

A BIOPHYSICAL UNDERSTANDING OF
MUTATIONS IN THE HUMAN THYROID
HORMONE RECEPTOR ALPHA

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

Beatriz Romartínez Alonso MSc (Huesca)
Department of Molecular and Cell Biology
University of Leicester

June 2017

A biophysical understanding of mutations in the human Thyroid Hormone Receptor alpha

Abstract

Thyroid Hormone Receptor alpha (TR α) is a transcription factor involved in the regulation of the T3 target genes expression in response to T3. TR α plays multiple critical roles in development and growth in addition to regulate the metabolism in adult organisms. TR α regulates the transcription of T3 target genes by recruiting corepressor or coactivator factors to the promoters of the target genes in response to T3. Corepressors such as SMRT are recruited by unliganded TR α , whereas coactivators such as GRIP1 are recruited in response to T3 binding to the receptors. Several mutations in the *THRA1* gene have been recently identified in patients showing symptoms of tissue-specific hypothyroidism, delayed growth and development, and severe constipation. These mutations lead to a disorder called Resistance to Thyroid Hormone due to mutations in TR α (RTH α) whose degree of impairment depends on the location and severity of the mutation. Three mutations affecting the C-terminal or AF-2 domain of the TR α generate truncated TR α proteins that show aberrant interactions with coregulator proteins and a complete absence of T3 response.

The aims of this thesis are to understand the molecular pathology of the mutant TR α as well as to obtain a deeper insight into the structural basis for repression by the mutant TR α using structural and biochemical approaches.

The biophysical characterization demonstrates that mutant TR α impairment is mainly due to the inability of the mutant TR α to recruit coactivators in response to T3. Consequently, TR α mutants are constitutively bound to corepressors and repress T3 target gene transcription. X-rays structural data indicated that TR α mutants are able to bind and accommodate T3 in the hydrophobic pocket in an identical place and conformation as in the WT.

Acknowledgements

I wish to sincerely thank my PhD supervisor, Professor John Schwabe for providing me this opportunity to improve my skills and contribute to his scientific research. Also Dr. Andrew Jamieson, Dr. Cyril Dominguez and Dr. Ralf Schmid for their advice and support throughout all these years. I would like to extend my gratitude as well to Dr. Louise Fairall, Dr. Peter Watson and Dr. Christopher Millard for their invaluable technical help in the lab and their experienced advice.

I also want to thank Professor Krishna Chatterjee and Dr. Maura Agostini for the warm welcome and stay I received in the Metabolic Research Labs in Cambridge. It was an amazing experience which I enjoyed as much as I learned from.

I would like to extend my gratitude to Schwabe's lab for the friendly and helpful environment. I am particularly grateful to Dr. Gregg Hudson, Dr. Yun Song and almost Dr. Almutasem Saleh for their support in the lab. And to Ms. Irene Nigi for all the fun we shared in our countless and unforgettable trips around Europe.

I would also like to acknowledge Veronica Lacasta, María Grasa, Noelia Salinas, Sheila Vivas, Imma Cambra and Carmen Arnal for their huge passion for life, great mood and infinite support even though I have been living outside our dear home city for as long as I can remember. Thank you for always remembering me.

My most sincere gratitude for the unconditional support received goes to my parents María Pilar and José María, and my brother, Rubén. Thank you for tolerating all my defects and for encouraging me despite the difficulties. Thank you Ru for all the laughs and moments shared in the UK, as well as all the parties and trips.

Finally, I wish to thank my marvelous boyfriend, Daniel Bastaros, for accompanying me in this terrifying adventure, for his making these the happiest years of my life, and for, why not, his invaluable help in English pronunciation. Thank you xikinin, because without you and your endless love none of this would have ever been possible.

Dedicated to

José María Romartínez Infante

Siempre con nosotros

List of Contents

<i>Abstract</i>	<i>ii</i>
<i>Acknowledgements</i>	<i>iii</i>
<i>List of Contents</i>	<i>v</i>
<i>Table of Contents</i>	<i>v</i>
<i>List of Figures</i>	<i>ix</i>
<i>List of Tables</i>	<i>xi</i>
<i>Abbreviations</i>	<i>xii</i>

Table of Contents

Chapter 1 - Introduction	1
1.1 - Nuclear Hormone Receptors	2
1.1.1 - Classes of nuclear receptors	4
1.1.2 - NR structure	6
1.1.3 - The DNA-binding domain	7
1.1.4 - The response elements recognition by the DBD	7
1.1.5 - The ligand-binding domain	10
1.2 - Nuclear Hormone Receptor Mechanism of Action	14
1.2.1 - Ligand-induced NR activation	15
1.2.2 - Coactivators and transcriptional activation	18
1.2.3 - Corepressors and transcriptional repression	22
1.2.4 - Full-length NR complexes behaviour on the DNA	26
1.3 - Genetic disorders of nuclear receptors	30
1.4 - Mechanism of thyroid hormone action	31
1.4.1 - Human Thyroid Hormone Receptor	33
1.4.2 - Ligand binding pocket of Thyroid Hormone Receptor	35
1.4.3 - Human Thyroid Receptor: mutations and disease	38
1.5 - Aims and objectives	40
Chapter 2 - Materials and Methods	44
2.1 - Materials	44
2.1.1 - Plasmid constructs.....	44
2.1.2 - Primers	45
2.1.3 - Bacterial cell lines	45
2.1.4 - Mammalian cells	46

