COOPERATIVE BINDING MECHANISMS LEADING TO THE SPECIFIC ANDROGEN RECEPTOR REGULATION OF TARGET GENES

by

KIMBERLY JANE REID

B.Sc. Double Major, Royal Roads Military College, 1989
B.Sc. Honours, The University of Victoria, 1997

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR A MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Department of Experimental Pathology and Laboratory Medicine)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

June 2001

© Kimberly Jane Reid, 2001
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Experimental Pathology

The University of British Columbia
Vancouver, Canada

Date Aug. 27, 2001
Abstract

Genes uniquely regulated by the androgen receptor (AR) typically contain multiple androgen response elements (AREs) that in isolation are of low DNA binding affinity and transcriptional activity. However, specific combinations of AREs in their native promoter context results in highly cooperative DNA binding by AR and high levels of transcriptional activation. Within this study, we demonstrate that the natural androgen-regulated promoters of PSA and Probasin contain two classes of AREs dictated by their primary nucleotide sequence that function to mediate cooperativity. Class I AR-binding sites display conventional guanine contacts. Class II AR-binding sites have distinctive atypical sequence features and upon binding to AR the DNA structure is dramatically altered through allosteric interactions with the receptor. Class II sites stabilize AR-binding to adjacent Class I sites and result in synergistic transcriptional activity and increased hormone sensitivity. The specific nucleotide variation within the androgen receptor binding sites dictate was determined to dictate differential functions to the receptor. The potential role of individual nucleotides within Class II sites and predicted consensus sequences for Class I and II sites was also identified. Our data suggest that this may be a universal mechanism by which AR achieved unique regulation of target genes through complex allosteric interactions dictated by primary binding sequences.
**Table of Contents**

Abstract
Table of Contents
List of Figures
List of Tables
List of Abbreviations
Acknowledgements
Preface

Chapter 1. Introduction

1.1 The Androgen Receptor and Prostate Cancer
1.2 Normal Androgen Action
1.3 The Steroid Receptor Subfamily Makes Common and Discriminatory Protein-DNA Contacts
1.4 The Steroid Receptors Cooperatively Bind to a Common Steroid Response Element and Homodimerize Upon DNA Binding
1.5 Five Levels Through Which AR-Specific Gene Regulation Can Occur
1.6 The Architectural Context of AR-Regulated DNA Regions Underlies AR Cooperative Binding
1.7 Cooperativity Arises From Protein-Protein and Protein-DNA Interactions
1.8 The Probasin Promoter and the PSA Enhancer Possess Multiple Cooperatively Interacting AREs
   1.8.1 The Rat Probasin Promoter
   1.8.2 The Human PSA Enhancer
1.9 Objectives of this Study
Chapter 2. Experimental Procedures p. 29

2.1 Methylation Protection and Methylation Interference of the AR DBD p. 29

2.1.1 Preparing Recombinant DNA Binding Protein p. 32

2.1.2 Generation of Radioactively Labelled DNA Probes p. 33

2.1.2a Generating Single-End Labelled DNA Probes Using PCR Amplification p. 34

2.1.2b Generating Single-End Labelled DNA Probes Through Successive Enzymatic Digestion of Recombinant Plasmid Vectors p. 36

2.1.2c Generating Single-End Labelled DNA Probes by Annealing Synthesized Oligos p. 38

2.1.3 Methylation Protection Assay p. 38

2.1.4 Methylation Interference Assay p. 41

2.2 Determining the Binding Affinities of the Probasin Promoter AREs p. 42

2.2.1 Generating Radioactively Labelled Probasin Promoter Fragments With High Specific Activity p. 42

2.2.2 Determining Binding Affinity Through EMSA p. 43

2.3 Transfection Studies to Determine Effects of Probasin Promoter Fragments on Transcriptional Activation p. 44

Chapter 3. Results and Discussion p. 46

3.1 Methylation Interference Assay Failed to Footprint a Cooperative System p. 46

3.2 A Modified Methylation Protection Assay Revealed Atypical AR Binding Sites p. 46

3.3 Local DNA Structural Distortion Resulting From AR DBD Allosteric Binding was Intrinsic to the DNA Sequences of the G-1 and G-2 AREs p. 53
3.4 Hypersensitive Guanines in Class II AREs are Necessary for AR DBD Binding

3.5 The Human PSA Enhancer Contained Class I and Class II AREs and the Two Classes Possess Distinctive Nucleotide Features

3.6 Mutation of Key Nucleotides Transforms the MeP Pattern of a Class II Half-Site to a Class I Half-Site

3.7 Specific Arrangements of Class I and Class II AREs Increase Cooperative Binding

3.8 Class II AREs Contribute to Synergistic Transactivation as Observed in Transient Transfection Assays

3.9 Cooperative Protein-Protein Interactions Between AR Homodimers Could Involve Stabilizing Disulfide Bridges

3.10 Recombinant Sox-4 Protein Increases the Number of AR DBD Monomers that Bind the Probasin Promoter

3.11 MeP of Other AR-Regulated Promoter and Enhancer Regions Reveal Novel Protection Patterns that Suggest Monomer and Direct Repeat AR-DNA Binding

Chapter 4. Future Directions and Conclusion

4.1 Summary of Project Results

4.2 Future Directions

4.3 Conclusion

Bibliography
List of Figures

Figure 1. The process of androgen action in a prostate epithelial cell p. 4
Figure 2A and B. Rat androgen receptor DBD structure and DNA half-site contacts p. 5
Figure 3A and B. Structure of the ER homodimer bound to an ERE and the specific AR contacts with the SRE p. 7
Figure 4. The transcription factor-DNA multicomplex or enhanceosome recruits the basal transcriptional machinery p. 14
Figure 5. The rat probasin promoter from −426 to +28 base pairs p. 16
Figure 6. The human PSA enhancer from −4267 to −4026 base pairs p. 17
Figure 7. Schematic of the methylation protection assay p. 30
Figure 8. Schematic of the methylation interference assay p. 31
Figure 9. Failed MeP of AR-Histag on the probasin −426 to +28 base pair promoter fragment using the conventional MeP assay p. 48
Figure 10. 1.0 M ß-mercaptoethanol disrupted the AR DBD-probasin promoter complex p. 49
Figure 11. Modified MeP of the AR DBD-Histag on the probasin promoter from −269 to −77 base pairs p. 51
Figure 12. The probasin promoter sequence from −426 to +24 base pairs p. 52
Figure 13. MeP of the probasin promoter sequence from −269 through −77 base pairs p. 54
Figure 14. MeP of the probasin promoter sequence from −269 through −164 base pairs p. 55
Figure 15. MeP of the probasin promoter sequence from -150 through -77 base pairs p. 56
Figure 16. MeP of the individual G-site probasin promoter AREs p. 58
Figure 17. Mel of the G-1 and ARE2 from the probasin promoter p. 59
Figure 18A. MeP of the human PSA enhancer from -4267 to -4062 base pairs p. 61
Figure 18B and C. MeP of the PSA enhancer ARE IIIA and the PSA enhancer sequence from -4267 to -4062 base pairs p. 62
Figure 19A-D The change in the relative DMS methylation hypersensitivity and protection pattern of the probasin G-1 element after site-directed mutagenesis p. 66
Figure 20. Binding Affinity Curves p. 70
Figure 21A-C. Transient Transfection Curves p. 72
Figure 22A and B. EMSA showing β-mercaptoethanol disruption of the AR DBD-DNA binding complex p. 76
Figure 22C. 1.0 M of β-mercaptoethanol had negligible effect on AR DBD binding to the canonical SRE p. 77
Figure 23. Sox-4 protein increases the occupancy of AR DBD on the probasin promoter p. 79
Figure 24A. Methylation protection of the AR DBD on the mE-RABP promoter from -543 to -166 base pairs p. 82
Figure 24B. Methylation protection of the AR DBD on the mE-RABP promoter from -195 to +26 base pairs p. 83

Figure 25. MeP of the AR DBD on the rat AR enhancer from 2028 to 2333 base pairs p. 84

Figure 26. MeP of the human p21 promoter from -285 to -1 base pairs p. 85

Figure 27. Methylation protection of the AR DBD on the upstream proximal probasin promoter from -705 to -426 base pairs p. 86
List of Tables

Table 1. Nuclear receptor P Box and D box sequence comparison p. 9

Table 2. Alignment of AREs into proposed Class II and Class I consensus sequences p. 64
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>-alanine</td>
</tr>
<tr>
<td>AR</td>
<td>-androgen receptor</td>
</tr>
<tr>
<td>ARE</td>
<td>-androgen receptor response element</td>
</tr>
<tr>
<td>BL13</td>
<td>-human transitional cell bladder carcinoma</td>
</tr>
<tr>
<td>BMR</td>
<td>-BioMax Resolution (Kodak film)</td>
</tr>
<tr>
<td>BMS</td>
<td>-BioMax Sensitive (Kodak film)</td>
</tr>
<tr>
<td>CAT</td>
<td>-chloramphenicol acetyl transferase</td>
</tr>
<tr>
<td>CHO</td>
<td>-Chinese hamster ovary cell line</td>
</tr>
<tr>
<td>CIP</td>
<td>-calf intestinal phosphatase</td>
</tr>
<tr>
<td>DBD</td>
<td>-DNA binding domain</td>
</tr>
<tr>
<td>D Box</td>
<td>-Distal Box</td>
</tr>
<tr>
<td>DHT</td>
<td>-dihydrotestosterone</td>
</tr>
<tr>
<td>DMS</td>
<td>-dimethyl sulfate</td>
</tr>
<tr>
<td>dNTP</td>
<td>-dideoxynucleotide triphosphate</td>
</tr>
<tr>
<td>DU145</td>
<td>-prostate cancer metastasis to the brain cell line</td>
</tr>
<tr>
<td>EMSA</td>
<td>-electrophoretic mobility shift assay</td>
</tr>
<tr>
<td>ER</td>
<td>-estrogen receptor</td>
</tr>
<tr>
<td>ERE</td>
<td>-estrogen receptor response element</td>
</tr>
<tr>
<td>FCS</td>
<td>-fetal calf serum</td>
</tr>
<tr>
<td>FRET</td>
<td>-fluorescence resonance energy transfer</td>
</tr>
<tr>
<td>GR</td>
<td>-glucocorticoid receptor</td>
</tr>
<tr>
<td>HepG2</td>
<td>-human hepatic cell line</td>
</tr>
<tr>
<td>hGK-1</td>
<td>-human glandular kallikrein1</td>
</tr>
<tr>
<td>HisTag</td>
<td>-histidine tagged</td>
</tr>
<tr>
<td>HMG</td>
<td>-high mobility group</td>
</tr>
<tr>
<td>Hsp70</td>
<td>-heat shock protein 70</td>
</tr>
<tr>
<td>Hsp90</td>
<td>-heat shock protein 90</td>
</tr>
<tr>
<td>IFN-β</td>
<td>-interferon beta</td>
</tr>
<tr>
<td>Ig-κB</td>
<td>-immunoglobulin-κB</td>
</tr>
</tbody>
</table>
KLK-1 - kallekrein-1
LBD - ligand binding domain
LNCaP - Lymph Node (metastases) Prostate Cancer cell line
LPB - large probasin promoter (~12 kbp)
Matα/MCM1/STE6 - multicomplex of mating-type α protein, minichromosome maintenance-1 protein and the α-cell-specific transporter gene
MCF-7 - human breast carcinoma cell line
MeI - methylation interference
mE-RABP - mouse epididymus retinoic acid binding protein
MeP - methylation protection
MMTV - mouse mammary tumour virus
MR - mineralocorticoid receptor
NAPS - nucleic acid and protein synthesis
NFAT/Fos-Jun/ARRE2 - multicomplex of the nuclear factor of activated T cells, the nuclear phosphoproteins Fos and Jun, and the ARRE2 element within the interleukin-2 promoter
NF-κB/HMG I/PD II - multicomplex of the nuclear factor-κB protein, the high mobility group I protein, and the positive regulatory domain II within the endothelial leukocyte adhesion molecule 1 gene promoter
Oct-1 - octamer-1 protein
PAGE - polyacrylamide gel electrophoresis
P Box - proximal box
PC3 - prostate cancer 3 cell line
PCR - polymerase chain reaction
PNK - polynucleotide kinase
PR - progesterone receptor
Probasin - prostatic basic protein
PSA - prostate specific antigen
RLU - relative luciferase unit
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1881</td>
<td>synthetic androgen methyltrienolone</td>
</tr>
<tr>
<td>S1p</td>
<td>sex limited protein</td>
</tr>
<tr>
<td>Sox-4</td>
<td>Sry-related HMG box protein 4</td>
</tr>
<tr>
<td>SRC-1</td>
<td>steroid receptor coactivator-1</td>
</tr>
<tr>
<td>SRE</td>
<td>steroid receptor response element</td>
</tr>
<tr>
<td>SV40</td>
<td>simian virus 40</td>
</tr>
<tr>
<td>T</td>
<td>testosterone</td>
</tr>
<tr>
<td>Thr</td>
<td>threonine</td>
</tr>
<tr>
<td>TR</td>
<td>thyroid hormone receptor</td>
</tr>
<tr>
<td>WT</td>
<td>wildtype</td>
</tr>
<tr>
<td>293</td>
<td>human embryonic kidney cell line</td>
</tr>
</tbody>
</table>
Acknowledgements and Dedication

I would like to gratefully acknowledge and thank my supervisor, Dr. Colleen Nelson whose enthusiasm and dedication to science has been a tremendous example. Within and without the Prostate Centre I would like to thank the following people for their encouragement, ever-present support and willingness to critically read this thesis: Jody Saito, Doug Hoffart, Stephen Hendy, Helen Cheng, Dr. Pepita Gimenez-Bonafe, Mira Rao, Lillian Yeung, Dr. Jason Read, Dr. Paul Rennie, Tanis Whitmore and Dr. Sujata Presard.

This thesis is dedicated to my husband, Phillip A. Rennison, for his limitless patience and encouragement.
Preface.

Results and conclusions contained within this study have been published previously in three separate journals. The methylation protection, binding affinity and transient transfection results pertaining to the proximal probasin promoter and upstream PSA enhancer were published in the Journal of Biological Chemistry. The methylation protection results pertaining to the mouse epididymus-retinoic acid binding protein promoter were published in the journal, Biology of Reproduction. Finally, the novel methylation protection technique described in this study was previously published in the journal, Biotechniques.

The publications in detail are as follows:


1. Introduction

1.1 The Androgen Receptor and Prostate Cancer

The prostate gland is a male accessory sex organ that is extremely sensitive to androgens, such as testosterone and dihydrotestosterone (DHT), and to their cognate receptor, the androgen receptor (AR) (Petrow et al, 1986; Frick et al, 1991). A correctly functioning AR and sufficient androgen hormone are necessary for the normal growth and development of prostate epithelial tissue (Isaacs et al, 1994; Thomas et al, 2001). Mediated by their action through AR, androgens appear to control prostate epithelial cell number by both stimulating the rate of cell proliferation and inhibiting the rate of cell death (Isaacs et al, 1994). Removal of androgens will invoke programmed cell death and, because of their reliance on androgens for survival, prostate epithelial cells are referred to as androgen dependent (Isaacs et al, 1994). From the above, it is apparent that androgens and the AR are critically involved in a complex cell-signaling network that is responsible for prostate epithelial cell growth, development and maintenance.

Like healthy prostate, cancerous prostate epithelial cells require androgens for proliferation. Although a causative role for the androgen/AR complex in the onset of prostate cancer has yet to be shown, upregulation and increased mutation frequency of the AR, leading to a gain in function, is a hallmark of advanced staged disease and could contribute to metastasis and progression (Hobisch et al, 1995; Taplin et al, 1995; Marcelli et al, 2000). Mutated AR can be activated by steroids or agonists other than androgens: within the cultured cell line LNCaP, which is derived from a prostate cancer metastasis to the lymph node, the endogenous AR contains a Thr→Ala mutation at position 877 that permits a broadened specificity beyond androgens for other steroids and steroid-like molecules such as progesterone, oestradiol and hydroxyflutamide (MacDonal et al, 2000). This particular AR mutation and others that lead to ligand promiscuity, were found in several prostate cancer metastases (Suzuki et al, 1993; Gaddipato et al, 1994). In addition, the AR was amplified in prostate cancers such that low levels of androgens are able to induce a biologically active response (Hobisch et al, 1995; Koivisto et al, 1995; Visakkorpi et al, 1995; Linja et al, 2001). The adrenal androgens, androstenedione and dehydroepiandrosterone, normally
contribute less than 10% of the precursor androgens to the prostate gland, however this low level may be enough to be biologically active if the number of AR molecules is highly amplified (Culig et al, 1996, Tan et al, 1997). AR activity appears to be critical to not only normal prostate growth but also to the growth and progression of prostate cancer.

The primary action of steroid hormones, such as androgens, is to stimulate target gene expression by stimulating rates of mRNA production (Higgins and Gehring, 1978). The selective and specific binding of steroid hormone-receptor complexes to target gene regulatory regions regulates transcription of mRNA (Gronemeyer and Pongs, 1980; Payvar et al, 1981). Surprisingly, it is not clear how the AR specifically and uniquely regulates the transcription of target genes that ultimately controls the growth, development and homeostasis of prostate cells. Bayes’ Theorem states that the more we know about a system, the better we can predict its outcome. If we apply Bayes’ Theorem to the role of androgens and the AR in prostate cancer, then the more we know about normal androgen action, the better we will gain insight in how prostate cancer progresses and how it may be controlled.

1.2 Normal Androgen Action

Androgens, such as testosterone, belong to the steroid hormone family that includes progesterone, estradiol, glucocorticoids and mineralocorticoids. Androgens are responsible for the normal development, differentiation, growth and homeostasis of several organs including, among others, the human male prostate. It is accepted within the steroid field that androgen action is primarily responsible for the above cellular effects. Androgen action is a molecular and biological process whereby the extracellular androgenic signal culminates in direct transcriptional regulation by the AR, a transcription factor that primarily employs the androgens, testosterone or DHT, as activating ligands (Brinkmann et al, 1989; Govindan et al, 1991).

Due to their lipophilic nature, androgens are able to easily enter the plasma membrane that separates the cell’s cytoplasm from the extracellular environment. How lipophilic androgens enter into the aqueous environment of the cytosol is not clearly understood. However, once androgens do enter the cell they are often converted to a more active form (Figure 1). Within prostate epithelial cells the enzyme 5α-reductase converts testosterone to
DHT, which has a higher affinity for the ligand-binding domain (LBD) of the AR than its precursor, testosterone.

The cognate receptor for androgens is the AR, which is a ligand-activated transcription factor that belongs to the superfamily of nuclear hormone receptors. The AR can be further assigned to the steroid receptor subfamily that includes the receptors for glucocorticoid, mineralocorticoid and progesterone (respectively GR, MR and PR) (Laudet et al, 1992). Like the GR, correctly folded AR is normally chaperoned within the cytoplasm by the heat shock proteins hsp70 and hsp90 along with other co-chaperone molecules (Ohara-Nemoto et al, 1988; Veldscholte et al, 1992) (Figure 1). After DHT binds the AR LBD, the AR/DHT complex is thought to dissociate from the chaperone multiplex and translocate to the nucleus. Once inside the nucleus the AR regulates transcription by specifically binding to androgen response elements (AREs) located within the promoter and enhancer regions of target genes.

