonally Upregulated Neu-Associated Kinase (Gardner et al, 2000)</td>
<td>Yes</td>
<td>Alpha 2 Globin</td>
<td>?</td>
</tr>
<tr>
<td>Chr 21 ORF 45</td>
<td>?</td>
<td>Smad Specific E3 Ubiquitin Protein Ligase 1 (Gielen et al, 2008)</td>
<td>Yes</td>
</tr>
<tr>
<td>Brain-Specific Angiogenesis Inhibitor 1 Precursor</td>
<td>?</td>
<td>Similar To mCG7611</td>
<td>?</td>
</tr>
<tr>
<td>Activity-Regulated Cytoskeleton-Associated Protein</td>
<td>?</td>
<td>Zinc Finger Protein 519</td>
<td>?</td>
</tr>
<tr>
<td>Testin</td>
<td>?</td>
<td>Mitogen-Activated Protein Kinase Organizer 1</td>
<td>?</td>
</tr>
<tr>
<td>CD180</td>
<td>?</td>
<td>Retinol Binding Protein 2, Cellular (Matsuda et al, 2004)</td>
<td>Yes</td>
</tr>
<tr>
<td>PI3KR1 (Spender et al, 2002)</td>
<td>Yes</td>
<td>LanC Antibiotic Synthetase Component C-like 2</td>
<td>?</td>
</tr>
<tr>
<td>Solute Carrier Family 12 (Potassium/Chloride Transporter)</td>
<td>?</td>
<td>Kalirin, RhoGEF Kinase Isoform 1</td>
<td>?</td>
</tr>
<tr>
<td>T-Cell Lymphoma Breakpoint Associated Target 1</td>
<td>?</td>
<td>TNFR2 (Deb, S et al, 2004)</td>
<td>Yes</td>
</tr>
<tr>
<td>FAM19A5</td>
<td>?</td>
<td>Zinc Finger Protein 179</td>
<td>?</td>
</tr>
<tr>
<td>Brain Abundant, Membrane Attached Signal Protein 1</td>
<td>?</td>
<td>Solute Carrier Family 47, Member 1</td>
<td>?</td>
</tr>
<tr>
<td>RoBo (Staton et al, 1998)</td>
<td>Yes</td>
<td>Similar to Acetyl-CoA Carboxylase 1</td>
<td>?</td>
</tr>
<tr>
<td>Interleukin 6 (Stein et al, 1997)</td>
<td>Yes</td>
<td>Jumonji Domain Containing 2C</td>
<td>?</td>
</tr>
<tr>
<td>Zinc Finger CCCH-Type Domain Containing 1</td>
<td>?</td>
<td>Karyopherin ?</td>
<td>?</td>
</tr>
</tbody>
</table>

Table 3.3 List of Genes in Close Proximity to the Isolated Sequences, and whether or not these Genes are Regulated by Oestradiol
The isolated sequences were analysed to detect the presence of ERα binding sites such as ERE full and half sites or other transcription factor binding sites to which ERα is known to bind co-operatively. These include SP-1 and members of the Ap-1 family, for example, c-Jun and ATF-2.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>ERα Half Site</th>
<th>SP-1</th>
<th>Ap-1</th>
<th>c-Jun</th>
<th>ATF-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethanol 2 (V2)</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Ethanol 17 (V17)</td>
<td>3</td>
<td>-</td>
<td>1</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>Ethanol 30 (V30)</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Oestradiol 12 (E12)</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>Oestradiol 30 (E30)</td>
<td>-</td>
<td>4</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Tamoxifen 14 (T14)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>Tamoxifen 32 (T32)</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

*Table 3.4* Isolated Sequences Containing Known ERα Receptor Binding Sites
3.4 Functional Analysis of Isolated Sequences

3.4.1 Step 9: Generation of a Reporter vector containing the isolated sequences

Following sequencing of the isolates, information was gathered on the origin of the sequence, human or plasmid, the DNA sequence of the isolates, whether they contained ERα binding sites and also any additional transcription factor binding sites. The next step was to determine the functional activity of these sequences. To do this the sequences would have to be cloned into a reporter vector. As pCR8/GW/TOPO (Fig. 2.5 Materials and Methods) is not a reporter vector, a LR recombination would have to take place between the pCR/GW/TOPO entry vector containing the cloned sequences and a destination vector, which has reporter capabilities. To do this a gateway vector conversion system with One-shot ccdB Survival competent was used to convert the reporter vector of choice into a gateway vector (Fig. 3.26), as per Materials and Methods section 2.8.1.

![Diagram of the newly constructed pGL3 Basic Destination vector with Cassette insert](image)

Fig. 3.26  Diagram of the newly constructed pGL3 Basic Destination vector with Cassette insert
Following successful generation of the pGL3-Basic destination vector (Fig. 3.26), a LR recombination was carried out as per Material and Methods section 2.9.1. The LR recombination reaction facilitates the transfer of the DNA inserts contained in the pCR/GW/TOPO entry vector into the newly generated pGL3-Basic reporter vector. Following the LR recombination reaction, the products were transformed into competent bacteria and a plasmid preparation was carried out on overnight cultures allowing the isolation of the pGL3-Basic destination vectors containing DNA inserts transferred from the pCR/GW/TOPO entry vectors.

