Investigating Molecular Interactions of Thyroid Hormone and Androgen Receptors

Yi Zhao

Student-ID: 20340983

Thesis submitted to the University of Nottingham for the degree of Master by Research

September 2022

Abstract

Nuclear receptors (NRs) are a large family of ligand regulated transcription factors that play important roles in mammalian physiology. Many NRs function as heterodimer complexes, allowing crosstalk between different ligand signaling pathways. Malfunction of NRs is important in cancer and other diseases. This study focused on a potential interaction between ligand binding domains (LBDs) of Androgen Receptor (AR; residues 647-920) and Thyroid receptor (TRβ residues 211-461). The androgen receptor signaling pathway plays an important role in both normal prostate development and prostate cancer progression. Recent studies indicate that expression levels of the thyroid hormone-binding protein μ-crystallin (CRYM) tumors influences tumor progression. This and other evidence suggest a possible interaction between AR and TR signaling, possibly through heterodimeric complexes.

As crystal structures were available for monomeric LBDs of both AR and TRβ, and a homodimer of the AR, molecular docking experiments were performed using multiple protein docking software (HDOCK, HADDOCK, GRAMM-X, ZDOCK). Cluster analysis then was performed to identify the most stable predicted conformers and to model residues in TRβ predicted to participate in binding to AR. The AR:TR interface was predicted to involve residues, F755, W752, Y764, P767 and V770 in AR, which form the AR homodimer interface in the homodimer crystal structure. For TRβ, the binding interface in the predicted model was

mainly located in helix 9 and helix 10, which is involved in retinoid X receptor (RXR) binding in TR/RXR heterodimers. The major contributors for heterodimerization with AR were residues, T756, V758, P802, S759, in AR and residues, S207, P384, P419, M423, M430, in TRβ.

Yeast two hybrid assays successfully detected interaction of AR and TR LBDs, but only in the presence of agonist ligands for both receptors. Moreover, mutation of residues, F755, W752, Y764, P767 and V770 in AR, which form the AR homodimer interface, disrupted interactions with TRβ in yeast. To investigate interactions between AR and TRβ LBDs in mammalian cells, the aim was to generate SPLIT yellow fluorescence protein (YFP) expression vectors for TRβ LBD. Difficulties in amplification of TRα1, TRα2 and TRβ1 cDNAs may be linked to GC content in each sequence. The construct TRβ1 LBD - vYFPC156 was generated and sequence validated and was functional in forming ligand dependent interaction with the nuclear receptor interacting domain (NID) of steroid receptor cofactor (SRC1) (construct SRC1 NID - vYFPN155) when co-transfected with in HEK293 cells. However, preliminary attempts to detect AR-TR LBD heterodimers using this assay were not successful and needs to be confirmed in future studies.

Acknowledgements

I've had some incredible adventures over the year while working on my project in Nottingham. Firstly, I would like to thank my supervisor, Prof. David Heery, for allowing me to work in his laboratory and for his invaluable advice. I'd also like to thank my supervisor, Dr. Hilary Collins, for her help and patience throughout the year; Dr. Christopher Roberts, for teaching me the experiments; Ryan Morris, for unending help in experiments; and Luc Costello Heaven, for helping me to settle in the life here. In addition, I am thankful to all members of the gene regulatory group for their guidance and support throughout my time here. Finally, I want to thank my parents for their understanding and support to finish my program. Finally, thanks to all the friends in the Nottingham who helped me to overcome difficulties throughout year.

Table of contents

Abstract *2*

Acknowledgements *4*

Table of contents *5*

List of figures *8*

List of tables *12*

List of Abbreviations *13*

Chapter 1 Introduction *15*

1.1 Nuclear Receptors- A diverse superfamily of gene regulators **15**

1.3 Androgen Receptor (AR) **27**

1.2 Prostate cancer progression and treatment **30**

1.4 Thyroid Hormone Receptors and their potential role in PCa **39**

1.5 Thyroid hormones **50**

1.6 Crosstalk of thyroid and androgen signaling: Potential for TR and AR heterodimerization? **52**

1.7 Project aim and objectives **53**

Chapter 2 Materials and Method *55*

2.1 Sources of materials **55**

2.1.1 Bacterial culture reagents 55

2.1.2 Tissue culture reagents 55

2.1.3 Yeast reagents 56

2.1.4 Nuclear hormone receptor ligands 56

2.1.5 Molecular biology reagents 56

2.1.6 Protein chemistry, western blotting reagents and equipment 57

2.1.7 Protein assay reagents 57

2.2 Bacterial preparation and culture **57**

2.2.1 Recipe of solutions and media used 57

2.2.1 LB + Amp plates 58

2.2.2 Preparation of competent *E. coli* 59

2.2.4 Measurement of competence 60

2.2.5 Transformation of competent cells 60

2.2.6 Culture of *E. coli* 61

2.2.7 Long term storage of bacterial 61

2.3 Molecular Biology Techniques **61**

2.3.1 Small-scale purification from bacterial cells 61

2.3.2 Large-scale plasmid purification from bacterial cells 62

2.3.3 Spectrophotometric quantification of DNA 63

2.3.4 Ethanol precipitation 63

2.3.5 Polymerase chain reaction (PCR) 63

2.3.6 Agarose gel electrophoresis 65

2.3.7 Purification of DNA fragments from agarose gels 66

2.3.8 Restriction digest of DNA....66
2.3.9 Ligation of DNA fragments....67
2.3.10 Colony PCR screening....67
2.3.11 Sequencing of plasmid DNA....69
2.3.12 TOPO cloning....69
2.4 Yeast methods 2.4.1 Recipe of solutions and media used for yeast methods....69
2.4.2 Yeast plasmids....70
2.4.3 Chloramphenicol amplification....72
2.4.4 Growth of yeast....72
2.4.5 Lithium acetate transformation of yeast....73
2.4.6 X-Galactosidase Assay....74
2.4.7 Preparation of yeast protein extracts....74
2.4.8 Bradford assay....75
2.4.9 β-Galactosidase Assay....75

2.5 Cell culture....76
2.5.1 Maintenance of cell lines....76
2.5.2 RNA extraction and DNA synthesize....77
2.5.3 Bimolecular fluorescence complementation....77

2.6 Bioinformatics....78
2.6.1 Sequence alignment and structure superposition....78
2.6.2 Protein structure predictions....79
2.6.3 Molecular docking....80
2.6.4 Conformer cluster....82
2.6.5 Virtual site-directed mutagenesis....82
2.6.6 Molecular mechanics with generalized Born and surface area solvation (*MM/GBSA*)....83
2.6.7 Analysis of interaction interface....84

Chapter 3 Evidence for Expression of TRα and TRβ in Human Prostate and Prostate Cancer Cells. 85

Chapter 4 Prediction for AR-TRβ heterodimer structure. 92

4.1 Structure Choice for Modelling....92

4.2 Interaction mode analysis....97

4.3 Summary....104

Chapter 5 Generation of Expression Vectors to Explore Interactions between AR and TRβ using Split YFP system....107

5.1 Cloning Strategy for Human TRα and TRβ LBD Sequences....111

5.2 Optimization of PCR Amplification of TRα and TRβ LBD sequences....114

5.3 Subcloning TR LBDs into the pCR4-Blunt-TOPO Cloning Vector....121

5.4 Ligation of plasmid TRβ1 LBD - vYFPC156....125

5.5 Bimolecular fluorescence complementation....133

5.6 Summary....136

Chapter 6 Yeast two hybrid assays....139

6.1 X gal Reporter assays....142

6.2 Quantitative β-gal assays....146

6.3 Summary....153

7. Concluding Comments.....155

References.....161

List of figures

Figure 1.2 The general structural functional domain of NRs 18

Figure 1.3 A: Cartoon 2D representation of the conserved NR DBD showing the two zinc fingers and positions of the P box, D box and CTE sequences. B. Representation of the folded DBD of a steroid receptor. C: Crystal structure of the ERa DBD dimer bound to a palindromic ERE (PDB:4AA6). Adapted from Weikum et al 20186.....20

Figure 1.4: Schematic showing the canonical fold of the NR LBD, with the central ligand binding pocket and cofactor binding surface (AF2). Adapted from Weikum et al 2018[6] 22

Figure 1.5 Ligand dependent activation of NRs. Adapted from Heery et al 2021[9] 23

Figure 1.6 Crystal structure of LBD dimers A: ERα LBD homodimer (PDB:1ERE); B: GR LBD homodimer (PDB:1M2Z) ; C: LXR-RXR heterodimer (PDB: 1UHL). Adapted from Weikum et al 2018[6] 25

Figure 1.7 Crystal structures showing the dimerization interface of the LBD homodimers of TRβ or AR 26

Figure 1.8 Schematic representation of the AR domain structure (generated by David M Heery) 28

Figure 1.9 The published structure of AR homodimer (PDB: 5JJM). The key residues identified at the dimer interface are shown in red and H refers to α helix. 30

Figure 1.10 The schematic of advanced prostate cancer and CRPC. 5-hydroxy testosterone is the natural engine for AR and KLK 3 gene is used in prostate cancer patients as a marker of disease. 31

Figure 1.11 Patient samples with metastatic PCa tumours in cBioportal 36

Figure 1.12: Lollipop plot showing recurrent AR mutation hotspots in prostate cancer (Image generated by cBioPortal database. 38

Figure 1.13: Schematic representation of the TRβ2 domain structure (generated by David M Heery) 41

Figure 3.1. TRA gene transcripts listed in Gencode as viewed in the Human Genome Browser (GRCh38/hg38). The major isoforms are TRα1 and TRα2. 86

Figure 3.2. TRB gene transcripts listed in Gencode as viewed in the Human Genome Browser (GRCh38/hg38). The major isoforms are TRβ1 and TRβ2 are indicated. 87

Figure 3.3. Expression levels in a range of normal tissues for the genes indicated represented as Normalised transcripts per million (median TPM). Obtained from Gtexportal.org. 87

Figure 3.4 TCGA expression data showing TRB expression in different human tumour types. Data is presented as FKPM (fragments per kilobase of exon per million mapped fragments).....88

Figure 3.5 Expression data for 7 prostate cancer cells in the CCLE dataset showing relative expression of TRA and TRB transcripts. Both genes were diploid in all cases indicating no gene copy number variation. Expression in the metastatic AR+ PCA cells line LNCaP is highlighted, with VCAP (ARV7+) and PC3 (AR-) shown for comparison. 89

Figure 3.6 A) RNA Seq data from LNCAP C4-2 cells showing mapped reads visualised in the integrated genome viewer (IGV). The regions shown are the canonical exon 2 containing the AUG start codon for all TRα isoforms, and the canonical exon 1 unique to the TRβ1 isoform 90

Figure 4.1 The predicted docking structure of (A 2AX9-2AX9 and (B 2NW4-2NW4. The similarity between these predicted AR homodimers and the published AR homodimer, 5JJM, was validated by calculating the RMSD values. 5JJM, 2NW4 and 2AX9 was shown in cyan, yellow and green, respectively. 94

Figure 4.2 The crystal structure chosen for modelling (2NW4). α-helix was represented by H while β-sheet was represented by B.... 错误!未定义书签。

Figure 4.3 (A The sequence of modelled TRβ. (B The predicted structure of modeled TRβ. The conserved residues are shown in blue while the added residues are shown in orange. The green and yellow circles represent residues in the core and permitted areas, respectively. (C The model's and its components' QMEAN scores of Cβ, all-atom, solvation, and torsion. (D Estimating the model's global quality by comparing the normalized QMEAN scores (denoted by a red asterisk) to a non-redundant collection of pdb structures of similar size. Ramachandran plots showing (E General residues, (F Glycine, (G Proline and (H Pre-Proline. 96

Figure 4.4 The predicted structure for TRβ. α-helix was represented by H while β-sheet was represented by B 97

Figure 4.5 A) Kelley penalty plot and B) AR-TR clustering distance matrix. The structure of 3 most populated clusters C)1, D)2, E)8 represented by HADDOCK 5, HADDOCK 7 and ZDOCK 1, respectively, corresponding to AR-TRβ binding. AR and TRβ are colored in green and cyan, respectively. 99

Figure 4.6 Interaction interface analysis for A) AR and B) TRβ by PDBePISA. Inaccessible residues are shown in dark, solvent-accessible residues are shown in light blue and interfacing residues are shown in yellow. H refers to residues, making hydrogen link, ASA refers to Accessible Surface Area (A^2), BSA refers to Buried Surface Area (A^2), ΔiG refers to solvent energy effect (kcal/mol), bars refer to buried area percentage (one bar per 10%). .. 错误!未定义书签。

Figure 4.7 Picture above is the predicted structure of AR-TR LB. The chain for AR is on the left while the chain for TR LBD is on the right. MM/GBSA results

for each residue in predicted heterodimer through hawkdock. A is for AR while B is for TRβ. The helix region for each receptor is shown in yellow and Beta sheets are shown in blue. Both of regions are numbered as above....104

Figure 5.1. The principal of split vYFP system. Potential interaction partners are shown by yellow circles. They are fused to fluorescent protein fragments, which are depicted by green and gray rectangles (full length YFP and splite vYFP). If the proteins interact, they may enhance the association of fluorescent protein fragments and show fluorescent (green rectangle).. 错误! 未定义书签。

Figure 5.2 The diagram of the interactions with construct to exam. The formation of a heterodimers such as AR LBD and TRβ LBD, AR LBD and TRα LBD, TRβ LBD and TRα LBD. As positive controls, some receptors with cofactor motifs, containing the sequence LXXLL, were chosen for interaction with AR LBD, TRβ LBD and TRα LBD. Cofactor motifs could interact strongly with all of nuclear receptor LBD's and should give a positive response in the split vYFP system. 110

Figure 5.3 Schematic representation of TR LBDs fused in frame with the C-terminal portion of the YFP protein (156-239). The amino acid boundaries for each LBD are indicated. The AR LBD YN155 construct which was already available is also shown. Sequences of forward and reverse primers are also shown below the schematic. 113

Figure 5.4 The schematic of the plasmid with C-terminus vYFP (156-239) which is used as the vector for cloning. 113

Figure 5.6 Gradient PCR showing amplification of TRα1, TRβ1 LBD at different annealing temperatures. Following amplification for 30 cycles (final volume 50ul), 5ul was subject to gel electrophoresis. Unincorporated primers are also visible. 115

Figure 5.7 Effect on varying DMSO concentration on yield of the TRα1 (A, TRα2 (B and TRB1 (C LBD PCR products. Note that the amount of template used in this assay was lower than in Figure 5.6. 116

Figure 5.7 Effect on varying Magnesium ions concentration on yield of the TRB1 LBD PCR products. 118

Figure 5.8 Synthesized cDNA was validated by PCR using the TRα1 and TRα2 primers for HEK293 cells and TRβ1 primers for LNCaP C4-2 cells. The aim sizes of TRα1, TRα2 and TRβ1 were 744 bp, 984 bp and 734 bp. 120

Figure 5.9 vector map of TA cloning vector pCR4TOPO.....122

Figure 5.10 Colony PCR showing inserts for the pCR4-Blunt-TOPO TRα1 (A, TRα2 (B and TRβ1 (C LBD clones.....123

Figure 5.11 BLAST search alignment indicating the deletion of an adenosine nucleotide in TRα transcripts in C4-2 cells. 124

Figure 5.12 Restriction digest check for ERa-LBD-YC156 using EcoRI-HF and XhoI. 126

Figure 5.13 Restriction digest for TRβ1-TOPO using EcoRI-HF and XhoI. 127

Figure 5.14 PCR results for TRβ1 LBD using for ligation with vYFPC156.....129

Figure 5.15 The digested TRβ1 LBD and vYFPC156.....129

Figure 5.16 A) 300 ug DNA input for ligation of Ladder and TRβ1 - vYFPC156. Right lane: The same amount of digested TRβ1 and vYFPC156 without T4 ligase. B) 30 ug DNA input for ligation of TRβ1 - vYFPC156. The ligation is processed by mixing the specific insert, vector (with the ratio 3:1), 0.5 ul T4 ligase, 1 ul ligase buffer and water to make a final volume of 10 ul. the mixture was incubated at 90 times (10 °C for 30 seconds and 30 °C for 30 seconds) 130

Figure 5.17 Colony PCR of clones TRβ1 - vYFPC156, using primers CMD forward and TRβ1 reverse. CMD is a part of the vector vYFPC156. The aim size of the amplified sequence is 999 bp.....131

Figure 5.18 The plasmid sequence of constructed TRβ1-vYFPC156. TRβ1 LBD was shown in green while vYFPC156 was shown in cyan.....132

Figure 5.19 The gel of plasmid of AR - vYN, TRβ - vYC, SRC1 NID - vYN and MED1 LXM1 - vYC 134

Figure 5.20 Complementation of SRC1 NID - vYFPN155 and TRβ1 LBD-vYFPC156 in the presence of 100 nM T3. SRC1 NID - vYFPN155 and TRβ1 LBD-vYFPC156 constructs were expressed in HEK293 cells and examined for YFP expression using confocal microscopy. Cells were stained by DRAQ-5. 136

Figure 6.1 The X gal assays result for yeast extract of colonies, co-transfected with LexA- AR LBD and VP16- TRβ LBD D355R, incubated with or without 100 nM mib and T3. The assays of colonies, co-transfected with LexA- AR LBD and VP16- AR LBD, incubated with or without 100 nM mib worked as positive control.....145

Figure 6.2 The X gal assays result for yeast extract of colonies, co-transfected with LexA- AR LBD and VP16- TRβ LBD, incubated with or without 100 nM mib and T3. The assays of colonies, co-transfected with LexA- AR LBD and VP16- AR LBD, incubated with or without 100 nM mib worked as positive control.....146

Figure 6.3 The stand curve for protein concentration. 147

Figure 6.4 β gal assays result for colonies, co-transfected with LexA- AR LBD and VP16- TRβ LBD, incubated with or without 100 nM mib and T3. The assays of colonies, co-transfected with LexA- AR LBD and VP16- AR LBD, incubated with or without 100 nM mib worked as positive control. 149

Figure 6.5 β gal assays results for colonies, co-transfected with LexA- AR LBD and VP16- TRβ LBD, incubated with or without 100 nM mib and T3. error bars showed the maximum activity.....151

Figure 6.6 β gal assays results for colonies, co-transfected with LexA- AR LBD and VP16- TRβ LBD, incubated with or without 100 nM mib and T3, 1 mM NH3. error bars showed the maximum activity. 153

List of tables

Table 1.1 Human NR superfamily 16

Table 2.1 Conditions to be used in standard PCR using Phusion DNA polymerase 64

Table 2.2 Conditions to be used in standard PCR using TAQ DNA polymerase 68

Table 2.3 Plasmids being used in yeast methods 71

Table 2.4 BSA protein standards being used for Bradford assay 75

Table 2.6 Summary of the score function and sampling algorithm of each web used in the project 81

Table 4.1 Information of PDBs representing the AR and TR structure 93

Table 4.2. Cluster component analysis for each AR-TRβ complex cluster. ... 99

Table 4.3 mCSM predicted Gibbs free energy between original and mutated complex. ΔΔG > 0 indicate mutations stabilized the heterodimerization shown in red, where Gibbs free energy ΔΔG < 0 destabilized heterodimerization shown in green. 102

Table 4.4 SDM predicted Gibbs free energy between original and mutated complex. ΔΔG > 0 indicate mutations stabilized the heterodimerization shown in red, where Gibbs free energy ΔΔG < 0 destabilized heterodimerization shown in green. 103

Table 4.5 Residues with the highest contribution to the AR-TRβ binding affinity..... 错误!未定义书签。

Table 5.1 PCR conditions using to amplify TRα1, TRα2 and TRβ1. The conditions, not specified, were followed section 2.3.5.....120

Table 6.1 Plasmids being used in yeast methods 142

Table 6.2 The values of different concentration of BSA.....147

Table 6.3 β gal assays record for colonies, co-transfected with LexA- AR LBD and VP16- TRβ LBD, incubated with or without 100 nM mib and T3. The assays of colonies, co-transfected with LexA- AR LBD and VP16- AR LBD, incubated with or without 100 nM mib worked as positive control. 148

Table 6.4 β gal assays record for colonies, co-transfected with LexA- AR LBD variants and VP16- TRβ LBD, incubated with or without 100 nM mib and T3. 150

Table 6.5 β gal assays record for colonies, co-transfected with LexA- AR LBD variants and VP16- TRβ LBD, incubated with or without 100 nM mib and T3, 1 mM NH3. 152

List of Abbreviations

AAD	Acidic-activation domain
ADT	Androgen deprivation therapy
AF-1	Ligand in-dependent activation function
AF-2	Ligand dependent activation function
AR	Androgen receptor
ARE	Androgen response element
BSA	Bovine serum albumin
CoCoA	Coiled-coil coactivator
CRPC	Castration resistant prostate cancer
CRYM	μ-crystallin
CSM	Cutoff Scanning Matrix
DBD	DNA binding domain
dH_2O	deionized water
DNAPK	DNA-dependent protein kinase
E.coli	Escherichia coli
ECACC	European Collection of Cell Cultures
EGF	Epidermal growth factor
Erk	Extracellular regulatory protein kinase
ESSTs	Employs environment-specific substitution tables
GL3	Gleason pattern 3
GL4	Gleason pattern 4
HAT	Histone acetyltransferase activity
HDAC3	Histone deacetylase 3 activity
HREs	Hormone response elements
hs DNA	Herring sperm DNA
HSP	Heat shock proteins
IGF-1	Islet-like growth factor
LBD	Ligand binding domain
LBP	Ligand binding pocket
mCRPC	metastatic castration resistant prostate cancer
MM/GBSA	Molecular mechanics with generalized Born and surface area solvation
NCoR	Nuclear receptor co-repressor
NEB	New England Biolabs
nmCRPC	non-metastatic castration resistant prostate cancer
NR	Nuclear receptor
NTD	N-terminus domain
ONPG	O-nitrophenyl-β-D-galactopyranoside
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PSA	Prostate-specific antigen

RSA	Relative solvent accessibility
RXR	Retinoic acid x receptor
SDS-PAGE	Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis
SRC	Steroid receptor coactivator
SRC2	Steroid receptor coactivator 2
ssDNA	salmon sperm DNA
T3	3,5,3'-Triiodothyronine
T4	Tetraiodothyronine
TBP	TATA box binding protein
TF	Transcription factors
TFIIB	Transcription factor IIB
TH	Thyroid hormone
TR	Thyroid hormone receptor
TRBP	TR-binding protein
TRE	Thyroid hormone receptor response element
YPD	Yeast extract peptone dextrose
YSM	Yeast synthetic complete media

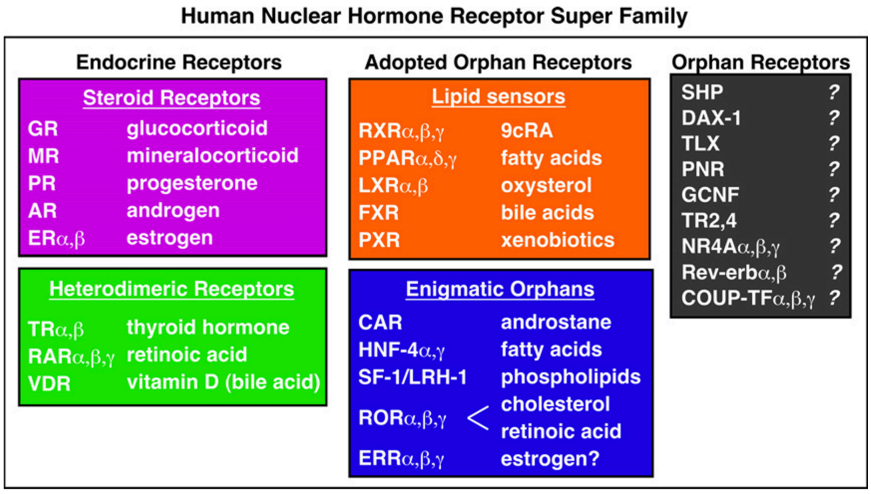
Chapter 1 Introduction

1.1 Nuclear Receptors- A diverse superfamily of gene regulators

Mammalian cells adapt to changes in their internal and external environments by altering their cellular functions or metabolism, usually through changes in gene expression. A key aspect of this ability to sense changes and maintain homeostasis is mediated through specialised receptor proteins. These proteins function in the detection and binding of specific small molecule ligands such as ions, metabolites, hormones, signalling molecules, xenobiotics, or even other proteins. Binding to a specific ligand normally induces a conformational change in the target receptor and as a consequence there is a change in function which leads to signal transduction and alterations in gene expression. While many receptors are present in the cell membrane, allowing them to respond to external stimuli, a variety of intracellular receptors are also known. One of the largest families of intracellular receptors are the nuclear receptors (NR), which in humans[1] comprise 48 distinct receptor genes. NRs are unique multifunctional proteins, as in addition to their receptor function, they also act as DNA-binding transcription factors, and thus transduce their ligand signals directly. The NRs have been subclassified based on ligand specificity, dimerization properties and sequence similarity into 7 subfamilies (NR0-NR6), although these also have common names based on their ligand binding preference e.g. oestrogen receptor, thyroid hormone receptor,

vitamin D receptor etc. Approximately 50% of the NRs have no known ligands, so these are currently referred to as orphan receptors (Table 1.1).

Table 1.1 Human NR superfamily



Most NRs recognize specific sequences containing at least one hexametric motif based on the consensus AGGTCA, organized as direct or inverted repeats (palindromes). NRs bind to these motifs as dimers to regulate a specific set of target genes. The steroid receptors are believed to be mainly present in the cytosol until the bind ligand, when they translocate to the nucleus to bind their target genes. In contrast most RXR heterodimers are thought to be already bound to DNA in the presence or absence of ligands, and ligand binding causes a switch from a gene repressive function to gene activation. This is due to their interaction with cofactors on DNA to help regulate gene transcription.

NRs have been reported to interact with many different nuclear proteins which function as coactivators or corepressors of gene transcription. Many of these cofactors are epigenetic regulators such as histone modification enzymes that add or remove acetyl or methyl groups to histone tails to modify their function. In the absence of ligand, DNA-bound NRs recruit corepressors such as NCOR/SMRT histone deacetylase (HDAC) complexes. in response to ligand binding the corepressors are displaced due to the conformational change in the LBD, and new interaction surfaces are exposed to recruit LXXLL-motif containing coactivators such as the SRC1/ NCOA family and associated histone acetyltransferases (CBP, p300, PCAF, GCN5). Due to this dynamic ability to recruit cofactor complexes in response to signals, NRs control many biological processes including growth, development, metabolism, sex, response to hormones, homeostasis. Similarly they are important in many human diseases such as inflammatory disease, developmental abnormalities, metabolic diseases, diabetes, kidney disease, obesity, liver disease, cancer, leukaemia and lymphoma [2] . Given that for many NRs LBD function is controlled by small molecule agonist and antagonist ligands, this also makes them attractive targets for drug development. Approximately, 15% of all FDA-approved drugs currently in use target NRs.

Figure 1.2 depicts the conventional modular structure shared by all NRs, which consists of five to six homology domains. The N-terminal A/B region has at least one ligand-independent activation function (AF1). The DBD / C region is highly conserved and is linked to a

moderately conserved LBD / E region by a flexible hinge domain / D region. The latter includes the ligand binding pocket as well as the ligand-dependent activation function 2 (AF2). The C-terminal region of several nuclear receptors has a F region with considerable sequence variability, whose function is unknown but may be considered part of the LBD. The DBD and LBD maintain functioning when isolated from the full-length protein, which is valuable for biochemical research.



Figure 1.2 The general domain organisation of NRs including the conserved regions A-F. The DNA Binding Domain (DBD) and Ligand Binding domain are indicated, as well as the regions associated with transactivation functions (AF1 and AF2)

The N-terminal Domain (NTD) of NRs

NR family members share a similar overall structure [3] (Figure 1.2) The N-terminal domain (NTD) shows the lowest level of sequence conservation and its size is also very variable between NRs. The NTD is generally thought to be unstructured unless forming a complex with cofactors. Some NRs have little to no NTD whereas the Androgen Receptor (AR) progesterone receptor (PR) and mineralocorticoid receptor (MR) have large NTDs of >500 amino acids. Early studies showed the many NR NTDs can activate transcription independently of the LBD. In fact, removal of the LBD generates a truncated receptor

with constitutive activation activity, referred to as ligand-independent activation function (AF-1)[4]. The AF1 acts independently as it can drive reporter activation when fused to any heterologous DBD. Thus, the principal function of the NTD appears to be to help to activate gene transcription either through recruitment of cofactors or direct interactions with the RNA polymerase II complex.

The Zinc Finger DNA Binding Domain of NRs

With two exceptions (Dax1 and SHP), all NRs share a strongly conserved DNA Binding Domain (DBD) consisting of two zinc finger domains. These motifs are “C4” type Zn fingers meaning they contain four cysteine residues. These cysteines stabilise the overall fold of the DBD by tetrahedrally coordinating a zinc atom. Each zinc finger structure contains DNA reader α-helices, the first of which can insert into the major groove of the DNA double helix to make contacts with specific bases that are the target sequences recognised by the NR. The sequence of first helix between the 3rd and 4th cysteines (termed the P box) determines the preference of the NR for the target sequence (also termed the hormone response element or HRE). The second zinc finger generates another helix that makes non-specific contacts with the DNA backbone. A sequence between the 1st and 2nd cysteines (the D-box) is involved in dimerization of the DBDs on DNA. These two zinc fingers are contained with a highly conserved 66 amino acid sequence. Finally, a third α-helix is encoded by the C-terminal extension (CTE) sequence extending into the neighbouring hinge region. This is more variable

amongst NRs and can perform several functions e.g. by making additional contacts with DNA backbone or to determine the distance between HRE half sites favoured by heterodimers, and thus determines selective for palindromes or direct repeats[5]. A 6 amino acid sequence termed the T box is present in NRs that form heterodimers with RXR region. Schematic representations of the DBD are shown in Figure 1.3.

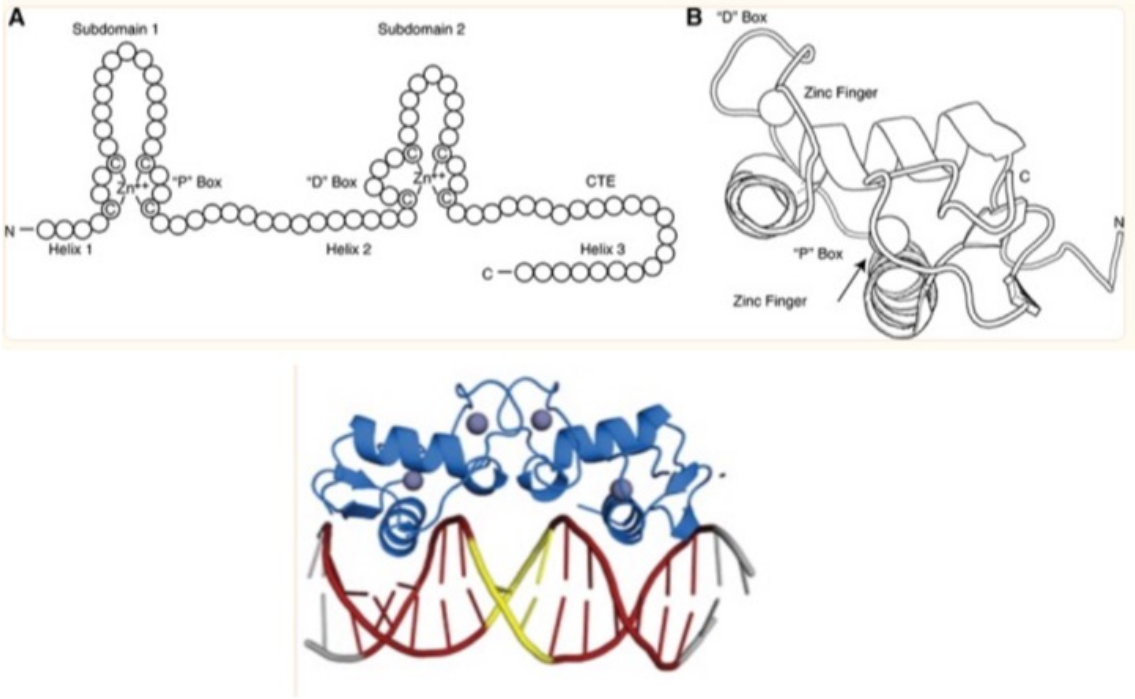


Figure 1.3 A: Cartoon 2D representation of the conserved NR DBD showing the two zinc fingers and positions of the P box, D box and CTE sequences. B. Representation of the folded DBD of a steroid receptor. C: Crystal structure of the ERα DBD dimer bound to a palindromic ERE (PDB:4AA6). Adapted from Weikum et al 2018[6]

The Canonical NR LBD.

The NR LBD undergoes a canonical fold consisting of 11-13 α-helices. These form a central hydrophobic pocket that is specifically adapted to

accommodate binding of the specific small ligand or ligands recognised by the NR (referred to as the 'cognate' ligand). A schematic representation of the NR LBD is shown in Figure 1.4). The LBD can undergo conformational changes that are induced by ligand binding and this may is influenced by nature of the ligand (agonist or antagonist). For many NRs the LBD in the unliganded or antagonist-bound state adopts a conformation that enables corepressor proteins to be selectively bound to the LBD surface. This is due to the presence of hydrophobic channels that can fit the corepressor motif, a leucine / isoleucine rich amphipathic α-helix present in some corepressors. Most NR ligands are small lipophilic molecules that can cross cell and nuclear membranes by passive diffusion, hence the ligand binding pocket tends to be hydrophobic. However, each NR will have evolved to accommodate only its own cognate ligands, ensuring specificity of signalling responses. Thus, the volume of the pocket varies between NRs and has adapted through evolution to accommodate the cognate ligand of each NR. Some orphan NRs, which have no known ligand appear to lack a ligand-binding pocket, thus they may be regulated by other mechanisms. The ligand binding pocket is surrounded by helices 3, 5, 7 and 10 and the binding with ligand induces the conformational changes of nuclear receptors, finally influencing the recruitment of cofactors and possibly dimerization. Thus, as an attractive pharmaceutical target, many drugs have been designed to selectively target this pocket in specific NRs [7, 8].

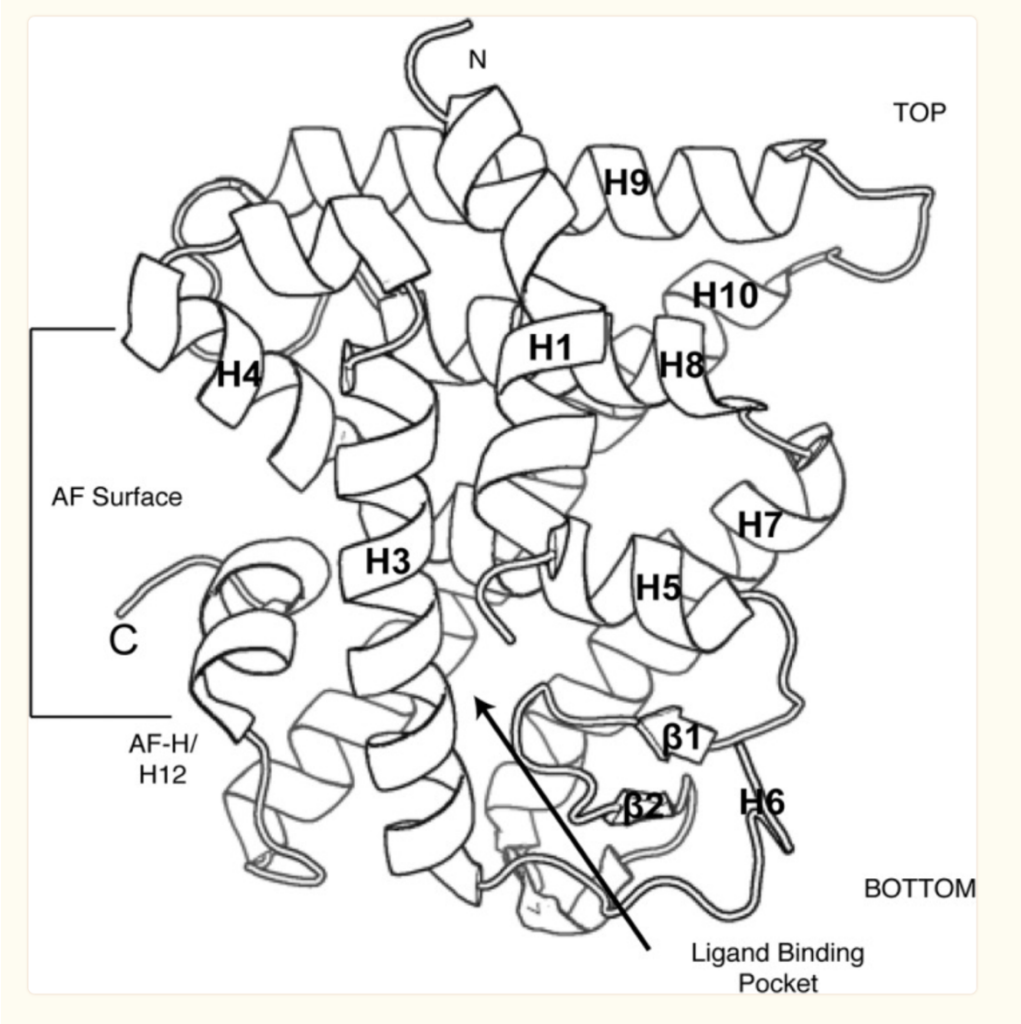


Figure 1.4: Schematic showing the canonical fold of the NR LBD, with the central ligand binding pocket and cofactor binding surface (AF2). Adapted from Weikum et al 2018[6]

In response to binding of agonist ligands, movement of residues within the ligand binding pocket induces a change in the conformation of the LBD surface so that binding sites are exposed that enable recruitment of coactivator proteins containing LXXLL type sequences. This involves the displacement of H12 also referred to as the activation helix or AF2 helix, which in some NRs normally blocks the coactivator binding site. The assembly of cofactor complexes on the gene promoter are thus stabilised by the interaction of the cofactor with both

LBDs in the dimer. Indeed, many coactivators such as Steroid Receptor Activator (SRC1,2,3) and Thyroid Receptor Associated Protein (TRAP220; a mediator complex component) contain at least two LXXLL motifs enabling contacts to both the LBD surfaces. Similarly, repressors such as SMRT and NCOR contain multiple LXXXI/LXXXI/L corepressor motifs to bind both LBDs.

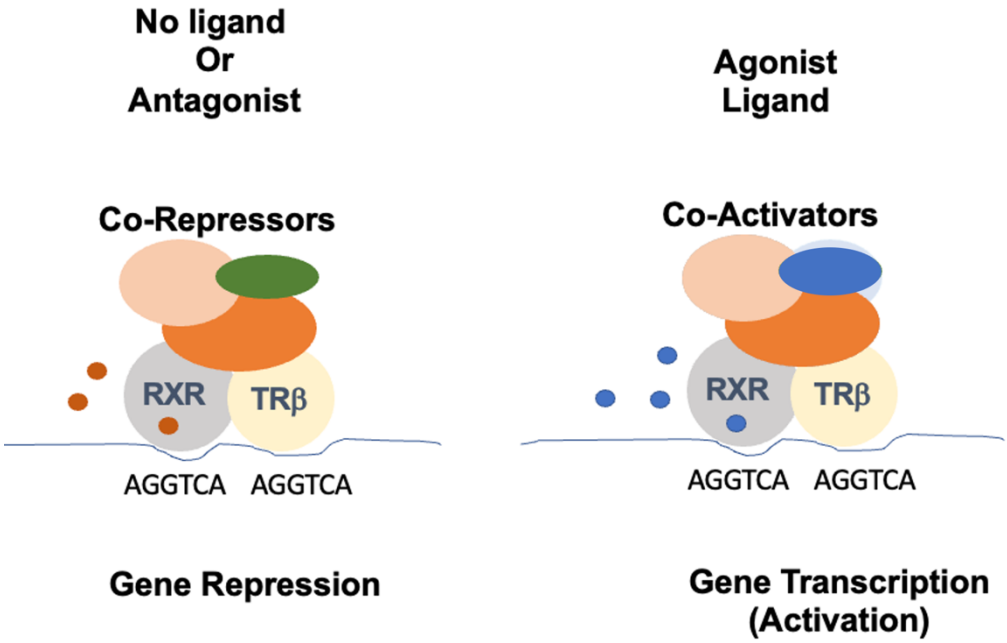


Figure 1.5 Ligand dependent activation of NRs. In the absence of ligand, or when bound to antagonist ligands, (orange circles) NR dimers can form corepressor complexes at target gene promoters. Binding of agonist ligands (blue circles) causes displacement of the corepressors and recruitment of coactivator complexes. Adapted from Heery et al 2021[9]

NR Dimerization

An important function of NRs is the ability to form dimers. Most NRs have been shown to function as dimers while only a few (LRH-1, NGF1-

B, and SF-1) act as monomers. Steroid receptors (AR, ERα, ERβ, PR, GR, MR) are mainly described as homodimers, whereas many other NRs such as thyroid hormone receptors (TRα;TRβ), peroxisome proliferator activated receptors (PPARα,β,γ,δ), Vitamin D (VDR), retinoic acid receptors (RARα,β,γ) Liver X receptors (LXR α,β, function mainly as heterodimers with retinoid X receptors (RXR α,β,γ). These dimerization interfaces can involve contacts between the DBDs to help orient the dimer on DNA, although the main interface appears to involve contacts between the LBDs, and in the cases for which there are crystal structures available these mainly involve hydrophobic contacts between an extended helix (often termed H10 or H10/11). The DBD and LBD are linked by a hinge region which enables both domains to form dimerization with less steric hindrance. In many NRs, the hinge region can also contain nuclear localisation signals (NLS) and can undergo protein-protein interactions with co-factors[10, 11].