1.3 The Steroid Receptor Subfamily Makes Common and Discriminatory Protein-DNA Contacts

All steroid receptors are capable of homodimerizing in a DNA-dependent manner through contacts mediated by their DNA binding domain and some are capable of homodimerizing in solution through contacts mediated by their LBD or DNA hinge region (Tsai et al, 1988; Dahlman-Wright et al, 1990; de Vos et al, 1993; Savory et al, 2001). In solution and not bound to DNA, the AR is able to homodimerize at concentrations well above physiological conditions (> 0.2 μM), however, AR primarily exist as a monomer in solution and below stochiometric concentrations (Liao et al, 1999). After activation by ligand, a steroid receptor monomer located in the nucleus will recognize a six base pair DNA half site with the consensus sequence AGAACA (Scheidereit et al, 1983; Cato et al, 1988; Roche et al, 1992). All four steroid receptors share a highly homologous DNA binding domain (DBD) that contains two zinc-binding motifs where each motif is closely followed by a short amphipathic α-helix (Figure 2A). The α-helix immediately following the first zinc motif is called the DNA recognition helix and contains the steroid receptor DNA recognition motif GSCKV called the Proximal Box (P Box) (Umesono and Evans, 1989).
Figure 1. The process of androgen action. Testosterone (T) enters the cytoplasm, is converted to DHT, which in turn binds the AR. The Hsp70/90 chaperone complex releases the AR/DHT complex, which then translocates to the nucleus where it binds an ARE and regulates gene transcription.
**Figure 2. Rat androgen receptor DBD structure and DNA half-site contacts.**

**A.** The DNA recognition P Box and dimerization D Box are outlined with the contacting and dimerizing residues indicated in bold (Umesono and Evans 1989). **B.** An AR monomer will interact with a steroid receptor canonical DNA half-site. The lysine (K) and arginine (R) residues make hydrogen bonds with the indicated guanines within the half-site. The Van der Waal contact between the valine (V) residue and the thymidine is discriminatory for the steroid receptor subfamily.
In general, the zinc motif and α-helical structure of the DBD is highly conserved among members of the nuclear receptor superfamily.

Upon recognition of the AGAACA half site, the steroid receptor DNA recognition α-helix makes base pair contacts using amino acid residues that are universally conserved throughout the nuclear receptor superfamily. The conserved lysine located in the steroid receptor P Box GSCKV makes a hydrogen bond essential for binding with the first guanine in the half site AGAACA (Figure 2B) (Umesono and Evans, 1989). This conserved lysine is also present in the estrogen receptor/thyroid hormone receptor (ER/TR) subfamily of nuclear receptors within the P Box sequences EGCKA, EGCKG, EGCKS, and EACKA. A second essential protein-DNA contact conserved throughout the nuclear receptor superfamily is made by an arginine, which is located outside of the P-box but within the DNA recognition α-helix. This arginine makes a hydrogen bond to a conserved guanine base paired to cytosine in the steroid receptor half site AGAAC(A) (Figure 2B).

The X-ray crystallographic structure of the AR bound to its DNA half site is still forthcoming, however, such an assembly has been achieved for the GR although the resolved structure contained DNA half sites that were separated by four nucleotides rather than the requisite three (Luisi et al, 1991). Since the steroid receptors, except ER, share a highly homologous DNA recognition α-helix and appear to recognize and bind to the same DNA half site, it is feasible to extrapolate from the GR/DNA crystal structure those protein-DNA contacts that are specific to the steroid receptor subfamily. The crystal structure of the GR homodimer bound to DNA has shown that each GR monomer DNA recognition α-helix lies within the major groove at right angles to the DNA helix and makes contacts with DNA through weak interactions such as hydrophobic and hydrogen bonds (Luisi et al, 1991). The same recognition α-helix/major groove orientation has been demonstrated for the ER/DNA crystal structure (Figure 3A) (Schwabe et al, 1993). The protein-DNA interactions revealed by both the GR/DNA and the ER/DNA crystal structures include the lysine-guanine and arginine-cytosine hydrogen bonds that are common to all nuclear receptors (Figure 2B). In addition, the GR/DNA structure illustrated that a valine found within the steroid receptor P Box GSCKV makes a specific and discriminating van der Waals contact with the methyl group of the thymidine base paired to A in the half site AGAAC(A) (Luisi et al, 1991; Schwabe et al, 1993).
Figure 3. A. Structure of the ER homodimer bound to a consensus ERE. The ER recognition α-helices lie at right angles to the DNA axis within the major groove. The D Boxes located within the second zinc motif mediate dimerizing salt bridges between the two ER monomers. Structure was visualized using Chime software and obtained from the Protein Data Bank (Schwabe et al, 1993). B. Two AR monomers bind two half sites within the SRE. The K and R residues make hydrogen bonds with the indicated guanines and the V residue makes a hydrophobic interaction with the thymidine methyl group.
This valine-thymidine contact is discriminating for the steroid receptor subfamily because the alanine within the ER P Box \textsc{EGCKA} sterically restricts the ER from binding to a DNA sequence containing an adenine in the fourth position of the half site AGA\textsc{ACA}.

1.4 The Steroid Receptors Cooperatively Bind to a Common Steroid Response Element and Homodimerize Upon DNA Binding

Members of the ER/TR subfamily of nuclear receptors also recognize a common half site AGGTCA. Specificity of response, however, is primarily obtained through the spacing and orientation of pairs of half sites found within gene regulatory regions (Umesono and Evans, 1989, Umesono et al, 1991). In contrast, the steroid receptors all commonly recognize, bind to, and homodimerize on an imperfect palindrome comprised of two six base pair half sites in reverse or head-to-head orientation. These two half sites are usually separated by three base pairs as represented by the steroid response element (SRE) 5' GGTACAnnnTGTTCT 3', which is an optimized sequence derived from the mouse mammary tumor virus (MMTV) promoter (Figure 3B) (Beato et al, 1987; Ham et al, 1988). Upon binding to the first half site in the SRE, the receptor monomer undergoes a conformation change exposing the dimerizing Distal Box (D Box) (Holmbeck et al, 1998; van Tilborg et al, 2000) (Figure 2A). Unlike the ER/TR subfamily, the amino acid residues within the D Box are highly homologous among members of the steroid receptor subfamily, which may account for the rigidity of half site spacing and orientation (Table 1).

The allosteric conformation change that results from steroid receptor binding to a DNA half site subsequently facilitates the binding of a second monomer to the neighbouring half site in reverse orientation. Once bound to the SRE, the two receptors homodimerize by forming salt bridges between the arginine and aspartic acid residues found in the newly exposed D Boxes within each steroid receptor monomer, similarly depicted in the DNA-bound ER homodimer crystal structure (Figure 3A) (Luisi et al, 1991; Schwabe et al, 1993).
## Sequence of P and D Boxes from select members of the steroid receptor and ER/TR subfamilies

<table>
<thead>
<tr>
<th>Receptor</th>
<th>P Box</th>
<th>D Box</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Steroid Receptors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GR, MR, PR</td>
<td>GSCKV</td>
<td>AGRND</td>
</tr>
<tr>
<td>AR</td>
<td>GSCKV</td>
<td>ASRND</td>
</tr>
<tr>
<td><strong>ER/TR Subfamily</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER</td>
<td>EGCKA</td>
<td>PATNQ</td>
</tr>
<tr>
<td>TRα</td>
<td>EGCKG</td>
<td>KYDSC</td>
</tr>
</tbody>
</table>

**Table 1.** Nuclear receptor superfamily P Box and D Box sequence comparison. Amino acid sequence variation and similarity of the P Boxes and D Boxes among select members of the nuclear receptor superfamily are shown (adapted from Umesono and Evans, 1989).
The facilitation of DNA binding between two steroid receptor monomers to an SRE is termed cooperative binding. Essentially, cooperative interactions between the two steroid receptor monomers are facilitated by the two DNA half sites within an SRE, resulting in greater-than-additive protein occupation (Tsai et al, 1988; Holmbeck et al, 1998). Cooperative binding has the effect of increasing overall binding affinity since protein occupation of the DNA binding site will occur more rapidly than if occupation was merely additive or linear. Furthermore, the DNA-facilitated cooperative binding of two steroid receptor monomers will ensure that the two half sites are saturated at a lower concentration of protein than if there was no facilitation between DNA half sites. All four steroid receptors (AR, PR, GR and MR) will bind cooperatively to an individual SRE with high affinity and sensitivity (Majors et al, 1983; Ham et al, 1988; Rundlett et al, 1995). There appears to be, therefore, a universal mechanism by which the members of the steroid receptor family recognize and cooperatively bind to a common DNA binding site.

Early data on the DNA binding characteristics of the steroid receptors GR, MR, PR and AR revealed that all four are capable of binding in vitro with high affinity to the SRE and all four are capable of transactivating from the SRE in transient transfection assays (Majors et al, 1983; Ham et al, 1988; Rundlett et al, 1995). These observations should not come as a surprise since the steroid receptor DNA binding domains (DBDs) possess remarkably high homology to each other. As well, the MMTV viral promoter, from which the SRE is derived, would likely benefit from exploiting a variety of endocrine signaling pathways. Such promiscuity of response is likely a desirable trait for a viral promoter in order to capitalize on a host’s cellular system. However, in mammalian cells each of the four steroid receptors specifically target a unique set of genes and the resulting gene expression culminates in quite different phenotypes. The dilemma, therefore, is in understanding how the AR specifically and uniquely regulates genes when it appears to recognize and bind to the same DNA binding element as the other three steroid receptors.
1.5 Five Levels Through Which AR-specific Gene Regulation Can Occur.

There are at least five levels through which AR-specific gene regulation can occur in light of the above binding site dilemma. They are as follows: (1) the availability of constituents such as activated hormone and its activated cognate hormone receptor; (2) the role of tissue specific coregulators; (3) the accessibility to chromatin; (4) the role of other transcription factors, and finally; (5) the role of the DNA architecture of natural AR-regulated promoter and enhancer regions.

Firstly, the availability of constituents would be an obvious solution to AR-specific activation. If only androgen and activated AR were available within a cell then receptor competition for DNA binding sites would be nonexistent. Reality denies this simple solution, however, since PR, GR and AR and their cognate ligands are found to coexist in many cell environments and activated GR tends to be present in a much higher concentration than AR (Stumpf et al, 1976; Chaudhuri et al, 1981; Saatok et al, 1984; Konrad et al, 1998; Yamashita et al, 2001).

The second solution is that within the cellular environment, coregulators exist that are capable of directing an AR-specific response. In other words, unique coregulators direct and modulate not only AR-DNA interactions but also the subsequent protein contacts that are necessary for an AR-specific transcriptional response. The search for unique AR-specific and tissue-specific coregulators is the current focus of a tremendous amount of research within the steroid receptor field. Many proteins have been found that interact with the AR although there has yet to be discovered a coregulator that uniquely and specifically interacts with the AR. Furthermore, the majority of research thus far involves using molecular techniques that fail to take into consideration the contribution of DNA to the AR-coregulator interaction. From inference from other nuclear receptors such as ER and GR, we know that upon binding to DNA the AR undergoes a conformation change that has functional significance, primarily the effect of receptor homodimerization and cooperative binding to an ARE (Liao et al, 1999). Using an immunoprecipitation assay or GST-pull down assay will not reveal critical coregulators that interact with the AR while it is bound to an ARE. Furthermore, such assays will not reveal critical coregulator interactions while the AR and the DNA are in different and transcriptionally relevant conformations.
One study, however, has examined the effect of the DNA-bound AR conformation on subsequent coregulator and other transcription factor interactions. This study showed that the AR will interact with the octamer transcription factor (Oct-1) when both transcription factors are bound to their respective response elements within the mouse sex limited protein (Slp) gene and that this interaction is dependent on the DNA-bound state of both proteins (Gonzalez and Robins, 2001). Furthermore, this group showed that the steroid receptor coactivator-1 (SRC-1) is actively recruited to the Oct-1/AR/DNA complex. These results were specific to the AR since the DNA-bound GR was unable to interact with Oct-1 and subsequently failed to recruit SRC-1 even though GR was capable of binding the Slp ARE. While these results show a role for coregulators in effecting and modulating an AR-specific response they also suggest that the DNA context is essential for the correct hormonal response. Therefore, the exclusive role of coregulators in AR-specific transactivation has yet to be shown and is probably unlikely.

The third solution pertains to DNA accessibility. AR-specific regulation of target genes relies on the accessibility of regulatory regions. GR, PR or AR can moderately transactivate from the MMTV promoter in transient transfection assays where the promoter is subject to only randomly arrayed nucleosomes (Richard-Foy H and Hagar GL 1987, Archer TK 1992, Truss M 1995). However, when the MMTV promoter is stably transfected, encased within the genome, and assembled with an ordered nucleosome array, the accessibility and subsequent transactivation response can be narrowed to a single steroid receptor type depending on the site of integration (Lambert and Nordeen, 1998; List et al, 1999). This suggests that chromatin and accessibility to regulatory regions can play at least a restrictive role in a selective response.

The fourth solution pertains to the role of other transcription factors in effecting an AR-specific response. Careful examination of the regulatory regions of several AR-regulated genes has revealed that the binding of transcription factors other than the AR are critical to the AR-specific response (Scarlett et al, 1995; Darne et al, 1997; Ning et al, 1999; Lu et al, 2000). The binding of the transcription factor Sp1 to the p21 promoter has been shown to be essential to the AR-specific regulation of this gene (Lu et al, 2000). Like coregulators, however, the likelihood is small that a specific transcription factor will be found that is solely responsible for the observed AR-specific response. So far, the transcription factors like Oct-
1 and Sp1 that interact with the AR are generally ubiquitous among cell types and involved in the regulation of several genes.

The final proposed solution to the observed AR-specific response lies within the architectural context of the natural promoter and enhancer regions found within AR-regulated genes. Architectural context refers to the primary and flanking nucleotide sequence, spacing, and relative orientation of the protein binding sites found within gene regulatory regions. The ‘talking DNA’ hypothesis suggests that the architectural context of DNA conveys information to a bound protein through allostery that will dictate conformation and function (Stephen Hendy, personal communication). Allosteric interactions are necessarily two-way, which means that a specific protein bound to DNA will also impart a DNA conformation change that has functional significance. This hypothesis incorporates the roles of coregulators and other transcription factors in generating an AR-specific response: DNA sequence could direct not only the specific transcription factors that bind to neighbouring sites but also the consequent correct conformation of the protein-DNA transcription regulatory multiplex, termed the enhanceosome, which is necessary to recruit coregulators that effect maximal transactivation of the gene in question (Figure 4) (Thanos et al., 1995; Merika and Thanos, 2001). These DNA sequence-directed allosteric interactions could function to increase protein binding affinity, sensitivity and specificity through cooperative binding, which will ultimately lead to the synergistic transcriptional activation or repression of targeted genes (Carey et al, 1998; Ellwood et al, 1999; Wang et al, 1999).

1.6 The Architectural Context of AR-Regulated DNA Regions Underlies AR Cooperative Binding

When natural promoter and enhancer regions of known AR-regulated genes are examined they are generally found to comprise of multiple AREs that individually vary in primary nucleotide sequence from the high affinity SRE (Kasper et al, 1994; Cleutjens et al, 1996; Grad et al, 1999; Lin et al, 2000). An ARE found in a mammalian gene regulatory region possesses nucleotide divergence that results in a relatively low affinity for the AR when compared to the high affinity SRE (Kasper et al, 1999; Verrijdt et al, 2000). This is seemingly a contradictory observation since intuition would suggest that individual AREs
The enhanceosome complex comprises of transcription factors and coregulators cooperatively recruited to the DNA regulatory region. The multicomplex that results is termed the ‘enhanceosome’, which will recruit the basal transcriptional machinery to the promoter.
found within the promoter and enhancer regions of AR-regulated genes would naturally have a high affinity for the AR. This interesting contradiction was the basis of a theory that the nucleotide divergence found within an ARE could have evolved to increase AR-specificity in a competitive steroid receptor environment but at a cost to binding affinity (Nelson et al., 1999). It was proposed that natural AREs not only evolved through the selection of nucleotides that contributed additional bond energy but also through the selection of nucleotides that discouraged binding of inappropriate receptors.

A binding site selection assay showed that, when in direct competition with either GR or PR, the AR would selectively bind to lower affinity response elements consequently restricting specificity to the AR (Nelson et al., 1999). This study and others also showed that if the third nucleotide in the core sequence AGAACA is altered from an adenine to a thymidine or a guanine, then AR- and PR-mediated transcriptional activity increased despite a lowering of DNA binding affinity for both AR and PR (Tan et al., 1992; Lieberman et al., 1993; Nelson et al., 1999). Together these observations suggest that nucleotides, particularly those that are evolutionarily conserved, that do not engage in base-specific bonds can nevertheless affect function. In other studies it has been shown that particular nucleotides are discriminated against. This is likely due to the incompatibility of a receptor DBD structure with a given sequence, which provides another level of discrimination (Nelson et al., 1999; Zilliacus et al., 1994; Gewirth et al., 1995; Nelson et al., 1995). Overall, there appears to be general core nucleotide requirements for an ARE, but there is also nucleotide variation within the core binding site and flanking sequence that provides function apart from receptor binding affinity.

The AREs that are found within natural AR-regulated regions individually have relatively low affinity albeit high specificity for the AR. Interestingly, like other steroid receptor response elements, when multiple AREs are examined collectively and in their natural promoter context they cooperatively interact to dramatically increase overall affinity for the AR (Tsai et al., 1989; Schule et al., 1988; Strahle et al., 1988). The rat probasin gene proximal promoter and the human prostate specific antigen (PSA) gene enhancer are two well-studied AR- and prostate-specific gene regulatory regions that contain multiple cooperatively interacting AREs (Figure 5 and 6) (Kasper et al., 1994; Huang et al., 1999). As observed with other
Figure 5. The rat probasin promoter from -426 to +28 base pairs. Two AREs were discovered that cooperatively interacted to increase overall binding affinity for the AR and they are ARE1 at position -241 to -227 base pairs and ARE2 at position -136 to -122 base pairs (Kasper et al, 1994).
Figure 6. The human PSA enhancer from -4267 to -4062 base pairs. Four AREs were discovered that cooperatively interact to increase AR binding affinity and to direct AR-specific PSA expression. These AREs are as follows; IIIA from -4065 to -4079 base pairs; III from -4133 to -4148 base pairs; IV from -4175 to -4189 base pairs; and V from -4220 to -4134 base pairs (Huang et al, 1999).
steroid receptors, the binding of one AR homodimer to an ARE will facilitate the binding of AR homodimers to neighbouring AREs (Tsai et al, 1989; Kasper et al, 1994; Cleutjens et al, 1996; Grad et al, 1999; Huang et al, 1999; Lin et al, 2000).