3.4.2 Step 10 : Functional Analysis of the newly Generated Expression Vectors

The isolated sequences had now been transferred from the pCR8/GW/TOPO entry vector into the newly generated pGL3-Basic Luciferase reporter destination vector by using the Gateway cloning system. As the sequences were now in a reporter vector this would allow us to identify if the sequences were in fact able to bind ERα and thus cause transactivation of the luciferase gene.

To determine the functionality of the individual sequence containing reporter vectors, HeLa cells were co-transfected with individual pGL3-Basic luciferase reporter constructs and pCR3.1-hERα which is a wild type ERα expression vector. The cells were also treated with ethanol, oestradiol and Raloxifene for 24 hours as per Materials and Methods section 2.10.1. A luciferase assay was then carried on all transfected cells. Luciferase gene expression would occur if the isolated sequences contained ERα binding sites, thus allowing transactivation via ERα mediated recruitment of coactivators and the pol II transcription machinery. Different hormone treatments were
also administered to determine if ERα was able to differentially regulate gene expression.

A co-transfection of pCR3.1 hERα and the original pGL3-Basic vector was also carried out in HeLa cells. This is a control transfection as there should be no luciferase gene expression due to the fact that the pGL3-Basis does not contain a binding site for ERα to allow gene transactivation. A co-transfection of pERE-Elb-Luc and pCR3.1 hERα was also carried out as a positive control. The pERE-Elb-Luc reporter vector contains an ERE derived from the Xenopus vitellogenin A2 enhancer and therefore luciferase gene expression should occur following co-transfection with pCR3.1 hERα.
Fig. 3.27  Luciferase Assay Results following co-transfection of newly generated pGL3 Basic Luciferase Reporter Constructs, pGL3 Basic and pERE-E1b-Luc with pCR3.1 hERα followed by treatment with Ethanol, Oestradiol and Raloxifene (Luciferase assay performed by David Lonard on constructs)
The Luciferase Assay results in Fig. 3.27 show that the sequences that have been isolated by the DamID method do in fact bind ERα, either directly or indirectly. They allow gene transactivation to occur possibly via the recruitment of coactivators and the transcription machinery. This confirms that the ERα:Dam fusion protein was successful in binding to ERE or related sequences isolated following transfection.

The luciferase assay results of co-transfection of pGL3-Basic and WT-ERα show that the original pGL3-Basic vector is unable to facilitate luciferase gene transactivation. This confirms that, in the cases of the newly generated pGL3-Basic reporter vectors containing the isolated sequences, it is the inserted DNA sequences that allow luciferase transcription via the binding of ERα. The results of co-transfection of pERE-EIb-Luc with pCR3.1 hERα demonstrate that the transfection assays were carried out at optimal conditions because as expected luciferase gene expression occurred in this case following WT ERα binding to the ERE contained in the pERE-EIb-Luc reporter vector.

Also it is shown that, following co-transfection of the pGL3-Basic destination vectors with WT ERα, the expressed ERα is able to regulate gene expression differentially upon the binding of various ligands. The expression levels following ethanol treatment represent the basal levels of gene expression as it has been demonstrated that ERα is able to bind DNA in the absence of ligand activation and thus mediate gene expression. In the case of oestradiol treatment, gene expression increases up to a maximum of 14 fold for the pGL3-Basic destination vector containing isolated sequence T32, while the lowest increase in fold induction is ~4 for the pGL3-Basic destination vector containing isolated sequence E30. This is due to the fact that oestradiol is an agonist and responsible for increasing activation of ERα and its binding to DNA.
Raloxifene has partial agonist and antagonist activity. Raloxifene treatment has been shown to have an agonistic effect on the gene expression in the majority of cases. The highest levels of induction are associated with V37, which shows a 7-fold increase in gene expression over the basal levels of gene expression. Only in the case of E30 has Raloxifene been shown to have an antagonistic effect on gene expression as levels have fallen to ~0.9 of the basal level of gene expression.

The above luciferase assay results indicate that ERα is in fact able to bind the isolated sequences and cause gene transactivation. Also via binding of the isolated sequences, ERα can differentially regulate gene expression following treatment by various hormones.
3.4.3 Determination of Direct or indirect DNA binding

ERα is known to regulate gene transcription by direct DNA binding and also indirect DNA binding. ERα binds DNA directly via its response element, the ERE, and by recruitment of coactivators, the polymerase II transcriptional machinery can regulate target gene transcription. ERα can also regulate gene transcription without interacting directly with the DNA by co-operatively binding other transcription factors such as SP-1 and AP-1.

