Functional Characterization of Nuclear Receptor and Co-activator Binding Loci in the Human Genome

Kim Maguire
Technological University Dublin

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

Part of the Medical Sciences 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
Functional Characterization of Nuclear Receptor and Co-activator Binding Loci in the Human Genome

By

Kim Ruth Maguire BSc

MPhil Biomedical Science

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

In conjunction with FAS Science Challenge Internship programme at Baylor College of Medicine, Houston, Texas.
“Somewhere, something incredible is waiting to be known.”

Dr Carl Sagan
ABSTRACT

Steroid hormones, such as oestrogen, mediate their effects via activation of oestrogen regulated genes using nuclear receptors. Selective oestrogen receptor modulators (SERMs), such as Tamoxifen, are used to treat oestrogen responsive breast cancers, functioning to act as oestrogen antagonists, preventing the oestrogen receptor binding DNA and blocking gene expression. However, Tamoxifen has been identified as an oestrogen agonist in other tissues which can often lead to secondary tumors in the years following the treatment. Identification of the genomic regions where SERMs can act as oestrogen agonists can possibly lead to the development of gene targeted therapies or other alternatives to prevent this from occurring.

Investigations of a previously constructed ERE reporter library were directed towards the isolation of SERM responsive plasmids. Identified SERM responsive plasmids were compared to oestrogen responsive reporter plasmids on the basis of their ability to mediate transcription in conjunction with other nuclear receptors and coactivators to establish potential novel interactions. ERα mutation analysis was performed to attempt to identify the mechanism by which these sequences have the ability to be activated with the addition of SERMs. It was found that SRC-3, a member of the SRC family of coactivators, had the ability to corepress SERMs via oestrogen receptors, a response not documented previously.

Further to results obtained from the investigations of the SERM responsive reporter plasmids, DamID technology was employed to attempt to construct two libraries of sequences that associated, indirectly, with SRC-1A or SRC-3, via ligand bound oestrogen or progesterone. Gateway technology was used to transfer cloned sequences.
in to reporter plasmids to enable the luciferase assay to be employed in order to determine the functionality of the isolated sequences.

Transactivation assays suggested that the attempt to direct the SRC-1A:Dam fusion protein to bind to DNA via either ligand bound oestrogen or progesterone failed, however following reporter plasmid sequencing bioinformatical studies were performed and a number of possible transcription factor binding motifs were identified in the sequences. Further transactivation assays indicated SRC-1A:Dam had, in a subset of the reporter plasmids, bound via the orphan nuclear receptors RORα and COUP TFI.

Results suggest that SRC-1A may mediate this association and coactivate gene transcription; the mechanism behind this novel finding remains to be elucidated.
DECLARATION

I certify that this thesis which I now submit for examination for the award of MPhil (Biomedical Science), 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 in 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  Kuni Maguire  Date  5th May 2010
Primarily, I would like to take this opportunity to express my sincere gratitude to my supervisors. To Dr. Bert O’Malley for giving me the opportunity become a member of his team and to carry out research in the Molecular and Cellular Biology Department of Baylor College of Medicine, Houston, Texas, and also for sharing all of his wisdom. To Dr. David Lonard for his patience, encouragement and guidance throughout my time spent in Texas. To Dr. Alice McEvoy her invaluable guidance, support, time and patience given to me throughout the writing process. And finally, to Dr. Fergus Ryan for all his encouragement and enthusiasm for the FAS Science Challenge, enabling DIT students to avail of such a prestigious and valuable programme, and for support throughout the process.

I am also especially thankful to my laboratory colleagues, Dr. Yung Yu, Dr. Steve Settle, Dr. Vladimir Stanisic and Dr Amber Johnson for all the great chats and for all the support throughout my research. I wish them all the best with their future research.

I would like to give a huge thank you to Dr. Pauline Ward, my surrogate mother while in Texas! Thank you for everything, I cannot begin to list all of the things you did.

I would like to thank FAS for giving me the opportunity to study in Baylor College of Medicine through the FAS Science Challenge programme, and especially to John Cahill, Manager of the FAS Science Challenge programme, Grainne Timlin, Dr. Denis Headon and Dr. Austin Cooney and all who had helped to set up this great programme for Irish students providing us the chance to study in world class institutions. It is a shame that this programme has been discontinued.

Finally, I would like to give a special thank you to all my friends and family who supported me over the year. Especially to Mum, Mike, Robin and Brendan.
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>ACTR</td>
<td>activator of thyroid receptor</td>
</tr>
<tr>
<td>AF-1/2/3</td>
<td>activation function 1/2/3</td>
</tr>
<tr>
<td>AIB-1</td>
<td>amplified in breast cancer-1</td>
</tr>
<tr>
<td>AP-1/2</td>
<td>activation protein-1/2</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>BCM</td>
<td>baylor college of medicine</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>cAMP</td>
<td>cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CARM-1</td>
<td>Coactivator-associated arginine methyltransferase 1</td>
</tr>
<tr>
<td>CBP</td>
<td>CREB-binding protein</td>
</tr>
<tr>
<td>cDNA</td>
<td>complimentary DNA</td>
</tr>
<tr>
<td>ChIP</td>
<td>chromosome immuno-precipitation</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>COUP TF-1/2</td>
<td>chicken ovalbumin upstream promoter transcription factor-1/2</td>
</tr>
<tr>
<td>CREB</td>
<td>cAMP response element binding proteins</td>
</tr>
<tr>
<td>DamID</td>
<td>DNA adenine methylation identification</td>
</tr>
<tr>
<td>DBD</td>
<td>DNA binding domain</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxy ribonucleic acid</td>
</tr>
<tr>
<td>DMEM</td>
<td>dulbecco's modified eagle medium</td>
</tr>
<tr>
<td>EtOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>ER</td>
<td>oestrogen receptor</td>
</tr>
<tr>
<td>ERs</td>
<td>oestrogen receptors</td>
</tr>
</tbody>
</table>
ERα  oestrogen receptor-alpha
ER β  oestrogen receptor-beta
ERE  oestrogen responsive elements
E1    oestrone
E2    oestrodiol
E3    oestriol
FSH   follicle stimulating hormone
gDNA  genomic DNA
GPR30 G protein coupled receptor
Grip1 glucocorticoid receptor interacting protein 1
GR    glucocorticoid receptor
HAT   histone acetyltransferase
HBD   hormone binding domain
HRE   hormone responsive elements
HSP   heat shock protein
HGP   human genome project
IGF-1 insulin like growth factor-1
IL-6/7 interleukin-6/7
kDa   kilo Dalton
LB    lisogeny broth
LBD   Ligand binding domain
LH    luteinizing hormone
LM-PCR ligand mediated PCR
Luc   luciferase
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPK</td>
<td>mitogen activated protein kinase</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NCBI</td>
<td>national centre for biotechnology information</td>
</tr>
<tr>
<td>NCoA-2</td>
<td>nuclear receptor coactivator 2</td>
</tr>
<tr>
<td>NEB</td>
<td>new england biolabs</td>
</tr>
<tr>
<td>NR</td>
<td>nuclear receptor</td>
</tr>
<tr>
<td>PAX4/6</td>
<td>paired box 4/6</td>
</tr>
<tr>
<td>PCIP</td>
<td>p300/CBP cointegrator associated protein</td>
</tr>
<tr>
<td>PCG-1α</td>
<td>peroxisome proliferator-activated receptor gamma coactivator-1 alpha</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PIP</td>
<td>phosphatidylinositol 3,4,5 triphosphate</td>
</tr>
<tr>
<td>PR</td>
<td>progesterone receptor</td>
</tr>
<tr>
<td>PR-A</td>
<td>progesterone receptor-A</td>
</tr>
<tr>
<td>PR-B</td>
<td>progesterone receptor-B</td>
</tr>
<tr>
<td>PRE</td>
<td>progesterone responsive element</td>
</tr>
<tr>
<td>RAC-3</td>
<td>receptor-associated coactivator 3</td>
</tr>
<tr>
<td>Ral</td>
<td>raloxifene</td>
</tr>
<tr>
<td>RARα</td>
<td>retinoic acid receptor-alpha</td>
</tr>
<tr>
<td>RNA</td>
<td>ribo nucleic acid</td>
</tr>
<tr>
<td>RORα</td>
<td>retinoic-related orphan receptor-alpha</td>
</tr>
<tr>
<td>RT</td>
<td>room temperature</td>
</tr>
<tr>
<td>RXR</td>
<td>retinoid x receptor</td>
</tr>
<tr>
<td>SCA1</td>
<td>spinocerebellar ataxia type 1</td>
</tr>
<tr>
<td>SERM</td>
<td>selective oestrogen receptor modulator</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SF-1</td>
<td>steroidogenic factor 1</td>
</tr>
<tr>
<td>SNPs</td>
<td>small nuclear polymorphisms</td>
</tr>
<tr>
<td>SRC-1/2/3</td>
<td>steroid receptor coactivator 1/2/3</td>
</tr>
<tr>
<td>TIF2</td>
<td>transcription intermediary factor 2</td>
</tr>
<tr>
<td>TNFα</td>
<td>tumor necrosis factor alpha</td>
</tr>
<tr>
<td>TR</td>
<td>thyroid receptor</td>
</tr>
<tr>
<td>TRAM-1</td>
<td>thyroid receptor activator molecule 1</td>
</tr>
<tr>
<td>UV</td>
<td>ultra violet</td>
</tr>
<tr>
<td>17β HSD</td>
<td>17 beta hydroxysteroid dehydrogenase</td>
</tr>
<tr>
<td>4HT</td>
<td>tamoxifen</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

1.0. INTRODUCTION ............................................................................................................. 16

1.1. Gene Expression ........................................................................................................... 16

1.2. Steroid Hormones ....................................................................................................... 16

1.3. Nuclear Receptors ...................................................................................................... 23

1.4. Coregulators: Coactivators and Corepressors .......................................................... 44

1.5. Steroids, Nuclear Receptors, Steroid Receptor Coactivators and Cancer ................. 54

1.6. Identification of Transcription Factor Binding Sites .................................................. 62

2.0. AIMS ............................................................................................................................... 69

3.0. MATERIALS AND METHODS ..................................................................................... 72

3.1. Investigation of ERE (oestrogen responsive element) Reporter Library .................. 72

3.2. Construction and Investigation of Library of sequences which interact indirectly to SRC-1A and SRC-3 .......................................................... 76

3.3. Locating Transcription Start Sites .............................................................................. 108

3.4. Statistical Analysis ...................................................................................................... 116

4.0. RESULTS ......................................................................................................................... 117

4.1. ERE Reporter Library Investigation ......................................................................... 117

4.2. SRC-1A and SRC-3/ERα and PR-B Reporter Library Generation and Investigation .. 129

4.3. Locating Transcription Start Sites in Reporter Plasmids ............................................ 180

5.0. DISCUSSION .................................................................................................................... 182

5.1. Oestrogen Responsive Element (ERE) Reporter Library Investigation .................. 183

5.2. SRC-1A and SRC-3 Library Construction and Investigation .................................... 185

5.3. Locating Transcription Start Sites .............................................................................. 195

5.4. Future Work ................................................................................................................ 195

6.0. REFERENCES .................................................................................................................. 197

APPENDIX I ........................................................................................................................ 211

APPENDIX II ...................................................................................................................... 216

APPENDIX III .................................................................................................................... 220

APPENDIX IV ...................................................................................................................... 222
**List of Figures**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Fluctuating levels of estrogen and progesterone</td>
</tr>
<tr>
<td>1.2</td>
<td>Molecular structure of Oestrone, Oestradiol and Oestriol</td>
</tr>
<tr>
<td>1.3</td>
<td>Molecular structure of Progesterone</td>
</tr>
<tr>
<td>1.4</td>
<td>Synthesis of Oestrogen</td>
</tr>
<tr>
<td>1.5</td>
<td>Illustration Outlining Oestradiol Synthesis</td>
</tr>
<tr>
<td>1.6</td>
<td>Synthesis of Oestrogen and progesterone</td>
</tr>
<tr>
<td>1.7</td>
<td>The activation of Steroid Receptors</td>
</tr>
<tr>
<td>1.8</td>
<td>Representation of human ERα and β and their respective isoforms</td>
</tr>
<tr>
<td>1.9</td>
<td>Overall distribution of ER in different tissues</td>
</tr>
<tr>
<td>1.10</td>
<td>Functional Domains of Human ERα</td>
</tr>
<tr>
<td>1.11</td>
<td>Physiological Role of Oestrogen Receptors</td>
</tr>
<tr>
<td>1.12</td>
<td>Sequence similarities between human PR-A and PR-B</td>
</tr>
<tr>
<td>1.13</td>
<td>Isoforms of ROR</td>
</tr>
<tr>
<td>1.14</td>
<td>Physiological functions of ROR</td>
</tr>
<tr>
<td>1.15</td>
<td>Selected Coregulator Posttranslational Modifications</td>
</tr>
<tr>
<td>1.16</td>
<td>Common structural and functional domains of SRC family</td>
</tr>
<tr>
<td>1.17</td>
<td>Multiple members of the SRC family</td>
</tr>
<tr>
<td>1.18</td>
<td>Signaling to SRCs</td>
</tr>
<tr>
<td>1.19</td>
<td>Chemical Structure of Tamoxifen</td>
</tr>
<tr>
<td>1.20</td>
<td>Chemical Structure of Raloxifene</td>
</tr>
<tr>
<td>1.21</td>
<td>Step by step methods for performing Chip-on-Chip analysis</td>
</tr>
<tr>
<td>1.22</td>
<td>Methylation of GATC sequences</td>
</tr>
<tr>
<td>1.23</td>
<td>MEME and TOMTOM Results</td>
</tr>
<tr>
<td>3.1</td>
<td>Transfection protocol: Plate organisation – Test plate</td>
</tr>
<tr>
<td>3.2</td>
<td>Transfection protocol: Plate organisation Control plate</td>
</tr>
<tr>
<td>3.3</td>
<td>Hormone and EtOH Dilutions</td>
</tr>
<tr>
<td>3.4</td>
<td>Forward and Reverse Primers for Dam PCR</td>
</tr>
<tr>
<td>3.5</td>
<td>Vector map of pCR3.1</td>
</tr>
<tr>
<td>3.6</td>
<td>ChargeSwitch Pro Plasmid Mini Prep Kit Protocol</td>
</tr>
<tr>
<td>3.7</td>
<td>HeLa Cell Transfection Layout</td>
</tr>
<tr>
<td>3.8</td>
<td>Formation of double stranded adapter and its ligation</td>
</tr>
<tr>
<td>3.9</td>
<td>Double stranded adapters sequences</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1.1. Consequences of SRC knockout studies .......................................................... 59
Table 3.1. PCR Programme Details for amplification of Dam from E. coli gDNA. ..... 78
Table 3.2. Plasmid Premix Preparation ........................................................................... 90
Table 3.3. Addition of Hormones to the transfected HeLa cells .................................... 92
Table 3.4. PCR Programme for amplification of methylated sequences ....................... 97
Table 3.5. PCR Programme for Amplification of cDNA .............................................. 114
Table 4.1. Observation of Agar Plate Growth ............................................................... 134
Table 4.2. Spectrophotometric analysis of isolated genomic DNA ............................... 144
Table 4.3. Number of colonies following overnight growth ......................................... 150
Table 4.4. Results of spectrophotometric analysis of TOPO Libraries ....................... 151
Table 4.5. Resultant colonies containing pGL3-Basic plasmids ................................. 154
Table 4.6. Genes that flank the 3' and 5' ends of isolated sequences ......................... 162
Table 4.7. All Jasper Clover Motifs ............................................................................... 164
Table 4.8. All Nuclear Receptor binding Motifs ........................................................... 165
Table 4.9. Transcription Factor motifs found in the reportor plasmid sequences ......... 166
Table 4.10. TF Binding Motifs in MEME Results found using TOMTOM ................ 168
1.0. INTRODUCTION

1.1. Gene Expression

The expression of a genome is the mechanism whereby information contained in the genome is synthesised into a gene product.

Steroid hormones mediate their functions by activating the expression of specific subsets of genes to elicit a physiological response. The activation and expression of genes is a highly regulated series of events which depend on the interaction of RNA Polymerase II with the genome. The interaction of RNA Polymerase II with the genome in response to steroid hormones is mediated by nuclear receptors. The interaction of nuclear receptors with the genome is hindered by the chromatin environment which can be altered to either enhance the genomic interaction of inhibit it via activators and inhibitors, co-enhancers and co-repressors (Reviewed in McKenna and O’Malley, 2002).

1.2. Steroid Hormones

Steroid hormones are generally synthesised by cholesterol in the gonads or adrenal glands. Being synthesised from cholesterol they are lipophillic and can readily transverse the cell membrane (Olefsky, 2001) and combine with a specific protein receptor which, when activated by the steroid hormone, interacts with the cellular transcription machinery or specific genomic elements to either activate or inhibit a
target gene. Steroid hormones are transported through the body via the circulatory system bound to specific carrier proteins.

### 1.2.1. Oestrogen and Progesterone

The ovarian steroids, oestrogen and progesterone, are steroid hormones which function to regulate a variety of aspects of female reproduction. Oestrogen is produced from developing ovarian follicles, the corpus luteum, the placenta during pregnancy and in smaller amounts by other tissues (secondary tissues; important source of oestrogen in post menopausal females) such as liver, adrenal glands and mammary glands. Progesterone, on the other hand, is produced in the ovaries, the brain and also in the placenta during pregnancy. Levels of oestrogen and progesterone in menopausal females vary depending on the stage of the cycle (*Figure 1.1.*).

*Figure 1.1. Fluctuating levels of estrogen and progesterone during the menstrual cycle.*
Positive and negative regulation of these steroid hormones are controlled by a series of feedback mechanisms incorporating the pituitary hormones: follicle stimulating hormone (FSH) and luteinizing hormone (LH) (Biology of Human Reproduction, R Pinon).

1.2.2. Sex Hormone Synthesis

Oestrogen occurs naturally in three major forms; oestrone (E1), oestrodiol (E2) and oestriol (E3) (Figure 1.2.). Estradiol is widely considered to be the principal form present in non pregnant females.

![Molecular structure of Oestrone, Oestradiol and Oestriol](image)

**Figure 1.2. Molecular structure of Oestrone, Oestradiol and Oestriol**

Estrogens are present in both males and females, usually at a significantly higher level in females of reproductive age. All forms promote the development of female secondary sex characteristics, such as mammary development and growth of body hair, and also in aspects of regulating the menstrual cycle, such as thickening of the endometrium. Oestrogen also has functions in males including regulation of sperm maturation and release (Hess *et al.* 1997).
Progesterone (Figure 1.3.) has a number of physiological roles all of which are augmented in the presence of oestrogen. Progesterone converts the endometrium to its secretory stage to prepare the uterus for the arrival of a developing embryo in case fertilization occurs, it also accelerates the movement of the oocyte or embryo to the uterus and causes enlargement of mammary glands (in combination with other steroid hormones such as estradiol).

![Progesterone](image)

*Figure 1.3. Molecular structure of Progesterone*

Follicular theca and granulosa cells in the ovary are the most important site of oestrogen synthesis in females of reproductive age where production is stimulated by FSH and LH (Figure 1.4.) (Wang et al. 1980; Simpson et al. 1994; Michael et al. 1995; Bulum et al. 2000).

![Early Follicular Phase](image)

*Figure 1.4. Synthesis of Oestrogen (Adapted from figure in Biology of Human Reproduction, Pinon)*
Follicular maturation is characterized by the proliferation and functional differentiation of ovarian granulosa and theca cells. In the early stage of follicular development, granulosa cells undergo rapid proliferation with limited capacity for hormonal production. As the follicles grow and develop, the cells demonstrate an increased ability to secrete hormones (Gougeon, 1993). Cross-talk between granulosa and theca cells is essential for the maintenance of the physiologic function and structure of follicular cells and for the production of the ovarian hormones (Yada et al, 1999).

Synthesis of oestrogen is initiated by increasing levels of the pituitary hormones LH and FSH. Oestrogen is synthesized from androstenedione, a substance of moderate androgenic activity, which is synthesized from cholesterol (Figure 1.5).

**Figure 1.5. Illustration Outlining Oestradiol Synthesis (Ray et al, 2006)**
Cholesterol is the ‘building block’ of all steroid hormones and undergoes double oxidation to produce 20,22-dihydroxycholesterol to initiate the synthesis of both oestrogens and progesterone, this is then further oxidized with loss of the side chain starting at position C-22 to produce pregnenolone. These conversions occur in ovarian mitochondria. 17α-hydroxypregnenolone is then formed by the action of a smooth endoplasmic reticulin enzyme, 17α-hydroxylase. Dehydroepiandrosterone is formed by another smooth endoplasmic reticulin enzyme, 17,20 lyses, which is converted to androstenedione by 3β-hydroxysteroid dehydrogenase which is then converted to estrone by aromatase. Alternatively, dehydroepiandrosterone is converted to androstenediol by 17β-hydroxysteroid dehydrogenase and then to testosterone by 3β-hydroxysteroid dehydrogenase which is then converted to estradiol by aromatase. Oestrone and oestradiol can further be converted to estriol in the liver and placenta.

Theca cells convert cholesterol to androgens under LH stimulation from the pituitary; however theca cells lack the enzyme aromatase and so cannot convert the androgens to oestrogens. The androgens produced in the theca cells diffuse to the granulosa cells which then convert them to oestrogens, primarily oestradiol, under FSH stimulation. Therefore, the ability of the ovarian follicle to produce oestrogens is dependant firstly on the androgen producing ability of the theca cells, followed by the aromatase activity of the granulosa cells.

Intracellular cyclic adenosine monophosphate (cAMP) regulates aromatase expression in the ovary. Upon FSH binding to its G-protein-coupled receptor located in the granulosa cell membrane, intracellular cAMP levels increase which functions to enhance binding of two critical transcription factors (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).

Progesterone is also synthesised from cholesterol, production being stimulation by LH in vivo (Carlson et al., 1971) and in vitro (Armstrong and Black, 1966) (Figure 1.6.). Oestrogen action on both the pituitary and the hypothalamus is required for the LH surge that occurs at the end of the follicular phase. FSH and oestrogen action on the granulosa cells initiates the synthesis of LH receptors and the combination of LH and FSH action on the granulosa calls triggers the initiation of progesterone synthesis.

**Late Follicular Phase**
(beginning of luteinization)

![Diagram](image_url)

**Figure 1.6.** Synthesis of Oestrogen and progesterone (Adapted from figure in Biology of Human Reproduction, Pinon)

Similarly, progesterone synthesis begins with the conversion of cholesterol to pregnenolone by a cholesterol side chain cleavage enzyme present in the mitochondria (Figure 1.5.). Pregnenolone is then converted to progesterone by 3β-hydroxysteroid dehydrogenase. Progesterone can also be converted to oestrogens by 17α-hydroxylase.
conversion to 17α-hydroxy progesterone and subsequent conversion to androstenedione and testosterone by 17,20 lyases and 17β-hydroxysteroid dehydrogenase.

The major sites of oestrogen biosynthesis in postmenopausal women include 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.

During pregnancy the placenta becomes the main source of progesterone.

1.3. Nuclear Receptors

Oestrogen and progesterone carry out their functions by interacting, via their specific receptors, with promoter elements in the genome thereby inducing gene transcription and protein production. The specific oestrogen and progesterone receptors are just two of a major group of protein receptors known as the ‘nuclear receptor superfamily’. These receptors are responsible for detecting lipophilic hormones and certain other molecules in the cell cytosol. Lipophilic hormones, such as oestrogens and progesterone, have the ability to transverse the cell membrane (Olefsky, 2001), in contrast to hydrophilic hormones which are unable to cross the cell membrane. Hydrophillic hormones generally interact with cell surface receptors which cause a
conformational change inside the cell initiating a cell signalling pathway inducing gene transcription of repression.

Lipophillic hormones however, enter the cell cytosol, bind to the appropriate nuclear receptor which causes dissociation of receptor bound heat shock proteins and induces dimerization, translocation to the nucleus and DNA binding to specific motifs in, or close to, the promoter element of the target genes (Beato, 1989) (Figure 1.7).

![Diagram of steroid receptor activation](image)

**Figure 1.7.** The activation of Steroid Receptors. Steroid hormones enter the cell cytoplasm and bind to nuclear receptors which induce a conformational change in the receptor and dissociation of nuclear receptor associated heat shock proteins. Ligand bound nuclear receptors dimerize which enables translocation to the nucleus and interaction with specific motifs in the promoter regions of target genes.

This binding can act to either induce or repress gene transcription. Therefore, the nuclear receptor superfamily can act as ligand-dependant transcription factors.
Members of the nuclear receptor superfamily work with other proteins to coordinate and control gene expression, and in effect play an important role in the growth, differentiation, metabolism, reproduction and morphogenesis in higher organisms and humans (Aranda et al, 2001).

The nuclear hormone receptor superfamily includes receptors for thyroid and steroid hormones, retinoids and Vitamin D and also a subgroup of transcription factors, orphan receptors, for which the ligands have not yet been identified (Aranda et al, 2001).

1.3.1. Oestrogen Receptors

Oestrogen receptors (ERs) refer to a group of receptors that are activated by oestrogens. These include the well known ER, members of the nuclear hormone family of intracellular receptors which has up to 60 members (Lander et al, 2001) as well as a trans-membrane intracellular receptor known as GPR30 (G-protein coupled receptor) (Revankar, 2005).

ERs are localized to many sites within the cell including the nucleus and the plasma membrane as well as cellular mitochondria (Chen et al, 2004) and endoplasmic reticulum (Revankar, 2005).

There are two known nuclear hormone oestrogen receptors, ERα and ERβ. Each are encoded by separate gene; ERα by ESR1 gene located on chromosome 6 (q24-q27) while ERβ is encoded by ESR2 gene located on chromosome 14 (q21-q22). There are 8 isoforms of ERα and β (Figure 1.8.) (Matthews and Gustafsson, 2003).

ERα and ERβ are Class I nuclear receptors that bind to elements in the 5' flanking regions of DNA as homodimers. The main function of the oestrogen receptor is as a
DNA binding transcription factor that regulates the expression of genes controlled by oestrogen. However, the oestrogen receptor has been found to have additional functions which are independent to DNA binding.

![Diagram of human ERα and β and their respective isoforms.](image)

**Figure 1.8.** Representation of human ERα and β and their respective isoforms. Both hERβ2 and hERβΔ5 can have either the hERβ1 long or the short N-termini. In addition, there is no evidence of full-length hERβ4 or hERβ5. The different fill patterns of the 3' end of hERβ2, 4, and 5 represent the divergent C-terminal regions of these isoforms. (Matthews and Gustafsson, 2003)

The binding of oestrogen to ERα or ERβ either activates or represses gene transcription. The action carried out is a result of the steroid-receptor complex binding to specific regions in the promoters of target genes, known as oestrogen responsive elements (EREs). The resulting protein products determine the cell biological actions of the sex steroid.

There is also a second mechanism of action for oestrogen receptors which involves the interaction of nuclear estradiol (E2) oestrogen receptor complexes with transcription factors, such as activator protein 1 (AP-1), that in turn bind their specific DNA binding elements (Kushner et al, 2000). DNA binding via EREs or AP-1 elements leads to
chromatin remodelling, histone unwinding, recruitment of basal transcription factors and co-activators, displacement of co-repressors and finally gene transcription and protein expression (discussed in review by Levin, 2005).

Additionally, a small amount of ERs localize to the plasma membrane whereby a response is elucidated via cellular signalling mechanisms such as kinase cascades, calcium, and other second messenger signalling pathways which function to regulate transcription (Qin et al, 2004; Watters et al, 2000).

GPR30 is uniquely localized to the endoplasmic reticulum, where it specifically binds oestrogen (Revankar, 2005). Upon oestrogen binding intracellular calcium is mobilized and phosphatidylinositol 3,4,5-trisphosphate (PIP3: involved in the PIP2 lipid signalling second messenger) synthesis is activated in the nucleus where it initiates many non-genomic signalling events.

Human ERα has a molecular weight of 66 to 70 kDa, comprising of 595 amino acids (Pavão et al, 2001). ERα was isolated in 1962 by Jensen et al, in 1968 it was found to be a ligand activated transcription factor (O'Malley et al, 1968) and it was cloned in 1986 by Greene et al (Greene et al, 1986). ERα is localised in central nervous system, cardiovascular system, urogenital tract, bone and breast (Gustafsson, 1999). It is also the major ER subtype in the liver and the uterus (Gustafsson, 1999; Taylor et al, 2000) (Figure 1.9).

ERβ was discovered in 1996 by Kuiper et al in rat prostate and ovary and cloned in 1997 by Enmark et al. Full-length ERβ is 530 amino acids in length and has a molecular weight of 58 to 62 kDa although it has been shown to be produced in multiple isoforms, ERβ1 and ERβ2 (Figure 1.8.) (Peterson et al, 1998). ERβ is seen at
highest levels in the ovary and prostate with lower expression levels seen in the cardiovascular system, bone, breast, uterus and testis (Mosselman et al, 1996; Tremblay et al, 1997; Kuiper et al, 1997; Gustafsson, 1999). It is also the major ER subtype in the gastrointestinal tract, brain and prostate (Figure 1.9.) (Gustafsson, 1999; Taylor et al, 2000).

**Figure 1.9. Overall distribution of ER in different tissues (Gustafsson, 1999)**

ERα has 6 main structural and functional domains (named A-F from N to C terminus) (Figure 1.10.): a variable or regulatory domain containing a transcriptional activation function (A/B), a DNA binding domain (C), a hinge region (D), a ligand binding domain (E) and a region that plays a role in distinguishing agonists vs. antagonists (F). In vitro site mutagenesis assays, deletion mutation studies and domain swapping experiments have confirmed the functional roles of these domains (Green et al, 1987; Kumar et al, 1986; Kumar et al, 1987)
The regulatory A/B domain (aa 1-184) of ERα exhibits very little conservation between species (Krust et al, 1986) and among other nuclear receptor superfamily members (Seagraves, 1991). This domain shows promoter- and cell-specific activity and is likely to contribute to the specificity of action among isoforms and that it could interact with cell type specific factors (Aranda et al, 2001). A portion of this N-terminal regulatory domain A/B (aa 41-150) is known as the Activation Function-1 (AF-1) domain, associated with the modulation of transcriptional activity (Klinge, 2000). The AF-1 domain of ERα is phosphorylated at serine or threonine residues via the mitogen-activated protein kinase (MAPK) *in vitro*, and in cells treated with growth factors that stimulate the Ras-MAPK cascade, and this phosphorylation has been shown to enhance transcriptional activity (Kato et al, 1995; Patrone et al, 1996; Aranda et al, 2001).