The cooperative assembly of protein-DNA complexes is widely observed in both prokaryotes and eukaryotes. It is generally accepted that cooperative binding is universally involved in mediating the assembly of specific transcription factor complexes, which culminate in the specific, synergistic and maximal transcriptional activation of gene expression (Ankenbauer et al, 1988; Tsai et al, 1989; Ohmori et al, 1997; Vashee et al, 1998 pp 530; Vashee et al, 1998 pp452; Ellwood et al, 1999; Merika et al, 2001). The specific facilitation that characterizes cooperativity exhibits three features that contribute to synergistic transactivation and they are as follows: (1) 

**Increased binding affinity of the transcription factor for its DNA binding element.** DNA binding sites that exhibit cooperativity facilitate the binding of transcription factors and architectural proteins such that the binding of one protein to its DNA response element assists the binding of other proteins to neighbouring DNA sites. This facilitation results in greater-than-additive or non-linear binding site occupation and directly translates into increased binding affinity (Mao et al, 1994; Liu et al, 1998; Senear et al, 1998); (2) 

**Increased sensitivity of the DNA binding site to transcription factor concentration.** Cooperatively interacting sites become occupied at a much lower concentration of transcription factor compared to binding sites that are occupied in a linear and independent manner (Mao et al, 1994; Senear et al, 1998). Independent or unassisted occupation of a DNA binding site would occur at a much higher protein concentration than that required for those proteins assisted by neighbouring DNA-bound proteins. Additionally, this increased sensitivity functions to serve as a molecular switch where small changes in factor concentration will result in the full cooperative occupation of the DNA regulatory region (Liu et al, 1998). Increased sensitivity is a hallmark of true cooperative interactions and its observation will help discriminate cooperative binding from DNA binding sites that possess high affinity but do not interact cooperatively (Carey et al, 1998); and (3) 

**Increased transcription factor specificity for DNA binding elements.** Cooperative interactions between transcription factors are typically specific suggesting that only the correct combination of transcription factors will occupy the gene regulatory region (Mo et al, 1998; Bhoite et al, 1998; Senear et al, 1998; Szymczyna and Arrowsmith, 2000).
This specificity of interaction arises from allostery. When a protein factor and DNA interact there is a conformation change in both the protein and the DNA that will affect the type of subsequent protein-protein and protein-DNA interactions (Holmbeck et al, 1998; Kerppola et al, 1998). The resultant protein-DNA conformation change would ensure that only the correct factors are recruited and encouraged to bind to DNA.

1.7 Cooperativity Arises From Protein-Protein and Protein-DNA Interactions

There are several well-studied eukaryotic cooperative systems that include the Matα2/MCM1/STE6 yeast α-specific gene promoter complex, the NFAT/Fos-Jun/ARRE2 mammalian cytokine gene promoter complex, the NF-κB/HMG I/PDII human interferon-β gene (IFN-β) promoter complex, and the ER/ERE vitellogenin gene promoter complex (Thanos and Maniatis, 1992; Zhong et al, 1997; Diebold et al, 1998; Wood et al, 1998; Zhang and Verdine, 1999). All of these cooperative systems have been shown to share several characteristics and they are: (1) protein and DNA conformational changes resulting from allosteric protein-protein and protein-DNA interactions; (2) a flexible protein-protein interaction interface; and (3) an extended protein-DNA interaction interface.

In general, investigations of the molecular mechanisms underlying cooperative binding of protein to DNA have focused on the role of protein-protein interactions. Target DNA nucleotide sequence is often described as a protein-docking site. Any DNA structural distortion that results from the DNA-protein interaction has usually been explained as a consequence of protein-protein interactions rather than concluding that DNA sequence is a contributing mechanism to the resulting protein/DNA structure. For example, the DNA bending observed in both the yeast MATα2/MCM1/STE6 and mammalian NFAT1/Fos-Jun/ARRE2 cooperative complexes have been considered the result of cooperative protein binding rather than as an underlying contributing molecular mechanism (Diebold et al, 1998; Kerppola et al, 1998). It is thought that upon binding DNA, NFAT1 undergoes a conformation change that compels Fos-Jun to bind DNA through protein-protein interactions. Due to this interaction with NFAT1, the Fos-Jun heterodimer apparently adopts a new conformation that in turn forces the DNA to bend to accommodate specific Fos-Jun/DNA contacts.
This protein-centric perspective arises from the static crystallographic method that is often employed to examine protein-DNA complexes and the dogma that proteins are the functional molecules whereas DNA is merely the genetic code. A crystallographic X-ray structure invariably fails to convey the dynamic and temporal contribution of each molecule to the formation of the overall complex. Additionally, functional studies typically examine the effect of protein mutations on complex formation and often ignore the contribution of DNA sequence. If we consider the protein-DNA interactions that occur within cooperative complexes to be allosteric, then the DNA would conceivably also undergo a conformation change upon protein binding and this DNA conformation change likely has biological significance (Senear et al, 1998). It has also been suggested that cooperative binding of transcription factors in vivo can occur by several mechanisms, some of which do not require direct protein-protein interactions (Vashee et al, 1999).

A few researchers have examined the contribution of DNA nucleotide sequence to the formation of cooperative complexes and subsequent synergistic transcriptional activation. The *in vitro* selection of non-consensus DNA binding site sequences in the presence of known cooperatively interacting proteins has shown that variation in nucleotide sequence affects the formation of cooperative complexes. An *in vitro* binding site selection assay constrained the transcription factor MATα2 to select DNA half sites in the presence of the cooperatively interacting MCM1 protein (Zhong et al, 1997). This assay revealed that the selected Mata2 half sites varied in nucleotide sequence from the consensus high affinity MATα2 half site. Intriguingly, the half sites selected in the presence of MCM1 resembled the asymmetric natural MATα2 half sites found within the cooperative MATα2-MCM1 binding sites of yeast a-site genes, such as STE6. The MATα2 half site nucleotides selected in this assay did not necessarily make specific contacts with MATα2 although their mutation significantly decreased both the *in vitro* cooperative formation of the MATα2/MCM1/STE6 complex and its *in vivo* synergistic transcriptional effects. These results suggest that nucleotides that do not specifically contact MATα2 and are not essential to MATα2 binding nevertheless contribute to the formation of a cooperative MATα2/MCM1 protein-DNA complex and affect synergistic transactivation.
The NF-κB transcription factor is an example of how variation in DNA binding site sequence can structurally alter the bound protein and that this protein conformation change can affect function. NF-κB binds DNA as a dimer that can consist of any combination of five different subunit proteins that are members of the Rel family (p50, p65, p52, c-Rel, and RelB) (Lenardo and Baltimore, 1989; Gilmore et al, 1992). Binding studies of NF-κB binding to DNA revealed that the NF-κB/DNA complex is reliant on a specific ion concentration that is dependent on the combination of NF-κB dimer subunits for a given DNA binding element (Menetski et al, 2000). The apparent dissociation constant of the protein complex is a function of ion concentration and pH. Measuring the ion release after the protein-DNA complex has come to equilibrium provides information on the structural differences for the same NF-κB dimer bound to DNA sites that vary in nucleotide sequence (deHaseth et al, 1977, Menetski et al, 2000). When the formation of a p50 homodimer NF-κB/DNA complex was compared using two different DNA binding site sequences, the number of ions released from the protein-DNA complex was significantly different depending on the DNA binding sequence (Metetski et al, 2000). This change in ionic strength upon complex formation implied that there was a change in structure of the bound p50 homodimer protein that depended on the sequence of the DNA binding site (Menetski et al, 2000). No changes were observed in DNA structure upon NF-κB binding, although the technique used (the net change in end-to-end distance of an oligonucleotide upon NF-κB binding) may not have been sensitive enough to detect local structural changes in DNA. Protein structural changes that were dependent on the nucleotide sequence of the DNA binding site were also observed within the p65 homodimer NF-κB – DNA X-ray crystal structure (Chen et al, 1998).

The variable effect of DNA sequence on NF-κB protein structure could explain why some NF-κB DNA-binding sites but not others promoted cooperative interactions with other transcription factors (Thanos and Maniatis, 1992). The human IFN-β and immunoglobulin-κB (Ig-κB) are two genes that are regulated by NF-κB. However, cooperative interactions between NF-κB and the architectural protein HMG I are observed within the IFN-β but not the Ig-κB regulatory region. HMG I belongs to the high mobility group (HMG) protein superfamily whose members recognize and bind to either specific DNA sequence or DNA-
tertiary structure (Laudet et al, 1993). HMG I is a sequence-specific architectural protein that specifically binds AT-rich DNA but does not have any intrinsic transactivation function. Instead these architectural proteins bind the minor groove of AT-rich DNA and intercalate a hydrophobic residue into the DNA structure, which causes the DNA to bend towards the minor groove (Churchill and Travers, 1991; Werner et al, 1995; Love et al, 1995). This architectural rearrangement of the DNA functions to open the major groove facilitating the access of a transcription factor, such as NF-κB (Yie et al, 1997). HMG I minor groove binding to the IFN-β promoter greatly facilitates the binding of NF-κB to the major groove (Thanos and Maniatis, 1992). This cooperative interaction was dependent on the AT-rich flanking and core binding sequence of the HMG I DNA-binding site within the IFN-β promoter (Zhang and Verdine, 1999). This AT-rich sequence is absent within the NF-κB core binding site and flanking regions of the Ig-κB gene, which does not display cooperativity between NF-κB and HMG I.

HMG I and NF-κB have been shown to interact using a GST pull-down assay. The mutation of several HMG I alanine residues, which are not necessary for DNA binding, will not only disrupt the interaction between HMG I and NF-κB but also the cooperative DNA-binding between these two proteins (Du et al, 1993; John et al, 1995; Leger et al, 1995; Zhang and Verdine, 1999). Intriguingly, the separate NMR structures of the HMG I bound to its DNA element and the NF-κB bound to its DNA element contradict the above NFκB-HMG I interaction data. These two separate NMR structures reveal that the proposed NFκB-HMG I interactions would be highly improbable if the proteins were within proximity of each other and bound to a common DNA molecule (Huth et al, 1997; Zhang and Verdine, 1999). Therefore, the protein-protein interactions that appear to be necessary for cooperative binding likely arise from the DNA structural changes that result from both NF-κB and HMG I binding to DNA (Zhang and Verdine, 1999). The crystal structure of NF-κB and HMG I together bound to the same DNA molecule has yet to be published, yet the above observations suggest that NF-κB and HMG I could mutually cooperate by reciprocally altering DNA structure. This implies that the specific nucleotide sequence of the binding site will favor NFκB-HMG I cooperative interactions whereas a different nucleotide sequence would fail to adopt the correct structure.
Within the nuclear receptor field, DNA binding site sequence has been shown to dictate the conformation of a bound ER homodimer (Wood et al, 1998). Partial digestion of the bound ER homodimer using trypsin and chymotrypsin revealed distinct cleavage patterns that varied with the estrogen response element (ERE) nucleotide sequence. Additionally, using naturally occurring EREs, three different ER-specific antibodies interacted differentially with the ER depending on the nucleotide sequence of the ERE. These data show that the nucleotide sequence of the ERE dictated the conformation of the bound ER. While functional studies were not done on these variant bound ER conformations, the results suggest that nucleotide sequence could dictate function. Conceivably the variations in bound ER conformations could permit or restrict subsequent transcription factor and/or coregulator interactions that will ultimately contribute to transcriptional regulation.

1.8 Probasin Promoter and the PSA Enhancer Possess Multiple Cooperatively Interacting AREs

1.8.1 The Rat Probasin Promoter

Rat prostatic basic protein (probasin) was first isolated from to the dorsolateral lobes of the rat prostate (Matuo et al, 1982). This region of the rat prostate is analogous to the peripheral zone of the human prostate, which is the most common site for prostate cancer. Since its first isolation, a homologue of rat probasin was found in mouse dorsolateral prostate although a human homologue has yet to be found (Genbank AF005204). Probasin expression was further delineated to the lateral lobe where it has the highest expression then, in descending order of expression, the dorsal lobe, anterior lobe, ventral lobe, and seminal vesicles (Matuo et al, 1986, Matusik et al, 1986). Although probasin is primarily secreted in high levels, weak protein staining was found in the nuclei of luminal epithelial cells (Matuo et al, 1985). There is only one mRNA transcribed from the probasin gene but there are alternate translation initiation codons that explain the observed different protein localizations. One translation initiation codon permits the translation of a signal peptide permitting membrane localization whereas the second initiation codon does not. After post-translation modification, however, the mature secreted and nuclear probasin proteins are identical.
suggesting a bifunctional role for the protein that is dependent on location rather than amino acid sequence (Spence et al, 1989).

A minimal probasin promoter region from −426 to +28 base pairs was enough to direct expression specifically to prostate cancer cell lines such as LNCaP, PC3 and DU145, whereas reporter gene expression was very low in non-prostate cell lines such as HepG2, BL13 (bladder cancer) CHO, MCF-7 (breast), 293 and MRC5 (Brookes DE 1998). Expression driven by the −426 to +28 base pair probasin promoter fragment was shown to be androgen-specific in the PC-3 human prostate cancer cell line in that cotransfected AR with androgen, but not GR with dexamethasone, resulted in chloramphenicol acetyl transferase (CAT) gene expression (Rennie et al, 1993; Kasper et al, 1994). However, in LNCaP cells reporter gene expression from this probasin promoter fragment was inducible using GR with dexamethasone, although levels of transcription were much lower than those achieved with equivalent concentrations of AR and R1881, the synthetic androgen methyltrienolone (Rennie et al, 1993).

The proximal probasin promoter and the large probasin (LPB) promoter (~12 kilobase pairs) were also shown to be prostate-specific in vivo. Using transgenic mice, the −426 to +28 base pair probasin promoter region or LPB promoter successfully targeted several reporter gene products, (such as the bacterial reporter gene CAT or the simian virus 40 (SV40) large T-antigen), to mouse prostate epithelial cells (Greenberg et al, 1994; Green et al, 1998).

Probasin gene expression is androgen regulated in that castration results in a dramatic drop in probasin mRNA levels (Dodd et al, 1983). Transgene expression driven by the LPB promoter increased with increasing serum androgen levels and stabilized when the mice reached sexual maturity. Like probasin mRNA expression levels, transgene expression levels driven by the LPB promoter dropped off after castration and levels were rescued after treatment with androgens (Greenberg et al, 1994).

DNase I protection assays revealed two androgen response elements (AREs) within the probasin promoter between positions −236 to −233 and −140 to −117 base pairs (ARE1 and ARE2 respectively) (Figure 6) (Rennie et al, 1993; Kasper et al, 1994; Kasper et al, 1999). These AREs varied considerably from the canonical SRE nucleotide sequence and Scatchard analyses revealed that the AREs individually have weak AR binding affinity.
compared to the high affinity SRE, although ARE2 had markedly higher binding affinity for the AR relative to ARE1 ($K_d$ values of 6.7 nM versus 20 nM respectively) (Kasper et al, 1994).

Together, the probasin ARE1 and ARE2 were shown to interact cooperatively such that the AR binding affinity was increased dramatically and occupation of these two sites occurred at a much lower concentration of AR relative to either ARE1 or ARE2 individually (Kasper et al, 1994; Kasper et al, 1999). ARE2 was shown in particular to be specifically and uniquely recognized by the AR although GR was capable of binding both ARE1 and ARE2 but with extremely low binding affinity (Kasper et al, 1999; Schoenmakers et al, 1999). When the nucleotide sequence of the steroid response element within the C(3) gene was changed to that of the probasin ARE2, then C(3) expression narrowed from dual AR/GR control to AR-specific control. This phenomenon suggested that nucleotides within ARE2 conferred AR-specificity by preventing GR from binding (Claessens et al, 1996). Another interpretation of this data is that the nucleotide sequence of ARE2 within the C(3) promoter context dictated specific function to the bound AR, but not to a bound GR, which permitted subsequent coregulator and/or other transcription factor interactions that mediated the observed AR-specific transcriptional response.

Disruption of either ARE1 or ARE2 destroyed cooperativity, substantially reduced AR binding to both elements, and also reduced androgen-dependent expression by more than 95% from a transiently transfected probasin promoter (~426 to +28 base pairs) driven CAT reporter gene (Kasper et al, 1994; Kasper et al, 1999). Finally, removal of the intravening wildtype sequence also disrupted androgen-dependent expression in transfection studies suggesting that either the spacing between the AREs was critical or that another binding site existed that was necessary for an androgen-directed synergistic transcriptional response (Kasper et al, 1994).

1.8.2 **The Human PSA Enhancer**

Human prostate specific antigen (PSA) is expressed at high levels in the prostate luminal epithelial cells and is either absent or expressed at very low levels in other tissues (Aumuller et al, 1990). Clustered with two other members of the glandular kallikrein gene

Within the proximal PSA gene promoter, ARE I was discovered at position -170 base pairs (Reigman et al, 1991) and a 35 base pair androgen responsive region was identified starting at position -400 base pairs, which contains ARE II at position -392 base pairs (Cleutjens et al, 1996). This proximal PSA promoter region conferred androgen-responsiveness, however, it was not AR-specific. GR with dexamethasone was able to stimulate transcription from the PSA promoter fragment to levels similar to those achieved with AR and R1881 in transient transfection assays although AR and GR induced expression was reasonably specific to the prostate cancer LNCaP cell line (Cleutjens et al, 1996). In contrast to the proximal probasin promoter, the proximal PSA promoter starting at -630 base pairs was inadequate to confer androgen-responsiveness or target gene expression to the prostate of transgenic mice (Cleutjens et al, 1997 pp 1256).

An enhancer element located approximately 4.2 kilobase pairs upstream of the PSA gene transcription start site, in concert with the proximal PSA promoter, was found to be necessary for both androgen-responsiveness and prostate-specific expression in transgenic mice (Cleutjens et al, 1997 pp 148; Schuur et al, 1996; Pang et al, 1997; Cleutjens et al, 1997 pp 1256). The core AR-regulated enhancer region spans 455 base pairs and contains at least four AREs (III, IIIA, IV and V) at positions -4079, -4143, -4179, -4225 base pairs (Figure 7) (Cleutjens et al, 1997 pp 148; Huang et al, 1999). Each ARE was found to intrinsically possess weak affinity for the AR as determined by DNase I quantitative footprinting, however, collectively these AREs cooperated to enhance AR-binding affinity and sensitivity to AR concentration and to contribute to synergistic transactivation (Huang W 1999). When any single ARE out of the four was mutated to destroy AR binding, transcriptional activity decreased by 50-75% relative to the wildtype enhancer. The PSA enhancer region, when coupled with the proximal PSA promoter, was determined to be AR-specific in that AR was able to induce transcriptional activity in both prostate cancer cell lines and non-prostate cell lines whereas the PR failed to induce transcription from this construct (the GR was not tested) (Huang et al, 1999).
1.9 Objectives of this Study

The cooperativity studies described in section 1.7 illustrate that nucleotide variation within transcription factor binding sites can dictate conformation of the bound protein and that the nucleotides involved need not make base-specific contacts to do so. The objective of this project was to characterize the allosteric protein-DNA interactions that underlie the cooperativity observed between AREs and to determine the critical architectural ARE context that results in synergistic AR-specific transcriptional regulation.

The probasin promoter and the PSA enhancer are specifically regulated by the AR and specifically expressed in the prostate. Both gene regulatory regions possess multiple AREs that cooperatively interact to ensure an AR-specific response. These characteristics recommend both the probasin promoter and PSA enhancer as excellent systems for the study of cooperative AR-ARE interactions.