It has been shown above that ERα is able to bind the DamID isolated sequences, and thus regulate DNA transcription (Fig. 3.27); however, it is not known if ERα is interacting with the sequences directly via its DBD, or indirectly via co-operative binding of other transcription factors to cause gene transactivation. To determine if ERα binds the isolated sequences directly or indirectly a co-transfection of the pGl3-Basic destination vectors with the ERα DNA binding domain mutant C202H / C205H was carried out, as per Materials and Methods section 2.10.2, followed by a luciferase assay.

The ERα C202H / C205H expression vector encodes a hERα mutant containing the double mutation C202H / C205H within the DBD. This mutant possesses two mutations to two of the Zinc coordinating cystines in the first Zinc finger, more specifically the P-Box. This form of the receptor is transcriptionally inactive when assayed on an ERE-containing luciferase reporter, as it cannot bind to an ERE.
Fig. 3.28  Luciferase Assay Results following co-transfection of newly generated pGL3 Basic Luciferase Reporter Constructs, pGL3 Basic and pERE-E1b-Luc with the ERα DBD Mutant C202H / C205H expression vector followed by Ethanol, Oestradiol and Raloxifene hormone treatment. (Luciferase assay performed by David Lonard on constructs)
The Luciferase assay results in Fig. 3.28 indicate that the ERα DNA binding domain is necessary for ERα mediated gene expression to occur. In all the cases following treatment with oestradiol, ligand activation does not result in a major increase in fold expression of the luciferase gene; this indicates that even though the receptor is activated, the mutation in its DBD and more specifically the P-Box means it is unable to bind the isolated sequences and mediate target gene expression.

If ERα was binding co-operatively to another transcription factor to mediate gene expression, the DBD mutation may not cause such a reduction in fold induction as is shown here. This may be due to the fact that, in the case of ERα regulating gene expression by binding indirectly to the DNA, a mutation in the DBD may have little effect on this regulation.

These set of results indicate that ERα may be binding directly to these isolated sequences to mediate gene transcription, possibly through the recruitment of coactivators and the Pol II transcription machinery. These results are a further indication that the initial ERα:Dam fusion protein was able to bind ERα DNA binding sequences when expressed in vivo following transfection of the HeLa cells; this facilitated targeted methylation of GATC sequences by Dam following recruitment to ERα binding sites via its ERα fusion partner.
3.4.5 ERβ Mediated Gene Expression

Studies have shown that ERα and ERβ have a similar binding specificity to EREs. However, ERα binds to EREs with an approximately 2 fold greater affinity than ERβ. Despite their differences in affinity, ERα utilizes the same nucleotides as ERβ to bind to EREs. This similar binding specificity may be explained by the fact the C regions of ERα and ERβ, which contain the DBD of ERα and ERβ, are highly homologous, showing 96% similarity. The DBD consists of two zinc finger motifs, CI and CII, which are involved in DNA binding specificity, so a high degree of similarity in these regions would explain ERα and ERβ’s similarity in binding specificity.

To test the ability of ERβ to bind the DamID isolated ERα interacting DNA sequences a co-transfection of pCR3.1 flag hERβ and the pGL3-Basic destination vector constructs was carried out in HeLa cell, as per Materials and Methods section 2.10.3, followed by subsequent luciferase assays. Luciferase gene expression would occur if the over expressed ERβ was able to bind the DamID isolated sequences and hence recruit coactivators and the Pol II transcription machinery.

A co-transfection of pCR3.1 flag hERβ and the original pGL3-Basic vector was also carried out. This is a control transfection as there should be no luciferase gene expression due to the fact that the pGL3-Basis does not contain a binding site for ERβ to allow gene transactivation. A co-transfection of pERE-E1b-Luc and pCR3.1 hERβ was also carried out as a positive control. The pERE-E1b-Luc reporter vector contains an ERE derived from the Xenopus vitellogenin A2 enhancer; therefore, luciferase gene expression should occur following co-transfection with pCR3.1 hERβ.
**Fig. 3.29** Luciferase Assay Results following co-transfection of newly generated pGL3 Basic Luciferase Reporter Constructs, pGL3 Basic and pERE-E1b-Luc with the pCR3.1 flag ERβ expression vector followed by Ethanol, Oestradiol and Raloxifene hormone treatment (Luciferase assay performed by David Lonard on constructs)
The above luciferase expression assay results indicate that over expressed ERβ is able to bind the isolated sequences contained in the newly generated pGL3-Basic destination vectors and, via this binding, regulate gene expression. ERβ is also shown to differentially regulate expression following different hormone treatments. Treatment with ethanol represents the basal level of luciferase gene expression. Following treatment with the agonist oestradiol it is shown that there is a large increase in the fold induction of gene transcription, due to ERβ ligand activation.

The largest fold increase in gene expression (Fig. 3.29) is associated with T29 isolated sequence which displays a ~14 fold increase in gene expression compared to the basal level, while the lowest is associated with sequence E30 which displays a ~5 fold increase in gene expression compared to the basal level. Treatment with the partial agonist / antagonist Raloxifene shows an antagonist response in all cases with the most antagonist effects associated with sequences E30 which display a 2 fold decrease in gene expression compared to basal levels.