The DNA binding domain, located in region C (aa185-263), is a highly conserved region (Kumar et al, 1987) containing two zinc fingers. The zinc fingers (both 1 and 2) are responsible for the direct interaction with the DNA double helix (Kumar et al, 1987). Each zinc finger contains four cysteine residues that coordinate the binding of a zinc atom. The ‘P Box’ is a term given to three amino acids at the base of the first zinc finger and is responsible for DNA binding specificity (Nelson et al, 2005), the three amino acids at the base of the second zinc finger is known as the ‘D Box’ and is
involved in dimerisation (Aranda et al, 2001). The core DBD contains two α-helices: the recognition helix located at the third conserved cysteine residue which binds the major groove of DNA making contact with specific bases and a second which spans the COOH terinus of the second zinc finger forming a right angle with the recognition helix (Aranda et al, 2001).

The hinge region, region D (aa 264-302) serves as a hinge between the DBD and the ligand binding domain (Aranda et al, 2001) containing sequences for receptor dimerisation and nuclear localization (Klinge, 2000).

The ligand binding domain (aa 303-553) (region E) is located in the C-terminal section of the receptor. This is a multifunctional domain that, in addition to ligand binding, mediated homo- and hetero-dimerisation, interaction with heat shock proteins, ligand-dependant transcriptional activity and in some cases hormone reversible transcriptional repression (Aranda et al, 2001). A portion of this domain (aa 530-553), AF-2 (a second activation functional domain), is also associated with transcriptional regulation via association with co-regulatory proteins (Henderson et al, 2003).

Region F (aa 554-595), (Figure 1.10.), is located at the C-terminal region. Its function is currently unknown but it shows very little sequence conservation similar between receptor isoforms (Mosselman et al, 1996). However, it is thought to play a role in distinguishing agonist vs. antagonist binding to the receptor (Pavao et al, 2001; Henderson et al, 2003).

Both ERα and ERβ share similar domain structure, ERβ also containing 6 functional domains (A-F). However, the A/B domain (aa 1-148) contains a repressor domain rather than an AF-1 activation domain. The A/B domain is the least conserved of the domains between ERα and ERβ with only 30% similarity at the amino acid level.
(Pavao et al., 2001). DNA binding domain is 96% and the ligand binding domain (aa 304-500) (containing a dimerisation domain, transactivation domain and a nuclear translocation domain) is 58% homologous, with an identical 'P Box', to the ERα sequence (Mosselman et al., 1996). There is a nuclear translocation domain located between the DNA binding domain and Hinge region. The E region (aa 304-500) contains a (dimerisation domain, transactivation domain and a nuclear translocation domain). The F region is found between amino acids 500 – 530.

1.3.1.1. Physiological Roles of Oestrogen Receptors

Oestrogen Receptors control a wide variety of physiological processes (Figure 1.11.), and therefore, are implicated in many diseases. Below is a brief overview of the major roles of the oestrogen receptors.

Reproduction

Oestrogens are essential for fertility, ERs being involved in ovulation, implantation, pregnancy maintenance and childbirth. It has been proposed the ERα mediates proliferative effects while ERβ mediates the differentiating effects of oestrogen within the follicles.

Additionally, during puberty, the lobular portions of the mammary terminal ducts are highly responsive to oestrogen. In breast tissue, oestrogens function to 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 distinct roles and distributions in the mammary glands. The density of
Oestrogen receptors in breast tissue is highest during the follicular phase of the menstrual cycle and gradually decreases following ovulation (Gruber et al, 2002).

Oestrogen receptors also have a role in male reproduction, ERα being the primary receptor responsible being localised in the efferent ductal epithelium. Its roles include the regulation of fluid reabsorption in the efferent ductules and maintaining a differentiated epithelial morphology. ERα knockout mice display dilation of cauda epididymal sperm, disruption of sperm morphology, inhibition of sodium transport and subsequent water reabsorption and eventual decreased fertility (Hess, 2003).

*Figure 1.11. Physiological Role of Oestrogen Receptors (Gruber et al, 2002)*
Skeletal System

Oestrogens have an important role in bone metabolism and homeostasis (reviewed in Turner et al, 1994) with effects on skeletal growth and also bone maturation.

In adolescence oestrogens are involved in the bone modelling, initiating 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 the stimulation of osteoblasts, and repressing bone resorption by repressing 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).

The action of oestradiol on bone maintenance and homeostasis is evident in women during menopause as decreasing oestrogen levels is associated with a higher incidence of osteoporosis (Turner et al, 1990; 1992).

Central Nervous System

Oestrogens are also thought to have neuroprotective actions (Behl and Holsboer, 1999). In brain tissue from adult rats, estrogens induce synaptic and dendritic remodeling and cause glial activation. In neurons of the hippocampus, an area involved in memory, estrogens increase the density of N-methyl-D-aspartate receptors and increase neuronal sensitivity to input mediated by these receptors. In cultured human neuroblastoma cells, estrogens have neuroprotective effects and reduce the generation of beta-amyloid peptides (reviewed in Gruber et al, 2002).
**Vascular Effects**

Oestrogens act as natural vasoprotective agents, and this is ER dependant (Bakir et al, 2000). Oestrogens cause short-term vasodilation by increasing the formation and release of nitric oxide and prostacyclin in endothelial cells (reviewed in Gruber et al, 2002).

### 1.3.2. Progesterone Receptors

There are two natural isoforms of the progesterone receptor: PR-A (94 kDa) and PR-B (104 kDa) (Horwitz & Alexander, 1983) (**Figure 1.12.**). One gene, PGR on chromosome 11 position q22, codes for both isoforms of the progesterone receptor, unlike the oestrogen receptor, which uses separate promotors (both oestrogen receptor inducible) and translational start sites to generate the two isoforms (Kastner et al, 1990).

**Figure 1.12.** Sequence similarities between human PR-A and PR-B

The progesterone receptor has multiple discrete, differentially conserved structural and function regions, designated A-F, similar to the oestrogen receptors (Beato, 1989). The non-conserved region A/B contains the transcription activation function region 1 (AF-1) which is present in both PR-A and PR-B. Region C contains the DNA binding domain (DBD), consisting of 2 zinc fingers that function to bind to DNA and enable recognition of the progesterone responsive element (PRE) located in the promoter elements of
progesterone responsive genes. Region E contains the hormone binding domain (HBD) which functions to prevent the DBD from binding to the responsive element. Ligand binding induces receptor dimerization prior to DNA binding and creates an 'active surface', within the HBD, thus generating a second transcriptional activating function (Webster et al., 1988, 1989; Tora et al., 1989). Region D is a hinge region, and Region F's function is currently unknown, both described in more detail in Section 1.2.1.

Transcription of one gene from multiple promoters provides advantages in gene expression; one or more of the resultant transcripts may be truncated causing the different transcripts to convey a variety, and sometimes opposing, biological actions. Thereby one gene is used to transcribe a number of isoforms, which can be expressed by different transcription factors, the products of which elicit a number of varied biological responses. The progesterone receptor gene is oestrogen inducible yet the promoter regions for PR-A and PR-B do not contain consensus oestrogen responsive elements (EREs), however the PR-A promoter region does contain a half-palindromic ERE site which may be involved in the oestrogen responsiveness of PR-A (Kastner et al, 1990). PR-B differs from PR-A by an additional stretch of 164 amino acids at the N-terminal end of PR-B, known as the PR-B specific domain which encodes a third transactivation function domain (AF-3) which is absent from PR-A (Figure 1.12.) (Sartorius et al, 1994; Wen et al, 1994). Additionally, they both share important structural and functional domains, with similar DNA and ligand binding affinities yet they are not functionally identical and have been shown to exhibit functionally distinct transcriptional activation properties specific to cell types, ligands and promoters (Tora et al, 1988; Vegeto et al, 1993; Meyer et al, 1990; Kastner et al, 1990).
PR-A and PR-B isoforms can have been isolated in both homodimer and heterodimer forms, the heterodimer form may have transcriptional activation properties that differ from those of PR-A or PR-B homodimers (Meyer et al, 1990; Mohamed et al, 1994).

PR-A has been identified as a cell and promoter specific functional repressor of PR-B. PR-A has been seen to act as a potent trans-dominant repressor of PR-B mediated transcription. Additional PR-A was also capable of inhibiting glucocorticoid, androgen and mineralocorticoid receptor-mediated gene transcription. This suggests a specific role for the hPR-A isoform in this regulatory process. This trans-dominant activity of PR-A originates from the extreme N-terminal 140aa region.

The different biological effects of PR-A and PR-B may function to generate tissue specific regulation of progesterone, whereby different cell types can provide a differing response to progesterone depending on the balance of PR-A and PR-B present in the cell.

Allan et al identified a hormone-specific conformational change in the LBD of PR which is required for events that follow PRE binding (such as transcriptional activity or receptor recycling). The conformational change may trigger HSP dissociation as well as DNA binding. However this conformational change is similar to that seen in nuclear receptors that bind to their response elements in DNA independently of a ligand (Allan et al, 1992).

1.3.2.1. Physiological Roles of Progesterone Receptor

Progesterone receptor has a major role in the functioning and maintaining of the menstrual cycle (Figure 1.1.) as well as the functions mentioned below.
**Mammary Glands**

Progesterone has a major developmental role in the normal mammary tissue, and it is hypothesised that it controls mammary gland proliferation and development (in partnership with oestrogen) via the secretion of paracrine growth factors (Anderson, 2002). It also is believed to be involved in the formation of lobular alveolar structures during pregnancy (Topper and Freeman, 1980). The influence of progesterone is likely to be proliferative in this process, mediated by progesterone regulation of cell cycle genes, growth factors, and growth factor receptors. Progesterone also exerts a differentiating effect on the breast through its role in lactation (reviewed in Graham and Clarke, 1997).

**Brain**

Oestrogen and progesterone control specific brain functions involved in reproductive behaviour. Progesterone regulates signals in the brain involving sexually responsive behaviour. The most well defined aspect of progesterone effects on this process are PR mediated effects in the hypothalamus and pre-optic area (reviewed in Graham and Clarke, 1997).

**Bone**

Expression of PR in normal human osteoblast like cells has been reported (Wei et al, 1993), and it may have a role in bone matrix regulation, via its effect on metalloproteinases.
Anti-oestrogen Action of Progesterone

Many of the effects of progesterone are thought to be due to its ability to oppose the action of oestrogen, particularly in the uterus. Progesterone abolishes oestrogen induction of many of the known hormone-responsive genes (reviewed in Clarke and Sutherland, 1990; Green and Clarke, 1997).

Progesterone directly reduces ER concentration, and additionally opposes ER-mediated gene-regulatory events, although the molecular mechanisms of this antagonism are not clear. However, the repressive effects of progesterone appear to be promoter- and cell-specific, and there is considerable variability between reports (reviewed in Graham and Clarke, 1997).

PR-A has been demonstrated to have similar repressive effects on other members of the nuclear receptor family, including those for androgens, mineralocorticoids, and glucocorticoids (McDonnell et al, 1994a; 1994b; Wen et al 1994), although the physiological significance of this observation remains to be determined.

1.3.3. Additional Transcription Factors

1.3.3.1. PAX4

The Paired box 4 (PAX4) gene was first identified in mice, and was subsequently cloned from humans (Dohrman et al, 2000). PAX4 is a member of the PAX family of transcription factors (Dohrman et al, 2000; Chi and Epstein, 2002)
PAX4 contains a paired box domain (a highly conserved DNA binding domain), an octapeptide and a paired-type homeodomain (homeodomain found only in PAX4 and PAX6) (Dahl et al, 1997; Dohrman et al, 2000). Pax proteins are key regulators of vertebrate organogenesis since they play major roles in embryonic pattern formation, cell proliferation and cell differentiation (Chi and Epstein, 2002; Dahl et al, 1997; Dohrman et al, 2000; Epstein et al, 1994).

PAX4 is a transcriptional repressor that binds to a common element in the glucagon, insulin and somatostatin promoters and it has been found to bind to an 8 bp consensus sequence: AA(T/A)AATTA (Smith et al, 1999)

Additionally, it is also involved in pancreatic islet cell development, specifically the differentiation of insulin producing beta cells (Dohrman et al, 2000; Soso-Pineda, 2004).

1.3.3.2. SNAIL

SNAIL is a zinc finger protein transcription factor transcribed from the Snail gene and is 264 amino acids in length. The Snai1 gene was first isolated from Drosophila (Simpson, 1983) after which the protein was isolated from human cell lines (Okaru et al, 2001). SNAIL belongs to the SNAIL C2H2-type zinc-finger protein family (which includes Slug) and contains 4 C2H2-type zinc fingers. It is involved in DNA, metal and zinc binding. SNAIL is expressed in a variety of tissues, including placenta and adult heart, lung, brain, liver, and skeletal muscle, with the highest expression in kidney (Paznekas et al, 1999).
Drosophila SNAIL is a transcriptional repressor (Hemavathy et al, 2000; Nieto, 2002) that acts to maintain the proper germ layers by repressing the expression within the mesoderm of regulatory genes involved in ectodermal development (Leptin, 1991). Snail has an NH2-terminal domain containing a seven-amino acid SNAG domain. Okaru et al used site-directed mutagenesis experiments to confirm that this domain is important for the repressor activity of SnaH (Okaru et al, 2001).

SNAIL zinc-finger protein was thought to be required zygotically for mesoderm formation (Boulay et al, 1987; Simpson, 1983).

A SNAIL binding site was identified in the CREaroCF region of the human aromatase gene (CTGATGAAGT) and, when SNAIL interacts with this site, was found to repress the activity of this promoter (Okary et al, 2001) thereby repressing the production of aromatase, one of the enzymes involved in the synthesis of oestrogen (Section 1.2.2.). Quenching is a form of gene regulation whereby repressors occupy neighbouring sites in a target promoter and prevents the ability of the activator to gain contact with the DNA bound transcription complex (Gray et al, 1994). Gray et al reported that this is the mechanism used by Snail.

Normal breast epithelial cells and fibroblasts express SNAIL at relatively high levels, however these levels drop significantly in breast cancer tissue (Okaru et al, 2001). Okaru et al suggest that high expression of SNAIL in normal breast epithelial cells and fibroblasts prevents proteins from binding to CREaro, the promoter element in the aromatase gene. This prevents excessive synthesis of estradiol. However, low levels of SNAIL expression in breast cancer tissue results in low levels of SNAIL binding to the CREaro site in the aromatase promoter, enabling transcription factors to build up and expression of aromatase which leads to synthesis of estradiol.
Additionally SNAIL has been implicated in the repression of Vitamin D receptor expression in human colon cancers (Palmer et al, 2004).

1.3.3.3. RORα

Retinoid-related Orphan Receptor-α, or RORα, is an orphan nuclear receptor encoded by the RORA (chromosome 15 q22.2) gene which encodes a protein between 468 and 548 aa in length depending on the isoform.

The RORα gene generates four isoforms that share common DBDs and putative LBDs but are distinguished by different N-terminal domains (Gigue’re et al, 1994) (Figure 1.13.).

![Diagram of ROR isoforms](image)

**Figure 1.13. Isoforms of ROR (Adapted from Jetten, 2009)**

The RORα isoforms bind to monomeric RORE (RORα responsive element) found in the regulatory region of target genes. These ROREs are generally composed of a 5' 6 base pair A/T-rich sequence that precedes a 3' AGGTCA core half-site motif (Gigue’re et al, 1994). It is thought that RORα binds DNA as a monomer, however it remains a possibility that it could form heterodimeric complexes with an unidentified partner in vivo (Gigue’re et al, 1995).
Most RORα isoforms are tissue-specific and are involved in the regulation of different physiological processes and have different target genes. Human RORα3 is only found in human testis (Steinmayr et al, 1998). RORα is found ubiquitously expressed in mouse tissue with highest levels of expression have been observed in the Purkinje cells of the cerebellum, retina, lens, spleen, skeletal muscle, and testis (Retnakaran et al, 1994; Matsui et al, 1995; Hamilton et al, 1996; Sashihara et al, 1996; Nakagawa et al, 1997; Koibuchi et al, 1998; Koibuchi et al, 1998).

The structure of RORs is typical of nuclear receptors. They consist of four major functional domains: an N-terminal domain in the A/B region followed by a highly conserved DBD, a hinge region and a C-terminal ligand binding domain (Figure 1.13.). The A/B region has been demonstrated to play a critical role in conferring DNA binding specificity on the various isoforms (Gigue’re et al, 1994; Gigue’re et al, 1995). The promoter context, in conjunction with the RORE sequence and the amino terminus, may be responsible for which of the isoforms is recruited.

RORα has been implicated in a number of physiological roles (Figure 1.14.).

Cholesterol (7-dehydrocholesterol and cholesterol sulphate) have been identified as RORα agonists (Kallen et al, 2004; Kallen at al, 2002), and have been found to regulate the expression of genes involved in lipid metabolism (Kallen et al, 2002). These factors suggest that RORα has a role in the regulation of lipid metabolism. High levels of RORα has been seen in cerebellar Purkinje cells (Matsui et al, 1995), and knockout studies have demonstrated that RORα deficient mice have display severe cerebellar atrophy, characterized by a significant reduction in Purkinje cells and a loss of cerebellar granule cells (Hamilton et al, 1996). Additionally, RORα has been implicated in spinocerebellar ataxia type 1 (SCA1), an autosomal dominant inherited neurodegenerative disorder
characterized by progressive loss of motor coordination, speech impairment, and problems swallowing (reviewed in Jetten, 2009).

Figure 1.14. Physiological functions of ROR (Jetten, 2009)

1.3.3.4. COUP TF I and II

Chick Ovalbumin Upstream Promoter-Transcription Factor (COUP-TF) is one of the best characterized orphan nuclear receptors. It was first identified as a homodimer, found to bind to a direct repeat regulatory element, containing an imperfect direct repeat of AGGTCA, in the chicken ovalbumin promoter (Pastorcic et al, 1986; Sagami et al, 1986; Wang et al, 1989). Each factor has the ability to generate its own distinct expression profile during mammalian development (Qiu et al, 1994). COUP TF I is identical to EAR2 (Miyajima et al, 1988) while COUP TF II has been found to be identical to ARP-1 (Ladias and Karathanasis, 1991).
Both COUP TF I and II are expressed in many human tissues, being involved in many biological regulation processes such as neurogenesis, organogenesis, determination of cell fate and metabolic homeostasis (Tsai et al, 1994; Ladias and Karathanasis, 1991; Wang et al, 1991; Mangeldorf et al, 1995; Ritchie et al, 1990; Qui et al, 1994).

COUP TF contains a 66-amino cis region which contains two conserved Zinc finger motifs, this is the DBD region.

COUP TF I and II are identical but for one amino acid difference in within the DBD which is a conservative change from Ser to Thr (Miyajima et al, 1988; Wang et al, 1989).

1.4. Coregulators: Coactivators and Corepressors

As stated previously in Section 1.3., nuclear receptors comprise of ligand-regulated (e.g. estrogen receptors and progesterone receptors) and orphan transcription factors (a total 48 receptors in humans). These proteins play a central role in the body's ability to transduce steroid, retinoid and other lipophilic endocrine hormone signals to the genome (Tsai and O'Malley, 2004; Mangelsdorf et al, 1995). Nuclear Receptors have the ability to recognise specific sequence motifs within the promoter, or enhancer, regions of their target genes, generally associating with the ligand binding domain and with the AF-1 and AF-2 domains of certain nuclear receptors. However, the ability of the NR receptor to bind to the specific sequence motifs is dependant on, firstly, the ability of the NR to access the motif through chromatin and, secondly, the relationship of the NRs with the RNA Polymerase II holocomplex. (Roeder, 2005). These difficulties are overcome, or exacerbated, by coregulators. Coregulators are molecules
that are directly recruited by nuclear receptors and either activate (coactivator) or repress (corepressor) the transcription of specific genes (Lonard and O’Malley, 2005).

Coactivator recruitment is generally ligand dependant while corepressors repress gene expression primarily via interaction with NRs not bound to ligand to repress transcription (Glass and Rosenfeld, 2000). Corepressor associated proteins such as histone deacetylases, prevents transcription by creating a local chromatin environment which hinders association of transcriptional machinery, this opposes the transcriptional promoting activities of coactivators (such as acetyltransferases). The opposing actions between pro-transcription and anti-transcription creates an equilibrium which can be tipped either way according to the signals received (Reviewed in Lonard and O’Malley, 2007; 2006).

Core coregulators, those that interact directly with nuclear receptors, exist in large steady-state complexes with multiple secondary co-coregulator partners. Each component may possess multiple enzymatic capabilities such as acetyltransferase, methyltransferase, phosphokinase, ubiquitin ligase, and ATPase activities, ultimately making these complexes versatile enzymatic ‘machines’ for regulating gene expression (Lonard and O’Malley, 2006) (Figure 1.15.).

Currently, approximately 300 coactivators and corepressors have been reported (http://www.NURSA.org), some examples are listed in Figure 1.15..
<table>
<thead>
<tr>
<th>Coregulator</th>
<th>PTM</th>
<th>Regulatory Consequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-terminal binding</td>
<td>Phosphorylation</td>
<td>Loss of repressor activity</td>
<td>Barnes et al, 2003</td>
</tr>
<tr>
<td>protein 1</td>
<td>N-CoR</td>
<td>Nuclear export</td>
<td>Hermanson et al, 2002</td>
</tr>
<tr>
<td>SUMOylation</td>
<td>Increased repressive activity</td>
<td>Tiefenbach et al, 2006</td>
<td></td>
</tr>
<tr>
<td>Ubiquitination</td>
<td>Increased protein turnover</td>
<td>Zhang et al, 1998</td>
<td></td>
</tr>
<tr>
<td>PCBP1</td>
<td>Phosphorylation</td>
<td>Loss of translational repression, conversion to a transcriptional coactivator, and control of alternative splicing</td>
<td>Meng et al, 2007</td>
</tr>
<tr>
<td>PGC-1a</td>
<td>Phosphorylation</td>
<td>Increased transcriptional activity</td>
<td>Puigserver et al, 2001</td>
</tr>
<tr>
<td>Methylation</td>
<td>Decreased transcriptional activity</td>
<td>Teysier et al, 2005</td>
<td></td>
</tr>
<tr>
<td>Acetylation</td>
<td>Decreased transcriptional activity</td>
<td>Rodgers et al, 2005</td>
<td></td>
</tr>
<tr>
<td>RIP140/NRIP1</td>
<td>Vitamin B6 conjugation</td>
<td>Huq et al, 2007</td>
<td></td>
</tr>
<tr>
<td>Phosphorylation</td>
<td>Increased repressive activity</td>
<td>Gupta et al, 2005</td>
<td></td>
</tr>
<tr>
<td>Arginine methylation</td>
<td>Loss of repressive activity</td>
<td>Mostaqul Huq et al, 2006</td>
<td></td>
</tr>
<tr>
<td>SHARP</td>
<td>Phosphorylation</td>
<td>Increased transcriptional repression</td>
<td>Vadlamudi et al, 2005</td>
</tr>
<tr>
<td>SMRT</td>
<td>Phosphorylation</td>
<td>Nuclear export</td>
<td>Jonas and Privalsky 2004</td>
</tr>
<tr>
<td>SRC-1</td>
<td>Phosphorylation</td>
<td>Increased transcriptional activity</td>
<td>Rowan et al, 2000</td>
</tr>
<tr>
<td>SRC-2</td>
<td>Phosphorylation</td>
<td>Increased transcriptional activity</td>
<td>Lopez et al, 2001</td>
</tr>
<tr>
<td>SRC-3</td>
<td>Phosphorylation</td>
<td>Increased transcriptional activity, transcription factor-dependent transcription</td>
<td>Wu et al, 2002, 2004</td>
</tr>
<tr>
<td>Arginine methylation</td>
<td>Transcriptional dynamics, protein dissociation</td>
<td>Feng et al, 2006; Naeem et al, 2007</td>
<td></td>
</tr>
<tr>
<td>Acetylation</td>
<td>Protein dissociation</td>
<td>Chen et al, 1999</td>
<td></td>
</tr>
<tr>
<td>Ubiquitination</td>
<td>Increased transcriptional activity and protein turnover</td>
<td>Wu et al, 2007</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1.15. Selected Coregulator Posttranslational Modifications (from Lonard and O’Malley, 2007)

1.4.1. Nuclear Receptor Coactivators

Coactivators associate with ligand bound activated NRs (or other DNA binding transcription factors) and are required both for the transcriptional activation process to
occur (Drapkin et al, 1993; Tijan and Maniatis, 1994; Guarente, 1995; Verrijzer and Tijan, 1996), and for enhanced gene expression.

Coactivators are diverse, both structurally and through their biochemical processes. Coactivators function via numberous enzymatic means, including acetylation, methylation, ubiquitination and phosphorylation. These enzymatic properties of the coactivators enable them to function as chromatin remodelers (reviewed by Lonard and O’Malley, 2006).

A family of coactivators, known as the p160/SRC family, currently consists of 3 members, described below. The common domains of this family of coactivators can be seen in Figure 1.16.

![Figure 1.16](image-url)

**Figure 1.16.** Schematic diagram of the common structural and functional domains of SRC family of coactivators. The lines above the bar identify the regions containing conserved structural and functional features among SRCs. These include: the basic-helix-loop-helix/Per/ARNT/Sim (bHLH/PAS); the serine/threonine-rich regions (S/T); the LXXLL motifs responsible for interaction with nuclear receptors (RID); the intrinsic transcription activation domains of SRCs (AD1 and AD2); the glutamine-rich regions (Q); and the HAT domain identified in SRC-1 and SRC-3.

The lines under the bar indicate regions that are responsible for interaction with transcription factors or other coregulators. MEF, Myocyte enhancer factor-2C; TEF, transcriptional enhancer factor; RID, receptor-interacting domain (Wu et al, 2005).
There are 3 members of the p160/SRC family of coactivators, namely; SRC-1, SRC-2 and SRC-3 (*Figure 1.17*).

*Figure 1.17. Multiple members of the SRC family. The proteins have been aligned according to major structural similarities and to emphasize both the structural divergence of the carboxy termini and the conservation of the amino-terminal domains of SRC family members. Regions to which specific functions of individual coactivators have been assigned are indicated. Only SRC-1 contains a consensus LXXLL/NR box motif in this region. (McKenna et al, 1999b)*

**1.4.1.1. SRC-1**

SRC-1 was the first NR coactivator to be successfully cloned (Onate et al, 1995). This coactivator exhibits a broad range of specificity in the coactivation of ligand dependant transactivation of nuclear receptors, including PR, GR, ER, TR and RXR (Onate et al, 1995). The interaction of hSRC-1 with the LBD of the progesterone receptor is ligand
dependent (Onate et al, 1995) and is abolished in the presence of the antiprogestin RU486. In addition, a hSRC-1 mutant, containing only the C-terminal receptor-interacting domain, suppresses PR coactivation by hSRC-1 in a dominant-negative fashion, both in transient transfection and by in vitro transcription assay Onate et al, 1995; McKenna et al, 1999b) indicating that the C-terminal domain is not involved in the coactivation capabilities of the molecule.

Previous studies indicate that hSRC-1 has the ability to mediate functional interactions between the N-terminal AF-1 (not conserved in sequence in nuclear receptors) and C-terminal AF-2 (located within the hormone-binding domain, highly conserved among the nuclear receptors) domains of nuclear receptors. Individual hSRC-1 domains are required for full functional cooperation between AF-1 and AF-2 of the progesterone receptor (Onate et al, 1998), as well as the oestrogen receptor (McInerney et al, 1996). This suggests that that the efficient assembly of the preinitiation complex by steroid receptors is dependant on an SRC-1 assisted interaction between their individual AF domains (McKenna et al, 1999b).

Additionally, research conducted by Xu et al suggested that peripheral steroid target organs of mice containing a targeted disruption of the SRC-1 gene have a decreased response to steroid hormones (Xu et al, 1998). Therefore, SRC-1 is responsibe for causing an effective steroid hormone response.

In addition to the full-length SRC-1 (termed SRC-1A), several splice variants have been described: SRC-1A, SRC-1C, SRC-1D and SRC-1E (Kamei et al, 1996; Yao et al, 1996; Takeshita et al, 1996). SRC-1A and SRC-1E are most studied.

SRC-1e is a more potent coactivator for oestrogen receptor than SRC-1A (Robyr et al, 2000). For instance, the oestrogen-regulated rat oxytocin promoter is coactivated by
SRC-1E but not by SRC-1A, whereas both SRC-1 isoforms stimulate oestrogen receptor-mediated transcription from an artificial ERE-containing promoter. Therefore, SRC-1A coactivation appears to rely on the promoter context of the receptor target gene (Robyr et al, 2000).

Both isoforms contain three nuclear receptor interacting motifs (LXXLL) found in many co-factors (Heery et al, 1997). SRC-1A however possesses a fourth LXXLL motif at its C terminus (Kalkhoven et al, 1998). Mutation of the fourth LXXLL in SRC-1A does not affect transcription and its function is currently unknown (Robyr et al, 2000).

Additionally, SRC-1 has histone acetyltransferase activity (HAT) (Spencer et al, 1997). HATs are enzymes that acetylate conserved lysine amino acids on histone proteins by transferring an acetyl group from acetyl CoA to form e-N-acetyl lysine. Acetylation of histone molecule in DNA is generally linked to transcriptional activation.

1.4.1.2. SRC-2

SRC-2, also known as GRIP1 (glucocorticoid receptor-interacting protein 1), TIF2 (transcription intermediary factor 2) and NCoA-2, are 160-kDa nuclear receptor-interacting proteins with considerable sequence and functional similarity to SRC-1 (Voegal et al, 1996; Hong et al, 1996, 1997; Torchia et al, 1997).

GRIP1(mSRC-2) and TIF2 (hSRC-2) associate in a ligand-dependent manner in vitro with several receptor LBDs (Voegal et al, 1996) and, in vivo, with RARα, oestrogen receptor and progesterone receptor in the presence of hormone, but not hormonal antagonists (Voegal et al, 1996; Hong et al, 1996, 1997).
1.4.1.3. SRC-3

The identification of a third member of the SRC/p160 family was a highly polymorphic protein isolated as p/CIP (p300/CBP cointegrator-associated protein) (Torchia et al, 1997), ACTR (activator of thyroid receptor) (Chen et al, 1997), RAC-3 (receptor-associated coactivator 3) (Li et al, 1997), AIB-1 (amplified in breast cancer-1) (Anzick et al, 1997), TRAM-1 (thyroid receptor activator molecule 1) (Takeshita et al, 1997) and SRC-3 (Suen et al, 1998), illustrates complex nomenclature in the SRC family (McKenna et al, 1999b).

hSRC-3 interacts with and coactivates a wide variety of nuclear receptors in a ligand-dependent manner, including RAR, thyroid receptor, RXR, glucocorticoid receptor (Chen et al, 1997), progesterone receptor (Li et al, 1997) and oestrogen receptor (Anzick et al, 1997).

mSRC-3 exhibits greater promiscuity than other SRC family members by enhancing the transcriptional activity of a number of different activators, including interferon-α and cAMP regulatory element binding protein (CREB) (Torchia et al, 1997), which were previously shown to be dependant primarily on CREB-binding protein (CBP) for efficient activation.