2.1.5 - Standard Chemicals and Reagents	46
2.2 - Generating protein constructs and mutagenesis: cloning, DNA sequencing and DNA purification methods	47
2.2.1 - Primer design	47
2.2.2 - Cloning procedure	48
2.2.3 - Site Directed Mutagenesis	48
2.2.4 - Small-scale plasmid DNA purification	49
2.2.5 - Large-scale plasmid DNA purification	50
2.2.6 - Determination of DNA concentration	51
2.2.7 - Plasmid sequencing	51
2.3 - Mass spectrometry analysis	51
2.4 - Protein expression and purification	52
2.4.1 - Cell transformation and growth	52
2.4.2 - GST-tagged protein purification	52
2.4.3 - Ionic exchange chromatography	53
2.4.4 - Size exclusion chromatography	53
2.4.5 - Protein concentration and quantification	53
2.4.6 - Protein analysis by SDS-PAGE	54
2.5 - Peptide synthesis	54
2.5.1 - Measurement of peptide concentration by NMR	57
2.5.2 - Peptide coupling	58
2.6 - Fluorescence anisotropy assay	58
2.7 - Freeze-drying	59
2.8 - Circular dichroism	60
2.9 - Crystallization trials	61
2.9.1 - Protein sample preparation	61
2.9.2 - Cryoprotection, collection and freezing of crystals	62
2.9.3 - Data collection	63
2.10 - JEG-3 mammalian cells: tissue culture and transient transfection	63
2.10.1 - Preparation and culture of JEG-3 cells	63
2.10.2 - Seeding and passaging of JEG-3 cells	63
2.10.3 - Transient transfection and cotransfection of JEG-3 cells	64
2.11 - Luciferase and β-galactosidase dual reporter assay	64
2.12 - Western blot analysis	66
Chapter 3 - Biophysical characterization of the interaction between WT and mutant TRα LBDs with their coregulator proteins	68

3.1 - Introduction	68
3.2 - Design of WT and LBD TRα constructs	73
3.3 - Optimization of protein expression and purification	74
3.4 - Fluorescence anisotropy assays to study the interaction between the TRα LBD and coregulator peptides	77
3.4.1 - Optimization of fluorescence anisotropy assays	79
3.4.2 - WT and mutant LBDs bind to SMRT peptides with similar affinity.....	80
3.4.3 - Neither WT nor mutant LBDs are able to interact with coactivator peptides in the absence of T3.....	82
3.4.4 - Mutant LBDs do not release corepressor peptides in response to T3	83
3.4.5 - Mutant LBDs are not able to interact with coactivator peptides even in the presence of T3	85
3.5 - Circular dichroism study of protein folding and peptide binding interactions.....	87
3.5.1 - T3 increases the thermal stability of the mutant TR α LBDs	88
3.5.2 - Study of the ability of mutant LBDs to bind coregulator peptides	90
3.5.3 - Liganded mutant LBDs are able to interact with corepressor peptides	92
3.6 - Discussion	94
Chapter 4 - Action of stapled peptides and T3 analogues in corepressor and coactivator interaction.....	98
4.1 - Introduction	98
4.2 - Receptor binding and corepressor displacement properties of T3 analogues	100
4.2.1 - Most of T3 analogues increase the thermal stability of the WT and mutant LBDs	102
4.2.2 - T3 analogues show moderate potency to release corepressor peptides	104
4.3 - Corepressor displacement properties of stapled corepressor peptides	106
4.3.1 - Design of stapled corepressor peptides	107
4.3.2 - WT and mutant LBDs bind stapled corepressor peptides with great affinity	109
4.4 - None of the T3 analogues show considerable effect in coactivator recruitment....	111
4.5 - Discussion	113
4.5.1 - Design of TR α improved analogues	116
Chapter 5 - Investigating corepressor and mutant LBDs interaction: a structural approach	119
5.1 - Introduction	119
5.2 - Crystallization trials.....	121
5.2.1 - Initial crystallization studies	121
5.2.2 - Optimization of initial crystallization conditions.....	123