The luciferase assay results of co-transfection of pGL3-Basic and WT-ERβ show that the original pGL3-Basic vector is unable to facilitate luciferase gene transactivation. This proves that, in the cases of the newly generated pGL3-Basic reporter vectors containing the isolated sequences, it is the inserted DNA sequences that allow luciferase transcription via the binding of ERβ. The results of co-transfection of pERE-EIb-Luc with pCR3.1 flag hERβ act as a control and, as expected, luciferase gene expression occurred in this case following WT ERα binding to the ERE contained in the pERE-EIb-Luc reporter vector.
In general the level of gene expression associated with ERβ is much lower than that shown with ERα. There is roughly 2-3 fold increase in gene expression associated with ERα over ERβ. The largest increase in gene expression associated with ethanol treatment is with sequence E30 which has a > 4 fold increase in gene expression following ERα binding compared to ERβ binding. The largest increase in gene expression associated with oestradiol treatment is with sequence V37 which has a > 4 fold increase in gene expression following ERα binding compared to ERβ binding. In the case of sequence T14 the level of gene expression is the same following binding of ERα and ERβ.

The largest increase in gene expression associated with raloxifene treatment is with sequence V37, which has a 42-fold increase in gene expression following ERα binding compared to ERβ binding. This is due to raloxifene acting as an agonist when co-transfection of the pGL3-Basic destination vector containing sequence V37 with ERα is carried out, but raloxifene serves as an antagonist when co-transfection of ERβ occurs.

These results indicate that ERα and ERβ have similar binding specificity as both bound all sequences and initiated gene expression. Equally, ERα and ERβ can differentially regulate gene expression following their activation by various hormones, as both showed various gene expression levels following treatment with ethanol, oestradiol and Raloxifene. ERα displays a greater increase in expression levels compared to ERβ, roughly around 2-3 fold when treated with ethanol and oestradiol. Raloxifene also seems to act as an agonist in the majority of cases when associated with ERα mediate gene expression, while in the case of ERβ mediated gene expression it is an antagonist.
Discussion

The aim of this project was to generate a reporter library of genomic ERE cis elements that can be functionally evaluated for their ability to regulate ERα mediated gene expression. This would allow a greater understanding of ERα mediated transcription and gain an insight into some of the following questions:-

- Which method of isolating transcription factor binding, ChIP or DamID, is most efficient for isolation of ERα interacting sites?
- Which type of ERE is most commonly isolated, ERE full site, half site or full and half site?
- What adjacent transcription factor binding motifs are found co-associated with the isolated EREs?
- Do the levels of ERα and ERβ mediated transcription vary depending on ERE type?
- Does treatment with agonists or partial agonist/antagonists cause a variation in ERα mediated transcription following binding to isolated EREs?
- What is the difference in gene expression levels following ERE mediated binding of ERα and ERβ?
4.1 Current Research on ERα Binding Sites

ERs interact either directly with genomic targets encoded by the ERE (5’GGTCAnnnTGACC 3’) or indirectly by tethering to nuclear proteins such as AP1 and SP1 transcription factors. The methods of ER binding site specificity, however, are not clear since these binding sequence motifs are ubiquitous in the genome, and there is no discernable differences between functional and nonfunctional sites by computational modelling approaches. This ambiguity is likely due to a lack of systemic information on binding site usage and architecture and mechanistic complexity involving additional transcription factors and epigenetic modifications (Lin et al, 2007).

To obtain a global map of ERα binding sites in breast cancer cells Lin et al applied a ChIP-PET technique to generate a library of ERα binding sites in MCF-7 cells following oestrogen treatment and 1,234 high quality ERα binding sites were identified. These ERα binding sites identified by ChIP-PET were located in every chromosome of the human genome, apart from the Y chromosome, which is not present in MCF-7 cells derived from a female breast cancer patient. Scanning of the 1,234 binding regions indicated that 884 ~71 % binding regions contained at least one ERE-like sequence, 25 % were ERE half sites and the remaining 4% contained no ERE sequence motifs. The majority of the sites were located in distal or intragenic regions relative to the nearest regulated transcripts and only 5% were in 5 kDa of the transcription start site (Lin et al, 2007).