Additionally, SRC-3 selectively enhances the transcriptional activity of ERα over that of ERβ, possibly reflecting a 60% difference in homology between the LBDs of these isoforms (Suen et al, 1998).

SRC-3 is both regulated by phosphorylation and distinct phosphorylation patterns on the SRC-3 molecule determine its specificity for different transcription factors (Wu et al, 2004).
Phosphorylation is required for SRC-3 activity, and the specificity of SRC-3 for different transcription factors further depends on where phosphorylation is induced on SRC-3 by a respective stimulus (Wu et al., 2004). It is proposed that a specific phosphorylation-dependent code allows SRC-3 to activate gene expression specifically and accurately (Wu et al., 2005).

1.4.3. Mechanism of Coactivator Action

Figure 1.18. demonstrates the various signalling pathways the SRC family are activated by and their preferential targets.

The p160/SRC family of coactivators have multiple methods of activating genomic transcription. Their two main mechanisms of action is to manipulate chromatin to enable access to the transcription machinery and also to activate the nuclear receptor via AF-1 and AF-2 domains in the nuclear receptor.

SRC-1 and SRC-3 has been identified as possessing HAT activity (Spencer et al., 1997; Lemon and Freedman, 1999; McKenna et al., 1999b).

CARM-1 (Coactivator-associated arginine methyltransferase 1) functions as a secondary coactivator through its association with the carboxyl-terminal region of p160 coactivators. CARM-1 can methylate histone H3 in vitro, and though mutation analysis was deemed necessary for the full coactivation activity of the p160 family. Thus, coactivator-mediated methylation of proteins in the transcription machinery may contribute to transcriptional regulation (Chen et al., 1999).
Figure 1.18. Signaling to SRCs. Kinase-mediated signaling pathways activated by a variety of stimuli (including steroid hormones, cytokines, and growth factors) that can communicate with SRCs, are displayed above. Although seven major phosphorylation sites (serines372, 395, 517, 569, 1033, 1185 and threonine1179) were identified on SRC-1, it is unclear whether phosphorylation is affected by steroid hormones (as indicated by ?; same is true for SRC-2), and only phosphorylation at threonine1179 and serine1185 induced by cAMP were shown. SRC-3 was used as a model molecule as there is a deeper understanding of SRC-3 phosphorylation. However, it is hypothesized that this same mode of operation from SRC-3 could also apply to SRC-1 and SRC-2. Although each SRC is able to coactivate with the various receptors in transient transfection assays, the preferred SRC partner for different receptor signaling is connected by an arrow. (Wu et al, 2005)
1.5. Steroids, Nuclear Receptors, Steroid Receptor Coactivators and Cancer

1.5.1. Breast Cancer

Breast cancer is the most common malignancy of women residing in industrialized countries and it is the most common cause of cancer associated morbidity and mortality (Giacinti et al, 2006). Sixty five percent of primary breast cancers are ER positive (Giacinti et al, 2006).

Oestrogen is involved in the growth and differentiation of epithelial cells in normal mammary tissue. The mitogenic effects of oestrogens on breast epithelial cells are due to, at least in part, increased expression of genes involved in the regulation of the cell cycle (Vendrell et al, 2004). However, very little is known about the actual mechanism of oestrogen mediated proliferation.

Target genes have been identified, through studies on oestrogen regulated gene expression in breast cancer cells, which may be involved in mammary cell proliferation. NOV, a gene that encodes for a negative regulator of cell proliferation, is down regulated by oestrogen in breast cancer cells (Vendrell et al, 2004). WNT2 is a gene whose expression is associated with abnormal proliferation in human breast tissue (Huguet et al, 1994). WNT2 expression has been seen 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). The TACC gene family contain a TACC domain, which forms a 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 a down regulation in TACCI mRNA has been observed in breast tumors (Conte, 2002).

Prolonged stimulation of oestrogen receptor on breast ductal epithelium contributes to the development and progression of breast cancer. Reducing oestrogen receptor levels or altering its activity has been found to induce tumour regression. In these cases treatment that are designed to block the effects of oestrogens are important in halting cancer progression (Osborne et al, 2000). (Discussed further in Section 1.5.2.)

While sixty five percent of primary breast cancers are ER positive, the remaining thirty five are ER negative. ER negative breast tumours are associated with more malignant and more drug resistant tumours.

Additionally, breast tumours that are initially identified as ER positive can loose their ER expression during tumour progression (Yang et al, 2001). This can indicate that the tumour is becoming more malignant.

Some ER positive breast cancers also contain progesterone receptors, whose presence may indicate that the ER remains responsive to oestrogen and can still control specific protein synthesis (Horwitz and McGuire, 1978).

1.5.2. **Selective Oestrogen Receptor Modulators – A treatment and preventative agent for Breast Cancers**

Treatments designed to block the effects of oestrogens are important in halting cancer progression in ER positive breast cancers (Osborne et al, 2000).
Selective oestrogen receptor modulators (SERMs) perform this function and 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 to the oestrogen receptor, alter its conformation, and thereby facilitate the binding of coregulatory proteins that can activate or repress the transcriptional activation of oestrogen target genes (Osborne et al, 2000).

1.5.2.1. Subdivisions of SERMs

SERMs can be divided into three major categories: Triphenylethylene derivatives, other nonsteroidal compounds and steroidal compounds that have more complete antioestrogenic activity. Tamoxifen and Raloxifene are the two major SERMs currently used as a treatment for hormone responsive breast cancer.

Tamoxifen

Tamoxifen (Figure 1.19.), a member of the Triphenylethlenes group of SERMs, was first synthesized in the 1960s, demonstrated to have antiproliferative effects in breast tissue (Jordan, 1994) and shown to be an effective therapy in patients with ER positive metastatic breast cancer (Osborne, 1998). The major metabolites of tamoxifen are N-desmethytamoxifen and trans-4-hydroxytamoxifen, which has an affinity for ER similar to that of 17-β oestradiol (Buckley et al, 1989).
Tamoxifen has both oestrogenic and anti-oestrogenic effects. It has agonist effects on bone (Love et al, 1992), blood lipids and also the endometrium (Osborne, 1998) while having anti-oestrogenic effects on the CNS and vaginal mucosa. The mixed oestrogenic and anti-oestrogenic effects of Tamoxifen are species, tissue and cell dependent. Tamoxifen has also been shown to have both oestrogenic and anti-oestrogenic effects on specific subsets of genes within a single cell type (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 and inhibiting cancer growth (Osborne, 1998). Tamoxifen's dual activities provide some advantages for patients 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 are associated with its oestrogenic activity in the endometrium, which can result in endometrium hyperplasia and also low-grade endometrial cancers (Wilking et al, 1997). Additionally, due to its antioestrogenic effects on the vaginal mucosa, treatment may result in menopausal symptoms (Love et al, 1991).

**Raloxifene**

Raloxifene (*Figure 1.20.*) is in the group of nonsteroidal compounds. It is a benzothiophene derivative and it binds to the ER with an equal affinity to that of oestradiol (Black et al, 1983). Raloxifene is an inhibitor of cultured breast cancer cells and possesses antitumor activity in rat mammary tumour models similar to that of Tamoxifen (Poulin et al, 1989).

![Figure 1.20. Chemical Structure of Raloxifene (www.chemblink.com)](image.png)

**1.5.3. Nuclear Receptor Coregulators and Human Disease**

Nuclear receptor coregulators are central to the regulation of NR-mediated transcription and in the coordination of intercompartmental metabolic processes (Lonard et al, 2007).
Knockout studies have demonstrated the importance of SRC-1, SRC-2 and SRC-3 in a wide variety of biological functions (*Table 1.1.*).

Consequently, disruptions in the SRC family of coactivators can lead to a wide variety pathological states. The following is an overview of the pathological states which occur as a result of these disruptions.

<table>
<thead>
<tr>
<th>Knockout</th>
<th>Phenotype</th>
</tr>
</thead>
</table>
| SRC-1    | *moderate motor dysfunction*  
          | *delayed development of purkinje cells*  
          | *control of energy balance*  
          | *loss of skeletal response to oestrogen*  
          | *altered hypothalamic-pituitary-adrenal axis function*  
          | *altered hepatic function* |
| SRC-2    | *control of energy balance in white and brown adipose tissue*  
          | *essential for the progesterone-dependant uterine and mammary morphogenesis*  
          | *testicular dysfunction*  
          | *spermatogenesis defect and placental hypoplasia* |
| SRC-3    | *animals of smaller size*  
          | *delayed puberty*  
          | *reduced mammary gland development*  
          | *reduced adipogenesis and inhibition of neointima formation by oestrogen*  
          | *decreased response to IGF-I* |

*Table 1.1. Consequences of SRC knockout studies (Adapted from Lonard et al, 2007)*

1.5.3.1. Cancer

Coactivators have been found to be over- and underexpressed in a number of cancers. This is the largest group of coactivator related disease.
Coactivator overexpression may invoke carcinogenesis, enhance its progression or alter the biological activities of therapeutic NR ligands (Shang and Brown, 2002). Coactivator misexpression can be thought of as a pervasive agent in the progression, or etiology, of human cancers.

SRC-1 has been found to be overexpressed in prostate, breast and gastric cancers. A MOZ-SRC-2 fusion protein has been identified in acute myeloid leukaemia. Also, there is a brain and breast cancer correlation with oestrogen and progesterone receptor and SRC-2 expression.

SRC-3 has been associated with numerous carcinomas, including breast, pancreatic, ovarian, endometrial, oesophageal squamous cell and colorectal. Overexpression of SRC-3 has been demonstrated in oral squamous cell carcinomas and a smaller isoform of the protein has been identified in certain breast and gastric cancers. Additionally, some polymorphisms can protect against breast cancer and influence calcaneal bone density. (All reviewed by Lonard et al, 2007)

1.5.3.2. Metabolic Syndromes

Knockout studies have found SRC-1 and SRC-2 to play opposing roles in energy metabolism (Feige and Auwerx, 2007). SRC-1<sup>−/−</sup> knockout mice are prone to obesity due to decreased energy expenditure while SRC-2<sup>−/−</sup> knockouts are leaner due to the reduced transcription of PPARγ2 (a protein essential for the differentiation of adipocytes). SRC-2<sup>−/−</sup> knockout mice also display a subsequent increase in PGC-1α/SRC-1 interaction which enhances the thermogenic actions of PGC-1α in brown adipose tissue. SRC-3's role in metabolism has been identified in the promotion of the
formation of white adipose tissue. SRC-3-/- knockout mice have a decreased adipose tissue mass (Louet et al, 2006; reviewed in Lonard et al, 2007).

1.5.3.2. Coactivator Fusion Proteins

Genetic disruptions leading to coregulator fusions with other proteins exist in certain cancers (Lonard et al, 2007). A MOZ-SRC-2 fusion protein predisposes an individual to acute myeloid leukaemia (Troke et al, 2006).

1.5.3.3. Coregulator Gene Polymorphisms

Coregulators function to control the extent of gene expression from nuclear receptors so that fluctuations in their expression or small changes in their biological activity will lead to significant differences in target tissue responses to hormone ligands. Individual variations in secondary sex traits, obesity and susceptibility to cancer may be attributed to alterations in the primary protein amino acid sequence and the cellular concentration of coregulators (Lonard et al, 2007).

Generally, coregulator SNPs (small nuclear polymorphisms) are located in the coregulator gene promoters or intronic or synonymous non-coding variants, however, SNPs can, if infrequently, occur and affect the coregulator amino acid sequence (Lonard et al, 2007).

There are 10 SNPs found in SRC-3 that affect its coding sequence and hence the translated protein (Sherry et al, 1999; Lonard et al, 2007). A Q586H variant allele, one
of the 10 SNPs that affects the coding sequence of the SRC-3 protein, confers a protective effect against breast cancer (Burwinkel et al, 2005)

1.5.4. Relationship of SERMs and SRCs

According to Lonard et al, SERMs have the ability to upregulate the expression of SRC-1 and SRC-3 in cells and, additionally stimulate the transcription of other nuclear receptors. They hypothesize that the differential agonist or antagonist action of SERMs is dependant on the coactivator expression in oestrogen target tissues (Lonard et al, 2004). Shang and Brown have demonstrated that elevated expression of SRC-1 in the uterus-derived Ishikawa cell line contributes to the increased agonist action of Tamoxifen in the uterus, while lower levels of SRC-1 in the breast derived MCF-7 cell line allow for Tamoxifen to function as an antagonist (Shang and Brown, 2002).

1.6. Identification of Transcription Factor Binding Sites

Current techniques available to enable the isolation of transcription factor binding sites include chromatin immunoprecipitation combined with microarray detection (ChIP-on-chip) and DamID. These two techniques are based on fundamentally different principles.
1.6.1. ChIP-on-chip

ChIP-on-chip analysis utilizes antibodies to identify genomic transcription factor binding sites. An overview of the methodology can be viewed in Figure 1.21.

*Figure 1.21. Step by step methods for performing Chip-onChip analysis* (http://genomics.nchresearch.org/gene_regulation_analysis.html)

63
Cells are isolated from tissue and fixed using 1% Formaldehyde which functions to cross link and preserve transcription factor DNA binding. Cells are then lysed and the nuclei isolated, following isolation the nuclei are lysed and the resultant lysate is sonicated or enzyme treated to fragment the DNA into 300 to 1000 bp fragments.

An antibody directed against the transcription factor of interest is then added to the sonicated lysate followed by cross linking reversal to release the antibody bound proteins, proteins are removed by Proteinase K treatment, DNA is purified, amplified using LM-PCR and labelled. They are then hybridized to a microarray after which image and data analysis is performed.

The microarray used contains fragments of the full human genome and the locations of the labelled sequences can be further analysed using a variety of techniques, including bioinformatics.

1.6.2. DamID

The DamID technique relies upon the low level in vivo expression of a fusion protein consisting of a protein of interested fused to Dam (DNA adenine methylation) isolated from Escherichia coli. Dam methylates adenines in the sequence GATC, a trait unique to this prokaryotic protein. When the Dam fusion protein is bound to DNA the Dam will methylate nearby GATC sequences (Figure 1.22).

Once Dam methylation has occurred the transcription factor binding sites can be isolated. Genomic DNA is isolated and methylated GATC sequences are cut between GAme and TC nucleotides using the methylation-sensitive restriction enzyme DpnI. This digestion step creates a pool of blunt-ended DNA fragments that have 5’ TC and
3′GA\(^{inc}\) sequences. A double stranded adapter oligonucleotide with a 32 bp 5′ overhang (dsAdR; formed using oligos AdRt and AdRb) is ligated to the blunt ends 5′ overhang ensuring directional ligation. Double strand adapter is used as double stranded DNA is ligated more efficiently than single stranded oligonucleotides. This is followed by DpnII digestion, a methylation sensitive enzyme that only cuts unmethylated GATCs. A PCR primer identical to the 15 most 3′ nucleotides of the AdRt nucleotide sequence and the 5′TC nucleotides of the gDNA fragment is used to amplify adaptor-ligated sequences (Vogel et al, 2007).

**Figure 1.22.** The methylation of GATC sequences of gDNA adjacent to Dam fusion protein binding sites.

### 1.6.3. Bioinformatics

Bioinformatics encompasses a wide range of techniques that utilise statistics, mathematics and IT to analyse biological information, such as genetic sequences. The internet is the backbone to a variety of bioinformatical tools and following the
completion of the HGP (human genome project) a range of information can be obtained from just one DNA sequence.

1.6.3.1 MotifViz

MotifViz is an interactive web server which serves to identify overrepresented known transcription factor binding motifs (Fu et al, 2004). The identification of short sequence motifs, such as transcription factor binding sites, in a set of given sequences may aid in the understanding of transcriptional regulation. Overrepresented regulatory motifs in promoter elements of genes often reflect the factors to which the promoter is responsive to. The MotifViz programme uses algorithms to identify these overrepresented motifs.

MotifViz offers four algorithms to analyse genomic sequences; Rover, Clover, Motifish and Possum (Fu et al, 2004). These four algorithms use different methods to find the overrepresented motifs.

According to Fu et al, Clover uses a thermodynamic model with permutation or background input for statistical evaluation of cis-element over-representation. Clover screens a set of DNA sequences against a precompiled library of motifs and assesses statistical over- or under-representation in the sequences (Frith et al, 2004). One of the precompiled libraries that can be used is called JASPAR, which is a database of eukaryotic transcription factor binding sites (Sandelin et al, 2004). The JASPAR profiles contain non-redundant transcription factor binding sites derived from published collections of experimentally defined sites (Sandelin et al, 2004). JASPAR profiles cover a variety of transcription factor classes; including the nuclear receptors. The MotifViz algorithms, including Clover, have the option of using the JASPER profiles
when searching for over-represented sequences. Therefore, searches can be performed to isolate any one of the JASPAR profiles, for example Nuclear Receptors. Utilizing the Clover algorithm with the JASPER nuclear receptor profile enables identification of overrepresented nuclear receptor binding sites in the genome sequences of interest.

1.6.3.2. MEME

The MEME Suite comprises of a set of Motif-based sequence analysis tools, which includes MEME, TOMTOM and GOMO. The MEME Suite enables elucidation of over-represented sequence motifs in a set of given sequences using MEME (Bailey and Elkan, 2004). MEME analyses a set of given sequences and using a defined algorithm. The algorithm discovers one or more motifs applying a technique known as “expectation maximization to fit a two-component finite mixture model” to the set of sequences (Bailey and Elkan, 2004). The motifs discovered by MEME (an example displayed below) can then be analysed further using other MEME Suite programmes, such as TOMTOM.

TOMTOM uses the MEME motifs to search for transcription factor binding motifs within them (Figure 1.23.).
Figure 1.23. A: MEME motif, B: TOMTOM Oestrogen Receptor binding motif, C: TOMTOM

Oestrogen Receptor binding motif found in MEME motif (A)
Critical cellular processes rely upon protein DNA interactions. Protein DNA binding may trigger transcription initiation or repression, DNA repair or may alter the structural conformation of chromatin an enable, or hinder, the binding of other proteins to the DNA. Elucidation of the regions of DNA where specific proteins bind to could enable us to gain a greater understanding of the proteins being investigated and also of the genes they may regulate.

Oestrogen mediates its effects through the activation of its respective nuclear receptor. The activated receptor then has the ability to translocate to the nucleus and regulate the expression of genes controlled by oestrogen. The genes regulated by oestrogen are determined by the cell type and the conformation of the chromatin.

Initially, the aim of this project was to investigate a reporter library of oestrogen responsive elements (EREs) (Jennings, 2009). This library was constructed by Jennings (Jennings, 2009) using DamID technology. Investigations on this library were to focus on the identification of SERM responders and perform further tests to analyse the isolated reporters. This would provide a deeper understanding into the role of Tamoxifen as an oestrogen antagonist. An overview of the aims are listed below:

- To isolate SERM responsive reporter plasmids from an ERE reporter library
- To assess the ability of the isolated sequences to mediate transcription in conjunction with other nuclear receptors and coactivators to establish potential novel interactions
To investigate the mechanism by which SERMs have the ability to activate certain genes via the oestrogen receptor using mutant oestrogen receptor expression plasmids.

The extent of gene activation by oestrogen is also controlled by coactivators and corepressors. Coactivators, such as the SRC family of coactivators, bind to DNA bound nuclear receptor ligand complexes and further increase the rate of gene transcription.

Due to interesting results obtained during initial investigations into the SERM responders, it was decided to perform novel investigations into the SRC family of coactivators. The aims of the investigation into the SRC family of coactivators is described below.

- To isolate sequences from the genome that associate with SRC-1A:Dam and SRC-3 in association with either ligand bound oestrogen or progesterone receptors
- To use bioinformatics to identify DNA cis elements, repetitive sequences and overrepresented transcription factor binding sites in these isolated sequences
- To assess the ability of the isolated sequences to mediate transcription in conjunction with other nuclear receptors and coactivators to establish potential novel interactions

Performing these investigations would enable a greater understanding in the relationship between the effect of the coactivator molecules SRC-1A and SRC-3 and oestrogen and progesterone receptor mediated transcription.

An additional aim, encompassing reporter plasmids from both the ERE, SRC-1A and SRC-3 libraries, was to attempt to identify the location of transcription start sites in
selected isolated reporter plasmid sequences. The ability to obtain this information would enable a deeper understanding of the binding sites that a variety of transcription factors of transcriptional coregulators associate with.
3.0. MATERIALS AND METHODS

3.1. Investigation of ERE (oestrogen responsive element)

**Reporter Library**

Construction of the ERE library was performed previously using DamID (Jennings, 2009). A hERα:Dam expression plasmid was constructed to enable the expression of an ERα:Dam fusion protein when transfected into a cell line.

EREs were isolated under three conditions; those that interact with ERα:Dam and EtOH (so to isolate EREs that associate with ERα not bound to ligand), those that interact with ERα:Dam and Oestrogen (to isolate EREs that associate with oestrogen bound ERα) and those that interact with ERα:Dam and Tamoxifen (to isolate EREs that associate with tamoxifen bound ERα). Transfection of pCR3.1 ERα:Dam to HeLa cells was performed and cells were treated with above mentioned treatments (EtOH, oestrogen or tamoxifen) to generate three separate libraries of sequences. The Dam portion of ERα:Dam, when DNA bound, methylated GATC sites adjacent to the binding sites. This was followed by DpnI and DpnII digestion to isolate the methylated GATC sequences, adapters were ligated to these isolated sequences and LM-PCR was performed. This was followed by insertion of the amplified sequences to a TOPO library where analysis was performed including sequencing of the inserts. Further analysis was enabled by transferring the TOPO library insert to a pGL3 Basic library. Upon the transfer trans-activation assays were performed to gain information of the promoter activity of the sequences.
Detailed information on the materials and methods used to generate these ERE reporter libraries can be found in Jennings, 2009.

3.1.1. Transactivation Assays

Transactivation assays were used to further analyse the ERE reporter library. The methods for transactivation assays are described below, the example provided is for the investigation of ERα and oestrogen on the ERE reporter library using HeLa Cells.

3.1.1.1. HeLa Cell Preparation for Transfection

HeLa Cells (Baylor College of Medicine, Houston, Texas) were plated in 24 well plates (x10), with 50,000 cells/well in DMEM (with 4.5g/L D-Glucose and L-Glutamine, without Sodium Pyruvate or Phenol Red) with 5% Charcoal Stripped Serum (Gibco) and grown overnight at 37°C in 5% CO₂. All initial cell cultures were prepared by the BCM cell culture core. Prior to transfection cells were checked, using Inverted Microscope, to assess cell health, growth and distribution.

3.1.1.2. Plasmid Preparation for Oestrogen Transfection

Plasmid premixes were prepared as follows. Ependorfs were labelled appropriately with 24 tubes for test (with reporter plasmid, hormone receptor and hormone) and 24 for controls (reporter plasmid only). Reporter plasmid was added to each of the control tubes at a concentration of 200ng/well while 200ng/well reporter plasmid was added to
the test ependorfs as well as 5ng/well hERα. These were left to incubate at 4°C overnight.

A 1:50 lipofectamine dilution was prepared using DMEM (with 4.5g/L D-Glucose and L-Glutamine, without Sodium Pyruvate or Phenol Red) without 5% Charcoal Stripped serum (Gibco). To the test samples 200μl was added and 50μl to the control. These were then left to incubate at RT °C for 30 mins.

3.1.1.3. Hela Cell Transfection

Hela Cell cultures were removed from the incubator and placed in a Laminar Flow Hood where the medium was aspirated and replaced with 200μl DMEM (same as used in premixes). Each well also received 50μl of either test or control reporter plasmid premix (Figure 3.1. and Figure 3.2.).

![Figure 3.1. Transfection protocol: Plate organisation – Test plate](image-url)

Figure 3.1. Transfection protocol: Plate organisation – Test plate
Plates were placed on a plate rocker for 10 mins and then moved into a 37°C incubator with 5% CO₂ for 4 hrs. During the incubation hormone/EtOH premixes were made using DMEM (with 4.5g/L D-Glucose and without L-Glutamine, Sodium Pyruvate or Phenol Red) with 5% Charcoal Stripped serum. EtOH was made to a final concentration of $10^{-7}$M in 12mls while 24mls E2 was prepared at a concentration of $10^{-9}$M, control wells were given only DMEM with 5% Charcoal Stripped Serum.

![Plate organisation Control plate](image)

**Figure 3.2.** Transfection protocol: Plate organisation Control plate

When the incubation was complete plates were removed from the incubator, medium aspirated and 500μl of the required hormone/EtOH premix (pre-warmed to 37°C) was added to each well (**Figure 3.3.**). Plates were incubated for 24 hrs at 37°C with 5% CO₂.

![Hormone and EtOH Dilutions](image)

**Figure 3.3.** Hormone and EtOH Dilutions
3.1.1.4. Luciferase Assay

If the plasmid contains a sequence with a promoter element and the relevant transcription activators, oxyluciferin will be produced. The amount of oxyluciferin produced is proportional to the strength of the promoter elements (See Section 4.1.1.1. for further details).

**Procedure**

Following 24 hr incubation with hormone/EtOH, luciferase assay was used to assess transcriptional activity in the HeLa cell line.

Medium was removed following the incubation and HeLa cells were lysed using 100μl Luciferase Lysis Buffer (Appendix I) per well. The plates were rocked on a plate rocker for 10 mins. Cell lysate (25μl) from each well was pipetted into 96 well plates. One-Glo Luciferase Assay System (Promega) was used to assess the luciferase activity in each of the wells. The firefly luciferase enzyme (stored away from light at -20°C) was brought to RT°C in darkness. The plates were read using a Luciferase Microplate Luminometer and the results analyzed.

3.2. Construction and Investigation of Library of sequences which interact indirectly to SRC-1A and SRC-3

A library of sequences that interacts with SRC-1A and SRC-3 via ligand bound ERα or PR-B was constructed and investigated as follows.

pCR3.1 SRC-1A:Dam and pCR3.1 SRC-3:Dam fusion plasmids were constructed to enable production of the SRC-1A:Dam or SRC-3:Dam fusion protein upon transfection to cell cultures.

3.2.1.1. Generation of Dam PCR product

Genomic template DNA was obtained from Escherichia Coli bacteria (gDNA previously isolated); 10ng was added to 1μl 10X Accuprime Buffer I (Invitrogen), 2μl 1μM 50:50 primer mix (Figure 3.4.) (Invitrogen), 5.5μl dH2O and 0.25μl Taq DNA Polymerase (Invitrogen). Primers were designed to amplify the full length Dam gene generating a PCR product of 800bps.

**Figure 3.4. Forward and Reverse Primers for Dam PCR**

PCR was performed (Table 3.1.) for 30 cycles. PCR products were stored at 4°C until required.
Table 3.1. PCR Programme Details for amplification of Dam from E. coli gDNA.

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialization</td>
<td>94°C</td>
<td>2 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94°C</td>
<td>15 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>56°C</td>
<td>15 sec</td>
</tr>
<tr>
<td>Elongation</td>
<td>68°C</td>
<td>1 min</td>
</tr>
<tr>
<td>Hold</td>
<td>4°C</td>
<td>infinity</td>
</tr>
</tbody>
</table>

3.2.1.2. Purification of PCR Product

QIAquick PCR Purification Kit (Qiagen) was used to purify the PCR product. QIAquick Kits contain a silica membrane assembly for binding of DNA in high-salt buffer and elution with low-salt buffer or water. The purification procedure removes primers, nucleotides, enzymes, mineral oil, salts, agarose, ethidium bromide, and other impurities from DNA samples. Silica-membrane technology eliminates the problems and inconvenience associated with loose resins and slurries. Specialized binding buffers are optimized for specific applications and promote selective adsorption of DNA molecules within particular size ranges.

The QIAquick system uses a simple bind-wash-elute procedure. Binding buffer (PBI) is added directly to the PCR sample or other enzymatic reaction, after which the mixture is applied to the QIAquick spin column. The binding buffer contains a pH indicator, enabling easy determination of the optimal pH for DNA binding. Nucleic acids adsorb to the silica-gel membrane in the high-salt conditions provided by the buffer while impurities are washed away. Pure DNA is eluted with a small volume of low-salt buffer (EB Buffer).
**Procedure**

The Microcentrifuge protocol was followed and all reagents were from the Qiagen Kit. PBI Buffer (100μl) was added to 10μl PCR product. This was applied to a 2ml QIAquick spin column, inserted into a QIAquick collection tube, centrifuged at 13,000 rpm for 45 secs and the flow through discarded. PE Buffer (75μl) was applied to the column and centrifuged as before. Flow through was discarded and sample centrifuged as before. The column was then placed in a clean eppendorf tube and DNA eluted by adding 50μl EB (Elution Buffer) to the column, incubating for 1 min and centrifuging as before.

3.2.1.3. Restriction Digestion of pCR3.1 SRC-1A:Luc, pCR3.1 SRC-3:Luc and amplified Dam PCR product

Construction of pCR3.1 SRC-1A:Luc and pCR3.1 SRC-3:Luc plasmids have been previously described (Lonard et al, 2004). A vector map can be viewed in **Figure 3.5**.
Figure 3.5. Vector map of pCR3.1. (Invitrogen). Isolated genes for SRC-1A or SRC-3 were inserted upstream of an isolated Luc gene in the highlighted ‘PCR Product’ region of the polylinker sequence (Lonard et al, 2004).

Apal digestions of pCR3.1 SRC-1A:Luc and pCR3.1 SRC-3:Luc were set up as follows: 1μl pCR3.1 SRC-1A:Luc or pCR3.1 SRC-3:Luc (from stock with a concentration of 0.2μg/ml) was added to 1μl NEB Buffer 4 (New England BioLabs), 7μl dH2O and 1μl ApaI (10U/μl) (Invitrogen). The Dam PCR product digestion consisted of 30μl purified Dam PCR product, 3.4μl NEB Buffer 4 (New England BioLabs) and 1μl ApaI (10U/μl) (Invitrogen). These were mixed, briefly spun in a
centrifuge to ensure all liquid was in the base and incubated for 1 hr at 25°C. Both samples were frozen at -30°C until required.