The allosteric protein-DNA interaction that results in a protein conformation change will also affect a conformation change in the DNA. Studying protein-DNA interactions usually involves resolving the complex either through X-ray crystallography or through employing solution-based chemical or enzymatic footprinting techniques. Crystallographic studies are time consuming and labour intensive but nonetheless provide critical information concerning the weak interactions between protein and DNA. The limitations of crystallography however, concern the resulting static nature of the protein-DNA complex and the confined conditions that are necessary to generate a resolved structure. Chemical or enzymatic probing of protein-DNA contacts, in contrast, permit the resolution of weak interactions in a realistic solution-based and dynamic system. Enzyme probes, such as DNase I, can be employed to probe global structural distortion and show protected and hypersensitive regions of DNA that result from protein binding. Chemical probes, such as dimethylsulfate (DMS), can be employed to probe specific nucleotide contacts that are involved in protein binding in both the major and minor grooves of DNA. Additionally, employing DMS to probe protein-DNA contacts also provides information on DNA structural distortion that results from protein-binding (Chan et al, 1990, Clark et al, 1990, Frappier et al, 1992; Espinas et al, 1995; Ramesh and Nagaraja, 1996; Reid and Nelson, 2001).
To examine cooperative protein-DNA interactions within the rat probasin promoter and human PSA enhancer, methylation protection (MeP) and interference (Mel) footprinting assays were employed. In the process of footprinting the cooperative probasin promoter from -426 to +28 base pairs, it was discovered that neither Mel nor the conventional MeP assay was successful in generating a footprint. An alternative and novel MeP assay was devised, which permits the visualization of protein-DNA contacts within highly cooperative systems (Reid and Nelson, 2001). Using this modified MeP protocol, two novel AREs were discovered within the proximal probasin promoter. The novel AREs, called G-1 and G-2, displayed an unusual protection and DMS hypersensitivity pattern compared to conventional AREs that suggested a structural change in DNA upon AR binding. Binding assays were employed to determine the dissociation constants of individual AREs and the relative binding affinity for multiple AREs. Transient transfection assays, using an ARE-driven luciferase reporter construct in LNCaP cells, were employed to determine whether the in vitro results had any significance in vivo.

Additional work concerning AR-ARE interactions was completed that contributes to the evolving understanding of how AR specifically regulates its target genes. Upon closer examination of the nucleotide sequence of both G-1 and G-2, there exists a potential overlapping binding sequence for the architectural protein Sox-4, which belongs to the sequence-specific HMG-box architectural protein superfamily described earlier. Applying the MeP assay revealed that the addition of Sox-4 to the AR binding reaction increased the number of AR occupied sites within the probasin promoter. Lastly, the AR-DNA contacts within other AR-regulated gene regions such as the mouse epididymous retinoic-acid binding protein (mE-RABP) promoter, the enhancer region within the rat AR gene, the upstream rat probasin promoter, and the human p21 proximal promoter were examined using the modified MeP assay.
2. Experimental Procedures

2.1 Methylation Protection and Methylation Interference of the AR DBD

_In vitro_ methylation-based footprinting (MeP) is used to identify the protein-DNA contacts at the nucleotides guanine or adenine, made by a DNA-binding protein (Figure 7) (Brunelle and Schleif, 1987). MeP can also be used to detect local structural DNA distortion caused by allosteric interactions with the transcription factor (Chan et al, 1990; Clark et al, 1990; Frappier et al, 1992; Espinas et al, 1995; Ramesh and Nagaraja, 1996; Reid and Nelson, 2001). Dimethylsulfate (DMS) is typically employed since this chemical will uniformly methylate guanines in the N7 position in the major groove and, less efficiently, adenines in the N3 position in the minor groove, unless the nucleotides are successfully protected from methylation by protein contact. Protection from DMS methylation does not directly measure weak interactions between amino acids and nucleotides. However, since the DMS molecule is very small, if its methylation activity is blocked then a weak interaction such as a hydrogen bond is inferred. Using MeP, guanines and adenines are protected from DMS attack after the protein-DNA complex is formed. Local structural distortion that results from protein binding will manifest as hypersensitivity to DMS methylation (Chan et al, 1990; Clark et al, 1990; Frappier et al, 1992; Espinas et al, 1995; Ramesh and Nagaraja, 1996; Reid and Nelson, 2001).

_In vitro_ methylation-interference (Mel) is used to identify those guanine or adenine contacts that are necessary for protein binding (Figure 8). In Mel the DNA is pre-methylated with DMS using one-hit kinetics, purified, then bound to the protein of interest. The free and bound DNA are segregated using an electrophoretic mobility shift assay (EMSA). If a particular pre-methylated guanine or adenine is necessary for the DNA-protein interaction then that DNA molecule will be segregated to the free fraction of DNA and will be correspondingly rejected from the bound fraction of DNA.

Unlike MeP, Mel does not provide information about the DNA structural distortion that results from allosteric interactions with transcription factors since the DNA was methylated before its introduction to protein. In addition, Mel cannot always be used for cooperative DNA binding interactions because a solitary disrupted DNA-protein interaction can be
Figure 7. Schematic of the methylation protection assay. The MeP assay involves the methylation of protein-bound DNA followed by chemical cleavage of methylated guanines or adenines. Cleaved DNA is visualized using denaturing PAGE.
Figure 8. Schematic of the methylation interference assay. The Mel assay requires the DNA to be pre-methylated using DMS then the purified probe is used in a protein-binding assay. Bound and free DNA populations are separated by PAGE followed by chemical cleavage. Cleaved DNA is visualized using denaturing PAGE where those nucleotides important for protein-DNA binding are enhanced in the free population.
compensated for by protein-protein and other neighbouring protein-DNA interactions. Therefore, the methylation of a single necessary guanine or adenine is not enough to dislodge the protein from the DNA. For multiple DNA-binding elements that exhibit cooperative binding, MeI will not provide DNA-protein contact information whereas MeP is very effective at investigating cooperative DNA binding complexes.

2.1.1 Preparing Recombinant DNA Binding Protein

The rat AR DBD (amino acids 524 through 648) and the full length human Sox-4 protein (52 kDa) were cloned into the EcoRI and BamHI sites of the plasmid vector pTrcHisC (Invitrogen) and transfected into Escherichia Coli (strain JM109) cells. To express the recombinant six-histidine N-terminal fusion protein, 5 mLs of L-Broth with 100 µg/µL of ampicillin were inoculated with transformed JM109 cells containing the recombinant pTrcHisC vector. After incubating overnight while shaking at 250 RPM at 37°C, 200 µL (50:1) of the overnight culture were sub-cultured into 100 mLs of L-Broth with 100 µg/µL of ampicillin and incubated at 37°C shaking at 250 RPM. When log growth was reached (approximately 3.0 hours), as measured by an optical density (O.D.600) measurement of 0.6, expression of the recombinant protein was induced by adding 0.1 M of IPTG (isopropylthio-β-D-galactoside). After 3.0 hours induction, the JM109 cells were pelleted and resuspended in 3.0 mL/gram Lysis Buffer (50 mM NaH2PO4 (pH 8.0); 300 mM NaOH; 10 mM imidazole) with 1.0 mg/mL Lysozyme (Gibco). Cells were lysed by sonication, pelleted, and then the cleared lysate was resuspended in 1.0 mL of Ni-NTA slurry (Qiagen) for every 4.0 mLs of cleared lysate.

To purify the recombinant HisTag protein from the other proteins in the cell extract, a nickel column is used to specifically bind the histidine residues located on the N-terminus of the recombinant protein of interest. To elute the recombinant histidine tagged protein off the column, imidazole is used to compete for the nickel-binding sites. Briefly, the cleared lysate and Ni-NTA slurry was mixed gently for 1.0 hour at +4°C then transferred to an empty column. The flow-through was permitted to exit by gravity flow, then the column was washed once with 50 mM NaH2PO4 (pH 8.0), 300 mM NaCl, and 20 mM imidazole to rid the column of non-specifically bound proteins. The column was washed once with 50 mM
NaH₂PO₄ (pH 8.0), 100 mM KCl, and 20 mM imidazole and then once again with 20 mM HEPES (pH 7.9), 100 mM KCl, and 20 mM imidazole. The recombinant protein was finally eluted off the nickel-NTA column using 20 mM HEPES (pH 7.9), 100 mM KCl, 25% glycerol and 250 mM imidazole. Cell extract, washes and recombinant protein fractions were visualized and confirmed by SDS-PAGE followed either by staining with Coomassie Brilliant Blue (BRL) or Western blotting using a polyclonal antibody (Invitrogen) directed against the six histidines located at the N-terminus of the recombinant protein. Protein concentrations were determined using the Bradford Assay and BCA protein assay and the protein fractions were stored at -85°C (BioRad).

2.1.2 Generation of Radiatively Labelled DNA Probes

To generate single-end labelled probes of the rat probasin promoter, the human PSA enhancer, the rat AR enhancer, the human p21 promoter, or the mE-RABP promoter one of three general methods was employed. The different protocols were as follows: (1) Single-end labelled fragments were generated using polymerase chain reaction (PCR) amplification with one of the two oligonucleotide primers end-labelled with ³²P (see section 2.1.2a). (2) Single-end labelled probes of the probasin promoter, p21 promoter, PSA enhancer, or mE-RABP promoter were generated by sequentially cutting the fragment of interest out of a recombinant plasmid vector (see section 2.1.2b). The plasmid was initially linearized using a restriction endonuclease then end-labelled with ³²P using either polynucleotide kinase (PNK) (NEB) or Klenow DNA polymerase (NEB). Lastly, the fragment insert was cut out using another restriction endonuclease; or (3) Finally, single end-labelled DNA probes of either the wildtype or mutated G-1 element derived from the rat probasin promoter were generated by annealing synthesized complementary oligos. Wildtype or mutated G-1 oligos were synthesized at the Nucleic Acids and Protein Synthesis laboratory (NAPS) at the University of British Columbia. Oligos were end-labelled with ³²P and annealed to a cold complementary oligo to generate a single-end labelled double-stranded DNA probe. All three end-labelling protocols employed polyacrylamide gel electrophoresis (PAGE) to purify the probes.
2.1.2.a Generating Single-end Labelled DNA Probes Using PCR Amplification

Eight different probasin promoter probes and one rat AR enhancer probe were generated using PCR amplification. The -426 to +28 base pair probasin promoter insert in the pBluescript plasmid vector and the entire rat AR gene (~3.4 kbp) in the pBluescript plasmid vector were used as DNA templates (Stratagene). The probasin promoter fragments generated by PCR were also inserted into the pTKLuc expression plasmid and used to generate the cloned versions of the fragments and for transfection studies (refer to section 2.3). The PCR primers were synthesized at NAPS UBC. The primers and resulting fragments are as follows (note: to facilitate cloning, the probasin fragments that were subsequently cloned into the pTKLuc expression vector were PCR generated using the following primers synthesized with BamHI/HindIII linkers): (1) Probasin ARE1-G1-ARE2-G2: -269 through -77 base pairs using forward and reverse primers 5' CAT CTA CCA TTC CAG TTA AG 3' and 5' CAG TTG TAC ACA GAC ATT GA 3' (2) Probasin G1-ARE2-G2: -229 through -77 base pairs using forward and reverse primers 5' TCT TAG TCT TTT TCT TAA TAG G 3' and 5' CAG TTG TAC ACA GAC ATT GA 3' (3) Probasin ARE1-G1: -269 through -164 base pairs using forward and reverse primers 5' CAT CTA CCA TTC CAG TTA AG 3' and 5' CTG TCC CAT TCT TCA GGC 3' (4) Probasin ARE2-G2: -150 through -77 base pairs using forward and reverse primers 5' GTC CAT GCC TAG TAA AG 3' and 5' CAG TTG TAC ACA GAC ATT GA 3' (5) Probasin ARE1: -269 through -210 base pairs using forward and reverse primers 5' CAT CTA CCA TTC CAG TTA AG 3' and 5' CTG TCC CAT TCT TCA GGC 3' (6) Probasin ARE2: -150 through -105 base pairs using forward and reverse primers 5' GTC CAT GCC TAG TAA AG 3' and 5' CAT CTA GTA TAC AAA TA 3' (7) Probasin G1: -229 through -164 base pairs using forward and reverse primers 5' GTC CAT GCC TAG TAA AG 3' and 5' CAT CTA GTA TAC AAA TA 3' (8) Probasin G-2: -121 through -77 base pairs using forward and reverse primers 5' TAT TTA GAA AAA GAC TA 3' and 5' CAG TTG TAC ACA GAC ATT GA 3'; and (9) rat AR Enhancer 2028 to 2333 base pairs using forward and reverse primers 5' CAT TGA AGG CTA TGA ATG TC 3' and 5' CGG TAC TCA TTG AAA ACC AG 3'.

Before generating single-end labelled PCR generated fragments, either the forward or the reverse primer had to be end-labelled using PNK (New England Biolabs (NEB)). Briefly,
10 μM of primer was incubated with 50 μCi of γ³²P-ATP (Amersham) with 1X PNK buffer (NEB) and 1.5 Units of PNK (NEB) for 1 hour at 37°C. The end-labelling reaction was then incubated at 75°C for 20 minutes to heat denature the kinase. To rid the reaction of unincorporated radionucleotides, the sample was centrifuged (2800Xg for 2 minutes) through a G25 tris-sepharose column (Amersham).

To PCR amplify each probe, 100 ng of the DNA template was added to a 25 μL volume containing 200 μM of deoxynucleotide triphosphates (Gibco), 1.5 mM MgCl2, 1X PCR Buffer (Gibco), 10 pM each of the radioactively labelled primer and the cold reverse primer, and 2.5 Units of Taq DNA polymerase (Gibco). Using the MJ Research DNA Engine Thermocycler, each sample was initially denatured at 95°C for 2.0 minutes then for 25 cycles the following steps: 95°C for 1.0 minute, 50°C for 30 seconds, and 72°C for 30 seconds. After all the cycles were completed the reaction was incubated at 72°C for 1.0 minute to ensure all fragments were fully extended. After adding loading buffer containing bromophenol blue and xylene cyanol, the entire reaction volume was loaded onto a 0.75 mm 5% PAGE (29:1 acrylamide:bis (BioRad)) with 1X TBE (90 mM Tris Base, 90 mM Boric Acid, and 2 nM Disodium EDTA). The PAGE was run at 16 V/cm at room temperature until the expected fragment size had run well into the gel as measured by the migration of the marker dyes - bromophenol blue and xylene cyanol. The gel was covered in plastic film and exposed to Kodak BioMax Sensitive (BMS) film for approximately two minutes at room temperature. Using the developed film as a template, the fragments were cut out of the gel and the gel slices were placed into a fresh 1.7 mL silanized tube and crushed using a blue P1000 pipette tip. To elute the labelled DNA from the gel slice, 500 μL of Maxam and Gilbert Elution Buffer (0.6 M ammonium acetate, 1 mM EDTA and 0.1% SDS) was added to the crushed gel slice and the tube was rotated overnight at room (Maxam AM and Gilbert W 1977).

The following day the acrylamide gel slice was briefly pelleted in a tabletop centrifuge and the eluant was transferred to a fresh 1.7 mL silanized tube using a blue P1000 pipette tip with its end clipped off. The eluant was then centrifuged through a BioRad miniprep spin column to filter out any remaining acrylamide pieces. To ethanol-precipitate the end-labelled DNA, 1.0 mL of ice-cold 95% ethanol was added to the eluant. The tube was inverted to mix and placed in a -20°C freezer for ten minutes to help precipitate the
DNA. The tubes were then centrifuged at maximum RPM (>10,000 Xg) on a tabletop centrifuge for thirty minutes at 4°C. The ethanol was carefully removed and the DNA pellet was washed twice with 70% ethanol with five-minute centrifuges between washes. After careful removal of all traces of ethanol, the DNA pellet was air-dried briefly then resuspended in DNA Binding Buffer D (20 mM HEPES (pH 7.9), 100 mM KCl, and 10% glycerol). The radioactive disintegrations per minute (dpm) for 1.0 μL of end-labelled DNA fragment were determined using a Beckman scintillation counter.

2.1.2.b Generating Single-end Labelled DNA Probes Through Successive Enzymatic Digestion of Recombinant Plasmid Vectors

To generate end-labelled rat probasin promoter probes (derived from the -426 to +28 base pair fragment insert), the pBluescript or pTKLuc plasmid (plasmid courtesy of Dr. S. Nordeen, University of Colorado) constructs had to be initially linearized using either HindIII or BamHI (NEB). To generate the -705 to -426 base pair end-labelled rat probasin promoter probe from the -705 to +28 base pair fragment insert, the pGEMteasy plasmid construct (courtesy of Dr. R. Matusik, Vanderbilt University, Te) had to be linearized using either NheI or HindIII (NEB). To generate the downstream end-labelled mE-RABP promoter from the -198 to +28 base pair fragment insert, the pBluescript plasmid construct (construct courtesy Dr. J. Lareyre, Vanderbilt University, Te) had to be linearized using either HindIII or XbaI (NEB). To generate the upstream end-labelled mE-RABP promoter from the -543 to -166 base pair fragment insert, the pBluescript plasmid construct (Dr. Lareyre) had to be linearized using either HindIII or XbaI (NEB).

If end-labelling using PNK, the 5' phosphorous had to be removed and this was done post-linearization by adding 1.5 Units of calf intestinal phosphotase (CIP) (NEB) to the restriction digest reaction. The dephosphorylation reaction was incubated for one hour at 37°C then the CIP and restriction enzymes were heat denatured for 20 minutes at 75°C. To help rid the sample of enzyme protein, a phenol/chloroform extraction was employed. The digest was brought to a volume of 100 μL using sterilized distilled deionised water (sddH2O) and then 100 μL of phenol/chloroform/isoamyl alcohol (29:1:1) was added. The organic and aqueous solutions were vigorously vortexed then centrifuged 10,000Xg for five minutes.
The top aqueous layer was transferred to a fresh 1.7 mL silanized tube and 10 μL of 3.0 M sodium acetate (pH 5.2) was added (300 mM final sodium acetate concentration). The linearized dephosphorylated DNA was ethanol-precipitated as described above. After the DNA pellet was air-dried, it was resuspended in 20 μL of sddH2O.

To end-label using PNK, the linearized and dephosphorylated recombinant plasmid vector containing the promoter or enhancer fragment of interest was incubated with 50 μCi of γ\(^{32}\)P-ATP (Amersham), 1X PNK buffer (NEB), and 1.5 Units of PNK (NEB) at 37°C for 30 minutes. The PNK enzyme was then heat denatured at 75°C for 20 minutes to destroy PNK activity. To rid the reaction of unincorporated radionucleotides, the labelling reaction was passed through an S200 spin column (Amersham). The labelled DNA was ethanol-precipitated as described above then resuspended in 25 μL of sddH2O.