Caroll et al mapped the association of the ER with the complete non-repetitive sequence of human chromosomes 21 and 22 by ChIP with tiled microarrays and showed that
forkhead factor binding sites are present near authentic EREs significantly more frequently than those that do not bind ER. Their research also proved that a Forkhead factor (FoxA1) was essential for ER-chromatin interactions and subsequent expression of oestrogen gene targets. FoxA1 is thought to facilitate euchromatic conditions and as such it is possible that the presence of FOXA1 identifies specific regions within chromatin to facilitate the association of the ER transcription complex. Down regulation of FOXA1 inhibits the ability of ER to associate with its binding sites, confirming the requirement for Forkhead-directed association of ER with chromatin, despite the fact that these sites contain sufficient information in the form of an ERE for ER docking (Caroll et al, 2005).
4.2 ChIP Isolation of ERα Interacting Sequences

ChIP was the first method that was used to try and generate a reporter library of genomic ERE cis elements allowing functional evaluation of their ability to regulate ERα mediated gene expression. Using the ChIP technique only three DNA fragments were isolated. The ability of these ChIP isolated sequences to enable ERα binding and allow subsequent gene transactivation was tested and results indicated that only two of the sequences, no. 44 and no.7 might actually be genuine ERα interacting sequences. This was because the level of ERα mediated gene expression increased over the basal level of expression for both these sequences on treatment with the agonist oestradiol. For sequence no.7, no increase in gene expression occurred following treatment with oestradiol, indicating that ERα is in fact not binding to this sequence in vivo.

In our hands the ChIP assay was not efficient in the isolation of ERα interacting sequences as, despite repetition of the technique on numerous occasions, only three sequences were isolated, of which only two may actually bind ERα. This low efficiency might be because ChIP works better when using a high concentration of DNA so a larger reservoir of cells was required. Also the antibody used for the interaction step may not have been specific enough to allow the isolation of the ERα:DNA complexes.

To try and improve the efficiency of the ChIP assay a larger volume of cells was used, and modifications were made to the interaction protocol such as increasing the amount of antibody used and allowing the interaction reaction to incubate for longer. However, these effects did not improve the efficiency of the ChIP reaction. Another downfall of the ChIP reaction was that, following ligation of the interaction products with the vector
and subsequent transformation, a large number of colonies resulted, even in the negative control ligation reaction. The colonies in the negative control occurred due to vector religation, and despite dephosphorylation of the vector numerous times, the vector religation problem persisted. This increased the workload in screening of plasmids because the majority of plasmids did not contain inserts.
4.3 DamID Isolation of ERα Interacting Sequences

DamID was the second method used to generate a reporter library of genomic ERE cis elements allowing functional evaluation of their ability to regulate ERα mediated gene expression. A number of steps were involved in using DamID to isolate the ERα interacting sequences. Initially the ERα:Dam DNA fusion had to be generated. This was done via initial PCR amplification of the individual ERα and Dam PCR products, followed by a ligation PCR reaction, thereby generating the ERα:Dam fusion. The PCR product was then ligated into the pCR3.1 hERα expression vector to allow functional testing of the newly generated ERα:Dam fusion vector using a luciferase assay after transient transfection. The results indicated that the ERα:Dam expressed in vivo fusion protein was able to bind an ERE and enable gene transactivation as expression increased 3-4 fold when treated with the agonist oestradiol.

The next step involved the in vivo targeted methylation of GATC sequences in close proximity to ERα binding sites by transiently transfecting HeLa cells with the ERα:Dam expression protein. Following isolation of the HeLa cell genomic DNA, methylation specific PCR was carried out to amplify the methylated sequences. After amplification the sequences were cloned using pCR8/GW/TOPO and subsequently transformed into competent bacteria. Restriction analysis was carried out on all resulting colonies to identify the presence of inserts. All plasmids containing inserts were sequenced to identify the presence of ERα interacting sites such as EREs, Ap1 sites and Sp1 sites.

Sequencing results of DamID isolated sequences

The sequencing results showed that the DamID method did allow isolation of ERα sequences as a number contained ERE half sites which may allow ERα binding and Ap1
and Sp1 sites which have been shown to facilitate ERα binding through cooperative binding. Furthermore, the DamID isolated sequences contained GATC sequences at their 5 ends showing these PCR amplified sequences resulted from the in vivo methylation of GATC sequences by the ERα:Dam fusion protein.

The sequences were also blasted against the human genome to indicate what chromosome the sequences were from and also what genes they were located near. A review of current literature was carried out to determine if these genes are regulated by oestradiol/ERα. Certain sequences such as V2 and V17 were found to blast against intragenic regions of USP22 and T-Cell Lymphoma Breakpoint Target 1 respectively, while sequence E12 blasted to a site 5 kb from Karyopherin 7 and 70 kb from Smad Specific E3 Ubiquitin Protein Ligase 1, which has previously been shown to be regulated by oestrogen (Gielen et al, 2008).
4.4 Functional Analysis of Isolated Sequences

_DamID Isolated Sequences enable ERα mediated Gene Expression_

The next step involved the functional analysis of these isolated sequences to determine if, in fact, ERα was able to bind the sequences and initiate gene transactivation in vivo. The sequences were transferred from the pCR8/GW/TOPO expression vector into a pGL3 basic luciferase reporter vector using the gateway cloning system. They were then functionally tested following transfection of HeLa cells and subsequent luciferase assay. Luciferase assay demonstrated that wild type human ERα was binding to the DamID isolated sequences in the newly generated luciferase reporter vectors and initiating gene transactivation. The level of luciferase gene expression in RLU, following co-transfection of the pCR3.1 hERα expression vector and the newly constructed luciferase reporter vectors containing the DamID isolated sequences, was in most cases as high if not higher than the level of luciferase gene expression associated with the co-transfection of pCR3.1 hERα and the pERE-E1b-Luc reporter vector, which contains an ERE derived from the Xenopus vitellogenin A2 enhancer.