QIAquick PCR Purification Kit (Qiagen) was used to purify the digested plasmid. Microcentrifuge protocol was followed as described in Section 3.2.1.2. DNA was eluted using 30μl EB.

Sall digestions were set up as follows. DNA (previously ApaI digested and purified pCR3.1 SRC-1A:Luc, pCR3.1 SRC-3:Luc or Dam PCR product) was added at a quantity of 30μl to 3.5μl NEB Buffer 3 (New England BioLabs) and 1.5μl Sall (10U/μl) (Invitrogen). These were again mixed, centrifuged briefly to ensure all liquid was at the base and incubated at 37°C for 1 hr.

3.2.1.4. Excision of Luc fragment from pCR3.1 SRC-1A:Luc and pCR3.1 SRC-3:Luc

Luc removal was performed to enable ligation of full length Dam to the plasmid in place of Luc to generate a SRC-1A/SRC-3:Dam fusion construct. To each digestion (Section 3.2.1.3.) 3.5μl 10X DNA loading dye was added. Agarose gel was made to a concentration of 1% with 10% Ethidium Bromide (Sigma) and 1X TAE Buffer poured over when set (detailed in Appendix 1). Each sample was applied to the gel as well as 2μl 1 Kb Plus DNA Ladder (Invitrogen) to assess fragment size and the gel was run for 45 mins at 90V. A UV trans-iluminator (photographic) was used to detect the bands. A UV trans-iluminator (non-photographic) was used to view bands and enable removal. Required bands (pCR3.1 SRC-1A, pCR3.1 SRC-3 and Dam PCR product) are excised
using a sharp blade and purified using QIAquick PCR Clean Up Kit (Qiagen). Bands not required (~1600 bp Luc) are discarded.

3.2.1.5. Isolation and Purification of gel-bound DNA

QIAquick PCR Clean Up Kit (Qiagen) was used to isolate and purify the gel-bound DNA. Gel slices are dissolved in a buffer containing a pH indicator, allowing easy determination of the optimal pH for DNA binding, and the mixture is applied to the QIAquick spin column. Subsequent to this step the kit has the same principle as the QIAquick PCR Purification Kit.

Procedure

Protocol was as follows; 450µl QC buffer was added to the agarose gel bound DNA and heated to 37°C until the agar melted. The solution was then transferred to a QIAquick spin column, inserted into a collection tube and centrifuged at 13,000 rpm for 1 min.

QIAquick PCR Purification Kit (Qiagen) was used to purify the isolated DNA. Microcentrifuge protocol was followed as described in Section 3.2.1.2. DNA was eluted using 30µl EB.

3.2.1.6. Ligation of isolated Dam PCR product with Plasmid DNA

The following reactions were set up to ligate the Dam PCR product to pCR3.1 SRC-1A and pCR3.1 SRC-3. Digested pCR3.1 SRC-1A or pCR3.1 SRC-3 plasmid to a volume of 1µl was mixed with 3µl Dam PCR product, 3µl dH₂O, 2µl T4 DNA Ligase Buffer
(Invitrogen) and 1μl T4 DNA Ligase (Invitrogen). A control reaction was also performed which involved replacing 3μl PCR Dam insert with an additional 3μl dH2O. These were both incubated at RT°C for 1 hr. Following incubation the four samples were heated to 70°C to inactivate enzyme activity.

3.2.1.7. Transformation of Plasmid DNA

Each ligation reaction, from Section 3.2.1.6., was added to 20μl One Shot® TOP10 chemically competent E.Coli (Invitrogen), flicked gently, put back on ice immediately and left for 30 min. The cells were then placed in a 42°C H20 bath for 45 secs and returned to ice. After 1 min on ice 500μl SOC Medium (Invitrogen) was added to each transformation reaction and incubated at 37°C for 1 hr. The samples were centrifuged at 7500 rpm for 1 min, 450μl prewarmed SOC Medium was removed and cell pellets resuspended in the remaining SOC Medium. Each cell suspension was plated Agar plates with 50μg/μl Ampicillin (Appendix I), prewarmed to 37°C and grown overnight at 37°C.

Individual colonies were picked from the pCR3.1 SRC-1A:Dam, the pCR3.1 SRC-3:Dam and grown overnight in 3ml LB Broth with 100μg/μl Ampicillin (Appendix I) on a shaking incubator at 37°C.

3.2.1.8. Isolation of Plasmid DNA

Plasmid DNA was isolated from the cultures using ChargeSwitch Pro Plasmid Mini Prep Kit (Invitrogen).
ChargeSwitch® Technology employs a novel mechanism that combines the strengths of current plasmid prep methods while overcoming the inherent limitations. The use of harsh reagents such as chaotropic salts, ethanol, or organic reagents, and isopropanol precipitation, are replaced with a simple ion exchange procedure. In this method, plasmid is bound to a surface ligand that is positively charged at low pH, and eluted when the pH is raised to 8.5 (Figure 3.6).

Procedure

The LB Broth culture was centrifuged for 10 mins at 7500 rpm, supernatant discarded and pellet resuspended in 250µl chilled Resuspension Buffer premixed with RNase A. This was followed by the addition of 250µl Lysis Buffer and samples were mixed by inversion until the solution became viscous (3 min). Precipitation Buffer (250µl) was added and samples were mixed by inversion until a white precipitate formed. Samples were centrifuged for 1 min at 15000 rpm to pellet cellular debris. Supernatant was transferred onto a ChargeSwitch Protein Mini Prep Column inserted into a collection tube and centrifuged at 15000 rpm for 1 min. The flow though was discarded and 750µl Wash Buffer I added and centrifuged at 15000 rpm for 1 min. Wash Buffer II (250µl) was added to the column followed by the same centrifugation procedure. The column was placed into a clean, labelled Eppendorf tube where 50µl Elution Buffer was added to elute the purified plasmid DNA. This was centrifuged at 15000 rpm and the column was discarded.
3.2.1.9. Plasmid Analysis

The plasmid DNA was digested with BamHI and EcoRI to determine if the plasmid had the correct insert.

Restriction digestions were set up as follows; 7μl purified plasmid DNA, was added to 1μl NEB 4 Buffer (New England BioLabs), 1μl BamHI (10U/μl) (Invitrogen) and 1μl EcoRI (10U/μl) (Invitrogen). This was incubated at RT°C for 1 hr, after which it was run on 1% Agarose gel (section 3.2.1.4.).
3.2.1.10. Amplification of pCR3.1 SRC-1A/SRC-3:Dam Plasmids

E.Coli cells transformed with pCR3.1 SRC-1A:Dam and pCR3.1 SRC-3:Dam plasmid DNA grown in Section 3.2.1.7. were used for further analysis.

100μl of the original culture of E.Coli cells containing pCR3.1 SRC-1A:Dam or pCR3.1 SRC-3:Dam plasmids was transferred to 3ml pre-warmed LB broth with 100μg/μl Ampicillin (Appendix I). These were swirled gently and incubated overnight at 37°C on a shaking incubator.

3.2.1.11. Purification of Plasmid DNA

Highspeed Plasmid Midi Kit (Qiagen) was used to purify the plasmid DNA. Midi-preps are used to obtain what is effectively a permanent supply of a certain plasmid, either for transformation, or for later modification. Additionally, midi-preps remove endotoxins; toxins found in bacteria which may lead to cell cytotoxicity during transfection. Endotoxin-free DNA will improve transfection into sensitive eukaryotic cells and is essential for gene therapy research.

QIAGEN Plasmid Kits clear bacterial lysates by centrifugation. The cleared lysate is then loaded onto the anion-exchange tip where plasmid DNA selectively binds under appropriate low-salt and pH conditions. RNA, proteins, metabolites, and other low-molecular-weight impurities are removed by a medium-salt wash, and ultrapure plasmid DNA is eluted in high-salt buffer. The DNA is concentrated and desalted by isopropanol precipitation and collected by centrifugation.
Procedure

The LB culture containing pCR3.1 SRC-1A:Dam and pCR3.1 SRC-3:Dam plasmids in E.Coli cells was centrifuged at 10,000 rpm for 5 mins after which the supernatant was removed and the cell pellet resuspended in 4mls of Resuspension Buffer PI. P2 Lysis Buffer (4mls) was added, mixed gently and left incubate for 5 mins. P3 (4mls) was added to the sample which was immediately poured into the barrel of the QIAfilter cartridge, and left for 10 mins. Buffer QC (4mls) was placed into HiSpeed Tip (qualification column) and left to flow through to calibrate it. After the incubation the plunger was used to push the cell lysate into the previously calibrated HiSpeed Tip. This was washed twice with 10mls Wash Buffer. DNA was eluted from the column using 5mls Buffer QF and collected in a 10ml tube. DNA was precipitated by adding 5.5mls Isopropanol (Fisher Scientific) to the eluate and this was incubated for 15 mins at RT°C. This was centrifuged for 20 mins at 3500 rpm at 4°C. The supernatant was discarded, 1ml 70% ethanol (Pharmco-Aaper) added, sample was centrifuged for 2 mins at same speed and temperature as previously described, 70% ethanol was aspirated and the pellet was left to dry carefully. The cell pellet was dissolved in 500μl TE buffer (Appendix I).

3.2.1.12. Determination of DNA Concentration

The concentration of DNA was then analyzed in a spectrophotometer using a programme for double stranded DNA.

A 1 in 50 dilution of all 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 entered
into the machine and the absorbance set at 260 nm (A260). The $A_{260}/A_{280}$ ratio was also measured to determine the purity of the DNA contained in the samples.

3.2.1.13. Comparing pCR3.1-SRC1A-Dam and pCR3.1-SRC1A-Luc

Two digests were set up to compare pCR3.1-SRC1A-Dam and pCR3.1-SRC1A-Luc. 7μl pCR3.1-SRC1A-Dam was added to 1μl NEB Buffer 4 (New England BioLabs), 1μl ApaI (10U/μl) (Invitrogen) and 1μl SphI (10U/μl) (Invitrogen). The other digestion contained 1μl pCR3.1-SRC1A-Luc, 7μl dH2O, 1μl NEB 4 (New England BioLabs) and 1μl SphI (10U/μl) (10U/μl) (Invitrogen). These were both incubated for 1 hr at 37°C and run, along with 1 Kb Plus DNA Ladder (Invitrogen), on a 1% Agarose gel with 10% Ethidium Bromide (Sigma) for 30 mins at 90V (Section 2.2.1.4.).

3.2.1.14. Concentration of DNA for Sequencing

Sequencing was used to confirm the presence of Dam in the plasmid and to confirm that it was inserted in the correct direction. A minimum DNA concentration of 0.1μg/μl was required for sequencing. DNA (500μl) was added to 1ml 100% EtOH (Pharmco-Aaper) and 30μl 3M Sodium Acetate pH5.5 (Ambion) and incubated at -20°C for 30 mins. The sample was centrifuged at 15000 rpm for 5 mins and the supernatant aspirated. 200μl 70% EtOH was added and centrifuged at 15000 rpm for 30 secs and the supernatant aspirated. The pellet was left until dry after which it was dissolved in a sufficient volume of TE Buffer pH8.0 (Appendix I) to make approximately 0.2μg/μl DNA concentration. A sample of this along with the reverse Dam primer (Figure 3.4.) was sent to the BCM sequencing core for analysis.
3.2.2. Construction of Library of genomic sequences which interact indirectly with SRC-1A:Dam

This process utilizes the SRC-1A:Dam expression vector (pCR3.1 SRC-1A:Dam) constructed according to protocol described in the Section 2.2.1. to make a genomic library of sequences which interact (indirectly) to SRC-1A:Dam via, either, Oestrogen Receptor α (ERα) or Progesterone Receptor-B (PR-B).

HeLa cell transfections, using the SRC-1A:Dam expression vector co-transfected with ERα or PR-B expression vectors, with subsequent treatment of the corresponding hormone were used to cause in vivo methylation of GATC sequences in genomic DNA. Only GATC sequences next to where the SRC-1A:Dam fusion protein bound would be methylated and could subsequently be isolated and amplified using PCR.

3.2.2.1. Hela Cell Preparation for Transfection

HeLa Cells (Baylor College of Medicine, Houston, Texas) were plated in 6 well plates (x2), with 200,000 cells/well in DMEM (with 4.5g/L D-Glucose and L-Glutamine, without Sodium Pyruvate or Phenol Red) with 5% Charcoal Stripped Serum (Gibco) and grown overnight at 37°C in 5% CO₂. All initial cell cultures were prepared by the BCM cell culture core. Prior to transfection cells were checked, using Inverted Microscope, to assess cell health, growth and distribution.
3.2.2.2. HeLa Cell Transfection

Transfections were carried out in duplicate. Premixes were prepared prior to transfection.

Each pCR3.1 SRC-1A:Dam plasmid and hormone receptor plasmid premix was prepared with enough in each for 4 wells. They were made by adding the appropriate volume of plasmids to separate Eppendorf tubes and 400μl DMEM (with 4.5g/L D-Glucose and L-Glutamine, without Sodium Pyruvate or Phenol Red) without 5% Charcoal Stripped Serum (Gibco) (Table 3.2.).

<table>
<thead>
<tr>
<th>Well</th>
<th>Total Vol</th>
<th>Plasmid – SRC 1A</th>
<th>Hormone Receptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 4</td>
<td>400μl</td>
<td>1000ng pCR3.1 SRC-1A:Dam</td>
<td>100ng pCR3.1 hERα</td>
</tr>
<tr>
<td>2 &amp; 5</td>
<td>400μl</td>
<td>1000ng pCR3.1 SRC-1A:Dam</td>
<td>100ng pCR3.1 hPRB</td>
</tr>
<tr>
<td>3 &amp; 6</td>
<td>400μl</td>
<td>1000ng pCR3.1 SRC-1A:Dam</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3.2. Plasmid Premix Preparation. Column number 1 represents the wells the premix is required for (all in duplicate). Column number 2 represents the total volume of the premix. Column 3 represents the final concentration of plasmid required in the total volume (Column 2). Column 4 represents the final concentration of Hormone Receptor, or the absence of it in the negative control for each well.

Lipofectamine premix was made by adding 12μl Lipofectamine (Invitrogen) to 1,200μl DMEM without 5% Charcoal Stripped Serum (Gibco).

Prior to transfection 400μl Lipofectamine premix was added to each pCR3.1 SRC-1A:Dam plasmid and Hormone Receptor plasmid premix giving a total volume of
800μl. All ependorfs were centrifuged at 7,500 rpm for 10 secs to get the liquid to the bottom and incubated at RT°C for 30 mins.

Cells were removed from incubator and under Laminar Flow Hood medium was aspirated from each well. This was replaced with 200μl of each appropriate premix (Figure 3.7.). Plates were rocked for 30mins and then incubated for 4 hrs at 37°C in 5% CO₂.

![Figure 3.7. HeLa Cell Transfection Layout. Figure demonstrates the final transfection concentration of all expression plasmids used.](image)

3.2.2.3. Hormone Treatment of Transfected Cells

During the incubation hormone/EtOH premixes were prepared in DMEM (with 4.5 g/L D-Glucose and L-Glutamine, without Sodium Pyruvate or Phenol Red) with 5% Charcoal Stripped Serum (Gibco) to a final volume of 10mls for Oestrodiol and Progesterone and 20mls for EtOH at the following concentrations: EtOH 10⁻⁷M, E2 10⁻⁶ M and Progesterone 10⁻⁷M.
Following incubation the media was aspirated from each well in a Laminar Fume Hood and 5mls of the appropriate hormone premix was added to the appropriate wells (Table 3.3). The plates were then incubated at 37°C with 5% CO₂ overnight.

<table>
<thead>
<tr>
<th>Well</th>
<th>Plasmid – SRC 1A</th>
<th>Hormone Receptor</th>
<th>Hormone/Blank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>250ng pCR3.1 SRC 1A Dam</td>
<td>25ng pCR3.1 hERα</td>
<td>EtOH</td>
</tr>
<tr>
<td>4</td>
<td>250ng pCR3.1 SRC 1A Dam</td>
<td>25ng pCR3.1 hERα</td>
<td>E2</td>
</tr>
<tr>
<td>2</td>
<td>250ng pCR3.1 SRC 1A Dam</td>
<td>25ng pCR3.1 hPRB</td>
<td>EtOH</td>
</tr>
<tr>
<td>5</td>
<td>250ng pCR3.1 SRC 1A Dam</td>
<td>25ng pCR3.1 hPRB</td>
<td>Progesterone</td>
</tr>
<tr>
<td>3</td>
<td>250ng pCR3.1 SRC 1A Dam</td>
<td>-</td>
<td>EtOH</td>
</tr>
<tr>
<td>6</td>
<td>250ng pCR3.1 SRC 1A Dam</td>
<td>-</td>
<td>EtOH</td>
</tr>
</tbody>
</table>

*Table 3.3. Addition of Hormones to the transfected HeLa cells.*

3.2.2.4. Isolation of genomic DNA

The gDNA of the HeLa cells should be methylated following the previously described transfections. In principle SRC-1A:Dam fusion protein will have bound indirectly to specific genomic sequences and methylated the nearby GATC sequences. Genomic DNA isolation was performed using an alcohol based precipitation technique.

**Procedure**

HeLa cells were harvested with HeLa Cell Lysis Buffer (Appendix I). To each well 500µl of the Lysis Buffer was added. Duplicate wells were combined to give a total of 5 groups. The lysate from each group was transferred to individual ependorf tubes,
mixed and incubated at 55°C for 2 hrs after which 20µl 5M NaCl (Ambion) was added to each tube.

Phenol:chloroform:isoamyl alcohol extraction was performed, 500µl Phenol:chloroform:isoamyl alcohol (25:24:1) (Invitrogen) was added to each tube, agitated gently for 5 min and centrifuged for 10 min at 10000 rpm. The top phase was removed and placed in a clean ependorf tube. This step was repeated; the top phase was separated and 200µl chloroform (Fisher-Scientific) was added to each tube. The samples were placed in 55°C heating block for 1hr to evaporate all chloroform and the samples refrigerated.

When all chloroform had evaporated 10µg RNase A (Invitrogen) was added to the tubes and they were incubated at 37°C for 1 hr. Phenol:chloroform:isoamyl alcohol extraction was performed again using 200µl Phenol:chloroform:isoamyl alcohol (25:24:1) (Invitrogen), agitating gently for 5 mins and centrifuging for 10 mins at 10000 rpm. The top phases were removed and placed in clean ependorf tubes. DNA was precipitated by adding 1ml 100% EtOH (Pharmco-Aaper) and ‘spooling’ the fine threads of DNA that formed with a pipette tip. The tips were then washed in 70% EtOH (Pharmco-Aaper) for 5 mins and transferred to 100µl TE Buffer (Appendix I). The tips were left in the Buffer overnight at RT°C to enable the DNA to dissolve fully. Following incubation the samples were put in a 55°C water bath to ensure all DNA was dissolved before the tips were removed.
3.2.2.5. Analysis of DNA Concentration

Samples were analysed for DNA concentration using a spectrophotometer (Section 3.2.1.12).

3.2.2.6. DpnI Digestion

The methylated GATC sequences in the isolated gDNA are digested by DpnI, which specifically recognizes methylated GATC sequences, 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**

DpnI digestion was performed by adding 0.5µl DNA from each of the samples into 2µl NEB Buffer 4 (New England BioLabs), 2µl DpnI (New England BioLabs) and the appropriate amount of dH\textsubscript{2}O to make the total reaction volume to 20µl. These were incubated at 37°C for 2 hrs followed by 80°C for 20 mins to inactivated DpnI.

3.2.2.7. Adapter Ligation

The purpose of the adapter ligation step was to ligate double stranded linkers to the DpnI digested DNA. To allow PCR amplification of the DpnI digested sequences, a double stranded adaptor nucleotide needs to be ligated to the blunt ends. The double stranded adaptor 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 of the DNA sequences. A double stranded adaptor was used in this step as the ligation of double stranded DNA is more efficient than that of single stranded DNA (Figure 3.8).

1. **Formation of Adaptor**

   AdRT 5'CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGA 3'
   AdRb 3'GCCGGCTCTC 5'

2. **Adaptor Ligation**

   5'CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGATChWNN
   3'GCCGGCTCTCAGNNNN
   Me

*Figure 3.8. Formation of double stranded adapter and its ligation*

DpnI digested DNA (20µl) was added to 4µl 5X T4 DNA Ligase Buffer (Invitrogen), 2µl ds adaptor AdR (comprising of Oligonucleotide AdR and AdRb, *Figure 3.9.*, details of procedure to construct in Appendix I) (50µM) (Sigma) and 2.9µl T4 DNA Ligase (Invitrogen), giving a total reaction volume of 29µl. A control reaction was also set up using 20µl Undigested DNA from one of the samples rather than the DpnI digested DNA. These were left to incubate for 2 hrs at RT°C and then heated to 80°C for 20 mins.

*Figure 3.9. Double stranded adapters sequences*
3.2.2.8. PCR Amplification of Methylated Sequences

PCR amplification of the methylated sequences with ligated double stranded adaptors is performed using an individual PCR primer (Figure 3.10.) that is identical to the 15 most 3' nucleotides of the AdRt oligonucleotide sequence and the 5' TC nucleotides.

![PCR Primer](image)

**Figure 3.10. Sequence of PCR Primer and Annealing and Extension Step**

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 di-nucleotide located at the ends of the DpnI digested fragments. This ensures that only these fragments are amplified in the PCR thus suppressing amplification of non-specific ligation products, which generally will not have a GA di-nucleotide at their ends.

**Procedure**

PCR (Table 3.4.) was used to amplify isolated sequences, and was set up as follows; 4µl Linkered DNA, 2µl Accuprime Buffer I (Invitrogen), 0.5µl Primer AdR PCR (Invitrogen), 13µl dH2O and 0.5µl Taq DNA Polymerase (Invitrogen).
<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialization</td>
<td>94°C</td>
<td>3 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94°C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>58°C</td>
<td>15 sec</td>
</tr>
<tr>
<td>Elongation</td>
<td>72°C</td>
<td>1 min</td>
</tr>
<tr>
<td>Hold</td>
<td>4°C</td>
<td>infinity</td>
</tr>
</tbody>
</table>

Table 3.4. PCR Programme for amplification of methylated sequences

3.2.2.9. Analysis of PCR products on Agarose Gel

The products were run on 1% agarose gel with 10% Ethidium Bromide (Sigma) at 90V for 45 mins (Section 3.2.1.4.).

3.2.2.10. Purification of PCR Products

PCR products were purified as per protocol described in Section 3.2.1.2. DNA was eluted using 50μl EB Buffer.

3.2.3. Cloning of PCR Products using pCR8/GW/TOPO TA Cloning Kit

The pCR8/GW/TOPO TA Cloning Kit (Invitrogen) was used due to its gateway cloning compatibility. This combines TOPO cloning and gateway technology to facilitate a 5 min, one step cloning of Taq Polymerase amplified PCR products into a plasmid vector.
with greater than 95% efficiency. The resultant clones can then be sequenced and characterized without difficulty. Also, the clones can be transferred from the pCR8/GW/TOPO entry vector to a Gateway destination vector of choice to enable expression of the gene of interest in a number of systems. Figure 3.11. describes the process flow.

**Figure 3.11. Schematic of Topoisomerase Method.** Topoisomerase I recognition sites are prepared on the Gateway Cloning Vector which ligates to the Taq amplified PCR products in a 5 min incubation at RT°C. The resultant ligation then causes the release of the Topoisomerase I.

The pCR8/GW/TOPO vector (Figure 3.12.) 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.

![Diagram](image)

**Figure 3.12. Vector map of pCR8/GW/TOPO**

3.2.3.1. Cloning reaction using pCR8/GW/Topo Vector

DpnI digested DNA (1μl) were mixed with 1μl Salt Soultion (Invitrogen), 3μl dH₂O and 1μl TOPO Vector (Invitrogen). This was mixed and incubated for 15 mins at RT°C and then placed on ice.
3.2.3.2. Transformation of Cloning Reactions

Transformation was performed using 2μl TOPO cloning reaction and adding it to 20μl Chemically Competent E. Coli (Invitrogen) and incubating on ice for 30 mins. This was followed by a 30 sec heat shock at 42°C and then the samples were placed back on ice. SOC medium (250μl) was added and the samples were incubated at 37°C for 1 hr. The bacteria were then spread on labelled Agar plates with 100μg/ml Spectinomycin (Appendix I) pre-warmed to 37°C and incubated at 37°C overnight. Following incubation, one of two steps took place.

Either individual colonies were picked (for analysis) and grown in 3mls LB Broth with 100μg/ml Spectinomycin followed by overnight incubation on agitating incubator at 37°C. Alternatively plates were washed with 5mls LB Broth with 100μg/ml Spectinomycin (Appendix I).

3.2.3.3. Purification of Cloned Plasmids

Charge Switch-Pro Plasmid MiniPrep Kit (Invitrogen) was then used to purify the plasmid DNA (Protocol described in Section 3.2.1.8.).

3.2.3.4. Concentration analysis

The purified plasmids were analysed using the spectrophotometer (Section 3.2.1.12.) to assess the DNA concentration and purity.
3.2.3.5. Plasmid Analysis

The plasmids were analyzed for the presence of an insert by enzyme digestion using the restriction enzyme. The digest was set up as follows; 8μl purified DNA was added to 1μl EcoRI (10U/μl) (Invitrogen) and 1μl NEB Buffer 4 (New England BioLabs).

3.2.3.6. Visualization of Digested Plasmid Products

The reactions containing the digested vectors were run on a 1% Agarose gel with 10% Ethidium Bromide (Sigma) at 90V for 45 mins (Section 3.2.1.4.).

3.2.4. Creation of Expression Vector Containing Isolated Sequences

3.2.4.1. Generation of Destination Vector

In order to functionally test the isolated genomic sequences they had to be transferred from the pCR8/GW/TOPO cloning vector to an adapted destination vector with reporter capabilities. 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 (Figure 3.13.).
The destination vector of choice is the pGL3-Basic vector, a vector 4118bp in length which also contains a luciferase gene and lacks a promoter element. The isolated sequences are shuttled from the entry vector to the multiple cloning site of the pGL3-Basic vector which is located upstream of the Luc gene. If the sequence of interest contains a promoter element it will enable Luc expression when transfected to HeLa cells. It is this feature, plus its Gateway compatibility, that makes this a desirable destination vector.

Initially, the vector linearized by digestion with the restriction enzymes NheI and Hind III, followed by an alkaline phosphatase treatment to remove 5' phosphates, thus decreasing the chance of vector religation. The reading frame shown below is then ligated to the vector following a ligation reaction using fast link ligase.

**Procedure**

(This procedure was performed by Dr. David Lonard and Cormac Jennings)
To enable ligation of the cassettes the pGL3-Basic vector was digested with restriction enzymes NheI and HindIII as follows: 1µl pGL3-Basic, 1µl HindIII (10U/µl) (Invitrogen), 1µl NheI (10U/µl) (Invitrogen) and 7µl NEB Buffer 4 (New England BioLabs).

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

3.2.4.2. Cassette Ligation

(This procedure was performed by Dr. David Lonard and Cormac Jennings)

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, 2µl Gateway Reading Cassette, 1µl Fast Link Ligation Buffer, 0.75µl 10mM ATP, 1µl Fast-Link Ligase and 10.25µl H2O. The ligation reaction was incubated for 1 hr at RT°C.

3.2.4.3. Transformation of One Shot ccdB Survival Competent Cells

(This procedure was performed by Dr. David Lonard and Cormac Jennings)

Following the previous steps, the ligation reaction was transformed into One Shot ccdB Survival Competent Cells. 1µl 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 mins followed by heat shock for 30 secs at 42°C. The vials were then placed back into the ice for another 2 mins. 250μl of SOC medium was added to the vial that was then incubated for 1 hr at 37°C. 25μl 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 30μg/ml chloramphenicol.

3.2.4.4. Plasmid DNA Isolation

(This procedure was performed by Dr. David Lonard and Cormac Jennings)

The plasmid DNA was isolated using the Qiagen ChargeSwitch-Pro Plasmid MiniPrep Kit, protocol according to Section 3.2.1.8. 1.5ml of the overnight bacterial culture was processed.

3.2.5. Transferring the isolated cloned sequences to the adapted Destination Vector

Once the destination vector is prepared, the gene of interested can be easily moved from the entry vector to the destination vector. 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. This reaction is mediated by the LR Clonase mix, consisting of integrase, integration host factor, and the phage excisionase which catalyses the LR recombination reaction. The end result is the excision of the gene of interest from the entry clone, pCR8/GW/TOPO entry vector, and integration of
it into the prepared destination vector, pGL3 Basic Vector, which then becomes the expression clone (Figure 3.14.).

\[\text{PCR8/GW/TOPO Entry Vector (Containing isolated sequences)} \rightarrow \text{Adapted pGL3 Destination Vector} \rightarrow \text{PGL3 Basic Vector, now an expression clone containing the isolated sequences} \rightarrow \text{PCR8/GW/TOPO Donor Vector (now empty)}\]

**Figure 3.14. Transferring the isolated sequences from the entry clone to the destination vector**

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. LR recombination reactions were set up for all pCR8/GW/TOPO vectors.

3.2.5.1. LR Recombination Reaction

The pCR8/GW/TOPO Libraries (50ng) were mixed with 50ng adapted pGL3-Basic Dest (Promega), added to 4μl dH2O and 1μl 5X LR Clonase Reaction Buffer (Invitrogen) and incubated at RT°C for 2 hrs. Following the incubation 0.5μl Proteinase K (2μg/μl) (Roche) was added and incubated at 37°C for a further 10 mins. The
Clonase reaction (1µl) was transformed to 20µl Chemically Competent E. Coli (Invitrogen) and incubated on ice for 30 mins followed by 30 sec heat shock at 42 °C and then placed back on ice. SOC medium (250µl) (Invitrogen) was added and the samples were incubated at 37°C for 1 hr. This was then spread on LB plates with 100µg/ml Ampicillin prewarmed to 37°C and grown overnight in a 37°C incubator.

3.2.5.2. Growth of Library Plasmids

Each individual colony growing on the LB agar plates (Appendix I) was picked (using a pipette tip) and placed in a 10ml Falcon Tube with 5mls LB Broth (100µg/ml Ampicillin) (Appendix I). These were grown overnight at 37°C on an agitating incubator.