If labelling using Klenow DNA polymerase, no dephosphorylation step was required. The linearized recombinant vector (~ 4 pmoles of 5' ends) was incubated with 50 μCi (16 pmoles) each of α\(^{32}\)P-dATP, α\(^{32}\)P-dTTP, α\(^{32}\)P-dGTP, and α\(^{32}\)P-dCTP (Amersham), 1X EcoPol Buffer (NEB), and 2.0 Units of Klenow (NEB) at 25°C for 15 minutes. To ensure that the 5’ overhangs were completely filled-in with nucleotides, a cold-chase of 200 μM of dNTPs was added and incubated at 25°C for 5 minutes. The Klenow enzyme was heat denatured at 75°C for 20 minutes. To rid the reaction of unincorporated radionucleotides, the labelling reaction was passed through an S200 spin column (Amersham). The labelled DNA was ethanol-precipitated as described above then resuspended in 34 μL of sddH2O.

To generate a single-end labelled fragment, the −426 to +28 probasin probe was excised from the plasmid vector using a second restriction enzyme – either BamHI or HindIII (NEB) or in the case of the −705 to −426 base pair probasin fragment, either NheI or HindIII. To generate a single-end labelled fragment, the mE-RABP probes were excised from the plasmid vector using a second restriction enzyme – either HindIII or XbaI. After a two-hour digest at 37°C, 10 μL of 4X loading buffer containing 5% bromophenol blue and xylene cyanol was added to the reaction. The single-end labelled DNA fragments were purified by PAGE and eluted as described above. After resuspension in Buffer D, the dpm for 1.0 μL of resuspension was counted using a Beckman scintillation counter.
2.1.2. Generating Single-end labelled DNA Probes by Annealing Synthesized Oligos

Six 29mer oligos were synthesized at NAPS UBC along with their complementary strand. The sequence of these oligos corresponds to either the wildtype (WT) G-l element from the probasin promoter region (-209 to -195 base pairs) or mutated derivations of this element. The synthesized oligos, with their mutated nucleotides underlined, are as follows:

(1) G-l WT: 5' CTT AAT AGG GAC ATA AAG CCC ACA AAT AA 3';
(2) G-l A195C: 5' CTT AAT AGG GAC ATA AAG CCC CCA AAT 3';
(3) G-l C197T: 5' CTT AAT AGG GAC ATA AAG CTC ACA AAT AA 3';
(4) G-l G207A: 5' CTT CCT AGG AAC ATA AAG CTC ACA AAT AA 3';
(5) G-l T215G: 5' CGT AAT AGG GAC ATA AAG CCC ACA AAT AA 3';
(6) G-l Spacer mutant: 5' CTT AAT AGG GAC ACGGAG CCC ACA AAT AA 3'.

Single-end labelled double-stranded oligos were generated first by incubating either the single-stranded top strand or complementary bottom strand with 50 μCi of γ^32P-ATP (Amersham), 1X PNK buffer (NEB), and 1.5 Units of PNK (NEB) in a reaction volume of 30 μL for one hour at 37°C. The PNK enzyme was heat denatured at 75°C for 20 minutes. Unincorporated radionucleotides were removed by centrifuging the labelling reaction through a G25 column (Amersham). To generate a single-end labelled double-stranded probe, the cold complementary oligo was annealed to the ^32P-labelled oligo. The annealing step involved a slow temperature ramp from 90°C to 4°C at 1°C/second. Once the labelled oligo was annealed to its unlabelled complement, the double-stranded probe was isolated and purified using PAGE as described above. The air-dried labelled oligo was resuspended in Buffer D and the dpm/μL counted as described above.

2.1.3 Methylation Protection Assay

As mentioned earlier, the MeP assay employs DMS to methylate the DNA using one-hit kinetics after the protein-DNA complex has come to equilibrium. Two different protocols were employed with contrasting degrees of success. Both methods add DMS to the protein-DNA complex in solution. The difference between the protocols is in the method employed to stop the methylation reaction. The first protocol follows the conventional method and
employs β-mercaptoethanol to quench DMS activity after the reaction has proceeded to the desired point. The second protocol employs PAGE to separate the negatively charged protein-bound DNA and unbound (protein-free) DNA from the neutral DMS molecule (Reid KJ and Nelson CC 2001).

For both protocols, the protein-DNA binding reaction and DMS methylation procedures were the same. Briefly, the recombinant HisTag protein (7.2 µg or 13.6 µM of AR-Histag, HMG-Histag, or Sox-4-Histag) was incubated at room temperature for 15 minutes with 2.0 µg Poly dl-dC (Amersham) in DNA Binding Buffer (DBB: Buffer D and 1 mM DTT freshly added). To each reaction, 350 000 dpm (26.5 fM) of 32P-single-end labelled DNA probe was added to the final volume of 30 µL and the binding reaction was brought to equilibrium at room temperature for 10 minutes. Once binding equilibrium was reached, 3.0 µL of a 2% DMS solution was added to the binding reaction to a final DMS concentration of 0.18% (19 mM) and incubated at room temperature for exactly 2 minutes. DNA treated in the same manner, but without HisTag protein, was used as a control.

For the first protocol the DMS methylation activity was quenched with β-mercaptoethanol after 2 minutes by adding 50 µL of DMS Stop Solution (1.5 mM sodium acetate (pH 7.0) and 1.0 M β-mercaptoethanol). After vortexing vigorously, the reaction was ethanol-precipitated as described above. After the last of the ethanol was carefully removed, the pellet was air-dried then resuspended in sddH2O. The volume of water used depended on the ensuing cleavage reagent.

For the second protocol loading the reaction onto a PAGE stopped the DMS methylation activity. After the 2 minutes incubation time each methylation reaction was loaded onto a 5% (29:1 acrylamide:bis) 0.5X TBE PAGE while the current was running at 16 V/cm. Since the DMS is an overall neutral molecule, the negatively charged DNA will move away from the DMS and into the gel. Additionally, PAGE permitted the separation of bound and unbound DNA populations. After the AR-bound DNA had moved well into the gel, the wet gel was exposed to Kodak BMS film for 1 hour at room temperature. Using the film as a template, the bound (protein) and unbound (no protein) DNA gel fractions were excised and the DNA eluted and ethanol precipitated as described above. Again, the volume of water used to resuspend the DNA depended on the ensuing cleavage reagent.
Piperidine is a chemical that will specifically cleave those guanines that are methylated in the N7 position. If piperidine was used then the air-dried pellet was resuspended in 90 μL of sddH2O. The cleavage reaction required 1.0 M piperidine, therefore 10 μL of stock piperidine was added to the resuspended methylated DNA probe to achieve this concentration. The methylated guanines were cleaved by incubating the piperidine reaction at 90°C for 30 minutes. Since piperidine is a volatile chemical, lyophilization is a good method to remove piperidine from the sample. The 100 μL reaction volume was snap-frozen in liquid nitrogen and lyophilised under vacuum for ~1.5 hours in a Savant speedvacuum employing an oil condensation pump. The dried pellet was washed twice with 100 μL of sddH2O, snap-frozen, and lyophilised for ~1.5 hours in between washes. The final dried pellet was resuspended in 10 μL of formamide buffer overnight at 4°C.

The following day, 1.0 μL of the piperidine-cleaved samples was counted using a Beckman scintillation counter. The samples were diluted to 2000 dpm/μL, heated to 90°C for five minutes, placed on ice, and then 2.0 μL of each sample (4000 dpm per lane) was loaded onto a pre-run (20 V/cm for 1.0 hour) 6% (29:1 acrylamide:bis), 1X TBE, 8.3M Urea PAGE. The denaturing PAGE was run at 20 V/cm at room temperature until the bromophenol blue and xylene cyanol reached the appropriate distance for the DNA fragment size of interest. For those fragments over 100 base pairs, a salt gradient gel was used to slow down the faster and lighter fragments, preventing them from running off the gel while allowing the heavier fragments to move farther into the gel, thereby improving band resolution. In contrast to the regular 1X TBE denaturing acrylamide gel, the salt gradient gel utilized 0.5X TBE and the bottom chamber buffer contained 100 mM sodium phosphate (pH 8.0). To establish the salt gradient and ensure that the salt-front would be ahead of the DNA fragments, the acrylamide gel was pre-run at 20 V/cm for exactly 40 minutes before loading the samples. For both types of PAGE, the denaturing gel was dried under vacuum at 80°C then exposed at -80°C to Kodak BioMax Sensitive (BMS) or BioMax Resolution (BMR) film in a High Energy Transcreen enhancer cassette or a regular cassette respectively.

NaOH can also be employed to cleave modified nucleotides. This cleavage method is more flexible than piperidine since NaOH will also cut methylated adenines as well as guanines. Since adenines are methylated by DMS less efficiently than guanines, adenines
will necessarily appear fainter than guanines on an autoradiograph. To cleave using NaOH, the dried pellet of methylated DNA probe was resuspended in 100 μL of sddH2O. After 100 μL of 20 mM of sodium phosphate (pH 7.0) was added, the reaction was vortexed vigorously and incubated at 95°C for 15 minutes. The reactions were placed briefly on ice and 20 μL of 1.0 M fresh NaOH was added to a final concentration of 90 mM, vortexed vigorously, and incubated at 95°C for 1.0 hour. To neutralize the reaction, 20 μL of 1.0 N HCl was added and the reaction was vortexed vigorously. To help precipitate the DNA fragments, 10 μg of tRNA was added to the mixture. To reduce the salt concentration the volume was increased to 500 μL with sddH2O. The NaOH cleavage reaction was then ethanol-precipitated as described above. The air-dried DNA fragments were resuspended in 10 μL of formamide buffer, counted using a scintillation counter, and separated using a 6% denaturing PAGE as described above. Acrylamide gels were dried and exposed to autoradiograph film as described above. To compare relative band intensity, the exposed autoradiograph film was scanned using a Hewlitt Packard 1200 dpi resolution scanner and the bands quantified using ImageQuant software.

2.1.4 Methylation Interference Assay

As mentioned, the MeI assay requires the DNA to be pre-methylated and purified before use in the protein-DNA binding reaction. Briefly, the labelled DNA fragment of interest (200 fM) in DMS Buffer (50 mM sodium cacodylate (ph 8.0), 10 mM MgCl2, and 1.0 μg of calf thymus DNA) was incubated with 45 mM DMS for exactly two minutes. The reaction was stopped with 50 μL of DMS Stop Solution and the DNA was ethanol-precipitated as described above. The methylated DNA probe was then used in a protein-binding reaction with AR DBD-HisTag exactly as described in the MeP protocol. After separating the bound (protein-bound), free (rejected in the presence of protein), and unbound (no protein) DNA populations using PAGE, the bands were excised from the acrylamide gel and the DNA probes eluted as described above. The selected DNA populations were then ethanol-precipitated and cleaved at the methylated guanines and/or adenines either using piperidine or NaOH cleavage methods as described in the MeP protocol. Unbound, free, and
bound populations of DNA probes were visualized using denaturing PAGE and autoradiography as described in the MeP protocol.

2.2 Determining the Binding Affinities of the Probasin Promoter AREs

*Note to Reader: Stephen Hendy, Research Associate with the Prostate Centre at Vancouver General Hospital, kindly performed the following EMSA and binding assays to determine the relative and actual binding constants for the probasin promoter AREs.*

To determine the binding affinity of individual AREs located within the probasin – 426 to +28 base pair region, the dissociation constant \((K_d)\) for each binding site was determined using Scatchard Analysis. A constant amount of AR-Histag protein was incubated with increasing concentration of radiolabelled DNA. The probe concentration from the bound and free fraction, as determined by the radioactive signal, at each titration point was plotted according to the Scatchard plot;

\[
\frac{[P]{\chi}}{[\chi]} \text{ versus } [P]{\chi}
\]

where \([P]{\chi}\) is the protein-bound DNA concentration and \([\chi]\) is the free DNA concentration. The slope, \(-K\), provides the equilibrium constant.

2.2.1 Generating Radioactively Labelled Probasin Promoter Fragments With High Specific Activity

For quantitative analysis of DNA binding affinity it is necessary to radioactively label the DNA probe with high efficiency and specific activity. Using a labelled probe with high specific activity ensures that variation between experiments is minimized and duplication of results is possible because the effects of cold or unlabelled probe on binding specificity will be minimized. For this reason, the DNA probe used in these experiments was labelled by incorporating radioactive \(\alpha^{32}\)P-dCTP into the amplified strand through the PCR extension
step. This contrasts the labelling method employed for the MeP and MeI assays where pre-
labelling one of the primers prior to amplification was necessary to generate single-end
labelled DNA probes. Using an isotope-incorporation method, the amplification conditions
can be manipulated such that each generated DNA probe will have a high average radioactive
specific activity. Using the following PCR amplification conditions, DNA probes were
generated with approximately 2 incorporated radioactive cytosines for every 20 cytosines
present within the DNA with a specific activity of 20 μCi/mMol.

To generate radioactively labelled DNA probes for binding studies, the PCR
amplification reaction contained 100 ng of the template plasmid −286 to +28 probasin
promoter in pBluescript with 200 μM of dNTPs, 20 μCi of α³²P-dCTP (specific activity of
3000 Ci/mMol), IX PCR Buffer, 1.5 mM of MgCl₂, 1.0 μM of forward and reverse primers
(refer to section 2.1.1 for primer sequence and fragments generated), and 2.5 Units of Taq
DNA polymerase. Using the MJ Research DNA Engine Thermocycler, the amplification
program was 95°C for 1.0 minute then 25 cycles of 95°C for 30 seconds, 45°C for 30
seconds, and 72°C for 30 seconds. A final extension step at 72°C for 5 minutes ensured that
all the amplification products were fully extended. Labelled amplified DNA probes were
purified using PAGE and ethanol-precipitated as described above. The air-dried DNA pellet
was resuspended in Buffer D and the dpm/μL was determined using a Beckman scintillation
counter.

2.2.2 Determining Binding Affinity Through EMSA

EMSA was employed to determine the AR-HisTag binding affinity for the individual
probasin promoter AREs and various combinations of these AREs. Briefly, 10 pmol of AR-
HisTag was pre-incubated for 15 minutes at room temperature with 1.0 μg of Poly dI-dC
(Amersham) in DBB in a 10 μL volume. Increasing amounts of the radiolabelled probe in a
2.0 μL volume of DBB was added and the binding reactions were further incubated at room
temperature for 10 minutes. After the reaction was brought to equilibrium, the titrations
were loaded onto a pre-run (16 V/cm for 3 minutes) 5% (29:1 acrylamide:bis) 0.5X TBE
PAGE and run at 16 V/cm until the bound and free DNA populations are well separated. The
gels were transferred to Whatman chromatography paper and dried under vacuum at 80°C
then exposed to Kodak BMR film. Using the developed film as a template, the bands corresponding to the bound protein-DNA complex and free DNA were excised from the dried gel. To determine the activity of each band, the dpm of each dried gel slice was counted using a scintillation counter. For the DNA probes containing only one ARE, the binding constants were determined by Scatchard analysis. For the DNA probes containing more than one ARE, the relative binding affinity was compared by plotting bound versus free values averaged from three independent binding experiments for each probe.

2.3 Transfection Studies to Determine Effect of Probasin Promoter Fragments on Transcriptional Activation

Note to Reader: Jody L. Saito, UBC Department of Pathology PhD student with the Prostate Centre at Vancouver General Hospital, kindly performed the following transfection experiments to determine the contribution of specific probasin promoter fragments to transcriptional activation.

To determine whether the in vitro footprinting and binding affinity assay results had any relevance within an in vivo system, transient transfection assays were carried out using the prostate cancer cell lines LNCaP and PC-3. Briefly, luciferase reporter plasmids were created by introducing BamHI/HinDIII ends to each fragment of the probasin promoter by PCR (please refer to section 2.1.2a) followed by cloning the amplified product into the BamHI/HinDIII sites of the pTK-Luc expression plasmid.

LNCaP or PC-3 cells were grown to 60% confluence in 24-well plates in RPMI media (Gibco) supplemented with 5% fetal calf serum (FCS). Cells were then transfected with 0.2 μg of luciferase reporter plasmid, 1.2 μg of rat AR in the pRLTK expression plasmid (Invitrogen), and 8 ng of renilla expression plasmid pRLTK (Promega) using Lipofectin (Gibco). Cells were incubated for 22 hours in 5% charcoal-stripped medium with 0 to 5 nM of R1881 (Perkin Elmer Gibco). Cells were washed and harvested with passive lysis buffer (Promega) and the luciferase activity of 20 μL cell lysate aliquots were determined using the Dual-Luciferase Reporter assay system (Promega) on a luminometer (Berthold, Germany). Luciferase activity was normalized for transfection efficiency using
renilla activity. Experiments were done in triplicate, averaged, and expressed both in relative luciferase units (RLU) and as fold induction.
3. Results and Discussion

3.1 Methylation Interference Assay Failed to Footprint a Cooperative System

Initial attempts at footprinting the rat probasin promoter (-426 through +28 base pairs) using Mel were problematic because of the cooperativity that occurred between the AREs within the promoter fragment. DMS methylation was carried out at one-hit kinetics, which meant that the methylation rate was empirically established at an average of one methylated guanine or adenine per molecule of DNA. Success using Mel was predicated on the theory that if a major groove guanine or minor groove adenine was necessary for protein binding and was methylated, then this DNA molecule would be rejected from the bound DNA population and relegated to the free DNA population. In a cooperative system, however, methylation of a single nucleotide necessary for protein binding would not guarantee the relegation of the DNA molecule to the free population. This was because other protein-protein or protein-DNA contacts involved in neighbouring cooperative interactions could compensate for any individual nucleotide disturbed through methylation. Using Mel to determine those guanines or adenines necessary for AR DBD binding to the cooperative probasin promoter unfortunately resulted in no information at all because of the failure of this assay to separate free and bound DNA populations (data not shown).

3.2 A Modified Methylation Protection Assay Revealed Atypical AR Binding Sites

By switching to the MeP assay it was possible to overcome the cooperativity-based problem intrinsic to the Mel assay since MeP did not require the separation of bound and free DNA populations. Instead, MeP involved the DMS methylation of the protein-DNA complex after the complex had reached equilibrium, therefore, all guanine and adenine contacts that were involved were detectable and cooperative interactions did not interfere with the information obtained.

Initial attempts at footprinting the probasin promoter region, from -426 to +28 base pairs, used the conventional method of employing β-mercaptoethanol to quench the methylation activity of DMS (Brunelle A and Schleif RF 1987). β-mercaptoethanol is a
hydrophobic molecule that is able to quench DMS activity but it is also able to disrupt protein-protein interactions – in particular disulfide bridges. The footprint result after employing β-mercaptoethanol was uninformative as no protected nucleotide contacts were observed when the protein-bound reaction (bound) was compared to the protein-free reaction (unbound) (Figure 9). There were three possibilities for this null result. Firstly, it was possible that the AR DBD-DNA contacts were not specific and therefore failed to be protected from the methylation activity of the small and sensitive DMS molecule. Secondly, it was possible that the experimental conditions were such that only a small proportion of the DNA population was bound by protein; therefore, the overwhelming unbound DNA population masked any protein-DNA contacts. Finally, it was possible that either DMS or β-mercaptoethanol disrupted the protein-DNA complex permitting DMS to freely and uniformly methylate the released DNA. β-mercaptoethanol is not entirely efficient at quenching DMS methylation activity, therefore, if the protein-DNA complex was disrupted due to the presence of β-mercaptoethanol, there was enough residual DMS remaining to methylate the released DNA fragments (Brunell A and Schleif RT 1987).