_ERα differentially Regulates Gene Expression following Ligand Binding_

These luciferase assay results showed that ERα was able to initiate gene expression following binding to the DamID isolated sequences; also that ERα was able to differently regulate the level of expression following the binding of various ligands. The cells had been treated with the hormones oestradiol, raloxifene as well as ethanol. The level of expression following treatment with ethanol represented the basal level of luciferase gene expression. Treatment of the cells with the ER agonist oestradiol, resulted in increased levels of gene expression for all isolated sequences, up to a maximum of 14 fold for the luciferase reporter vector containing isolated sequence T32,
while the lowest increase in fold induction was ~4 for the luciferase reporter vector containing isolated sequence E30.

Raloxifene has partial agonist / antagonist activity depending on the cell type. In the case of Raloxifene treated cells, gene expression increased from the basal level for all sequences, apart from the sequence containing the E30 DamID isolated sequence in which levels of gene expression were slightly lower than the basal levels of expression. The highest increase in fold induction of gene expression associated with cells treated with raloxifene over the basal level was following co-transfection with the luciferase vector containing isolated sequence v37. This displayed a 7 fold induction of gene expression over the basal level.

**ERα Gene expression is mediated via Direct Binding to DamID isolated Sequences**

A vital question in gaining an insight into how ERα regulates gene transactivation from the DamID isolated sequences is, whether it is through direct or indirect DNA binding? To determine if ERα was binding directly to the sequences, an ERα C202H / C205H expression vector which encodes a hERα mutant containing the double mutation C202H / C205H within the DBD which is transcriptionally inactive when assayed on an ERE-containing luciferase reporter, as it cannot bind to an ERE was transiently transfected with the pERE-E1b-LUC reporter vector, followed by a luciferase assay. The average RLU value associated with co-transfection of the DamID isolated sequences with wild type ERα and treatment with oestradiol, was 305,000 while the average RLU value following co-transfection of the ERα DBD mutant with the same sequences was 23,000 which represents ~13 fold reduction in gene expression. This clearly shows the DBD of ERα is required to drive gene expression from the DamID isolated sequences.
The Luciferase assay results indicated that ERα mediates gene transactivation through direct binding to the isolated sequences and not through cooperative binding to another transcription factor. This is due to the fact that the ERα C202H / C205H mutant expression vector possesses two mutations to two of the Zinc coordinating cystines in the first Zinc finger, more specifically the P-Box, which alters ERα’s ability to directly bind DNA. This is further evidence that the ERα:Dam fusion protein was able to bind directly to these ERα DNA interaction sites and methylate nearby GATC sequences.

**DamID isolated Sequences enable ERβ Mediated Gene Expression**

Studies have shown that ERα and ERβ have a similar binding specificity to EREs. However, ERα binds to EREs with an approximately 2 fold greater affinity than ERβ. The next question was whether ERβ was able to bind to the DamID isolated ERα responsive elements and mediate gene transactivation. A co-transfection of pCR3.1 hERβ and pERE-E1b-Luc was carried out followed by luciferase assay. Luciferase assay results showed that ERβ was able to bind the DamID isolated sequences and mediate gene expression. This is expected as ERα and ERβ have similar binding specificity, due to the high homology in the C regions, which contains the DBD so ERα responsive sequences which were also responsive to ERβ.

**ERβ differentially Regulates Gene Expression following Ligand Binding**

The luciferase assay results, following co-transfection of ERβ, also demonstrated the ERβ differentially regulated gene expression after treatment with the hormones
oestradiol, raloxifene, and ethanol. Treatment with ethanol represented the basal level of gene expression associated with ERβ. Upon treatment with the agonist oestradiol, levels of gene expression increases for all DamID isolated sequences, up to a maximum of 14 fold increase in gene expression for the T27 isolated sequence, and the lowest a 5 fold increase for the E30 sequence. Following treatment with the partial agonist/antagonist raloxifene the level of gene expression was reduced compared to the basal level in all cases, with the most agonistic results associated with sequence E30 showing a two fold reduction in levels of gene expression compared to basal levels of expression.

*Differences in ERα and ERβ mediated Gene Expression*

When comparing the degree of ERα and ERβ mediated luciferase gene expression, as measured in RLU, it was found that ERα had on average a 2-3 fold increase in expression levels in cells treated with oestradiol compared to that of ERβ. The level of ERα mediated gene expression has already been shown to be greater than that of ERβ mediated gene expression (Tremblay et al, 1997). This may be due to a two fold higher binding affinity of ERα to an ERE over ERβ. Because transcription is a dynamic process, ERα may initiate more cycles of gene expression by recruitment of co-factors and the transcription complex, as it is bound to the DNA longer than ERβ due to its higher affinity.