5.2.3 - Modification of LBD constructs	126
5.2.4 - Structure determination.....	129
5.2.5 - Asymmetric unit and crystal packing.....	135
5.2.6 - Overall architecture of the P393G T394X LBD	140
5.2.7 - Crystallization trials of T3 analogues liganded LBDs	146
5.3 - Discussion	151
Chapter 6 - Functional characterization of potentially pathogenic TRα variants	153
6.1 - Introduction	153
6.2 - Identification of novel pathogenic human TR α LBD variants by searching the ExAC	156
6.2.1 - Potentially pathogenic human TR α LBD variants functional analysis	162
6.2.2 - Expression of Flag-tagged TR α variants.....	168
6.3 - Discussion	169
Chapter 7 - Discussion and future perspectives.....	174
7.1 - Molecular pathology of the mutant TR α LBDs	174
7.2 - Structural basis of TR α mechanism of repression	175
7.3 - Functional characterization of potentially pathogenic TR α variants.....	176
7.4 - Stapled peptides and T3 analogues investigation in corepressor and coactivator interaction.....	177
7.5 - TR α as pharmacological target	178
Appendix 1: Table of Oligonucleotides.....	184
Appendix 2: Buffers and Solutions	187
Appendix 3: Fluorescence anisotropy results of T3 analogues.....	192
Chapter 8 - References	194

List of Figures

Figure 1.1: Schematic view of the structural and functional organization of nuclear receptor superfamily.....	6
Figure 1.2: ER DNA-binding domain structure as determined by NMR.....	7
Figure 1.3: Structural organization of two DNA-binding domains bound to DNA.....	8
Figure 1.4: Structural organization of the NR ligand-binding domain.	11
Figure 1.5: Typical dimerization surfaces that form between receptor LBDs.	13
Figure 1.6: Ligand-induced NR activation.....	15
Figure 1.7: Functional domains of SRC/p160 family members and LBD:CoA interaction interface.	20
Figure 1.8: Corepressor functional domains and LBD interaction.	24
Figure 1.9: Model of the assembly of the core SMRT/NCoR repression complex.....	25
Figure 1.10: Overall structural organization of full-length NRs in the DNA.	29
Figure 1.11: Schematic of thyroid hormone action.	33
Figure 1.12: Scheme of TR isoforms and T3.....	35
Figure 1.13: Superimposition of the LBP structure of TR α and TR β	37
Figure 1.14: Schematic representation of the TR α 1 and TR α 2.	39
Figure 1.15: TR α and TR β tissue distribution.	39
Figure 1.16: Molecular modelling of the mutant TR α LBDs and their mechanism of action.....	41
Figure 2.1: Plasmid maps.	45
Figure 2.2: Rosetta 2 <i>E. coli</i> cells plasmid.	46
Figure 2.3: Overview of overhanging PCR used in the production of mutant genes.	49
Figure 2.4: ¹ H NMR spectrum of p-nitrophenol.	58
Figure 2.5: CD spectrum of the purified E403X LBD mutant protein.	61
Figure 2.6: Schematic representation of the luciferase and β -galactosidase reporter assay.	66
Figure 3.1: DNA and amino acid sequence of C-terminal WT and mutant TR α	69
Figure 3.2: Disorder and secondary structure predictions of full-length TR α	74
Figure 3.3: Expression and purification of TR α constructs.	75
Figure 3.4: Purification of mutant TR α LBD constructs by HiTrap Q IEX.....	76
Figure 3.5: WT purification process.	77
Figure 3.6: Binding assays of native SMRT peptide to TR α	81
Figure 3.7: Binding assays of BODIPY-labeled GRIP1 coactivator peptides to TR α LBDs.	83
Figure 3.8: Binding assays of SMRT peptide to TR α LBDs.	85
Figure 3.9: Binding assays of GRIP1 peptide to TR α LBDs.....	86
Figure 3.10: CD spectra of the WT and mutant LBDs.....	88
Figure 3.11: Temperature induced denaturation experiments of the TR α LBDs.....	89
Figure 3.12: Temperature induced denaturation experiments of the TR α LBDs.....	91
Figure 3.13: Melting temperature values of the different LBDs.....	93
Figure 3.14: Summary of coactivator and corepressor binding assay results showing the K _d in the presence and absence of T3.....	97