As mentioned above, the LBD sequences of NRs show a good degree of conservation and similar overall structures which consists of 12 helices organised in 3 sheets, with 4 beta sheets and loop regions. However, it is not well understood why some NRs have a preference to form homodimers or heterodimers with RXRs or other NRs. The conformation of homodimers and heterodimers reveal high similarity in most cases where structures are available, with the residues of helix 7, 8, 9 and 10 and loops L8 and L9 contributing more than 75% of the dimer interface. The binding with ligands could affect the stability of dimerization and it is also shown that the dimerization interface and

ligand binding pocket are connected[12]. However, the LBD homodimers of GR or AR show an alternative dimer interface the involving hydrophobic contacts between H5, L3 and L4. Schematic representations are depicted in Figure 1.6.

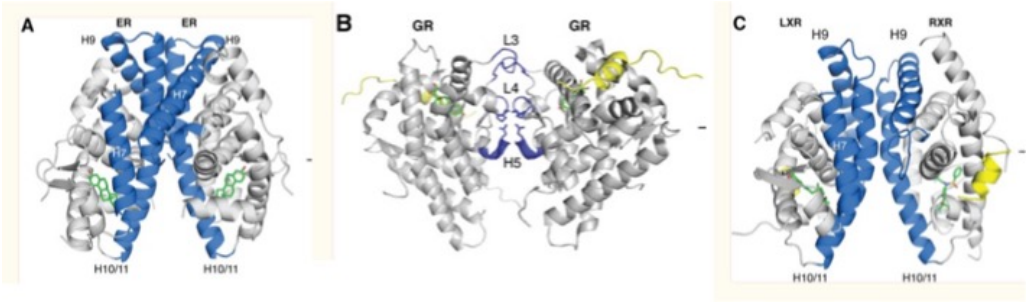


Figure 1.6 Crystal structure of LBD dimers A: ERα LBD homodimer (PDB:1ERE); B: GR LBD homodimer (PDB:1M2Z) ; C: LXR-RXR heterodimer (PDB: 1UHL). Adapted from Weikum et al 2018[6]

While TRs appear to have a preference to form heterodimers with RXRs, they can also function as weak homodimers. Comparison of the crystal structures of LBD homodimers of TRβ to AR or GR homodimers reveals similarities in their arrangement as shown in Figure 1.7. Thus, some NRs like TRs may be able to form different dimeric complexes through distinct dimer interfaces.

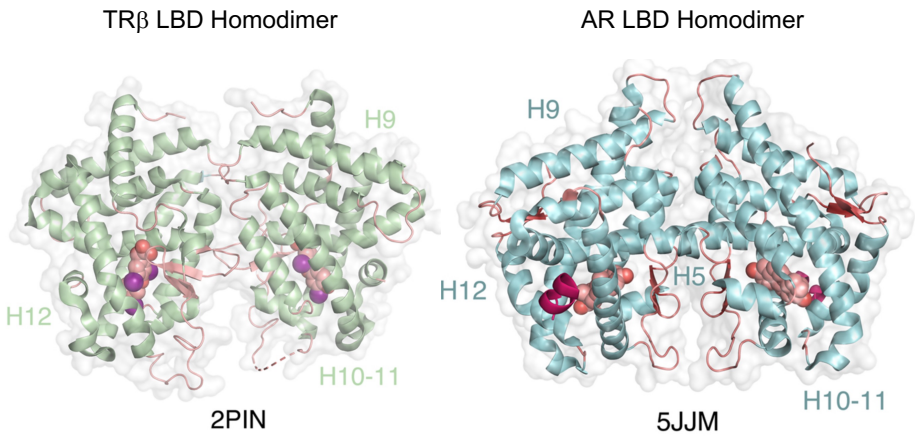


Figure 1.7 Crystal structures showing the dimerization interfaces of the LBD homodimers of TRβ (PDB: 2PIN) or AR (PDB: 5JJM).

While most reviews on NRs focus on homodimer formation by steroid receptors and RXR as the principal heterodimer partner for other NRs, systematic dimerization screens have not been performed. However, numerous studies in the literature have reported the existence of "atypical" NR heterodimers. For example, interactions between steroid receptors such as AR and ERα have been reported. The AR was also reported to heterodimerise with the orphan receptor NR2C2/TR4, but only in the presence of DHT. In our group, it was shown that the orphan receptor NR2E3 or PNR, which is expressed exclusively in the human photoreceptor cells, forms heterodimers with the PPARγ receptor. PNR LBD also showed a ligand-dependent interaction with TRβ. Similarly, interactions of AR with other NR LBDs were identified in yeast two hybrid screens. These observations suggest that other physiologically important interactions between NRs may exist that impact on their functions *in vivo*.

NRs regulate transcription by recruiting coactivators or corepressors [13]. Changes in conformation of the LBD in response to binding agonist or antagonist ligands strongly regulates the formation of these complexes. The LBD binding surface recognised by cofactors is usually provided by residues from helix 12[10, 14-16] and helix 5. A conserved sequence in cofactors is the motif LXXLL[17]. When not bound to ligand, or complexed with antagonist, NRs recruit corepressors, whereas NR binding to agonist ligands alters the LBD surface allowing it to recruit coactivators. The functions of cofactor complexes are discussed in more detail below.

1.3 Androgen Receptor (AR)

The androgen receptor (AR) (also known as NR3C4) belongs to the steroid receptor family [18] and signalling by this receptor plays an important role in both male and female development. The AR mediates the effects of testosterone, dihydrotestosterone, and other androgenic hormones on the body, including development of the skeletal, immune, cardiovascular, neural systems. In males AR drives the normal development of male sex characteristics including normal prostate development and prostate cancer progression. The NTD of AR comprises the transcriptional activation domain AF-1. Unusually for NRs, AF-1 is considered to be the major activation domain of AR, and efforts to map the sequence of AF1 responsible for transcriptional activity identified two transcriptional activation units (tau-1 and tau-5)[19,

[20]. The AR NTD is one of the largest within the NR family and also contains low complexity regions such as a polyglutamine (polyQ) repeats (i.e. CAG repeats) and polyglycine (polyG) or (GGC) repeats. While the wildtype AR allele has 22 polyQ repeats this is variable from 12-30 repeats among men. Interestingly, lower glutamine repeat numbers are associated with higher transcriptional activity of AR[21]. However, greater expansions of the polyQ region are associated with a condition known as spinal and bulbar muscular atrophy (SBMA) or Kennedy's disease, which can lead to the development of female sexual characteristics in genetically male individuals. The AR zinc finger domain DBD has a preference for AGAACA motifs arranged as a palindrome with 3 base pair spacing and the 2 DBDs in the dimer show strong cooperativity in bind to DNA. ChIP-Seq studies have identified considerable diversity in the array of sequences bound by AR and diverse genomic binding sites in different tissues. This may reflect the role of other transcription factors such as FoxA1 which act as pioneer factors to guide AR in the selection specific binding sites in the genome.

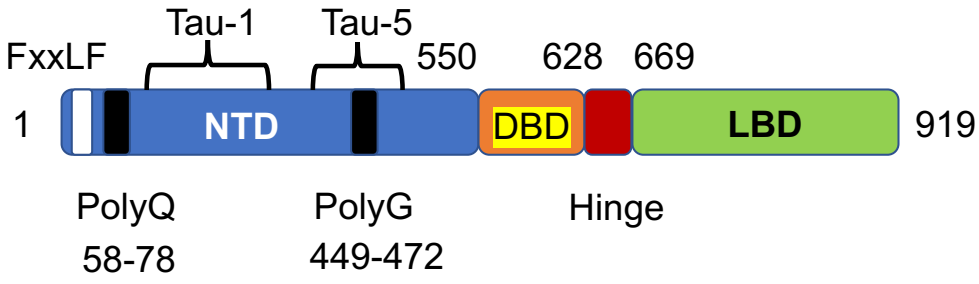


Figure 1.8 Schematic representation of the AR domain structure showing the NTD, DBD and LBD. The NTD regions containing the Tau-1 and Tau-5 activation domains, polyglutamine (PolyQ) and polyglycine rich (PolyG) and a AR-specific FXXLF motif are indicated (provided by David M Heery)

Adjacent to the DBD is the AR hinge region, which contains a bidirectional nuclear localization signal and mediates the transport of AR homodimers from the cytosol to the nucleus. Also, the hinge region is important in AR post-transcriptional regulation, including methylation, ubiquitination, and acetylation of lysines in this region [22, 23].

An unusual aspect of AR function is the direct interaction of the NTD with the LBD, which has been shown to be mediated by a motif FXXLL similar to that found in cofactors. This is also the motif (F/YXXLF/Y) for which AR has a preference in cofactors like ARA70/ NCOA4, which show a preference to bind AR over other NRs, While in isolation the AR LBD displays rather weak ability to activate reporters in response to ligand, it has been shown that in mouse models loss of dimerization or NTD-LBD interactions result in transcriptionally inactive receptors[24]. Thus, the LBD is important for AR function and is a major target for development of drugs to treat prostate cancer [25, 26].

Binding with agonist androgens induces a conformational change of AR which helps AR to form homodimers. Three α-helical structures help to recognize and bind hormone response elements to regulate the transcription of target gene [27]. Like other nuclear receptors, at inactive stage, AR interacts with heat shock proteins (HSP) such as HSP90, HSP70, etc., forming a complex to prevent it from binding to DNA. Previous work revealed the structure of AR LBD homodimer, showing a different binding mode of nuclear receptors. As shown in Figure 1.9, the major binding area was located in the end of α helix 4, α helix 5 and β sheet 1. And several helixes of AR LBD were closely contacted with

each other. While most nuclear receptors form dimers by the longest helix (H10/11).

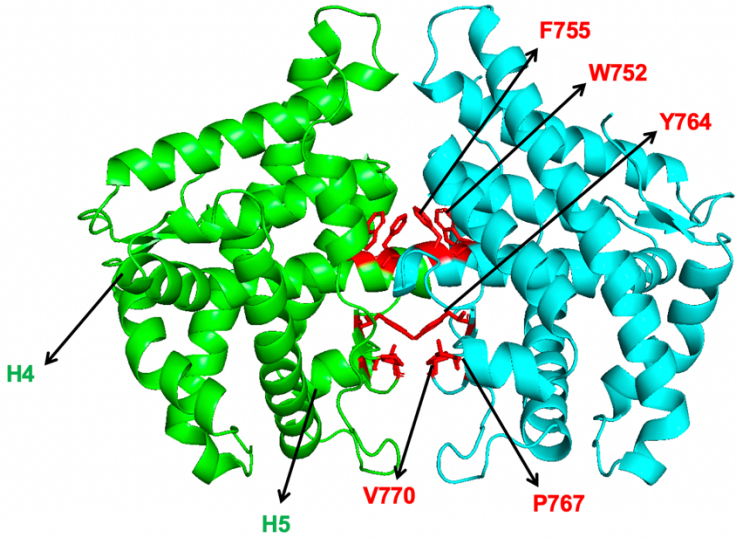


Figure 1.9 The published structure of AR homodimer (PDB: 5JJM). The key residues identified at the dimer interface are shown in red and H refers to α helix.

Unpublished data from our in our lab using yeast two hybrid assays has confirmed that mutation of the residues identified in the AR LBD homodimer interface (F755, W752, Y764, P767 and V770) but not other residues in H11, can result in loss of dimer interaction.

1.2 Prostate cancer progression and treatment

Prostate cancer is the second most common cancer in men and the fifth leading cause of cancer deaths worldwide. Cancers originating in the prostate are often initially dependent on androgens to drive their growth. Thus, treatment approaches may include the use of antihormone

therapies to block the tumor growth and slow the advance to more serious metastatic disease. Metastatic prostatic cancer or castration resistant prostate cancer (CRPC) has poor prognosis and develops resistance to androgen deprivation therapies (ADT). Thus, there is an urgent need for better understanding of the molecular pathways involved in metastatic disease so that novel therapies can be developed.

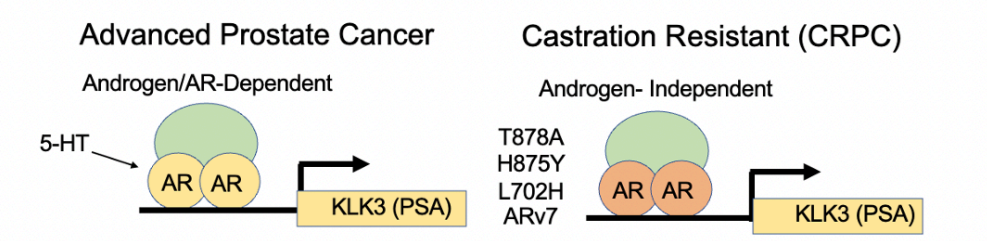


Figure 1.10 Schematic showing regulation of the KLK3 (Prostate Specific Antigen) gene in advanced prostate cancer or CRPC. 5-hydroxy testosterone is the natural agonist for AR and KLK 3 gene is used in prostate cancer patients as a marker of disease. Mutations in the AR LBD associated with hormone resistance in CRPC are also shown.

There are three types of cells in human prostate which include luminal cells, basal cells and neuroendocrine cells[28]. Monolayer columnar epithelial cells are responsible for promotion of protein secretion, expression of unique biomarkers (keratin 8 and 18) and contain high levels of AR. Basal cells express p64 and biomarkers, keratin 5 and 14, and low level of AR. Neuroendocrine cells mainly expressed endocrine related markers, including CHGA and SYP. Being similar to other tumors, tumorigenesis and progression of prostate cancer due to a combination of many factors, including activation of inflammatory

signaling pathways, oxidative stress, DNA damage and gene rearrangement [29-33]; and NKX3.1 down regulation[34], MYC activation[35], TMPRSS-EGR gene fusion[36, 37] and inactivation of tumor suppressor genes[37]. Prostate cancer can be also divided into three main stages [38, 39]. A) Intraepithelial neoplasia which is also called precancerous lesions. It could cause abnormal proliferation of luminal cells and disappearance of basal cells. B) Androgen-dependent adenocarcinoma which can be subdivided into two types, stealth and clinical. Stealth refers to those patients who show positive in lymph node biopsy or pathological analysis of bone puncture specimens but do not have symptoms of prostate related disease. Clinical patients can be verified by both clinical diagnosis and pathological biopsy. At this stage, basal cells completely disappear, luminal cells phenotype is enhanced, and tumor growth is largely dependent on androgens. Thus, the main clinical treatment is androgen deprivation therapy (ADT), supported with selective adjuvant radiotherapy or chemotherapy based on its severity. Most patients show improvement in symptoms after treatment but relapse after 12-18 months, after which tumors progress to the next stage.

C) Androgen-independent adenocarcinoma which is also known as castration prostate cancer (CRPC). At this stage, tumor growth is not entirely dependent on androgens and shows distant metastasis. Tumors mainly express genes of basal cells and stem cells. However, tumor cells still depend on androgens for survival. Clinical treatment for this stage is divided into two methods. One is for patients with non-

metastatic castration resistant prostate cancer (nmCRPC) [40, 41]. The other approach is for metastatic castration resistant prostate cancer (mCRPC). Treatment is based on ADT, combined with AR antagonist, and chemotherapy by docetaxel. Radium 233 is selectively used, according to bone metastases.

An important biomarker in prostate cancer is KLK3, an AR regulated gene encoding prostate-specific antigen (PSA). PSA levels are used to detect abnormal prostate and distinguish between benign disease and carcinoma. While prostate cancer is due to a tumour formation in the male prostate gland it is an androgen driven cancer. Thus, the tumor cells are initially dependent on androgen receptors to drive genes that allow it to spread and proliferate.

Dihydrotestosterone (DHT) is the major natural agonist for AR which controls the interaction between AR and cofactors. The combination of DHT and AR in the cytoplasm will lead to the dissociation of AR and HSP90, through the action of protein kinase A and phosphorylation, and then AR forms homodimers, translocate to the nucleus and interacts with the androgen response element (ARE) in the promoter regions of target genes. At this point, the AR homodimer will recruit co-regulatory proteins, co-activators or co-repressors to bind to the AR complex. Co-activators allow the AR complex to activate or repress the transcription of target genes, regulating proliferation and differentiation of prostate cells. In addition, a ligand-dependent AR non-genomic signaling pathway has been identified which promotes further effects on cells in response to androgens.

This complex drives the expression of many genes involved in the prostate tumour. including the KLK3 biomarker gene. Thus, prostate specific antigen (PSA) is a mainstay of PCa diagnosis and clinical classification and also an important indicator for curative effect and prognosis. Small molecule antagonists of AR or other relevant enzymes in steroidogenesis have been developed to block the growth of prostate tumors[42] which is known as ADT. However, tumours come to be resistant towards these drugs and develop into castration resistant prostate cancer (CRPC) after a few years.

Many CRPC tumours still express PSA, suggesting that the AR signalling pathway is still activated in PCa cells after ADT, thereby maintaining the expression of AR target gene PSA. At the same time, the majority of CRPC patients still express AR and most of the AR proteins are located in the nucleus. Clinical trials confirmed a new drug and a new second-generation antiandrogen that inhibits CYP17 of the cytochrome P450 family. The drug MDV3100 (aka Enzalutamide) is effective in the treatment of CRPC and was approved by the US Food and Drug Administration for clinical use in 2011, confirming that the androgen-AR in CRPC cells is indeed still in a state of continuous activation. It is suggested that the activation of AR signalling pathway depends on the existence of a considerable level of androgens in CRPC patients and the abnormality of AR itself.

It was observed that the testosterone level in the prostate of PCa patients will drop rapidly after chemical castration, but after the occurrence of castration resistance, the testosterone level will return to

the pre-castration level again[43]. The results of this study suggest that in addition to adrenal-derived testosterone, PCa cells themselves may be able to synthesize new testosterone and DHT in patients with CRPC. Montgomery et al[44] found that testosterone and DHT levels were significantly higher in PCa tissue from CRPC patients or castration-resistant mice than in non-castration-resistant mice. In PCa patients or mice, FASN, CYP17A1, HSD3B1, HSD17B3, CYP19A1 and UGT2B17 which mediates androgen metabolism are increased while the expression of SRD5A2 (steroid 5-alpha reductase) enzyme is decreased. HSD3B1 catalyses a rate-limiting step for DHT synthesis. CRPC tumors showed stable mutation in HSD3β1 which caused the accumulation of HSD3β1 and lead to DHT increase finally[45]. These findings suggest that the androgen metabolism pathway is altered in CPRC patients. Changes in the spectrum of enzymes that mediate androgen metabolism in the adrenal gland, especially the upregulation of enzymes, can enhance the synthesis of testosterone from adrenal origin. Then it binds and activates AR, allowing PCa cells to continue to grow under low serum testosterone levels after castration.

Meanwhile, it has been estimated that AR gene amplification may be present in the majority of recurrent CRPC specimens. For example, 61% of a cohort of 429 patients (SU2C/PCF 2019) with metastatic PCa tumours in cBioportal had amplification/overexpression or mutation of AR, whereas 37% of these tumors had lost PTEN gene, as shown in Figure 1.11.

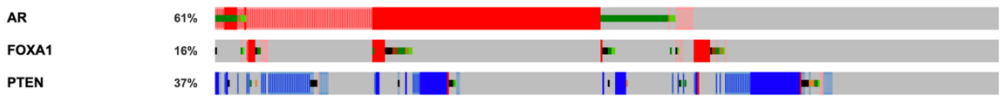


Figure 1.11 Recurrent gene alterations in Patient samples with metastatic PCa tumors in cBioportal. Red or pink indicates sample with gene amplifications or gene overexpression, blue denotes deep deletions and green/ black indicate mutations and gene fusions.

Amplification of AR gene at the DNA level, results in its expression level being significantly up regulated[46], so that the AR signalling pathway can be activated by a very small amount of androgens [47, 48].

Moreover, in recent years, a series of constitutively active AR splice mutants have been discovered. When designing AR shRNAs, Guo et al. [49] found that shRNAs designed for different exons had different knockout effects on AR with relative molecular weights of 110,000 and 80,000, suggesting that full-length AR and small molecular weight AR may originate from different transcripts. Afterwards, they cloned several different AR splice mutants, such as AR3, AR4, AR5, AR6, etc. At the same time, other scientific groups have also reported a series of AR splicing mutants. Dehm et al. [50] reported AREx1 / 2 / 2b and AR1 / 2 / 3 / 2b. AR-V1, AR-V2, AR-V3, AR-V4, AR-V5, AR-V6 and AR-V7 were reported by Hu et al. [51]; ARv567es was cloned in LNCaP cell xenografts by Sun et al. [52]. Watson et al. [53] reported AR-V8, AR-V9 and two mouse AR mutants mV-1 and mV-4. From them, where the most common splice variant is AR-V7, which lacks the LBD and thus can keep AR in a state of continuous activation (via AF-1 in the NTD) to

participate in the occurrence of drug resistance in prostate cancer [50-53]. Moreover, Watson et al. [53] demonstrated that PCa castration treatment induces the production of AR mutants that activate the ARE in the absence of androgen ligand stimulation. Although the expression of AR mutants in PCa is not high, many studies have suggested that it plays an important role in the occurrence and development of CRPC. Knockout of AR3 in CWR-R1 and 22Rvl cells significantly inhibited cell growth under androgen-induced conditions[49]. At the same time, AR3 or AR-V7 expression levels can be used to predict PSA recurrence and tumor specific survival time. In addition, studies have confirmed that AR mutants may affect full-length AR, and its own function may be affected by full-length AR. For example, ARv567es binds and enhances the stability of full-length AR and enhances androgen sensitivity of endogenous full-length AR[52]. Using full-length AR-specific siRNA and enzalutamide, it was confirmed that AR mutants require PCa growth-promotion via full-length AR. The expression of endogenous androgen synthase is increased in tumour tissue. Androgen synthase can promote the resynthesis of androgens and convert weak adrenal androgens into testosterone and dihydrotestosterone, thereby activating AR signal pathway.

Mutations in AR frequently occurred in CRPC (62.17%)[54]. 15% to 30% of CRPC patients show point mutations in the AR gene[55, 56] and these mutations occur mostly in LBD, followed by the NTD[57]. For example, the lollipop plot shown below identifies recurrent AR mutations in the SU2C study mentioned above. The 5 mutational hotspots are in

the LBD (L702H; W742C/L; H875Y; F877L; T878A/S) as shown in Figure 1.12.

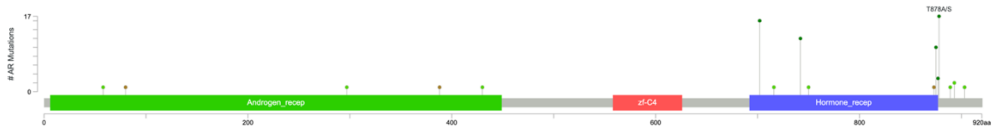


Figure 1.12: Lollipop plot showing recurrent AR substitution mutation hotspots in prostate cancer (Image generated by cBioPortal database).

These mutations can make AR resistant to the antagonist ligands. In some cases, the antagonists actually act as agonists for these mutant receptors. Mutated AR can recruit more AR chaperone proteins to enhance the stability of AR protein and improve the sensitivity to androgen, and sometimes other steroid hormones can also activate AR due to gene mutation [58]. Approximately 30% to 50% of CRPC patients show AR gene amplification which could cause the overexpression of AR[59, 60].

AR-mediated signaling pathways can also be activated in androgen-depleted environments. Islet-like growth factor (IGF-1), epidermal growth factor (EGF), etc. can activate AR mediated transcription of the PSA gene. Studies have shown that Src non-receptor tyrosine kinase-mediated tyrosine phosphorylation at Y534 of AR is an important mechanism for the activation of AR by EGF. In addition, the elevated expression of IL6 in metastatic PCa is closely related to serum PSA level and the occurrence of CRPC. IL6 promotes the growth of LNCaP cells through the paracrine pathway and promotes the growth of DU145 and PC-3 cells through the autocrine pathway. IL-

6 can activate AR via the STAT3 pathway, activate the expression of AR target genes, promote the expression of PSA, and make LNCaP cells resistant to apoptosis.

An alternative approach to CRPC therapy might target AR associated cofactors. Androgen deprivation treatments have been shown to result in increased expression of the acetyltransferases CBP (CREB Binding protein) and p300 in prostate cancer tissues. It has been reported that inhibitors of CBP and p300 can block AR activity, suppress tumor growth and proliferation of CRPC[61].

1.4 Thyroid Hormone Receptors and their potential role in PCa

Thyroid hormone receptors TRα and TRβ, are major regulators of growth, differentiation, metabolism, and homeostasis in many different tissues. TRs are intracellular NRs, and their main action is in the nucleus via their function as ligand regulated transcription factors. TRα is widely expressed, whereas TRβ is more restricted to certain tissues including retina, pituitary, liver & brain. TRs mediate the effects of triiodothyronine (T3) and its precursor tetraiodothyronine (T4) *in vivo* which have important roles in the development of brain, skin, bone and muscle, reproductive and visual system as well as homeostatic control in liver kidney pancreas and intestine. For example, TRs/T3 are required mammalian brain development such as migration and differentiation of nerve cells, synapse formation and myelination. Loss of TRα1, or expression of dominant-negative mutant receptors causes

behavioral changes in adult mice, resulting in severe anxiety and altered hippocampal morphology[62]. Isoforms of thyroid hormone receptors control neural development through different pathways, and the control of TRs is most sensitive in the development of the hearing and visual systems [63].

The THRA gene encoding TRα gene is located at 17q11.2. and encodes two major isoforms generated through differential splicing in the 3' region, resulting in different C-terminal protein sequences. Only TRα 1 has the ability to binding function to T3 and DNA. The TRβ gene is located at 3p24.2 and transcribes three mRNAs, TRβ1, TRβ 2 and TRβ 3 in mammals. Due to alternative promoters and differential splicing in the 5' region, three TRβ proteins with different N-terminal sequences are generated, all of which have full receptor function. As with other NRs, TR can be structurally divided into 6 conserved regions. Regions from the amino terminus to the carboxyl terminus are A~F regions, which constitute 4 functional regions. The A/B region is related to the transcriptional activation regulation function of the receptor, also known as the AF-1 region. The C region forms the DBD, which is located in the center of TR and it is the most conserved region of the receptor. It has the functions of recognizing specific target gene sequences and activating the transcription of target genes, as well as promoting proteasomal degradation and increasing the non-genomic effect of β-actin activation [64, 65]. TR DBD has about 100 amino acid residues and is composed of two zinc finger structures and a carboxyl-terminal extension. Among them, the amino acids between the 3rd and

4th cysteine of the first zinc finger structure constitute a P box (P.box), which is the site for specific recognition of hormone-responsive cis-elements. The D region is the hinge region. The E/F region contains the LBD. The LBD contains the AF-2 region, which is important for ligand-dependent transcriptional activation. The binding of hormones mainly changes the LBD structure of TR, especially the change of helix 12, which leads to the change of the interaction between TR and co-regulators. The AF-2 region is located within the helix 12 of the C-terminus of TR, which interacts with co-activators. Mutational analysis identifies ligand-dependent transcriptional activation of the AF-2 region is especially important.

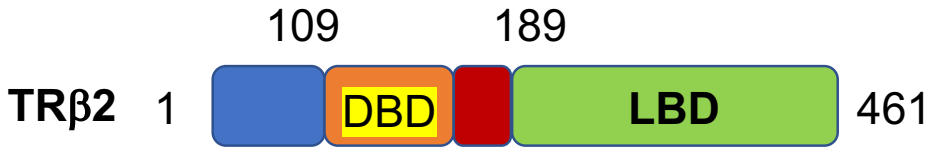


Figure 1.13: Schematic representation of the TRβ2 domain structure (generated by David M Heery)

A comparison of the crystal structures of non-ligand and ligand-bound receptors suggests that the LBD adopts a conformational change upon ligand binding, such as helix 12 being placed over the entrance to the ligand-binding pocket of the LBD, assisting in the formation of a hydrophobic groove on the LBD surface. Methods such as the yeast two-hybrid system and immunoprecipitation have identified a number of potential coactivators such as SRC1 family and TRAP200/ MED1. Recent X-ray crystal structure studies of TR/RXR heterodimers in

complex with T3 indicate that helices 3, 5, 6 and 12 play a role in generating the coactivator surface.

Different subtypes of TR are widely expressed in tissues, but each subtype also has certain tissue specificity. TRα 1 and TRα 2 are expressed in almost all tissues. Among them, TRα 1 is highly expressed in skeletal muscle and brown fat. TRβ 1 is highly expressed in the brain, liver, kidney, and has a circadian rhythm in the liver. TRβ2 is tissue-specifically expressed in the anterior pituitary and hypothalamus, as well as during brain and inner ear development. The expression of TRβ 2 is also associated with retinal development in chickens, mice and in humans [66].

Both confocal microscopy and cellular component analysis showed that TR was present in the cytoplasm and nucleus. Thyroid hormones facilitate the shuttling of TR from the cytoplasm to the nucleus. TRβ1 isoform and other proteins (e.g. mitogen-activated protein MAPK PI3-K p85 regulatory subunit and nuclear receptor coactivator) are found in thyroid hormone-treated cytoplasm and form a complex. The formation of the complex may be beneficial to promote TR entry into the nucleus to initiate gene transcription and cell proliferation. Studies have shown that in the absence of TH, TR not only binds to the thyroid hormone receptor response element (TRE), but also inhibits (or silences) the essential transcription of up-regulated target genes, a feature that is different from steroid hormone receptors. Thyroid hormone enters the nucleus by free diffusion or transported by nuclear membrane carriers and binds to its specific nuclear receptors.

Abnormal changes in thyroid hormone concentrations, energy insufficiency, certain thyroid related diseases, and certain chemical substances have different effects on the activity of thyroid hormone nuclear receptors. After the formation of the ternary complex of thyroid hormone-nuclear receptor-DNA, it initiates gene transcription, enhances mRNA synthesis, increases the translation level of mRNA, increases the corresponding protein synthesis, and then exerts a series of specific physiological and biochemical effects.

Three TR-associated protein complexes have been found to function to repress transcription upon binding to ligand-deficient TRs: Nuclear receptor co-repressor (NCoR), silent regulator of RAR and TR (SMRT), and small single nuclear receptor co-repressor (SUN-CoR). The first two are large nuclear proteins that share 40% of amino acids and play an important role in regulating basic transcriptional repression. SUN-CoR is a 16 kDa nucleoprotein that binds N-CoR and SMRT. N-CoR and SMRT, which preferentially interact with TRs and RAR without ligand binding, are important components in a complex that inhibits basal transcription of target genes with histone deacetylase 3 activity (HDAC3). Studies have shown that N-CoR and SMRT contain a sequence similar to the LXXLL motif (NR box), called the CoRNR box, which is the site of interaction with nuclear receptors. This similarity suggests that the differential recognition of ligand-bound or non-binding nuclear receptors by co-activators or co-repressors in hormone action is due to small differences between the NR box and the CoRNR box. The LXXLL sequence interacts with coactivators and nuclear hormone

receptors. Crystal structure analysis has demonstrated that the hydrophobic α-helix formed by the second LxxLL motif of steroid receptor coactivator 2 (SRC-2) can fit into the hydrophobic groove of TR LBD generated by AF-2 after ligand binding to TR[67]. The interaction sites of different lengths and specific sequences of co-repressors and co-activators, and the structural changes of helix 12 in the AF-2 region caused by co-ligands, determine whether TRs bind to co-repressors or co-activators.

Co-repressors can form a large complex with other repressors, such as Sin3 and histone deacetylase 3. Histone deacetylation of chromatin near TREs of target genes may help chromatin structure maintain a basal transcriptionally off state. Basic repression of transcription requires specific chromatin composition and TR/RXR binding, and the addition of TH activates and causes chromatin reorganization. Therefore, histone deacetylation, histone acetylation after ligand binding, and regulation of chromatin structure and nucleosome position are all important for the transcription of target genes. In addition, DNA methylation may also play an important role in basal repression, since methylated CpG-binding proteins can associate with co-repressor complexes including Sin3 and histone deacetylases. In addition, TRs without ligands can directly interact with TFIIB, an essential transcription factor that promotes gene silencing. TRs with or without TH alter gene transcription with important implications for thyroxine action.

TRs with or without TH alter gene transcription with important implications for thyroxine action. At low hormone levels, such as in hypothyroidism, receptors that do not bind their ligands can suppress expression rather than just exist as inactive receptors. Recent studies have shown that mice knocked out of the TRα and TRβ genes have less clinical symptoms than congenital hypothyroidism. In the absence of the hormone, removal of the receptor may terminate its function as a repressor of the target gene, resulting in a lighter dominance due to "missing" basal transcription of the target gene[68].

Although cofactors that interact with nuclear receptors and enhance transcriptional activation continue to be discovered, ligand-dependent nuclear hormone receptor transcriptional activators are primarily two complexes: the steroid receptor coactivator (SRC) complex and the mediator complex. The SRC family includes SRC-1/2/3. SRCs include multiple nuclear hormone receptor binding sites and contain the iconic LXXLL sequence motif[17, 69]. This sequence is important for the coactivator to bind to the coactivator action site of the TR LBD (helices 3, 5, 6, 12). SRCs also interact with CREB-binding proteins (cyclic AMP response element binding protein, CBP) interacts with the related protein p300. CBP/p300 can bind tightly to sequence-specific transcription factors (TFs), nuclear receptors, TATA box binding protein (TBP) and transcription factor IIB (TFIIB). TFIIB does not bind to DNA, but can form a complex with RNA polymerase II[70]. CBP/p300 also interacts with PCAF (p300/CBP-associated factor), PCAF has histone acetyltransferase activity (HAT), mainly targeting histones H3

and H4, PCAF is part of a complex that includes TBP-related (TAF) which can interact with SRC-1 and SRC-3. Both PCAF and CBP play an important role in the engagement of nuclear receptors to the basal transcription machinery, and in their roles as enzymes of HAT activity that alter chromatin structure [71]. CBP/p300 can act as a co-activator of CREB, p53 and AP-I, so it is also a mediator of many cellular signals. CBP/p300 can act as a co-activator of CREB, p53, AP-I and NF-κB, so it is also a mediator of many cellular signals. SRC can also recruit histone methyltransferases (such as CARM1 and PRMT1) and other downstream coactivators, such as coiled-coil coactivator (CoCoA). CCAR1 may be related to the components of the mediator complex or nuclear receptors, suggesting that CCAR1 may act as a link to coordinate the activities of SRCs and mediator complexes[72]. The mediator complex is a group of related coactivator complexes (including TRAP, DRIP, SMCC, PC2, NAT, CRSP, and ARC) composed of approximately at least 30 subunits that interact with ligand-activated nuclear receptors and also Interacts with other types of transcriptional activators including p53, VP16 and SP1. One of the important subunits is DRIP205/TRAP220, a mediator complex interacts directly with nuclear receptors through DRIP205/TRAP220 when nuclear receptors activate gene transcription. This subunit contains an LXXLL motif similar to SRC. But none of the DRIP/TRAP subunits are members of the SRC family or SRC-related proteins. Additionally, the mediator complex has no intrinsic HAT activity. But some components are mammalian homologues of the yeast regulatory complex, which is

related to RNApol II in yeast. It is suggested that TR recruits the DRIP/TRAP complex, which may help recruit or stabilize the RNA polII core enzyme. Recently, another coactivator, TR-binding protein (TRBP), was found to interact with TR through the LXXLL motif. It can also interact with CBP/p300, as well as with DNA-dependent protein kinase (DNA-PK). DNA-PK subunits Ku70 and Ku86 can also interact with the NCoR/SMRT co-repressor complex, possibly enhancing HDAC activity. However, the precise mechanism of action of TRBP, DNA-PK and other co-activator and co-repressor complexes is unclear[73]. Recent studies have shown that the mammalian homologues of Sw-1/Snf and BRG-1/BRM-1 can interact with nuclear hormone receptors and activate transcription *in vitro*. The DRIP/TRAP complex can recruit RNApolII core enzymes to promote transcription of target genes. Co-immunoprecipitation experiments revealed some proteins bound to hormone response elements (HREs), suggesting a possible mechanism for cyclic recruitment of receptor and coactivator complexes to response elements. Therefore, the recruitment of specific coactivators and their transient recruitment forms may be important determinants of transcriptional specificity and strength[74].

Generally, TRs bind target DNA in the form of homodimers or heterodimers. However, studies have shown that TRβ0 (the homolog of TRβ 1 in birds, reptiles, and amphibians) can form trimers, and trimers have an enhanced ability to recruit co-activators and transcriptional activators in vitro compared to dimers. The trimeric form possesses

isoform specific DNA recognition and transcriptional regulatory features not found in TR dimers[75].

Thyroid hormone response elements (TREs) can be divided into two categories: positive and negative. Ligand-bound TRs activate transcription through positive TREs and inhibit transcription when bound to negative TREs. In contrast to positively regulated target genes, the transcriptional activity of negatively regulated genes can be activated in the absence of T3. The DBD of TR plays an important role in DNA binding to positive target genes, but not in negative T3-regulated target gene[76]. The role of TR in regulating negative target genes has been reported to be independent of its ability to bind DNA. Whether TR-binding DNA is necessary in negative regulation is currently unclear[77]. Since TR is lightly associated to the TREs of these genes, it is not known whether the regulation of TRs is via direct TR binding or the interaction of these proteins with other factors. For example, TRs can inhibit the activity of AP-1, a heterodimeric transcription factor containing Jun and Fos. It is speculated that T3 mediated repression of the prolactin promoter may prevent AP-1 activity. TRs can also interact with several other classes of transcription factors, including NF-1, Oct-1, Sp-1, p53, Pit-1, and CTCF. A study used a gene chip containing a promoter region to identify 81 genes which binds with thyroid hormone receptor from 5000 genes. About half of the genes have binding sites in introns, and 30% are within 1 kb 5' upstream of the transcription start point[78].

In addition to nuclear receptor transcriptional regulation, TRs have some special features. TRs mainly form heterodimers with the retinoic acid x receptor (RXR) and binds to the T3 response element (TRE) of the target gene promoter. This is also known as DR-4, a direct repeat of the sequence AGGTCA separated by 4 nucleotides. Experiments that prevent homo- and/or heterodimer formation by base mutation reveal that TR/RXR is selective for TRE in mammals. The interaction of specific TR oligomers and different elements will play a role in the TRE selectivity of ligand-bound TR[79]. A chip-Seq analysis of genomic TR binding sites in mouse liver confirmed that TRs do occupy binding sites in the absence of ligand; however, exposure to T3 induced additional de novo binding sites and T3-dependent chromatin remodelling as shown through DNAse hypersensitivity mapping. Analysis of the TR binding sites found an enrichment of DR4 motifs, and also FOXA1 motifs suggesting TRs cooperate with other TFs to regulate transcription[80]. Thus, TRs have both T3-dependent and T3-independent binding sites, and the precise mechanisms involved need to be determined.

Furthermore, it has been suggested that TR-RXR heterodimers activate transcription through the TR component of the heterodimer, during which RXR does not bind its ligand nor interact with the TR in the heterodimer. The transcriptional regulation of TR is mainly the switch of co-repressor and co-activator. Although RXR and these co-regulators share a common TR recognition site, RXR does not play the major role in regulating TR transcription. However, other studies suggest that

RXR is not a silent partner in this heterodimer, RXR can bind its ligands and regulate the activity of TR-RXR[81]. One study used a gene chip approach to detect the expression of 54 genes regulated by T3 in Neuro2a cells[82]. After knockdown of endogenous RXR by RNA interference, these genes switched from highly RXR-dependent to RXR-independent upon T3 induction. Thus, while biochemical studies can reveal the dynamic effects of ligand on assembly and disassembly of TR/RXR cofactor complexes, it remains to be determined how these complexes function in cells and tissues.