3.2.5.3. Plasmid Purification

The Zuppy Plasmid Miniprep Kit (Zymo Research) was used to extract plasmid DNA from 3mls of each culture, the remaining culture was stored, in an air-tight tube, at 4°C until required. The 3ml culture was centrifuged at 7500 rpm for 5 mins, supernatant aspirated and the bacterial cell pellet re-suspended in 600µl of TE buffer (Appendix I). Cells were lysed by the addition of 100µl 7X lysis buffer and mixed by gentle inversion for 2 mins. Neutralization Buffer was chilled to 4°C and 350µl of was added and mixed by inversion until the entire sample turned yellow, this was then centrifuged at 15000 rpm for 4 mins. The supernatant was transferred into a Zymo-Spin™ II column which was placed into a collection tube and centrifuged at 15000 rpm for 15 secs. The supernatant was discarded, 200µl of Endo-Wash Buffer added and the samples were
centrifuged at 15,000 rpm for 30 secs. Zuppy Wash Buffer (400μl) was added and centrifuged at 15,000 rpm for 30 secs. The column was then removed from the collection tube and placed in a clean, labelled eppendorf tube after which 30μl of Zuppy Elution Buffer was added directly to the column matrix, left to stand for 1 min at RT°C and centrifuged at 15,000 rpm for 15 secs.

3.2.5.4. Concentration analysis

The purified plasmids were analysed using the spectrophotometer (Section 3.2.1.12.) to assess the DNA concentration and purity.

3.2.5.5. Sequencing

A selection of the library constructs were sent for sequencing at BCM sequencing facility.

3.2.6. Functional Analysis of newly constructed SRC-1A:Dam Libraries

Libraries were analysed using Luciferase Assays. Details of the methodology used can be found in Materials and Methods Section 3.1.1. Additional information of plasmid concentrations in the luciferase assays performed can be found in Appendix II. Expression plasmids used were constructed previously by David Lonard and are as follows:
pCR3.1 hERα — encodes full length human ERα

pCR3.1 hPR-B — encodes full length human PR-B

pCR3.1 COUP-TFI — encodes full length human COUP TFI

pCR3.1 COUP-TFII — encodes full length human COUP TFII

pCR3.1 RORα — encodes full length human RORα

pCMV5 SNAIL — encodes full length human SNAIL

pCR3.1 PAX4 — encodes full length human PAX4

3.3. Locating Transcription Start Sites

Reporter plasmids were analysed to assess where the transcription start site is located. A modified GeneRacer Kit (Invitrogen) protocol was followed.

3.3.1. Modified Gene Racer Protocol

GeneRacer uses an advanced RACE (rapid amplification of cDNA ends) technique and SuperScript® III RT to improve the efficiency of amplifying full-length 5' and 3' cDNA ends. GeneRacer ensures that only transcripts containing full-length cDNA ends are amplified. The advanced protocol begins at the RNA level by specifically targeting only 5' capped mRNA, subsequent steps result in the removal of the cap which is then replaced with the GeneRacer™ RNA Oligo. During reverse transcription this RNA oligo sequence is incorporated into the cDNA and only cDNA that is completely reverse
transcribed will contain this known sequence. 5' RACE PCR is then performed using the homologous GeneRacer™ primer that is specific to the RNA oligo sequence and a gene-specific primer. The result is amplified DNA that contains the full-length 5' cDNA sequence. In addition to capturing the full-length 5' end of genes with known transcriptional start sites GeneRacer™ captures additional 5' end sequence of genes with unknown start sites.

3.3.1.1. Transfection Protocol

An experiment was performed to locate the position of transcription start sites in the COUP-TF responsive plasmid sequences (P11, P26, P27, P33) as well as 10 reporter plasmids from the hERa library (-2p, -17, T29, T32, F30, T22, T31, T14, -30, E12) and also pERE-E1b-LUC as a control. A modified GeneRacer (Invitrogen) protocol was used to accomplish this.

HeLa Cells (Baylor College of Medicine, Houston, Texas) were plated in 5 x 6 well plates at a density of 200,000 cells per well in DMEM (with 4.5g/L D-Glucose and L-Glutamine, without Sodium Pyruvate or Phenol Red) (Gibco) with 5% Charcoal Stripped Serum and grown overnight at 37°C with 5% CO₂. All initial cell cultures were prepared by the BCM cell culture core. Prior to transfection cells were checked, using Inverted Microscope, to assess cell health, growth and distribution.

Each reporter plasmid was prepared in duplicate; 8μl lipofectamine was added to 1,500ng reporter plasmid and 1.6ml DMEM (with 4.5g/L D-Glucose and L-Glutamine, without Sodium Pyruvate or Phenol Red) (Gibco) without 5% Charcoal Stripped Serum. COUP TFII (20ng) was also added to each COUP TFII responsive reporter plasmid and
20ng pCR3.1 hERα was added to the other 11 reporter plasmid samples. These were incubated at RT°C for 30 mins. Following incubation plates were removed from the incubator and remainder of the protocol was carried out in a Laminar Flow Hood. Medium was aspirated from each of the wells and replaced with 800µl of each reporter plasmid premix (duplicate). The plates were returned to the 37°C with 5% CO₂ incubator and left for 4 hrs. Medium was removed, again in a Laminar Flow Hood, and 2 ml DMEM (with 4.5g/L D-Glucose and without L-Glutamine, Sodium Pyruvate or Phenol Red) (Gibco) with 5% Charcoal Stripped serum was added to each well. The plates were returned to the incubator and left overnight.

3.3.1.2. RNA Isolation

The following day RNA was isolated as follows. HeLa cells were harvested using 750µl TEN Buffer (Appendix I) per well (TEN Buffer added, plates were checked using a microscope to ensure all cells had dislodged from the base of the wells), duplicate wells were combined into a single ependorf tube. Tubes were spun at 7,500 rpm for 20 sec to form a cell pellet. Liquid was aspirated and 250µl Trizol (Invitrogen) was added to 2 tubes at a time and immediately vortexed to resuspend the pellet. Chloroform (50µl) (Fisher Scientific) was added to each tube and vortexed for 30 secs. Tubes were centrifuged at 15,000 rpm for 10 mins. The top phase was removed, avoiding interface or lower layer, and transferred to a clean, labelled ependorf. 125µl Isopropyl alcohol (Fisher Scientific) was added to each tube, vortexed and centrifuged at 12,000 rpm for 5 mins. Liquid was aspirated and pellet was rinsed with 250µl 70% EtOH (Pharmco-Aaper), vortexed and centrifuged at 12,000 rpm for 20 secs. EtOH was
removed and pellet was air dried at RT°C. RNase-free H₂O (50µl) was added to each tube, incubated for 30 mins and vortexed.

3.3.1.3. Concentration analysis

RNA concentration was quantified using a Spectrohotopmeter (Section 3.2.1.12.). Sample T14 had no RNA and so was removed from the remainder of the GeneRacer protocol.

3.3.1.4. Dephosphorylating RNA

This step involves RNA treatment with calf intestinal phosphatise (CIP) to dephosphorylate non-mRNA or truncated mRNA.

Procedure

The RNA samples were combined, using 0.43µl from each of the 14 RNA samples (total 6µl). To this 1µl 10X CIP buffer, 1µl RNaseOut, 1µl CIP and 1µl DEPC (diethylpyrocarbonate) water (all Invitrogen) were added to give a total reaction volume of 10µl. This was mixed gently by pipetting, briefly vortexed and centrifuged for 2 secs to collect the fluid. This was incubated at 50°C for 1 hr. Following incubation it was vortexed briefly and placed on ice.
3.3.1.5. RNA Precipitation

When chilled, 50 µl Trizol was added and vortexed, then 10 µl chloroform was added and the sample was again vortexed. The upper phase was removed and transferred to a clean, labelled eppendorf and 25 µl Isopropyl Alcohol (Fisher Scientific) was added, sample vortexed and placed at -20°C for 30 mins. The sample was then centrifuged at 15,000 rpm for 5 mins and the supernatant aspirated. The pellet was rinsed with 50 µl 70% EtOH (Pharmco-Aaper), vortexed and EtOH aspirated. The pellet was dried at RT°C and resuspended in 7 µl RNase free H2O.

3.3.1.6. Decapping Reaction

A 10 µl decapping reaction was set up using 7 µl of the dephosphorylated RNA, 1 µl 10X TAP Buffer, 1 µl RNaseOut and 1 µl TAP (all Invitrogen). This reaction was mixed gently by pipetting, briefly vortexed, centrifuged to collect the fluid and incubated for 1 hr. Following incubation the sample was placed on ice. RNA precipitation protocol was repeated.

3.3.1.7. Ligation Reaction

Dephosphorylated, decapped RNA (7 µl) was added to the tube (from the GeneRacer Kit) containing pre-aliquoted, lyophilized GeneRacer Oligo (0.25 µg) (Invitrogen). This was mixed gently by pipetting up and down to resuspend the RNA Oligo and centrifuged briefly to collect the fluid. A 5 min incubation at 65°C was performed to relax the RNA secondary structure, this was then placed on ice to chill for 2 min and
centrifuged to collect the fluid. When chilled 1µl 10X Ligase Buffer, 1µl 10mM ATP, 
1µl RNaseOut (40U/µl) and 1µl T4 RNA ligase (5U/µl) (all Invitrogen) were added to 
the tube (total volume of 10µl) and mixed together gently by pipetting. This was 
incubated at 37°C for 1 hr, centrifuged gently and placed on ice. RNA precipitation 
protocol was repeated.

3.3.1.8. Reverse Transcription

The ligated RNA (10µl) was added to 1µl Luc Reverse primer (50µM) (Figure 3.15.) 
(Invitrogen), 1µl dNTP mix and 1µl RNase free H2O (all Invitrogen) and incubated at 
65°C for 5 mins. Then, 4µl 5X First Strand Buffer, 1µl DTT 0.1M, 1µl RNaseOut and 
1µl Superscript III RT°C was added (total reaction volume 10µl) (all Invitrogen), 
centrifuged briefly and incubated at 55°C for 45 mins.

5' CGTATCTCITTACCATGCTATTGCA 3'

Figure 3.15. Luc Reverse primer sequence

To inactivate the enzymes the sample was then placed at 70°C for 15 mins, placed on 
ice for 2 mins and centrifuged briefly to collect the fluid. To the reaction 1µl RNase H 
(2 U) (Invitrogen) was added and incubated at 37°C for 20 mins. This was stored at 
-20°C until ready for PCR.
3.3.1.9. PCR Amplification of cDNA

PCR was performed using 1μl cDNA, 0.5μl Luc Reverse Primer (detailed previously) (10mM) (Invitrogen), 0.5μl 5' GeneRacer Primer (Figure 3.16.), 2.5μl 10X Accuprime Buffer I, 0.25μl Accuprime Taq HIFI and 20.25μl dH2O (total volume 25μl) (all Invitrogen). The PCR programme details are outlined in Table 3.5.

![5' CGACUGGAGCGAGGACACUGA 3'](image)

_Figure 3.16. GeneRacer™ 5' primer_

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialization</td>
<td>94°C</td>
<td>2 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94°C</td>
<td>15 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>62°C</td>
<td>15 sec</td>
</tr>
<tr>
<td>Elongation</td>
<td>68°C</td>
<td>15 sec</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94°C</td>
<td>15 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>56°C</td>
<td>15 sec</td>
</tr>
<tr>
<td>Elongation</td>
<td>68°C</td>
<td>15 sec</td>
</tr>
<tr>
<td>Hold</td>
<td>4°C</td>
<td>infinity</td>
</tr>
</tbody>
</table>

Table 3.5. *PCR Programme for Amplification of cDNA*

3.3.1.10. Analysis of PCR Products

PCR products and 1 Kb Plus Ladder (Invitrogen) were run on 3% MetaPhor Agarose Gel (Appendix I) at 90V for 45 mins (Section 3.2.1.4.).
3.3.1.11. Purification of PCR Products

PCR products were purified using QIAquick PCR Purification Kit (Qiagen) following protocol described previously (Section 3.2.1.2). DNA was eluted from the column using 30μl EB.

3.3.1.12. Amplification of Plasmid

Purified DNA (2μl) was added to 20μl One Shot® TOP10 Chemically Competent E.Coli (Invitrogen), flicked gently, put back on ice immediately and left for 30 min. The cells were then placed in a 42°C H2O bath for 45 secs and returned to ice. After 1 min on ice 500μl SOC Medium (Invitrogen) was added to each transformation reaction and incubated at 37°C for 1 hr. The samples were centrifuged at 7500 rpm for 1 min, 450μl SOC Medium was removed and cell pellets resuspended in the remaining SOC Medium. Each cell suspension was plated on Agar plates with 50 μg/μl Ampicillin (Appendix I), prewarmed to 37°C, and grown overnight at 37°C. Following incubation the plates were observed. Colonies growing were picked, using a pipette tip, and the tip placed in 5mls LB Broth with 50μg/μl Ampicillin (Appendix I). This culture was grown on an agitating incubator at 37°C overnight.

3.3.1.13. Plasmid Purification

Following incubation 3mls of the culture was purified using Zuppy Plasmid MiniPrep Kit following protocol described in Section 3.2.5.3.
3.3.1.14. Analysis of DNA

The eluate was analysed for DNA concentration using a spectrophotometer and diluted to give a concentration of 200ng/μl.

3.3.1.15. Sequencing

The diluted DNA was sent to the BCM sequencing core for sequencing analysis.

3.4 Statistical Analysis

Analysis was conducted using Microsoft Excel. Student’s t-test was used for analysis of transfection data in cell lines.
4.0. Results

This study aimed to examine an oestrogen responsive element (ERE) reporter library generated by DamID to identify and investigate SERM responders. Subsequent to this, DamID was employed to create a library of human genomic regions where SRC-1A associates with ligand bound ERα or PR-B.

4.1. ERE Reporter Library Investigation

The oestrogen receptor binds to EREs to mediate gene expression. An ERE reporter library was generated using the DamID technique described in the Materials and Methods Section, using a ERα:Dam fusion protein (Jennings, 2009). The ERα:Dam fusion protein was used in conjunction with hormone treatment with EtOH (V Library), Oestrogen (E Library) or Tamoxifen (T Library) to generate three separate reporter libraries.

The function of this was to address questions such as; what adjacent transcription factor binding motifs are found co-associated with EREs, and also whether the level of gene expression varied between ERα and ERβ binding.

Initial screening investigations were performed on these libraries (20 reporter plasmids from each library were investigated) to assess the differential response of reporter plasmids in this library to Oestrogen, Tamoxifen and Raloxifene. Tamoxifen and Raloxifene are selective oestrogen responsive elements (SERMs) which are used in the treatment of Oestrogen responsive tumours such as breast. The results of these investigations are described below.
4.1.1. Identification and Investigation of SERM Responders from ERE reporter Libraries

Initial ERE Library construction was previously performed using transfection protocol to transfec an ERα:Dam plasmid to HeLa Cells as described by Jennings (Jennings, 2009). This was followed by cell treatment with Oestradiol, Tamoxifen or Ethanol. The cells were incubated and during this time the ERα:Dam protein formed a nuclear receptor complex with either oestradiol or tamoxifen, this enabled in vivo targeted methylation of GATC sequences close to ERα binding sites. Selective amplification of isolated sequences was performed using LM-PCR, cloned using a TOPO entry vector (pCR8/GW/TOPO) and the sequences sub-cloned to an adapted gateway compatible reporter vector (pGL3-Basic) via LR recombination reaction (Figure 4.1.).

pGL3-Basic was the reporter vector of choice, which contains no promoter element, and the isolated DNA sequence is inserted into a multiple cloning site, upstream from a Luc gene (Figure 4.2.). The absence of a promoter site and the presence of a Luc gene upstream from the insert enables the use of assays to examine the functional activity of the insert.
Figure 4.1. Overview of the construction of the ERE reporter libraries (Jennings, 2009)

Figure 4.2. pGL3-Basic Vector displaying Ampicillin gene for selectivity, the location of where the genomic insert is ligated to the vector and the location of the Luciferase gene, upstream of the genomic insert.
The result of these steps is three separate ERE reporter libraries. Each is obtained from transfecting ERα:Dam to HeLa cells and treating these cells with EtOH, Oestradiol (E2) or Tamoxifen (4HT). These served as the starting point for further investigations described below.

4.1.1.1. Initial Screening of ERE Reporter Plasmid Libraries

Initial screening trans-activation assays were performed on 20 plasmids from each of the three ERE Reporter Libraries (Ethanol, Oestrogen and Tamoxifen), described previously. These assays were used to identify reporter plasmids whose promoter responded more favourably to Tamoxifen (4HT) or Raloxifene (Ral) (SERMs) than to Oestrogen. This response may indicate the presence of a SERM responsive sequence, where Tamoxifen or Raloxifene has the ability to induce gene expression when it associates with the particular region of DNA that is isolated in the reporter plasmid.

The luciferase assay was then used to quantify functional activity for each plasmid. As stated previously, the pGL3 Basic reporter vector does not contain a promoter site and contains a luciferase gene upstream of the insert. Therefore, if the isolated sequence that was inserted into the pGL3 Basic vector contains a functional promoter sequence it will drive the expression of luciferin in the HeLa cells, the amount of oxyluciferin expressed will be proportional to the strength of the promoter (Figure 4.3.). The Luciferase Assay was performed as per Materials and Methods Section 3.1.1.4.
The luciferase assay is based on the fact that luciferin (or fluoroluciferin), when mixed with oxygen and firefly luciferase, produces oxyluciferin (or oxyfluoroluciferin) (Figure 3.4.), levels of which can be quantified using a luminometer.

Figure 4.4. Chemical reaction that occurs during the Luciferase Assay (promega.com)

Figure 4.5. displays the results from initial luciferase assays.
Figure 4.5. Luciferase assay results following co-transfection of pGL3-Basic reporter constructs with pCR3.1 hERa followed by 24hr hormone treatment: 10-7M EtOH (Ethanol), 10-9M E2 (Oestradiol), 10-7M 4HT (Tamoxifen) or 10-7M Ral (Raloxifene).

Nine reporter plasmids were found to display this SERM response (E5p, E2, E73, E78, E82, T4p, T8, T77, T100). All nine were sent to sequencing to confirm they are of genomic origin. However, upon sequencing only two were found to be of genomic origin, E5p and T4p (highlighted in purple boxes in Figure 4.5.b. and Figure 4.5.c.). The plasmids of non-genomic origin did not undergo further investigations.
**Figure 4.6.** displays more detailed results of the transactivation assay of the two genomic SERM responsive plasmids (highlighted in purple boxes in **Figure 4.5.**: T4p and E5p).

![Graph A](image1.png)

**Figure 4.6.** Luciferase assay results following co-transfection of pGL3-Basic reporter T4p or E5p with pCR3.1 hERa followed by 24hr hormone treatment. (Same assay as in **Figure 4.5.- close up of results**)

Oestrogen responsive reporter plasmids were used as a comparison (highlighted in blue boxes in **Figure 4.5.**: E12 and V17). **Figure 4.7.** displays more detailed results of two E2 responsive reporter plasmids, of genomic origin, which were to be used as a comparison.
The oestrogen responsive reporter constructs, E12 and V17 (Figure 4.7.), were used as a comparison to the SERM responsive reporter constructs T4p and E5p (Figure 4.6.). The difference in the two responses can be seen from observation of Figure 4.6. and Figure 4.7. In the SERM responsive plasmids Tamoxifen and Raloxifene response is much greater than that of the Oestrogen response. This is the opposite of the Oestrogen responders. The response observed in the Oestrogen responders is one that would be expected due to the antagonistic activity of Tamoxifen and Raloxifene. Locating these SERM responsive sequences could be important in identifying the factors that result in Tamoxifen resistance during the treatment of breast cancers and also could be used to identify which genes are activated following Tamoxifen binding and in which tissues these SERM responsive sequences are active in.

The two SERM reporter plasmids, T4p and E5p, were investigated further to observe the effects of other transcription factors and ERα mutations on their transcriptional activity.
4.1.1.3. Investigation of Mutant ERα on SERM Responders

Mutant ERα reporter plasmids were employed to assess which sites are required for the SERM responsive plasmids to continue to cause transcription. C205H is a DBD (DNA Binding Domain) mutant, L539A is an AF-2 defective mutant and S118A is an AF-1 defective mutant.

![Graph A](image)

**Figure 4.8.** Luciferase assay results following co-transfection of pGL3-Basic reporter constructs (T4p in Panel A; E5p in Panel B) with pCR3.1 hERα or hERα expression mutants (pCR3.1 C205H, pCR3.1 L539A or pCR3.1 S118A). This was followed by 24hr hormone treatment with 10⁻⁷M EtOH (Ethanol), 10⁻⁷M E2 (Oestradiol), 10⁻⁷M 4HT (Tamoxifen) or 10⁻⁷M Ral (Raloxifene). *p<0.05, **p<0.01 mutant ERα compared to wild-type ERα

A significant reduction in transcriptional activity was observed when mutant ERα was compared to wild-type ERα following corresponding hormone/drug treatment (p<0.05, p<0.01) (i.e.: wild type ERα transfection followed by Raloxifene treatment is compared to C250H transfection followed by Raloxifene). Furthermore, analysis of the wild-type ERα result itself indicated the significance of the SERM responsive plasmids response.
to Raloxifene \((p<0.05)\) and Tamoxifen \((p<0.01)\), in the case of \(T4p\), and of Raloxifene \((p<0.05)\) alone in the case of \(E5p\) compared to \(E2\) treatment.

The results (Figure 4.8.) demonstrate that all ER\(\alpha\) mutant domains tested are required for promoter activity in the SERM responsive reporter plasmids indicating a fully functional ER\(\alpha\) molecule is required for this SERM response to occur.

4.1.1.4. Investigation of Transcription Factors and Co-Factors on SERM Responders

SNAIL and PAX4 are nuclear receptors, while SRC-3 is a nuclear receptor co-activator. These nuclear receptors and nuclear receptor co-activators were used due to information obtained from bioinformatic studies (Lonard, 2009; unpublished observations). PAX4 binding sites were found in a number of the isolated ERE reporter plasmids, as was SNAIL. SRC-3 was investigated due to its proven co-activation of ER\(\alpha\) (Anzick et al, 1997).

Figure 4.9. shows the results of the two SERM responsive reporter plasmids transfected to HeLa cells with pCR3.1 (control; blank plasmid), pCR3.1 \(hER\alpha\), pCR3.1 \(hER\alpha\) with pCMV5 SNAIL, pCR3.1 \(hER\alpha\) with pCR3.1 PAX4 and pCR3.1 \(hER\alpha\) with pCR3.1 SRC-3. Each transfected group of cells were incubated for 4 hrs and then underwent a 24 hr treatment with either EtOH, Oestrogen, Tamoxifen or Raloxifene.

Statistical analysis on this data was investigated as was Figure 4.8.. The results in this case indicate that \(E5p\) and \(T4p\) both result in a significant response to Raloxifene and Tamoxifen when compared to \(E2\) treatment \((p<0.05, p<0.01)\).
Figure 4.9. Luciferase assay results following co-transfection of pGL3-Basic reporter constructs with pCR3.1 (empty vector), pCR3.1 hERα, pCR3.1 hERα and pCMV5 SNAIL, pCR3.1 hERα and pCR3.1 PAX4 or pCR3.1 hERα and pCR3.1 SRC-3. This was followed by 24hr hormone treatment: 10-7M EtOH (Ethanol), 10-9M E2 (Oestradiol), 10-7M 4HT (Tamoxifen) or 10-7M Ral (Raloxifene).*p<0.05, **p<0.01 hERα alone compared to hERα co-transfected with respective nuclear receptor/co-activator

It is apparent from Figure 4.9, that the isolated SERM responsive sequences are activated by the transfection of pCR3.1 hERα alone followed by hormone/drug treatment. pCR3.1 PAX4 and pCMV5 SNAIL exert a significant co-repressive effect on the sequences (p<0.05, p<0.01) and pCR3.1 SRC-3, interestingly, does not co-activate transcription in these sequences, in fact SRC-3 appears to significantly co-repress these SERM responsive reporter plasmids (p<0.05, p<0.01).

The effect of SNAIL and SRC-3 on oestrogen responsive reporter constructs (E12 and V17) was also investigated and results can be seen in Figure 4.10. The repressive SNAIL response seen in SERM responsive reporter plasmids (Figure 4.9) is mimicked
in the oestrogen responsive reporter plasmids (*Figure 4.10.*) with a significant reduction in transcriptional activity observed in both constructs (p<0.05, p<0.01)

The effect of SRC-3 on the oestrogen responders was investigated to identify whether there is differential regulation of SRC-3 on the two groups of reporter constructs. It can be seen (*Figure 4.10.*) that SRC-3's transcriptional coactivator ability is restored in these plasmids, cooperating only with E2 treated cells co-transfected with ERα and SRC-3, not with Raloxifene. These results correlate to previous findings that SRC-3 is a coactivator of ligand bound ERα (Anzick et al, 1997).

*Figure 4.10.* Luciferase assay results following co-transfection of pGL3-Basic oestrogen responsive reporter constructs, V17 and E12, with pCR3.1 hERα and pCR3.1 (empty expression vector) or pCR3.1 hERα and pCR3.1 SRC-3 followed by 24hr hormone treatment: 10-7M EtOH (Ethanol), 10-9M E2 (Oestradiol), 10-7M or 10-7M Ral (Raloxifene). *p<0.05, **p<0.01 hERα alone compared to hERα co-transfected with SNAIL.
Two SERM responsive plasmids were identified from the ERE reporter library constructed by Jennings (Jennings, 2009). Transactivation assays were performed to assess the mechanism by which these SERMs (Tamoxifen and Raloxifene) function to activate, rather than repress, ERα. The initial investigative assay used mutant ERα expression plasmids to determine if transcription could still be activated by a mutant ERα. The results indicated that the DNA-Binding, AF-1 and AF-2 domains were all necessary in order to activate transcription, indicating that the SERMs are interacting with ERα in a similar mechanism to Oestrogen.

Transcription factors and transcriptional coactivators were investigated next to assess if there was any differential response found to SERMs as compared to the oestrogen responsive constructs. A differential SRC-3 response was observed. SRC-3 has been previously reported to be an ERα co-activator. The results obtained indicated that, for the SERM responsive reporter constructs, SRC-3 has the ability to co-repress activities of ERα.

It was decided to investigate the SRC family of ‘co-activators’ further.

4.2. SRC-1A and SRC-3/ERα and PR-B Reporter Library Generation and Investigation

Due to the interesting results obtained from the investigation of SRC-3's effect on SERM responsive ERE reporter constructs it was decided to investigate the SRC family of coactivators in further depth, generating a library of reporter plasmids that bind
indirectly to SRC-1A or SRC-3 via ligand bound ERα or PR-B. SRCs bind to promoter elements in DNA via ligand bound nuclear receptors (Onate et al, 1995). More information on this family of molecules, and coregulators in general, can be found in Section 1.4. Investigations focused on SRC-1A and SRC-3, chosen due to their ability to bind to a range of nuclear receptors, including ERα and PR-B.

The purpose of this study was to generate a reporter plasmid library of genomic sequences which associate, via steroid hormone receptors (ERα or PR-B), with SRC-1A or SRC-3 and further to this, evaluate the sequences ability to regulate transcription. This should enable a greater understanding of the involvement of these two SRC molecules in the regulation and transcription of SRC regulated genes. An overview of the steps involved in this process can be viewed in Figure 4.11.

**Figure 4.11.** Overview of the DamID mediated SRC-1A gDNA interacting sequence isolation process.

Reporter plasmid library construction firstly began with the generation of an SRC-1A (or SRC-3) Dam fusion protein. A pCR3.1 SRC-1A Luc Vector underwent SalI and ApaI digestion (Materials and Methods Section 3.2.1.3), effectively cutting out the Luc gene. The Dam gene was amplified from genomic Escherichia coli by PCR, purification of PCR product and subsequent complimentary restriction digestions (ApaI and SalI) (Materials and Methods Section 3.2.1.1. to 3.2.1.3.). This left corresponding digested ends which during a ligation reaction combining the isolated Dam gene and the SRC-1A vector minus the Luc gene resulted in a vector which was composed of the SRC-1A gene located upstream and next to the Dam gene (Figure 4.12.).

![Diagram](https://via.placeholder.com/150)

*Figure 4.12. The process of constructing pCR3.1 SRC-1A Dam plasmid*
4.2.1.1. Generation of Dam PCR product, pCR3.1 SRC-1A and pCR3.1 SRC-3

The Dam gene was amplified from E. coli by PCR. The Dam forward primer contained a SalI restriction recognition site at the 5' end and the Dam reverse primer contained an ApaI restriction recognition site at the 5' end (Materials and Methods Section 3.2.1.1.). Following PCR amplification and DNA purification (Materials and Methods Section 3.2.1.2.) the PCR products were digested with SalI and ApaI, as were plasmids pCR3.1 SRC-1A:Luc and pCR3.1 SRC-3:Luc (Materials and Methods Section 3.2.1.3.).

Figure 4.13. presents a 1% agarose gel which displays the bands of the isolated Dam PCR products (Lanes 2 and 3) along with digested pCR3.1 SRC-1A:Luc (Lanes 4 and 5) and pCR3.1 SRC-3:Luc (Lanes 6 and 7).

![Figure 4.13. 1% Agarose Gel. Lane 1 is molecular weight marker. Lane 2 and 3 are Dam PCR Products. Lane 4 and 5 are the restriction digested pCR3.1 SRC-1A:Luc plasmids and lane 6 and 7 are digested pCR3.1 SRC-3:Luc plasmids.](image)

The gel demonstrates that both the PCR amplification of Dam and the digestion of pCR3.1 SRC-1A Luc were successful as the Dam PCR product (Lane 2 and 3) was present at the expected size (800 bp) and the SRC-1A:Luc (Lane 4 and 5) vector
backbone is present at 7500 bp with the excised Lec insert at 1600 bp. The gel also shows the unsuccessful digestion of pCR3.1 SRC-3 Luc (Lane 6 and 7) as there is no band at the expected 1600 bp site, which would be indicative of successful restriction digestion and release of Luc fragment.

4.2.1.2. Removal of the Luc fragment

The Luc fragment (1600 bp) was removed from the gel and discarded (Materials and Methods Section 3.2.1.4. and 3.2.1.5.).

4.2.1.3. Ligation of Amplified Dam PCR Product and pCR3.1 SRC-1A/SRC-3

The pCR3.1 SRC-1A plasmid minus the Luc fragment was ligated to the Dam PCR product (Materials and Methods Section 3.2.1.6.). A control ligation was also performed for each plasmid which involved performing a second ligation for each plasmid, excluding the amplified PCR product.