<i>Figure 4.1: T3 analogues</i>	101
<i>Figure 4.2: Temperature induced denaturation experiments of the TRα LBDs</i>	103
<i>Figure 4.3: Apparent dissociation constant of T3 analogues</i>	105
<i>Figure 4.4: Design and CD study of stapled peptides</i>	109
<i>Figure 4.5: Binding assays of stapled SMRT peptide to TRα LBDs</i>	110
<i>Figure 4.6: Apparent dissociation constant of T3 analogues</i>	112
<i>Figure 4.7: Fitting of T3 analogues in the LBD binding cavity as if they were occupying the same position as the hormone</i>	114
<i>Figure 4.8: NH-3 and CO23 chemical formula and NH-3 fitting in the LBP of mutant LBD</i>	117
<i>Figure 5.1: Purification and crystallization of the P393G T394X LBD: T3 (1: 5) complex</i>	134
<i>Figure 5.2: P6₄22 space group description</i>	136
<i>Figure 5.3: Molecules in the crystallographic asymmetric unit</i>	137
<i>Figure 5.4: Crystal packing arrangement</i>	139
<i>Figure 5.5: Structure of P393G T394X LBD construct</i>	140
<i>Figure 5.6: Alignment of the sequences for the P393G T394X LBD and WT LBD</i>	141
<i>Figure 5.7: Significant differences between P393G T394X LBD and the WT LBD</i>	142
<i>Figure 5.8: Surface view of the binding cavity</i>	143
<i>Figure 5.9: Structural model of P393G T394X LBD in interaction with corepressor and coactivator peptides</i>	145
<i>Figure 5.10: Biophysical characterization of P393G T394X LBD</i>	146
<i>Figure 5.11: P393G T394X LBD:JM13 (1:5) complex</i>	150
<i>Figure 6.1: Crystallographic structure of TRα LBD showing the position of selected variants</i>	161
<i>Figure 6.2: Scheme of the functional assays performed in mammalian cells</i>	163
<i>Figure 6.3: T3-dependent transcriptional activation and dominant negative results of the S330X, L367P and P399Q variants</i>	164
<i>Figure 6.4: T3-dependent transcriptional activation and dominant negative activity of the D166G, P193S, A214V, I222V, V394I, S330L, V353I and T314P variants</i>	165
<i>Figure 6.5: T3-dependent transcriptional activity and dominant negative results of the A344V, R356C, H358P and M369L variants</i>	167
<i>Figure 6.6: Expression of N-terminal flag-tagged TRα variants</i>	168
<i>Figure 6.7: Position and function of the L369</i>	170
<i>Figure 6.8: Position and function of the P399</i>	172
<i>Figure 7.1: Model of the SMRT corepressor complex interacting with RXRα - TRα heterodimer bound to chromatin</i>	182

List of Tables

<i>Table 1.1: Human Nuclear Receptors: nomenclature and ligands.</i>	5
<i>Table 1.2: Chromatin remodeling factors recruited by SRC/p160 coactivator proteins.</i>	21
<i>Table 2.1: An overview of the plasmids.</i>	44
<i>Table 2.2: Peptides used in this project.</i>	56
<i>Table 3.3: Summary and comparison of the clinical features, genotype and phenotype of RTHα patients.</i>	71
<i>Table 3.4: Melting temperature values of the different LBDs.</i>	92
<i>Table 5.1: Initial crystallization experiments.</i>	122
<i>Table 5.2: Optimization of initial crystallization conditions.</i>	124
<i>Table 5.3: Crystals formed in the Stura and MacroSol and NR LBD screens using a rMMS experimental set up.</i>	125
<i>Table 5.4: Optimization process of initial hits conditions.</i>	128
<i>Table 5.5: Data collection and refinement statistics.</i>	132
<i>Table 5.6: Pictures of the P393G T394X LBD crystals containing T3 analogues.</i>	147
<i>Table 5.7: Data collection and refinement statistics.</i>	148
<i>Table 6.1: TRα variants found in the ExAC database.</i>	157
<i>Table 6.2: Output of the THRA gene searching in ExAC database.</i>	158

Abbreviations

A382PfsX7	Frame-shift mutation caused by a single nucleotide deletion that shifts the reading frame at codon 382 and alters the six subsequent residues and then introduces a premature stop codon, truncating the protein 22 residues before the C-terminal.
aa	Amino acid
AF-1	Activation function-1 domain
AF-2	Activation function-2 domain
AR	Androgen receptor
ARC	Activator-recruited cofactor
ATP	Adenosine triphosphate
Amp	Ampicillin
Boc-K	Boc-acetyl-lysine
bp	Base pair
BSA	Bovine serum albumin
°C	Celsius degrees
cAMP	Cyclic AMP
CCP4	Collaborative computational project number 4
CD	Circular dichroism
cDNA	Complementary deoxyribonucleic acid
ChIP	Chromatin immunoprecipitation
Chl	Chloramphenicol
CO₂	Carbon dioxide
C-terminal	Carboxyl terminal
CoREST	Corepressor to REST (RE1 Silencing Transcription Factor/Neural Restrictive Silencing Factor)
CREB	cAMP response element
CTE	C-terminal extension
D₂O	Deuterium oxide
DAD	Deacetylation activation domain
DBD	DNA binding domain
DCM	Dichloromethane or methylene chloride
DIPEA	N,N-Diisopropylethylamine