1.5 Thyroid hormones

Thyroid hormones (TH) play a wide range of roles in ontogeny, maintaining homeostasis, cell proliferation and differentiation. TH is secreted by follicular cells in thyroid. These hormones are compounded by the tyrosine residues in the thyroglobulin synthesised in the thyroid gland and play an important role in the human body[83, 84]. Thyroid hormones play an important role in mammalian brain development such as migration and differentiation of nerve cells, synapse formation and myelination. Thyroid hormones are vital in children's physical development and brain growth. Low and deficient thyroid hormone not only decrease normal epiphysis, growth retardation, but also can lead to less mental activity and even dementia. Inadequate thyroid function in adults will also result in memory loss and slow response. Thyroid hormone accelerates the synthesis and catabolism of the human body,

accelerates the oxidation process of tissue cells, accelerates the heartbeat, increases heat production and heat dissipation. Insufficient thyroid function slows the heartbeat, reduces cardiac output, fears cold, feels dry, and shows no sweat. Thyroid hormones also increase the decomposition and synthesis of proteins, fats and carbohydrates. Mucoid edema, and serous effusion, including pericardial effusion, pleural effusion and joint effusion, would occur when thyroid function declines and is insufficient. These hormones enhance reactivity, including enhanced muscle contraction, and increase blood sugar. Thus, Thyroid hormones are critical for healthy development, growth, neuromodulation and metabolic regulation [85-87]. These effects are more pronounced when thyroid hormone is deficient, such as maternal iodine deficiency or symptoms of untreated congenital hypothyroidism, manifested by severe neurological deficits and growth retardation[83]. Biologically active thyroid hormones include tetraiodothyronine (T4) and 3,5,3'-Triiodothyronine (T3). T3 is generated from T4 and shows 10 to 15 times higher binding affinity for thyroid nuclear receptors[88]. The effect of T3 is mainly exerted by binding with thyroid receptors[88]. Among the isoforms, TRα1 and TRβ are expressed in both normal and cancerous prostate cell lines[9, 89, 90] and shows high binding affinity toward T3. Further study showed that these interactions are not regulated by androgen[91].

LNCaP cells are widely used for studies of prostate cancer which is a representative model for cancer in early stage. LNCaP cell line is more sensitive towards T3 for proliferation, comparing with PC-3 line[90-

[92]. It is also reported that T3 increased the amount of AR in LNCaP cells and emphasized the androgenic proliferative response[9, 91, 92]. But the interaction is blocked under high concentration of agonist for AR[91]. Meanwhile, T3 stimulates the secretion of PSA[91, 92], which is regulated by AR and androgens.

1.6 Crosstalk of thyroid and androgen signaling: Potential for TR and AR heterodimerization?

Thyroid hormones are known to be important for the male reproductive system as revealed by hypothyroid or hyperthyroid states and their impact on puberty in rodent models and in clinical studies. TRs are expressed in normal and cancerous prostate tissues and appear to influence androgen signalling by both indirect and direct mechanisms as reviewed in Aguiano et al Endocrine Connections 2022 [93]. T3 can stimulate proliferation of prostate cancer cell lines such as LNCaP in vitro and enhance KLK3/PSA expression. Indeed, the KLK3 promoter region has been shown to contain functional positive TREs, which are activated by androgens suggesting direct actions between TRs and AR. Other key genes including AR, THRs, DIO2, SRD5A2 also contain adjacent AREs and TREs suggesting coregulation by these hormones. The CRYM gene encodes μ-crystallin, a ketoimine reductase enzyme that is a T3 binding protein expressed in prostate and other tissues. CRYM contributes to the control of metabolism by regulating the bioavailability of thyroid hormones in tissues, thus dampening signalling

via thyroid receptors. CRYM was shown to be an androgen-regulated gene and it was noted that CRYM expression was reduced in hormone refractory PCa[94]. A recent collaborative study involving our group confirmed that low CRYM expression in prostate tumours shows an association with poor prognosis and provided further evidence for an interaction between thyroid and androgen signaling[9]. By binding CRYM, T3 is not available to activate the thyroid receptor, so the TR cannot interact with the AR. CRYM overexpression resulted in reduced expression of androgen regulated genes such as KLK3, and RNA Seq studies in LNCap cells revealed a significant overlap between genes regulated by androgens or CRYM. For example, it was found that about 30% of the genes that are downregulated by CRYM that are up regulated by androgens are down regulated by CRYM. indicating co-regulation by androgen and thyroid signaling[9].

1.7 Project aim and objectives

AR regulates many genes expression in both normal prostate development and prostate cancer progression and plays an important role in the therapy of prostate cancer. Antagonists have been developed to treat prostate cancer by blocking the AR signaling pathway. However, patients came to be resistant because of the mutations in AR after a few years treatment. Studies in the literature indicate that thyroid hormones collaborate with androgen signalling and work on CRYM confirmed another aspect of how this is regulated. An

important question that remains unanswered is whether the coregulation of genes by AR and TRs involves direct heterodimeric interactions of these receptors. Previous unpublished data from our group detected LBD interactions of AR and TR in yeast two hybrid assays. This interaction between AR and TRβ LBDs was only observed in the presence of agonists for both AR (mibolerone) and TR (T3) and both of the ligands increased interactions in dose manner. The antithyroid drug methimazole was found to partially block these interactions in Y2H.

The goal of this project was to explore the interactions between AR and TRβ. Using both *in silico* and *in vitro* approaches. To begin, structural predictions, molecular modeling, MM/GBSA calculation, as well as structural superposition, were utilized to predict the potential structure of AR and TRβ complex. Following *in silico* research, attempts were made to explore interactions in mammalian cells using the split yellow fluorescence protein (YFP) complementation system and in yeast two hybrid assays.

Chapter 2 Materials and Method

2.1 Sources of materials

All general laboratory chemicals were of analytical grade and provided by Fisher Chemicals or Sigma-Aldrich, unless otherwise specified. Phosphate buffered saline (PBS) was made using PBS tablets (1 tablet/100 ml double deionised water) from OXOID Ltd. All solutions were made using double deionised water (ddH_2O), unless RNase free water was needed, in which case deionised water was cleaned by passing it through a Purite select / Neptune water purification system. The Jenway 3510 pH metre was used to test the pH of the solutions. Sterilisation was performed by filtering via a 0.22 micron filter or autoclaving at 120 $^{\circ}$C.

2.1.1 Bacterial culture reagents

OXOID Ltd. provided tryptone and yeast extract for bacterial growth media. Stratagene provided the *Escherichia coli* (*E.coli*) strain DH5α for all molecular cloning applications in this investigation [95].

2.1.2 Tissue culture reagents

HEK293 and LNCaP C4-2 cell lines were purchased from the European Collection of Cell Cultures (ECACC). All tissue culture supplements were purchased from Lonza BioWhittaker® and plasticware was obtained from Helena Biosciences.

2.1.3 Yeast reagents

The yeast strain used was L40 which was a gift from S. Hollenberg. The genotype is *MATa his3Δ200 trp1-901 leu2-3112 ade2 LYS::(4lexAop-HIS3) URA3::(8lexAop-LacZ)GAL4.* Amino acids were obtained from Sigma and used in yeast selective media and agar. Herring sperm DNA (hs DNA) was purchased from Promega. Dropout media contains 10mg Adenine, 50mg L-Arginiene, 80mg L-Aspartic Acid, 20mg L-Histidine, 50mg L-Isoleucine, 50mg L-Lysine, 20mg L-Methionine, 50mg L-Phenylalanine, 100mg L-Threonine, 50mg L-Tyrosine, 140mg L-Valine for 1L of medium, selected for pASV3 mod and pBTM116 mod.

2.1.4 Nuclear hormone receptor ligands

T3 and mibolerone were obtained from Sigma. NH3 was a gift from Prof Lukas Kenner (University of Vienna).

2.1.5 Molecular biology reagents

All enzymes and their buffers were purchased from Roche Applied Science or New England Biolabs (NEB). 1 kb Plus DNA ladder was purchased from Invitrogen Life Sciences. Polymerase enzymes and their respective buffers and supplements were purchased from NEB. Gel extraction kits were acquired from York Biosciences, Nucleospin® plasmid kits from Macherey-Nagel, and maxi-DNA kits was purchased from QIAGEN Ltd.

2.1.6 Protein chemistry, western blotting reagents and equipment

Bio-rad Laboratories provided the SDS-PAGE, western blot transfer apparatuses, as well as the Nitrocellulose transfer membrane. National Diagnostics provided a 30 percent (w/v) acrylamide mix. NEB provided *Precision plus* protein standards. was Lysozyme was purchased from Pharmacia Biotech. Complete Protease Inhibitor cocktail tablets were acquired from Roche. Primary antibody and secondary antibody were purchased from Sigma-Aldrich.

2.1.7 Protein assay reagents

Protein assay reagent was purchased from Bio-rad and bovine serum albumin (BSA) for protein standards was purchased from First Link (UK) Ltd. O-nitrophenyl-β-D-galactopyranoside (ONPG) was purchased from Thermo Scientific.

2.2 Bacterial preparation and culture

2.2.1 Recipe of solutions and media used

CCMB80 Buffer: 0.098% (w/v) KOAc, 1.18% (w/v) $CaCl_2{\cdot}H_2O$, 0.4% $MnCl_2{\cdot}4H_2O$, 0.2% $MgCl_2{\cdot}6H_2O$, 1% (w/v) Glycerol.

SOB media: 2% (w/v) Tryptone, 0.5% (w/v) yeast extract, 0.05% (w/v) NaCl, 0.0186% (w/v) KCl, 0.2033% (w/v) $MgCl_2{\cdot}6H_2O$, 0.2465% (w/v) $MgSO_4$. SOB agar was solidified by the addition of 2% (w/v) bacteriological agar where required.

100 mg/ml Ampicillin (1000X) stock solution: 1g Ampicillin, 5 ml di H_2O, and 5 ml Ethanol

Luria-Bertani (LB) medium: 1% (w/v) tryptone, 1% (w/v) NaCl (pH 7.0) and 0.5% (w/v) yeast extract at pH 7.0. LB agar was solidified by the addition of 2% (w/v) bacteriological agar where required.

10X Tris-borate-EDTA (TBE): 10.8% (w/v) Tris-base, 5.5% (w/v) boric acid, 40 ml 0.5 mM EDTA (pH 8.0)

SOC medium: 0.5% (w/v) yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose.

1000x Ampicillin stock solution: 100 mg/ml in ddH_2O.

Primers:

TRα1 Forward:

AAAAGAATTCCACCATGCCTGAAGAGTGGGATCTGATC

TRα2 Forward:

AAAAGATATCCACCATGCCTGAAGAGTGGGATCTGATC

TRα1 Reverse:

AAAAGCGGCCGCGAGACTTCCTGATCCTCAAAGACCTC

TRα2 Reverse: AAAACTCGAGGGGAGAGGCTGCATTGCCTGC

TRβ1 Forward: AAAAGAATTCCACCATGGACGAGGAATGGGAGCTC

TRβ1 Reverse: AAAACTCGAGATCCTCGAACACTTCCAAGAA

2.2.1 LB + Amp plates

LB agar was molten in microwave under a middle power setting for a few mins. The solution was cooled and added in 0.1% (v/v) 1000X Amp.

20 ml LB agar was poured in each plate. Antibiotics and the date were labeled on the container. The plates were dried in sterile culture hood and then stored at 4°C until required. These plates were stored no more than 3 weeks due to antibiotic degradation.

2.2.2 Preparation of competent *E. coli*

E. coli cells were streaked from a glycerol stock stored at -80 °C on an SOB agar plate and they were put in the 30 °C incubator overnight. A single colony was picked into 2ml of SOB medium and sterile filtered glycerol was added to the culture to make 15% (w/v). Then, it was shake overnight at 30°C. On the other day, 1 ml samples were aliquoted into cryotubes. These seed stocks were stored in - 80°C freezer indefinitely.

1 ml vial of seed stock was inoculated into 250 ml SOB media and grown at 25 °C to an OD_{600} of 0.3. The flask was then incubated on ice for 10 min and consequently harvested by centrifuged at 3000 rpm at 4°C for 10 min in pre-chilled tubes. After centrifuging as above, supernatant was discarded by pouring out slowly and pipetting out the remaining. On ice, each cell pellet was gently resuspended in 10 ml of ice cold CCMB80 buffer separately then combined in one tube. This suspension was incubated on ice for 20 mins. Then cell suspension was centrifuged again at 4 °C. Supernatant was discarded as above. Pellets were resuspended in 10 ml of ice cold CCMB80 buffer. 100 µl of cell suspension was aliquoted into chilled 500 µl tubes. Finally, these

tubes were flash frozen in liquid nitrogen 3 times and stored in -80°C indefinitely.

2.2.4 Measurement of competence

50 µl of DH5α and 10pg pUC19 plasmid were mixed in 500 µl tubes and held on ice for 30 min. It was heat shocked at 42°C for 1 min and held again on ice for 5 min. Then, 200 µl SOC medium was added in and incubated at 37 °C at 220 rpm for 1 hour in a shaking incubator. 30 µl was plated on Amp plates and incubated overnight. Colonies on the plates were counted out. The competence of DH5α is calculated by following formula.

Transformation efficiency (CFU/ug)=$N/M*V_1/V_2$

N:	Number of colonies	M:	Added DNA mass (µg)
V_1	Volume of transformation	V_2	Volume plated

2.2.5 Transformation of competent cells

The prepared competent cells and plasmid were thawed on ice for 10 mins. 50µl of competent cells were pipetted into a 1.5 ml tube with 1 ng of the relevant plasmid and mixed gently for each transformation. The mixture was left on ice for 30 min and then it was heat shocked at 42 °C in a water bath for 2 min. It was placed in ice for another 5 min and then added antibiotic free LB medium, to make a total volume of 500 µl. Later, it was incubated at 37 °C at 220 rpm for 1 hour in a shaking incubator. The cells were pelleted by centrifuge at 13000 rpm for 1 min.

400 µl supernatant was removed and the pellets were resuspended in the remaining liquid. Bacterial suspension was plated on an agar plate containing the relevant antibiotics and incubated at 37 °C overnight.

2.2.6 Culture of *E. coli*

A single colony was grown in 10 ml LB medium with appropriate antibiotics. Media was solidified by addition of 2% bacteriological agar. Cultures were grown by incubation in an orbital shaker at 37°C, 220 rpm, or on plates by incubation at 30 °C overnight.

2.2.7 Long term storage of bacterial

Glycerol stocks containing a 75 percent to 25 percent combination of bacteria in LB medium and glycerol were generated and kept at - 80 °C for long-term storage.

2.3 Molecular Biology Techniques

2.3.1 Small-scale purification from bacterial cells

10 ml *E. coli* cultures were grown at 37 °C at 220 rpm overnight. Cells were pelleted by centrifugation at 4000 rpm at 4 °C for 5 min and the plasmid DNA in cells was purified by the NucleoSpin® Plasmid kit (Machery-Nagel) according to the manufacturer's protocol. First, the pelleted cells were resuspended in 250 µl pre-chilled resuspension buffer A1. 250 µl plasmid lysis buffer A2 was added and then tubes were inverted 6-8 times until the solution getting clear. Later, 300 µl plasmid neutralization buffer A3 was added, and tubes were inverted 6-

8 times until the colour changed. The lysate was clarified by spinning at 13500 rpm for 5 min. The supernatant was transferred to the NucleoSpin column with collection tube and spun down at 13500 rpm for 1 min. The plasmid was washed by 600 µl washing buffer and 500 µl buffer A4 twice and spun down at 13500 rpm for 1 min. Columns were re-insert in 1.5 ml tubes. 50 µl elution buffer was added at the centre of the column and spun down at 13500 rpm for 1 min to elute DNA.

2.3.2 Large-scale plasmid purification from bacterial cells

QIAGEN plasmid Maxi kit was used to obtain a large-scale preparation of plasmid DNA. Single colonies were inoculated in 1ml LB media with the appropriate antibiotic at 37 °C at 220 rpm for 8 hours and 150 µl of it were incubated in 250 ml LB medium with the appropriate antibiotic at 37 °C at 220 rpm overnight. The bacterial cells were harvested by centrifugation at 4000 rpm at 4 °C for 30 mins. The pelleted cells were resuspended in 10 ml pre-chilled Buffer P1. Then, 10 ml Buffer P2 was added in the tubes. Tubes were inverted 4-6 times and incubated at room temperature for 5 mins. Later, 10 ml pre-chilled Buffer P3 was added in, mixed immediately and thoroughly by inverting 4-6 times. Tubes were incubated on ice for 20 mins. Solutions were centrifuged at 4000 rpm at 4 °C for 45 mins. The supernatant was centrifuged again at 4000 rpm at 4 °C for 30 mins. At the same time, QIAGEN-tip 500 was equilibrated by applying 10 ml Buffer QBT. The supernatant was then applied to the QIAGEN-tip and the tips were emptied by gravity flow. QIAGEN-tip were washed by 30 ml Buffer QC twice. The DNA was

eluted by 15 ml Buffer QF. Then, DNA was precipitated by adding 10.5 ml iso-propanol and centrifuged at 4000 rpm at 4 °C for 30 mins. The pelleted DNA was washed by 5 ml 70% ethanol and centrifuged at 4000 rpm for 5 mins. The pelleted DNA was air-dry and redissolved in water.

2.3.3 Spectrophotometric quantification of DNA

The quantification and quality of DNA were detected by measuring the absorbance under 260 nm using the Nanodrop® ND-1000 UV-Vis spectrophotometer. The ratio of OD_{260}:OD_{230} and OD_{240}:OD_{260} were used to evaluate the purity of DNA where pure DNA having OD_{260}/OD_{230} between 2.0 and 2.0 and OD_{260}/OD_{280} between 1.7 and 2.0.

2.3.4 Ethanol precipitation

0.1 volumes of 3 M NaOAc (pH 5.2) and 2 volumes of 100% ethanol were used to precipitate DNA. They were mixed thoroughly and incubated at - 80 °C overnight. On the following day, the mixture was centrifuged at 4000 rpm at 4 °C for 30 mins. The pelleted DNA was washed by 75% ethanol and centrifuged at 4000 rpm at 4 °C for 30 mins to remove excess salt ions. The pelleted DNA was then air dried and resuspended in H_2O.

2.3.5 Polymerase chain reaction (PCR)

PCR was used to synthesis DNA fragments from cDNA generated from cells to construct the recombinant plasmids. Phusion® High Fidelity

DNA polymerase (NEB) was used to amplify DNA to be used in cloning. Typically, 50 µl and 20 µl reactions were taken out. The components are shown in Table 2.1. All the reaction components were assembled on ice and quickly transferred to a preheated Applied Biosystems 2720 thermal cycler. The Phusion DNA polymerase was added at last to prevent the degradation of primers due to the 3′→ 5′ exonuclease activity. The template DNA was firstly denatured at 98 °C for 30 seconds and run 25-35 thermal cycles as follows. Template DNA was denatured at 98 °C for 10 seconds. Primers were annealed to the region to be amplified at 55-65 °C for 30 seconds. The DNA chain was extended at 72 °C for 90 seconds. After cycles, the final extension was set to be at 72 °C for 7 mins. And they were held at 10 °C before collection. After completion, 10 µl out of 50 µl PCR reaction were used for agarose gel electrophoresis to check that the produced DNA fragments were the correct size.

Table 2.1 Conditions to be used in standard PCR using Phusion DNA polymerase

Components	20 µl reaction	50 µl reaction	Final concentration
5X Phusion HF	4 µl	10 µl	1X
10 mM dNTPs	0.4 µl	1 µl	200 µM
10 µM Forward Primer	1 µl	2.5 µl	0.5 µM

10 μM Reverse Primer	1 μl	2.5 μl	0.5 μM
Template DNA	25 ng	25 ng	
DMSO	0.6 μl	1.5 μl	3%
Phusion DNA Polymerase	0.2 μl	0.5 μl	1.0 units/50 μl PCR
Nuclease Free Water	To 20 μl	To 50 μl	

2.3.6 Agarose gel electrophoresis

Agarose gels were prepared according to the size of fragments being resolved. Usually, agarose gels are made between 0.7 % and 2 % agarose (w/v). So, 1 gram of agarose and 100 ml of running buffer can prepare a 1 % agarose gel. A low-concentration agarose gel was prepared for larger fragments, while a high-concentration gel was prepared for smaller fragments. To prepare an agarose gel, an appropriate amount of agarose was weighed into a flask and then 1X TBE was added[96]. The mixture was dissolved by middle heat in microwave. 0.5 μg/ml ethidium bromide was added when the mixture cooled down[96]. During the cooling process, the gel tank was prepared. The dissolved agarose gel was poured into the tank, and it was solidified at room temperature and place in the electrophoresis tank. 1X TBE was added to the electrophoresis tank to cover the gel surface. DNA samples were mixed with 6X gel loading buffer. Ladders were made by mixing 1 kb DNA ladder with 6X gel loading buffer. Samples were loaded in the wells and run for 40-60 mins at 100 V. Bands were

then visualised using an ultraviolet transilluminator and BioRad Quality One software. The approximate size (kb) and concentration (ng/µl) of DNA were determined by comparison with standard molecular weight markers.

2.3.7 Purification of DNA fragments from agarose gels

DNA was extracted from the gel using NEB Monarch® DNA Gel Extraction Kit. DNA bands were visualised by an ultraviolet (UV) light and cut out from the gel. 4 volumes of Monarch Gel Dissolving Buffer were added to the tube with the gel slice. The samples were incubated at 55 °C, until the gel slice was dissolved. Dissolved samples were loaded in column with collection tubes and spun at 13000 rpm for 1 min. They were washed by 200 µl Wash Buffer twice and spun at 13000 rpm for 1 min. DNA Elution Buffer were add to the centre of the matrix and spun at 13000 rpm for 1 min to elute the DNA.

2.3.8 Restriction digest of DNA

Plasmid DNA was digested with restriction endonucleases to yield compatible ends for cloning and to confirm the insertion of the gene fragment. The restriction enzymes were selected based on the primer design used in the cloning technique, and digests were performed in the buffers and at the temperature suggested by the manufacturer (NEB). Double digests were performed jointly when the buffers and temperatures were compatible (enzymatic activities were 100%), otherwise they were performed individually.

Digests were typically performed in a 50 µl reaction volume containing 1 reaction buffer, 2 units restriction endonuclease, 1 µg template DNA, and sterile water brought up to volume. Unless otherwise stated by the manufacturer, digests were typically performed at 37 °C for 2 hours.

2.3.9 Ligation of DNA fragments

Cut DNA fragments were annealed by T4 DNA ligase to form recombinant plasmids. Before the ligation, the quantity of the cut vector and insert were checked by agarose gel electrophoresis (Section 2.3.6) and spectrophotometric quantification (section 2.3.3). Typically, 10 µl reaction was set up with 1 µl T4 ligase, 1µl 10X T4 ligase buffer and varied amount of vector and insert. The ratio of vector: insert was set to 1:3 for ligation reaction under the the manufacturer's instructions. Vector only control ligations were set up to represent self-ligation of the vector in the absence of insert DNA. Insert only control ligations were set up to check the ligation ability of the insert. Ligation of ladder was used to check the activity of T4 ligase. The ligated plasmid was presented by agarose gel electrophoresis (Section 2.3.6) and then transformed in DH5α (Section 2.2.5).

2.3.10 Colony PCR screening

After transformation, DH5α were plated on the LB agar plates containing selective antibiotic and incubated at 37 °C overnight. Single colonies were inoculated in 30 µl LB medium containing selective antibiotic and mixed thoroughly. 5 µl of the culture was used as the

template DNA to run a PCR reaction. For the colony PCR, TAQ DNA polymerase was used to screen the plasmid the proper DNA insert. The forward primer was complementary for the vector DNA while the reverse primer was complementary for insert DNA which insured PCR product held the correct sequence of both vector and insert. The PCR conditions were shown in Table 2.2. The template DNA was firstly denatured at 95 °C for 10 mins and run 25-35 thermal cycles as follows. Template DNA was denatured at 95 °C for 30 seconds. Primers were annealed to the region to be amplified at 55 °C for 30 seconds. The DNA chain was extended at 68 °C for 90 seconds. After cycles, the final extension was set to be at 68 °C for 7 mins. And they were held at 10 °C before collection. After completion, 10 µl out of 20 µl PCR reaction were used for agarose gel electrophoresis to check that the produced DNA fragments were in proper size.

Table 2.2 Conditions to be used in standard PCR using TAQ DNA polymerase

Components	20 µl reaction	Final concentration
10X Standard *Taq* Reaction Buffer	2 µl	1X
10 mM dNTPs	0.4 µl	200 µM
10 µM Forward Primer	0.4 µl	0.2 µM
10 µM Reverse Primer	0.4 µl	0.2 µM
Template DNA	variable	<1,000 ng
Taq DNA Polymerase	0.1 µl	1.25 units/50 µl PCR
Nuclease-free water	to 20 µl	

2.3.11 Sequencing of plasmid DNA

5 µl (100 ng/µl) of Plasmid DNA samples with good quality were sent to Gene service Ltd (Source Bioscience Ltd Nottingham, UK), and sequenced using Big Dye v3.1 chemistry and an AB13730xl automated sequencer.

2.3.12 TOPO cloning

2 µl out of 50 µl PCR product was picked to a 500 µl tube, 0.5 µl TOPO vector and 0.5 µl salt solution was added in. This tube was left at room temperature for 15 mins. At the same time, the DH5α cells were thawed on ice for 10 mins. 50 µl of competent cells were pipetted into TOPO vector and mixed gently for each transformation. The mixture was left on ice for 30 min and then heat shocked at 42 °C in water bath for 2 min. It was placed in ice for another 5 min and then added antibiotic free LB medium, to make a total volume of 500 µl. Later, it was incubated at 37 °C at 220 rpm for 1 hour in a shaking incubator. The cells were pelleted by centrifuge at 13000 rpm for 1 min. 400 µl supernatant was removed and the pellets were resuspended in the remaining liquid. Bacterial suspension was plated on an agar plate containing the relevant antibiotics and incubated at 37 °C overnight.

2.4 Yeast methods

2.4.1 Recipe of solutions and media used for yeast methods

Yeast extract peptone dextrose (YPD) media: 1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) glucose. Solidified where required by the

addition of 2% bacteriological agar.

Yeast synthetic complete media (YSM): 0.67% (w/v) yeast nitrogen base without amino acids, 2% glucose (w/v), 0.2% drop-out mix (w/v). Solidified where required by the addition of 2% bacteriological agar.

Drop-out mix: A combination of adenine, lysine, histidine and uracil (Le Douarin et al., 2001).

Z buffer: 60 mM Na_2HPO4, 40 mM NaH_2PO4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercaptothion, filter sterilised.

1 M LiAc: 6.6 % (w/v) LiAc, filter sterilized.

50% PEG (w/v): Dissolve 50 g PEG, 3015-3685 (Sigma no. P-3640) in 30 mL deionized water (dH_2O) and adjust the final volume to 100 mL. Mix well by vertexing. Wait at room temperature until it gets clear. Then, filter sterilized.

Phosphate-buffered saline (PBS) buffer (pH 7.4): 500 µg/mL X-gal, 0.5% (w/v) agarose, and 0.05% (v/v) β-mercaptoethanol

O-nitrophenyl-β-D-glucopyranoside (ONPG) (4 mg/mL): Dissolve in Z buffer.

2.4.2 Yeast plasmids

AR and TRβ were cloned into the vectors containing the DBD of the bacterial repressor protein LexA and the acidic-activation domain (AAD) from the herpes simplex virus protein VP16. LexA-fusion expression plasmids are based on pBTM116mod[97], a derivative of pBTM116[98]. The VP16-fusion expression plasmids are based on pASV3mod[97], a modified version of pASV[97]. These plasmids were from the plasmid

stocks of Dr. Jonathan Whitchurch and Dr. Christopher Roberts as shown in Table 2.3.

Table 2.3 Plasmids being used in yeast methods

Number	Plasmid	Origin
10094	pBTM116mod-AR LBD T851D	Dr. Jonathan Whitchurch
10095	pBTM116mod-AR LBD V770A	Dr. Jonathan Whitchurch
10096	pBTM116mod-AR LBD W752R	Dr. Jonathan Whitchurch
10097	pBTM116mod-AR LBD F755S	Dr. Jonathan Whitchurch
10098	pBTM116mod-AR LBD P767A	Dr. Jonathan Whitchurch
10099	pBTM116mod-AR LBD Y764C	Dr. Jonathan Whitchurch
22	pASV3mod AR LBD	Dr. Christopher Roberts
23	pBTM116mod AR LBD	Dr. Christopher Roberts
29	pBTM116mod TRβ LBD	Dr. Christopher Roberts
30	pBTM116mod TRβ LBD D355R	Dr. Christopher Roberts

2.4.3 Chloramphenicol amplification

Chloramphenicol amplification was used to get a high amount of pASVmod suitable for yeast transformation which showed low yield in miniprep. A single colony from an LB- ampicillin plate with freshly transformed DH5α colonies was inoculated into 20 ml of LB- ampicillin (100 µg/mL) and grown overnight at 37 °C at 220 rpm shaking. The following day, the culture was spun down at 4000 rpm for 10 mins. 20 ml fresh LB media with 100 µg/mL ampicillin and 170 µg/mL chloramphenicol were added in the tube. These cultures were incubated overnight at 37 °C at 220 rpm shaking, then the pellets were spun down at 4000 rpm for 10 mins. The plasmids in pellets were purified by the NucleoSpin® Plasmid kit according to the manufacturer's protocol (section 2.3.1).

2.4.4 Growth of yeast

Single colonies of Saccharomyces cerevisiae were maintained by streaking on the YPG agar plates, where selected plasmids were maintained in synthetic complete medium plate with appropriate supplements. Plates were incubated in 30 °C incubator for 3 days until individual colonies appeared and stored at 4 °C for several weeks. For liquid cultures, a single colony was inoculated in 20-30 µl synthetic complete media with 100 nM ligands in falcon tubes. Tubes were incubated at 30 °C at 220 rpm overnight.

2.4.5 Lithium acetate transformation of yeast

A modified version of the high efficiency lithium acetate method of yeast transformation was used to transform L40, as described[99]. A single colony of S. cerevisiae was inoculated in 5 ml YPD media and grown overnight as described in Section 2.4.4. On the next day, the culture was diluted in 100 ml YPD to an initial OD_{600} of about 0.05 and grown 4-6 hours until OD_{600} reached 0.8-1.0. Meanwhile, 50% PEG 3350, 1M LiAc, 0.1 M LiAc, 1 M DTT and H_2O were put on ice. 2 mg/ml salmon sperm (ssDNA) was denatured at 100 °C for 5 mins and snapped cool on ice. Cells were pelleted by centrifugation at 4000 rpm for 5 min. Cells were resuspended in 10 ml ice-cold water and transferred to 1.5 ml Eppendorf tube. Cells were pelleted by centrifugation at 13000 rpm for 2 min and washed by 0.5 ml ice-cold 0.1 M LiAc. Cells were collected by centrifugation at 13000 rpm for 2 min. Then, pellets were resuspended in 20 µl ice-cold water and stored at - 80 °C.

30 µl 2 mg/ µl ssDNA, 140 µl 50% PEG 3350, 22 µl 1M LiAc, 12 µl 1M DTT and 500 ng plasmid DNA were mixed on ice and vortexed well. 2 µl yeast were then added in. Mixture was mixed again by vertexing gently until homogeneous. Then, it was incubated at 30 °C for 30 mins and heat shocked at 42 °C for 20 mins. Cells were collected by centrifugation at 5500 rpm for 30 seconds and resuspended in H_20. The suspension was plated on YSM plates for two days. Later, 3 colonies of each plate were re-plated onto a new plate.

2.4.6 X-Galactosidase Assay

Colonies were picked into 10 ml YSM and cultured at 30 °C overnight. 500 µl culture was taken out with 500 µl YSM for OD_{600} detection. One absorbance unit of yeast was pelleted by centrifugation at 13000 rpm for 5 min. Pellets were placed 3 times 3 min in liquid nitrogen and 3 min in 37 °C water bath. Pellets were subsequently resuspended in 20 µL water, transferred to a transparent flat bottom 96-well microplate, mixed with 100 µL phosphate-buffered saline (PBS) buffer (pH 7.4), containing 500 µg/mL X-gal, 0.5% (w/v) agarose, and 0.05% (v/v) β-mercaptothion. An hour later, 100 µl 1M Na_2CO_3 were added to stop the reaction.

2.4.7 Preparation of yeast protein extracts

A full loop of yeast was inoculated in 30 ml YSM medium and vortexed to make it an equal suspension. 15 ml of the suspension was decanted into the other labelled tube. Ligands were added to make a final concentration of 10^{-6} M in the + ligand tube and an equal volume of YSM was added in the – ligand tube. Both tubes were covered with foil and incubated at 30 °C with 220 rpm shaking overnight.

On the next day, foil sleeves were removed, and all samples were centrifuged at 4000 rpm for 5 minutes at 4 °C. The pellets were resuspended in 1 ml sterile H_2O and transferred to separate 2 ml microcentrifuge tubes. Tubes were briefly centrifuged for 15 seconds and resuspend in 150 µl Z buffer. 150 µl of glass beads were added to each sample. Samples were lysed in the disruptor using a cycle of 1

minute disrupting, 1 minute on ice, a total of four times. All samples were centrifuged for 30 minutes at 14,000 rpm at 4 °C. Once complete, supernatants (protein extract) were transferred as much as possible to separate 500 µl tubes.

2.4.8 Bradford assay

Bradford assay was taken to detect the concentration of protein extract. The protein concentration of all yeast protein extracts was determined by comparing OD_{595} with the following BSA protein standards as shown in Table 2.4.

Table 2.4 BSA protein standards being used for Bradford assay

Component (µl)	0 mg/ml (Blank)	2 mg/ml	4 mg/ml	6 mg/ml	8 mg/ml	10 mg/ml	Yeast protein extract samples
diH_2O	800	798	796	794	792	790	800
1 mg/ml BSA	0	2	4	6	8	10	0
Z Buffer	2	2	2	2	2	2	0
Yeast protein extract	0	0	0	0	0	0	2
Bradford reagent	200						

2.4.9 β-Galactosidase Assay

20 µl of protein extract were added in 480 µl of Z buffer and transferred to a cuvette. The mixture was briefly vortexed to ensure even distribution of protein.

To begin the β-galactosidase reaction, 100 µl of 4 mg/ml ONPG (solid ONPG dissolved in Z buffer) was added. The time of each addition was recorded. Each cuvette was briefly vortexed upon the addition of

ONPG. Each reaction was waited until a pale-yellow colour begins to appear. 250 µl of 1 M Na_2CO_3 was added to the relevant cuvette and the time of addition was recorded. Each cuvette was briefly vortexed upon the addition of Na_2CO_3. The absorbance of the resulting mixtures was measured at OD_{420} once the reactions are complete. Any cuvettes which had not produced a colour change 30 minutes after ONPG addition had sodium carbonate added and the OD_{420} measured. Specific activity of the extracts was calculated using the formula:

Specific activity (U) = OD_{420} × 0.85/ (0.0045 × Protein × Volume ×Time)

Protein is the protein concentration of the yeast extract in mg/ml. Volume is the extract volume assayed in ml. Time is the time of reaction in minutes.

2.5 Cell culture

2.5.1 Maintenance of cell lines

MCF-7 and HEK293 cells were routinely maintained in Dulbecco's Modified Eagle's Medium (DMEM) containing phenol red and supplemented with 1% (v/v) 0.2 M L-glutamine, 1% Penicillin/Streptomycin antibiotics and 10% (v/v) heat inactivated fetal calf serum (FCS). Cultures were grown in sterile plasticware at 37 °C in a humidified atmosphere containing 5% CO_2. Adherent cells were passaged at 90% confluency. To split cells, DMEM was aspirated. The cells were resuspended and subsequently diluted in fresh DMEM in T75 flask (for maintenance).

2.5.2 RNA extraction and DNA synthesize

Total RNA was extracted using the RNeasy Mini Kit (Qiagen). 10^7 cells were centrifuged at 15000 rpm for 5 mins and the supernatant was discarded. The pelleted cells were resuspended in 350 µl buffer RLT and added 350 µl 70% ethanol. The mixture was transferred to column and spun at 13000 rpm for 15 seconds. Then, 700 µl buffer RW1 were and spun as above. Finally, 500 µl buffer RPE were added and spun at 13000 rpm for 2 mins. The RNA was collected by adding 30 µl RNase free water and stored at -80 °C. Reverse transcription was carried out with 250 ng of total RNA using the QuantiTect Reverse Transcription Kit (Qiagen). 250 ng RNA was mixed with 2 µl gDNA wipeout buffer and 10.9 µl RNase free water and incubated at 42 °C for 2 mins. Then, 1 µl reverse transcriptase, 4 µl RT buffer and 1 µl RT primer mix were added in the mixture. The mixture was incubated at 42 °C for 15 mins and 95 °C for 3 mins. Finally, the cDNA was stored at -20 C.

2.5.3 Bimolecular fluorescence complementation

Bimolecular fluorescence complementation (BiFC) was performed in HEK293 cells grown in DMEM in the presence of the appropriate ligand, antagonist, or vehicle. Adherent cells were planted at a density of 2 X 10^5 cells per well in 8 wells iBidi plates in 300 µl DMEM per well and incubated for 24 hours before transfection. The next day, 3-4 hours before transfection, DMEM was changed with new medium. The calcium phosphate co-precipitation approach was used to transfect 50-80 percent sub confluent cells for transfections. Cells were transfected

with 25-300 ng of -vYN and -vYC plasmids per well. With the inclusion of pcDNA3.1(+) empty vector DNA, the total quantity of DNA per well was standardised. 16 hours after transfection, cells were washed twice in sterile PBS and new medium containing the appropriate ligand or vehicle was added. Cells were fixed in 4 percent paraformaldehyde for 10 minutes after transfection before being washed 4-5 times in PBS and stained with DRAQ5 (1:25000) for 10 minutes to stain the nucleus. Cells were then rinsed 4-5 times in PBS and imaged on a confocal microscope (Zeiss LSM510).

2.6 Bioinformatics

2.6.1 Sequence alignment and structure superposition

In order to analyze the sequence of constructed plasmid, the sequence result was sent to standard nucleotide blast server[100] (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) to query the somewhat similar sequence using standard database, nucleotide collection (nr/nt). Some sequence results which showed less similarity with target genes were sent to blastx server (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) to analyze the similarity in translated sequence.

Structural alignment was performed based on the sequence of the AR (Uniport[101] code: P10802) and TRβ (Uniport code: P10275) in FASTA format using the Discovery Studio (DS) 3.0 package. In brief, the crystal

structures (PDB ID: 2NWS[102] for AR and PDB ID: 1XZX[103] for TRβ) in complexes with co-crystallized ligands were retrieved from the RCSB Protein Data Bank (PDB, http://www.rcsb.org)[104], and aligned to compare the similarity of the protein structures using pymol.

2.6.2 Protein structure predictions

The hinge region between residues 252-263 were absent in all the reported PDB structures of TRβ with co-crystallized agonist. Thus, Swiss-model[105-107] (https://swissmodel.expasy.org/interactive#structure) was used to fix the missing part. FASTA sequence of TRβ was uploaded as the target sequence. The server will search for some highly homologous sequences in the SWISS-MODEL Template Library [108-114]. The selected sequences were then multiple matched to the target sequence motif to get the initial set of template structure equivalence point sets[115-118]. If two different templates meet certain overlapping criteria in space, the relationship between them is the equivalence point. Many equivalence points constituted an equivalence point set. Based on the determined initial set of template-structure equivalence sites, the structures of each template were rotated to make their positions overlap each other as much as possible[119, 120]. After the superposition, the structural conserved regions (SCRs) of homologous proteins and the corresponding framework were obtained. And the highly similar regions on the target sequence were selected and defined as the SCRs of the target protein[121, 122]. Once the template structure was matched, the first

sequence of the obtained homologous SCRs was used to match the target sequence. The model was assessed and sorted basing on QMEAN and QMEANDisCo global model quality[123-129]. Once the modelling process was finished, structure was chosen by the modeling scores and the templates being used.

2.6.3 Molecular docking

The accuracy of protein-protein docking depends on the quality of the protein structures involved and also the sampling algorithm, scoring function and degree of flexibility of web servers being used. Thus, we firstly docked the structure of AR LBD to itself and compared the predicted structure with 5JJM[130], the crystalized AR-AR homodimer. The PDB structures of AR LBD, showing lowest Cα RMSD values between predicted and the crystalized AR LBD homordimer, were chosen for docking with TRβ. To improve accuracy of the model , 4 common web servers were using for the prediction of AR-TR LBD heterodimerization, including HADDOCK (https://wenmr.science.uu.nl/)[131, 132], HDOCK (http://hdock.phys.hust.edu.cn/)[133, 134], GRAMM-X (http://vakser.compbio.ku.edu/resources/gramm/grammx/)[135] and ZDOCK (https://zdock.umassmed.edu/)[136, 137]. The selected AR and TRβ pdb structures were uploaded for docking. And residues, W751, F754, Y763, P766 and V769 in AR, were selected as active residues which are directlyinvolved in the interaction. In HADDOCK, the minimum percentage of relative solvent accessibility (RSA) to consider

a residue as accessible was set to 15%. Center of mass restraints and Surface contact restraints were set off. For FRAMM-X, free docking was selected as docking method. For ZDOCK, we used the version ZDOCK 3.0.2.