To determine which of the above possibilities was responsible for the null footprinting result, the AR DBD-DNA complex was examined by EMSA after exposure to DMS and/or β-mercaptoethanol. If the protein-DNA complex is disrupted by either DMS or β-mercaptoethanol then an EMSA would allow visualization of the disrupted complex. Additionally, an EMSA would reveal the proportion of DNA bound by the AR DBD relative to the free population.

The EMSA revealed that the entire DNA probe was bound by AR DBD for the given experimental conditions and that DMS minimally disturbed this particular protein-DNA complex (Figure 10). However, 1.0 M β-mercaptoethanol completely disrupted the AR DBD-DNA complex either when it was added after the protein-DNA complex was brought to equilibrium or after the addition of DMS. An alternate method, therefore, was required to stop DMS methylation activity without disrupting the AR DBD-DNA complex.

Using EMSA to visualize the AR DBD-DNA complex provided an alternative method to quench DMS activity and recover the bound DNA population. Since DMS is overall a neutral molecule, EMSA permitted the electrophoretic separation of the negative
Figure 9. Failed MeP of AR-Histag on the probasin -426 to +28 base pair promoter fragment using the conventional MeP assay, which employs β-mercaptoethanol to quench DMS activity. The 'UNBOUND' lane refers to methylated DNA that has not been bound by the AR DBD protein. The 'BOUND AR DBD' lane refers to DNA that has been bound by the AR DBD prior to DMS methylation. The positions of ARE1 and ARE2 are indicated.
Figure 10. 1.0 M β-mercaptoethanol disrupted the AR DBD-probasin promoter (-426 to +28 base pairs) complex whereas 19 mM DMS had little or no effect on the protein-DNA complex.
DNA from DMS relinquishing the need to add a quenching chemical such as β-mercaptoethanol.

Using this modified MeP method revealed AR DBD-guanine contacts within the -426 to +28 base pair probasin promoter fragment (Figure 11). As anticipated, the previously identified AREs (ARE1 at -241 to -227 base pairs and ARE2 at -136 to -122 base pairs) were protected in the pattern expected, with guanines in both half sites making the obligatory major groove contacts with the AR DBD. The sequences of the known probasin promoter AREs with the resulting underlined protein contacts as revealed through MeP are as follows: ARE1-5' ATAGCA TCT TGTTC T 3' and ARE2-5' AGTACT CCA AGAA CC 3'.

Unexpectedly, two additional AREs were identified using the MeP assay at positions -209 to -195 base pairs and -107 to -93 base pairs (Figure 11 and 12). These novel AR binding sites within the probasin promoter displayed an unusual methylation pattern in that the guanines located at either end of each respective ARE were hypersensitive to, rather than protected from, DMS methylation. Due to this unusual hypersensitivity these AREs were termed G-sites reflecting the hypermethylated guanines present within the ARE. The sequences of these novel AREs identified using MeP are as follows (with the protected guanines underlined and the hypermethylated guanines in bold): G-1 -5' GGGACA TAA AGCCCA 3' and G-2 -5' ATGACA CAA TGTCAA 3'.

The hypersensitivity observed within G-1 and G-2 belied the uniform methylation activity normally observed for DMS. For a single guanine to be subjected to a higher DMS methylation rate than neighbouring guanines inferred that the guanine was more available for methylation than its neighbour. This has been observed with other protein-DNA interactions employing DMS as a chemical probe (Espinas ML 1995, Ramesh V and Nagaraja V 1996). Local structural distortion of the DNA resulting from AR DBD binding could have exposed guanines to a higher methylation rate. This suggested that the allosteric binding of AR DBD to either the G-1 or G-2 ARE resulted in a local DNA structural conformation change.
Figure 11. Modified MeP of the AR DBD-Histag on the probasin promoter from −269 to −77 base pairs revealed the previously described ARE1 and ARE2 (Rennie PS 1994) and two additional novel AREs; G-1 and G-2, which displayed hypersensitivity to DMS.
Figure 12. The probasin promoter sequence from -426 to +24 base pairs. MeP revealed two novel AREs, G-1 and G-2, at positions -195 to -209 base pairs and -93 to -107 base pairs respectively, which displayed unconventional protection patterns and hypersensitivity to DMS. Protected guanines are indicated by an open circle, whereas hypersensitive guanines are indicated by a dark circle.
3.3 Local DNA Structural Distortion Resulting From AR DBD Allosteric Binding was Intrinsic to the DNA Sequences of the G-1 and G-2 AREs

Local DNA structural distortion, as measured by the DMS hypersensitive sites in the MeP assay, could manifest from one of two possible mechanisms. The probasin promoter fragment examined by MeP contained four AREs that cooperatively interact where the protein-DNA interface likely comprises of protein-protein and protein-DNA interactions. The first mechanism that could underlay the observed guanine hypersensitivity was that the resultant local DNA structural distortion had arisen from the collective interaction of neighbouring DNA-bound AR DBD homodimers. This buttressing effect of the AR DBD homodimers could have modified the AR DBD contacts at either G-1 or G-2. Secondly, the specific nucleotide sequence of both G-1 and G-2 could have affected how the AR DBD bound DNA, adopting a different conformation than that observed with conventional AREs. This last proposed mechanism implies that the primary DNA nucleotide sequence dictates the conformation of the bound AR homodimer and therefore its function.

To determine which proposed mechanism was responsible for the observed methylation hypersensitive sites found within G-1 and G-2, the individual probasin promoter AREs were removed from the larger promoter context and the influence of neighbouring AREs and subjected to the MeP assay. If the hypersensitivity observed was due to the buttressing effect of neighbouring DNA-bound AR DBD homodimers, then we would expect the G-1 and G-2 sites to lose their hypersensitivity and revert to a more conventional protection pattern. If the hypersensitivity to DMS methylation was intrinsic to the nucleotide sequence of G-1 and G-2, then we would expect that the hypersensitive guanines would be maintained despite having removed G-1 and G-2 out of the larger context of the probasin promoter.

The MeP assay was applied to the fragments of the probasin promoter containing combinations of AREs in their natural promoter context. These combinations were G1-ARE2-G2, ARE1 – G1 and ARE2 – G2 (Figure 13, 14 and 15). The results from these smaller combinations of probasin promoter AREs did not alter from the original observation in that ARE1 and ARE2 invariably revealed conventional AR DBD contacts whereas G-1
Figure 13. MeP of the probasin promoter sequence from -269 through -77 base pairs. MeP of this fragment revealed that the hypersensitivity to DMS methylation within G-1 and G-2 was retained in the presence of ARE2.
Figure 14. MeP of the probasin promoter sequence from -269 through -164 base pairs. MeP of this fragment revealed that the hypersensitivity to DMS methylation within G-l was retained in the presence of ARE1.
Figure 15. MeP of the probasin promoter sequence from −150 through −77 base pairs. MeP of this fragment revealed that the hypersensitivity to DMS methylation within G-2 was retained in the presence of ARE2.
and G-2 maintained their unconventional hypersensitivity to DMS methylation. Applying the MeP assay to the individual G-1 and G-2 sites revealed that hypersensitivity to DMS was also maintained (Figure 16A and 16B). These results support the proposal that the nucleotide sequence of both G-1 and G-2 dictate how the AR DBD binds DNA and, consequently, imply that nucleotide sequence of the ARE dictates function to the bound AR homodimer. This was the first time that two different classes of AREs have been described based on the allosteric effects of AR binding. AREs that display conventional guanine contacts were termed Class I AREs whereas those AREs that displayed hypersensitivity to DMS methylation in place of protected guanines were termed Class II AREs.

### 3.4 Hypersensitive Guanines in Class II AREs are Necessary for AR DBD Binding

A MeP assay will generally show those guanines that make weak interactions with the bound protein. What was not clear from the MeP results was whether the observed G-1 and G-2 hypersensitive guanines had any functional role in the AR DBD-DNA interaction. The hypersensitive guanines could either be integral to the AR DBD binding to DNA or incidental in that their presence was not required for the stability or conformation of the bound homodimer. To determine the necessity of the hypersensitive guanines, a Mel assay was performed on the individual G-1 and G-2 AR binding sites.

A successful Mel assay will enhance in the free DNA population, those nucleotides required for protein-DNA interactions whereas these rejected nucleotides will be absent from the bound DNA population (Figure 8). Mel involved pre-methylating the DNA before performing the protein-DNA binding reaction, followed by a separation of the AR DBD bound from the free DNA populations using EMSA. As described earlier, the Mel assay was unsuccessful in probing nucleotide contacts of AREs that interact cooperatively because of the buttressing effect of neighbouring AR DBD homodimers. Individual AREs, however, would have no such buttressing effect; therefore, the Mel assay remained a useful tool.

The Mel results revealed that the G-1 guanines that were found to be hypersensitive to DMS methylation in the MeP assay were also relegated to the free DNA population in the Mel assay just as the conventionally protected guanine contacts in ARE2 were found to be necessary and therefore enhanced in the Mel assay (Figure 17). These results revealed that
Figure 16. MeP of the individual G-site probasin promoter AREs. A. The promoter fragment from -229 through -164 base pairs reveals that the hypersensitive guanines were intrinsic to the G-1 nucleotide sequence. B. The promoter fragment from -121 through -77 base pairs reveals that the hypersensitive guanines were intrinsic to the G-2 nucleotide sequence.
Figure 17. MeI of the G-1 and ARE2 from the probasin promoter. To differentiate from the MeP assay, those guanines found to be necessary for AR DBD-Histag binding are indicated by an open triangle. The G-1 guanines found to be hypersensitive in the MeP assay, indicated in bold, are clearly important for AR-Histag binding since they are absent from the bound fraction and slightly enhanced in the free fraction.
the hypersensitive guanines were indeed necessary for the AR DBD interaction with Class II AREs. Therefore, in addition to the unusual sequence-dependent DMS hypersensitivity revealed in the MeP assay, these guanines in Class II AREs appear to also play a functional role in binding site recognition, subsequent conformation change, and/or stabilization of the AR-DNA interaction. It is possible that the guanines are initially involved in hydrogen bonds with the AR in the recognition of Class II sites, as they are in Class I sites. After this initial recognition, there is an allosteric conformational change to the Class II binding sites in which the interaction with the guanines is altered and subsequently become hypersensitive to DMS methylation.

3.5 The Human PSA Enhancer Contained Class I and Class II AREs and the Two Classes Possess Distinctive Nucleotide Features

The results pertaining to Class II AREs, which displayed local structural distortion upon AR DBD binding, were thus far limited to the rat probasin promoter region. It is possible that the hypersensitive sites were merely artefacts of the probasin promoter and not translatable to other androgen-regulated gene regions. To investigate the universality of Class II AREs and their trademark hypersensitive guanines, the human PSA enhancer region from -4267 to -4062 base pairs was examined using the MeP assay modified for cooperative systems. This region was shown to harbour at least four AREs, referred to in an earlier study as V, IV, III, and IIIa. All four AREs interact in a cooperative manner and each site contributes to full androgen-specific induction (Figure 6) (Huang W 1999).

MeP results showed that the PSA enhancer region possess both Class I and Class II AREs respectively as indicated by conventional AR DBD-ARE contacts and hypersensitivity to DMS methylation (Figure 18A, 18B and 18C). The previously characterized AR binding sites, V and IIIa, within the PSA enhancer possessed the same distinctive hypersensitive guanines as reported above for the probasin promoter. The high affinity PSA enhancer site III made similar conventional contacts as observed within Class I elements. Two classes of AREs, therefore, were present in at least two known androgen-regulated promoters that display cooperative binding. The PSA enhancer site IV appeared similar to a Class II element.
Figure 18A. MeP of the human PSA enhancer from -4267 to -4062 base pairs. The previously identified AREs III and IV possess conventional protection pattern whereas ARE V clearly possesses unconventional hypersensitive guanines.
Figure 18B and 18C. B. MeP of the PSA enhancer ARE IIIA possesses hypersensitivity that resembles the probasin G-1 and G-2 AREs. C. The PSA enhancer sequence from -4267 to -4062 base pairs with protected guanines indicated by open circles and hypersensitive guanines indicated by dark circles.
but did not have enough guanines to confidently classify. Finally, MeP analysis of the PSA AREs V and IIIa further revealed that additional guanines flanking Class II sites were hypersensitive suggesting that structural distortion resulting from AR DBD binding could extend over an area of at least 17 base pairs (Figure 18A and 18B).

A comparative table was drawn up to determine a consensus sequence for each of the two classes of AREs (Table 2). Class II AREs from the probasin promoter and the PSA enhancer were aligned in order to determine whether these novel AREs possessed specific representative or consensus nucleotides when compared to aligned conventional Class I AREs. Along with verified Class I and II AREs from the probasin promoter and PSA enhancer, proposed Class I and II AREs based on sequence similarity from known AREs from other androgen-regulatory regions such as the mouse Slp promoter, mouse E-RABP promoter and proximal PSA promoter were also used to generate a consensus sequence. The nucleotide numbering scheme employed to describe the ARE sequence assigns the number 0 to the central nucleotide in the spacer region of the inverse palindrome. From this starting point the nucleotide position number ascends to +7 towards the 3’-half site and descends to -7 towards the 5’-half site (Table 2).

A comparison of the two consensus sequences reveals common and distinctive features that distinguish Class I and Class II AREs. Both Classes have a cytosine at position -3 and a guanine at +3 where the guanines at these positions in each half site make a conserved hydrogen bond with the arginine residue found within the AR recognition α-helix (Figure 3B and Table 2). Both Class I and II AREs possess an adenine at position -4 in the 3’ half site where a conserved Van der Waals contact is made between the thymidine in this position and a valine found within the AR recognition P Box (Figure 3 and Table 2). Unlike the Class I consensus sequence, there is no conserved adenine-thymidine base pair at position +4 in the 5’ half site of the Class II consensus sequence. The absence of this Van der Waal’s contact in the 5’ half site suggests that, in general, the AR homodimer binds asymmetrically to the Class II ARE. A pronounced feature of a Class II site is a unique guanine at position -5 and cytosine at position +4 and/or +5. The most prominent feature of Class II sites is a purine at position +7 whereas Class I sites posses a highly conserved pyrimidine at this
Table 2. Alignment of AREs into proposed Class II and Class I consensus sequences. AREs from the probasin promoter, PSA enhancer, PSA promoter, Slp enhancer, and mE-RABP promoter were employed to generate consensus. The consensus nucleotides that are common to both Class I and Class II sites are underlined. The consensus nucleotides that are distinctive to each class type are indicated by an open circle for Class I nucleotides and a black-filled circle for Class II nucleotides. Italicized elements are proposed by sequence similarity to the Class designation.
position. Outside of the Class II consensus sequence, a thymidine was found conserved at position -13 in all cases and an adenine was found at position +13 in all but one case. Finally, there is a tendency to have a higher AT-rich flanking region and spacer in Class II AREs in comparison to Class I AREs.

3.6 **Mutation of Key Nucleotides Transforms the MeP Pattern of Class II Half Site to a Class I Half Site**

The nucleotide sequence of Class II AREs was shown to dictate the bound AR DBD conformation and the results from the MeI assay had shown that the hypersensitive guanines appear to be critical for DNA binding site recognition, subsequent allosteric conformational changes, and/or stability of the protein-DNA complex. To further delineate the role of specific nucleotides within the sequence of Class II AREs, key residues within the Class II element G-1 were altered such that the modified sequence partially resembled a Class I consensus sequence. The MeP assay was then used to determine whether these mutations affected the DMS methylation pattern and consequently the bound-AR DBD conformation.

The distinctive features of Class II AREs include a relatively AT-rich spacer region, a thymidine at position -13 relative to the central spacer nucleotide, and a purine at position +7 in the 3' half site in contrast to the consensus pyrimidine found at this position in Class I sites. These distinct features were examined for their contribution to the DMS hypersensitivity pattern through site-directed mutagenesis.

The results from the DMS MeP assays demonstrate that conversion of the thymidine to a guanine at position -13, six base pairs upstream of the 5' half site, decreased the guanine hypersensitivity in the 5' half site compared to wildtype G-1 but did not affect the binding pattern of the 3' half site (Figure 19A and 19B). When the adenine at position +7 of the 3' half site was converted to a cytosine, the DMS protection pattern of the 3' half site reverted to the pattern observed with a conventional Class I binding site but did not alter the hypersensitivity in the 5' half site (Figure 19C). Finally, converting the palindrome spacer nucleotides from TAA to CGG also decreased the hypersensitivity in the 5' half site alone (Figure 19D). These results suggest that Class II type binding by the AR is primarily distinguished by the non-consensus purine at position +7 within the 3' half site but is also
Figure 19. The change in the relative DMS methylation hypersensitivity and protection pattern of the probasin G-1 element after site directed mutagenesis. Dark bars represent those guanines hypersensitive to DMS methylation and grey bars represent those guanines protected from DMS methylation relative to unbound DNA. A. G-1 wildtype oligo. B. T215G oligo where the thymidine at position -13 was converted to a guanine. C. A195C oligo where adenine at position +7 was altered to a cytosine. D. Spacer oligo where the TAA spacer region was converted to CGG.
influenced by the identity of other nucleotides flanking the ARE and within the spacer region.

3.7 Specific Arrangements of Class I and Class II AREs Increase Cooperative Binding

Within the steroid receptor field, it is accepted that there is a conventional steroid receptor DNA response element whose nucleotide sequence will vary slightly in nature depending on receptor subtype, but nevertheless will maintain conventional contacts between the steroid receptor and the DNA response element. Contrary to this supposition, we have discovered two classes of AREs based on the primary nucleotide sequence within the probasin promoter and PSA enhancer. The newly discovered Class II AREs display unconventional local DNA structural distortion along with the absence of previously assumed obligatory protein-DNA contacts. We have shown that the nucleotide sequence of an ARE dictates how the AR DBD binds in that, upon AR DBD binding, specific guanines within Class II AREs were hypersensitive to DMS methylation and, furthermore, that these guanines were necessary for the AR DBD-DNA interaction.

The next question was whether the nucleotide sequence, which dictated the conformation of the bound AR DBD homodimer, also dictated function. Mammalian AR-regulated promoters such as the probasin and the PSA gene contain multiple low-affinity AREs that cooperatively interact to dramatically increase overall affinity for the AR. The molecular mechanisms that underlie cooperative binding are unknown. However, the existence of two classes of AREs that result in different conformations of bound AR homodimers imply that the interplay between these two classes of AREs could provide a molecular basis for cooperative binding.

It is accepted that cooperative binding likely involves both protein-protein and protein-DNA interactions (Senear DF 1998). To determine the contribution of protein-DNA interactions to cooperativity, the relative binding affinities of multiple binding sites were compared to the specific binding affinity of individual binding sites. A distinctive feature of cooperative systems is that a much lower concentration
of protein is required to fully occupy the multiple binding sites that make up a cooperative system compared to a single binding element (Mao C 1994, Kerppola T 1998, Liu Z 1998, Senear DF 1998). Using this feature, it was possible to determine the contribution of each ARE class type to cooperativity by comparing binding affinities. Subsequently this comparison can determine the contribution of nucleotide variations in each ARE class type and the concomitant differences in bound-AR DBD conformation to the cooperativity function.