It has also been shown that ERα and ERβ may regulate levels of expression differently upon ligand binding. Following ERα co-transfection with the DamID sequences in the luciferase reporter vectors and treatment with the partial agonist/antagonist raloxifene, it was shown that for all sequences, apart from the E30 isolated sequence, raloxifene...
increased gene expression levels compared to the basal levels of expression, thus having an agonistic effect. In the case of ERβ, co-transfection with the DamID sequences in the generated luciferase reporters, raloxifene has an antagonistic effect in all cases as levels of gene expression are reduced compared to the basal levels of expression. One possible mechanism by which raloxifene has an agonistic effect upon binding to ERα is by the subsequent recruitment of coactivators, while raloxifene binding to ERβ may cause recruitment of corepressors, resulting in the differences in gene expression levels resulting from raloxifene binding ERα and ERβ.
4.5 ChIP versus DamID

DamID has been the most useful method for the isolation of ERα interacting sequences and subsequent generation of a reporter library. The advantages of DamID over ChIP, and vice versa, for the isolation of the ERα interacting sequences were:-

1) DamID was not dependent on the availability of high quality antibodies to ERα.

ChIP relies on a high-quality antibody against ERα and this may have led to an insufficient yield of ERα interacting sequences following the interaction reaction, as the antibody may not have bound to all ERα-DNA complexes. Also if the antibody was not specific it may have resulted in cross-reactivity, which may be the reason sequence no.7 above was isolated using the ChIP technique. As DamID is not dependent on antibodies, this ruled out cross-reactivity.

2) DamID was performed on cells in a single well of a six well plate, while the ChIP experiments required a whole plate of cells. This allowed multiple DamID experiments to be set up on the six well plates, while the ChIP required multiple plates, enabling a saving on cost and time. The ChIP experiment did not work well; this may have been due to a lower concentration of cells than was required, indicating that even the plate of cells that ChIP was carried out on may not have been sufficient.

3) ChIP can be performed on the endogenous ERα protein while DamID required the generation of an exogenous fusion protein. This exogenous fusion may sometimes result in a protein losing its genomic binding specificity. In the case of the ERα:Dam fusion protein this specificity was retained as, when in-vivo
expressed with the pERE-E1b-LUC reporter vector, the ERα:Dam fusion was shown not only to bind the ERE and mediate gene transactivation, but to also differentially regulate this gene expression following the binding of various ligands. This meant that, in this instance, the exogenous fusion protein did not prove a disadvantage in the isolation of the genomic ERα interacting sequences.

4) DamID requires the expression of the fusion protein at low levels which is a disadvantage. This is because Dam is a highly active enzyme and expression of the fusion protein, even at moderate amounts, would cause nontargeted methylation. To reduce nontargeted DNA methylation a serious of co-transfections of the ERα:Dam fusion protein at various concentrations with the pERE-E1b-Luc reporter vector was carried out. The lowest concentration of ERα:Dam at which the fusion protein was still mediate gene transactivation and also differentially regulate this expression following the binding of various ligands was chosen as the concentration of fusion protein that would be used for the subsequent DamID experiments.
4.6 Future Work

Currently only a small reporter library of genomic ERE cis elements has been generated. Future work includes the PCR screening of more of the cloned DamID isolated sequences contained in the TOPO vectors as well as the transfer of sequences of human origin into a luciferase reporter vector using the gateway cloning system. This would result in a much larger reporter library of ERα interacting sequences. Sequencing of the larger library would allow a statistically greater determination of the type of ERE present and also the type of adjacent transcription factor binding sites present. Functional analysis on this larger ERα interacting library would enable determination of whether the type of ERE and additional transcription factor binding sites had influences on the level of gene expression.

By using this larger library of ERα interacting sequences, additional information could be gained on the difference in levels of ERα and ERβ mediated gene expression. Further analysis would allow a greater insight into ERα and ERβ mediated gene expression following treatment with a variety of ligands including tamoxifen and ICI 182,780.

Further information on ERα mediated gene expression, following binding to the DamID sequences, could also be gained by co-transfection assays with various types of ERα mutants such as ERα AF-1 and ERα AF-2 mutants. This would provide a better insight into how important these domains are in ERα mediated gene expression since levels of expression could be compared between the wild type and mutant proteins.
In addition, an insight could be gleaned into whether or not EREs facilitate gene expression in a distinctly different manner in cancer and normal cells. This could be achieved by carrying out transient transfections with the reporter vectors containing the ERα interacting sequences, followed by luciferase assay on a variety of cell lines. This would enable comparison of gene expression in, not only normal and cancerous cells, but also between different cancer cell lines. For example, differences in gene expression may result due to elevated coactivator levels in certain cell lines.
Addendum

5.1 Estrogen Receptor Knockout Mice

ER knock-out mice have provided invaluable evidence for the biological functions of ERα and ERβ.