The score function and sampling algorithm of each web server were shown in Table 2.5. 10 most potential structures from each web server were selected out according to the docking scores.

Table 2.6 Summary of the score function and sampling algorithm of each web used in the project

Web server	Sampling algorithm	Score function
HADDOCK	Flexible Monte Carlo search	Van der Waals, electrostatics, binding site restriction, and buried surface area
HDOCK	Rigid body Fast Fourier transform (FFT).	Long range shape complementarity and statistical mechanics-based
GRAMM-X	Rigid body FFT	Shape complementarity and hydrophobicity
ZDOCK	Rigid body FFT	Shape complementarity, Van der Waals,

	electrostatics, desolation energy

2.6.4 Conformer cluster

Conformer cluster module in Maestro Schrodinger (Schrödinger Release 2020-2: Maestro, Schrödinger, LLC) was used to quickly remove additional identical structures from the 40 predicted structures, based on energy or geometric considerations. These structures were clustered by comparing RMSD values of all heavy atoms. The optimum clusters was determined by Kelley penalty plots[138].

2.6.5 Virtual site-directed mutagenesis

Site directed mutator (SDM http://marid.bioc.cam.ac.uk/sdm2/)[139-141] and mCSM (http://biosig.unimelb.edu.au/mcsm_ppi2/) server were used to calculate the energy changes of single point mutations on protein-protein interaction between the wild-type and the mutant. SDM calculated the difference in the stability scores (ΔΔS) of the unfolded (ΔS^U) and folded (ΔS^F) states for the wild-type and mutant protein structures, respectively, using employs environment-specific substitution tables (ESSTs). ESSTs were created via an investigation of amino acid substitutions permitted within families of homologous proteins with known 3D structures[142].

The mCSM computer software family estimated the influence of point mutations on protein stability and protein-protein interactions using a

graph-based technique based on Cutoff Scanning Matrix (CSM)[143, 144]. Feature vectors, also known as mCSM signatures, were utilized to represent the structural environment of the wild-type protein. They were defined as inter atomic distance patterns surrounding the altered residue. Relying on the PMapper software, mCSM employs a pharmacophore count vector to represent the rates of eight atom categories (hydrophobic, positive, negative, hydrogen acceptor, hydrogen donor, aromatic, Sulphur, and neutral). [145] The difference in pharmacophore count vectors between mutant and wild type was utilized to define the pharmacophore change vector, which was subsequently attached to the mCSM signature. The mCSM signature also included experimental parameters such as pH, temperature and the RSA of wild-type residue. The thermodynamic data sets from ProTherm[146], SKEMPI[147], and ProNIT[148] were used to train the mCSM prediction models for protein stability, protein-protein, and protein-nucleic acid interactions.

2.6.6 Molecular mechanics with generalized Born and surface area solvation (*MM*/*GBSA*)

The binding free energy of AR-TRβ heterodimers was evaluated using MM-GBSA by hawkdock server (http://cadd.zju.edu.cn/hawkdock/)[149-152].

The MM/GBSA computation was based on the ff02 force field[153] and the GB^{OBC1} model[154]. As a result, these characteristics were used to the study of critical residues in protein-protein binding surfaces. The tleap

module in Amber16[155] first inserted all missing hydrogens and heavy atoms of the protein-protein complex, and then the ff02 force field parameters were assigned to the proteins. Following that, the complex was optimized in vacuo using 2000 cycles of steepest descent minimization and 3000 cycles of conjugate gradient minimization. Finally, the modified GB (GBOBC1) models established by Onufriev et al. were used to determine the polar desolation energy. The dielectric constants for the outside and interior/solute were adjusted at 80 and 1, respectively.

2.6.7 Analysis of interaction interface

The AR-TRβ heterodimer binding chains and binding sites were identified using PDBePISA (https://www.ebi.ac.uk/pdbe/pisa/). The binding chains for these two proteins' interactions were determined from the interface result, while the binding sites were obtained from the buried residues analysis findings

Chapter 3 Evidence for Expression of TRα and TRβ in Human Prostate and Prostate Cancer Cells.

The aim of the project was to investigate potential heterodimer interactions between the LBDs of the Androgen Receptor with Thyroid Hormone receptors TRα and TRβ, using a split YFP interaction assay in mammalian cells, or in yeast two hybrid assays. Evidence for potential interaction of these NRs is based on previous observations in the literature suggesting that androgen and thyroid signaling may coregulate genes in prostate cancer cells[9, 93]. Therefore, we first set out to research the evidence for expression of thyroid receptor genes in prostate and prostate cancer tissues.

The human TRα gene is located on chromosome 17 and five transcripts have been described in the NCBI gene data base (Figure 3.1). The two major transcripts encode TRα1 and TRα2 which differ at their C-terminal ends. The last exon of TRα2 overlaps with the NR1D1 gene encoding Reverbα located on the opposing strand.

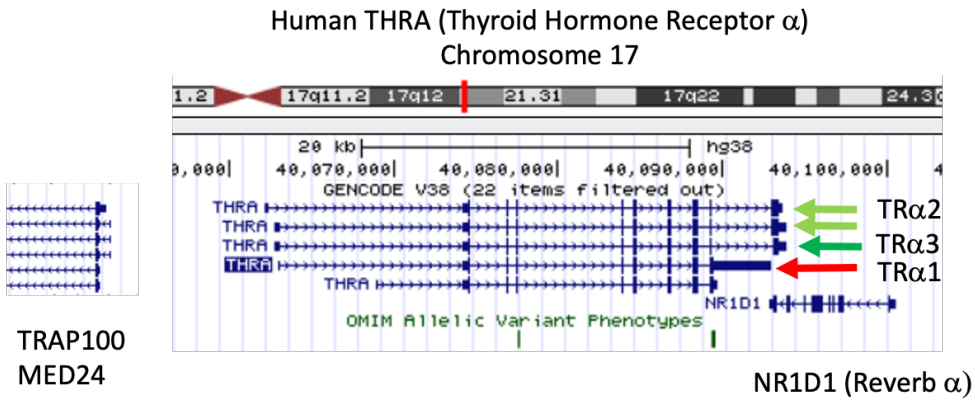


Figure 3.1. TRα gene transcripts listed in Gencode as viewed in the Human Genome Browser (GRCh38/hg38). The major isoforms are TRα1 and TRα2.

The TRβ gene is located on chromosome 3 and multiple transcripts have been described as shown in Figure 3.2. These transcripts differ mainly in their N-terminal sequences, and show tissue specific expression, but TRβ1 is the most widely expressed isoform. The TRβ gene is adjacent to but does not overlap the NR1D2 gene encoding Reverbβ.

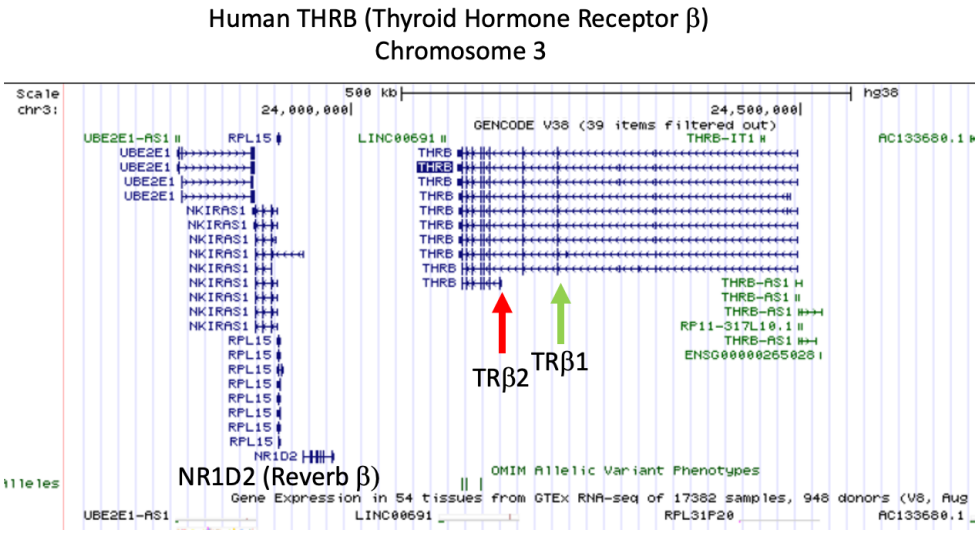


Figure 3.2. TRβ gene transcripts listed in Gencode as viewed in the Human Genome Browser (GRCh38/hg38). The major isoforms are TRβ1 and TRβ2 are indicated.

A search of the GTEX portal confirmed that both TRα and TRβ transcripts are expressed in normal prostate tissue (Fig. 3.3) with median TPM values (normalised transcripts per million) of 6 and 32, respectively, compared to AR (21) NR1D1 (61) NR1D2 (21) and CRYM (13).

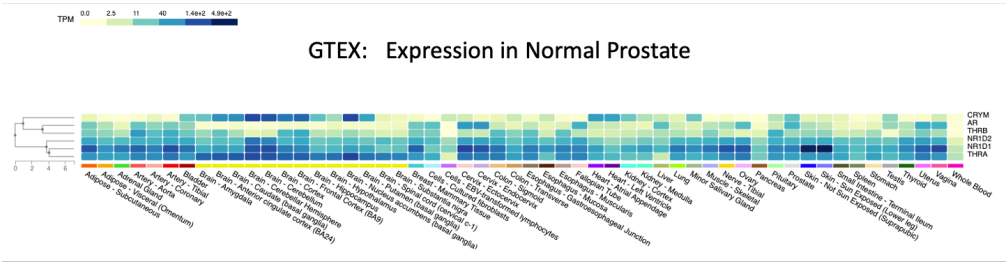
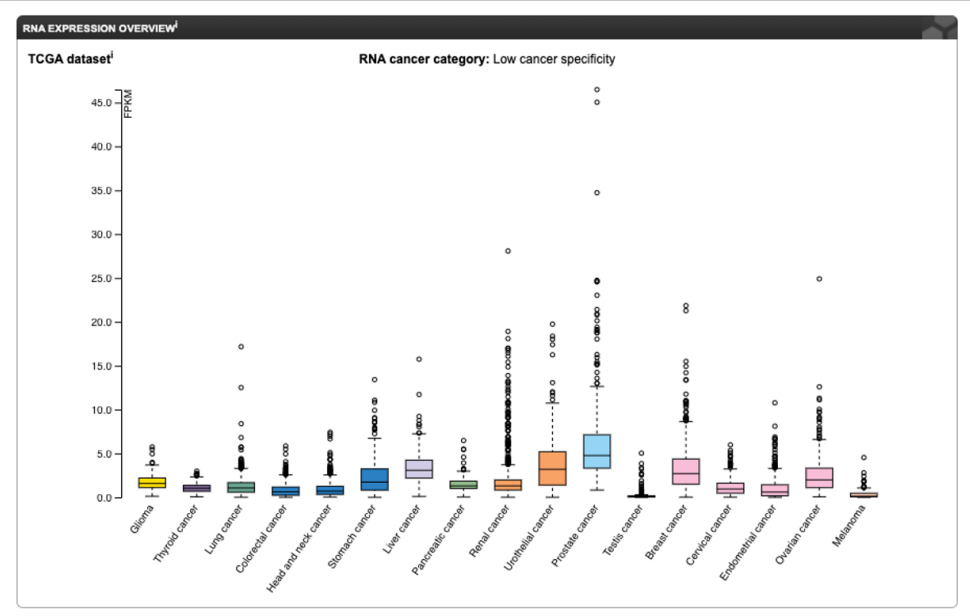


Figure 3.3. Expression levels in a range of normal tissues for the genes indicated represented as normalised transcripts per million (median TPM). Obtained from Gtexportal.org.

Analysis of cancer-enhanced expression derived from TCGA data and accessed via the *proteinatlas.org* shows that while changes in TRβ expression compared to normal tissue are not significantly associated with any particular tumor type, there is a tendency towards enhanced expression in prostate cancer (Fig 3.4). No such trend was observed for TRα.

THRB expression enhanced in prostate cancer (TCGA data)



https://www.proteinatlas.org/ENSG00000151090-THRB/pathology

Figure 3.4 TCGA expression data showing TRβ expression in different human tumour types. Data is presented as FKPM (fragments per kilobase of exon per million mapped fragments).

We also accessed the cBioPortal database for expression information for TRα and TRβ in prostate cancer cell lines. As shown in Fig.3.5, no alterations in gene copy number of TRα or TRβ were observed as both genes were found to be diploid. The prostate adenocarcinoma LNCaP and VCAP showed highest expression of TR genes, whereas expression in the AR negative cell line PC3 was lower. This may be related to the observation that AR regulates expression of the TRβ gene, as mentioned in the Introduction[9].

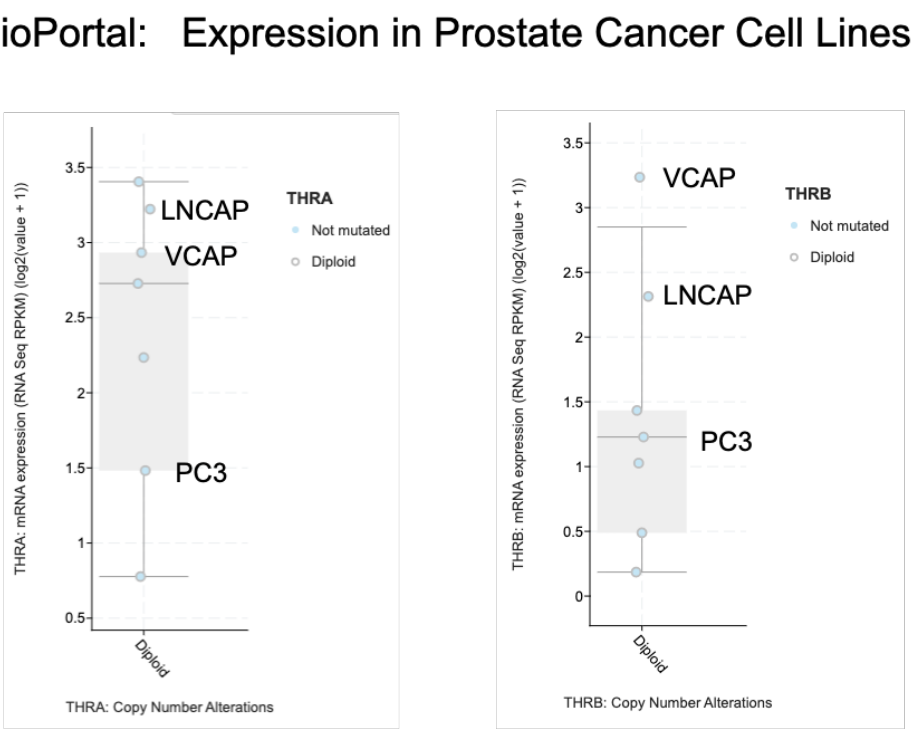


Figure 3.5 Expression data for 7 prostate cancer cells in the CCLE dataset showing relative expression of TRα and TRβ transcripts. Both genes were diploid in all cases indicating no gene copy number variation. Expression in the metastatic AR+ PCA cells line LNCaP is highlighted, with VCAP (ARV7+) and PC3 (AR-) shown for comparison.

Finally, as RNA Sequence data was available in the group for the LNCaP C4-2 cell line, we confirmed that the TRβ1 isoform is expressed and represented in the RNA seq reads (Figure 3.6). A Sashimi plot generated in IGV also indicated that both TRα1 and TRα2 as well as other transcripts appear to be expressed in C4-2 cells (Figure 3.6B)

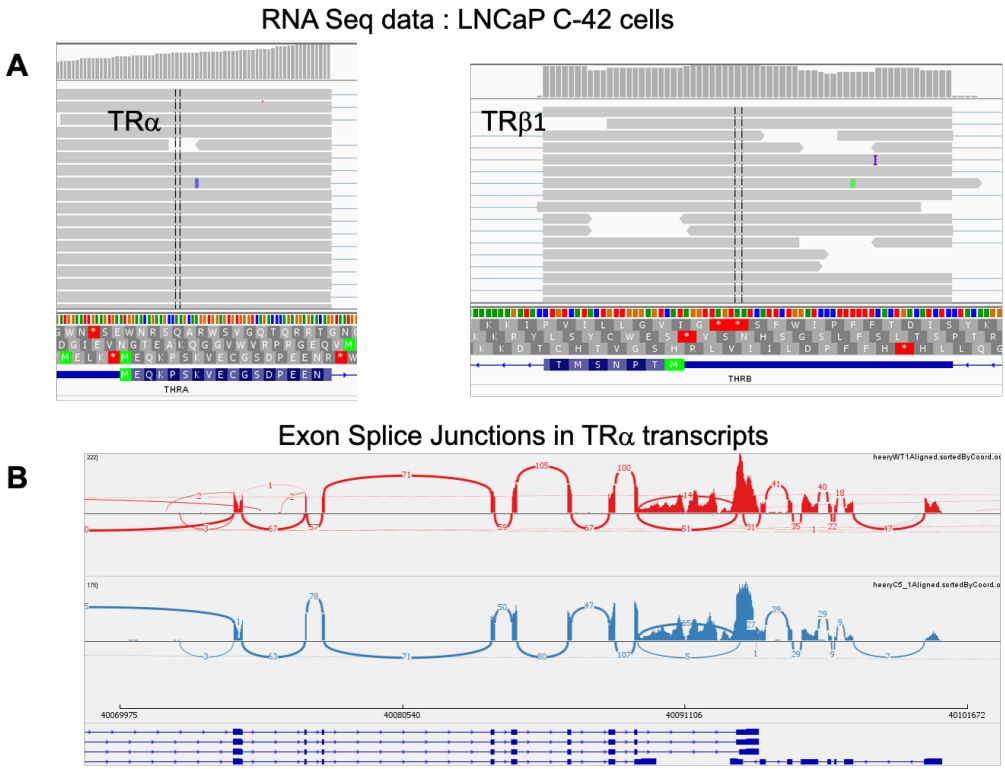


Figure 3.6 A) RNA Seq data from LNCAP C4-2 cells showing mapped reads visualised in the integrated genome viewer (IGV). The regions shown are the canonical exon 2 containing the AUG start codon for all TRα isoforms, and the canonical exon 1 unique to the TRβ1 isoform

In summary, the above analyses confirm that both TRα and TRβ genes are expressed in human prostate tissues, which also expresses the AR gene. TRα two major transcripts encode TRα1 and TRα2 which differ at their C- terminal ends. TRβ1 is the most widely expressed isoform for TRβ. While both genes are expressed in normal prostate TR*β* expression showed a tendency to have enhanced expression in prostate cancers and prostate cancer cell lines. Co-expression of TRs and AR would be a requirement for the formation TR/AR heterodimers, should they exist *in vivo*. In addition, as prostate adenocarcinoma

LNCaP and VCAP cells showed highest expression of TR genes according to cBioPortal database, it was decided to use LNCaP C4-2 cell cDNA as a template to PCR amplify TR LBD coding sequences to generate expression vectors for use in this study.

Chapter 4 Prediction for AR-TRβ heterodimer structure.

The aim of the project was to investigate potential heterodimer interactions between AR LBD and TRβ LBD, as a novel target in CRPC. Having confirmed that AR and TRs are co-expressed in prostate and prostate cancers, the next step was to undertake an in-silico approach to explore these interactions. Homology modelling of the TRs and AR proteins was performed to add missing loops, assuring the completeness of the protein structures. In order to produce "consensus-based" projected docking complexes, we decided to utilize a meta-strategy that combined the benefits of Four of the most widely used protein-protein docking servers (HADDOCK, HDOCK, GRAMM-X, and ZDOCK) were utilized to predict AR-TRβ complex. The resulting complexes were grouped using the "clustering of conformer" module implemented in Maestro Schrodinger (Schrödinger Release 2020-2: Maestro, Schrödinger, LLC. Kelley penalty plots were used to identify the optimal number of clusters, N. The final docking pose was chosen by comparing the binding affinity of specific residues through virtual site directed mutagenesis. The binding interface was analyzed by PDBePISA and HAWKDOCK server.

4.1 Structure Choice for Modelling

As explained in Chapter 1.3, binding of agonists may cause conformational change of the AR LBD to promote homodimer formation.

Moreover, according to previous yeast two hybrid study, the interaction between AR and TR is ligand-dependent and mediated via the LBD. Thus, we scanned all the PDB structures of AR and TR in the PDB databank (https://www.rcsb.org/) and chose 2AM9, 2AMA, 2AMB, 2AX9, 2AXA, 2IHQ and 2NW4 to present AR LBD monomers, while 1XZX and 3GWS are chosen for TR which are LBD domains monomers co-crystalized with agonist (Table 4.1).

Table 4.1 Information of PDBs representing the AR and TR structure

NR	PDB	Method	Resolution	Chain	Position	Domain	
AR	2AM9	X-ray	1.64 Å	A	655-920	LBD + testosterone	156
AR	2AMA	X-ray	1.90 Å	A	655-920	LBD + dihydrotestostero	156
AR	2AMB	X-ray	1.75 Å	A	655-920	LBD + tetrahydrogestrinon	156
AR	2AX9	X-ray	1.65 Å	A	665-920	LBD + R-3	157
AR	2AXA	X-ray	1.80 Å	A	665-920	LBD+S-1	157
AR	2IHQ	X-ray	2.00 Å	A	647-902	LBD + N-Aryl-Hydroxybicyclohydantoin	158
AR	2NW4	X-ray	3.00 Å	A	647-902	LBD + BMS-564929	102
TR	1XZX	X-ray	2.50 Å	X	202-461	LBD+T3	103
TR	3GWS	X-ray	2.20 Å	X	202-460	LBD+T3	159

The AR LBD has been shown to form homodimers upon androgen stimulation [130] and the crystal structure has been resolved. Thus, we docked the AR LBD monomer structures with themselves using HADDOCK, HDOCK, GRAMM-X, and ZDOCK and compared the RMSD value of the docked AR homodimer structure with the published crystal AR homodimer structure, 5JJM, to judge the docking and

structure accuracy. As illustrated in Figure 4.1, the docked structure of 2NW4 dimer show high similarity with the AR homodimer crystal structure, supporting the suitability of this approach to explore binding modes.

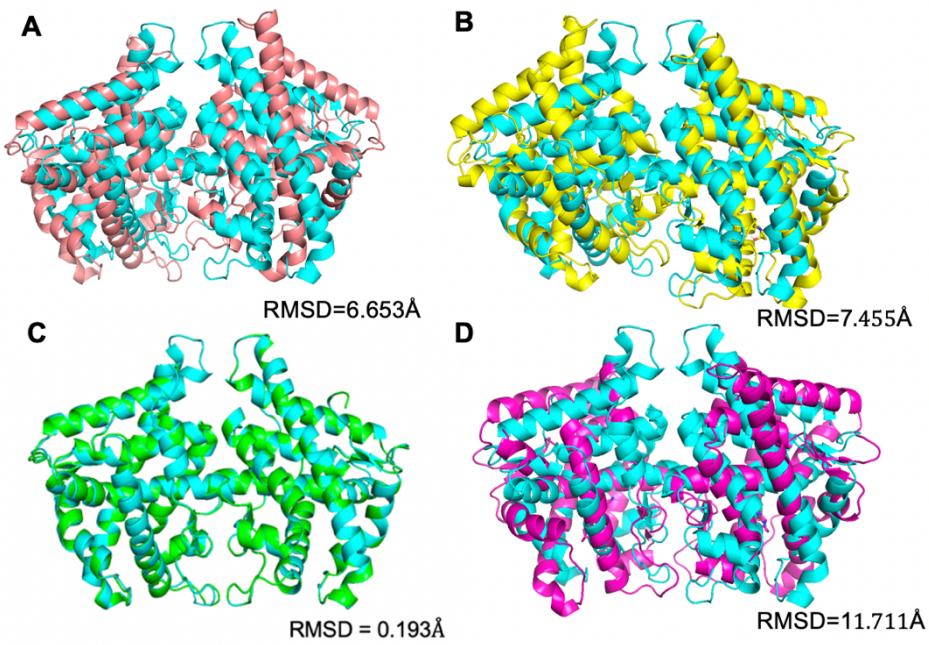


Figure 4.1 Superposition of predicted docking structure of AR LBD monomer 2NW4-2NW4 using A) HADDOCK, B) HDOCK, C) GRAMM-X, and D) ZDOCK, with the AR LBD homodimer, 5JJM. The similarity was validated by calculating the Cα RMSD values. 5JJM was shown in cyan.

For the TRβ LBD structures, it was observed that the sequences of both 1XZX and 3GWS lacked a hinge region (residues: 252-263). The missing region of TRβ was added by homology modelling. To fix the missing region, we downloaded the FASTA sequence of TRβ LBD (Residues: 36-280) from Uniprot (https://www.uniprot.org/) and used it as the target sequence to build the full length TRβ LBD structure using ProMod3 [123] section in swiss-model server

(https://swissmodel.expasy.org/) [160]. The initial input spans the sequences 211-460. Several structures, mostly ligand binding domains, were predicted with a confidence of more than 90%. Predictions for the AA 211-460 area were created using 1XZX (Residues: 22-280) as template, as shown in Figure 4.3. Conserved residues are copied directly from the template to the model while the insertions were remodeled based on the fragment library, the geometry of resulting model is optimized by a force field[123]. QMEAN Z-scores of Cβ, all-atom and solvation were close to 0 which suggested a strong match between the predicted model and the template (Figure 4.3C). However, the score for torsion is not so relevant. The predicted model quality is close to non-redundant pdb structures of similar size (Figure 4.3D). Almost all the residues of predicted model were located in the allowed area according to Ramachandran plots (Figure 4.3E-H). Thus, the predicted model was considered acceptable to represent TRβ LBD.

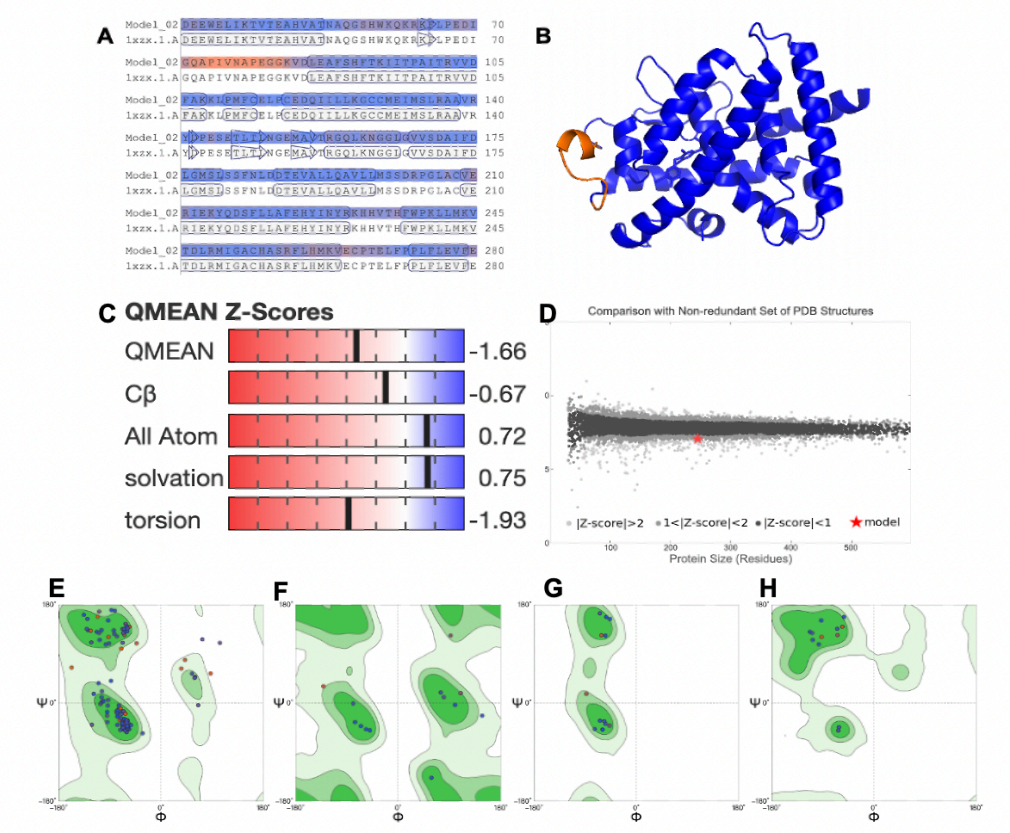


Figure 4.2 A) The sequence of modelled TRβ LBD. The missing hinge region sequence is highlighted. B) The predicted structure of modeled TRβ. The conserved residues are shown in blue while the added residues are shown in orange. The green and yellow circles represent residues in the core and permitted areas, respectively. C) The model and its components' QMEAN Z-scores of Cβ, all-atom, solvation, and torsion. D) Estimating the model's global quality by comparing the normalized QMEAN scores (denoted by a red asterisk) to a non-redundant collection of pdb structures of similar size. Ramachandran plots showing E) General residues, F) Glycine, G) Proline and H) Pre-Proline.

The final structure, as shown in Figure 4.4, was used for further analysis. The predicted structure consists of 10 α-helices (from H1 to H10) linked by short loops. One β-sheet was predicted between H4 and

H5. A short C-terminal α-helix was also incorporated in certain versions (H10). The sequence 52-74 had moderate disorder levels for the secondary structure and might be found in the loop between H1 and H2.

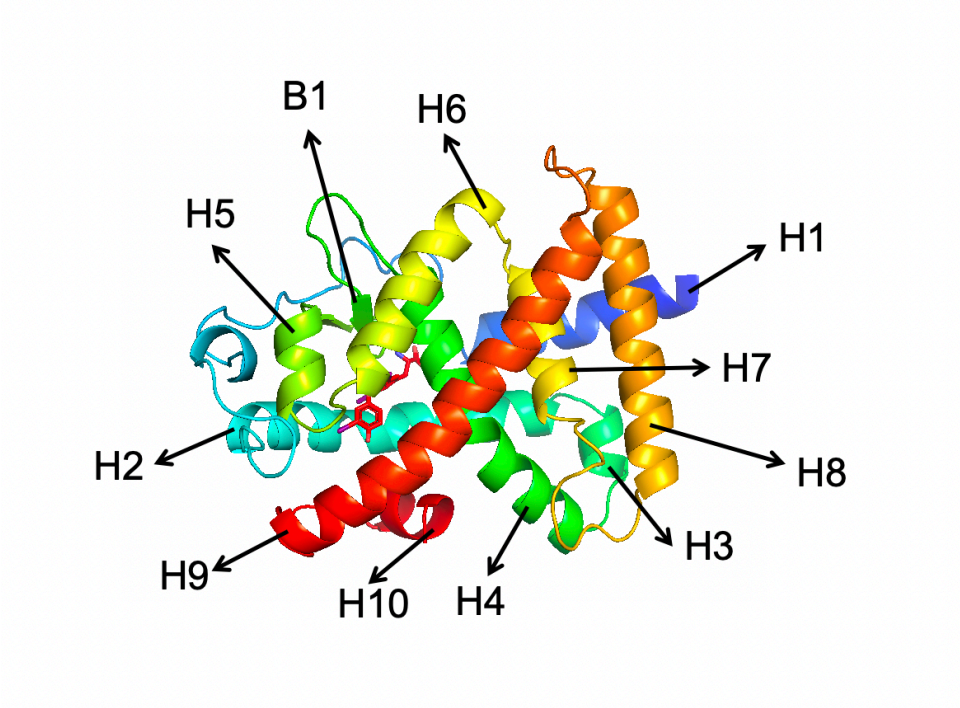


Figure 4.3 The predicted structure for TRβ. α-helix was represented by H while β-sheet was represented by B

4.2 Interaction mode analysis

Parallel docking analyses were performed for the interaction between the model TRβ LBD and the AR LBD, represented by 2NW4, as mentioned above by using HADDOCK, HDOCK, GRAMM-X and ZDOCK servers. In the reported AR LBD homodimer structure[129], the major interface involves hydrophobic and Van der Waals interactions stabilized by a series of hydrophobic side chains, including W752, F755, Y764, and V770 and additional van der Waals of hydrogen

bonding interactions with P767, P802 and polar residues. In previous unpublished work from our group, site-directed mutagenesis of the AR LBD confirmed that the residues W752, F755, Y764, P767 and V770 are crucial for the formation of AR-AR LBD homodimers in yeast two-hybrid assays. In the protein-protein docking computations, they were thus designated as “interacting residues”. The Conformer Cluster module in Maestro Schrodinger (Schrödinger Release 2020-2: Maestro, Schrödinger, LLC) was used to cluster identical structures of 5 lowest scored predicted structures from each server. Figure 4.5 shows the results of the conformation clustering. The optimal number of clusters from the Kelley penalty plot is 10, as shown in Figure 4.5A, and the corresponding distance matrix is shown in Figure 4.5B. Figure 4.5C-E depicts the 3 most populous cluster, which has 12, 7 and 6 members, respectively. ZDOCK occupied 6 (100%) members in cluster 8. HADDOCK took 6 (50%) members in cluster 1 and 4 (57%) members in cluster 2. GRAMM-X had 6 (50%) members in cluster 1, and 3 (43%) members in cluster 2. HADDOCK 5, HADDOCK 7 and ZDOCK 1 were the components closest to cluster 1, 2, 8. The components of each model in the clusters were shown in Table 4.2.

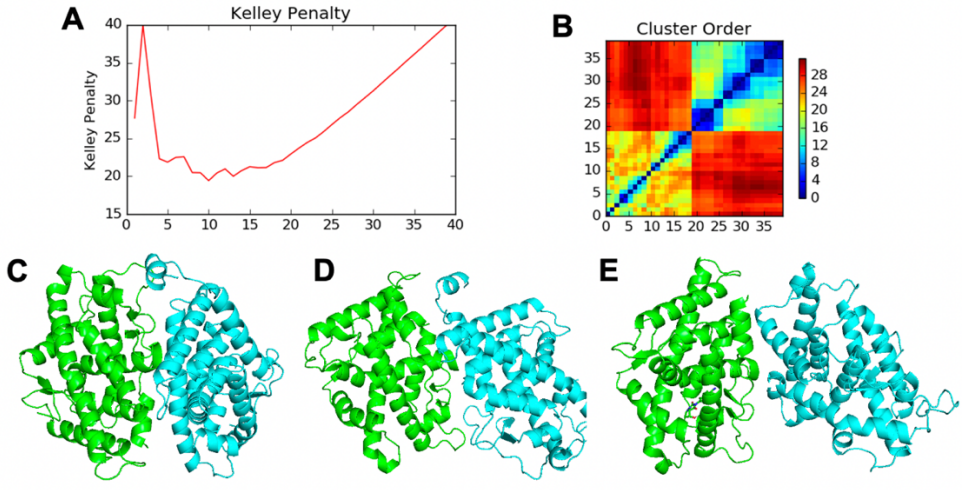


Figure 4.4 A) Kelley penalty plot and B) AR-TR clustering distance matrix. The structure of 3 most populated clusters C)1, D)2, E)8 represented by HADDOCK 5, HADDOCK 7 and ZDOCK 1, respectively, corresponding to AR-TRβ binding. AR and TRβ are colored in green and cyan, respectively.

Table 4.2. Cluster component analysis for each AR-TRβ complex cluster.

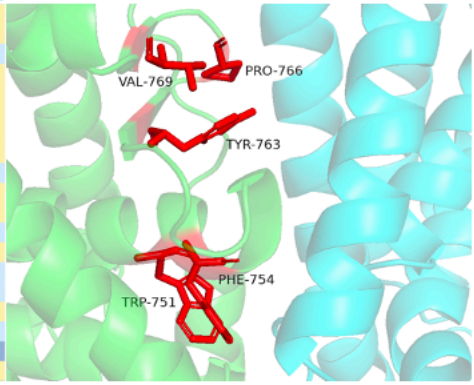
Cluster	Docking Server	Component
1	HADDOCK	1, 4, 5, 6, 9, 10
	GRAMM-X	1, 4, 5, 6, 7, 8
2	HADDOCK	2, 3, 7, 8
	GRAMM-X	2, 9, 10
3	HDOCK	1, 10
4	HDOCK	3, 6
5	HDOCK	2, 5, 8, 9
6	HDOCK	4
7	HDOCK	7
8	ZDOCK	1, 3, 6, 8, 9, 10

9	ZDOCK	2
10	ZDOCK	5, 7

PDBePISA was used to analyze the interaction interface of predicted heterodimer clusters. As shown in Figure 4.6, residues from W752 (numbered W751 in figure) to V770 (numbered V769 in figure) in AR, are in contact with TRβ LBD in cluster 1. These residues are only partially contacted with TRβ LBD in cluster 2 and cluster 8.

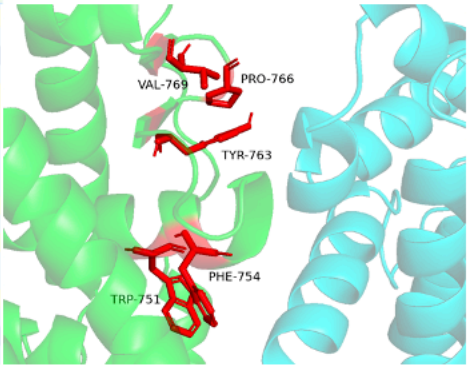
Cluster 1

Structure 1	HSDC	ASA	BSA	ΔG
A:TRP 751		34.60	15.63 |||||	0.07
A:ARG 752		62.58	1.84 |	0.03
A:SER 753		0.17	0.00	0.00
A:PHE 754		43.14	42.51 ||||||||	0.62
A:THR 755	H	71.18	71.18 ||||||||	0.25
A:ASN 756	H	74.21	72.62 ||||||||	-0.42
A:VAL 757		31.24	31.24 ||||||||	0.41
A:ASN 758	H	103.02	79.63 |||||||	-0.29
A:SER 759		11.86	0.00	0.00
A:ARG 760		184.84	30.39 ||	0.36
A:MET 761		56.42	17.23 ||||	0.28
A:LEU 762		0.50	0.00	0.00
A:TYR 763		43.76	42.05 ||||||||	-0.02
A:PHE 764		19.35	0.00	0.00
A:ALA 765		0.34	0.00	0.00
A:PRO 766		53.54	24.51 |||||	0.14
A:ASP 767		30.10	0.00	0.00
A:LEU 768		0.00	0.00	0.00
A:VAL 769		34.92	13.73 ||||	0.22



Cluster 2

Structure 1	HSDC	ASA	BSA	ΔG
A:TRP 751		36.39	16.29 |||||	-0.13
A:ARG 752		56.98	0.00	0.00
A:SER 753		0.00	0.00	0.00
A:PHE 754		47.51	0.00	0.00
A:THR 755		74.03	11.89 ||	0.00
A:ASN 756		82.18	58.94 |||||||	-0.36
A:VAL 757		35.85	26.12 |||||||	0.42
A:ASN 758		110.20	0.00	0.00
A:SER 759		8.30	0.00	0.00
A:ARG 760		192.05	0.00	0.00
A:MET 761		51.29	1.67 |	0.03
A:LEU 762		0.17	0.00	0.00
A:TYR 763		39.45	33.96 ||||||||	-0.12
A:PHE 764		16.14	0.00	0.00
A:ALA 765		0.50	0.00	0.00
A:PRO 766		57.36	16.61 |||	0.21
A:ASP 767		28.30	0.00	0.00
A:LEU 768		0.00	0.00	0.00
A:VAL 769		38.92	0.00	0.00



Cluster 8

Structure 1	HSDC	ASA	BSA	ΔG
A:TRP 751		39.00	16.07 |||||	0.09
A:ARG 752		77.86	0.00	0.00
A:SER 753		0.27	0.00	0.00
A:PHE 754		52.06	37.38 ||||||	0.57
A:THR 755		78.53	42.11 |||||	0.55
A:ASN 756		91.98	0.00	0.00
A:VAL 757		40.60	0.00	0.00
A:ASN 758	H	105.15	78.16 |||||||	-0.71
A:SER 759		11.90	0.00	0.00
A:ARG 760		180.13	16.49 |	-0.11
A:MET 761		75.75	0.00	0.00
A:LEU 762		1.00	0.00	0.00
A:TYR 763		35.05	0.00	0.00
A:PHE 764		31.54	0.00	0.00
A:ALA 765		1.84	0.00	0.00
A:PRO 766		58.80	0.00	0.00
A:ASP 767		47.10	0.00	0.00
A:LEU 768		0.00	0.00	0.00
A:VAL 769		48.47	0.00	0.00

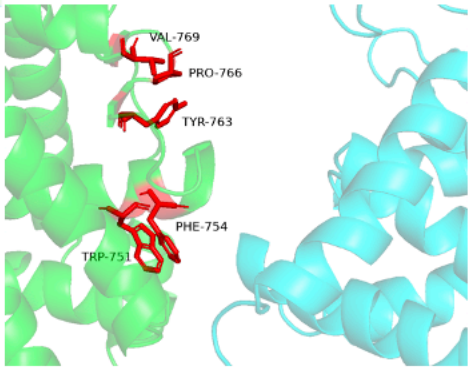


Figure 4.5 Interaction interface analysis for cluster 1, 2 and 8 by PDBePISA. Inaccessible residues are shown in dark, solvent-accessible residues are shown in light blue and interfacing residues are shown in yellow. H refers to residues, making hydrogen bonds, ASA refers to Accessible Surface Area (A2), BSA refers to Buried Surface Area (A2), ΔiG refers to solvent energy effect (kcal/mol), bars refer to buried area percentage (one bar per 10%).