It has been established that the concentration of protein required to fully occupy the multiple AREs found within the −426 to +28 base pair probasin promoter fragment, was at least 10 fold lower than that required to individually occupy either ARE1 or ARE2 (Rennie PS 1994). From the above data, there exists two additional AREs within the probasin promoter and it is apparent that two different classes of AREs likely contribute to the cooperativity of this system. To determine the contribution of these two classes of ARE to cooperativity, the specific binding affinity of each of the four individual AREs was determined using Scatchard Analysis. These analyses were performed using a constant amount of purified DNA in an EMSA to avoid inconsistencies caused by protein dilution over large ranges of concentration. This affinity measurement was then compared to the relative binding affinity of different combinations of AREs that contained both Class I and Class II AREs in their natural promoter context.

The Scatchard Analysis revealed that the four AREs found within the probasin promoter individually had a weak specific binding affinity for the AR DBD (Figure 20). The relative binding affinities of the AR DBD were 6.3 nM for ARE2, 11 nM for ARE1, 14 nM for G-1, and 18 nM for G-2. Although the isolated Class I sites of the probasin promoter have slightly higher affinity than the Class II sites, all elements displayed measurable binding activity. The low affinity for the AR DBD for individual AREs was expected since many of the multiple AREs found within androgen-regulated promoters individually displayed weak affinity for the AR.

Within the probasin promoter context, the natural arrangement of the AREs alternated the two classes of AREs in the following order: 5′ ARE1 – G1 – ARE2 – G2 3′ (Figure 12). To determine the functional contribution of Class II AREs to
cooperative AR DBD binding to the probasin promoter, a series of DNA fragments were created containing combinations of ARE1, G-1, ARE2, and G-2 in their native context. These combinations of Class I and II elements were analyzed to determine the relative contribution of individual binding sites to cooperativity. This was performed by measuring AR DBD-DNA complex formation as a function of increasing DNA concentration.

In all combinations tested a substantial level of cooperativity was observed in that the concentration of AR DBD required to fully occupy multiples of AREs was lower than that needed to occupy any of the individual AREs (Figure 20). These analyses revealed that the combination of ARE1 and G-1 (with $K_d$ values of 11 nM and 14 nM in isolation respectively) resulted in a cooperative interaction that half-saturated the DNA at more than 10-fold lower concentration (0.9 nM). Similarly ARE2 and G-2 (with $K_d$ values of 6.3 nM and 18 nM in isolation respectively) interacted in a cooperative manner to shift the rate of complex formation an order of magnitude lower in concentration. The promoter fragment containing G-1, ARE2, and G-2 interacted with the highest degree of DNA binding cooperativity resulting in half-saturated binding at 0.11 nM (Figure 20). This last probasin promoter fragment containing G1-ARE2-G2 displayed over 50 times stronger binding compared to ARE2, the individual element with the highest DNA binding affinity.

The combination of ARE1-G1-ARE2-G2 resulted in half-saturated binding at 1.2 nM, which intriguingly demonstrates that the addition of ARE1 to the extremely cooperative G1-ARE2-G2 promoter fragment weakened the overall strength of the DNA binding complex (Figure 20). Further examination of the ARE1-G1-ARE2-G2 promoter fragment revealed that the binding curve was biphasic. The order of complex formation appears to be occupation of G1-ARE2-G2 at low concentrations of protein, followed by additional binding to ARE1 as protein concentration increases.

Overall these results indicate that Class II elements are instrumental in providing the dramatic level of DNA binding cooperativity observed on the probasin promoter. Furthermore it is apparent that the four individual elements interact in a complex manner where G-1 can interact individually with ARE1 or ARE2-G2.
Figure 20. Individual AREs displayed weak binding affinity for the AR. Cooperative binding was displayed between the AREs within the ARE1-G1 and ARE2-G2 probasin promoter fragments. The highest cooperative probasin promoter fragment contained G1-ARE2-G2. The addition of ARE1 to this highly cooperative region decreased overall binding affinity and cooperativity for the fragment. The X axis denotes amount of free DNA probe in nM whereas the Y axis denotes the log of AR DBD bound probe in nM.
However, if all four elements are present on the same DNA fragment then the interaction between G-1 and ARE2-G2 is weakened, presumably through G-1’s interaction with ARE1, thereby affecting the stability of the overall complex.

3.8 Class II AREs Contribute to Synergistic Transactivation as Observed in Transient Transfection Assays

Thus far the probasin proximal promoter-AR DBD binding characteristics have been described using in vitro methods of analysis. To determine whether the in vitro data had biological significance, it was necessary to examine the above results in the context of an in vivo system. Cooperative binding of transcription factors to a promoter region did not necessarily imply that AR-specific regulation or synergistic transcription would result. This is because cooperative binding and synergistic transcription are two different albeit linked processes. The presumption is that a specifically regulated promoter segment would not only be occupied by the correct transcription factors as promoted by cooperative binding, but that the resultant conformation of the transcription factor-DNA multicomplex would recruit essential coactivators that in turn assist in recruiting the basal transcriptional machinery. The resultant transcription factor–coactivator-DNA complex, termed the ‘enhanceosome’ and its formation likely depends on the primary nucleotide sequence of the gene regulatory region, which not only dictates the specific type of transcription factor that binds but also the resultant conformation of the recruited protein-DNA multicomplex.

Transient transfection assays were employed to determine whether the cooperative binding observed for the various combinations of AREs used in the in vitro binding assays, translated into AR-mediated synergistic transcriptional activation. Using the human prostate cancer cell lines LNCaP and PC3, promoter constructs of individual AREs or the four different ARE combinations described above were attached to a luciferase reporter gene and transiently transfected along with increasing concentrations of R1881, a synthetic androgen. As expected the individual AREs displayed minimal luciferase expression (less than 2000 RLUs) that
Figure 21. The X axis denotes R1881 concentration in nM. The Y axis denotes relative luciferase units corrected for Renilla expression. A. The individual AREs all displayed weak transactivation ability. B. The combination of either ARE1-G1 or ARE2-G2 resulted in additive transactivation relative to the individual AREs. C. The most synergistic combination was the ARE1-G1-ARE2-G2 promoter fragment.
were maximal at 0.5 nM R1881, which corresponds to 12-, 23-, 25-, and 33- fold induction for G-1, ARE1, G-2, and ARE2 respectively (Figure 21A). Combining Class I and Class II sites resulted in an increase in sensitivity to hormone concentration which was maximal at ~0.5 nM R1881 (Figure 21B). However, the magnitude of transcriptional activation of ARE1-G1 and ARE2-G2 was at most additive when compared to the activity observed for the individual elements. The addition of G-1 to the ARE2-G2 binding element displayed a dramatic increase in transcriptional activation to a maximum of 169-fold induction (~20 000 RLUs), which can be described as synergistic (Figure 21C). The synergistic transcriptional activation observed with the G1-ARE2-G2 promoter fragment is consistent with the increase in DNA binding cooperativity observed when G-1 was combined with ARE2-G2 in the in vitro binding assays (20B).

The addition of ARE1 to the highly cooperative G1-ARE2-G2 DNA binding region resulted in a biphasic curve from 0.01 to 0.05 nM of R1881 followed by a sharp increase in the level of activation with increasing hormone concentration. The ARE1-G1-ARE2-G2 promoter fragment culminated in a 5-fold enhanced transcriptional response (~100 000 RLUs) compared with the G1-ARE2-G2 fragment and more than 30-fold greater than any individual ARE (Figure 21C). The observed high relative transcriptional activity of the ARE1-G1-ARE2-G2 promoter fragment contrasts the cooperative binding data, which demonstrated that the addition of the ARE1 element decreased the overall binding strength of the complex (Figure 20).

Together the in vitro binding data and the in vivo transient transfection data suggest that a biphasic curve of activity arises from the highly cooperative DNA binding complex at low concentrations of R1881, presumably corresponding to concentration of activated AR in the nucleus. At high levels of R1881 (or activated AR), however, the inclusion of the ARE1 binding site to the probasin promoter fragment is able to provide highly synergistic levels of transcriptional activation. Overall, these data imply that unique combinations of Class I and Class II elements are required for maximal transcriptional activation by the AR and that each ARE plays an unique functional role in transcriptional activation, DNA affinity, and complex stability. Intriguingly, these observations suggest that the AREs within the
probasin promoter function as a rheostat sensitive to AR concentration. A slight increase in androgen signal ensures that the highly cooperative arrangement of Class I and Class II AREs are occupied and transcription is activated. Another slight increase in androgen signal ensures full occupation of neighbouring AREs that result in maximal transcriptional activation.

In summary, the above analyses demonstrate that the AR binds to two structurally and functionally distinct classes of AREs that are directed by allosteric interactions of the binding complex. AR binding to conventional Class I sites have been previously recognized and employ conventional nucleotide contacts that are used by other nuclear receptors. In contrast, Class II binding sites result in DNA-structural alterations that display DMS hypersensitivity to guanines that are normally contacted in Class I sites, where they are protected from DMS methylation activity. In isolation both classes of AR binding sites are of low DNA binding affinity and transcriptional activity. Unique combinations of Class I and Class II sites result in dramatic cooperative DNA binding and a highly synergistic effect upon transcriptional activity. This complex and composite function that facilitates cooperative DNA binding and achieves a synergistic level of transcriptional response is dictated by the primary nucleotide sequence to which AR binds.

3.9 Cooperative Protein-Protein Interactions Between AR Homodimers Could Involve Stabilizing Disulfide Bridges

Initial attempts to employ β-mercaptoethanol to quench DMS methylation activity resulted in the total disruption of the cooperative AR DBD-DNA complex that formed on the probasin promoter segment (Figure 10). β-mercaptoethanol is a reducing agent that has been employed to reduce or disrupt disulfide bonds that can form between free cysteine residues located within or between proteins. The AR DBD possesses two presumably free cysteine residues: one in the first zinc motif and one in the hinge region (Figure 2A). Neither of these cysteines has been shown to be necessary for AR DBD recognition or binding to an individual ARE, however, it is
conceivable that either or both residues could be involved in cooperative interactions with neighbouring DNA-bound AR DBDs.

If β-mercaptoethanol disrupts AR protein-protein interactions that are necessary for cooperative DNA binding but does not disrupt the binding of an AR DBD homodimer to an individual ARE, then this can be examined using an EMSA. β-mercaptoethanol at 1.0 M concentration disrupted the highly cooperative G1 – ARE2 – G2 probasin promoter fragment as observed earlier with the −426 to +28 probasin promoter fragment (Figure 10 and 22A). To a lesser extent, 1.0 M β-mercaptoethanol also disrupted the binding of the AR DBD complex to the individual Class II G-1 element (Figure 22B). Intriguingly, AR DBD binding to the canonical SRE was not disturbed by 1.0 M β-mercaptoethanol (Figure 22C). These data suggest that novel protein-protein interactions arise from AR DBD binding to natural AREs but do not arise from AR DBD binding to the promiscuous SRE, supporting the argument that nucleotide sequence of natural AREs affects the conformation of the bound AR DBD.

3.10 Recombinant Sox-4 Protein Increases the Number of AR DBD Monomers that Bind the Probasin Promoter

The primary and flanking nucleotide sequence of the probasin promoter Class II AREs G-1 and G-2 also matched the sequence that potentially could be recognized by an HMG-box architectural protein from the Sox family (MatInspector 2.2). Sox proteins possess an HMG box domain that binds the minor groove of AT-rich sequence and intercalates a hydrophobic residue into the DNA structure. This intercalation results in bending the DNA towards the minor groove effectively opening the major groove and thereby facilitating transcription factor access. Architectural proteins have been implicated in the facilitation of many cooperatively interacting protein-DNA systems such as NF-κB protein and the IFNα promoter (see Introduction, Section 1.7).
Figure 22A and 22B. EMSA showing β-mercaptoethanol disruption of the AR DBD-DNA binding complex. **A.** 1.0 M of β-mercaptoethanol disrupted AR DBD binding to the highly cooperative G1-ARE2-G2 probasin promoter fragment. **B.** 1.0 M β-mercaptoethanol moderately disrupted AR DBD binding to the Class II G-2 element.
**Figure 22C.** 1.0 M of β-mercaptoethanol had negligible effect on AR DBD binding to the canonical SRE.
Database research revealed that Sox-4 should recognize the minor groove sequence that coincides with both G-1 and G-2 (MatInspector 2.2). To determine whether the Sox-4 protein would specifically contact the downstream proximal probasin promoter, a MeP assay was employed using purified recombinant Sox-4 histidine tagged protein. Since piperidine would only cleave methylated guanines in the major groove, its use would not yield information on Sox-4 minor groove contacts. Instead, NaOH was employed since it will cleave both methylated guanines in the major groove and adenines in the minor groove. DMS methylates adenines at the N3 position in the minor groove thus exposing Sox-4 – DNA interactions. Initial footprinting results were negative despite Sox-4 successfully shifting the G1–ARE2–G2 probasin promoter fragment in an EMSA (data not shown). When Sox-4 was combined with the AR DBD in an MeP assay however, the number of AR DBD monomer binding sites within the proximal probasin promoter fragment increased (Figure 23). These additional binding sites were assumed to be AR DBD and not Sox-4 since the protected sites were guanines and therefore in the major groove and, furthermore, the nucleotide binding sequence resembled ARE half-sites.

3.11 MeP of Other AR-Regulated Promoter and Enhancer Regions Reveal Novel Protection Patterns that Suggest Monomer and Direct Repeat AR-DNA Binding

In addition to the –426 to +28 base pair probasin promoter and the PSA enhancer, three other AR-regulated regions were examined using the modified MeP assay. The mE-RABP promoter from -553 to +28 base pairs revealed two Class I-type AREs at -432 and -418 base pairs and -74 and -91 base pairs termed AR binding site-1 (ARBS-1) and AR binding site-0 (ARBS-0) respectively (Figure 24A and B) (Lareyre JJ 2000).
Figure 23. Sox-4 protein increases the occupancy of AR DBD on the probasin promoter. Protected guanines are indicated by open circles whereas hypersensitive guanines are indicated by dark circles. In the presence of Sox-4, additional guanines protected by the AR DBD and hypersensitive to DMS are indicated by an asterisk and an ‘H’ respectively.
The rat AR enhancer from position 2028 to 2333 base pairs revealed a Class II-type ARE at 2165 to 2179 base pairs that was previously described as an AR-responsive intravenous sequence (IVS) along with a potential Class I-type ARE at 2073 to 2087 base pairs that was previously described as a putative ARE (Grad JM 1999) (Figure 25). Within the AR exonic enhancer there are two more previously described AREs (AR enhancer ARE1 and ARE2), which require greater resolution to determine the exact AR DBD contacts, however, the MeP information acquired in this study suggests that ARE2 at 2250 to 2272 base pairs may comprise of at least two direct repeats of the AR DBD separated by four base pairs (Grad JM 1999) (Figure 25).

The proximal p21 promoter from position -288 to -1 base pairs revealed AR DBD protected and hypersensitive guanines at the previously described p21 ARE at -209 to -195 base pairs and a novel Class II-type ARE at -39 to -25 base pairs (Lu S 1999) (Figure 26). An additional novel yet unresolved ARE or cluster of AREs between -177 and -141 base pairs was also discovered although, like the rat AR enhancer, greater resolution is required to confirm and establish the specific AR DBD contacts within this region of the p21 proximal promoter (Figure 26).

Finally, the -705 to -426 base pair upstream proximal probasin promoter revealed an intriguing cluster of AR DBD monomer and homodimer binding sites between -529 and -509 base pairs, which corresponds to the AR binding site-3 (ARBS-3) (Matusik R, personal communication) (Figure 27). An additional novel cluster of AR DBD binding sites was revealed using the MeP assay between -577 and -554 base pairs (Figure 27).

The MeP results for the rat AR enhancer, human p21 promoter, and the rat upstream proximal probasin promoter (-705 to -426 base pairs) are preliminary results at best (Figures 25, 26, and 27). Smaller DNA fragments spanning each AR DBD binding region will yield greater resolution and, therefore, greater information on the specific AR DBD contacts within these AR DBD binding regions of DNA. Having stated that more detailed work is required, the preliminary results from these other AR-regulated DNA regions suggest that the AR DBD may not be restricted to a
homodimer structure bound to an inverse DNA palindrome. Others have suggested that the AR could bind as a homodimer to a direct repeat although this has yet to be shown either in an X-ray crystal structure or in functional studies (Zhou Z 1997, Claessens F 2001). The MeP results for the mE-RABP promoter, rat AR enhancer, human p21 promoter and the upstream proximal probasin promoter also suggest that the AR DBD could bind as a monomer (Figures 24A and B, 25, 26, and 27). It is not clear from this study whether this observation is the result of DNA-bound AR DBD homodimers recruiting AR DBD monomers to the DNA or if AR DBD monomer binding occurs regardless of the larger DNA context. The GR DBD is known to bind as a monomer to DNA half site sequences whereas the full length GR fails to bind as a monomer (Chalepakis G 1990, Segard-Maurel I 1996). Therefore, the AR DBD monomer binding observed in this study may be a similar artefact. Overall, these intriguing results suggest that AR may be able to bind in unconventional ways although this has yet to show any functional significance.
Figure 24A. Methylation protection of the AR DBD on the mE-RABP promoter from −543 to −166 base pairs. The previously described AR binding site-1 (ARBS-1) is indicated (Lareyre JJ 2000). Protected guanines are indicated with an open circle.
Figure 24B. Methylation protection of the AR DBD on the mE-RABP promoter from -195 to +26 base pairs. The previously described AR binding site-0 (ARBS-0) is indicated with the protected guanines represented by an open circle (Lareryre JJ 2000).
Figure 25. MeP of the AR DBD on the rat AR enhancer from 2028 to 2333 base pairs. Previously established ARE-2, IVS (intravening sequence), and the putative ARE reveal DMS protected and hypersensitive guanines (Grad JM 1999).
Figure 26. MeP of AR DBD on the p21 promoter from -285 to -1 base pairs. Previously described p21 ARE is indicated (Lu S 1999).
Figure 27. Methylation protection of the AR DBD on the upstream proximal probasin promoter from −705 to −426 base pairs. ARBS-3 is indicated (Matusik R personal communication).
4. Future Direction and Conclusion

4.1 Summary of Project Results

Previous observational research has shown that the AR will cooperatively bind to multiple AREs within the transcriptional regulatory regions of targeted genes such as the rat probasin promoter and the human PSA enhancer (Tsai et al, 1989; Schule et al, 1988; Strahle et al, 1988; Kasper et al, 1994; Huang et al, 1999). There are several previously studied systems that use cooperative binding of transcription factors to DNA as a mechanism to increase transcription factor specificity and sensitivity to transcription factor concentration (Thanos and Maniatis, 1992; Zhang and Verdine, 1999; Wood et al, 1998). Close examination of these systems has revealed that both protein-protein and protein-DNA allosteric interactions underlie cooperative binding and subsequent specificity and sensitivity.