E. coli cells transformed with ligated plasmids were plated on agar containing ampicillin (Materials and Methods Section 3.2.1.7.) and incubated overnight. The results of growth can be seen in Table 4.1..

There was an absence of growth on the pCR3.1 SRC-1A Control plate which was indicative of inefficient restriction digestion (which confirms the lack of a Luc band in Figure 4.11.). A small number of colonies were present on the pCR3.1 SRC-1A Dam plate. The high number of colonies on the pCR3.1 SRC-3 Control plate was an
unexpected result, further analysis was required to confirm ligation was successful in
the case of pCR3.1 SRC-1A:Dam, and to confirm all steps were successful in the case
of pCR3.1 SRC-3:Luc Apal/SalI digestion (as there was no evidence of initial Luc
removal from the pCR3.1 SRC-3:Luc plasmid).

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Colonies Observed on Agar Plate</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCR3.1 SRC-1A Dam</td>
<td>2</td>
</tr>
<tr>
<td>pCR3.1 SRC-1A Control</td>
<td>None</td>
</tr>
<tr>
<td>pCR3.1 SRC-3 Dam</td>
<td>2</td>
</tr>
<tr>
<td>pCR3.1 SRC-3 Control</td>
<td>17</td>
</tr>
</tbody>
</table>

*Table 4.1. Observation of Agar Plate Growth following ligation of amplified Dam fragment
and digested pCR3.1 SRC-1a/SRC-3 plasmid.*

Colonies of interest (two from pCR3.1 SRC-1A:Dam plate and two from pCR3.1 SRC-
3:Dam plate) were picked and cultured (Materials and Methods Section 3.2.1.7.).
Plasmid DNA was isolated as per Materials and Methods Section 3.2.1.8. and analysed.

4.2.1.4. Analysis of Constructed Plasmids

Plasmid DNA was assessed using EcoRI and BamHI enzyme digestions (*Figure 4.14.*)
to determine if the plasmid had the correct insert (Materials and Methods Section 3.1.9.).

It appears that the BamHI and EcoRI digestion did not successfully reveal a Dam insert
as no band at the expected 535 bp was observed.
Digestion analysis was then repeated using BamHI and EcoRI (Materials and Methods Section 3.2.1.9.), and only pCR3.1 SRC-1A Dam showed evidence of containing a Dam fragment at the expected size of 535 bp (Figure 4.15.).

Figure 4.15. 1% Agarose Gel, with products following BamH1 and Apa1 Digestion. Lane 1 is molecular weight marker. Lane 2 is pCR3.1 SRC-1A:Dam. Lane 3 and 4 is pCR3.1 SRC-3:Dam. Dam fragment faintly visible at 535 bp size in lane 2 only, no band visible for pCR3.1 SRC-3:Dam
4.2.1.5. Optimisation of Techniques to Construct pCR3.1 SRC-3:Dam

Due to unsuccessful construction of pCR3.1 SRC-3:Dam adjustments were made to the methods used.

Protocol was repeated from the initial digestion with ApaI and SalI (Materials and Methods Section 3.2.1.3). Electrophoresis was performed on the digested pCR3.1 SRC-3:Luc plasmid (Figure 4.16). There was no evidence of the Luc gene at the expected 1,600 bp site.

Adjustments were made to the original protocol to enable successful construction of pCR3.1 SRC-3:Dam. The ApaI digestion step was adjusted as well as the ligation step. The ApaI digestion was performed with 3µl pCR3.1 SRC-3:Luc, 1µl NEB Buffer 4 (New England BioLabs), 5µl dH₂O and 1µl ApaI (Invitrogen) and the ligation time was extended from 1 hr to 2 hrs.

The ApaI digestion was performed with 3 times the amount of pCR3.1 SRC-3 Luc to provide more plasmid DNA and thus give a higher chance of obtaining sufficient amounts of plasmids, also ligation time was extended from 1 hr to 2 hrs to provide more
time for the PCR DNA Dam to ligate to the plasmid. Results from the gel (Figure 4.17.) confirmed that the Luc gene was cut from the pCR3.1 SRC-3:Luc plasmid.

![Agarose Gel Image]

**Figure 4.17.** 1% Agarose Gel; following Luc excision. Lane 1 is molecular weight marker. Lane 2 and 3 are pCR3.1 SRC-3:Luc. The Luc fragment has been excised.

However, following ligation and transformation to chemically competent E. coli cells and overnight incubation on agar plates neither the test or control plates obtained any growth. A 6 hr ligation at RT°C was performed on the digested pCR3.1 SRC-3 plasmid (no Luc gene) in an attempt to successfully ligate the Dam PCR product to the pCR3.1 SRC-3 plasmid. This also yielded no growth following transfection to chemically competent E. coli cells. Attempts to construct pCR3.1 SRC-3 Dam were suspended.

4.2.1.6. Further Analysis of pCR3.1 SRC-1A:Dam plasmid

The agar plates growing E. coli colonies with pCR3.1 SRC-1A:Dam plasmids were purified and plasmid DNA isolated (Materials and Methods Section 3.2.1.10 and 3.2.1.11). The resulting purified plasmid DNA was assessed using restriction digestion to compare pCR3.1 SRC-1A:Dam and pCR3.1 SRC-1A:Luc using the restriction
enzyme SphI (Materials and Methods Section 3.2.1.13). **Figure 4.18.** displays a 1% agarose gel following electrophoresis of the restriction digestions (of pCR3.1 SRC-1A:Luc and pCR3.1 SRC-1A:Dam) and confirms a difference in the banding pattern of the pCR3.1 SRC-1A:Dam and the pCR3.1 SRC-1A:Luc plasmids.

**Figure 4.18.** 1% Agarose Gel containing SphI restriction digested plasmids. Lane 1 contains molecular weight marker. Lane 2 contains the pCR3.1 SRC-1A:Dam digested plasmid. Lane 3 contains the pCR3.1 SRC-1A:Luc digested plasmid.

Sequencing was used to confirm the directionality of the gene and to confirm that the full length gene for both SRC-1A and Dam are present (Materials and Methods Section 3.2.1.14.).

A minimum DNA concentration of 0.1µg/µl was required for sequencing. Spectrophotometric analysis of the plasmid DNA revealed the concentration of DNA to be 0.068µg/µl with a $\text{A}_{260}/\text{A}_{280}$ ratio of 1.75. DNA had to be concentrated to enable sequencing to be performed (Materials and Methods Section 3.2.1.12.). Following concentration of the plasmid spectrophotometric analysis determined the DNA concentration to be 0.177µg/µl with a $\text{A}_{260}/\text{A}_{280}$ ratio of 1.44. DNA was more concentrated but was not as pure.
Sequencing confirmed presence of Dam downstream and adjacent to the SRC-1A gene (Figure 4.19).

<table>
<thead>
<tr>
<th>Key</th>
<th>Primer</th>
<th>SRC-1A</th>
<th>Dam Sequence</th>
<th>Unknown sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3'ACCGTGGGGTTCATCAAAAGCCAGCAAGTGATCCACCTTTTAGGTCGCTATCCGGCCGTTCGCTGCTTTGATCTACG</td>
<td>GCTGCTATCCAGCAATGCTGTGGTGAACGCTGCATGCCGTGCTTCTTCTCTTCATCCGCTTGCTGCTTTGATCTACG</td>
<td>GCTGCTATCCAGCAATGCTGTGGTGAACGCTGCATGCCGTGCTTCTTCTTCATCCGCTTGCTGCTTTGATCTACG</td>
<td>GCTGCTATCCAGCAATGCTGTGGTGAACGCTGCATGCCGTGCTTCTTCTTCATCCGCTTGCTGCTTTGATCTACG</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4.19.** Plasmid insert sequence displaying correct direction of SRC-1A and Dam genes. Red nucleotides depict the SRC-1A gene, underlined red indicated the location of the primer sequence used, blue indicates the Dam sequences and the black nucleotides are of unknown origin, possibly originating from the plasmid.

Confirmation of the Dam gene insertion downstream of the SRC-1A gene in the plasmid enabled the DamID procedure to begin. The newly created plasmid containing the SRC-1A:Dam fusion gene (Figure 4.20.) was used in HeLa cell transfections, which, upon transfection would express the SRC-1A:Dam fusion protein.
Figure 4.20. Vector map of pCR3.1. (Invitrogen). This figure details the insertion of the newly formed SRC-1A:Dam fusion gene to the 'PCR Product' region of the polylinker sequence. SRC-3:Dam fusion gene was inserted to the same site. The SalI digestion site is located in the end of the SRC-1A and SRC-3 sequence which was already present in the plasmid. The Dam is then linked to that SalI site in the SRC-1A or SRC-3 genes to the Apal site located in the polylinker sequence (highlighted in red).

4.2.2. Construction of Library of genomic sequences which interact indirectly with SRC-1A:Dam via hERα or PR-B

The purpose of this section was to investigate the involvement of SRC-1A in hERα and hPR-B mediated transcription. Figure 4.21. demonstrates the DamID process whereby the SRC-1A:Dam fusion plasmid is transfected to HeLa cells along with an expression
plasmid for hERα or hPR-B. Following 4 hr incubation the cells were treated with either oestradiol, progesterone or ethanol. In vivo binding of the steroid and its receptor occur, followed by translocation to the nucleus where the complex associates with the hormone response elements in the genome and SRC-1A:Dam interacts indirectly with the genome via the receptor ligand complex. When bound, the Dam can then methylate the proximal GATC sequences. The DNA is then isolated and undergoes analysis. UV spectrometry is used to analyse the concentration of the DNA and its purity. The end result of this step is isolated gDNA some of which contains methylated GATC sequences marking where the SRC-1A:Dam associated with the genome via either hERα and E2 or hPR-B and progesterone.
Figure 4.21. Construction of library of sequences found to associate indirectly to SRC-1A:Dam.
4.2.2.1. Transfection of HeLa Cells with SRC-1A:Dam Expression Plasmid

The following steps were performed to enable SRC-1A:Dam expression in HeLa cells, which via indirect binding of SRC-1A to the DNA will permit Dam methylation of nearby GATC sequences.

250ng SRC-1A:Dam was transfected to Hela cells with co-transfection of 25ng hERα or hPR-B expression vectors. The transfected HeLa cells were incubated for 4 hrs after which hormone treatment was performed with Oestradiol or Progesterone for 24 hours. A subset of the transfected cells were also treated with ethanol, which would enable the isolation of SRC-1A:Dam that binds to DNA bound nuclear receptors that bind in the absence of hormone treatment.

A transfection reaction was also performed transfecting SRC-1A:Dam without a nuclear receptor expression vector and treating these transfected cells with EtOH which would act as a control during LM-PCR mediated specific amplification of methylated GATC sequences, as no GATC specific methylation should occur when cells are transfected with SRC-1A:Dam alone. (As per Materials and Methods Section 3.2.2.1. to 3.2.2.3.)

4.2.2.2. Isolation of Genomic DNA

Genomic DNA was isolated from the transfected HeLa cells using an ethanol precipitation technique as per Materials and Methods Section 3.2.2.4 and the DNA concentration and purity analysed using a spectrophotometer as per Materials and Methods Section 3.2.2.5. The results of this can be seen in Table 4.2.
Table 4.2. Spectrophotometric analysis of isolated genomic DNA from HeLa Cells transfected with pCR3.1 SRC-1A:Dam co-transfected with the noted nuclear receptor followed by 16 hr treatment with the noted hormone/EtOH.

<table>
<thead>
<tr>
<th>Transfection &amp; Hormone Treatment</th>
<th>DNA Concentration mg/μl</th>
<th>DNA Purity A(^{260/280})</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERα &amp; EtOH</td>
<td>0.135</td>
<td>1.69</td>
</tr>
<tr>
<td>ERα &amp; E2</td>
<td>0.093</td>
<td>1.64</td>
</tr>
<tr>
<td>PR-B &amp; EtOH</td>
<td>0.047</td>
<td>1.63</td>
</tr>
<tr>
<td>PR-B &amp; Prog</td>
<td>0.030</td>
<td>1.69</td>
</tr>
<tr>
<td>Control</td>
<td>0.008</td>
<td>1.73</td>
</tr>
</tbody>
</table>

4.2.2.3. Specific Amplification of Methylated GATC Sequences

Isolation of the methylated GATC sequences was achieved with the use of the DpnI enzyme, which specifically recognizes methylated GATC sequences, thus generating a pool of blunt-ended DNA fragments with 5’TC and 3’GA\(^\text{methyl}\) (Materials and Methods Section 3.2.2.6).

The DpnI digested partially methylated sequences were amplified using a methylation specific LM-PCR method. The LM-PCR procedure (Materials and Methods Section 3.2.2.7) ensures only methylated DNA fragments are amplified. DpnII digestion, which is typically part of the protocol, was not performed. DpnII destroys (via digestion) all fragments containing unmethylated GATC sequences which excludes them from amplification, however this DpnII digestion this step was omitted from protocol. The activities of DpnI functions to digest only methylated GATC sequences therefore unmethylated GATC are not cut, leaving short DNA sequences with digested ends (Figure 4.22). The unmethylated GATC sequences should remain ‘hidden’ in long DNA sequences and therefore ‘ignored’ during the adapter ligation step. Therefore the
unmethylated GATC sequences will not interfere with the LM-PCR. In this protocol DpnII digestion was not performed.

![DpnI digestion of methylated GATC sequences](image)

**Figure 4.22. DpnI digestion of methylated GATC sequences in isolated gDNA**

Additionally, non specific DNA products which may be derived from random DNA breaks are excluded from the PCR due to the specificity of Primer design. The PCR primer overlaps with the GA dinucleotide located at the ends of the DpnI digested fragments, ensuring only these fragments are amplified thus suppressing amplification of non specific ligation products, which, generally, will not have a GA dinucleotide at their ends.

The LM-PCR reaction (**Figure 4.23.**) also includes a control for DpnI digestion. The control is gDNA from one of the samples which did not undergo DpnI digestion, therefore no adaptor mediated amplification of these methylated sequences should take place as they were never digested and should remain ‘hidden’ in long strands of gDNA (similar to the unmethylated GATC sequences).
Following LM-PCR (Materials and Methods Section 3.2.2.7 and 3.2.2.8) the samples are run on a 1% agarose gel (Materials and Methods Section 3.2.2.9.) to enable visualisation of LM-PCR products (Figure 4.24.).

The PCR products create a smear in the gel indicating that there is a variety of sequences isolated from the transfected HeLa cells. PCR product sequences all appear to be shorter than 500 bp. Transfections treated with EtOH appear to have a larger amount of DNA as the bands appear brighter. Primer Dimers are visible in the gel.

*Figure 4.23. LM-PCR procedure*
Figure 4.24. 1% Agarose Gel electrophoresis of (LM)-PCR products from LM PCR of DNA isolated from the HeLa cells transfected to SRC-1A:Dam with ERα/PR-B transfection followed by treatment with E2/Prog for 16hrs. Genomic DNA was then digested with DpnI, Adapters were ligated to the blunt ends of the isolated sequences and LM PCR was performed.

Lane 1 (Figure 4.24.) contains molecular weight marker. Lane 6 contains the control, that is a control reaction set up during the adapter ligation stage, this a control run as per the tests but with undigested (not DpnI digested) genomic DNA from one of the samples of isolated genomic DNA from the transfected HeLa cells.

Lane 2 (Figure 4.24.) contains LM-PCR generated PCR products resulting from the amplification of GATC methylated sequences following the transfection of HeLa cells with the SRC-1A:Dam expression vector, co transfected with an expression vector for ERα followed by treatment with ethanol. The lane shows LM-PCR products ranging in size from approximately 100 bp to 500 bp. These LM-PCR products show that not only is the fusion product able to indirectly bind to DNA but the Dam fusion protein of the
SRC-1A:Dam molecule is functioning and can cause GATC methylation. Even though the SRC-1A:Dam and ERα transfected HeLa cells were treated with ethanol there are still PCR products. It is known that ERα had the ability to bind DNA in the absence of hormone.

Lane 3 (Figure 4.24.) contains LM-PCR generated products resulting from the amplification of DNA containing methylated GATC sequences following the transfection of HeLa cells with the SRC-1A:Dam expression plasmid and cotransfection with ERα with subsequent treatment with oestradiol. PCR products in this lane range in size, similarly to Lane 2, from 100 bp to 500 bp. However there is much less LM-PCR product in this lane when compared to Lane 2. This could indicate that SRC-1A binds preferentially to ERα in the absence of Estradiol, whereas the ERα agonist, Oestradiol, enhances ERα genomic binding. Basically, there was less SRC-1A:Dam indirect DNA association during the transfection conditions of Lane 3 when compared to Lane 2.

Lane 4 and 5 (Figure 4.24.) contains LM-PCR generated products resulting from the amplification of DNA following the transfection of HeLa cells with the SRC-1A:Dam expression plasmid with cotransfection of PR-B followed by treatment with ethanol (Lane 4) or Progesterone (Lane 5). Both of these lanes contain less LM-PCR products when compared to Lanes 2 and 3 respectively, indicating that the SRC-1A:Dam transfections with cotransfection of ERα was more successful than that with SRC-1A:Dam and PR-B. Lane 5 contains the least amount of LM PCR product with LM-PCR product size ranging from approximately 100 bp to 400 bp while lane 4 product sizes range from 100 bp to 500 bp.

Lane 6 contains the DpnI undigested genomic DNA following LM-PCR protocol. The absence of a banding pattern is due to the fact that the DNA template in this case is
gDNA from HeLa cells that have been transfected with the SRC-1A:Dam expression plasmid, cotransfected with PR-B, and treated with progesterone but DpnI digestion was not performed. As no DpnI digestion has taken place, the methylated GATC sequences are not isolated and thus double stranded adapters can not ligate and LM-PCR cannot occur. Additionally, the primer dimer band at 100bp is the brightest of all lanes indicating no primers have been used in the LM-PCR procedure that was performed on the contents of this lane.

A faint band is visible at approximately 100 bp in Lanes 2 through 6 (strongest in Lane 6). This is a primer dimer band, which result when primers annealing to each other forming a dimer, due to the low concentration of target amplicons.

4.2.3. Cloning of PCR Products using pCR8/GW/TOPO TA Cloning

The isolated, amplified SRC-1A:Dam sequences were then sub-cloned into a gateway entry vector. pCR8/GW/TOPO was the vector of choice, chosen due to its function as a gateway entry vector (vector map can be seen in Materials and Methods Figure 3.13.). This vector contains a TOPO cloning site which can efficiently clone Taq amplified PCR products. Taq polymerase has non-template dependent terminal transferase activity that adds a single deoxyadenosine (A) to 3' ends of PCR products. The pCR8/GW/TOPO vector has overhanging deoxythymidine (T) residues which enables efficient ligation of the Taq amplified LM-PCR products into the TOPO vector (Figure 4.25.).
Figure 4.25. Process involved in inserting the sequences to the TOPO gateway vector and then to the specially designed destination vector.

The TOPO cloning reactions were set up as per Materials and Methods Section 3.2.3.1.

Sequences isolated from transformation:

<table>
<thead>
<tr>
<th>No. Colonies</th>
<th>Denoted</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRC-1A:Dam cotransfected with ERα, treated with EtOH</td>
<td>2</td>
</tr>
<tr>
<td>SRC-1A:Dam cotransfected with ERα, treated with oestradiol</td>
<td>None</td>
</tr>
<tr>
<td>SRC-1A:Dam cotransfected with PR-B, treated with EtOH</td>
<td>2</td>
</tr>
<tr>
<td>SRC-1A:Dam cotransfected with PR-B, treated with progesterone</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.3. Table displaying the number of colonies following overnight growth of Chemically Competent E. coli cells transformed with TOPO cloning reaction products.
The products of the cloning reactions were transformed into chemically competent E. coli according to Materials and Methods Section 3.2.3.2 and plated onto spectinomycin plates for overnight growth. The plates were examined for the presence of colonies (*Table 4.3.*).

Colonies that grew were individually picked and grown (to increase plasmid containing bacterial numbers) in LB Broth containing 100μg/ml Spectinomycin to select for the TOPO Cloning Vector which contains a spectinomycin resistance gene. The cultures were purified using a Mini-Prep system and the purified DNA was analysed using spectrophotometry (Materials and Methods Section 3.2.3.3. and 3.2.3.4.). The results can be seen in *Table 4.4.*.

Isolated plasmids were then screened for the presence of inserts by restriction digestion.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Transfection &amp; Hormone Treatment</th>
<th>DNA Concentration μg/μl</th>
<th>DNA Purity A&lt;sub&gt;260/280&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>ERα &amp; EtOH</td>
<td>0.079</td>
<td>1.81</td>
</tr>
<tr>
<td>1b</td>
<td>ERα &amp; EtOH</td>
<td>0.176</td>
<td>1.74</td>
</tr>
<tr>
<td>3a</td>
<td>PR-B &amp; EtOH</td>
<td>0.171</td>
<td>1.85</td>
</tr>
<tr>
<td>3b</td>
<td>PR-B &amp; EtOH</td>
<td>0.142</td>
<td>1.88</td>
</tr>
<tr>
<td>4a</td>
<td>PR-B &amp; Prog</td>
<td>0.107</td>
<td>1.94</td>
</tr>
</tbody>
</table>

*Table 4.4. Results of spectrophotometric analysis of TOPO Libraries.*

EcoRI restriction digestion sites flank the TOPO cloning sites in the pCR8/GW/TOPO cloning vector (*Figure 4.26.*). Restriction digestion of the plasmids using EcoRI enables further assessments of the insert (Materials and Methods Section 3.2.3.5.).
Figure 4.26. EcoRI restriction digestion sites on either site of the LM-PCR insert enable analysis of the inserts.

The restriction digested plasmids were then run on a 1% agarose gel (Materials and Methods Section 3.2.3.6.) to enable visualization of the presence of inserts (Figure 4.27.).

Due to the lack of colonies obtained using this protocol, the procedure was optimized.

To increase the number of TOPO inserts a reduced quantity of LM-PCR product was added and the volume replaced with dH₂O (1μl PCR Product and 3μl dH₂O rather than 4μl PCR Product). As a result of this adjustment to the TOPO cloning procedure approximately 2000 colonies were obtained on each of the 4 agar plates (ERα and EtOH, ERα and E2, PR-B and EtOH or PR-B and Prog).
Figure 4.27. 1% Agarose Gel containing a restriction digestion of TOPO vectors containing LM-PCR inserts with EcoRI. This step was to screen for colony inserts. Lane 1 contains molecular weight marker. The remaining lane contents are described above the gel.

Each plate was washed with LB Broth containing 100μg/ml Spectinomycin (Materials and Methods 3.2.3.2.), plasmid DNA was purified, isolated, and concentration was determined using a spectrophotometer (Materials and Methods Section 3.2.3.3 and 3.2.3.4).

4.2.4. Creation of Expression Vector Containing Isolated Sequences

Sequences needed to be cloned into a reporter vector to enable functional investigation of the promoter activity of the sequence that was isolated. As pCR8/GW/TOPO (Figure 3.11.) 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. In order to do this a gateway vector
conversion system with One-shot ccdB Survival competent was used to convert the reporter vector of choice (pGL3-Basic Vector) into a gateway vector (*Figure 3.12.*) according to Materials and Methods Section 3.2.4.1. to 3.2.4.4.

Upon successful generation of the pGL3-Basic destination vector, a LR Clonase reaction was performed on the TOPO Cloning reaction (Materials and Methods Section 3.2.5.1.). 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. The pGL3-Basic reporter vector was manipulated to enable its function as gateway destination vector.

Following the LR recombination reaction the products were transformed into competent bacteria and plated on LB Agar plates with 100μg/ml Ampicillin to select for the presence of the pGL3-Basic Dest vector (Materials and Methods Section 3.2.5.2.). A record of colony growth can be observed in *Table 4.5.*

<table>
<thead>
<tr>
<th>Plate No.</th>
<th>Initial Transfection</th>
<th>No. of Colonies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>From HeLa cells transfected with SRC-1A:Dam, cotransfected with expression plasmid for ERα and treated with Ethanol</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>From HeLa cells transfected with SRC-1A:Dam, cotransfected with expression plasmid for ERα and treated with E2</td>
<td>60</td>
</tr>
<tr>
<td>3</td>
<td>From HeLa cells transfected with SRC-1A:Dam, cotransfected with expression plasmid for PR-B and treated with Ethanol</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>From HeLa cells transfected with SRC-1A:Dam, cotransfected with expression plasmid for PR-B and treated with Progesterone</td>
<td>63</td>
</tr>
</tbody>
</table>

*Table 4.5. Resultant colonies containing pGL3-Basic plasmids following E. coli transfection*
Colonies were individually picked from plates 2 (E1-E60) and 4 (P1-P60), grown overnight in LB Broth containing 100μg/ml Ampicillin which was followed by mini prep (Materials and Methods Section 3.2.5.3.) to isolate and purify the pGL3 Basic reporter plasmids (Figure 4.28.).

Figure 4.28. Isolation and preparation of pGL3 Basic Colonies

Spectrophotometric analysis was performed to analyse DNA concentration and purity (Materials and Methods Section 3.2.5.4.). Once the concentration of each plasmid solution was identified they were diluted with TE Buffer to give a final concentration of 0.2μg/μl.
4.2.5. Functional Analysis of Newly Constructed Expression Vectors

The isolated sequences are now contained in a newly generated pGL3-Basic vector library. As the sequences are now in a reporter vector, it is possible to identify if the isolated sequences are able to bind ERα or PR-B, and if the addition of WT SRC-1A will result in increased or decreased transactivational activity.

To determine the functionality of the individual sequence contained in the pGL3-Basic reporter vectors trans-activation assays were performed using HeLa cells transfected with individual pGL3-Basic Luc reporter constructs and wild type hERα or hPR-B expression vectors. Four hours post transfection cells were treated with Oestrogen (for WT hERα transfected cells), Progesterone (for WT hPR-B transfected cells), EtOH or no hormone treatment and cells were incubated for a further 24 hrs. (Materials and Methods Section 3.1.1.4.). A luciferase assay (see Section 4.1.1.1. for more details of Luciferase assay) was then used to measure the functional activity of each plasmid.

4.2.5.1. Initial Transactivation Assays

The initial aim of this procedure was to screen the pGL3-Basic Libraries for plasmids for functional genomic inserts. Results obtained are displayed in Figure 4.29.
Figure 4.29. Luciferase Assay results following HeLa cell transfection with 200ng reporter plasmid (E1-E23; P1-P24) with cotransfection of 5ng hERα expression vector for the reporter plasmid library of sequences that associate with hERα and SRC-1A and 10ng hPR-B expression vector for the reporter plasmid library of sequences that associate with hPR-B. Following a 4 hr incubation the transfections were treated with either $10^{-9}$M Oestrogen or $10^{-7}$M Progesterone for 24 hrs. A subset of reporter plasmids were treated with EtOH and another subset received no hormone treatment (NHT) as controls. Graph A displays the results from the hERα transfection, while Graph B is from hPR-B transfection.

The function of these initial assays was to separate those reporter plasmids that had genomic sequences that indirectly bound to SRC-1A via hERα and Oestrogen or hPR-B and Progesterone from those that did not (ie: vector or bacterial sequences).

Theoretically the genomic sequences isolated should respond to hERα and Oestrogen or hPR-B and Progesterone if a relevant corresponding transcription start site is present.
The graphs show no differential increase of transcriptional activity in any of the plasmids, this can be seen in both of the graphs displayed above where expression with EtOH and Hormone remain similar throughout all of the plasmids. If the reporter plasmids had a sequence that had a promoter in it in which SRC-1A, hERa and Oestrogen bound to during the DamID assay then a selective activity would be seen.

Lack of activity could mean a number of things; that the sequences that are inserted in the plasmid do not contain a promoter and therefore Luc could not be transcribed, that the sequences have some other effect on transcription or that the sites isolated bind to SRC-1A in association with other proteins and not specifically the nuclear receptor and steroid that was added (endogenous transcription factors).

Sequencing was performed to differentiate genomic from non genomic sequences and to enable further investigation into these sequences.

4.2.6. Sequencing of Reporter Library Inserts

Sequencing was initially performed on a select number of reporter plasmids from each library. It was then determined that the PR-B reporter plasmid library worked at a higher efficiency than the ERα reporter plasmid library (i.e. there were more genomic sequences in the PR-B library) and so further sequencing analysis focused on the PR-B library.

Sequences were analysed using NCBI BLAST (National Centre for Biotechnology Information, Basic Local Alignment Search Tool, http://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequences identified as containing genes that encode for Progesterone Receptor, SRC-1A or Globin were excluded as they were
considered to be non-genomic (sequences not from sites that associated with SRC-1A when the library was being constructed). Sequencing was performed on the PR-B library until twenty one genomic sequences were obtained (P1, P2, P4, P10, P11, P15, P19, P20, P21, P22, P23, P24, P26, P27, P32, P33, P36, P37, P38, P40, P45). Sequences can be viewed in *Appendix III.*

4.2.7. **Further Analysis of Genomic PR-B Library**

A further trans-activation assay was performed on the twenty one selected PR-B plasmids using HeLa cell transfections, co-transfecting 21 different *hPR-B* reporter plasmid sequences and expression vectors for *hERα, hPR-B, hGRα* simultaneously with and without SRC-1A. Four hrs post transfection the cells were treated with a combination of Oestrogen, Progesterone and Dexamethasone. pGRE-E1b-Luc was used as a positive control to confirm that the cells were able to withstand such a high level of continued plasmid DNA and hormone treatment.

*Figure 4.30. and Figure 4.31.* display the results. The results show a similar pattern to that seen in previous transactivation assays. The control, pGRE-E1b-Luc, displayed a 15 fold increase in luciferase activity following the addition of the hormone mix; none of the reporter plasmids displayed the same, or even a similar result.
Figure 4.30. Luciferase assay results. HeLa cells were cotransfected with 50ng pCR3.1 (Empty Vector), 200ng SRC-1A:Dam PR-B reporter plasmids and expression vectors for hERα(5ng), PR-B (10ng) and GR(10ng). Four hrs post transfection cells were treated with a hormone mix of E2(10^{-9}M), Prog (10^{-7}M) and Dex(10^{-7}M) and incubated for 24 hrs.

The results of this trans-activation assay proved that the SRC-1A Dam protein did not bind to the genome in association with either hERα and Oestrogen or hPR-B and Progesterone during library construction. Results also confirm that the methods used were functional as pGRE E1b Luc (a reporter plasmid containing a sequence responsive
to GR). It was then necessary to go through other channels to decide which direction to go to next.

4.2.8. Bioinformatics

Genomic sequences were analysed using Bioinformatic Tools to assess the pattern of transcription factor binding motifs in the genomic reporter plasmids.