DLS	Diamond Light Source
DMF	Dimethyl formamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide triphosphate
DR	Direct repeats
DRIP	VDR-interacting protein
DTT	Dithiothreitol
E403X	Non-sense mutation that results in the introduction of a premature stop codon, truncating the protein 8 amino acids before the C-terminal.
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylene diamine tetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
ERα	Estrogen receptor alpha
F397fs406X	Frame-shift mutation derived from a single nucleotide insertion that results in a frame-shift from codon 397 and a premature stop codon introduction nine residues after the nucleotide change.
FCS	Fetal calf serum
FITC	Fluorescein isothiocyanate
Fmoc	Fluorenylmethyloxycarbonyl
FP	Fluorescence polarization
FA	Fluorescence anisotropy
FT3	Free or available circulating T3
FT4	Free or available circulating T4
g	Times gravity
GPS2	G protein pathway suppressor 2
GR	Glucocorticoid receptor
GRIP 1	Glucocorticoid receptor-interacting protein 1 or nuclear receptor coactivator 2 (NCoA-2)
GST	Glutathione S-transferase
H	Helix
HAT	Histone acetyltransferase

HCTU	2-(6-Chloro-1H-Benzotriazole-1-Y1)-1,1,3,3,-tetramethylamminium hexafluorophosphate
HDAC	Histone deacetylase
HEPES	4-(2-hydroxyethyl) piperazine-1-ethanesulphonic acid
HID	HDAC interaction domain
HiTrap Q HP	HiTrap Q ion exchange chromatography high performance
HRE	Hormone response element
IPTG	Isopropyl β -D-1-thiogalactopyranoside
IR	Inverted repeats
JEG-3	Choriocarcinoma cell line
K	Degrees Kelvin
Kd	Dissociation constant
kDa	Kilo Daltons
l	Liter
LAT 1,2	L-type amino acid transporters 1 and 2
LB	Lauria-Bertani
LBD	Ligand binding domain
LBP	Ligand binding pocket
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LRH-1	Liver receptor homologue 1
LSD-1	Lysine specific demethylase 1
L-T3/T3	3,3',5-Triiodo-L-thyronine
L-T4/T4	3,3',5,5'-Tetraiodo-L-thyronine or thyroxine
L-T2/T2	3,3'-Diiodo-L-thyronine
M	Molar
Maldi-Tof	Matrix-assisted laser desorption/ionization time of flight
MCT	Monocarboxylate transporter
MED1	Mediator of RNA polymerase II transcription subunit 1
β-MeSH	β -Mercaptoethanol
mg	Milligram
ml	Milliliter
mM	Millimolar
MR	Mineralocorticoid receptor

MRC	Medical research council
mRNA	Messenger RNA
N₂	Nitrogen
N-terminal	Amino terminal
NCoR	Nuclear receptor corepressor
ND	Non detected
NGFI-B	Nerve growth factor IB
NE	Nuclear extract
nl	Nanoliter
nm	Nanometre
nM	Nanomolar
NMP	N-methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
NCT	Non-crystallographic translation
NR	Nuclear Hormone Receptors
NuRD	Nucleosome remodeling deacetylase complex
p53	Protein 53 kDa
PR	Progesterone receptor
ONPG	O-nitrophenyl-β-D-galactopyranoside
Opti-MEM	Reduced serum media
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEG MME	Polyethylene glycol monomethyl ether
PNACL	Protein and nucleic acid chemistry laboratory
PPAR	Peroxisome proliferator-activated receptor
PSF	Penicillin, streptomycin and amphotericin B
PTM	Post translational modification
POL II	DNA polymerase II
RAR	Retinoic acid receptor
rMMS	Random microseed matrix-screening
RNA	Ribonucleic acid

RONN	Regional order neural network
ROR	RAR-related orphan receptor
rpm	Revolutions per minute
rT3	3,3',5'-Triiodothyronine or reverse T3
RTHα/β	Resistance to thyroid hormone α/β
RXR	Retinoid X receptor
S75 GL	Superdex TM 75 10/300 gel filtration column
SANT	Swi3, Ada2, NCoR, TFIIB domain
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SDS	Sodium dodecyl sulphate
SEM	Standard error of mean
SHGB	Serum sex hormone-binding globulin
Sin3a/b	Switch independent 3 a/b
SF-1	Steroidogenic factor 1
SMRT	Silencing mediator of retinoid or thyroid hormone receptor
SPPS	Solid phase peptide synthesis
SWI/SNF	Switch/sucrose non fermentable
Swi3	Switch 3
SWR1	SWI2/SNF2 related ATPase
T-X100	Triton-X 100
TAT	Transactivator of transcription of human immunodeficiency virus (HIV)
TCA	Trichloroacetic acid
TCEP	Tris (2-Carboxyethyl) phosphate
TES	Triethylsilane
TEV	Tobacco etch virus
TFA	Trifluoroacetic acid
<i>THRA1</i>	Thyroid hormone receptor alpha 1 gene
T_m	Melting temperature
TMF	Tetramethylsilane
TRAP	TRs associated protein
TRα/TRβ	Thyroid Hormone Receptor α/β
Triac	3,5,3'-Triiodothyroacetic acid