The binding affinity of certain residues in AR were further explored via comparing the binding free energy performed by HAWKDOCK. As shown in Table 4.3, cluster 1 (with -32.73 kcal/mol) showed much higher average binding affinity than other clusters (-14.91kcal/mol for cluster 2 and -16.42 kcal/mol for cluster 8). In agreement with previous yeast two hybrid result, the binding free energies of W752, F755, Y764, P767 and V770 in cluster 1 are quite high. While only Y764 showed high binding free energy in cluster 2. And all the selected residues showed less affection on the interaction between AR and TRβ.

Table 4.3 Binding free energy calculation for residues in cluster 1, 2 and 8 through HAWKDOCK

Cluster\(kcal/mol)	All	W752	F755	Y764	P767	V770	T851
1	-32.73	-1.00	-3.88	-2.01	-2.09	-1.87	0.00
2	-14.91	-0.50	-0.06	-2.43	-0.73	0.00	0.00
8	-16.42	-0.48	0.00	0.00	0.73	0.00	0.01

AR variants showing disruption in yeast two hybrid were then analysed by virtual site direct mutagenesis using mCSM and SDM web server. From mCSM results, mutations, W752R, F755S and P767A, geometrically were predicted to disrupt the heterodimerization between AR and TRβ LBDs while V770A , T851D and Y764C were predicted to be less important for heterodimerization.From the SDM *in silico* analysis, F755S and V770A are predicted to increase the solvent energy, resulting in decreased stability of heterodimer, while mutations, Y764C and P767A are predicted to decrease the solvent energy which may stabilise the conformation.

Table 4.4 mCSM predicted Gibbs free energy between original and mutated complex. $\Delta\Delta G > 0$ indicate mutations stabilized the heterodimerization shown in red, where Gibbs free energy $\Delta\Delta G < 0$ destabilized heterodimerization shown in green.

Cluster\(kcal/mol)	W752R	F755S	Y764C	P767A	V770A	T851D
1	-1.324	-1.539	-0.014	-1.141	-0.303	0.171
2	-0.812	-0.438	0.130	-0.382	-0.445	0.238
8	-0.796	-0.478	0.021	-0.138	-0.251	0.254

Table 4.5 SDM predicted Gibbs free energy between original and mutated complex. ΔΔG > 0 indicate mutations stabilized the heterodimerization shown in red, where Gibbs free energy ΔΔG < 0 destabilized heterodimerization shown in green.

Cluster\(kcal/mol)	W752R	F755S	Y764C	P767A	V770A	T851D
1	0.09	-0.42	0.38	0.67	-1.72	0.12
2	0.31	-0.42	0.38	0.67	-1.72	0.42
8	0.18	-0.53	0.38	0.67	-1.72	0.40

Thus, the cluster 1 model, which showed the best consistency between the in silico predictions and yeast two-hybrid data, was selected for interaction mode analysis.

The predicted heterodimer structures in cluster 1 were uploaded to Hawkdock, and the binding free energy of each residue in predicted heterodimers were calculated by MM-GBSA method. *Figure 4.6 showed the key part in the AR-TRβ interaction. The results were consistent with the protein surface analyses by PDBePISA. As shown in Figure 4.7A, the residues surrounded beta-sheet 1 in AR are major contributor for the binding with TRβ. This region contains the end of helix 4, beta sheet 1, and helix 5. The beginning of helix 7 also showed high binding affinity toward TRβ. The hinge region from residues 682 to 692 seems to be anchored by the interaction with TRβ. For TRβ, the binding region is mainly located in helix 9 and helix 10.*

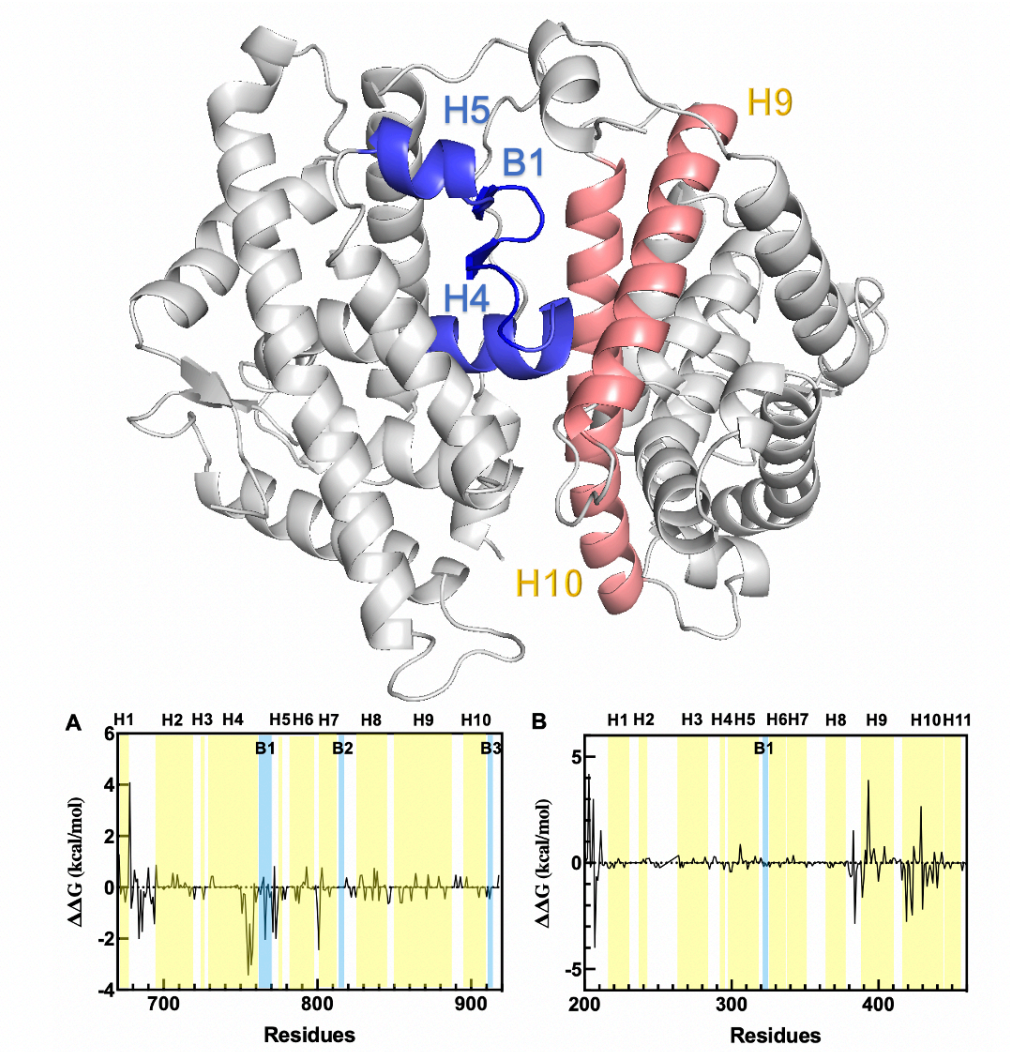


Figure 4.6 Picture above is the predicted structure of AR-TR LB. The chain for AR is on the left while the chain for TR LBD is on the right. MM/GBSA results for each residue in predicted heterodimer through hawkdock. A is for AR while B is for TRβ. The helix region for each receptor is shown in yellow and Beta sheets are shown in blue. Both of regions are numbered as above

4.3 Summary

According to previous yeast two hybrid study, the interaction between AR and TR is ligand dependent and promoted via the LBD. Thus, the projected domains in the AR are found between residues 647-902 and residues 211-460 for TRβ. Both projected nuclear receptor structures,

notably LBD part, exhibit a high level of conformational conservation, suggesting that interaction between AR and TR proteins might possess a conserved contact mode similar to the AR homodimer. The binding interface of AR homodimer involved a largely hydrophobic interaction via several helices and one beta sheet. In contrast, other nuclear receptor form dimers involving contacts between the longest α-helix (H10/H11).

As discussed in Chapter 1.3, binding with agonists could cause conformational change of AR which helps AR to form homodimer. AR The representative AR LBD monomer structure for docking was chosen based mainly on 2 principals. First, the PDB structure was co-crystalized with an agonist. The RMSD value between the chosen structure and 5JJM, the reported crystal AR homodimer, was low. The reported structure of TRβ missed a hinge region between residues :252-263. The missing part had been fixed by homology modelling, using swiss-model server. The selected AR structure (2NW4) and modelled TRβ structure were using for docking.

As previously stated, different docking servers performed protein-protein docking in a variety of methods so we combined 4 common docking servers, HADDOCK, HDOCK,GRAMM-X and ZDOCK, to analyze the interaction between AR and TRβ LBDs. The top 10 scored complexes from each server were clustered by Conformer cluster module in Maestro Schrodinger (Schrödinger Release 2020-2: Maestro, Schrödinger, LLC). All clusters were validated by surface interaction analysis through PDBePISA server, binding free energy calcualtion

through HAWKDOCK and then sent to SDM and mCSM server for virtual site directed mutagenesis analysis, where the mutated residues, W752, F755, P767 and V770 of AR, were predicted to decrease the stability of AR-TRβ complex. In support of this, these residues were found to be critical for the formation of AR homodimers in unpublished yeast two-hybrid data from our group.

The qualified cluster was re-validated by HAWKDOCK using MM/GBSA calculation. The residues surrounded beta-sheet 1 in AR are major contributor for the binding with TRβ. This region contains the end of helix 4, beta sheet 1, and helix 5. The beginning of helix 7 also showed high binding affinity toward TRβ. The hinge region from residues 682 to 692 seems to be anchored by the interaction with TRβ. For TRβ, the binding region is mainly located in helix 9 and helix 10.

Chapter 5 Generation of Expression Vectors to Explore Interactions between AR and TRβ using Split YFP system

Green fluorescent protein (GFP) and its yellow, red and cyan derivatives (YFP, RFP, CFP) are important tools for observing protein expression, localization and interactions in live cells [162]. The "split" vYFP system uses two expression vectors encoding separate and non-overlapping N- and C-terminal halves of the venus YFP protein. When transfected into mammalian cells, these separated portions are unable to assemble into intact vYFP protein, and thus lack the ability to generate a fluorescent signal. However, if the fragments are brought into close proximity e.g. through assembly of fused sequences, they may be able to refold to reconstitute a functional YFP protein, resulting in the acquisition of stable fluorescence (Figure 5.1). This complementation is achieved by fusion of the split vYFP fragments to bait and prey protein sequences, which restore fluorescence if they form a complex *in vivo*. The Venus YFP (vYFP) is a modified version of jellyfish GFP, that has been engineered to emit yellow fluorescence and be more stable at 37 ^{o}C, as well as being less sensitive to acidosis or chloride ions. This modification allows for more sensitive and reliable imaging of protein complexes with high spatial resolution inside a live cell, depending on the time necessary for fluorophore synthesis and complex stability through interaction of fluorescent protein fragments. Because the test may be performed in the natural cellular environment,

there is no need to dye, fix, or lyse the cells, as was the case with more intrusive assays.

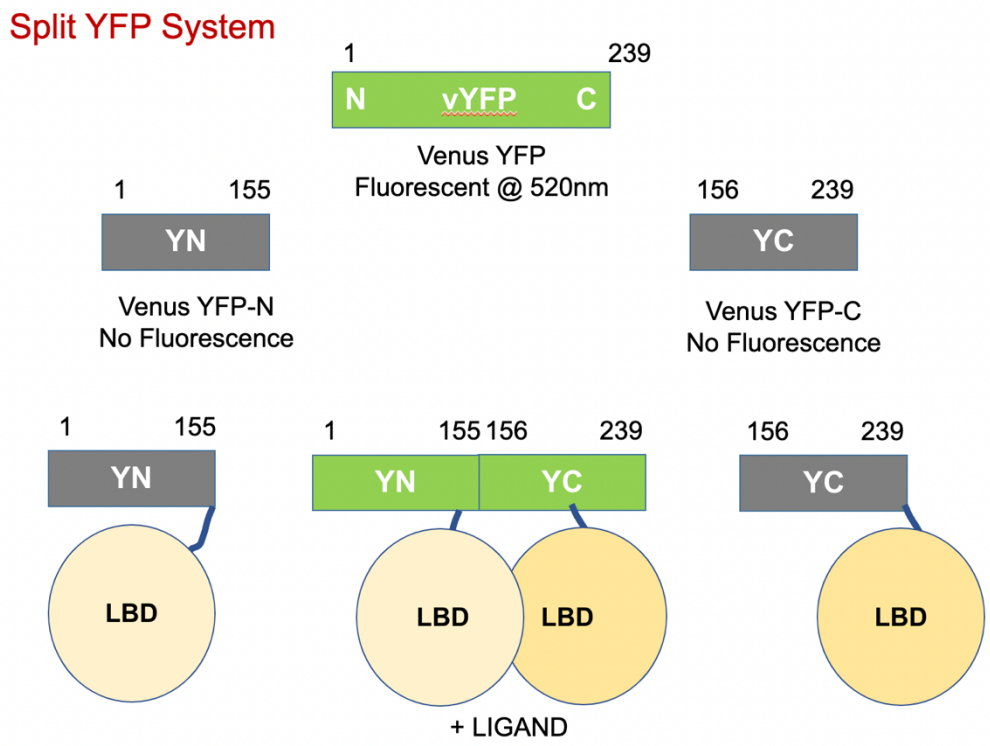


Figure 5.1. The principal of split vYFP system. Potential interaction partners are shown by yellow circles. They are fused to fluorescent protein fragments, which are depicted by green and gray rectangles (full length YFP and split vYFP). If the proteins interact, they may enhance the association of fluorescent protein fragments and show fluorescent (green rectangle)

As discussed earlier, androgen receptor plays an important role in both normal prostate development and prostate cancer progression. Most prostate tumours contain amplification or overexpression of the AR gene enhancing the oncogenic driver function in this disease (Fig. 1.11). Antihormone therapies such as flutamide and enzalutamide have been developed to compete with natural androgens for binding to the

LBD, thus blocking progression of this disease. Although many drugs targeting AR have been developed, in the case of metastatic disease the accumulation of mutations in the AR sequence associated with therapy resistance (CRPC), means that antihormone therapies usually become ineffective in patients within several years. Thus, research to find other targetable signaling pathways is crucial to the development of alternative therapies. In this context, our previous work has provided evidence that some genes in prostate cancer are coregulated by AR and TRs and their hormones. In addition, low expression levels of the thyroid hormone binding protein CRYM have been shown to be associated with poor prognosis. These observations prompted work in our laboratory to investigate potential direct interactions between AR and TRβ. Thus, an investigation of whether association of the LBDs of AR and TRβ using split vYFP system, would provide strong supporting evidence to confirm these interactions in mammalian cells.

To that purpose, we set out a plan to assess ligand-dependent complex formation between NR LBDs and using LBD- cofactor peptide interactions as controls. Steroid receptor coactivator (SRC1) and mediator protein (MED1; which is also known as thyroid receptor associated protein TRAP220) both contain multiple LXXLL motifs important for binding to NR LBDs. The SRC1 NR interaction domain (NID) contains 3 LXXLL motifs[163] whereas the MED1 sequence LXM1 shows strong interaction with TR LBDs. These cofactor sequences, as shown in Figure 5.2. would be used as positive controls for LBD binding. Expression vectors for vYFPN155-AR LBD, vYFPN155 SRC1

NID and vYFPC156-MED1 LXM1, were previously generated in our lab. We have previously detected ligand induced interaction of the estrogen receptor LBD with the SRC1 NID and MED1 LXM1 constructs using this system. As prostate cancer cell lines are difficult to transfect, we used HEK293 cells to achieve efficient expression of our split vYFP constructs.

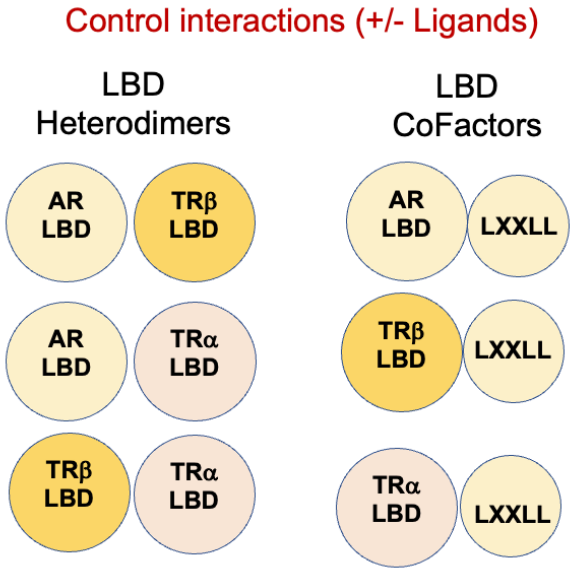


Figure 5.2 Plan of potential interactions to investigate in the split vYFP system . The formation of a heterodimers between the LBDs of AR and TRβ, AR and TRα , TRβ and TRα . As positive controls, LBD interactions with cofactor motifs, containing the sequence LXXLL, would be included.

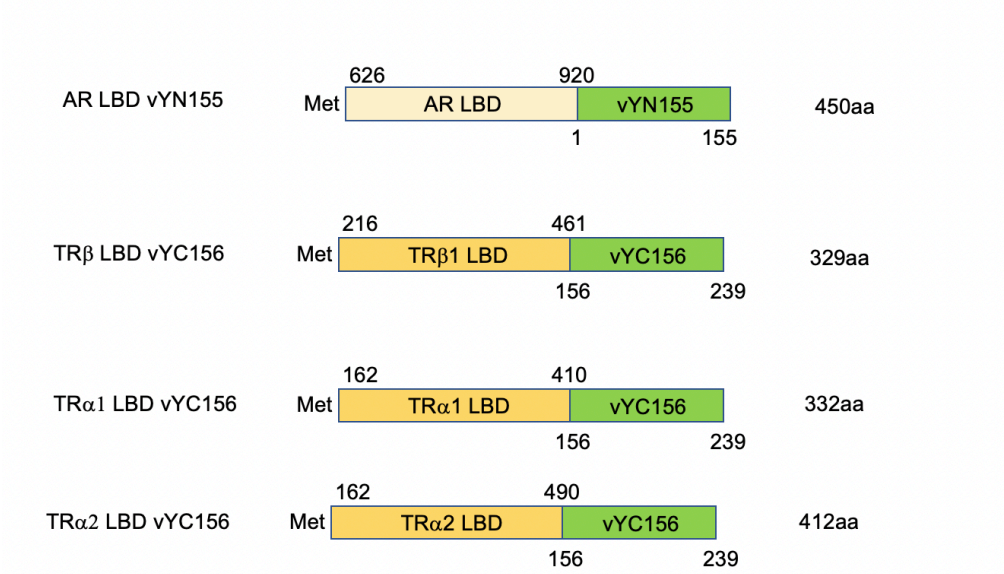
As vectors for testing TRs in this system were not available, it was first necessary to subclone LBD coding sequences for TRα and TRβ into the vYFPC156 vector, as described in the next section.

5.1 Cloning Strategy for Human TRα and TRβ LBD Sequences.

The split YFP biofluorescence complementation system utilises 2 vectors, the vYFP C terminus (156-239) and the vYFP N terminus (1-155). [164]. These contain multiple cloning sites to enable fusion in frame of 'bait' and 'prey' sequences, similar to yeast two hybrid interactions. The initial plan was to PCR amplify and clone the TRα1, TRα2 and TRβ1 LBDs cDNA sequences into the vector vYFPC156 (a gift from Dr Nick Holliday, University of Nottingham) so we could test interaction with AR LBD-YN155 in mammalian cells.

The cloning strategy chosen was to attempt to clone the thyroid receptor LBD sequences by PCR amplification from cell line cDNA, and use restriction digests to insert into target vectors, as this approach had been successful in the laboratory for cloning full length PIP5K1A and other sequences. Thus the constructs would consist of TRα1, TRα2 or TRβ1 LBDs fused in frame with the C terminus of venus YFP (156-239). PCR primers were designed for as shown in Figure 5.3. As the YFP sequence is tagged to the C-terminus of the LBD, the forward primers incorporated a restriction enzyme site followed by a Kozak sequence and ATG start codon to facilitate expression. Restriction enzyme sites were guided by absence in the target insert sequence. We chose the enzyme EcoRI (forward) and NotI (reverse) for TRα1; EcoRV (Forward), XhoI (Reverse) for TRα2; EcoRI (Forward), XhoI (Reverse) for TRβ1. An additional four adenosines were added to all primers at the beginning of 5' to improve efficacy of digestion by the restriction

enzymes. This is because some restriction enzymes work inefficiently if within 2-3 base pairs of the end of the DNA fragment. Thus, the forward primer from 5' end to 3' end includes a 4bp AAAA extension, RE sequence, KOZAK sequence, initiation codon (ATG) and at least 21 bp of homology to the target (sense strand) sequence of target gene. The Reverse primer from 5' end to 3' end includes AAAA and 21 bp of target homology (the antisense strand). The reverse primers also contained an in-frame restriction site (and no stop codon) to allow in-frame fusion to the C-terminal fragment of vYFP (156-239), as shown in the schematic in Figure 5.4.



```
                                  M   P   E   E   W   D   L   I
                      EcoRI KOZAK     162 163 164 165 166 167 168
hTRa1   F162: AAAA GAATTC CACC ATG CCT GAA GAG TGG GAT CTG ATC
                                  V   E   Q   D   E   F   V   E
                        NotI      410 409 408 407 406 405 404 403
hTRa1 R410 : AAAA GCGGCCGC GA GAC TTC CTG ATC CTC AAA GAC CTC

                                  M   P   E   E   W   D   L   I
                      EcoRV  KOZAK    162 163 164 165 166 167 168
hTRa2  F162 : AAAA GATATC CACC ATG CCT GAA GAG TGG GAT CTG ATC
                                V   E   Q   D   E   F   V
                        XhoI    410 409 408 407 406 405 404
hTRa2 R410 : AAAA    CTCGAG  GGG AGA GGC TGC ATT GCC TGC

                                  M   D   E   E   W   E   L
                     EcoRI KOZAK    216 217 218 219 220 221
hTRb1 F216 : AAAA-GAATTC CACC ATG GAC GAG GAA TGG GAG CTC

                              D   E   F   V   E   L   F
                     XhoI   461 460 459 458 457 456 455
hTRb1 R461 : AAAA-CTCGAG ATC CTC GAA CAC TTC CAA GAA
```

Figure 5.3 Schematic representation of TR LBDs fused in frame with the C-terminal portion of the YFP protein vYC165 (156-239). The amino acid boundaries for each LBD are indicated along with the sequence length of the fusion proteins. The AR LBD YN155 construct which was already available is also shown. Sequences of forward and reverse primers are also shown below the schematic.

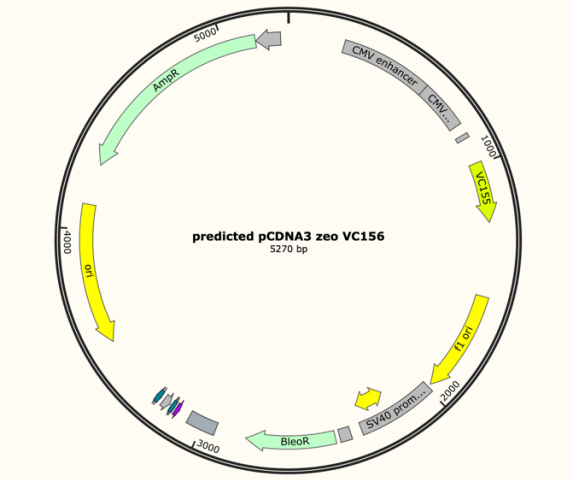


Figure 5.4 Plasmid map of vYC156 (vYFP; 156-239) which is used as the vector for cloning TR LBDs.

5.2 Optimization of PCR Amplification of TRα and TRβ LBD sequences

cDNA prepared from the prostate cancer cell line LNCaP C4-2 was initially provided by Ryan Morris. This cDNA was chosen because of the ample expression of TRs in this cell line as discussed in Chapter 3. Initial attempts to amplify products using 25ng/ul cDNA as input and an annealing temperature at the primer Tm, with a short cycling reaction (98^{o}C for 10s, Tm for 30s, 72^{o}C for 90s ; 30 cycles) were either unsuccessful or gave a very low amount of product. The poor yield could be due to a number of factors including rarity of the target template in the cDNA sample, suboptimal PCR reaction buffer, GC content of the template, the primer annealing temperature etc. However, human error was also a factor as my results were sometimes inconsistent between experiments due to pipetting errors, cross contamination of samples etc. To improve yield and specificity, a number of modifications were considered to optimize the conditions for each PCR reaction.

Annealing temperature is an important factor to consider in PCR process. Within a certain temperature range, the higher annealing temperature means the higher specificity of primers recognizing the specific area in template DNA. As the annealing temperature was increased, the primer requires a longer compatible sequence to bind with since the specificity will be higher. If the temperature is too low, a partial match between the primer and non-specific templates can lead to additional non-specific products on the gel.

To optimize the amplification of TRα1 (744 bp), TRα2 (984 bp) and TRβ1 (734 bp) LBD sequence from C4-2 cDNA, a PCR gradient was performed. The Tm for the TRα1 primer pair was calculated as 61°C. As shown in Figure 5.6, a weak band of the predicted size was obtained, with little change in yield observed between 59 °C and 62 °C, but yield was reduced at 63.1 °C. On the other hand, the yield of TRβ1 is decreased from 55 °C to 59 °C while the calculated Tm for TRβ1 primer pair is 60 °C.

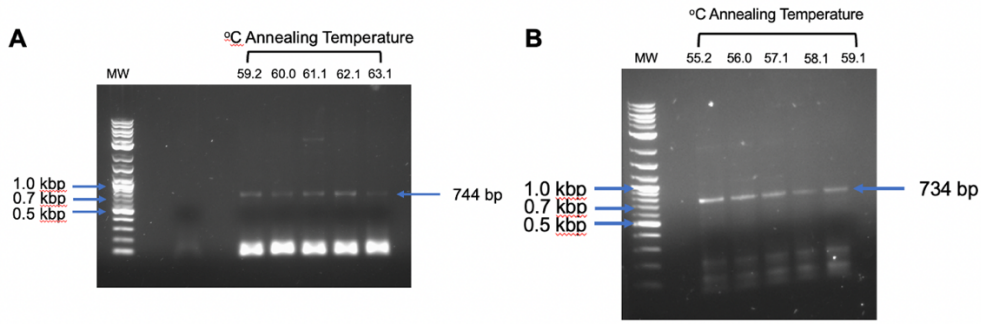


Figure 5.6 Gradient PCR showing amplification of TRα1, TRβ1 LBD at different annealing temperatures. Following amplification for 30 cycles (final volume 50ul), 5ul was subject to gel electrophoresis. Unincorporated primers are also visible.

To assess whether GC content was a factor, I performed PCRs for the TRα1 product under variation of DMSO content. Dimethyl sulfoxide (DMSO) is an organic solvent that acts to reduce the melting temperature (Tm) of GC-rich regions and destabilizing secondary structures in GC-rich DNA. The sequence of TRα1, TRα2 were discovered to hold a high percentage of GC (57% and 60% respectively) which might require a higher amount of DMSO. However, the percentage of GC in TRβ1 was only occupied 50%, which might

require a lower level of DMSO to be used. Thus, we set up a series of DMSO concentrations to optimize the PCR of TRα1, TRα2 and TRβ1. As shown in Figure 5.7 in the conditions used as described for amplification of TRα2 LBD, PCR yield was found to be considerably improved at the highest concentration of DMSO (6%). However, it should be noted that only 2.5 ng/υl of template was used in this experiment so it cannot be compared to other control experiments that used a different amount of starting template. It was later pointed out by my supervisor that optimizations of PCR conditions must be done more systematically, varying only one condition at a time.

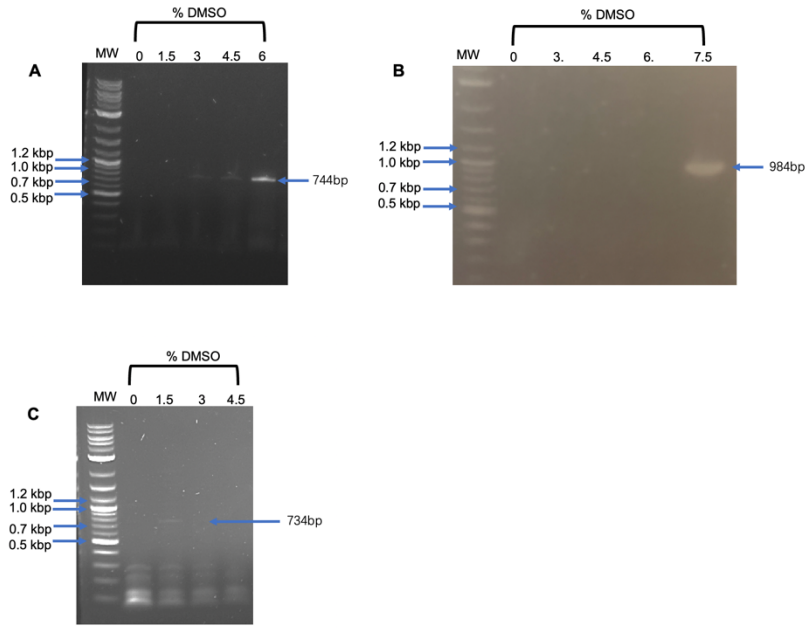


Figure 5.7 Effect on varying DMSO concentration on yield of the TRα1 (A, TRα2 (B and TRB1 (C LBD PCR products. Note that the amount of template used in this assay was lower than in Figure 5.6.

Another choice of optimization for TRβ1 PCR process was the concentration of magnesium ions. Magnesium ions are important in

many areas of cellular metabolism. They bind to the surface of macromolecules to maintain the structures of proteins, nucleic acids, and cell membranes [165-167] and enhance certain structural or catalytic activities of proteins, enzymes, or ribozymes. They also play a role in enzymatic processes in a variety of ways. They may form magnesium-substrate scaffolds to which enzymes attach, bind directly to enzymes, and modify their structure, and guide processes via particular catalytic functions. These crucial ions are used by DNA polymerases to duplicate and repair DNA. Different concentration of magnesium ions were tested, which normally needs low levels to facilitate binding of the polymerase and DNA of interest because high ionic strength inhibits the interaction. To optimize the concentration for magnesium ions, the magnesium free buffer was used instead of the standard Phusion buffer. The magnesium chloride was then added to the magnesium chloride free reaction buffer to achieve a final concentration of 0 - 4.5 mM in reaction. After that, the similar recipe of PCR reaction was set up with 10 ul 5X buffer, 1 ul 10 mM dNTP, 2.5 ul 10 uM TRβ1 Forward Primer, 2.5 ul 10 uM TRβ1 Reverse Primer, 25 ng cDNA and 0.5 ul Phusion polymerase. It was denatured at 98 ^{o}C for 5 mins and run on a short cycling reaction (98 ^{o}C for 10s, 55 ^{o}C for 30s, 72 ^{o}C for 90s ; 30 cycles). As shown in Figure 5.7, The DNA bands were much brighter when applying with 0 to 1.5 mM magnesium ions. However, there were no bands between the concentration higher than 1.5 mM for magnesium ions in TRβ1 amplification.

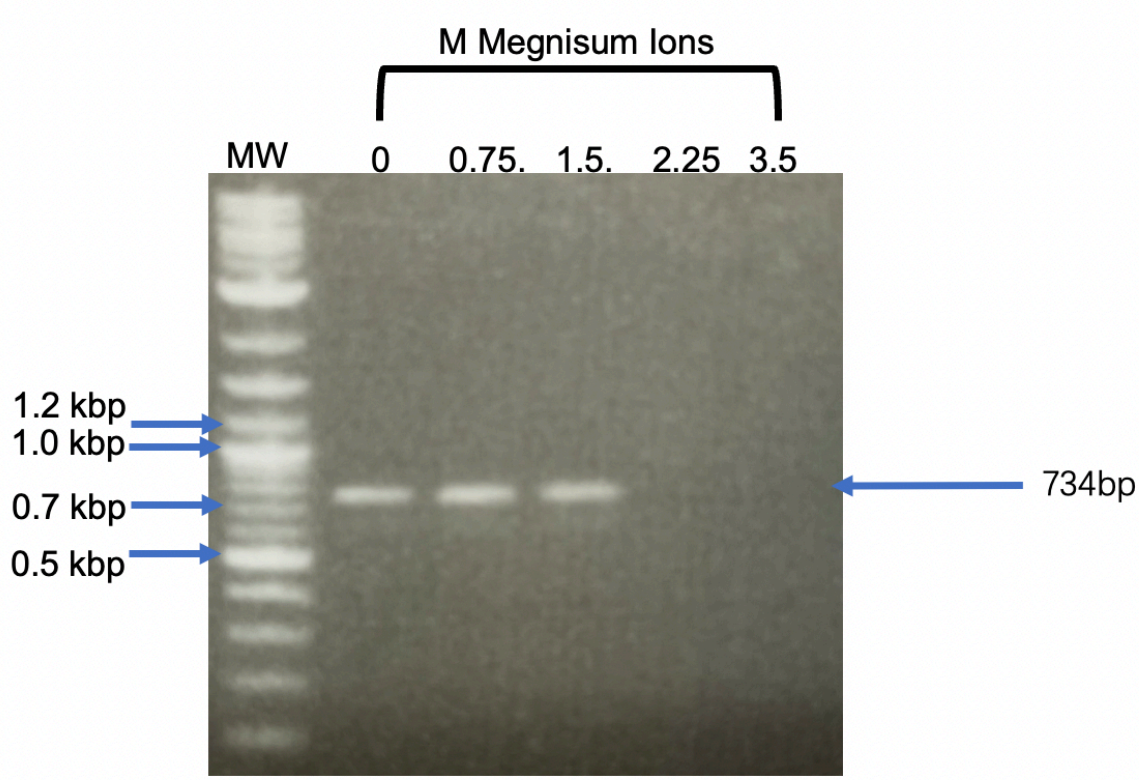


Figure 5.7 Effect on varying Magnesium ions concentration on yield of the TRβ1 LBD PCR products.

In conclusion, the best condition for TRα1 was 6 % DMSO with 61 ^{o}C annealing temperature. The amplification of TRα2 worked better in 7.5 % DMSO with 61 ^{o}C annealing temperature. The amplification of TRβ1 was much more effective in 1.5 % DMSO and 0-1.5 M magnesium ions with 55 ^{o}C annealing temperature. The conditions, that were not specified, were followed the protocol in section 2.3.5.

At this point, cDNA stock, provided from Ryan Morris, had run out. Therefore it was decided to prepare cDNA from HEK293 cells (which express abundant TRα1 and TRα2 but not TRβ). In addition, RNA Seq data indicated a single A deletion in TRα2 transcripts present

in LNCaP C4-2. LNCaP C4-2 cells cDNA was prepared for for amplification of TRβ1

LNCaP C4-2 and HEK293 cells were provided by Ryan Morris and Luc Costello Heaven and maintained in DMEM, supplied with 10% (v/v) heat inactivated FCS, 1% (v/v) 0.2 M L-glutamine and 1% (v/v) Penicillin/Streptomycin antibiotics. The 16th generation of LNCaP C4-2 cells and 21st generation of HEK293 cells were used for RNA extraction. First, cells were washed down by 5 ml DMEM. 10 ul of each culture was put on the counting board, showing that there were 6.9 X 10^5 cells per ml for HEK293 and 4.5 X 10^5 cells per ml for LNCaP C4-2 cell line. Following RNeasy Mini Kit (Qiagen) protocol, yields of about 1 υg RNA of each were obtained.

PCR analysis successfully amplified TRα2 sequence from HEK293 cDNA cells whereas the amplified band for TRα1 was very weak, possibly due to lower abundance. The TRβ1 LBD sequence from LNCaP C4-2 cDNA was also amplified as shown in Figure 5.8. The conditions were shown in Table 5.1.

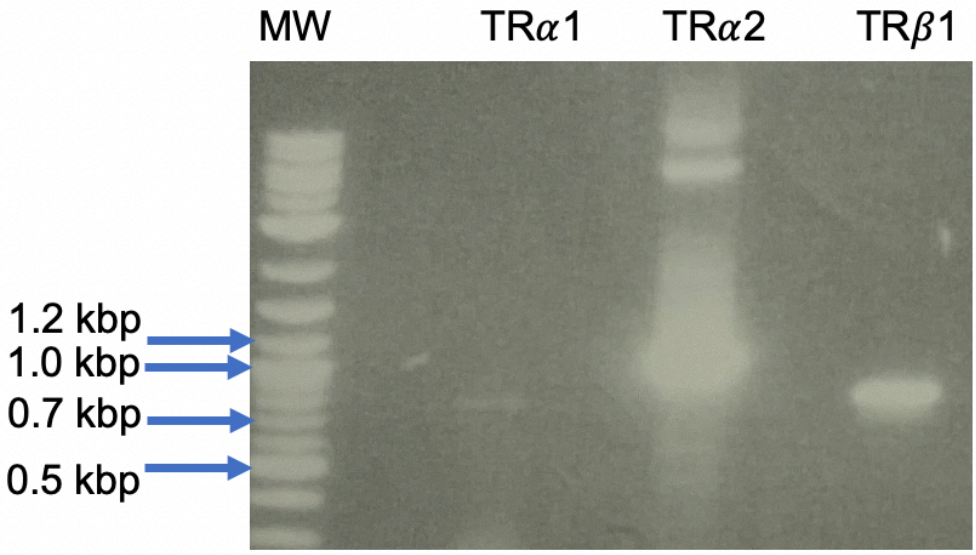


Figure 5.8 PCR amplification of TRα1 and TRα2 LBD sequences from HEK293 cDNA and TRβ1 primers for LNCaP C4-2 cells. The expected sizes of TRα1, TRα2 and TRβ1 were 744 bp, 984 bp and 734 bp.

Table 5.1 PCR conditions using to amplify TRα1, TRα2 and TRβ1. The conditions, not specified, were followed section 2.3.5

Sequence	Annealing Temperature (°C)	DMSO (%)	$MgCl_2$ (mM)
TR α 1	61	6	1.5
TR α 2	61	7.5	1.5
TR β 1	55	1.5	0.75

5.3 Subcloning TR LBDs into the pCR4-Blunt-TOPO Cloning Vector

Having successfully amplified PCR products (at low yield) for the LBDs of TRα1, TRα2 and TRβ1, fragments were purified from agarose gel slices using Qiaprep kit and digested with EcoRI/XhoI, EcoRV/NotI and EcoRI/NotI respectively. Vector pYC156 was digested with the same enzyme combinations, purified from gels. Several attempts were made to ligate these vector/ insert combinations. Although ampicillin resistant clones were obtained, none of these had the inserts as expected. To understand why the subcloning was unsuccessful, the vector from stocks was sent for sequence analysis. Unfortunately, this revealed that the pYC156 vector sequence logged into the laboratory stocks (10029) was not as expected as the multiple cloning sites did not contain the restriction sites in the correct order to allow cloning. Several other stocks of this vector were also found to be incorrect as they contained NLS sequences before the YFP sequence. These were likely to be additional split YFP vector variants provided from the Nick Holliday lab that were incorrectly labelled (10030).

As an alternative approach, due to the low yield of PCR fragments it was decided to attempt to clone the TR LBD PCR fragments into a TA cloning vector. pCR4 Blunt Topo vector kit was purchased from ThermoFisher. The map is shown in Figure 5.9.

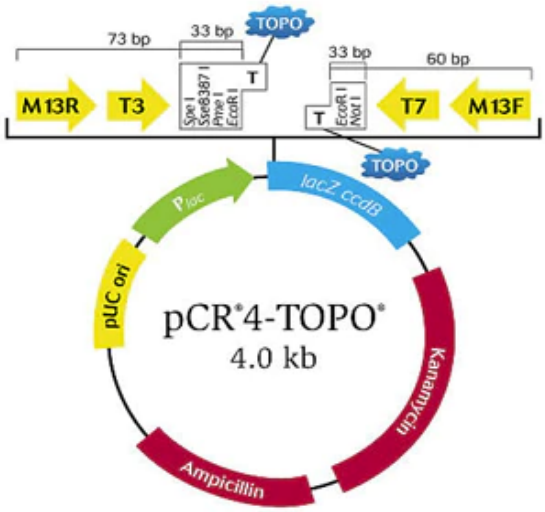


Figure 5.9 vector map of TA cloning vector pCR4TOPO

The TOPO TA cloning system was used to subclone the TRβ1 PCR fragment using a modified protocol (CJ Roberts) of that described by the manufacturer. The pCR4 Blunt TOPO treated vector (0.5 µl) was mixed with Salt solution (in kit) 0.5 µl and the Phusion polymerase PCR product (2.5 µl). Reaction mixtures were incubated at room temperature for 15 minutes, then moved to ice, and used in transformation reactions with chemically competent *E. coli* DH5α. Transformation without plasmid was taken as a control. After overnight growth at 37 ºC on LB Ampicillin (100 mg/ml), more than 50 colonies were obtained. Colony PCR analysis for each plasmid were performed on eight of these clones using the TRα1, TRα2 and TRβ1 primers, revealing that most clones tested positive for the insert, as shown in Figure 5.10.