4.2.8.1. NCBI Blast

NCBI Blast (mentioned in Section 4.2.6.) was the initial tool used to identify genomic sequences located in the reporter plasmids. This programme matches sequences provided with possible locations in the genome. It was also used to identify possible genes that flanked the isolated sequences (Table 4.6.).
<table>
<thead>
<tr>
<th>5' Sequence</th>
<th>3' Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>208780 bp at 5' side: hypothetical protein</td>
<td>P1  16023 bp at 3' side: RUN domain containing 2B</td>
</tr>
<tr>
<td>2099210 bp at 5' side: hypothetical protein</td>
<td>P2  140120 bp at 3' side: kelch repeat and BTB (POZ) domain</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>187762 bp at 5' side: hypothetical protein</td>
<td>P4  14648 bp at 3' side: golgi phosphoprotein 3</td>
</tr>
<tr>
<td>289876 bp at 5' side: collagen, type XXII, alpha 1</td>
<td>P10 445010 bp at 3' side: potassium channel, subfamily K,</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>19805 bp at 5' side: hypothetical protein</td>
<td>P11 70232 bp at 3' side: hypothetical protein LOC728912</td>
</tr>
<tr>
<td>668934 bp at 5' side: pyrin-only protein 2</td>
<td>P15 13358 bp at 3' side: fibroblast growth factor 12 isoform</td>
</tr>
<tr>
<td>49194 bp at 5' side: odd-skipped related 1</td>
<td>P19 493796 bp at 3' side: tetratricopeptide repeat domain 32</td>
</tr>
<tr>
<td>9792 bp at 5' side: peroxisome proliferative</td>
<td>P20 16378 bp at 3' side: Fanconi anemia, complementation</td>
</tr>
<tr>
<td>activated receptor, delta</td>
<td></td>
</tr>
<tr>
<td>33059 bp at 5' side: pleckstrin homology domain</td>
<td>P21 26743 bp at 3' side: ribosomal protein S13</td>
</tr>
<tr>
<td>containing, family A member 7</td>
<td></td>
</tr>
<tr>
<td>216018 bp at 5' side: similar to hCG1660378</td>
<td>P22 118448 bp at 3' side: longevity assurance homolog 6</td>
</tr>
<tr>
<td>erythrocyte membrane protein band 4.1 (elliptocytosis 1, ...)</td>
<td>P23 erythrocyte membrane protein band 4.1 (elliptocytosis 1, ...)</td>
</tr>
<tr>
<td>endothelin converting enzyme 1</td>
<td>P24</td>
</tr>
<tr>
<td>echinoderm microtubule associated protein like 4</td>
<td>P26</td>
</tr>
<tr>
<td>51687 bp at 5' side: hypothetical protein</td>
<td>P27 18054 bp at 3' side: protein tyrosine phosphatase,</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>187762 bp at 5' side: hypothetical protein</td>
<td>P32 14648 bp at 3' side: golgi phosphoprotein 3</td>
</tr>
<tr>
<td>98700 bp at 5' side: mannosidase, alpha, class</td>
<td>P33 458966 bp at 3' side: similar to hCG1645384</td>
</tr>
<tr>
<td>2A, member 1</td>
<td></td>
</tr>
<tr>
<td>ribosomal protein L27</td>
<td>P36</td>
</tr>
<tr>
<td>2099210 bp at 5' side: hypothetical protein</td>
<td>P37 140120 bp at 3' side: kelch repeat and BTB (POZ) domain</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1189 bp at 5' side: similar to facioscapulohumeral muscular dystrophy</td>
<td>P38</td>
</tr>
<tr>
<td>19805 bp at 5' side: hypothetical protein</td>
<td>P40 70232 bp at 3' side: hypothetical protein LOC728912</td>
</tr>
<tr>
<td>2498 bp at 5' side: solute carrier family 26,</td>
<td>P45 5736 bp at 3' side: mitogen-activated protein kinase 14</td>
</tr>
<tr>
<td>member 8 isoform b</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.6. Genes that flank the 3' and 5' ends of isolated sequences (P1-P45).
4.2.8.2. MotifViz: Clover

Clover (Frith et al, 2004) is a web-based bioinformatical programme that searches for known transcription factor binding sites within a given set of sequences. This programme was used to search for ‘ALL JASPER’ (all receptor sites) and ‘ALL Nuclear Receptors’ (all nuclear receptor sites) and obtained the data displayed in Table 4.7 and Table 4.8. A threshold is a filter that is put in so it is possible to filter the results so only those with the highest values are displayed. For example, a threshold of four would only display results that have a score of four or more. Motifs from the ALL Jasper Clover search were scored according to the best base pair match; a high score reflects a good motif match within the sequences. Scores can be seen in Table 4.9. Analysis of Clover results revealed a number of transcription factor binding motifs that could be investigated further.
Table 4.7. All Jasper Clover Motifs located with a threshold of 4.
Table 4.8. All Nuclear Receptor binding motifs present in the reporter plasmid sequences with a threshold of 4
Table 4.9. Transcription Factor motifs found in the sequences. The ‘Score’ indicates the extent to which the motif found in the sequence matches the actual transcription factor binding motif, 5.89 being the best matched.

<table>
<thead>
<tr>
<th>Motif</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snail</td>
<td>5.89</td>
</tr>
<tr>
<td>Thing1-E47</td>
<td>4.46</td>
</tr>
<tr>
<td>NF-kappaB</td>
<td>2.91</td>
</tr>
<tr>
<td>Gklf</td>
<td>2.1</td>
</tr>
<tr>
<td>SP1-B</td>
<td>2.08</td>
</tr>
<tr>
<td>e-REL</td>
<td>1.73</td>
</tr>
<tr>
<td>SP1</td>
<td>1.63</td>
</tr>
<tr>
<td>Athb-1</td>
<td>1.42</td>
</tr>
<tr>
<td>HMG-1Y</td>
<td>1.36</td>
</tr>
<tr>
<td>HFH-2</td>
<td>1.34</td>
</tr>
<tr>
<td>p65</td>
<td>0.94</td>
</tr>
<tr>
<td>Pbx</td>
<td>0.831</td>
</tr>
<tr>
<td>S8</td>
<td>0.642</td>
</tr>
<tr>
<td>Tall beta-E47S</td>
<td>0.593</td>
</tr>
<tr>
<td>Pax-4</td>
<td>0.494</td>
</tr>
<tr>
<td>n-MYC</td>
<td>0.458</td>
</tr>
<tr>
<td>deltaEF1</td>
<td>0.352</td>
</tr>
<tr>
<td>USF</td>
<td>0.306</td>
</tr>
<tr>
<td>Dof3</td>
<td>0.231</td>
</tr>
<tr>
<td>SOX17</td>
<td>0.179</td>
</tr>
<tr>
<td>HFH-1</td>
<td>0.166</td>
</tr>
<tr>
<td>Hen-1</td>
<td>0.113</td>
</tr>
</tbody>
</table>
4.2.8.3. MEME Suite

MEME identifies overrepresented motifs in a set of given sequences. MEME found three repetitive sequences throughout the genomic DNA library analysed. The one most prevalent motif (Motif 1) can be seen in Figure 4.32. Motif 1 was found in P2, P10, P11, P23, P26, P37, P40 and P45.

![Figure 4.32. MEME Motif 1 found in P2, P10, P11, P23, P26, P37, P40 and P45](image)

4.2.8.4. TOMTOM

Each motif MEME finds overrepresented can then be analysed further using TOMTOM. TOMTOM finds known transcription factor binding motifs within the overrepresented MEME sequence. The best matched TOMTOM transcription factor binding motif for MEME 1 Motif 1 can be seen in Figure 4.33..

![Figure 4.33. TOMTOM Result from MEME Motif 1 is a transcription factor binding motif for Gfi-1](image)
Other binding motifs found from TOMTOM are listed in Table 4.10.

<table>
<thead>
<tr>
<th>MOTIF 1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gfi1</td>
<td></td>
</tr>
<tr>
<td>Maf-1</td>
<td></td>
</tr>
<tr>
<td>SU_h</td>
<td></td>
</tr>
<tr>
<td>Dof3</td>
<td></td>
</tr>
<tr>
<td>Spzl</td>
<td></td>
</tr>
<tr>
<td>Ahr-ARNT</td>
<td></td>
</tr>
<tr>
<td>HMG-1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MOTIF 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MZF_5-13</td>
<td></td>
</tr>
<tr>
<td>GAMYB</td>
<td></td>
</tr>
<tr>
<td>DeltaEF1</td>
<td></td>
</tr>
<tr>
<td>Snail</td>
<td></td>
</tr>
<tr>
<td>MZF_1-4</td>
<td></td>
</tr>
<tr>
<td>Dorsal_1</td>
<td></td>
</tr>
<tr>
<td>Agamous</td>
<td></td>
</tr>
<tr>
<td>SP1</td>
<td></td>
</tr>
<tr>
<td>GR</td>
<td></td>
</tr>
<tr>
<td>HNF4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MOTIF 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Androgen</td>
<td></td>
</tr>
<tr>
<td>COUP-TF</td>
<td></td>
</tr>
<tr>
<td>SAP-1</td>
<td></td>
</tr>
<tr>
<td>GAMYB</td>
<td></td>
</tr>
<tr>
<td>HMG-1Y</td>
<td></td>
</tr>
<tr>
<td>Gfi1</td>
<td></td>
</tr>
<tr>
<td>PPARgammaRXRα</td>
<td></td>
</tr>
<tr>
<td>NF-Y</td>
<td></td>
</tr>
<tr>
<td>Hunchback</td>
<td></td>
</tr>
<tr>
<td>E74A</td>
<td></td>
</tr>
<tr>
<td>Dof2</td>
<td></td>
</tr>
<tr>
<td>PBF</td>
<td></td>
</tr>
<tr>
<td>Dof3</td>
<td></td>
</tr>
<tr>
<td>Elk-1</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.10. TF Binding Motifs in MEME Results found using TOMTOM

Bioinformatical data obtained from Clover and MEME was reviewed and further investigation focused of transcription factors that appeared in both sets of analysis tools. These included SNAIL, COUP-TF I and II and RORα.
4.2.9. **Further Trans-Activation Assays to Investigate unknown Transcription Factor SRC-1A association**

It was a possibility that SRC-1A Dam associated with nuclear receptors endogenously present in the HeLa cells when the library sequences were being selected. It was then necessary to identify which nuclear receptors were involved. Based on bioinformatical data range of nuclear receptors were investigated to assess what SRC-1A associated with on the genome. Bioinformatical data on the presence of transcription factor binding motifs influenced the choice of transcription factor used in the transactivation assays, the MotifViz programme Clover was the main programme used.

SNAIL was the first transcription factor to be analysed and was chosen due the presence of a very high score (good match to transcription factor motifs and present in multiple reporter plasmids) in the MotifViz results. The luciferase assay was performed to assess the influence of SNAIL on the transactivation of reporter plasmids in the presence or absence of SRC-1A (*Figure 4.34.*). It is apparent from the graph that some of the plasmids are responsive, regardless of the transfection, and the reporter plasmids that are cotransfected with SNAIL and SRC-1A simultaneously in the transactivation assay seem to have a similar or lower response than when transfected with SRC-1A alone. Levels of transcription induction are not strong enough to conclude that SNAIL associates with, and affects transcription of, the sequences contained in the reporter plasmids.
**Figure 4.34.** Luciferase assay results following HeLa cells cotransfection with 200ng isolated, sequenced reporter plasmid (P1-P45) with either; 1: pCR3.1 (empty vector), 2: pCR3.1 and pCMV5 SNAIL, 3: pCR3.1 and pCR3.1 SRC-1A or 4: pCMV5 SNAIL and pCR3.1 SRC-1A. pGL3 Basic (empty vector) was used as a control. The transfected HeLa cells were then incubated for 24hrs.

The next transcription factor to be analysed for its influence on the reporter plasmid library was RORα. This is an orphan nuclear receptor whose motif was seen in four of the reporter plasmid inserts (p10, p26, p38 and p45) using the MotifViz programme.

**Figure 4.35.** displays the results following transfection of the reporter plasmids (P1-P45) and demonstrated some interesting results. Increased transcriptional activity is seen in P2, P20, P37 and P45. Out of these responders, only P45 was seen to have a RORα binding site in MotifViz. Neither bioinformatical programme (MotifViz or MEME) identified the presence of a RORα binding motifs in P2, P20 or P37. Levels of transcription in pGL3-Basic is low indicating there is a high possibility of the presence of motifs in the reporter plasmids that cause transcription induction. However, overall levels of gene transcription are still relatively low.
Figure 4.35. Luciferase assay results following HeLa cell cotransfection with 200ng reporter plasmids (P1-P45) and either pCR3.1 or pCR3.1 RORα. pGL3 Basic (empty vector) was used as a negative control and pGRE-E1b-Luc was used as a positive control.

Normalized results (or fold induction) can be viewed in Figure 4.36. This further confirms the transcriptional activation of P2, P20, P37 and P45 by RORα.

Figure 4.36. Normalized results for RORα transfection described in Figure 4.35.

COUP TF motifs were also seen in the bioinformatical results, from both MotifViz and TOMTOM. Results obtained following transfection of COUP TF1 or COUP TFII
expression plasmid (Figure 4.37.) demonstrated that both COUP TFI and II induced transcription in all of the reporter plasmids to some degree while pGL3 Basic (empty vector) transcriptional response remained relatively low.

![Graph](image.png)

**Figure 4.37.** Luciferase assay results following cotransfection of HeLa cells with reporter plasmid (P1-P45) and either pCR3.1 (empty vector), pCR3.1 COUP TFI or pCR3.1 COUP TFII. pGL3 Basic was used as a negative control.

Reporter plasmids p26, p27, p33, p40 and p45 appeared to be the best responders following COUP TFI and COUP TFII transfection and all had levels of transcription higher than that seen in pGL3 Basic. These reporter plasmids were focused on in further tests.

Interestingly, p10 and p23 were the only two plasmid sequences seen in MotifViz to have COUP TF binding motifs and neither of these displayed a higher level of transcription induction with COUP.

To assess the concentration of COUP TFI and COUP TFII that were required for further transfections titres were performed. Initial COUP TF transfections were performed
using 25ng (data not displayed) and 10ng per well. Lower concentrations were used for the titres. The results can be seen in Figure 4.38.

![Figure 4.38. Luciferase assay results following HeLa cell cotransfection with 200ng reporter plasmid and 0.1ng pCR3.1(empty vector) / pCR3.1 COUP TFI / pCR3.1 COUP TFII or 0.25ng pCR3.1 / pCR3.1 COUP TFI / pCR3.1 COUP TF II or 1ng pCR3.1 / pCR3.1 COUP TFI / pCR3.1 COUP TFII. pGL3 Basic was used as a negative control.](image)

Following observation of the results in Figure 4.38, and previous results in Figure 4.36, it was decided to use 2ng COUP TFI and COUP TFII in further transfections, a concentration not tested, however, 1ng appeared to provide good specific results and it was suspected that 2ng would retain the specificity and provide additional activation capabilities.

COUP TF is an orphan nuclear receptor transcription factor. COUP TF has been seen previously to bind to an intermediate and not directly to the genome. It is this capability of COUP TF that led to investigations to determine if COUP TF was binding to an intermediate transcription factor. RORα in the presence and absence of SRC-1A was
investigated to see if it cooperated with COUP TFI or II to regulate transcription (Figure 4.39).

The results displayed are normalized fold inductions, obtained by dividing the reporter plasmid result by that obtained using an empty pGL3 Basic vector which provides a normalized result and then dividing the normalized result by the normalized result obtained from the empty pCR3.1 vector (expression plasmid blank).

**Figure 4.39.** Luciferase assay results following HeLa cell cotransfection of reporter plasmids (P26, P27, P33, P40, P45) with pCR3.1 / COUP TFI / COUP TFII followed by pCR3.1 / pCR3.1 and pCR3.1 RORa / pCR3.1 RORa and pCR3.1 SRC-1A. pGL3 Basic was used as a negative control, and to enable normalized values to be calculated. Statistical analysis was performed by comparing HeLa cell transfection with pCR3.1 to that with COUP TFI or COUP TFII and RORa and SRC-1A. *p<0.05, **p<0.01 negative control pCR3.1 compared to RORa & SRC1A co transfected with either COUP TFI or COUP TFII
The results indicated that COUP TFI and II, but particularly COUP TFI, were displaying an interaction with RORα and SRC-1A to induce gene transcription. This was particularly evident with reporter plasmid P45, where both COUP TFI and COUP TFII interacted with RORα and SRC-1A to significantly induce gene transcription (p<0.01) (Figure 4.39). There appears to be cooperational activity between COUP TFI/COUP TFII, RORα and SRC-1A. Certain COUP TF responders (P26 and P45) show greater transactivation. Additionally, the results confirm that RORα alone is unable to induce gene transcription.

The final set of transactivation assays performed functioned to assess a variety of concentrations of COUP TF I, COUP TF II, RORα and SRC-1A. The results can be seen in Figure 4.40., with normalized values in Figure 4.41.)
Figure 4.40. Luciferase results following HeLa cell cotransfection with 200ng COUP responsive reporter with transcription factors as described above. pGL3 Basic was used as a negative control.

The results demonstrate that, particularly in P27 and P45, these transcription factors work co-operatively in a concentration dependant and site specific manner.
**Figure 4.41.** Normalized values for Figure 4.40. (Colour Key can be viewed in Figure 4.40.)

**Figure 4.42.** is a simplified version and can summarize the results obtained from this transactivation assay. The pGL3-Basic reporter plasmid chosen to demonstrate the results is P45.

**Figure 4.42.** Simplified results from Graph E in Figure 4.40. using 10ng COUP TF/I/II.

It can be seen from this graph that the empty vector provides no response, this is the negative control. COUP TF I, COUP TF II and RORα did not cause a large increase in transcription when transfected alone, while SRC-1A, when co-transfected with RORα, displayed a slight increase in transcription. However, when COUP TF I was co-transfected with RORα it increased the transcriptional response, when compared to COUP TF I alone, by greater than 10 fold.
Interestingly, when COUP TF I was co-transfected with both RORα and SRC-1A it elicited the greatest response, indicating that there is some form of cooperativity between the three molecules. This is in contrast to COUP TFII which demonstrated only modest transcriptional activity when cotransfected with RORα and SRC-1A.

4.2.10. Bioinformatical Investigation of COUP TFI, RORα and SRC-1A Responsive Sequences

Bioinformatical analysis was performed on sequences P26, P27, P33, P40 and P45 to ascertain any commonality between these COUP TFI, RORα and SRC-1A responsive sequences. The analysis focused on the MEME suite, using both MEME and TOMTOM. These are described in more detail in Section 4.2.8.

Meme analysis presented with three motifs found in the sequences. Motif 1 (Figure 4.43) is a motif found in all five sequences.

![Figure 4.43. Motif 1](image-url)
TOMTOM analysis of Motif 1 revealed just one transcription factor binding motif (Figure 4.44.), Chop-cEBP.

Figure 4.44. Transcription factor binding site for Chop-cEBP

TOMTOM found no transcription factor binding sites in Motif 2. Motif 3 was found in sequences P26, P33, P40 and P45 (Figure 4.46.).

Figure 4.45. Motif 2 found in P26, P27, P33, P40 and P45

Figure 4.46. Motif 3, present in P26, P33, P40 and P45
TOM TOM isolated a transcription factor binding site for Broad Complex 1 in Motif 3 (Figure 4.47.).

Figure 4.47. TOMTOM results from Motif 3 revealed transcription factor binding site for Broad-complex 1

Further analysis is required on the selected reporter plasmid sequences however they provide interesting insight into possible cooperation between COUP TFI, RORα and SRC-1A.

4.3. Locating Transcription Start Sites in Reporter Plasmids

GeneRacer technique (Materials and Methods 3.2.7.) was used to attempt to locate the transcription start site in reporter plasmid sequences. This method proved unsuccessful in locating the transcription start sites of the COUP TF responsive reporter plasmids or reporter plasmids from the ERE library. Sequences obtained can be viewed in Appendix IV (TSS1 to 6 from trial 1, TSS7 to 12 from trial 2). It appeared that the LucRev primer was binding non-specifically to endogenous DNA and not to reporter plasmid DNA. Attempted optimization of this method still proved unsuccessful. A failure to produce PCR products can be seen in Figure 4.48.
Figure 4.48. 1% Agarose Gel electrophoresis of Gene Racer transcription start site PCR products. Lane 1 contains molecular weight marker. Lane 2 contains PCR products. No PCR products seen, only primer dimers.
5.0. **Discussion**

Initially, the aim of this project was to investigate a reporter library of oestrogen responsive elements (EREs) which constructed by Jennings (Jennings, 2009) using DamID technology. Investigations on this library were to focus on the identification of SERM responders and perform further tests to analyse the isolated reporters, including assessing the ability of the isolated sequences to mediate transcription in conjunction with other nuclear receptors and coactivators and to perform ERα mutation analysis. The purpose of this was to gain a deeper understanding into the role of Tamoxifen as an oestrogen antagonist.

SRC-3, a well documented coactivator and a member of the SRC family of coactivators, appeared to repress SERM responsive reporter activity and it was decided to investigate the influence of SRC family members further. This body of work aimed to isolate sequences from the genome that associate with SRC-1A:Dam and SRC-3 in association with either ligand bound oestrogen or progesterone receptors using DamID technology, the same technology that was used to construct the ERE library. Following construction of the library, further aims included the use of bioinformatics to identify DNA cis elements, repetitive sequences and overrepresented transcription factor binding sites in these isolated sequences and to assess the ability of the isolated sequences to mediate transcription in conjunction with other nuclear receptors and coactivators to establish potential novel interactions.

An additional aim, encompassing reporter plasmids from both the ERE, SRC-1A and SRC-3 libraries, was to attempt to identify the location of transcription start sites in selected isolated reporter plasmid sequences. The ability to obtain this information
would enable a deeper understanding of the binding sites that a variety of transcription factors of transcriptional coregulators associate with.

5.1. Oestrogen Responsive Element (ERE) Reporter Library Investigation

This portion of the investigations focused on the investigation of SERM responsive reporter plasmid sequences, comparing the responses of these sequences to that of oestrogen responsive sequences.

Transfections were performed on the ERE library to identify SERM responsive and oestrogen responsive reporter plasmids. Four were selected for further investigations following sequencing to ensure the plasmid inserts were of genomic origin. E5p and T4p were selected as SERM responsive plasmids, while V17 and E12 were selected as oestrogen responsive reporter plasmids to use as a comparison during the investigations.

Transfections were performed on these reporter plasmids to assess the ERα binding site requirements of the SERM responsive plasmids and the oestrogen responsive plasmids. These transfections were performed to determine if the differential response of the oestrogen and SERM responsive plasmids was attributed to the SERMs interacting with a different area of the ERα molecule perhaps causing ERα to bind in a different way to the DNA. It is known that when tamoxifen binds to ERα it causes a different conformational change to those that occur when oestrogen binds. The conformational change that occurs when Tamoxifen binds repositions Helix 12 of the ERα molecule and prevents the binding of coactivators (Shiau et al, 1998), preventing AF-2 activation.
Therefore Tamoxifen functions as an ERα antagonist to genes which rely on the AF-2 region for ER-mediated transcription (Ring and Dowsett, 2004).

It is known that there are three major domains (DNA binding domain, AF-2 domain and AF-1 domain) required in order for effective ERα mediated transcription to occur. A DNA Binding Domain mutant, an AF-2 defective mutant and an AF-1 defective mutant were used during the transfections and cells were treated with EtOH, oestrogen, tamoxifen and raloxifene. The results (Figure 4.8.) revealed that all three domains were required for effective ERα mediated transcription in both subsets of reporter plasmids. This contradicts the theory that Tamoxifen does not require the AF-2 domain to activate AF-2 independent genes and suggests that in these reporter plasmids there is an alternative mechanism operating.

Following ERα mutant analysis of the SERM responsive plasmids, further investigation was performed to assess their response to transfections with hERα alone and hERα with either SNAIL, PAX4 or SRC-3. These transfections were then treated with EtOH, oestrogen, tamoxifen or raloxifene.

The results revealed interesting results. The results obtained from transfections with SNAIL and PAX4 were as expected and both have been previously demonstrated to act as transcriptional repressors (Hemavathy et al., 2000; Nieto, 2002; Smith et al., 1999) and the levels of transcription from both transfections were lower than that of the ERα transfections (for all treatments, EtOH, oestrogen, tamoxifen and raloxifene). However, the result obtained from transfection with SRC-3 revealed SRC-3 also acted as a repressor when it associated with the sequences contained in the reporter plasmids E5p and T4p. This result was surprising as SRC-3 has only previously been reported to act as a transcriptional coactivator (Wu et al., 2002, 2004, 2007; Feng et al., 2006; Naeem et
al, 2007; Chen et al, 1999). Oestrogen responsive reporter plasmids E12 and V17 were then used to assess their responsiveness to SRC-3. These transfections produced a 'typical' response of a coactivator, producing approximately a 3 fold increased transcriptional response with the addition of SRC-3.

The interesting results obtained using SRC-3 led the project to further investigate the family of SRC coactivators.

5.2. SRC-1A and SRC-3 Library Construction and Investigation

SRC-1A and SRC-3 are two of the members of the SRC family of coactivators that were focused on in this section. The major aim of this section was to construct four reporter plasmid libraries (using DamID technology), containing sequences that interacted with SRC-1A or SRC-3 via either ERα and oestrogen or PR-B and progesterone. Construction of these libraries would enable further analysis of the libraries to be performed to gain a greater understanding in the relationship between the effect of the coactivator molecules SRC-1A and SRC-3 and oestrogen and progesterone receptor mediated transcription. Additionally, if successful, it would enable a new method for investigating a broad range of coactivators and reveal more about the sequences that these molecules influence.

The primary step was to construct SRC-1A: and SRC-3:Dam fusion plasmids. This process used a pCR3.1 SRC-1A:Luc plasmid or a pCR3.1 SRC-3:Luc plasmid, whereby the Luc gene was excised from the plasmid and replaced with a Dam gene, which was
inserted directly downstream of the SRC-1A or SRC-3 gene, forming a fusion gene. The Dam gene was isolated from E.Coli DNA and amplified using PCR.

Results of the Luc excision step revealed successful removal of the Luc fragment from the pCR3.1 SRC-1A:Luc plasmid, however, removal of the Luc fragment from pCR3.1 SRC-3 appeared to be unsuccessful. Protocol was continued however for both plasmids, the Luc fragment was removed and discarded and a ligation reaction was performed to ligate the amplified Dam gene to the pCR3.1 SRC-1A and pCR3.1 SRC-3 plasmids and transformed to chemically competent E.Coli cells and ncubated on agar containing ampicillin overnight. Controls were also transfected which involved transfecting the pCR3.1 SRC-1A or pCR3.1 SRC-3 plasmids without a Luc or Dam gene. The results demonstrated successful excision of the Luc gene from the pCR3.1 SRC-1A:Luc plasmid as there was no growth on the control agar due the the pCR3.1 SRC-1A being in linear form and thus unable to express its ampicillin resistant gene. pCR3.1 SRC-1A:Dam therefore was successfully constructed.

On the other hand, a high number of colonies observed on the pCR3.1 SRC-3 control plate indicated that the Luc excision step was unsuccessful, as the control plasmid was able to express the ampicillin resistant gene.

pCR3.1 SRC-1A:Dam and pCR3.1 SRC-3:Dam colonies were picked and purified for further analysis.

Further analysis confirmed the successful construction of pCR3.1 SRC-1A:Dam and the unsuccessful construction of pCR3.1 SRC-3:Dam. Sequence analysis confirmed the directionality of the Dam gene in the pCR3.1 SRC-1A:Dam plasmid.
Due to difficulties in constructing the pCR3.1 SRC-3:Dam expression plasmid the attempts to do so were suspended. However, as construction of the pCR3.1 SRC-1A:Dam was successful, this SRC-1A:Dam fusion plasmid was used to construct a genomic library containing sequences that the SRC-1A:Dam fusion protein interact indirectly with via either hERα and oestrogen or hPR-B and progesterone.

Once the pCR3.1 SRC-1A:Dam plasmid was constructed and processed DamID protocol could proceed.

The DamID protocol involved in vivo targeted methylation of GATC sequences adjacent to the indirect binding sites of the SRC-1A:Dam fusion protein. This was achieved by cotransfecting HeLa cells with the pCR3.1 SRC-1A:Dam plasmid along with expression plasmids for either hERα or hPR-B, followed by treatment with the respective hormones (either oestrogen, progesterone or EtOH).

Following the transfection protocol genomic DNA was isolated from the transfected HeLa cells and DpnI digestion was performed to enable isolation and amplification of methylated sequences.

LM-PCR amplification was performed and the resultant sequences were analysed to ensure a variety of sequences were amplified (successful LM-PCR products will appear as a smear on agarose gel). Smears were present for all four LM-PCR reactions, however the sequences obtained using the receptors plus EtOH contained more DNA than that obtained from receptors plus hormone. This could indicate that SRC-1A:Dam preferentially binds indirectly to DNA via unbound receptor, however this has not been previously reported and in fact the action of SRC-1 has been found to be ligand dependant (Onate et al, 1995). However, the appearance that there is more LM-PCR product in the lanes that did not contain hormone does not indicate that SRC-1A has the
ability to activate promoter elements contained in these sequences, it merely suggests that SRC-1A has the ability to bind to unligand bound ERα and PR-B.

Protocol was adjusted, using a reduced amount of LM-PCR product and as a result of this optimization the TOPO cloning procedure obtained approximately 2000 colonies. Rather than perform restriction digestions on this TOPO library, colonies containing sequences obtained with ERα with oestrogen and PR-B and progesterone were directly transferred to an expression vector library.

The expression vector used was pGL3-Basic, which was adapted to be gateway compatible by Jennings (Jennings, 2009) and enables functional analysis of the isolated sequences. This vector contains a Luc gene located downstream of the sequence insert site. Additionally, it does not contain a promoter sequence, thereby the sequence insert can be investigated to assess for promoter activity.

Isolated sequences were transferred from the TOPO vector to the pGL3-Basic vector using gateway cloning technology.

The pGL3-Basic reporter library was functionally tested using HeLa cell transfections followed by the luciferase assay.

This protocol was used initially to screen plasmids for genomic inserts. In theory, if the plasmids contained sequences that bound, indirectly, to the SRC-1A:Dam fusion protein via the receptor (ERα or PR-B) and hormone (oestrogen or progesterone) and contains a promoter element, following HeLa cell transfections with the relevant receptor and hormone, the luciferase assay will be able to detect the promoter activity.