Tris	Tris (hydroxymethyl) aminomethane
v/v	Volume/volume
v/w	Volume/weight
VDR	Vitamin D receptor
WD40	Tryptophan-aspartic acid repeat
μl	Microliter
μm	Micrometre
μM	Micromolar
WT	Wild type

Chapter 1 - Introduction

The ability to reproduce is perhaps the most significant property of living cells and organisms. In order to survive as a species, living organisms transmit their genetic information for countless generations with nearly perfect fidelity. This continuity in transmitting the fundamental structures and mechanisms of life requires an extremely stable molecule that remains practically the same over millions of years, deoxyribonucleic acid (DNA). This molecule contains the essential instructions to create a cell and is shared by all living organisms. The essential mechanisms that direct effective storage, expression, and division have remained nearly unchanged as well. In addition, the DNA molecule has been selected by evolution not only to store and transmit the genetic information of all living cells, but also the molecule that defines individual species and distinguishes them from one another.

However, effective survival requires adaptation. According to Charles Darwin's *Origin of Species*, "the preservation of favourable individual differences and variations, and the destruction of those which are injurious, I have called Natural Selection, or the Survival of the Fittest" (Darwin 1860). The environment changes over time and the species have to adjust to the new environments in order to survive. Accordingly, living organisms have to change along with the environment and, in order for these changes to be transmitted, the genetic instructions have to be modified. These modifications are due to infrequent mistakes in the DNA replication process and lead to changes in the nucleotide sequence of DNA, producing genetic mutations. Occasionally, a beneficial mutation arises and allows for an organism or cell to survive in the new environment and the change prevails. Therefore, the structure of the DNA not only allows the living cells to preserve their genetic material and duplicate it with near-perfect fidelity, but also it allows change and adaptation.

As evolution advances, new structures, processes, or regulatory mechanisms are acquired, derived from a series of small changes or mutations. Simple eukaryotic cells such as yeast evolved from the last common ancestor, developing new genes and regulatory pathways not present in prokaryotes (Margulis 1993). The advantages of cellular specialization led to the evolution of multicellular organisms, which contain

hundreds of different cell types, each specialized for some function that supports the entire organism and provides the basic units of tissues and organs. However, every single cell in the organism contains a virtually identical genome. This phenomenon cannot be explained by the simple scheme of the genetic code and requires an explanation of gene expression and control.

Understanding gene transcription regulation has become essential to understand how the expression of the same genome can produce such different outcomes. The ability to regulate transcription is important for a cell to respond to both internal and external signals and thus differentiate into the correct cell type. Although DNA carries all necessary information for producing a new cell, the expression of specific proteins determines the cell fate. The expression pattern is determined by the environment: sequential response to a combination of internal and external signals. Regulation of these signals determines the proteome of a specific cell type and moment of development. In adult organisms, cells with different functions exchange a wide variety of signals as well in order to exchange information and coordinate metabolic activities. In all these cases, the signal represents information that is detected by specific receptors and converted to a cellular response, which usually involves a change in gene expression.

These receptors, called nuclear receptors, are able to specifically recognize different signals and respond to them by altering the transcription pattern. Nuclear receptors directly interact with DNA, modifying the gene expression by recruiting large protein complexes. The ability to activate or repress gene expression in response to specific signals is critical to set the transcriptome and proteome of a cell and thus, coordinate the cell function.

1.1 - Nuclear Hormone Receptors

Nuclear receptors (NRs) represent a large superfamily of diverse but evolutionarily related DNA-binding transcription factors that regulate many biologically important processes such as growth, differentiation, development, homeostasis, and circadian rhythm (Mangelsdorf et al. 1995; Yang et al. 2006). NRs provide a direct link between

signaling molecules and the transcriptional response. NRs regulate the transcription of target genes by binding directly to a large variety of hydrophobic small molecules, which include lipophilic hormones, such as steroids, retinoids, thyroid hormones, and vitamin D₃, metabolites and signaling molecules (Gronemeyer et al. 2004). These ligand-dependent transcriptional factors serve to sense both the cellular and external environment by association with their cognate ligand. Ligands trigger changes in the conformational and dynamic behavior of the receptors that in turn regulate the recruitment of coregulators and chromatin-modifying machineries. Therefore, the ultimate action of liganded NRs on target genes, after site-specific DNA binding, is to enhance the recruitment and/or function of the general transcription machinery (Roeder 1996). Some NRs, including retinoic acid receptors (RARs) and thyroid hormone receptors (TRs), exhibit a dual functionality, being able to act as silencers of transcription in absence of ligands, due to their ability to recruit corepressor complexes at the promoters of target genes, in addition to activating transcription in the presence of ligand (Chen & Evans 1995). In contrast, many NRs are orphan receptors, for which a putative physiological ligand has not yet been identified (Gallastegui et al. 2015).