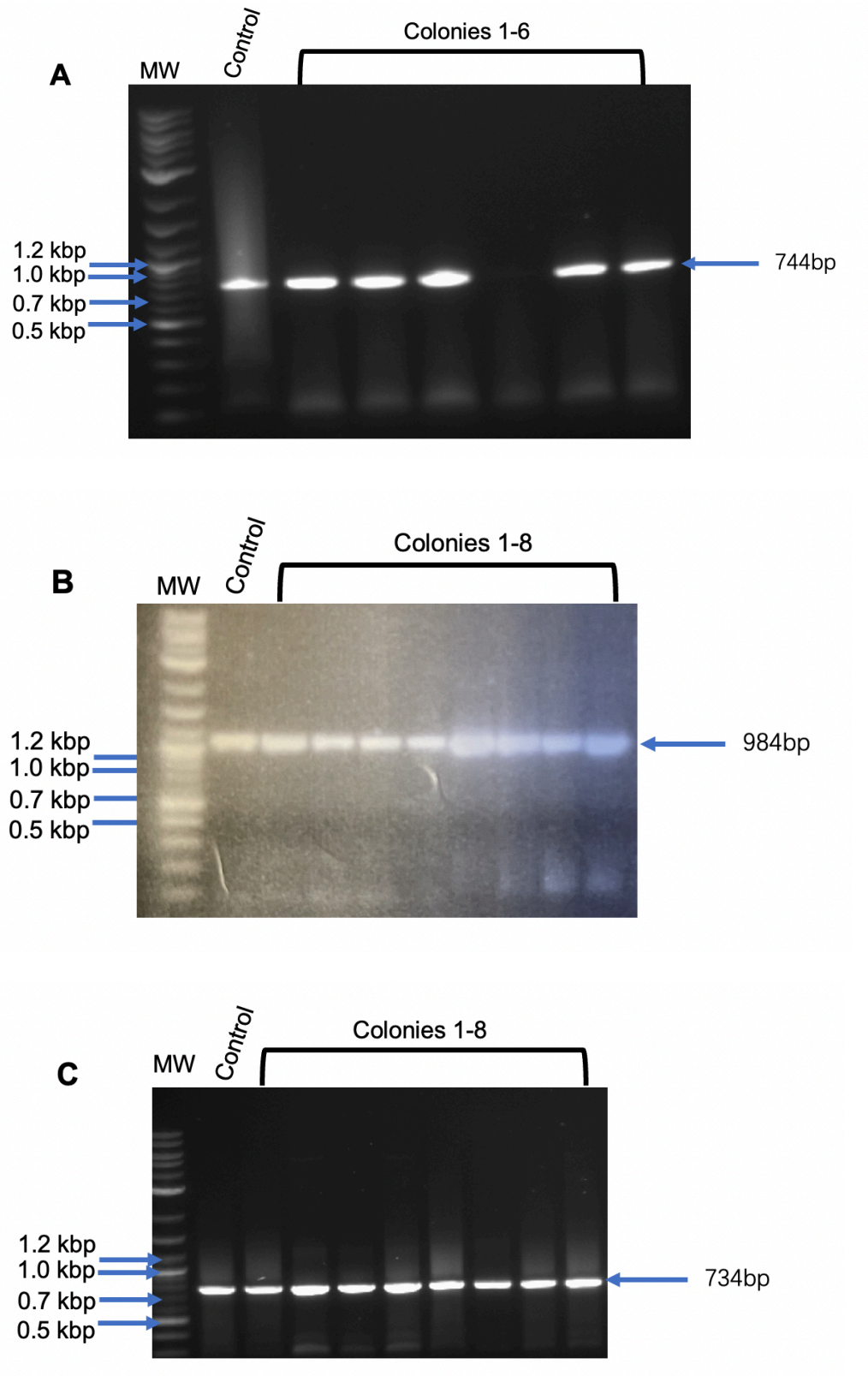


Figure 5.10 Colony PCR showing inserts for the pCR4-Blunt-TOPO TRα1 (A, TRα2 (B and TRβ1 (C LBD clones

Having obtained positives (by colony PCR) for TRα1 and TRα2 in TOPO vector, several clones were sent for sequencing. This confirmed the presence of the EcoRV/NotI and EcoRV/XhoI sites flanking the inserts as designed. Unfortunately, BLAST search of the insert revealed the presence of a single nucleotide deletion that causes a frame shift at position 176 in the LBD, as shown in Figure 5.11. This was also present in the other clones. In case this was due to a PCR error, a second PCR was performed and cloned in pCR4-TOPO. Once again, the same deletion was encountered. Thus, we concluded that C4-2 cells may carry a frameshift mutation in the TRα gene.



Figure 5.11 BLAST search alignment indicating the deletion of an adenosine nucleotide in TRα transcripts in C4-2 cells.

Given the mutations in the TRα1 and TRα2 LBD constructs, it was decided to focus on the TRβ1 LBD, as this was thought to be more likely to be able to interact with the AR.

5.4 Ligation of plasmid TRβ1 LBD - vYFPC156

Therefore, I next attempted to clone this in the split YFP system. As we had difficulty locating the correct empty vector, sequence analysis of another construct in the stocks ERα-LBD-YC156 (#10030) confirmed it to have the correct vector sequence with an insert between EcoRI and XhoI sites. Restriction analysis of this plasmid with EcoRI and XhoI released a 0.9kb insert as expected, as shown in Figure 5.12. TRβ1 LBD fragment was released by digesting the TRβ1-vYFPC156 plasmid with EcoRI and XhoI as shown in Figure 5.13. Following gel purification, ligation of the TRβ1 LBD and vYFPC156 was performed using T4 ligase.

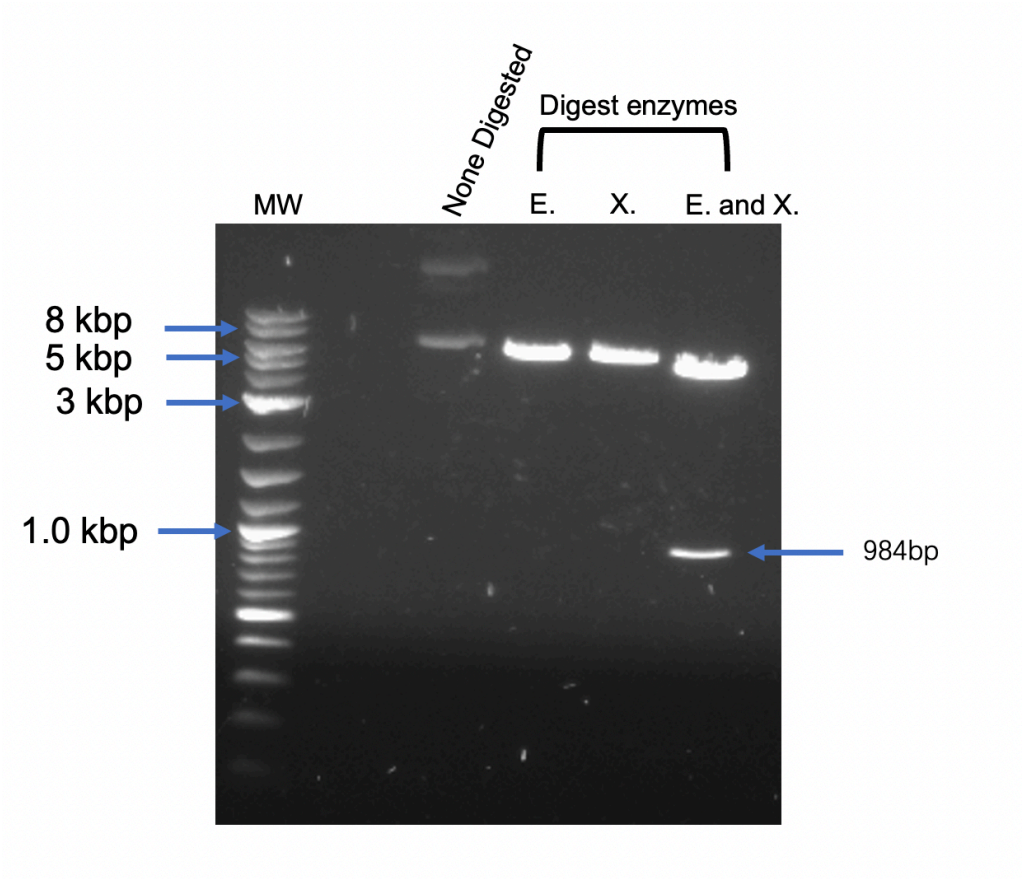


Figure 5.12 Restriction digest check for ERα-LBD-YC156 using EcoRI-HF and XhoI.

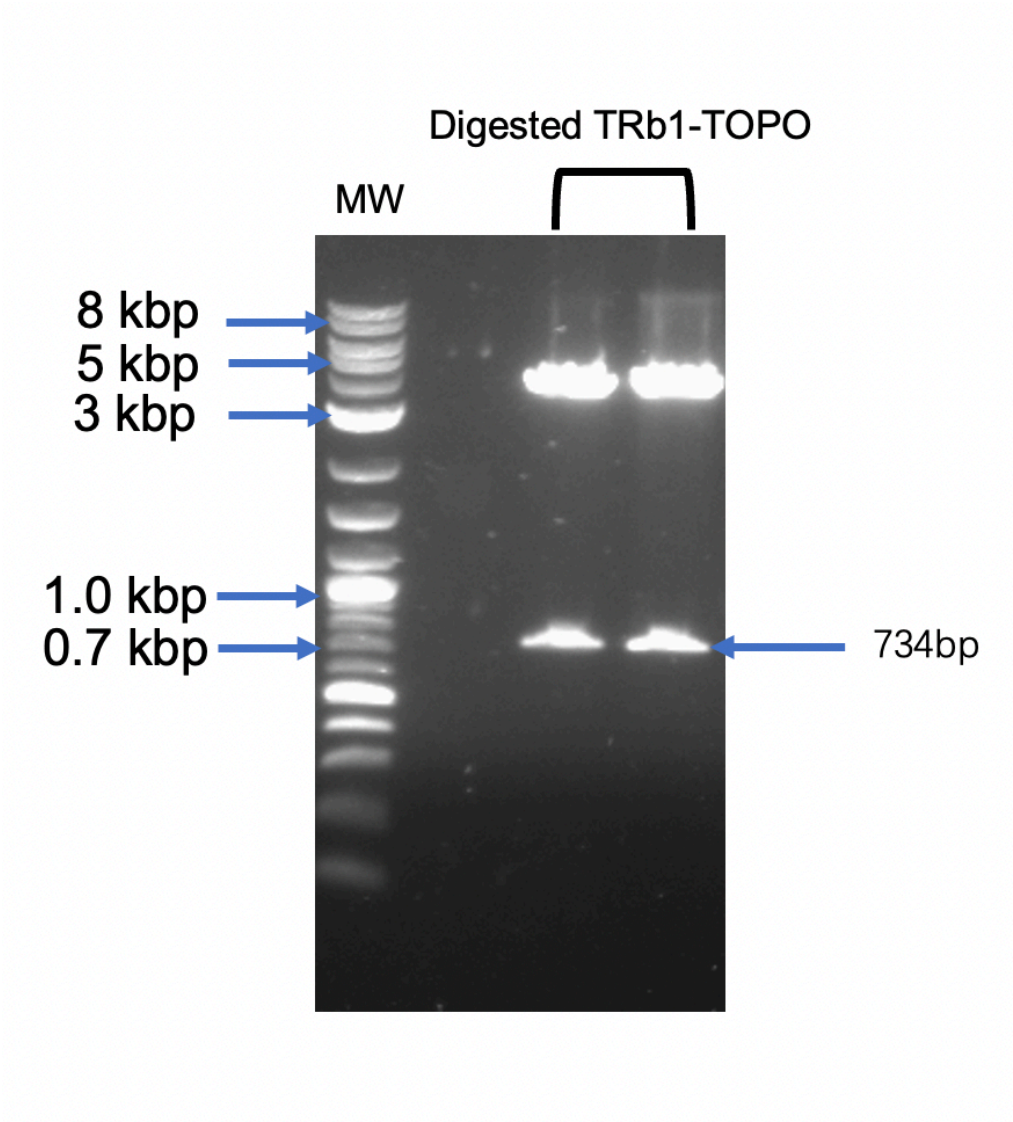


Figure 5.13 Restriction digest for TRβ1-TOPO using EcoRI-HF and XhoI.

Colony PCR showed positive results using TRβ1 primers, but sequencing revealed that the protein product had been contaminated with plasmid TRβ1-TOPO. Following the examination of several techniques, two strategies were shown for the removal of contaminating DNA. To identify the source of the contamination, several controls were set up including transformation of the digested vYFPC156, the digested TRβ1, T4 DNA polymerase and no plasmid. Unfortunately, all of the controls showed negative. Rather than the contamination from reagent,

it seemed that DNA was contaminated by equipment, where the problem was solved by replacing the pipette.

However, the number of colonies were much lower than in normal transformation after one day culture on LB plates and still none of them were shown to be consistent with the plasmid TRβ1 - vYFPC156 according to colony PCR using CMD forward and vYFPC156 reverse primer. To optimize the ligation process, the ratio of insert and vector, ligation efficiency of T4 ligase and the total amount of DNA to be used were considered to be important factors for the ligation process.

From here, the digested TRβ1 was obtained by digesting the TRβ1 product. 50 μl PCR was set up, using 25 ng cDNA from LNCaP C4-2 cells, following the optimized PCR condition in Section 5.2. 10 μl loading dye was added into PCR product and 20 ul of the mixture was running on each well of the 0.8% agarose gel, as shown in Figure 5.14. I cut the band out and extracted TRβ1 through gel extraction kit. About 1 ug TRβ1 was yielded, measuring by Nanodrop. The TRβ1 was incubated with 1 ul EcoRI-HF, 1 ul XhoI, 5 ul Cutsmart and water up to 50 ul, in 37 ^{o}C water bath overnight. Later, the quality of digested TRβ1 and vYFPC156 was validated by running 100 ng on the agarose gel as shown in Figure 5.15. The rest of the material was purified the concentrated by ethanol precipitation which could precipitate DNA, separating them from other water-soluble impurities.

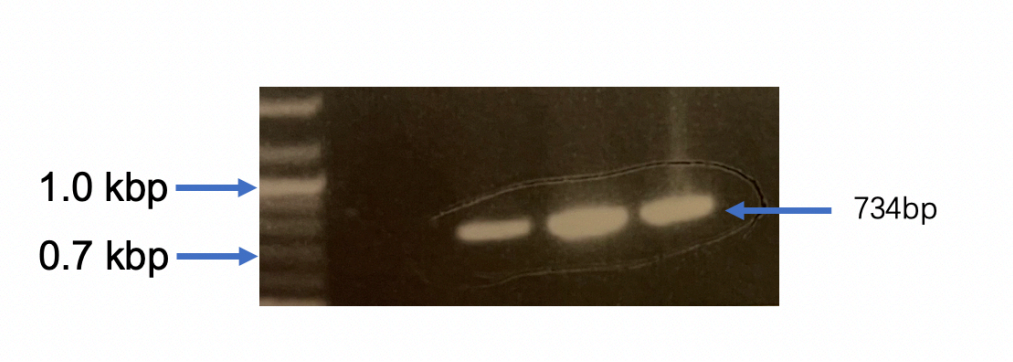


Figure 5.14 PCR results for TRβ1 LBD used for ligation with vYFPC156

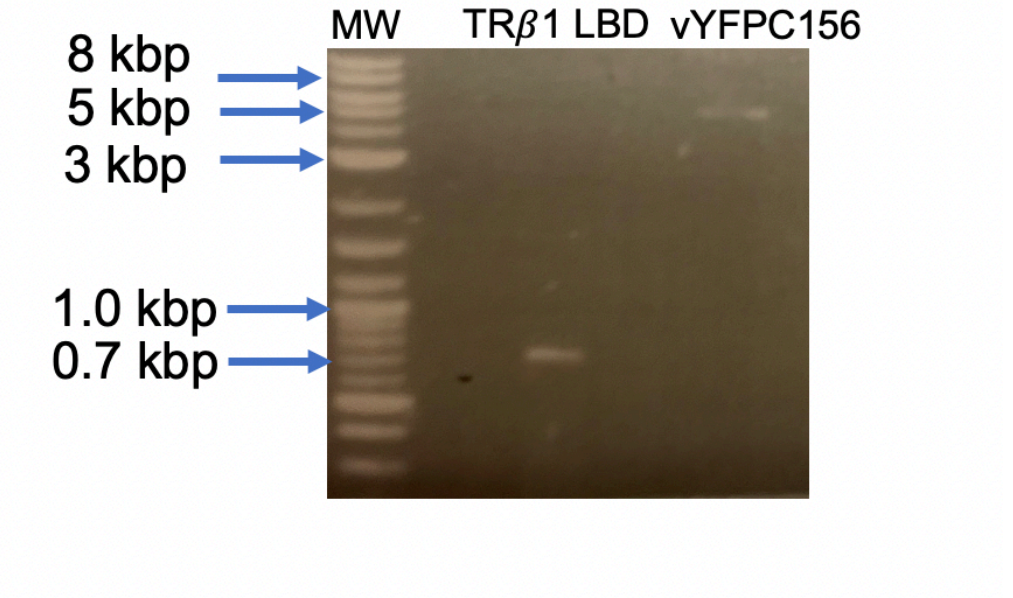


Figure 5.15 The digested TRβ1 LBD and vYFPC156

Furthermore, it was investigated if the amount of DNA, being used, could influence the efficiency of T4 ligase. Thus, we used ligation of 300ng of 1 kbp ladder, ligation of TRβ1 and vYFPC156 with total amount of DNA of about 300 ng and 30 ng. They were incubated in a thermal cycling reaction (10 ^{o}C for 30 min, 30 ^{o}C for 30 min) The products were tested for DNA content using agarose gel electrophoresis (Figure 5.16A). Figure 5.16 showed a direct comparison of the 1 kbp ladder and the ligated 1 kbp ladder (300 ng) that was under the same condition of the ligation for TRβ1 - vYFPC156. The band for

ligated 1 kbp ladder was above the unlighted 1 kbp ladder indicating that the T4 ligase worked efficiently. Higher amount of total DNA (100 ng TRβ1 and 200 ng vYFPC156 with insert: vector ratio of 3:1) in ligation process resulted in lower efficiency in ligation, where bands were quite smeared, as shown in Figure 5.16A. As a result, a lower amount of total DNA (11 ng TRβ1 and 24 ng vYFPC156 with insert: vector ratio of 3:1) was used for the ligation process, as shown in Figure 5.16B. The digested TRβ1 formed dimers at the size around 1.5 kbp. And there were several bands around 6 kbp which might refer to the bands of vector only and TRβ1 - vYFPC156.

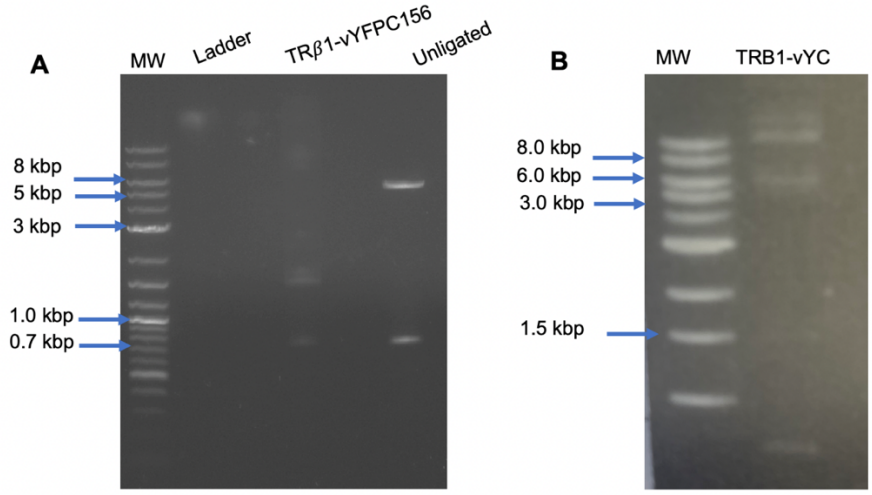


Figure 5.16 A) 300 ng DNA input for ligation of Ladder and TRβ1 - vYFPC156. Right lane: The same amount of digested TRβ1 and vYFPC156 without T4 ligase. B) 30 ug DNA input for ligation of TRβ1 - vYFPC156. The ligation is processed by mixing the specific insert, vector (with the ratio 3:1), 0.5 ul T4 ligase, 1 ul ligase buffer and water to make a final volume of 10 ul. the mixture was incubated at 90 times (10 °C for 30 seconds and 30 °C for 30 seconds)

The ligation solution (Figure 5.16B) was transformed to DH5α. And the further colony PCR using primer CMD forward and TRβ1 reverse confirmed the presence of TRβ1 and vYFPC156 in plasmid, as shown in Figure 5.17.

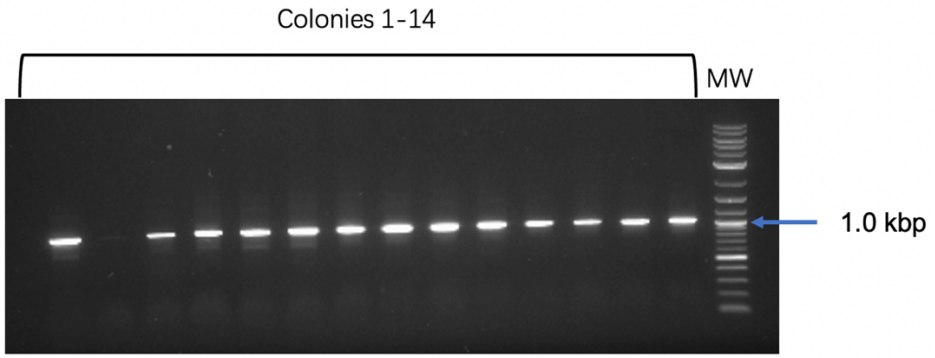


Figure 5.17 Colony PCR of clones TRβ1 - vYFPC156, using primers CMD forward and TRβ1 reverse. CMD is a part of the vector vYFPC156. The aim size of the amplified sequence is 999 bp

Several positive colonies were miniprepped and sent for sequence. All colonies showed correct sequence which is in frame with vYFPC156, as shown in Figure 5.18.

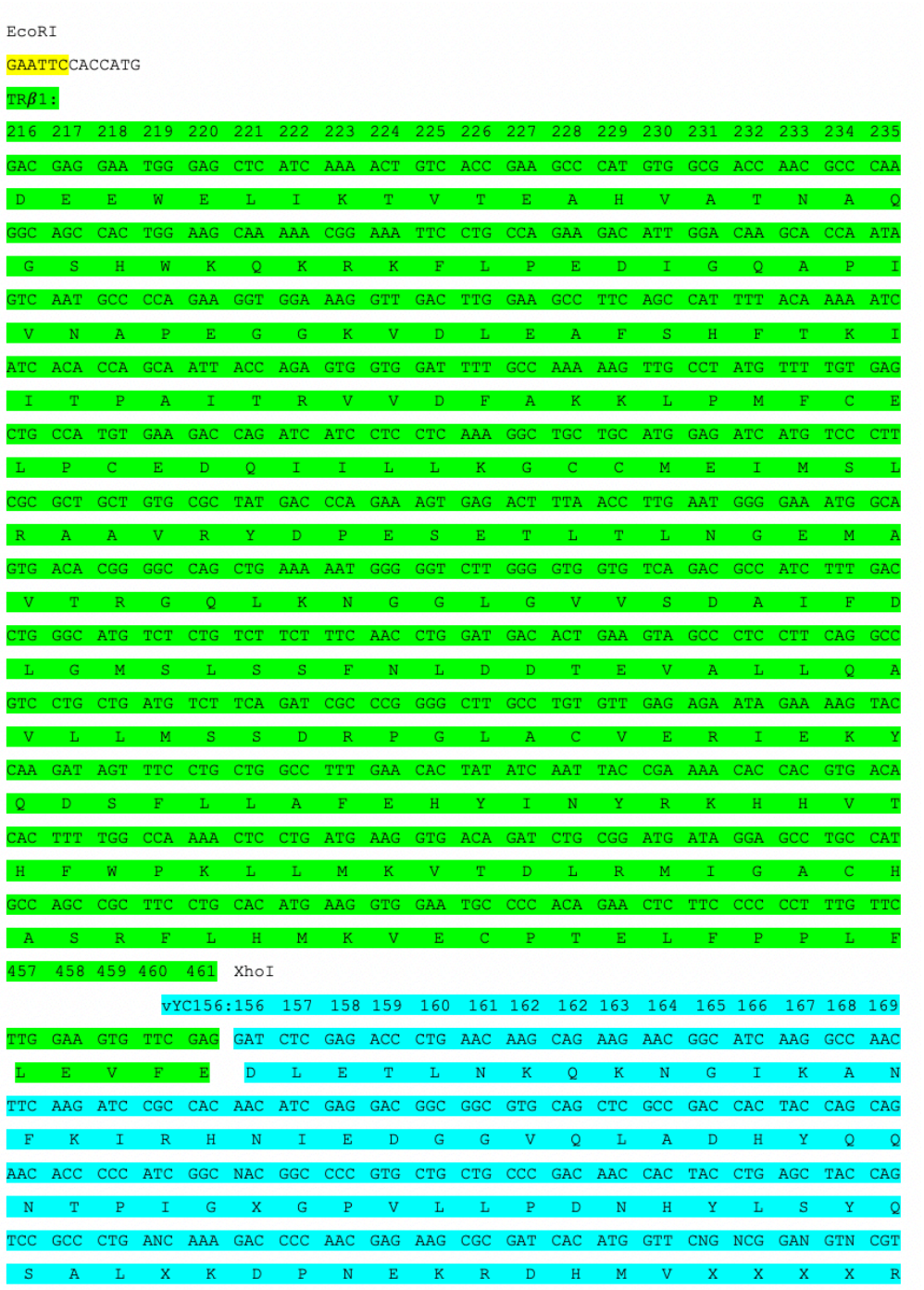


Figure 5.18 The plasmid sequence of constructed TRβ1-vYFPC156. TRβ1 LBD was shown in green while vYFPC156 was shown in cyan

5.5 Bimolecular fluorescence complementation

After successfully fusing TRβ1 LBD in vYFPC156 vector, it was employed to investigate its binding ability towards AR LBD in mammalian cells. As explained below, confocal imaging of the HEK293 cells, co-transfected with AR-LBD-vYFPN155 and TRβ-LBD-vYFPC156, was employed to test for the interaction between LBD domains of AR and TRβ. Transfections require relatively high amounts of DNA, therefore maxi-preps of AR LBD - vYFPN155, TRβ LBD – vYFPC156, MED1 LXM1 - vYFPC156 and SRC1 NID - vYFPN155 were prepared using QIAGEN plasmid Maxi kit. The pelleted DNA was washed by 5 ml 70% ethanol and centrifuged at 4000 rpm for 5 mins, air dried and redissolved in water. The DNA concentration of each plasmid was adjusted to 1 μg/μl. This yielded 137 μl AR LBD - vYFPN155, 285 ul SRC1 NID - vYFPN155, 456 μl MED1 LXM1 - vYFPC156 and 72 μl TRβ1 LBD - vYFPC156. The 260 nm/280 nm ratios were 1.87, 1.89, 1.74 and 1.77, respectively while the 260 nm/230 nm ratios were 2.04, 2.07, 2.01 and 2.03. 300 ng of each plasmid was running with Agarose gel electrophoresis and it revealed substantial quantities of the plasmid of interest (Figure 5.19).

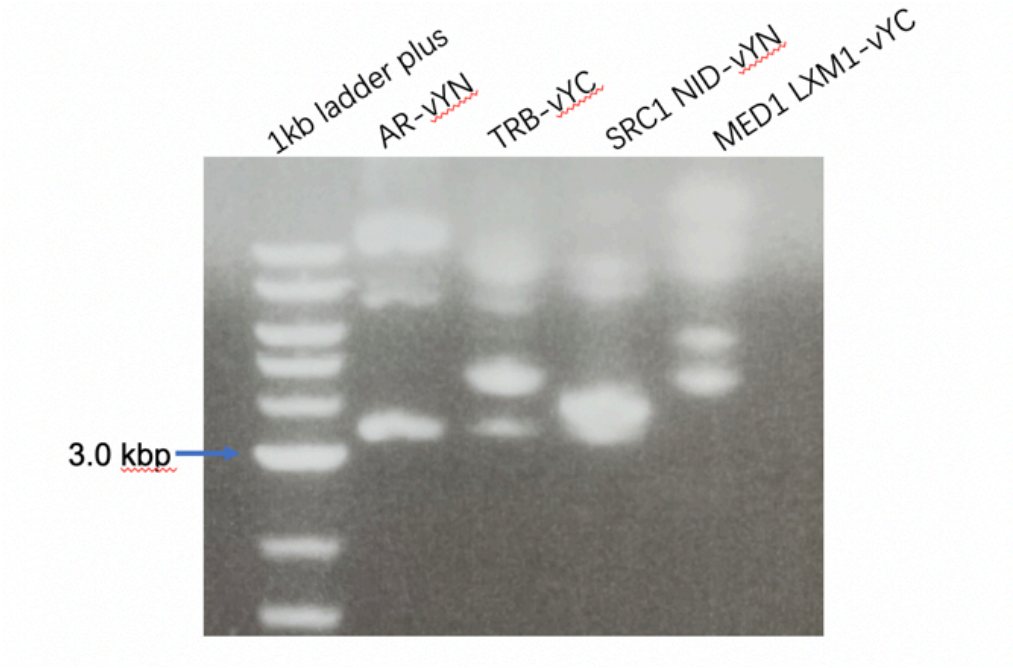


Figure 5.19 Agarose gel showing maxiprep plasmid of AR - vYN, TRβ - vYC, SRC1 NID - vYN and MED1 LXM1 - vYC

To optimize transfection conditions of HEK293 cells, different DNA concentrations (from 25 to 300 ng) were used for the co-transfection procedure. Full length GFP was used as a positive control for cell transfection- this resulted in a quite high efficiency detection of GFP+ cells and the GFP signal was present in both nucleus and cytoplasm. Cotransfection of AR LBD - vYFPN155 and TRβ1 LBD-vYFPC156 was performed in duplicate in 6 well plates in the presence or absence of 100nM cognate ligands. Mibolerone (mib) (a stable form of androgen) was used as an agonist for AR LBD while T3 was used as a ligand for TR LBD. As a control, AR LBD - vYFPN155 was co-transfected with MED1 LXM1 - vYFPC156 in the presence or absence of 100 nM mibolerone. Similarly, TRβ1 LBD-vYFPC156 was co-transfected with SRC1 NID - vYFPN155 in the presence or absence of 100 nM T3.

After 24 hours, the plates were examined for fluorescence complementation. Unfortunately, in this preliminary experiment, we did not detect any evidence for interaction of AR and TRβ1 LBDs in these conditions. However, the expected interaction between the AR LBD and MED1 LXM1 motif, did not yield any fluorescent cells either. Thus the positive control for AR LBD failed, thus we cannot make a safe conclusion about the formation of AR and TRβ1 LBD dimers. The reason for this is unknown, but it may simply require more careful replications of the experiment as AR LBD/ MED1 LXM1 interactions have been observed previously in our laboratory (Joel Fulton PhD thesis). As shown in Fig 5.20, the successful detection of YFP fluorescence in the nuclei of several cells indicated the T3-dependent interaction of the TRβ1 LBD with the SRC1 NID domain. Fluorescence was not observed in the absence of T3. This confirms the functionality of the TRβ1 LBD-vYFPC156 construct which is capable of recruiting a cofactor in response to ligand and allows split vYFP complementation. Further optimization of the protocol will be required to detect the function of the AR LBD in recruitment of cofactors or homo-heterodimeric interactions.

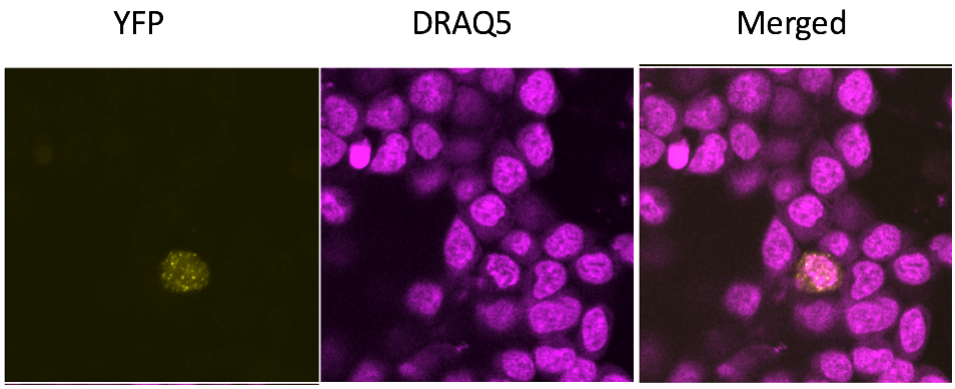


Figure 5.20 Complementation of SRC1 NID - vYFPN155 and TRβ1 LBD-vYFPC156 in the presence of 100 nM T3. SRC1 NID - vYFPN155 and TRβ1 LBD-vYFPC156 constructs were expressed in HEK293 cells and examined for YFP expression using confocal microscopy. Cells were stained by DRAQ-5.

5.6 Summary

The aim of this part of the project was to generate split vYFP fusion vectors to test interactions of AR /TR LBD heterodimers in mammalian cells. Primers were designed to generate TRα1, TRα2 and TRβ1 LBD fusion constructs in vYFPC156 vector. It was necessary to optimize the conditions for PCR from cDNA total, PCR conditions and cloning strategies. Human errors, contamination and technical issues meant it took several months to successfully achieve the desired clones. Initial PCR attempts, following the recommended conditions to amplify products using 25 ng/ul cDNA as input and an annealing temperature at the primer Tm, with a short cycling reaction (98 ^{o}C for 10 s, Tm for 30 s, 72 ^{o}C for 90s ; 30 cycles) were either unsuccessful or generated low yield of product. Thus, the amplifications of TRα1, TRα2 and TRβ1 were

optimized by testing the temperature and percentage of DMSO and magnesium ions. After several experiments, it was found that 6%, 7.5% and 1.5% DMSO might be beneficial to amplify TRα1, TRα2 and TRβ1, respectively, which was consistent with the percentage of GC in each sequence. Similarly, TRα1 and TRα2 required higher annealing temperature (61 ^{o}C) than TRβ1 (55 ^{o}C). Moreover, the amplification of TRβ1 preferred lower lever of magnesium ions (0.75 M).

The TRα1, TRα2 and TRβ1 LBD sequences were successfully cloned into TOPO vector. However, the sequence result suggesting a single base pair (A) deletion in the sequence of TRα2 present in the original cell lines or arising as a recurrent PCR artifact. As a result, we focused on the interaction between TRβ1 and AR only.

To clone the TRβ1 into vYFPC156, we used ERα LBD - vYFPC156 to release the vector and PCR amplified TRβ1. Several attempts were made to optimize the ligation reaction. Finally, the plasmid was successfully made by mixing the 30 ng total DNA (with the ratio of 3:1 for insert and vector), 0.5 μl T4 ligase, 1 ul ligase buffer and water to make a final volume of 10 μl. The mixture was incubated at 90 times (10 ^{o}C for 30 seconds and 30 ^{o}C for 30 seconds).

After confirmation of inserts by colony PCR and sequence analysis, the split vYFP AR LBD and TRβ LBD constructs were maxiprepped, along with vYFP cofactor controls (SRC1 NID; MED1 LXM1) for testing in co-transfections. The amount and quality of obtained plasmids were adequate to carry out the transfection tests to investigate the interaction between AR and TRβ1. Importantly, it was able to confirm that the TRβ-

LBD was functional in terms of interaction with cofactor LxxLL motifs in response to ligand, as was shown by successful fluorescence complementation. Unfortunately, there was not sufficient time to establish why the AR-LBD failed to show interactions with TRβ or the control MED1 LXM1, which was previously done in the lab. Sequencing confirmed the AR construct was correct and in frame. However, future experiments will be able to use these vYFP-TRβ1 reagent retest AR and TR interactions in this system.

Chapter 6 Yeast two hybrid assays

Fields and Song [168] created the yeast two-hybrid method as a sensitive genetic tool for detecting protein-protein interactions. The yeast two hybrid system was designed, according to following principles. Transcription Factors contain at least two distinct domains: a DNA binding domain (DBD) and a transcription activation domain (Activation domain, AD), which when separated can function independently [169]. However, DBD and AD expressed as separate domains are unable to activate the transcriptional response of a reporter gene. While the DBD can bind the promoter of a reporter gene, it has no ability to activate transcription. Similarly, while the AD domain is able to interact with other components of the RNA Polymerase transcriptional machinery, without the DBD it cannot be targeted to the reporter gene. However, if AD domains are brought into close contact with the DBD on a reporter gene promoter then transcriptional activation can be achieved. This can be exploited to test for protein-protein interactions. Thus, any pair of polypeptide sequences that can interact with each other may be fused to the DBD or AD domains, utilized to reassemble the individual DBD and AD to form a functioning trans activator. The two proteins to be studied (protein X and protein Y; or bait and prey proteins) are fused to the DBD and AD domains, respectively by molecular cloning., and co-expressed in yeast cells. If there is no interaction between the two proteins, the reporter gene will not be transcribed and expressed; if there is an interaction between the two proteins, the DBD and AD

domains are spatially close resulting in expression of the reporter. This approach has been used in our group to study protein-protein interactions of NRs such as the AR LBD and AR LBD variants[170]. It can be applied to study homo- or heterodimer complexes or interactions with cofactors.

Yeast two-hybrid plasmids carry replication maintenance sequences (centromeric or 2μl sequences), antibiotic selection sequences for *Escherichia coli* (e.g. ampicillin resistance) and auxotrophic selection markers (e.g. TRP1, LEU2) for selection in Saccharomyces cerevisiae. They are used in conjunction with a yeast reporter strain containing one or more reporter genes (typically HIS3 and lacZ) which have been engineered to contain binding sites for the DBD. The yeast transcription factor GAL4 [168, 171, 172] or the bacterial repressor protein LexA are most commonly used to generate DBD vectors [98, 173, 174]. This project employs the LexA-containing vector, pBTM116mod, a derivation of pBTM116 [98 175]. The DBD sequence in the vector is followed by a multiple cloning site (MCS), which enables the bait cDNA to be inserted in the vector in frame with the DBD, allowing expression of a DBD-fusion protein. This plasmid is selected in yeast via a TRP1 gene marker.

The most commonly used ADs in two-hybrid systems are the ADs of the yeast GAL4 protein, the AD of the herpes virus protein, VP16, or the B42 AD, an activating sequence from *E. coli*. Fusion proteins containing these ADs are created by subcloning suitable cDNAs in-frame into the MCS. This project employs pASV3mod, a

pASV3 [176] variant that includes the VP16 AD and a LEU2 selective marker. Many different prey sequences have been generated in our group for use in two-hybrid assays. These include VP16-fusions to full length TRβ1, TRβ1 LBD and AR LBD.

The yeast reporter strains used for two-hybrid screens include one or more auxotrophic markers. In descriptions of yeast strain genotypes capital letters indicate a functional gene, whereas lower case letters indicate an inactivate gene. Thus HIS3, URA3, TRP1, LEU2 are commonly used as selective markers in yeast cells carrying his3, ura3, trp1, leu2 auxotrophic markers. These genes are involved in essential amino acid synthesis or uracil synthesis and are thus lethal mutations when the yeast is grown on media lacking these nutrients. Plasmids carry functional genes so can complement the deficiency on selective media.

The L40 reporter strain was used in this project [177]. The L40 genotype is *MATa his3Δ200 trp1-901 leu2-3112 ade2 LYS::(4lexAop-HIS3) URA3::(8lexAop-LacZ)GAL4.* This means it is defective in trp, leu, ade and his synthesis. Two reporter genes HIS3 and lacZ regulated by Lexa binding sites were integrated into the LYS and URA3 genes, respectively, inactivating them. Thus, reporter activation can be measured by b-galactosidase assays or growth on histidine deficient medium. The TRP1 and LEU2 markers are respectively selective for yeast transformants carrying pBTM116mod and pASV3mod. Transformants thus have capacity to grow on leucine and tryptophan deficient medium. These colonies are purified by restriking on selective

medium and cultured for reporter assays as described in the Methods section.