5.1.1 Reproductive Phenotypes of ER knockout Models

Oestrogen has many roles in reproduction and the generation of the ERα and ERβ knockout (αERKO and βERKO) mice has further illustrated its roles and mechanisms. Both sexes of the αERKO are infertile, whereas, only the βERKO female shows impaired infertility. Infertility in the male αERKO mice is due to deficits at several points in the reproductive process, including severe reduction in sperm numbers and lack of sperm function. The seminiferous tubules of the αERKO testes show progressive dilation that is accompanied by degeneration of the seminiferous epithelium. In contrast, the testes of the βERKO appear normal and produce sufficient functional sperm to allow fertility (Hewitt et al, 2000).

The infertility of the female αERKO mouse is due in part to the insensitivity of the uterus to the mitogenic and differentiative actions of oestrogen. All expected tissues are present but appear immature. ERα is thus not necessary for the development of the uterus, but is necessary for complete maturation and function of the tissue. Wild type ERβ and βERKO uteri are indistinguishable and show normal organisation and development of the stromal, myometrial and epithelial layers, as well as glandular structures. It can be concluded that ERβ is, apparently, not required for normal development of the female reproductive tract (Hewitt et al, 2000).
The αERKO female does not ovulate, while the βERKO is subfertile with reduced litter numbers and smaller litter sizes compared with wild type littermates. The αERKO female has an enlarged hemorrhagic cystic ovary. This phenotype begins to develop progressively as the animal matures and is apparently caused by a lack of oestradiol feedback inhibition in the pituitary, which results in chronically elevated LH and subsequent hyperstimulation of the ovary (Hewitt et al, 2000).

5.1.2 Mammary Gland

The female mammary gland is immature at birth and consists of a mainly stromal tissue, with only a rudimentary epithelial duct structure emanating from the nipple. The duct elongates in response to ovarian and pituitary hormones, eventually filling the stromal tissue with a branched like tree structure. Lobular alveolar buds develop along the length of these ducts during pregnancy and differentiate into secretory lactational structures (Hewitt et al, 2000). ERα is highly expressed in the adult mouse mammary gland, whereas ERβ is only slightly detected. The αERKO mammary gland exhibits normal prenatal and prepubertal development, but remains rudimentary after puberty, lacking the epithelial branching and lobuloalveolar development evident in wild-type glands. βERKO females possess normal ductal structure, and the entire fat pad is filled as with wild type females. After pregnancy βERKO mammary glands differentiate normally and reveal the lobulated structures necessary for lactation (Walker et al, 2004).
5.1.3 Skeletal System

Oestrogens are important for bone metabolism and homeostasis. Oestradiol has been shown to be crucial for the maintenance of bone mass. Femurs in αERKO and αβERKO mice are shorter than in wild type mice but not in βERKO mice. Female αERKO mice have smaller bone diameters and males have a lower bone density. βERKO females have been reported to have an increase in bone mineral density content and no changes have been observed in the βERKO males. These observations may be due to a negative effect of ERβ on bone density in the female mouse (Walker et al, 2004).

5.2 Updated Research on ER Binding Sites

Liu et al used a ChIP on ChIP approach to allow a large scale identification of ERα and ERβ binding DNA regions in intact chromatin. Results indicated a high degree of overlap between DNA regions bound by ERα and ERβ, but there were also regions selectively bound by ERα in the presence of ERβ, as well as regions that were selectively bound by either receptor. The regions bound by ERα had distinct prosperities in terms of genomic landscape, sequence features and conservation compared with regions that were bound by ERβ. ERβ bound sequences were located closer to transcription start sites compared to those bound by ERα. This suggests that a higher number of ERβs are proximal transcription factors (Liu et al, 2008).

GC-rich patterns, such as SP-1, were found more often in regions binding ERβ, whereas AT rich patterns such as forkhead transcription factor binding sites were predicted more often in areas preferentially bound by ERα. As shown previously by Carroll et al, 2005, the ChIP experiments carried out by Liu et al suggested that ERα bound regions
selectively bind forkhead motifs, as compared to ERβ bound regions. A number of regions were also found to bind ERα only in the presence of ERβ. This would suggest that in such regions ERβ recruits ERα by direct or indirect interactions, by making the region available to ERα by affecting chromatin state (Liu et al, 2008).
References


Stein, B and Yang, M. Repression of the Interleukin 6 Promoter by Estrogen Receptor is Mediated by NFκB and C/EBPβ. Molecular and Cellular Biology, 4971-4979, 1995.


http://www.cellscience.com/reviews4/Molecular_Basis_Steroid_Action_Prostate.html

www.invitrogen.com

http://medchem.rutgers.edu

www.asus.net.pl

http://www.worldofmolecules.com/emotions/estrogen.htm

http://chemistry.about.com

http://www.pharmacology2000.com

http://www.chemblink.com

http://www.marquette.edu

http://www.activemotif.com