The observation that AR cooperatively binds to multiple AREs is an important contribution to understanding androgen action. Current knowledge of steroid receptors cannot explain how the AR specifically regulates its own set of unique genes while simultaneously sharing a common DNA binding site with other steroid receptors. It is logical to propose that cooperative binding contributes to AR-specific transcriptional regulation since cooperative binding has been implicated in the specificity and sensitivity of other transcription factor systems (Thanos and Maniatis, 1992; Zhang and Verdine, 1999; Wood et al, 1998). However, unlike other cooperative systems where interactions occur between different types of transcription factors, the AR interacts cooperatively with other AR molecules. This observation emphasizes the role of protein-DNA rather than protein-protein interactions in cooperative AR binding to DNA since variation in type of DNA-binding protein cannot be a contributing factor. Additionally, it is evident from this study and others that promoter context and nucleotide sequence contributes to specific AR-DNA interactions since specific combinations of AREs in their native context, which individually vary in primary nucleotide sequence from the canonical SRE, are
necessary for cooperative AR binding and AR-specific transcriptional regulation (Tan et al., 1992; Lieberman et al., 1993; Nelson et al., 1999).

In order to determine the contribution of nucleotide sequence to the underlying mechanisms of cooperative binding, this study examined in detail the AREs found within two AR-regulated gene regions; the rat probasin promoter and the human PSA enhancer. Initial attempts at using Mel to examine specific protein-DNA contacts between the AR DBD and multiple AREs within the probasin promoter were ineffectual. Failure likely occurred because the buttressing effect of neighbouring protein-protein and/or protein-DNA interactions within the AR DBD-DNA cooperative complex compensated for any single nucleotide disrupted through DMS methylation. The conventional MeP assay, which employs the reducing agent β-mercaptoethanol to quench DMS methylation activity, was also ineffectual in probing AR DBD-ARE contacts within the probasin promoter from -426 to +28 base pairs. This failure to footprint AR DBD on the probasin promoter was due to the perturbing effects of β-mercaptoethanol on the AR DBD-DNA cooperative complex presumably through the disruption of protein-protein interactions. In contrast, β-mercaptoethanol did not disturb an AR DBD homodimer bound to the canonical SRE suggesting that protein-protein interactions such as disulfide bridges are involved in cooperative AR DBD binding to the multiple AREs found within the probasin promoter. Unlike the canonical SRE, β-mercaptoethanol also disturbed AR DBD binding to the Class II ARE, G-1. This last observation tentatively suggests that protein-protein interactions that are disruptable by β-mercaptoethanol arise from AR DBD homodimer binding to individual AREs found in AR-specifically regulated genes but do not arise from AR DBD binding to the promiscuous viral-derived SRE. This last supports the argument that nucleotide sequence of natural AREs affects the conformation of the bound AR DBD.

Employing EMSA to stop DMS methylation activity instead of β-mercaptoethanol provided AR DBD-DNA contact information for the highly cooperative probasin promoter. In addition to the previously described ARE1 and ARE2, two novel AREs were discovered and were labelled G-1 and G-2 because of their unusual hypersensitivity to DMS methylation at guanines where protection from
methylation was expected. For a guanine to be hypersensitive to DMS methylation implied local structural distortion of the DNA that resulted from AR DBD binding (Espinas et al, 1995; Ramesh and Nagaraja, 1996; Reid and Nelson, 2001). The hypersensitivity to methylation was found to be intrinsic to the nucleotide sequence of both G-1 and G-2 and the hypersensitive guanines were found to be necessary for the AR DBD-DNA interaction. Here was physical evidence that primary nucleotide sequence of an ARE will dictate how the AR DBD binds DNA and there appeared to be at least two classes of AREs according to their MeP pattern upon AR DBD binding. Class I AREs are represented by the probasin promoter ARE1 and ARE2, which displayed conventional AR DBD contacts. Class II AREs are represented by the probasin promoter G-1 and G-2 binding elements, which display unusual hypersensitivity to DMS methylation upon AR DBD binding. Both Class I and Class II AREs were found within the human PSA enhancer implying that differential binding of the AR DBD to AREs that vary in primary nucleotide sequence is not merely a phenomenon isolated to the rat probasin proximal promoter.

When the nucleotide sequences of established Class I or Class II AREs were aligned, it became apparent that there were distinctive as well as common features between the class consensus sequences. Both classes of AREs possess a cytosine at position −3 and a guanine at +3 where the arginine within each AR DNA recognition α-helix within the AR DBD homodimer is expected to hydrogen bond to guanine. Additionally both classes possess an adenine at position −4 where a Van der Waal's contact is expected between valine and the methyl group of a thymidine. Relative to Class I AREs, Class II AREs distinctively possessed an AT-rich spacer region, a thymidine at position −13 (outside of the primary sequence of the ARE), and a purine at position +7 in the 3' half site in contrast to the highly consensus pyrimidine found within Class I AREs at this position. Site directed mutagenesis of these distinctive features revealed Class II type binding by the AR is primarily distinguished by the consensus purine at position +7 where its conversion to a pyrimidine changed the 3' half site MeP pattern to that of a Class I ARE but did not affect the 5' half site MeP pattern. Similarly, site directed mutagenesis of the Class II ARE spacer region or conversion of the −13 thymidine to a guanine appeared to decrease the
hypersensitivity to DMS methylation only within the 5' half site. Overall, the site
directed mutagenesis effects on the DMS protection and hypersensitivity pattern of
G-1 were confined to one particular half site implying that an AR DBD monomer,
within the larger AR DBD homodimer, will recognize and bind uniquely to its own
half site. AREs found within, AR-specifically regulated genes are invariably
asymmetric inverse palindromes suggesting that each half site nucleotide sequence
conveys specific and functionally relevant information to its own bound AR
monomer. The asymmetry noted in mammalian AREs could itself dictate function to
the bound homodimer apart from variation in nucleotide sequence of the ARE or its
relative location.

Scatchard analysis of the individual proximal probasin promoter AREs
revealed that each had a weak specific binding affinity for the AR DBD. When
combinations of Class I and Class II AREs were examined, the highest cooperatively
interacting permutation was the promoter fragment containing G1–ARE2–G2. Half-
saturated binding occurred at 0.11 nM for this highly cooperative fragment, which
was over 50 times stronger binding than ARE2, the individual element with the
highest DNA binding affinity. Intriguingly, adding ARE1 to this combination
weakened overall binding affinity and the binding curve for the larger ARE1-G1-
ARE2-G2 fragment appeared to be biphasic: at low concentrations, AR DBD appears
to first occupy G1-ARE2-G2 followed by occupation of ARE1 at higher AR DBD
concentration. These data suggest that Class I and Class II AREs interact in a
complex manner and that both classes are essential for the dramatic cooperativity
observed on the probasin promoter.

Transient transfection analyses of the probasin promoter fragment
combinations augmented and complemented the above binding affinity data.
Synergistic transcriptional activation was observed with the G1-ARE2-G2 and
ARE1-G1-ARE2-G2 promoter fragments. ARE1-G1-ARE2-G2 exhibited a biphasic
curve that resulted in a 5-fold higher transcriptional response relative to G1-ARE2-
G2. Complementing the binding affinity data, at low concentrations of R1881 the
lower part of the biphasic curve suggests that G1-ARE2-G2, the region with the
highest DNA binding affinity, is fully occupied by activated AR and synergistic
transcription ensues. At higher concentrations of R1881, however, the higher part of the biphasic curve suggests that ARE1 is now occupied and maximal transcription is permitted. Like the binding affinity data, these transfection data imply that both Class I and Class II AREs are required for maximal transcriptional activation by the AR.

Four supplementary AR-regulated promoter or enhancer regions were examined using the modified MeP assay: the mE-RABP promoter, the rat AR enhancer, the proximal human p21 promoter, and the upstream rat probasin promoter from -705 to -426 base pairs. In addition to Class I and Class II AR DBD patterns of methylation protection and hypersensitivity, other novel patterns of AR DBD binding were discovered. AR DBD monomer half site binding was observed along with clusters of nested AR DBD binding, particularly within the -705 to -426 base pair probasin promoter. These AR DBD footprints revealed both conventional and expected AR DBD-DNA contacts as well as unconventional hypersensitivity to methylation such as that observed within Class II AREs. Although the above novel AR DBD binding patterns have yet to show functional significance either in affinity binding or transfection studies, these last results imply that the AR DBD may not be restricted to binding DNA regulatory regions as a homodimer in a traditional head-to-head orientation. Instead, other AR DBD configurations and orientations may be possible depending on DNA sequence and promoter context.

When the sequence of the probasin promoter from -426 to +28 base pairs was analyzed using the internet based software, Matinspector 2.0, it was discovered that the nucleotide sequence of the Class II AREs, G-1 and G-2, coincided with the near perfect binding sequence for the HMG-box architectural protein Sox-4. While the EMSA of recombinant full-length histidine tagged Sox-4 resulted in a shifted protein-DNA complex, no footprint was discerned using the modified MeP assay using NaOH cleavage. When Sox-4 was combined with the AR DBD, however, the number of DNA half sites occupied by AR DBD increased. Further work is required to determine whether Sox-4 makes specific contacts with the probasin promoter and whether the apparent Sox-4 enhancement of AR DBD occupation has an authentic effect on the regulation of prostate-specific and AR-specific genes.
4.2 Future Projections

While other research groups have successfully purified active full length AR, this has yet to be accomplished in our lab although the ability to do so is close at hand. The problem appears to be solubility and maintaining long-term activity of the full length AR after purification from bacterial, mammalian or insect protein purification systems. Once purified, applying the MeP assay to the full length AR will augment the data we have generated using the AR DBD especially with respect to the observed AR DBD monomer DNA binding. When the DNA binding abilities of the full length GR are compared to the binding abilities of the GR DBD, it is clear that the GR DBD is more capable of binding DNA as a monomer than the full length GR (Chalepakis et al, 1990; Segard-Maurel et al, 1996). Given the high sequence homology between the DBD domains of the AR and GR it is reasonable to expect similar results with the full length AR. Illustrating that full length AR generates a similar MeP pattern as the AR DBD will also confirm our results.

To determine the in vivo protein occupation of the probasin promoter and other AR-regulated gene regions one could employ in vivo footprinting techniques using either DNase I or DMS supplemented with ligation-mediated PCR (LM-PCR) (Mueller and Wold, 1989). So far there has yet to be a published footprint revealing in vivo AR-DNA interactions although other DNA-binding proteins have been footprinted in vivo and shown to interact with AR-regulated gene regions (Scarlett and Robins, 1995). The lack of data on AR interactions with DNA may be because these in vivo interactions are transient and therefore unable to be captured using present footprinting techniques or the AR-DNA complex may be easily disrupted in vivo by DNase I or DMS treatment. However, in vivo footprinting using LM-PCR is a challenging technique to master and its use in probing in vivo AR-DNA interactions may yet prove informative.

An alternative to LM-PCR in vivo footprinting is the Chromatin Immunoprecipitation (ChIP) assay, which permits the examination of specific protein-DNA interactions within an in vivo environment (Strahl-Bolsinger et al, 1997). ChIP involves the initial covalent crosslinking of proteins bound to DNA in
vivo, usually using ultraviolet light. After covalently linking the proteins of interest to the DNA, the DNA is removed from the cell, purified from other nuclear and cellular components, and sheared into relatively large fragments. At this point the protein-DNA fragments are enriched for a particular protein of interest by immunoprecipitating with a specific antibody. The specific protein-DNA fragments are then subject to protease digestion and the released DNA fragment that was occupied by the immunoprecipitated protein is PCR amplified using primers specific to the DNA region of interest. The sequence of the amplified DNA reveals the in vivo binding location of the protein factor. Unlike in vivo footprinting, the CHIP assay requires prior knowledge of the type of proteins expected to bind to a given DNA sequence because it relies on a protein-specific antibody to pull out or enrich for the DNA binding sequence of that protein.

The in vitro binding affinity assays carried out in this study illustrate the important contribution of both Class I and Class II AREs to AR DBD cooperative DNA binding to the proximal probasin promoter. The in vivo transient transfection studies carried out in LNCaP and PC3 cells in this study also point to the important contribution of Class I and II AREs in AR-specific synergistic transcription. Similar contributions to cooperative binding and AR-specific synergistic transcription by Class I and II AREs located within the human PSA enhancer have also been previously demonstrated using full length AR (Huang et al, 1999). Comparing binding affinity results with the in vivo transcription results indicates a strong association between AR-specific cooperative binding to the proximal probasin promoter and the AR-specific synergistic transcription that results from these cooperatively interacting DNA binding elements. The next step in probing this association between cooperative AR-specific binding and AR-specific synergistic transcription is to mutate the individual AREs found with the proximal probasin promoter and examine the effects of mutation on both cooperative binding and synergistic transcription.

Mutational analysis of the Class II ARE, G-1, was accomplished in this study but only so far as to examine the change in MeP pattern of the G-1 half sites upon AR DBD binding. The binding affinity and transient transfection assays suggest that
altering Class II sites to resemble Class I sites likely will have an affect on observed cooperative binding and synergistic transcription. This is apparent since the DNA fragment that possessed the highest cooperativity was a unique combination of two Class II sites flanking a Class I site; G1-ARE2-G2. Furthermore, the DNA fragment that displayed the highest synergistic transcription was the unique combination of ARE1-G1-ARE2-G2. Generating the larger probasin promoter fragments that contain site directed mutations such as the spacer mutant (ATT>CGG), the +7 purine-pyrimidine mutant (A195C) and the -13 thymidine to guanine mutant (T215G), will provide interesting data on their effects on cooperative binding in the presence of the other unaltered probasin AREs and their effects on synergistic transcription.

The hypersensitivity to DMS methylation upon AR DBD binding to Class II AREs is indicative of structural DNA distortion, however, it would be useful to probe the relative effects of Class I and Class II ARE nucleotide sequences on the DNA-bound conformation of the AR DBD homodimer. This could be accomplished by applying the trypsin/chymotrypsin protein digest assay or the differential antibody-protein interaction assay perfected in Dr. Ann Nardulli's lab, which applied these techniques to the DNA-bound ER (Wood et al, 1998). If binding to a Class I ARE resulted in a different AR DBD conformation compared to the conformation of the AR DBD bound to a Class II ARE then one would expect a variation in trypsin/chymotrypsin digest pattern or antibody-interaction. This variation in AR DBD conformation would be dependent on the nucleotide sequence of the ARE.

Examining the allostery between the AR and DNA from the perspective of the protein conformation change would support the thesis that DNA sequence dictates specific protein conformation and therefore function. An even more conclusive albeit labour-intensive approach would be to generate separate X-ray crystal structures of the AR DBD bound to a Class I ARE and to a Class II ARE. Such structures would reveal not only the weak interactions that are made between the AR DBD and a specific class of ARE, but also the specific AR DBD-DNA conformation that results from AR DBD binding to a specific ARE class type. From the resultant ARE class-specific conformation structure, one may even be able to deduce function.
While the MeP assay was able to generate a DMS protection and hypersensitive pattern visible on an autoradiograph film, exact orientation of AR DBD homodimers was not always resolvable using this assay. Recent published work has suggested that the AR was capable of binding DNA as a direct repeat homodimer (Zhou et al, 1997; Claessens et al, 2001). To obtain such an orientation and conformation, the AR DBD would have to homodimerize through protein-protein interactions that involve amino acid residues outside the dimerizing D Box. Such dimerizing ability has been observed for the thyroid hormone nuclear receptor but not for members of the steroid receptor subfamily (Umesono et al, 1991). Within the steroid receptor subfamily, dogma states that the head-to-head steroid receptor homodimer is the primary mode of occupying DNA binding sites, with salt bridges linking monomers through their D Boxes. For the most part, the data in this study support the convention of steroid receptor homodimers occupying DNA binding sites that contain two half sites in inverse orientation separated by three base pairs. However, AR DBD monomer and direct repeat binding could not be ruled out within some MeP patterns generated for various AR-regulated gene promoter and enhancer regions.

Orientation and dimerization of the AR DBD bound to DNA could be resolved using X-ray crystallography. However, the large nucleotide variation observed in AREs found within AR-regulated gene regions would prohibit the generation of an AR DBD-ARE X-ray crystal structure for every permutation of nucleotide sequence. Alternatively, a modification of the fluorescent resonance energy transfer (FRET) assay could be used to determine the orientation of the AR DBD bound to a specific ARE (Ramirez-Carrozzi and Kerppola, 2001). The AR DBD and the DNA strand containing the specific ARE could be tagged at one end using two different fluorescent molecules. After the AR DBD-DNA binding reaction has been brought to equilibrium, the change (or lack of change) in fluorescent signal can be measured. A change in wavelength would occur when the two fluorescent molecules are in close proximity and therefore interacting. Close proximity between fluorescent molecules would mean a specific orientation, such as parallel AR DBD-DNA orientation (depending on the fluorescent molecule location). If there were no
change in wavelength one could assume that the fluorescent molecules are distant from one another and therefore the AR DBD-DNA orientation would be antiparallel, again depending on the location of the fluorescent molecule tags. Using fluorescent tagging in this way requires the FluorImagerFSI fluorescence scanner (Molecular Dynamics) or something similar. This method has been successfully applied to determine the DNA binding orientation of Fos-Jun heterodimers (Ramirez-Carrozzi and Kerppola, 2001).

4.3 Conclusion

The AR and its cognate ligand, androgen, are intricately involved in the growth, homeostasis, and development of normal prostate epithelial tissue. The complex control that androgens and the AR have over cell proliferation is also critical to the growth and proliferation of malignant prostate tissue. Understanding exactly how androgen action works to control proliferative molecular responses in prostate epithelium and other androgen-regulated tissues is essential to understanding how prostate cancer initiates, progresses and, most importantly, how it may be controlled or cured.

This study aimed to gain insight into how the AR specifically regulates target genes despite apparently recognizing the same DNA binding site as other steroid receptors. From the results of this study it is apparent that the architectural context of the promoter and enhancer regions of AR-regulated genes significantly contribute to AR-specific recognition and binding to DNA regulatory regions and to AR-specific transcriptional regulation. This architectural context refers to the nucleotide sequence of both the primary DNA binding element and the flanking sequence along with the relative spacing and orientation of AREs within the DNA regulatory region. Nucleotide sequence and DNA binding site orientation can dictate function to a bound transcription factor and the subsequent allosteric exchange between protein and DNA will permit the specific cooperative recruitment of other transcription factors and coregulators. The resultant protein-DNA enhanceosome then is capable of recruiting the transcriptional machinery and this interaction between critical DNA-
bound proteins and the transcriptional apparatus culminates in synergistic transcription.

Understanding how diseases like prostate cancer are able to harness molecular processes like androgen action will lead to novel therapies and prevention strategies that could either manage the disease or perhaps cure. Conceivably, basic research into the architectural components of gene regulatory regions that direct an AR-specific and tissue-specific response could result in complex and intricately engineered gene driven therapies that target malignant prostate tissue. At the very least, understanding how androgens and the AR control gene expression in prostate epithelial tissue will contribute to knowledge of molecular mechanisms that factor in prostate cancer onset and progression.
Bibliography


Holmbeck S, Dyson H and Wright P. 1998. DNA-induced conformational changes are the basis for cooperative dimerization by the DNA binding domain of the retinoid X receptor. J Mol Biol; 284(3):533-9