Previous yeast two hybrid assays in our lab have used yeast two hybrid to investigate AR LBD homodimers[178] or heterodimer formation between NR LBDs[170] In addition unpublished results have identified interaction between AR LBD and TRβ LBD which was dependent on the presence of both ligands mib and T3. Reporter activation was shown to be ligand dose dependent. The aim of this part of the project was to independently reproduce these data and explore the effects of mutations in the AR LBD on these interactions.

6.1 X gal Reporter assays

We used yeast two hybrid method to test the interaction between AR and TRβ LBDs. Expression plasmids are listed in Table 6.1. Wild type AR LBD and AR LBD variants were fused to LexA DBD using the pBTM116mod vector, while wild type AR LBD and TRβ LBD were fused to the VP16 AD in pASV3mod vector. The LexA- AR LBD and VP16-TRβ LBD were co-transformed into L40 yeast. LexA- AR LBD and VP16- AR LBD homodimer interactions were also tested as a positive control

Table 6.1 Plasmids being used in yeast methods

Number	Plasmid	Origin
10094	pBTM116mod-AR LBD T851D	Dr. Jonathan Whitchurch

10095	pBTM116mod-AR LBD V770A	Dr. Jonathan Whitchurch
10096	pBTM116mod-AR LBD W752R	Dr. Jonathan Whitchurch
10097	pBTM116mod-AR LBD F755S	Dr. Jonathan Whitchurch
10098	pBTM116mod-AR LBD P767A	Dr. Jonathan Whitchurch
10099	pBTM116mod-AR LBD Y764C	Dr. Jonathan Whitchurch
22	pASV3mod AR LBD	Dr. Christopher Roberts
23	pBTM116mod AR LBD	Dr. Christopher Roberts
29	pBTM116mod TRβ LBD	Dr. Christopher Roberts
30	pBTM116mod TRβ LBD D355R* (reported to have better stability)	Dr. Christopher Roberts

After the transformation, the yeast was plated onto the TSM plates, and three colonies of each co-transformant were selected out for further analysis. Colonies were then inoculated in 20 ml YFP media and vortexed to mix. For initial interaction assays, we used a qualitative X-gal color change assay. Each 20 ml culture was separated into 2 tubes of 10 ml, with ligand added to one tube while the other received vehicle (50% ethanol or DMSO). For LexA- AR LBD and VP16- AR LBD interactions the ligand added was 100 nM mib, or both 100 nM T3 and 100nM mib for LexA- AR LBD and VP16- TRβ LBD D355R. The D335R mutation on the TRβ LBD surface has been reported to stabilize the monomeric form of the TRβ LBD and support stable dimer formation[179]. All of the cultures were incubated at 30℃ overnight. The next day, 1 absorbance unit (at OD_{600}) of each culture were transferred and treated

by X-gal reagent as described in Methods. After 1 h incubation, the development of blue color in the presence of mib confirmed that AR LBD forms dimers in the presence of ligand (Figure 6.1). Similar reporter activity was detected in all three selected colonies. However, none of the colonies expressing VP16 - TRβ D355R with LexA- AR LBD showed any change in color either with or without ligand as shown in Figure 6.1. Although this assay successfully detected AR LBD homodimers in response to ligand binding, the failure to detect AR / TRb interactions might be due to the low sensitivity of the assay, or poor expression/ instability of the VP16 - TRβ D355R construct. This construct was provided by Dr Chris Roberts but had not been tested previously.

The expressions of VP16- AR LBD and VP16 - TRβ D355R (based on yeast protein extraction method, mentioned in in section 2.4.7) in colonies transfected with LexA- AR LBD and VP16- TRβ LBD D355R. Then, we used VP16 - TRβ LBD for interaction with LexA- AR LBD.

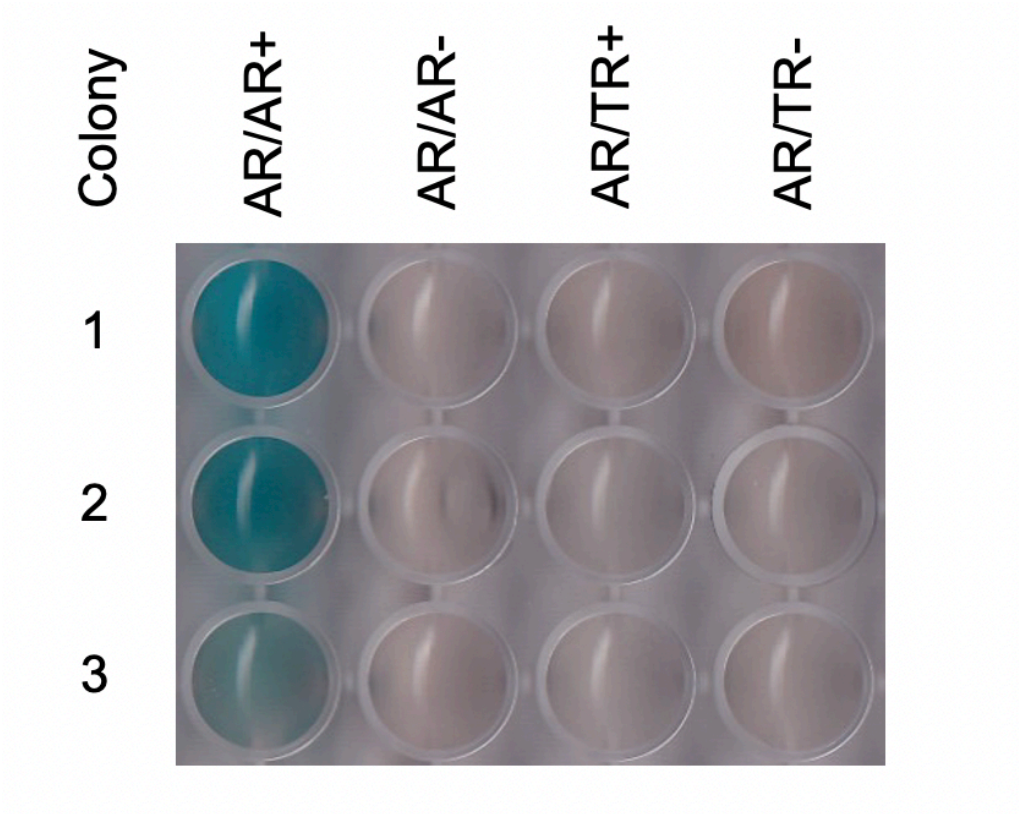


Figure 6.1 X gal assays to detect reporter activation in cell-free extracts of triplicate yeast L40 colonies, co-transformed with LexA- AR LBD and LexA- AR LBD in the presence of 100 nM mibolerone or vehicle. Also shown are triplicates for co-transformed LexA- AR LBD and VP16- TRβ LBD D355R, in the presence of 100 nM each mibolerone and T3 ligands, or vehicle.

We next repeated this experiment replacing the D335R mutant for a wild type construct i.e. using a VP16 - TRβ LBD construct that was previously used successfully in our group. As before the triplicates co-transformed with LexA- AR LBD and VP16- AR LBD displayed reporter activation in the presence of ligand, but not vehicle (Figure 6.2). Interestingly, in this experiment the LexA- AR LBD and VP16 - TRβ LBD co-transformants showed a faint color change indicating reporter

activation, but only in the presence of the two ligands mib Lerone and T3 (figure 6.2).

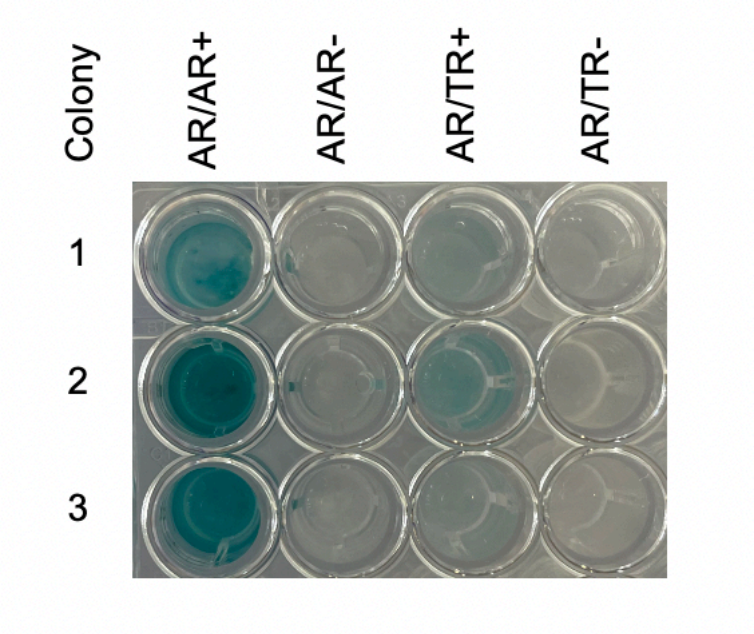


Figure 6.2 X gal assays to detect reporter activation in cell-free extracts of triplicate yeast L40 colonies, co-transformed with LexA- AR LBD and LexA-AR LBD in the presence of 100 nM mib or vehicle. Also shown are triplicates for co-transformed LexA- AR LBD and VP16- TRβ LBD WT, in the presence of 100 nM each mib and T3 ligands, or vehicle.

6.2 Quantitative β-gal assays

Having detected weak interactions of AR and TR LBDs in the qualitative X-gal assays, quantitative β-gal reporter assays was attempted. As done before, triplicate colonies of LexA- AR LBD co-transformed with VP16- AR LBD or VP16-TRβ LBD were used for the assay. Cell-free extracts were prepared from 10 ml cultures grown in the presence or absence of agonist ligands. To control for protein concentration, the

amount of protein in each extract was determined by Bradford assay using a BSA standard curve (Table 6.2 and Figure 6.3.)

Table 6.2 BSA Standard Curve of Absorbance vs BSA Protein Concentration

Concentration (ug/ml)	0	2.5	5	7.5	10	12.5	15	17.5	20	22.5
$OD_{bradford}$	0	0.095	0.182	0.324	0.411	0.485	0.625	0.853	0.986	1.112

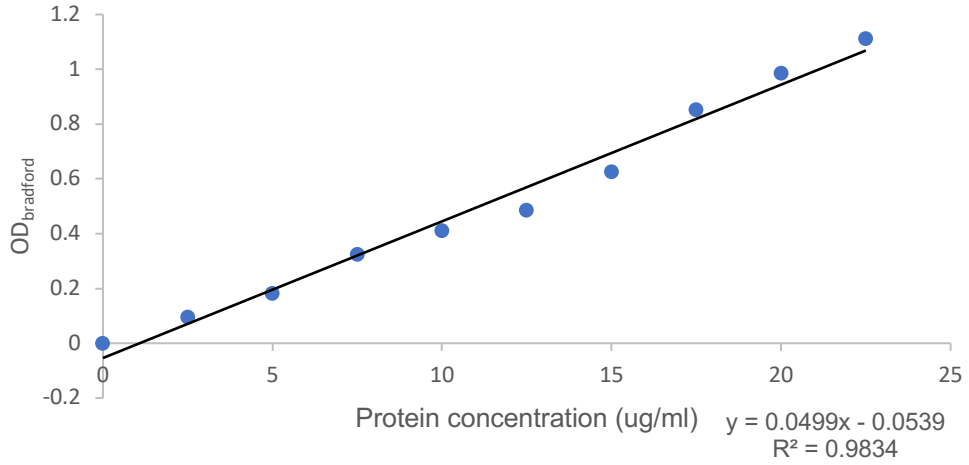


Figure 6.3 Standard curve for protein concentration. The R^2 value and linear regression equation are shown.

B-gal reporter activity was measured using the ONPG colorimetric assay. The data including assay time, protein conc of extracts, absorbance measurements are shown in Table 6.3. Activity is shown in Miller units. A graph showing the reporter activation of AR-AR and AR-TRβ LBD interactions is shown in Figure 6.4. The triplicates of LexA-AR LBD and VP16- AR LBD show a substantial (5-fold) ligand induced reporter activation in the presence of 100 nM mib. Similarly, all three clones of LexA- AR LBD and VP16- TRβ LBD, showed around 20-fold increase in reporter activation in the presence of ligands T3 and Mib.

These results are consistent with previous unpublished data from our group suggesting potential heterodimer complexes between the AR and TRβ proteins, mediated by the LBDs.

Table 6.3 β gal assays details for triplicates of LexA- AR LBD cotransormed with VP16- AR LBD or VP16- TRβ LBD, incubated in the presence or absence of cognate ligands 100 nM mibolerone and 100nM T3.

Description	Protein concentration of protein extract (µg/ml)	Time of ONPG addition (minutes)	Time of Na_2CO_3 addition (minutes)	Total time taken for reaction to occur (minutes)	OD_{420}	Volume of protein extract used (µl)	Activity
AR/AR1+	2.55	0.30	34.30	34	0.030	20	2.72
AR/AR2+	4.15	1.30	19.30	18	0.162	20	17.06
AR/AR3+	3.41	2.00	34.30	32.30	0.072	20	5.12
AR/AR1-	2.81	3.00	35.00	32	0.012	20	1.05
AR/AR2-	11.75	3.30	35.30	32	0.041	20	0.86
AR/AR3-	3.50	4.30	36.00	31.30	0.017	20	1.21
AR/TR1+	3.35	5.30	11.30	6	0.047	20	18.39
AR/TR2+	4.47	6.00	12.30	6.30	0.028	20	7.58
AR/TR3+	2.59	6.30	13.00	6.30	0.024	20	11.22
AR/TR1-	3.63	7.30	36.00	28.30	0.010	20	0.76
AR/TR2-	2.27	8.00	36.30	28.30	0.003	20	0.36
AR/TR3-	5.29	8.30	15.00	28.30	0.037	20	1.93

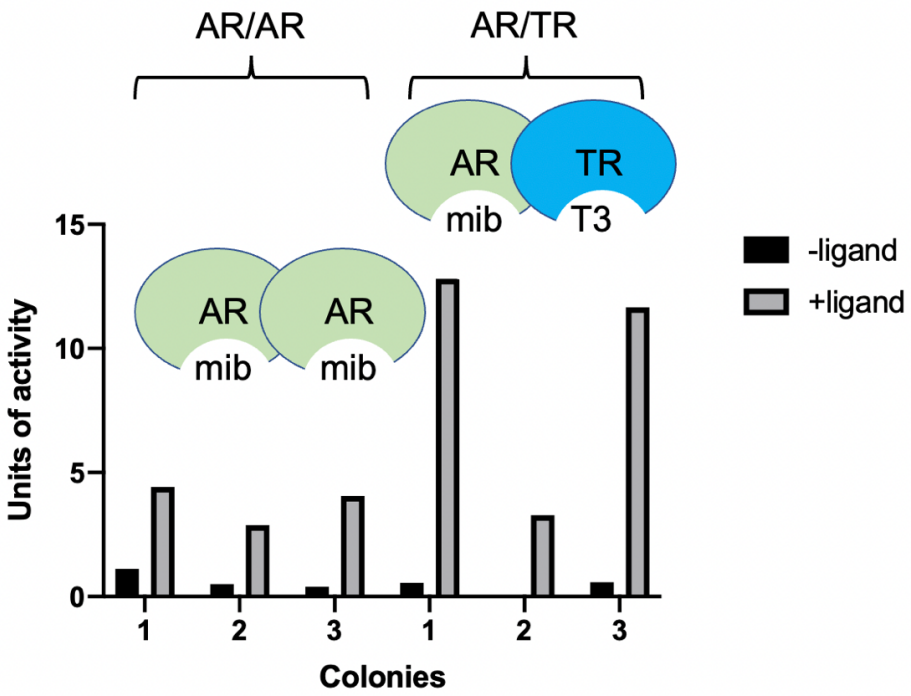


Figure 6.4 β gal reporter assays due to ligand induced interactions of LexA-AR LBD with VP16- AR LBD and VP16- TRβ LBD, incubated with or without 100 nM mib and T3, Triplicates are shown.

Having detected interactions of the AR and TRβ LBDs in yeast two hybrid assays, we next assessed the requirement of selected residues involved in interactions of the AR LBD homodimer (PDB: 5JJM). To do these we tested the interaction of a series of AR LBD substitution mutants for interaction with TRβ LBD in the yeast two hybrid assay. VP16- TRβ LBD was co-transformed with LexA- AR LBD Wt, LexA- AR LBD V770A, LexA- AR LBD W752R, LexA- AR LBD F755S, LexA- AR LBD Y764C and LexA- AR LBD P767A , all of which are implicated in stabilizing the interface of AR homodimers. As a control LexA- AR LBD T851D which is not involved in the dimer interface was included. Assay

data are shown in Table 6.4. As shown in Figure 6.5, In the absence of ligands, no interaction of WT AR LBD with TRβ. In the presence of 100 nM mib or 100 nM T3 a low level of reporter activity was observed. However, strong reporter activation was observed when ligands for both receptors were included. In contrast, mutations in the residues implicated in AR:AR LBD dimer interface showed a greatly reduced reporter activation, indicated loss of interaction with VP16- TRβ LBD. These activities were similar to the level obtained with mib only and suggests the activity observed with mib only may be due to LexA-AR LBD only, which can activate transcription in yeast[180 17, 69]. The Lexa-AR T851D mutant retained interaction with AR LBD as shown by the reporter activation, suggesting residue T851 was not important in the interaction between AR and TRβ.

Table 6.4 β gal assays record for colonies, co-transfected with LexA-AR LBD variants and VP16- TRβ LBD, incubated with or without 100 nM mib and T3.

Colonies		$OD_{Bradford}$	Protein concentration of protein extract (μg/ml)	Volume of protein extract used (μl)	Total time taken for reaction to occur (minutes)	OD_{420}	Activity
AR/TR-	1	0.957	19.310	20	30	0.002	0.027
	2	0.850	21.285	20	30	0.002	0.025
	3	1.055	21.154	20	30	0.002	0.025
AR/TR mib	1	0.908	18.417	20	30	0.084	1.196
	2	0.931	18.845	20	30	0.174	2.422
	3	0.927	18.771	20	30	0.166	2.320
AR/TR T3	1	1.008	20.279	20	30	0.002	0.026
	2	0.818	16.741	20	30	0.075	1.175
	3	0.852	17.374	20	30	0.054	0.815
AR/TR mib+T3	1	0.957	19.330	20	18.5	0.393	8.650
	2	0.850	17.337	20	24	0.453	8.568

	3	0.838	17.114	20	24	0.457	8.757
AR	1	1.019	20.484	20	30	0.001	0.013
W752R	2	0.910	18.454	20	30	0.001	0.014
/TR	3	0.907	18.398	20	30	0.003	0.043
AR	1	1.131	22.570	20	30	0.238	2.766
F755S	2	1.322	26.127	20	30	0.105	1.054
/TR	3	1.314	25.978	20	30	0.119	1.202
AR	1	1.163	23.166	20	30	0.124	1.404
Y764S	2	1.203	23.911	20	30	0.207	2.271
/TR	3	1.184	23.557	20	30	0.221	2.461
AR	1	0.961	19.404	20	30	0.025	0.338
P767A	2	0.899	18.250	20	30	0.067	0.963
/TR	3	0.895	18.175	20	30	0.052	0.750
AR	1	1.100	21.992	20	30	0.032	0.382
V770A	2	0.922	18.678	20	30	0.078	1.096
/TR	3	0.920	18.640	20	30	0.102	1.435
AR	1	1.068	21.397	20	20	0.412	7.577
T851D	2	1.032	20.726	20	20	0.409	7.765
/TR	3	1.028	20.652	20	20	0.381	7.260

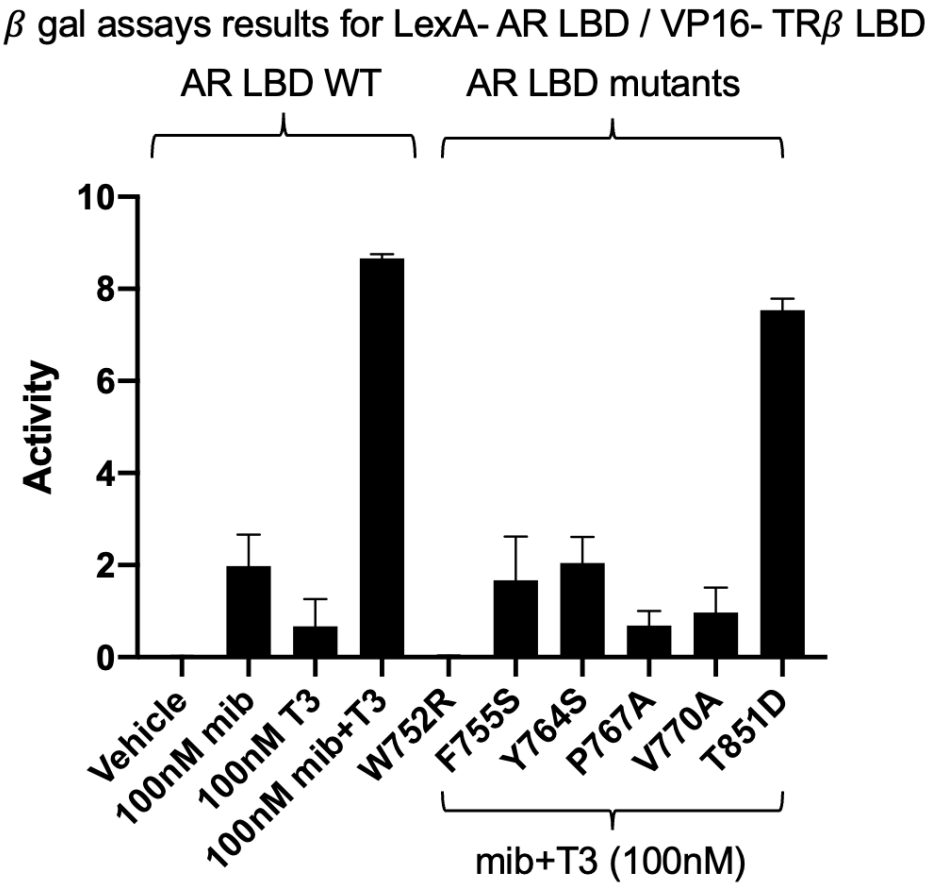


Figure 6.5 β gal assays showing reporter activation due to interactions of WT AR LBD or AR LBD mutants with TRβ incubated with or without 100 nM mib and T3. The data shown are the mean of triplicates and Error bars showed the standard error of the mean.

In this study we have shown that direct interactions of AR LBD and TRβ LBD can be induced when agonist ligands for both NRs are present. This supports the observations that AR and TR signaling pathways may cooperate in prostate tumor progression. To test whether antagonists ligands of TRβ might interfere with this interaction, we used a synthetic molecule (Compound NH3) obtained as a gift from prof Lukas Kenner (University of Vienna). NH3 was tested for the influence on the AR-TRβ interaction in quantitative β-gal assays. As shown in Table 6.5 & Figure 6.6, inclusion of 10-fold excess of NH3 reduced the reporter activity modestly. Further replicate experiments and dose curve assays will be required to confirm whether antagonists of T3 such as NH3 might be useful in reducing the AR:TR activity.

Table 6.5 β gal assays record for colonies, co-transfected with LexA- AR LBD variants and VP16- TRβ LBD, incubated with or without 100 nM mib and T3, 1 mM NH3.

Colonies	$OD_{Bradford}$	Protein concentration of protein extract (µg/ml)	Volume of protein extract used (µl)	Total time taken for reaction to occur (minutes)	OD_{420}	Activity (U/mg)
AR/TR-	0.241/0.248	6.061	20	30	0.072	3.116
AR/TR mib	0.192/0.193	5.875	20	30	0.062	2.768
AR/TR T3	0.166/0.154	4.488	20	30	0.047	2.747
AR/TR NH3	0.221/0196	5.391	20	30	0.073	3.552
AR/TR mib + T3	0.239/0.250	6.061	20	3	0.111	48.042
AR/TR mib+T3+NH3	0.192/0.193	5.093	20	5	0.105	32.451

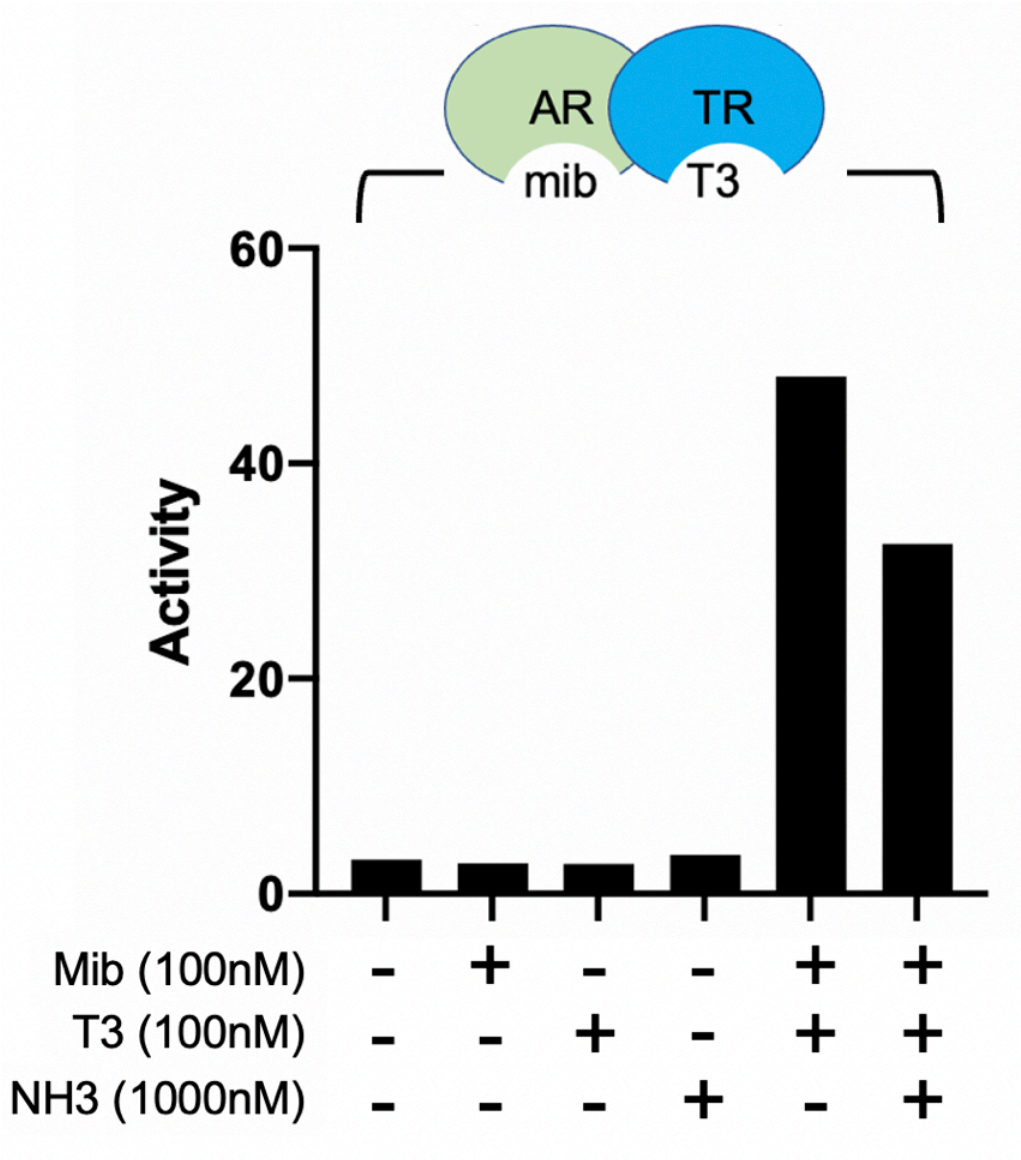


Figure 6.6 β gal assays showing reporter activation due to interaction of LexA-AR LBD and VP16- TRβ LBD, incubated with or without 100 nM mib, 100 nM T3, 1 mM NH3 as indicated. Data are the mean of triplicates and error bars showed standard error of the mean.

6.3 Summary

Previous unpublished data from our group suggested that AR LBD could interact with the LBD of TRβ, in a manner dependent on binding of agonist ligand to both LBDs. Using quantitative β gal assays it was

confirmed here that the interaction between AR LBD and TRβ LBD required both the ligand T3 and mib. Mutations of residues, W752, F755, Y764, P767, V770, which are critical for AR LBD homodimer formation, were also found to critical to the interaction between AR LBD and TRβ LBD. Moreover, preliminary data showed that an antagonist of TRβ (NH3) was able to reduce the maximal reporter activity, suggesting competition with T3 for TRb binding. Further work will be required to establish comparative affinities of T3 and NH3 for TRβ. However, our data support the hypothesis that cooperative effects of thyroid hormones and androgens in prostate tumors may involve direct interactions of AR and TRs mediated vi LBDs.

7. Concluding Comments

The study of the interaction between AR and TR is of biological and potential therapeutic importance for understanding the biology of prostate tumors. AR is a critical regulator of normal prostate function and is the major driver in proliferation of tumors of the prostate. AR mediates androgen regulation of gene expression including the KLK3 gene encoding prostate-specific antigen (PSA) a biomarker for this disease. TRs function in the control of metabolic processes in many tissues and mediate the regulation of genes by T3 and T4 hormones, either through TR/RXR heterodimers or as TR homodimers. Previous studies in the literature have revealed that TRs and thyroid hormones are active in the prostate and in prostate cancer, although the importance of this is not well understood. In a collaborative study[9] our group showed that expression of CRYM, which reduces the availability of T3 and T4 in tissues, has a negative correlation with disease progression in CRPC. This suggests that TRs may contribute to disease progression. Both TRβ and AR are expressed in prostate cancers and several studies have identified that they coregulate gene expression, including expression of the KLK3 gene which has binding sites for both AR and TRs in its promoter region. These observations provide strong evidence for crosstalk between androgen and thyroid receptor signalling, although whether this is vis direct or indirect interactions at gene promoters is unclear.

The aim of this project was to expand on unpublished data from our group that AR and TRs form novel heterodimeric complexes through their LBDs. Our lab has discovered novel heterodimer interactions in previous studies using yeast two hybrid screens e.g. The PNR/PPARγ heterodimer in the retina[170]. Therefore, a combined approach was used to study this potential heterodimer of AR and TR *in silico* and to explore these protein-protein interactions in yeast and mammalian cell models. Using available PDB structures for monomeric LBDS of AR co-crystalized with agonist, structure 2NW4 was selected for docking studies. This structure was selected because it had the lowest RMSD value when superimposed with 5JJM, the reported crystal AR homodimer[130]. A TRβ monomeric LBD structure in complex with agonist was also selected (1XZX), and hinge region sequence missing from the published structure were modelled using the Swiss-model server. These representative structures for AR and TRβ LBDs were docked by 4 common docking servers, HADDOCK、HDOCK、GRAMM-X and ZDOCK servers. The 10 top scoring complexes from each server were clustered by Conformer cluster module in Maestro Schrodinger (Schrödinger Release 2020-2: Maestro, Schrödinger, LLC). These complexes were divided into 4 clusters according to the Kelley plots. Cluster 1 was selected based on being consistent with site directed mutagenesis data from unpublished yeast two hybrid experiments from our group exploring requirements for the AR:AR dimer interface. In addition, interaction interface analysis, binding free energy calculation and virtual site directed mutagenesis analysis using PDBePISA,

HAWKDOCK, SDM and mCSM server also showed good agreement with the selected cluster. Thus, this cluster was chosen to represent the preferred model of the AR and TRβ LBD heterodimer.

To explore AR/TR interactions in cells, yeast two hybrid assays were performed which verified the interaction between AR LBD and TRβ LBD. This was only found to occur in the presence of agonists for both receptors (T3 for TRβ and mibolerone for AR). This suggests that heterodimer formation is promoted by the transcriptionally active conformations of these LBDs, at least in yeast cells. The low level of reporter activation observed in the presence of mib is likely due to LexA-AR LBD, as the Heery group and others have previously shown that NR LBDs fused to LexA DBD are transcriptionally active and can stimulate reporter gene expression in yeast cells in response to ligand [181, 182]. Interestingly, the predicted AR/TR structure exhibits a high level of conformational similarity to the AR LBD homodimer, and thus might share a similar interface. The two-hybrid data shown here confirmed that F755, W752, Y764, P767 and V770 in AR LBD (residues 647-902), which are the major binding residues shown in the reported crystal AR homodimer, are essential for interaction with TRβ (residues 211-460). The MM/GBSA calculation data through HAWKDOCK confirms that residues surrounding beta-sheet 1 in AR are major contributor for the binding with TRβ. This region contains the end of helix 4, beta sheet 1, and helix 5. The beginning of helix 7 also required for TRβ LBD. The hinge region from residues 682 to 692 seems to be anchored by the interaction with TRβ. For TRβ, the binding interface involves residues

mainly located in helix 9 and helix 10. Finally, we used a competitive inhibitor of T3 (NH3) which is an antagonist of TRβ function[183]. Preliminary data show that this reduced the ligand dependent association of AR and TRβ in yeast, providing further support for our hypothesis.

To confirm interactions in mammalian cells, we decided to use the split vYFP system which was used previously to explore AR/ cofactor interactions in our lab. This approach has advantages over using endogenous proteins and co-immunoprecipitation approaches, as it is difficult to rule out indirect interactions mediated through binding to DNA or via bound cofactors. The split YFP system required the generation of TR LBD fusion constructs by cloning into vYFPC156 vector. Progress on this part of the project was delayed as many technical issues had to be overcome to optimized PCR amplification of TR cDNAs . This report has described how each step was optimized and how the cloning strategy was modified to achieve generation of the TRβ1 LBD-vYFPC156 construct.

Although we successfully amplified and cloned human TRα1 and TRα2 LBD coding regions, sequence analysis revealed these to contain a single base pair deletion that alters the reading frame of the LBD. As this was also visible in RNA Seq data from the cell lines used, this appears to be a variant found in these cell lines. Alternatively, there may be a feature of this sequence that causes polymerase error leading to PCR artifact causing deletion of one base in the stretch of the six adenosine bases.

Having successfully cloned TRβ1 LBD- vYFPC156 vector we were able to test its interaction with AR -LBD vYFPN155 in HEK293 cells. As a control TRβ1 - vYFPC156 was also co-transfected with SRC1 NID - vYFPN155. In the presence of T3, we reconstituted fluorescent YFP in the nucleus of transfected cells, however no fluorescence was observed in the absence of T3. This result verified the functionality of the plasmid TRβ1 LBD- vYFPC156 construct as it is able to bind to a cofactor in response to ligand. However, the co-transfections using AR - vYFPN155 with TRβ1 - vYFPC156, MED1 LXM1 or SRC1 LXM3 failed to give a fluorescent signal as expected. It is possible this was due to human error in setting up the transfections. However, once this issue is resolved, follow on studies will be able to verify whether AR and TR heterodimer complexes can be formed in mammalian cells.

With regard to future work, if heterodimer complexes between AR and TRs can be established, this may have importance for considering therapeutic options for treatment of prostate cancer and possibly other tumours such as thyroid tumours. It has been shown that CRYM, which regulates bioavailability of T3, is a prognostic marker in CRPC, with loss of CRYM expression associated with advanced disease and worse prognosis[9].

In addition, thyroid cancer is more common in males and recent studies have highlighted aberrant expression of AR gene papillary thyroid carcinoma[184].

Understanding the structure of the AR/TR heterodimer, e.g. via crystallization studies would be important for understanding its function

in gene regulation. It would also raise a question as to which genes are coregulated by AR/TR dimers, which could be explored using chromatin immunoprecipitation (chip) assays to discover genomic binding sites. Finally compounds like NH3 could be evaluated in xenograft studies to determine whether these can block the growth of prostate tumours in vivo.

References

1. Robinson-Rechavi, M.; Carpentier, A. S.; Duffraisse, M.; Laudet, V., How many nuclear hormone receptors are there in the human genome? *Trends Genet* **2001,** *17* (10), 554-6.
2. Robinson-Rechavi, M.; Escriva Garcia, H.; Laudet, V., The nuclear receptor superfamily. *J Cell Sci* **2003,** *116* (Pt 4), 585-6.
3. Giguere, V.; Ong, E. S.; Segui, P.; Evans, R. M., Identification of a receptor for the morphogen retinoic acid. *Nature* **1987,** *330* (6149), 624-629.
4. Tora, L.; Gronemeyer, H.; Turcotte, B.; Gaub, M. P.; Chambon, P., The N-terminal region of the chicken progesterone receptor specifies target gene activation. *Nature* **1988,** *333* (6169), 185-8.
5. Umesono, K.; Evans, R. M., Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* **1989,** *57* (7), 1139-1146.
6. Weikum, E. R.; Liu, X.; Ortlund, E. A., The nuclear receptor superfamily: A structural perspective. *Protein Sci* **2018,** *27* (11), 1876-1892.
7. Jordan, V. C.; O'Malley, B. W., Selective estrogen-receptor modulators and antihormonal resistance in breast cancer. *J Clin Oncol* **2007,** *25* (36), 5815-24.
8. Oakes, N. D.; Kennedy, C. J.; Jenkins, A. B.; Laybutt, D. R.; Chisholm, D. J.; Kraegen, E. W., A new antidiabetic agent, BRL 49653, reduces lipid availability and improves insulin action and glucoregulation in the rat. *Diabetes* **1994,** *43* (10), 1203-10.
9. Aksoy, O.; Pencik, J.; Hartenbach, M.; Moazzami, A. A.; Schlederer, M.; Balber, T.; Varady, A.; Philippe, C.; Baltzer, P. A.; Mazumder, B.; Whitchurch, J. B.; Roberts, C. J.; Haitel, A.; Herac, M.; Susani, M.; Mitterhauser, M.; Marculescu, R.; Stangl-Kremser, J.; Hassler, M. R.; Kramer, G.; Shariat, S. F.; Turner, S. D.; Tichy, B.; Oppelt, J.; Pospisilova, S.; Hartenbach, S.; Tangermann, S.; Egger, G.; Neubauer, H. A.; Moriggl, R.; Culig, Z.; Greiner, G.; Hoermann, G.; Hacker, M.; Heery, D. M.; Merkel, O.; Kenner, L., Thyroid and androgen receptor signaling are antagonized by mu-Crystallin in prostate cancer. *Int J Cancer* **2021,** *148* (3), 731-747.
10. Horlein, A. J.; Naar, A. M.; Heinzel, T.; Torchia, J.; Gloss, B.; Kurokawa, R.; Ryan, A.; Kamei, Y.; Soderstrom, M.; Glass, C. K.; et al., Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* **1995,** *377* (6548), 397-404.
11. Norris, J. D.; Fan, D.; Kerner, S. A.; McDonnell, D. P., Identification of a third autonomous activation domain within the human estrogen receptor. *Mol Endocrinol* **1997,** *11* (6), 747-54.
12. Cheskis, B.; Freedman, L. P., Modulation of nuclear receptor interactions by ligands: kinetic analysis using surface plasmon resonance. *Biochemistry* **1996,** *35* (10), 3309-18.
13. Lonard, D. M.; Lanz, R. B.; O'Malley, B. W., Nuclear receptor coregulators and human disease. *Endocr Rev* **2007,** *28* (5), 575-87.
14. Barettino, D.; Vivanco Ruiz, M. M.; Stunnenberg, H. G., Characterization of the ligand-dependent transactivation domain of thyroid hormone receptor. *EMBO J* **1994,** *13* (13), 3039-49.
15. Danielian, P. S.; White, R.; Lees, J. A.; Parker, M. G., Identification of a conserved region required for hormone dependent transcriptional activation by steroid hormone receptors. *EMBO J* **1992,** *11* (3), 1025-33.