Due to their ability to respond to small lipophilic molecules, NRs are considered extremely good potential as therapeutic drug targets and many molecules have been successfully developed as prescription drugs. Among the most commonly used examples are tamoxifen, a ligand for estrogen receptors (ERs) used for breast cancer, thiazolidinediones, a ligand for peroxisome proliferator-activated receptor γ (PPAR γ) used for type 2 diabetes, mifepristone, a ligand for the progesterone receptor (PR) used for fertility, and dexamethasone, a ligand for the glucocorticoid receptor (GR) used for inflammatory diseases (Burriss et al. 2013). However, since NRs are involved in almost every aspect of mammalian development, metabolism and physiology, pharmacological targeting of these receptors is restricted by the complexity of NR signaling. Therefore, attempts to develop new drugs will rely on further characterization of their molecular mechanism of action.

1.1.1 - Classes of nuclear receptors

The human nuclear receptor superfamily contains 48 NRs, each of which has a crucial and non-redundant role. Despite the highly evolutionary conserved structural organization, the NR function and mode of action are very diverse. The superfamily, which appears to be specific for metazoans (Escriva et al. 2000), was classically divided into four distinct classes based on their dimerization and DNA-binding properties (Mangelsdorf et al. 1995). However, additional characterization of the NRs based on sequence alignment and phylogenetic tree construction determined the existence of six evolutionary groups (**Table 1.1**) (Nuclear Receptors Nomenclature Committee 1999; Escriva et al. 2000; Thornton & Desalle 2000).

Group 1 includes the receptors TRs, RARs, VDR, and PPARs, as well as orphan receptors such as RORs, Rev-Erbs, CAR, PXR, LXRs, and others; these receptors heterodimerize with RXR and characteristically bind to direct repeats (although some bind to symmetrical repeats as well). Group 2 contains RXRs, COUP-TF, and HNF-4, which bind primarily to direct repeats as homodimers. Group 3 includes steroid hormone receptors, which bind to DNA half-sites organized as inverted repeats as homodimers; this large group contains the receptors GR, MR, PR, AR, and ER. Group 4 is another small group that contains the nerve growth factor-induced clone B group of orphan receptors (NGFI-B, NURR1, and NOR1) that typically bind to extended core sites as monomers. There are other small groups: group 5 that contains the steroidogenic factor 1 (SF-1) and the receptors related to *Drosophila* FTZ-F1, group 6 that includes only the GCNF1 receptor and group 0 that consists of NRs lacking either the DBD or the LBD (Nuclear Receptors Nomenclature Committee 1999; Escriva et al. 2000; Thornton & Desalle 2000).

Each type of receptor includes different isoforms: for example the TR contains 2 isoforms TR α and TR β encoded by two different genes. The isoforms in turn, may also have different products of alternate splicing or promoter usage (e.g. TR α 1 and TR α 2).

Nomenclature name		Common abbreviation	Common name	Endogenous Ligands
Group 1	NR1A1	TR α	Thyroid receptor α	T3
	NR1A2	TR β	Thyroid receptor β	T3
	NR1BA	RAR α	Retinoic acid receptor α	Retinoic acid
	NR1B2	RAR β	Retinoic acid receptor β	Retinoic acid
	NR1B3	RAR γ	Retinoic acid receptor γ	Retinoic acid
	NR1C1	PPAR α	Peroxisome proliferator-activated receptor α	Fatty acids
	NR1C2	PPAR β	Peroxisome proliferator-activated receptor β	Fatty acids
	NR1C3	PPAR γ	Peroxisome proliferator-activated receptor γ	Fatty acids
	NR1D1	REV-ERB α	Reverse-Erb α	Heme
	NR1D2	REV-ERB β	Reverse-Erb β	Heme
	NR1F1	ROR α	RAR-related orphan receptor α	Oxysterols
	NR1F2	ROR β	RAR-related orphan receptor β	Oxysterols
	NR1F3	ROR γ	RAR-related orphan receptor γ	Oxysterols
	NR1H2	LXR β	Liver X receptor β	Oxysterols
	NR1H3	LXR α	Liver X receptor α	Oxysterols
	NR1H4	FXR	Farnesoid X receptor	Bile acids
	NR1I1	VDR	Vitamin D receptor	Vitamin D
	NR1I2	PXR	Pregnane X receptor	Xenobiotics
	NR1I3	CAR	Constitutive androstane receptor	Xenobiotics
Group 2	NR2A1	HNF4 α	Hepatocyte nuclear factor 4 α	Orphan
	NR2A2	HNF4 γ	Hepatocyte nuclear factor 4 γ	Orphan
	NR2B1	RXR α	Retinoid X receptor α	9- <i>Cis</i> -retinoic acid
	NR2B2	RXR β	Retinoid X receptor β	9- <i>Cis</i> -retinoic acid
	NR2B3	RXR γ	Retinoid X receptor γ	9- <i>Cis</i> - acid
	NR2C1	TR2	Testicular orphan receptor 2	Orphan
	NR2C2	TR4	Testicular orphan receptor 4	Orphan
	NR2E1	TLX	Tailless homolog orphan receptor	Orphan
	NR2E3	COUPTF1	Chicken ovalbumin upstream promoter-transcription factor 1	Orphan
	NR2F1	COUPTF2	Chicken ovalbumin upstream promoter-transcription factor 2	Orphan
NR2F6	EAR2	ErbA-related gene	Orphan	
Group 3	NR3A1	ER α	Estrogen receptor α	Estradiol
	NR3A2	ER β	Estrogen receptor β	Estradiol
	NR3B1	ERR α	Estrogen related receptor α	Orphan
	NR3B2	ERR β	Estrogen related receptor β	Orphan
	NR2B3	ERR γ	Estrogen related receptor γ	Orphan
	NR3C1	GR	Glucocorticoid receptor	Glucocorticoids
	NR3C2	MR	Mineralocorticoid receptor	Mineralocorticoids
	NR3C3	PR	Progesterone receptor	Progestins
NR3C4	AR	Androgen receptor	Androgens	
Group 4	NR4A1	NGFIB	Nerve-growth-factor-induced gene B	Orphan
	NR4A2	NURR1	Nur-related factor 1	Orphan
	NR4A3	NOR1	Neuron-derived orphan receptor 1	Orphan
Group 5	NR5A1	SF-1	Steroidogenic factor 1	Orphan
	NR5A2	LRH-1	Liver receptor homolog-1	Orphan
Group 6	NR6A1	GCNF	Germ cell nuclear factor	Orphan
Group 0	NR0B1	DAX1	Dosage sensitive sex reversal, adrenal hypoplasia critical region, on chromosome X, gene 1	Orphan
	NR0B2	SHP	Short heterodimeric partner	Orphan