The results from these transactivation assays were unsuccessful as all reporter plasmids showed no response. Lack of response could mean a number of things; that the
sequences that are inserted in the plasmid do not contain a promoter and therefore Luc
could not be transcribed, that the sequences have some other effect on gene expression
or that the sites isolated bind to SRC-1A in association with other proteins and not
specifically the nuclear receptor and hormone that was added (endogenous transcription
factors).

In total there were approximately 65 reporter plasmids in each library. Sequencing was
performed on 5 plasmids from each library to determine which sequence inserts are
from genomic origin (using NCBI Blast) and which are bacterial or vector sequences.
Initial sequencing determined that the library obtained from pCR3.1 SRC-1A:Dam
cotransfection with PR-B followed by treatment with progesterone contained a higher
percentage of sequences of genomic origin than the library obtained from pCR3.1 SRC-
1A:Dam cotransfection with ERα followed by treatment with oestrogen. As neither of
the library sequence inserts appeared to respond to the receptor and hormone (Figure
4.29.) it was constructed with further analysis was performed on the PR-B library as it
contained a higher amount of genomic inserts. The first 21 plasmids that contained
genomic inserts from the SRC-1A:Dam PR-B progesterone library were used for further
analysis.

Additionally, NCBI Blast revealed genes located upstream and downstream the isolated
sequences. Providing information about the genes possibly controlled by the isolated
sequences provides additional information regarding on the isolated sequences and their
function in the genome.

The initial assays performed on these plasmids involved two transactivation assays.
The first was a cotransfection of HeLa cells with the said plasmids and expression
vectors for SRC-1A, ERα, PR-B and GR. The cells were then simultaneously treated
with oestrogen, progesterone and dexamethasone. The second transactivation was set up the same except the SRC-1A expression vector was substituted with pCR3.1, the empty expression vector. The theory behind this was to assess if any of these ligand bound receptors interacted with the plasmid inserts and if SRC-1A had any affect on their action on the sequences. pGRE/E1b/Luc, a reporter plasmid containing a sequence responsive to GR was used as a control to ensure the procedure was working. The control responded to the transactivation assay as expected, with the addition of SRC-1A to the transfection the gene expression of the control increased by 13 fold. However the test plasmids showed no significant response to either transactivation assay. A significant response would be that similar to the control, in that there is a response to the addition of hormone and not to the addition of EtOH. The test plasmids showed approximately equal response to hormone and to EtOH indicating the isolated sequences do not interact with ERα, PR-B or GR.

It was a possibility that during DamID step whereby GATC methylation occurred SRC-1A:Dam indirectly bound to the DNA via proteins endogenously present in the HeLa cells. A bioinformatical approach was then taken to attempt to identify possible binding sites within the isolated sequences. Two programmes were used; MotifViz and The MEME Suite. Motifviz (Fu et al, 2004) identifies overrepresented regulatory motifs (i.e.: transcription factor binding sites) in a given set of sequences while MEME does not locate transcription factor binding sites but merely identifies over-represented sequence motifs in a set of given sequences (Bailey and Elkan, 2004). A second part of the MEME Suite is a programme, known as TOMTOM, which is used to then identify transcription factor binding sites within the overrepresented sequence motifs identified by the MEME programme.
Of the transcription factor binding motifs identified by the above mentioned programmes, the transcription factors SNAIL, RORα and COUP TFI & II were the ones most prevalent and these were focused on for further investigations.

Transactivation assays with SNAIL and SRC-1A and with RORα revealed no significant results.

Transactivation assays with COUP TFI (chicken ovalbumin upstream promoter transcription factor I) and COUP TFII did find 5 responsive plasmids. However, plasmids that responded to COUP TFI and COUP TFII did not, according to bioinformatical results, contain COUP binding motifs.

COUP TFI and COUP TFII belong to the steroid receptor superfamily (Wang et al, 1989). Currently, no specific ligands for these transcription factors have been identified and thus they are classed as orphan nuclear receptors.

COUP-TF is one of the best characterized orphan nuclear receptors. It was first identified as a homodimer, found to bind to a direct repeat regulatory element, containing an imperfect direct repeat of AGGTCA, in the chicken ovalbumin promoter (Pastorcic et al, 1986; Sagami et al, 1986; Wang et al, 1989). In mammals two genes have now been identified: COUP TFI/EAR3 and COUP TFII/ARP-1 (Miyajima et al, 1988; Wang et al, 1991; Ladias et al, 1991). COUP TFI and COUP TFII are closely related transcription factors however each factor has the ability to generate its own distinct expression profile during mammalian development (Qiu et al, 1994).

Both COUP TF I and II are expressed in many human tissues, being involved in many biological regulation processes such as neurogenesis, organogenesis, determination of cell fate and metabolic homeostasis (Tsai et al, 1994; Ladias and Karathanasis; 1991;

COUP TF contains a 66-amino acid region which contains two conserved Zinc finger motifs, this is the DBD region; COUP TF I and II are identical but for one amino acid difference within the DBD which is a conservative change from Ser to Thr (Miyajima et al, 1988; Wang et al, 1989).

COUP-TF has been seen to act as a transcriptional activator and repressor. It was initially identified as a transcriptional activator (Sagami et al, 1986) and further to this has been seen to demonstrate transrepression capabilities (Cooney et al, 1992; Tran et al, 1992; Park et al, 2003; Achatz et al, 1997; Leng et al, 1996), via DNA-independent heterodimerization of the ligand binding domain COUP TFs with other nuclear receptors, for example thyroid hormone receptors, retinoic acid receptors and retinoid x receptors (Leng et al, 1996; Kruse et al, 2008).

One possible mechanism of action hypothesised by Leng et al, suggests that the heterodimerization of the COUP TFs with the nuclear receptor LBD results in either suppression of the AF regions in these receptors or reduction of the nuclear receptors ligand binding ability (Leng et al, 1996).

The ability of COUP TF to heterodimerize to other nuclear receptors raised a number of questions relating to the results already found. The fact that none of the COUP TF responsive plasmids contained a known COUP TF binding sites (results from clover bioinformatical approach) lead to the question: perhaps COUP TF is interacting through another nuclear receptor and is operating via a novel binding site?
Initially, transactivation assays were performed to reduce background (non specific response from the empty vector) obtained in the plasmids response to COUP TFI and COUP TF II. It was then determined from these titre assays that 2ng was the optimal amount of COUP TFI and COUP TFII to use in further investigations.

As the affects of RORα on the reporter plasmids was studied previous to COUP TF, and COUP TF has been demonstrated to have RAR (retinoic acid receptor) interaction capabilities (Kliewer et al, 1992), RORα was the first investigated.

HeLa cell transfections were performed, cotransfecting reporter plasmids with COUP TFI or COUP TFII (or empty expression vector, pCR3.1) and then each of these groups with either expression vectors for RORα alone or with expression vectors for RORα and SRC-1A. pGL3 Basic was used as a control. The results demonstrated that the COUP TF responsive plasmids did respond more favorably to COUP TF, RORα and SRC-1A.

Figure 4.43. demonstrates a summary of these results obtained, this response was seen in P27 and P45 in particular. From this figure it can be seen that there is no response with the empty vector transfection alone (pGL3 Basic; V). Results from transfections with reporter plasmid with either COUP TFI, COUP TFII or RORα showed a slight response, but when compared to further responses it was negligible. Cotransfection reporter plasmid, RORα and SRC-1A showed a response approximately double to that seen in RORα alone. Cotransfection of reporter plasmid, COUP TFI and RORα showed a dramatic response, as did COUP TFI and RORα although this was not as significant. When SRC-1A was added to these transfections the response increased further.

It is clear from these results that P27 and P45 contain binding sites that facilitate some form of cooperational binding between COUP TFI and RORα, whose relationship with DNA is enhanced with the addition of SRC-1A.
As, in this situation, the association of COUP TF I with RORα results in transcriptional activity, it suggests that COUP TF is not associating with the ligand binding domain of RORα, as previous studies have found the binding of COUP TF to the ligand binding domain of other nuclear receptors (thyroid hormone receptors, retinoic acid receptors and retinoid X receptors) (Leng et al., 1996) to result in transcriptional repression. Rather than forming a COUP TF:RORα heterodimer, the two molecules could bind to a novel binding motif within these sequences at separate times and carry out separate functions. The action of each of these molecules alone may not be sufficient to induce transcription, however when the actions of each of the molecules are combined transcription is activated. The role of SRC-1A may involve mediation of the association of these molecules to the DNA, acting perhaps as a 'crane', moving one of the receptors in, enabling it to carry out its actions, removing it and bringing in the second.

However, these are only hypothesized models of how these factors are interacting and further work would have to be carried out to assess the exact mechanism of how these nuclear receptors are functioning to activate transcription and what the role of SRC-1A has in this mechanism.

Additionally, in order for this to have occurred during the construction of the library RORα and COUP TF would have to be endogenously present in the HeLa cells used during these steps, as the SRC-1A:Dam would have had to preferentially bind to the COUP-TFs and RORα than to the ligand bind progesterone receptor.

RORα, COUP TFI and COUP TFII can be found in cancerous cervical cells, which is the origin of HeLa cells (www.genecards.com). This confirms that cancerous cervical cells have the ability to produce all three of the orphan nuclear receptors, therefore there
is the possibility that instead of SRC-1A Dam binding to the respective hormone receptor which attached to its ligand (hERα and E2 or hPR-B and Progesterone), it therefore seemed that SRC-1A Dam had a higher affinity for COUP TF or RORα in the conditions provided than it had for ERα or PR-B.

The five COUP TF responsive plasmids were bioinformatically analyzed using the MEME Suite. TOMTOM revealed no COUP-TF or RORα binding sites in the sequences.

5.3. Locating Transcription Start Sites

It was attempted to locate the transcription start sites of a selection of reporter plasmids from both the ERE reporter library and the SRC-1A reporter library. This was unsuccessful despite optimisation of the techniques used.

The identification of transcription start sites within the reporter plasmid sequences would have enabled further investigation into the effect of the transcription start site to transcriptional activation.

Further optimisation would be required to enable this technique to be used to investigate transcription start sites.

5.4. Future Work

Further work still needs to be performed on the activity of SRC-3 as a transcriptional repressor in the SERM responsive plasmids identified.
The role of DamID to investigate this further was unsuccessful and once a SRC-3:Dam fusion plasmid has been constructed this may be an option to identify where SRC-3 association sites on the genome. It would prove interesting to identify sites in the genome where SRC-3 binds via Tamoxifen bound ERα, and to attempt to locate more SERM responsive sequences.

The DamID technique did not work as expected but with alternative methods the problems encountered could be avoided. Primarily the method did not produce a library of sequences that bound to either of the ligand bound receptors that were being investigated. However it did result in a selection of sequences that were isolated due to SRC-1A:Dam binding via endogenous transcription factors in HeLa cells, of which it was identified that COUP TF and RORα were just two that a selection of the reporter plasmids responded to. Further investigations, and perhaps trial and error, would be required to identify the endogenous transcription factors that the remaining sequences were isolated by.

Additionally, an interesting study would involve deducing the mechanism of action of the cooperation between COUP TFI, RORα and SRC-1A to activate gene transcription. Co-IP (Co Immunoprecipitation) followed by western blots could be used to identify if there is and heterodimerization occurring between COUP TFI and RORα or not and the results from those investigation would influence further work performed.
6.0. REFERENCES


Allan GF, Tsai SY, Tsai MJ, O'Malley BW. Ligand-dependent conformational changes in the progesterone receptor are necessary for events that follow DNA binding. Proc. Natl. Acad. Sci. 89:11750-11754, 1992


Armstrong DT and Black DL. Influence of luteinizing hormone on corpus luteum metabolism and progesterone biosynthesis throughout the bovine estrous cycle. Endocrinology. 78:937-944, 1966


Chauvet C, Bois-Joyeux B, Danan JL. Retinoic acid receptor-related orphan receptor (ROR) alpha4 is the predominant isoform of the nuclear receptor RORalpha in the liver and is up-regulated by hypoxia in HepG2 human hepatoma cells. Biochem J. 364:449-56, 2002

Chen D, Ma H, Hong H, Koh SS, Huang SM, Schurter BT, Aswad DW, Stallcup MR. Regulation of Transcription by a Protein Methyltransferase. Science. 284(5423):2174-2177, 1999


Clarke CL, Sutherland RL. Progestin regulation of cellular proliferation. Endocr Rev. 11:266-301, 1990


Cooney AJ, Tsai SY, O'Malley BW, Tsai MJ. Chicken Ovalbumin Upstream Promoter Transcription Factor (COUP-TF) Dimers Bind to Different GGTCA Response Elements, Allowing COUP-TF To Repress Hormonal Induction of the Vitamin D3, Thyroid Hormone, and Retinoic Acid Receptors. Molecular and Cellular biology. 12(9):4153-4163, 1992


Gray S, Szymanski P, Levine M. Short-range repression permits multiple enhancers to function autonomously within a complex promoter. Genes Dev. 8: 1829-1838, 1994


Hong H, Kohli K, Garabedian MJ, Stallcup MR. GRIP1, a transcriptional coactivator for the AF-2 transactivation domain of steroid, thyroid, retinoid, and vitamin D receptors. Mol Cell Biol. 17:2735-2744, 1997

Hong H, Kohli K, Trivedi A, Johnson DL, Stallcup MR. GRIP1, a novel mouse protein that serves as a transcriptional coactivator in yeast for the hormone binding domains of steroid receptors. Proc Natl Acad Sci USA 93:4948-4952, 1996


Jennings C. Generation of a Reporter Library of ERα Interacting DNA Sequences. Submitted to DIT for MSc, 2009

Jetten AM. Retinoid-related orphan receptors (RORs): critical roles in development, immunity, circadian rhythm, and cellular metabolism. Nuclear Receptor Signaling. 7:1-32, 2009


Kalkhoven E, Valentine JE, Heery DM, Parker MG. Isoforms of steroid receptor co-activator 1 differ in their ability to potentiate transcription by the oestrogen receptor. EMBO J. 17:232-243, 1998

Kallen J, Schlaeppe JM, Bitsch F, Delhon I, Fournier B. Crystal structure of the human RORα Ligand binding domain in complex with cholesterol sulfate at 2.2 Å J Biol Chem. 279:14033-14038, 2004

Kallen JA, Schlaeppe JM, Bitsch F, Geise S, Geiser M, Delhon I, Fournier B. X-ray structure of the hRORα LBD at 1.63 Å: structural and functional data that cholesterol or a cholesterol derivative is the natural ligand of RORalpha. Structure. 10:1697-707, 2002


201

Klinge C. Estrogen receptor interaction with co-activators and co-repressors. Steroids. 65:227-251, 2000


Kumar V, Green S, Staub A, Chambon P. Localization of the oestradiol-binding and putative DNA-binding domains of the human oestrogen receptor. EMBO J. 5:2231-2236, 1986


Leptin M. Twist and snail as positive and negative regulators during Drosophila mesoderm development. Genes Dev. 5:1568-1576, 1991


Li H, Gomes PJ, Chen JD. RAC3, a steroid/nuclear receptor-associated coactivator that is related to SRC-1 and TIF2. Proc Natl Acad Sci USA. 94:8479-8484, 1997


Matysia-Scholze U and Nehls M. The structural integrity of RORα isoforms is mutated in staggerer mice: cerebellar coexpression of RORα1 and RORα4. Genomics. 43:78-84, 1997


McKenna NJ, Lanz RB, O'Malley BW. Nuclear Receptor Coregulators: Cellular and Molecular Biology. Endocrine Reviews. 20(3):321-344, 1999b


Michael MD, Kilgore MW, Morohashi KI & Simpson ER. Ad4 BP/SF-1 regulates cyclic AMP-induced transcription from the proximal promoter (P1) of the human aromatase P450 (CYP19) gene in the ovary. Journal of Biological Chemistry. 270:13561-13566, 1995


Nelson CC, Hendry SC. and Romaniuk PJ. Relationship between P-box amino acid sequence and DNA binding specificity of the thyroid hormone receptor. The effects of half-site sequence in everted repeats. J. Biol. Chem. 270:16981-16987, 1995


Onate SA, Boonyaratankornkit V, Spencer TE, Tsai SY, Tsai MJ, Edwards DP, O’MalleyBW. The steroid receptor coactivator-1 contains multiple receptor interacting and activation domains that cooperatively enhance the activation function 1 (AF1) and AF2 domains of steroid receptors. J Biol Chem. 273:12101-12108, 1998

Onate SA, Tsai SY, Tsai MJ and O’Malley BW. Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. Science. 270:1354-1357, 1995


205


Qiu YH, Tsai SY, Tsai UT: COUP-TF: An orphan member of the steroid/thyroid hormone receptor superfamily. Trends Endocrinol Metab. 5:234-239, 1994


Ring A and Dowsett M. Mechanisms of tamoxifen resistance. Endocrine-Related Cancer. 11:643-658, 2004


Roeder RG. Transcriptional regulation and the role of diverse coactivators in animal cells. FEBS Lett. 579:909-915, 2005

Rowan BG, Weigel NL, O'Malley BW. Phosphorylation of steroid receptor coactivator-1. Identification of the phosphorylation sites and phosphorylation through the mitogen-activated protein kinase pathway. J. Biol. Chem. 275:4475-4483, 2000


Sartorius CA, Melville MY, Hovland AR, Tung L, Takimoto GS, Horwitz KB. Molecular Endocrinology. 8: 1347-1360, 1994


Tjian R, Maniatis T. Transcriptional activation: a complex puzzle with few easy pieces. Cell 77:5-8, 1994

Topper YJ, Freeman CS. Multiple hormone interactions in the developmental biology of the mammary gland. Physiol Rev 60:1049-1106, 1980


Tsai SY, Tsai MJ, O'Malley BW. Steroid receptor coactivator-1 is a histone acetyltransferase. Nature 389:194-198, 1997


208


Wang LH, Tsai SY, Cook RG, Beattie WG, Tsai M-J, O'Malley BW. COUP transcription factor is a member of the steroid receptor superfamily. Nature. 340:163-166, 1989


Webster NJ, Green S, Tasset D, Ponglikitmongkol M, Chambon P. The transcriptional activation function located in the hormone-binding domain of the human oestrogen receptor is not encoded in a single exon. The EMBO Journal. 8(5):1441-1446, 1989


Wu RC, Smith CL, O’Malley BW. Transcriptional Regulation by Steroid Receptor Coactivator Phosphorylation. Endocrine Reviews. 26(3):393-399, 2005


http://genomics.nchresearch.org/gene_regulation_analysis.html

www.chemblink.com

www.invitrogen.com

www.pharmacology2000.com

www.promega.com
APPENDIX I

MATERIALS

Machines

Agitating Incubator 37°C (for LB Broth Cultures)  New Brunswick Scientific
Incubator 37°C with 5% CO₂ (for Cell Line Cultures)  Nuaire
Incubator 37°C (for Agar Plates)  Lab-Line
Water Bath 37°C  Fisher Scientific
AccuBlock Digital Dry Bath (42°C/70°C)  Labnet
Mini Centrifuge  ThermoIEC
Centrifuge (Avanti J-25)  Beckman Coulter
Microwave  GE Appliances
Spectrophotometer  Eppendorf (Biophotometer)
UV Trans-illuminator (photographic) (Image Master VDS-CL)  Amersham Biosciences
UV Trans-illuminator (non-photographic) (Fisher Biotech: 312 nm Variable Intensity Trans-illuminator FBTIV-614)  Fisher Scientific
PCR Engine  MJ Research
Cell Plate Rocker  CMS
Luciferase Microplate Luminometer (Centro LB 960)  Berthold Technologies
Laminar Flow Hood (Biogard Hood)  Baker Company Inc.
Axiavert 25 Inverted Microscope  Zeiss

Equipment
<table>
<thead>
<tr>
<th>Products</th>
<th>Manufacturers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pipettes</td>
<td>Eppendorf Research</td>
</tr>
<tr>
<td>Pipette Tips</td>
<td>Corning Incorporated</td>
</tr>
<tr>
<td>Ependorf Tubes</td>
<td>Sarstedt</td>
</tr>
<tr>
<td>14ml Polypropylene Round-Bottom Tube</td>
<td>Falcon</td>
</tr>
<tr>
<td>Stripette Pipettes</td>
<td>Costar</td>
</tr>
<tr>
<td>Ultra High Performance Centrifuge Tubes 15ml</td>
<td>VWR</td>
</tr>
<tr>
<td>Ultra High Performance Centrifuge Tubes 45ml</td>
<td>VWR</td>
</tr>
<tr>
<td>Vortex</td>
<td>Scientific Industries</td>
</tr>
<tr>
<td>Powder Free Latex Examination Gloves</td>
<td>VWR</td>
</tr>
<tr>
<td>PCR Tubes</td>
<td>Discovery Scientific</td>
</tr>
<tr>
<td>6 Well Plate Culture Plate</td>
<td>Falcon</td>
</tr>
<tr>
<td>24 Well Cell Culture Plate</td>
<td>Falcon</td>
</tr>
<tr>
<td>96 Well Plate</td>
<td>Greiner Bio-One</td>
</tr>
</tbody>
</table>

**Reagents**

- 10X Accuprime Buffer I: Invitrogen
  - 200 mM Tris-HCl (pH 8.4), 500 mM KCl, 15 mM MgCl2, 2 mM dGTP, 2 mM dATP, 2 mM dTTP, 2 mM dCTP, thermostable AccuPrime™ protein, 10% glycerol

- NEB Buffer 3: New England Biolabs
  - 100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl2, 1 mM Dithiothreitol

- NEB Buffer 4: New England Biolabs
50 mM Potassium acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, 1 mM Dithiothreitol

Prepared Solutions

TEN Buffer (pH 8.0)
20 μl Tris (pH 8.0), 500 μl EDTA (pH 8.0), 15 ml NaCl.

TAE Buffer
50X TAE Buffer
242 g Tris, 57.1 ml Acetic Acid, 100 ml 0.5 M EDTA pH 8.0
(Final concentrations of 2 M Tris, 2 M Acetic Acid, 0.05 M EDTA pH 8.0)
Dilute 1:50 dH₂O to get working solution.
40 mM Tris, 20 mM NaOAc, 1 mM EDTA, pH 8

HeLa Cell Lysis Buffer
0.1 ml 10 mM Tris pH 8.0, 0.2 ml 10 mM EDTA, 0.5 ml SDS (0.5%), 9.2 ml dH₂O
Prior to use 25.7 μl Proteinase K (200 μg/ml) was added to 2.5 ml of the stock solution.

TE Buffer pH 8.0
10X TE Buffer
5 ml 1 M Tris-HCl, pH 8.0, 1 ml 0.5 M EDTA, pH 8.0, 494 ml of water
(Final concentrations of 10 mM Tris-HCl, 1 mM EDTA)
**Luciferase Lysis Buffer**

25 mM TRIS pH 8.0, 150 mM NaCl, 0.5% Triton X-100, 1 mM EDTA, 5% Glycerol and 1M DTT (DTT added immediately prior to use)

**Agarose Gel (1%)**

25ml 1% Agarose (1mg Ultrapure Agarose powder in 100ml TAE Buffer, dissolved in microwave) and 2.5 µg Ethidium Bromide

**MetaPhor Agarose Gel (3%)**

25ml 3% MetaPhor Agarose (3mg MetaPhor Agarose powder (Cambrex) in 100ml TAE Buffer, dissolved in microwave) add 2.5 µg Ethidium Bromide

**Ampicillin Agar (50 µg/µl)**

16g LB Agar (Lennox L Agar, Invitrogen)

500mls dH₂O

250 µl Ampicillin (from 100 mg/ml Stock Solution)

**Spectinomycin Agar (50 µg/µl)**

16g LB Agar (Lennox L Agar, Invitrogen)

500mls dH₂O

250 µl Spectinomycin (from 100 mg/ml Stock Solution)

**Ampicillin LB Broth (100 µl/ml)**

12g LB Broth Base (Lennox L Broth Base, Invitrogen)

600mls dH₂O
600 µl Ampicillin (from 100 mg/ml Stock Solution)

**Spectinomycin LB Broth (100 µl/ml)**

12g LB Broth Base (Lennox L Broth Base, Invitrogen)

600mls dH2O

600 µl Spectinomycin (from 100 mg/ml Stock Solution)

**Double Stranded AdR (dsAdR) 50mM**

Dissolve 100mM each of AdRt and AdRb in dH2O. Combine equal volumes of AdRt (100 mM) and AdRb (100 mM). Mix and place in a tightly closed tube in a beaker with water of about 90°C. Let the beaker cool to room temperature (20°C), to enable the adaptors to anneal slowly. Aliquots were stored at -20°C.
APPENDIX II

Trans-Activation Assays
<table>
<thead>
<tr>
<th>Reporter Library</th>
<th>Expression Plasmid</th>
<th>Hormone Treatment</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 PR SRC-1A-DAM (24)</td>
<td>10ng hPR-B</td>
<td>10⁻⁷M Prog</td>
<td>Substitute Prog for 10⁻⁷M EtOH for each transfection</td>
</tr>
<tr>
<td>2 ERα SRC-1A-DAM (24)</td>
<td>5 ng hERα</td>
<td>10⁻⁷M E2</td>
<td>Substitute E2 for 10⁻⁷M EtOH for each transfection</td>
</tr>
<tr>
<td>3 PR SRC-1A-DAM (29)</td>
<td></td>
<td>10⁻⁷M Dex</td>
<td>pGRE-E1b-LUC as control (known to be responsive)</td>
</tr>
<tr>
<td>4* PR SRC-1A-DAM (21 confirmed sequences)</td>
<td>5ng hERα, 10ng hPR-B, 10ng hGRα</td>
<td>10⁻⁷M E2, 10⁻⁷M Prog, 10⁻⁷M Dex</td>
<td>200ng pGL3-Basic Dest with all gousps.</td>
</tr>
<tr>
<td>5 PR SRC-1A-DAM (21 confirmed sequences)</td>
<td>50ng pCR3.1, 25ng pCR3.1, 25ng SNAIL, 25ng SRC-1A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 PR SRC-1A-DAM (21 confirmed sequences)</td>
<td>25ng pCR3.1, 25ng COUP-TF I, 25ng COUP-TF II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 PR SRC-1A-DAM (21 confirmed sequences)</td>
<td>10ng pCR3.1, 10ng COUP-TF I, 10ng COUP-TF II, 10ng RORα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 P26, 27, 33, 40, 45</td>
<td>2ng pCR3.1 or COUP TF I or COUP TF II and 50ng pCR3.1 or 10ng RORα or 50ng pCR3.1 SRC1A</td>
<td></td>
<td>pGL3 basic control</td>
</tr>
<tr>
<td>9 P26, 27, 33, 40, 45</td>
<td>2ng pCR3.1 or COUP TF I or COUP TF II and 50ng pCR3.1 or 25ng RORα with 25ng pCR3.1 or 25ng RORα with 25ng pCR3.1 SRC1A</td>
<td></td>
<td>pGL3 basic control</td>
</tr>
<tr>
<td>10 P26, 27, 33, 40, 45</td>
<td>0.1ng pCR3.1 COUP TF I COUP TF II 0.25ng pCR3.1 COUP TF I COUP TF II 1ng pCR3.1 COUP TF I COUP TF II</td>
<td></td>
<td>pGL3 Basic</td>
</tr>
</tbody>
</table>
* Transfection Number 4.
Plate was divided into 3 groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Reporter</th>
<th>Receptors</th>
<th>Additional</th>
<th>Hormone</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Yes</td>
<td>Yes</td>
<td>50 ng pCR3.1</td>
<td>Hormone Premix</td>
<td>Substitute Hormone for EtOH 10^-7M</td>
</tr>
<tr>
<td>Group 2</td>
<td>Yes</td>
<td>Yes</td>
<td>50 ng pCR3.1 SRC 1A</td>
<td>Hormone Premix</td>
<td>Substitute Hormone for EtOH 10^-7M</td>
</tr>
<tr>
<td>Group 3</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Reporter Library</td>
<td>Expression Plasmid</td>
<td>Hormone Treatment</td>
<td>Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------</td>
<td>------------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>ERα DAM (28)</td>
<td>2.5ng hERα</td>
<td>10⁻⁷ M Ral</td>
<td>Substitute Ral for 10⁻⁷ M EtOH for each transfection 200ng pGL3-Basic and pERE-E1b-LUC</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ERα – DAM 20 from EtOH, E2 and 4HT</td>
<td>5ng hERα</td>
<td>10⁻⁷ M E2 10⁻⁷ M 4HT 10⁻⁷ M Ral</td>
<td>Substitute Hormone for 10⁻⁷ M EtOH for each transfection</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>SERM induction reporters (from E2 and 4HT Libraries)</td>
<td>5ng hERα (in all transfections) 25ng pCR3.1 25ng SNAIL 25ng SRC-3</td>
<td>10⁻⁷ M E2 10⁻⁷ M 4HT</td>
<td>Substitute Hormone for 10⁻⁷ M EtOH for each transfection pERE-E1b-LUC as positive control.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>E5p, T4p, E10</td>
<td>5 ng pCR3.1 5ng pCR3.1 hERα 5ng pCR3.1 hERα &amp; 10ng SNAIL 5ng pCR3.1 hERα &amp; 10ng PAX4 pCR3.1 hERα &amp; 50ng pCR3.1 SRC3</td>
<td>10⁻⁷ M EtOH 10⁻⁷ M E2 10⁻⁷ M 4HT 10⁻⁷ M Ral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>E5p, T4p, E10</td>
<td>5 ng pCR3.1 5ng pCR3.1 hERα 5ng pCR3.1 hERα &amp; 10ng SNAIL 5ng pCR3.1 hERα &amp; 10ng PAX4 pCR3.1 hERα &amp; 50ng pCR3.1 SRC3</td>
<td>10⁻⁷ M EtOH 10⁻⁷ M E2 10⁻⁷ M 4HT 10⁻⁷ M Ral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>E5p, T4p, E10</td>
<td>5ng pCR3.1 5ng hERα 5ng C205H 5ng L539A 5ng S118A</td>
<td>10⁻⁷ M EtOH 10⁻⁷ M E2 10⁻⁷ M 4HT 10⁻⁷ M Ral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>E5p, T4p, E10</td>
<td>5ng pCR3.1 5ng hERα 5ng C205H 5ng L539A 5ng S118A</td>
<td>10⁻⁷ M EtOH 10⁻⁷ M E2 10⁻⁷ M 4HT 10⁻⁷ M Ral</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX III

P-Library reporter plasmid sequences
APPENDIX IV

Gene-Racer sequences