16. Kurokawa, R.; Soderstrom, M.; Horlein, A.; Halachmi, S.; Brown, M.; Rosenfeld, M. G.; Glass, C. K., Polarity-specific activities of retinoic acid receptors determined by a co-repressor. *Nature* **1995,** *377* (6548), 451-4.
17. Heery, D. M.; Kalkhoven, E.; Hoare, S.; Parker, M. G., A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **1997,** *387*, 733-736.
18. Germain, P.; Staels, B.; Dacquet, C.; Spedding, M.; Laudet, V., Overview of nomenclature of nuclear receptors. *Pharmacol Rev* **2006,** *58* (4), 685-704.
19. Kumar, R.; Betney, R.; Li, J.; Thompson, E. B.; McEwan, I. J., Induced alpha-helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIF. *Biochemistry* **2004,** *43* (11), 3008-13.
20. Brodie, J.; McEwan, I. J., Intra-domain communication between the N-terminal and DNA-binding domains of the androgen receptor: modulation of androgen response element DNA binding. *J Mol Endocrinol* **2005,** *34* (3), 603-15.
21. Cunha, G. R.; Donjacour, A. A.; Cooke, P. S.; Mee, S.; Bigsby, R. M.; Higgins, S. J.; Sugimura, Y., The endocrinology and developmental biology of the prostate. *Endocr Rev* **1987,** *8* (3), 338-62.
22. Cutress, M. L.; Whitaker, H. C.; Mills, I. G.; Stewart, M.; Neal, D. E., Structural basis for the nuclear import of the human androgen receptor. *J Cell Sci* **2008,** *121* (Pt 7), 957-68.
23. Clinckemalie, L.; Vanderschueren, D.; Boonen, S.; Claessens, F., The hinge region in androgen receptor control. *Mol Cell Endocrinol* **2012,** *358* (1), 1-8.
24. Eerlings, R.; Barbakadze, N.; Nguyen, T.; Nadaraia, N.; Smeets, E.; Moris, L.; Handle, F.; El Kharraz, S.; Devlies, W.; Voet, A.; Dehaen, W.; Claessens, F.; Helsen, C., Small-molecule profiling for steroid receptor activity using a universal steroid receptor reporter assay. *J Steroid Biochem Mol Biol* **2022,** *217*, 106043.
25. Matias, P. M.; Donner, P.; Coelho, R.; Thomaz, M.; Peixoto, C.; Macedo, S.; Otto, N.; Joschko, S.; Scholz, P.; Wegg, A.; Basler, S.; Schafer, M.; Egner, U.; Carrondo, M. A., Structural evidence for ligand specificity in the binding domain of the human androgen receptor. Implications for pathogenic gene mutations. *J Biol Chem* **2000,** *275* (34), 26164-71.
26. Sack, J. S.; Kish, K. F.; Wang, C.; Attar, R. M.; Kiefer, S. E.; An, Y.; Wu, G. Y.; Scheffler, J. E.; Salvati, M. E.; Krystek, S. R., Jr.; Weinmann, R.; Einspahr, H. M., Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone. *Proc Natl Acad Sci U S A* **2001,** *98* (9), 4904-9.
27. Shaffer, P. L.; Jivan, A.; Dollins, D. E.; Claessens, F.; Gewirth, D. T., Structural basis of androgen receptor binding to selective androgen response elements. *Proc Natl Acad Sci U S A* **2004,** *101* (14), 4758-63.
28. Testa, U.; Castelli, G.; Pelosi, E., Cellular and Molecular Mechanisms Underlying Prostate Cancer Development: Therapeutic Implications. *Medicines (Basel)* **2019,** *6* (3).
29. Begley, L.; Monteleon, C.; Shah, R. B.; Macdonald, J. W.; Macoska, J. A., CXCL12 overexpression and secretion by aging fibroblasts enhance human prostate epithelial proliferation in vitro. *Aging Cell* **2005,** *4* (6), 291-8.
30. Bavik, C.; Coleman, I.; Dean, J. P.; Knudsen, B.; Plymate, S.; Nelson, P. S., The gene expression program of prostate fibroblast senescence modulates neoplastic epithelial cell proliferation through paracrine mechanisms. *Cancer Res* **2006,** *66* (2), 794-802.

31. Bethel, C. R.; Chaudhary, J.; Anway, M. D.; Brown, T. R., Gene expression changes are age-dependent and lobe-specific in the brown Norway rat model of prostatic hyperplasia. *Prostate* **2009,** *69* (8), 838-50.
32. de Bono, J. S.; Guo, C.; Gurel, B.; De Marzo, A. M.; Sfanos, K. S.; Mani, R. S.; Gil, J.; Drake, C. G.; Alimonti, A., Prostate carcinogenesis: inflammatory storms. *Nat Rev Cancer* **2020,** *20* (8), 455-469.
33. Tan, B. L.; Norhaizan, M. E., Oxidative Stress, Diet and Prostate Cancer. *World J Mens Health* **2021,** *39* (2), 195-207.
34. Wang, X.; Kruithof-de Julio, M.; Economides, K. D.; Walker, D.; Yu, H.; Halili, M. V.; Hu, Y. P.; Price, S. M.; Abate-Shen, C.; Shen, M. M., A luminal epithelial stem cell that is a cell of origin for prostate cancer. *Nature* **2009,** *461* (7263), 495-500.
35. Kim, J.; Eltoum, I. E.; Roh, M.; Wang, J.; Abdulkadir, S. A., Interactions between cells with distinct mutations in c-MYC and Pten in prostate cancer. *PLoS Genet* **2009,** *5* (7), e1000542.
36. Albadine, R.; Latour, M.; Toubaji, A.; Haffner, M.; Isaacs, W. B.; E, A. P.; Meeker, A. K.; Demarzo, A. M.; Epstein, J. I.; Netto, G. J., TMPRSS2-ERG gene fusion status in minute (minimal) prostatic adenocarcinoma. *Mod Pathol* **2009,** *22* (11), 1415-22.
37. Kim, M. J.; Cardiff, R. D.; Desai, N.; Banach-Petrosky, W. A.; Parsons, R.; Shen, M. M.; Abate-Shen, C., Cooperativity of Nkx3.1 and Pten loss of function in a mouse model of prostate carcinogenesis. *Proc Natl Acad Sci U S A* **2002,** *99* (5), 2884-9.
38. Shen, M. M.; Abate-Shen, C., Molecular genetics of prostate cancer: new prospects for old challenges. *Genes Dev* **2010,** *24* (18), 1967-2000.
39. Wang, G.; Zhao, D.; Spring, D. J.; DePinho, R. A., Genetics and biology of prostate cancer. *Genes Dev* **2018,** *32* (17-18), 1105-1140.
40. Geller, J.; Albert, J.; Loza, D.; Geller, S.; Stoeltzing, W.; de la Vega, D., DHT concentrations in human prostate cancer tissue. *J Clin Endocrinol Metab* **1978,** *46* (3), 440-4.
41. Teo, M. Y.; Rathkopf, D. E.; Kantoff, P., Treatment of Advanced Prostate Cancer. *Annu Rev Med* **2019,** *70*, 479-499.
42. Denis, L. J.; Griffiths, K., Endocrine treatment in prostate cancer. *Seminars in Surgical Oncology* **2000,** *18* (1), 52-74.
43. Culig, Z.; Bartsch, G., Androgen axis in prostate cancer. *J Cell Biochem* **2006,** *99* (2), 373-81.
44. Montgomery, R. B.; Mostaghel, E. A.; Vessella, R.; Hess, D. L.; Kalhorn, T. F.; Higano, C. S.; True, L. D.; Nelson, P. S., Maintenance of intratumoral androgens in metastatic prostate cancer: a mechanism for castration-resistant tumor growth. *Cancer Res* **2008,** *68* (11), 4447-54.
45. Chang, K. H.; Li, R.; Kuri, B.; Lotan, Y.; Roehrborn, C. G.; Liu, J.; Vessella, R.; Nelson, P. S.; Kapur, P.; Guo, X.; Mirzaei, H.; Auchus, R. J.; Sharifi, N., A gain-of-function mutation in DHT synthesis in castration-resistant prostate cancer. *Cell* **2013,** *154* (5), 1074-1084.
46. Edwards, J.; Krishna, N. S.; Grigor, K. M.; Bartlett, J. M., Androgen receptor gene amplification and protein expression in hormone refractory prostate cancer. *Br J Cancer* **2003,** *89* (3), 552-6.
47. Chen, C. D.; Welsbie, D. S.; Tran, C.; Baek, S. H.; Chen, R.; Vessella, R.; Rosenfeld, M. G.; Sawyers, C. L., Molecular determinants of resistance to antiandrogen therapy. *Nat Med* **2004,** *10* (1), 33-9.
48. Scher, H. I.; Sawyers, C. L., Biology of progressive, castration-resistant prostate cancer: directed therapies targeting the androgen-receptor signaling axis. *J Clin Oncol* **2005,** *23* (32), 8253-61.
49. Guo, Z.; Yang, X.; Sun, F.; Jiang, R.; Linn, D. E.; Chen, H.; Chen, H.; Kong, X.; Melamed, J.; Tepper, C. G.; Kung, H. J.; Brodie, A. M.;

Edwards, J.; Qiu, Y., A novel androgen receptor splice variant is up-regulated during prostate cancer progression and promotes androgen depletion-resistant growth. *Cancer Res* **2009,** *69* (6), 2305-13.
50. Dehm, S. M.; Schmidt, L. J.; Heemers, H. V.; Vessella, R. L.; Tindall, D. J., Splicing of a novel androgen receptor exon generates a constitutively active androgen receptor that mediates prostate cancer therapy resistance. *Cancer Res* **2008,** *68* (13), 5469-77.
51. Hu, R.; Dunn, T. A.; Wei, S.; Isharwal, S.; Veltri, R. W.; Humphreys, E.; Han, M.; Partin, A. W.; Vessella, R. L.; Isaacs, W. B.; Bova, G. S.; Luo, J., Ligand-independent androgen receptor variants derived from splicing of cryptic exons signify hormone-refractory prostate cancer. *Cancer Res* **2009,** *69* (1), 16-22.
52. Sun, S.; Sprenger, C. C.; Vessella, R. L.; Haugk, K.; Soriano, K.; Mostaghel, E. A.; Page, S. T.; Coleman, I. M.; Nguyen, H. M.; Sun, H.; Nelson, P. S.; Plymate, S. R., Castration resistance in human prostate cancer is conferred by a frequently occurring androgen receptor splice variant. *J Clin Invest* **2010,** *120* (8), 2715-30.
53. Watson, P. A.; Chen, Y. F.; Balbas, M. D.; Wongvipat, J.; Socci, N. D.; Viale, A.; Kim, K.; Sawyers, C. L., Constitutively active androgen receptor splice variants expressed in castration-resistant prostate cancer require full-length androgen receptor. *Proc Natl Acad Sci U S A* **2010,** *107* (39), 16759-65.
54. Robinson, D.; Van Allen, E. M.; Wu, Y. M.; Schultz, N.; Lonigro, R. J.; Mosquera, J. M.; Montgomery, B.; Taplin, M. E.; Pritchard, C. C.; Attard, G.; Beltran, H.; Abida, W.; Bradley, R. K.; Vinson, J.; Cao, X.; Vats, P.; Kunju, L. P.; Hussain, M.; Feng, F. Y.; Tomlins, S. A.; Cooney, K. A.; Smith, D. C.; Brennan, C.; Siddiqui, J.; Mehra, R.; Chen, Y.; Rathkopf, D. E.; Morris, M. J.; Solomon, S. B.; Durack, J. C.; Reuter, V. E.; Gopalan, A.; Gao, J.; Loda, M.; Lis, R. T.; Bowden, M.; Balk, S. P.; Gaviola, G.; Sougnez, C.; Gupta, M.; Yu, E. Y.; Mostaghel, E. A.; Cheng, H. H.; Mulcahy, H.; True, L. D.; Plymate, S. R.; Dvinge, H.; Ferraldeschi, R.; Flohr, P.; Miranda, S.; Zafeiriou, Z.; Tunariu, N.; Mateo, J.; Perez-Lopez, R.; Demichelis, F.; Robinson, B. D.; Schiffman, M.; Nanus, D. M.; Tagawa, S. T.; Sigaras, A.; Eng, K. W.; Elemento, O.; Sboner, A.; Heath, E. I.; Scher, H. I.; Pienta, K. J.; Kantoff, P.; de Bono, J. S.; Rubin, M. A.; Nelson, P. S.; Garraway, L. A.; Sawyers, C. L.; Chinnaiyan, A. M., Integrative clinical genomics of advanced prostate cancer. *Cell* **2015,** *161* (5), 1215-1228.
55. Waltering, K. K.; Urbanucci, A.; Visakorpi, T., Androgen receptor (AR) aberrations in castration-resistant prostate cancer. *Mol Cell Endocrinol* **2012,** *360* (1-2), 38-43.
56. Grasso, C. S.; Wu, Y. M.; Robinson, D. R.; Cao, X.; Dhanasekaran, S. M.; Khan, A. P.; Quist, M. J.; Jing, X.; Lonigro, R. J.; Brenner, J. C.; Asangani, I. A.; Ateeq, B.; Chun, S. Y.; Siddiqui, J.; Sam, L.; Anstett, M.; Mehra, R.; Prensner, J. R.; Palanisamy, N.; Ryslik, G. A.; Vandin, F.; Raphael, B. J.; Kunju, L. P.; Rhodes, D. R.; Pienta, K. J.; Chinnaiyan, A. M.; Tomlins, S. A., The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **2012,** *487* (7406), 239-43.
57. Gottlieb, B.; Beitel, L. K.; Nadarajah, A.; Paliouras, M.; Trifiro, M., The androgen receptor gene mutations database: 2012 update. *Hum Mutat* **2012,** *33* (5), 887-94.
58. Culig, Z.; Klocker, H.; Bartsch, G.; Hobisch, A., Androgen receptor mutations in carcinoma of the prostate: significance for endocrine therapy. *Am J Pharmacogenomics* **2001,** *1* (4), 241-9.
59. Visakorpi, T.; Hyytinen, E.; Koivisto, P.; Tanner, M.; Keinanen, R.; Palmberg, C.; Palotie, A.; Tammela, T.; Isola, J.; Kallioniemi, O. P., In vivo

amplification of the androgen receptor gene and progression of human prostate cancer. *Nat Genet* **1995,** *9* (4), 401-6.
60. Brooke, G. N.; Bevan, C. L., The role of androgen receptor mutations in prostate cancer progression. *Curr Genomics* **2009,** *10* (1), 18-25.
61. Jin, L.; Garcia, J.; Chan, E.; de la Cruz, C.; Segal, E.; Merchant, M.; Kharbanda, S.; Raisner, R.; Haverty, P. M.; Modrusan, Z.; Ly, J.; Choo, E.; Kaufman, S.; Beresini, M. H.; Romero, F. A.; Magnuson, S.; Gascoigne, K. E., Therapeutic Targeting of the CBP/p300 Bromodomain Blocks the Growth of Castration-Resistant Prostate Cancer. *Cancer Res* **2017,** *77* (20), 5564-5575.
62. Bernal, J., Thyroid hormone receptors in brain development and function. *Nat Clin Pract Endocrinol Metab* **2007,** *3* (3), 249-59.
63. Berbel, P.; Navarro, D.; Roman, G. C., An evo-devo approach to thyroid hormones in cerebral and cerebellar cortical development: etiological implications for autism. *Front Endocrinol (Lausanne)* **2014,** *5*, 146.
64. Guigon, C. J.; Kim, D. W.; Zhu, X.; Zhao, L.; Cheng, S. Y., Tumor suppressor action of liganded thyroid hormone receptor beta by direct repression of beta-catenin gene expression. *Endocrinology* **2010,** *151* (11), 5528-36.
65. Guigon, C. J.; Fozzatti, L.; Lu, C.; Willingham, M. C.; Cheng, S. Y., Inhibition of mTORC1 signaling reduces tumor growth but does not prevent cancer progression in a mouse model of thyroid cancer. *Carcinogenesis* **2010,** *31* (7), 1284-91.
66. Faustino, L. C.; Ortiga-Carvalho, T. M., Thyroid hormone role on cerebellar development and maintenance: a perspective based on transgenic mouse models. *Front Endocrinol (Lausanne)* **2014,** *5*, 75.
67. Cheng, B.; Al-Essa, M.; Sequeira, F.; Thakkar, J., Detection of thiopental in the steroid fraction of serum from neonates following maternal exposure. *J Steroid Biochem Mol Biol* **2004,** *92* (1-2), 73-7.
68. Yen, P. M.; Feng, X.; Flamant, F.; Chen, Y.; Walker, R. L.; Weiss, R. E.; Chassande, O.; Samarut, J.; Refetoff, S.; Meltzer, P. S., Effects of ligand and thyroid hormone receptor isoforms on hepatic gene expression profiles of thyroid hormone receptor knockout mice. *EMBO Rep* **2003,** *4* (6), 581-7.
69. Heery, D. M.; Hoare, S.; Hussain, S.; Parker, M. G.; Sheppard, H., Core LXXLL motif sequences in CREB-binding protein, SRC1, and RIP140 define affinity and selectivity for steroid and retinoid receptors. *J Biol Chem* **2001,** *276* (9), 6695-702.
70. Tata, J. R., Signalling through nuclear receptors. *Nat Rev Mol Cell Biol* **2002,** *3* (9), 702-10.
71. Oetting, A.; Yen, P. M., New insights into thyroid hormone action. *Best Pract Res Clin Endocrinol Metab* **2007,** *21* (2), 193-208.
72. Kim, J. H.; Yang, C. K.; Heo, K.; Roeder, R. G.; An, W.; Stallcup, M. R., CCAR1, a key regulator of mediator complex recruitment to nuclear receptor transcription complexes. *Mol Cell* **2008,** *31* (4), 510-519.
73. Song, I. H.; Buttgereit, F., Non-genomic glucocorticoid effects to provide the basis for new drug developments. *Mol Cell Endocrinol* **2006,** *246* (1-2), 142-6.
74. Shi, Y. B.; Matsuura, K.; Fujimoto, K.; Wen, L.; Fu, L., Thyroid hormone receptor actions on transcription in amphibia: The roles of histone modification and chromatin disruption. *Cell Biosci* **2012,** *2* (1), 42.
75. Mengeling, B. J.; Lee, S.; Privalsky, M. L., Coactivator recruitment is enhanced by thyroid hormone receptor trimers. *Mol Cell Endocrinol* **2008,** *280* (1-2), 47-62.

76. Kimura, S., Thyroid regeneration: how stem cells play a role? *Front Endocrinol (Lausanne)* **2014,** *5*, 55.
77. Hage, M. P.; Azar, S. T., The Link between Thyroid Function and Depression. *J Thyroid Res* **2012,** *2012*, 590648.
78. Dong, H.; Paquette, M.; Williams, A.; Zoeller, R. T.; Wade, M.; Yauk, C., Thyroid hormone may regulate mRNA abundance in liver by acting on microRNAs. *PLoS One* **2010,** *5* (8), e12136.
79. Velasco, L. F.; Togashi, M.; Walfish, P. G.; Pessanha, R. P.; Moura, F. N.; Barra, G. B.; Nguyen, P.; Rebong, R.; Yuan, C.; Simeoni, L. A.; Ribeiro, R. C.; Baxter, J. D.; Webb, P.; Neves, F. A., Thyroid hormone response element organization dictates the composition of active receptor. *J Biol Chem* **2007,** *282* (17), 12458-66.
80. Printz, T.; Rosenberg, T.; Godballe, C.; Dyrvig, A. K.; Grontved, A. M., Reproducibility of Automated Voice Range Profiles, a Systematic Literature Review. *J Voice* **2018,** *32* (3), 273-280.
81. Li, D.; Yamada, T.; Wang, F.; Vulin, A. I.; Samuels, H. H., Novel roles of retinoid X receptor (RXR) and RXR ligand in dynamically modulating the activity of the thyroid hormone receptor/RXR heterodimer. *J Biol Chem* **2004,** *279* (9), 7427-37.
82. Diallo, E. M.; Wilhelm, K. G., Jr.; Thompson, D. L.; Koenig, R. J., Variable RXR requirements for thyroid hormone responsiveness of endogenous genes. *Mol Cell Endocrinol* **2007,** *264* (1-2), 149-56.
83. Zimmermann, M. B., Iodine deficiency. *Endocr Rev* **2009,** *30* (4), 376-408.
84. Rubio, I. G.; Medeiros-Neto, G., Mutations of the thyroglobulin gene and its relevance to thyroid disorders. *Curr Opin Endocrinol Diabetes Obes* **2009,** *16* (5), 373-8.
85. Cheng, S. Y.; Leonard, J. L.; Davis, P. J., Molecular aspects of thyroid hormone actions. *Endocr Rev* **2010,** *31* (2), 139-70.
86. Williams, G. R., Neurodevelopmental and neurophysiological actions of thyroid hormone. *J Neuroendocrinol* **2008,** *20* (6), 784-94.
87. Tata, J. R., The road to nuclear receptors of thyroid hormone. *Biochim Biophys Acta* **2013,** *1830* (7), 3860-6.
88. Bianco, A. C.; Dumitrescu, A.; Gereben, B.; Ribeiro, M. O.; Fonseca, T. L.; Fernandes, G. W.; Bocco, B., Paradigms of Dynamic Control of Thyroid Hormone Signaling. *Endocrine Reviews* **2019**.
89. Sanchez-Tusie, A.; Montes de Oca, C.; Rodriguez-Castelan, J.; Delgado-Gonzalez, E.; Ortiz, Z.; Alvarez, L.; Zarco, C.; Aceves, C.; Anguiano, B., A rise in T3/T4 ratio reduces the growth of prostate tumors in a murine model. *J Endocrinol* **2020,** *247* (3), 225-238.
90. Hsieh; M.-L., Cell Growth Effects of Triiodothyronine and Expression of Thyroid Hormone Receptor in Prostate Carcinoma Cells. *Journal of Andrology* **2005,** *26* (3), 422-428.
91. Esquenet, M.; Swinnen, J. V.; Heyns, W.; Verhoeven, G., Triiodothyronine modulates growth, secretory function and androgen receptor concentration in the prostatic carcinoma cell line LNCaP. *Molecular & Cellular Endocrinology* **1995,** *109* (1), 105-11.
92. Zhang, S.; Hsieh, M. L.; Zhu, W.; Klee, G. G.; Tindall, D. J.; Young, C. Y., Interactive effects of triiodothyronine and androgens on prostate cell growth and gene expression. *Endocrinology* **1999,** *140* (4), 1665-71.
93. Anguiano, B.; Montes de Oca, C.; Delgado-Gonzalez, E.; Aceves, C., Prostate gland as a target organ of thyroid hormones: advances and controversies. *Endocr Connect* **2022,** *11* (2).

94. Malinowska, K.; Cavarretta, I. T.; Susani, M.; Wrulich, O. A.; Uberall, F.; Kenner, L.; Culig, Z., Identification of mu-crystallin as an androgen-regulated gene in human prostate cancer. *Prostate* **2009,** *69* (10), 1109-18.
95. Hanahan, D., Studies on transformation of Escherichia coli with plasmids. *Journal of Molecular Biology* **1983,** *166* (4), 557-580.
96. Sambrook, J.; Fritsch, F. E.; Maniatis, T., Molecular Cloning: a Laboratory Manual. *CSH* **1982**.
97. Le Douarin, B.; Heery, D. M.; Gaudon, C.; vom Baur, E.; Losson, R., Yeast two-hybrid screening for proteins that interact with nuclear hormone receptors. *Methods Mol Biol* **2001,** *176*, 227-48.
98. Vojtek, A. B.; Hollenberg, S. M.; Cooper, J. A., Mammalian Ras interacts directly with the serine/threonine kinase raf. *Cell* **1993,** *74* (1), 205-214.
99. Manivasakam, P.; Schiestl, R. H., High efficiency transformation of Saccharomyces cerevisiae by electroporation. *Nucleic Acids Research* **1993,** *21* (20), 4856-4856.
100. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T. L., BLAST+: architecture and applications. *BMC Bioinformatics* **2009,** *10*, 421.
101. The UniProt, C., UniProt: the universal protein knowledgebase. *Nucleic Acids Res* **2017,** *45* (D1), D158-D169.
102. Ostrowski, J.; Kuhns, J. E.; Lupisella, J. A.; Manfredi, M. C.; Beehler, B. C.; Krystek, S. R., Jr.; Bi, Y.; Sun, C.; Seethala, R.; Golla, R.; Sleph, P. G.; Fura, A.; An, Y.; Kish, K. F.; Sack, J. S.; Mookhtiar, K. A.; Grover, G. J.; Hamann, L. G., Pharmacological and x-ray structural characterization of a novel selective androgen receptor modulator: potent hyperanabolic stimulation of skeletal muscle with hypostimulation of prostate in rats. *Endocrinology* **2007,** *148* (1), 4-12.
103. Sandler, B.; Webb, P.; Apriletti, J. W.; Huber, B. R.; Togashi, M.; Cunha Lima, S. T.; Juric, S.; Nilsson, S.; Wagner, R.; Fletterick, R. J.; Baxter, J. D., Thyroxine-thyroid hormone receptor interactions. *J Biol Chem* **2004,** *279* (53), 55801-8.
104. Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Feng, Z.; Gilliland, G. L.; Iype, L.; Jain, S.; Fagan, P.; Marvin, J.; Padilla, D.; Ravichandran, V.; Schneider, B.; Thanki, N.; Weissig, H.; Westbrook, J. D.; Zardecki, C., The Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* **2002,** *58* (Pt 6 No 1), 899-907.
105. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **2018,** *46* (W1), W296-W303.
106. Biasini, M.; Schmidt, T.; Bienert, S.; Mariani, V.; Studer, G.; Haas, J.; Johner, N.; Schenk, A. D.; Philippsen, A.; Schwede, T., OpenStructure: an integrated software framework for computational structural biology. *Acta Crystallogr D Biol Crystallogr* **2013,** *69* (Pt 5), 701-9.
107. Xu, J.; Jiao, F.; Berger, B., A tree-decomposition approach to protein structure prediction. *Proc IEEE Comput Syst Bioinform Conf* **2005**, 247-56.
108. Bienert, S.; Waterhouse, A.; de Beer, T. A.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T., The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res* **2017,** *45* (D1), D313-D319.
109. Li, W.; Jaroszewski, L.; Godzik, A., Sequence clustering strategies improve remote homology recognitions while reducing search times. *Protein Eng* **2002,** *15* (8), 643-9.

110. Steinegger, M.; Meier, M.; Mirdita, M.; Vohringer, H.; Haunsberger, S. J.; Soding, J., HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* **2019,** *20* (1), 473.
111. Jones, P.; Binns, D.; Chang, H. Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; Pesseat, S.; Quinn, A. F.; Sangrador-Vegas, A.; Scheremetjew, M.; Yong, S. Y.; Lopez, R.; Hunter, S., InterProScan 5: genome-scale protein function classification. *Bioinformatics* **2014,** *30* (9), 1236-40.
112. Wang, G.; Dunbrack, R. L., Jr., PISCES: a protein sequence culling server. *Bioinformatics* **2003,** *19* (12), 1589-91.
113. Finn, R. D.; Coggill, P.; Eberhardt, R. Y.; Eddy, S. R.; Mistry, J.; Mitchell, A. L.; Potter, S. C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; Salazar, G. A.; Tate, J.; Bateman, A., The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* **2016,** *44* (D1), D279-85.
114. Mirdita, M.; von den Driesch, L.; Galiez, C.; Martin, M. J.; Soding, J.; Steinegger, M., Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Res* **2017,** *45* (D1), D170-D176.
115. Guex, N.; Peitsch, M. C.; Schwede, T., Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* **2009,** *30 Suppl 1*, S162-73.
116. Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd, Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* **2004,** *25* (11), 1400-15.
117. Kabsch, W.; Sander, C., Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983,** *22* (12), 2577-637.
118. Jones, D. T., Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **1999,** *292* (2), 195-202.
119. Henikoff, S.; Henikoff, J. G., Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A* **1992,** *89* (22), 10915-9.
120. Cheng, J.; Randall, A. Z.; Sweredoski, M. J.; Baldi, P., SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res* **2005,** *33* (Web Server issue), W72-6.
121. Studer, G.; Tauriello, G.; Bienert, S.; Biasini, M.; Johner, N.; Schwede, T., ProMod3-A versatile homology modelling toolbox. *PLoS Comput Biol* **2021,** *17* (1), e1008667.
122. Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L., Jr., Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* **2009,** *77* (4), 778-95.
123. Studer, G.; Rempfer, C.; Waterhouse, A. M.; Gumienny, R.; Haas, J.; Schwede, T., QMEANDisCo-distance constraints applied on model quality estimation. *Bioinformatics* **2020,** *36* (6), 1765-1771.
124. Studer, G.; Biasini, M.; Schwede, T., Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane). *Bioinformatics* **2014,** *30* (17), i505-11.
125. Benkert, P.; Biasini, M.; Schwede, T., Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **2011,** *27* (3), 343-50.
126. Bertoni, M.; Kiefer, F.; Biasini, M.; Bordoli, L.; Schwede, T., Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep* **2017,** *7* (1), 10480.
127. Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T., lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **2013,** *29* (21), 2722-8.

128. Haas, J.; Barbato, A.; Behringer, D.; Studer, G.; Roth, S.; Bertoni, M.; Mostaguir, K.; Gumienny, R.; Schwede, T., Continuous Automated Model EvaluatiOn (CAMEO) complementing the critical assessment of structure prediction in CASP12. *Proteins* **2018,** *86 Suppl 1*, 387-398.
129. Velankar, S.; Dana, J. M.; Jacobsen, J.; van Ginkel, G.; Gane, P. J.; Luo, J.; Oldfield, T. J.; O'Donovan, C.; Martin, M. J.; Kleywegt, G. J., SIFTS: Structure Integration with Function, Taxonomy and Sequences resource. *Nucleic Acids Res* **2013,** *41* (Database issue), D483-9.
130. Nadal, M.; Prekovic, S.; Gallastegui, N.; Helsen, C.; Abella, M.; Zielinska, K.; Gay, M.; Vilaseca, M.; Taules, M.; Houtsmuller, A. B.; van Royen, M. E.; Claessens, F.; Fuentes-Prior, P.; Estebanez-Perpina, E., Structure of the homodimeric androgen receptor ligand-binding domain. *Nat Commun* **2017,** *8*, 14388.
131. Dominguez, C.; Boelens, R.; Bonvin, A. M., HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **2003,** *125* (7), 1731-7.
132. van Zundert, G. C. P.; Rodrigues, J.; Trellet, M.; Schmitz, C.; Kastritis, P. L.; Karaca, E.; Melquiond, A. S. J.; van Dijk, M.; de Vries, S. J.; Bonvin, A., The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol* **2016,** *428* (4), 720-725.
133. Yan, Y.; Tao, H.; He, J.; Huang, S. Y., The HDOCK server for integrated protein-protein docking. *Nat Protoc* **2020,** *15* (5), 1829-1852.
134. Yan, Y.; Zhang, D.; Zhou, P.; Li, B.; Huang, S. Y., HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res* **2017,** *45* (W1), W365-W373.
135. Tovchigrechko, A.; Vakser, I. A., GRAMM-X public web server for protein-protein docking. *Nucleic Acids Res* **2006,** *34* (Web Server issue), W310-4.
136. Chen, R.; Li, L.; Weng, Z., ZDOCK: an initial-stage protein-docking algorithm. *Proteins* **2003,** *52* (1), 80-7.
137. Pierce, B. G.; Hourai, Y.; Weng, Z., Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PLoS One* **2011,** *6* (9), e24657.
138. Kelley, L. A.; Gardner, S. P.; Sutcliffe, M. J., An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. *Protein Eng* **1996,** *9* (11), 1063-5.
139. Topham, C. M.; Srinivasan, N.; Blundell, T. L., Prediction of the stability of protein mutants based on structural environment-dependent amino acid substitution and propensity tables. *Protein Eng* **1997,** *10* (1), 7-21.
140. Worth, C. L.; Preissner, R.; Blundell, T. L., SDM--a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res* **2011,** *39* (Web Server issue), W215-22.
141. Rodrigues, C. H. M.; Myung, Y.; Pires, D. E. V.; Ascher, D. B., mCSM-PPI2: predicting the effects of mutations on protein-protein interactions. *Nucleic Acids Res* **2019,** *47* (W1), W338-W344.
142. Overington, J.; Johnson, M. S.; Sali, A.; Blundell, T. L., Tertiary structural constraints on protein evolutionary diversity: templates, key residues and structure prediction. *Proc Biol Sci* **1990,** *241* (1301), 132-45.
143. Pires, D. E.; de Melo-Minardi, R. C.; dos Santos, M. A.; da Silveira, C. H.; Santoro, M. M.; Meira, W., Jr., Cutoff Scanning Matrix (CSM): structural classification and function prediction by protein inter-residue distance patterns. *BMC Genomics* **2011,** *12 Suppl 4*, S12.
144. Pires, D. E.; Ascher, D. B.; Blundell, T. L., mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* **2014,** *30* (3), 335-42.

145. Schneider, G.; Lee, M. L.; Stahl, M.; Schneider, P., De novo design of molecular architectures by evolutionary assembly of drug-derived building blocks. *J Comput Aided Mol Des* **2000,** *14* (5), 487-94.
146. Kumar, M. D.; Bava, K. A.; Gromiha, M. M.; Prabakaran, P.; Kitajima, K.; Uedaira, H.; Sarai, A., ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. *Nucleic Acids Res* **2006,** *34* (Database issue), D204-6.
147. Jankauskaite, J.; Jimenez-Garcia, B.; Dapkunas, J.; Fernandez-Recio, J.; Moal, I. H., SKEMPI 2.0: an updated benchmark of changes in protein-protein binding energy, kinetics and thermodynamics upon mutation. *Bioinformatics* **2019,** *35* (3), 462-469.
148. Pires, D. E.; Blundell, T. L.; Ascher, D. B., Platinum: a database of experimentally measured effects of mutations on structurally defined protein-ligand complexes. *Nucleic Acids Res* **2015,** *43* (Database issue), D387-91.
149. Weng, G.; Wang, E.; Wang, Z.; Liu, H.; Zhu, F.; Li, D.; Hou, T., HawkDock: a web server to predict and analyze the protein-protein complex based on computational docking and MM/GBSA. *Nucleic Acids Res* **2019,** *47* (W1), W322-W330.
150. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model* **2011,** *51* (1), 69-82.
151. Sun, H.; Li, Y.; Tian, S.; Xu, L.; Hou, T., Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. *Phys Chem Chem Phys* **2014,** *16* (31), 16719-29.
152. Chen, F.; Liu, H.; Sun, H.; Pan, P.; Li, Y.; Li, D.; Hou, T., Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein-protein binding free energies and re-rank binding poses generated by protein-protein docking. *Phys Chem Chem Phys* **2016,** *18* (32), 22129-39.
153. Cieplak, P.; Caldwell, J.; Kollman, P., Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases. *Journal of Computational Chemistry* **2001,** *22* (10), 1048-1057.
154. Onufriev, A.; Bashford, D.; Case, D. A., Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **2004,** *55* (2), 383-94.
155. Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *J Comput Chem* **2005,** *26* (16), 1668-88.
156. Pereira de Jesus-Tran, K.; Cote, P. L.; Cantin, L.; Blanchet, J.; Labrie, F.; Breton, R., Comparison of crystal structures of human androgen receptor ligand-binding domain complexed with various agonists reveals molecular determinants responsible for binding affinity. *Protein Sci* **2006,** *15* (5), 987-99.
157. Bohl, C. E.; Miller, D. D.; Chen, J.; Bell, C. E.; Dalton, J. T., Structural basis for accommodation of nonsteroidal ligands in the androgen receptor. *J Biol Chem* **2005,** *280* (45), 37747-54.
158. Sun, C.; Robl, J. A.; Wang, T. C.; Huang, Y.; Kuhns, J. E.; Lupisella, J. A.; Beehler, B. C.; Golla, R.; Sleph, P. G.; Seethala, R.; Fura, A.; Krystek, S. R., Jr.; An, Y.; Malley, M. F.; Sack, J. S.; Salvati, M. E.;

Grover, G. J.; Ostrowski, J.; Hamann, L. G., Discovery of potent, orally-active, and muscle-selective androgen receptor modulators based on an N-aryl-hydroxybicyclohydantoin scaffold. *J Med Chem* **2006,** *49* (26), 7596-9.
159. Nascimento, A. S.; Dias, S. M.; Nunes, F. M.; Aparicio, R.; Ambrosio, A. L.; Bleicher, L.; Figueira, A. C.; Santos, M. A.; de Oliveira Neto, M.; Fischer, H.; Togashi, M.; Craievich, A. F.; Garratt, R. C.; Baxter, J. D.; Webb, P.; Polikarpov, I., Structural rearrangements in the thyroid hormone receptor hinge domain and their putative role in the receptor function. *J Mol Biol* **2006,** *360* (3), 586-98.
160. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C., SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* **2003,** *31* (13), 3381-5.
161. Helsen, C.; Nguyen, T.; Vercruysse, T.; Wouters, S.; Daelemans, D.; Voet, A.; Claessens, F., The T850D Phosphomimetic Mutation in the Androgen Receptor Ligand Binding Domain Enhances Recruitment at Activation Function 2. *Int J Mol Sci* **2022,** *23* (3).
162. Kerppola, T. K., Visualization of molecular interactions using bimolecular fluorescence complementation analysis: characteristics of protein fragment complementation. *Chem Soc Rev* **2009,** *38* (10), 2876-86.
163. Heery, D. M.; Kalkhoven, E.; Hoare, S.; Parker, M. G., A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **1997,** *387* (6634), 733-6.
164. Nagai, T.; Ibata, K.; Park, E. S.; Kubota, M.; Mikoshiba, K.; Miyawaki, A., A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol* **2002,** *20* (1), 87-90.
165. Cowan, J. A., *Inorganic biochemistry: an introduction*. John Wiley & Sons: 1997.
166. Cowan, J. A., *Biological chemistry of magnesium*. VCH Publishers: 1995.
167. Burgess, J., *Ions in solution: basic principles of chemical interactions*. Elsevier: 1999.
168. Fields, S.; Song, O., A novel genetic system to detect protein-protein interactions. *Nature* **1989,** *340* (6230), 245-6.
169. Ptashne, M., How eukaryotic transcriptional activators work. *Nature* **1988,** *335* (6192), 683-9.
170. Fulton, J.; Mazumder, B.; Whitchurch, J. B.; Monteiro, C. J.; Collins, H. M.; Chan, C. M.; Clemente, M. P.; Hernandez-Quiles, M.; Stewart, E. A.; Amoaku, W. M.; Moran, P. M.; Mongan, N. P.; Persson, J. L.; Ali, S.; Heery, D. M., Heterodimers of photoreceptor-specific nuclear receptor (PNR/NR2E3) and peroxisome proliferator-activated receptor-gamma (PPARgamma) are disrupted by retinal disease-associated mutations. *Cell Death Dis* **2017,** *8* (3), e2677.
171. Chien, C. T.; Bartel, P. L.; Sternglanz, R.; Fields, S., The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc Natl Acad Sci U S A* **1991,** *88* (21), 9578-82.
172. Wade Harper, J., The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* **1993,** *75* (4), 805-816.
173. Gyuris, J., Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* **1993,** *75* (4), 791-803.
174. Freedman, L. P., Increasing the Complexity of Coactivation in Nuclear Receptor Signaling. *Cell* **1999,** *97* (1), 5-8.
175. Le Douarin, B.; Heery, D. M.; Gaudon, C.; vom Baur, E.; Losson, R., Yeast Two-Hybrid Screening for Proteins that Interact with Nuclear Hormone Receptors. In *Steroid Receptor Methods: Protocols and Assays*, Lieberman, B. A., Ed. Humana Press: Totowa, NJ, 2001; pp 227-248.

176. Le Douarin, B.; Pierrat, B.; vom Baur, E.; Chambon, P.; Losson, R., A new version of the two-hybrid assay for detection of protein-protein interactions. *Nucleic Acids Res* **1995,** *23* (5), 876-8.
177. Pierrat, B.; Heery, D. M.; Lemoine, Y.; Losson, R., Functional analysis of the human estrogen receptor using a phenotypic transactivation assay in yeast. *Gene* **1992,** *119* (2), 237-245.
178. Lek, N.; Tadokoro-Cuccaro, R.; Whitchurch, J. B.; Mazumder, B.; Miles, H.; Prentice, P.; Bunch, T.; Zielinska, K.; Metzler, V.; Mongan, N. P.; Heery, D. M.; Hughes, I. A., Predicting puberty in partial androgen insensitivity syndrome: Use of clinical and functional androgen receptor indices. *EBioMedicine* **2018,** *36*, 401-409.
179. Jouravel, N.; Sablin, E.; Togashi, M.; Baxter, J. D.; Webb, P.; Fletterick, R. J., Molecular basis for dimer formation of TRbeta variant D355R. *Proteins* **2009,** *75* (1), 111-7.
180. Bevan, C. L.; Hoare, S.; Claessens, F.; Heery, D. M.; Parker, M. G., The AF1 and AF2 domains of the androgen receptor interact with distinct regions of SRC1. *Mol Cell Biol* **1999,** *19* (12), 8383-92.
181. Heery, D. M.; Zacharewski, T.; Pierrat, B.; Gronemeyer, H.; Chambon, P.; Losson, R., Efficient transactivation by retinoic acid receptors in yeast requires retinoid X receptors. *Proc Natl Acad Sci U S A* **1993,** *90* (9), 4281-5.
182. Heery, D. M.; Pierrat, B.; Gronemeyer, H.; Chambon, P.; Losson, R., Homo- and heterodimers of the retinoid X receptor (RXR) activated transcription in yeast. *Nucleic Acids Res* **1994,** *22* (5), 726-31.
183. Lim, W.; Nguyen, N. H.; Yang, H. Y.; Scanlan, T. S.; Furlow, J. D., A thyroid hormone antagonist that inhibits thyroid hormone action in vivo. *J Biol Chem* **2002,** *277* (38), 35664-70.
184. Chou, C. K.; Chi, S. Y.; Chou, F. F.; Huang, S. C.; Wang, J. H.; Chen, C. C.; Kang, H. Y., Aberrant Expression of Androgen Receptor Associated with High Cancer Risk and Extrathyroidal Extension in Papillary Thyroid Carcinoma. *Cancers (Basel)* **2020,** *12* (5).