Table 1.1: Human Nuclear Receptors: nomenclature and ligands.

1.1.2 - NR structure

NR proteins show a characteristic modular arrangement of five to six domains (designated A to F, from the N-terminal to the C-terminal) based on sequence conservation and function (**Figure 1.1**) (Giguere et al. 1986). The domain structures vary in length and sequence among the receptors, although the DNA-binding domain (DBD, region C) and the ligand-binding domain (LBD, region E) have remained highly conserved through evolution. These two regions are the most important and can function independently (Green et al. 1986; Hollenberg et al. 1995). The variable N-terminal A/B domain and the D region are less conserved and the C-terminal F region is not present in all receptors.

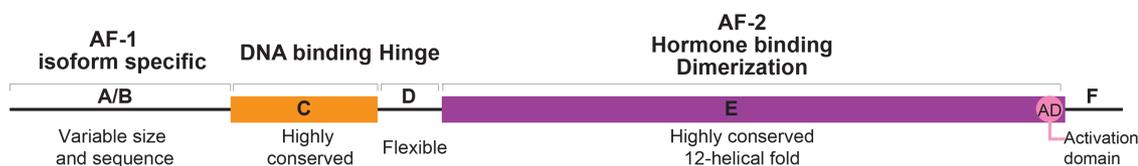


Figure 1.1: Schematic view of the structural and functional organization of nuclear receptor superfamily. The evolutionarily conserved regions C and E are indicated as boxes, and a black bar represents the diverse A/B, D, and F regions.

The extreme N-terminal A/B domain contains a flexible and highly variable region with a transcriptional activation function, termed the AF-1 or activation function-1 domain. For some receptors, including PPAR γ and AR, this domain can operate autonomously through recruitment of activating cofactors or other transcription factors (Hu et al. 1996; McEwan 2004). In addition, this domain can be the target of post-translational modifications, for instance the RARs A/B domains include several consensus phosphorylation sites (Taneja et al. 1997).

The hinge or D region is a poorly conserved domain situated between the DBD and the LBD. It is highly variable in amino acid length and sequence and is considered to serve as a hinge between the most important domains to enable them to adopt different conformations.

1.1.3 - The DNA-binding domain

The DBD is the central C region of the NRs responsible for binding to specific DNA sequences, called hormone response elements (HREs). The DBD constitutes an independent domain of approximately 70 amino acids which is highly conserved (Green & Chambon 1987; Kumar et al. 1987; Umesono & Evans 1989). The DBD structure was first determined by two-dimensional ^1H NMR using recombinant estrogen receptor (Schwabe et al. 1990) and glucocorticoid receptor (Härd et al. 1990). The structure consists of two amphipathic helices packed at $\sim 90^\circ$ to each other forming an extensive hydrophobic core. Eight conserved cysteine residues coordinate the two zinc ions which maintain the overall domain conformation. These two zinc-binding sites lie at the N-terminal of each helix (**Figure 1.2**).

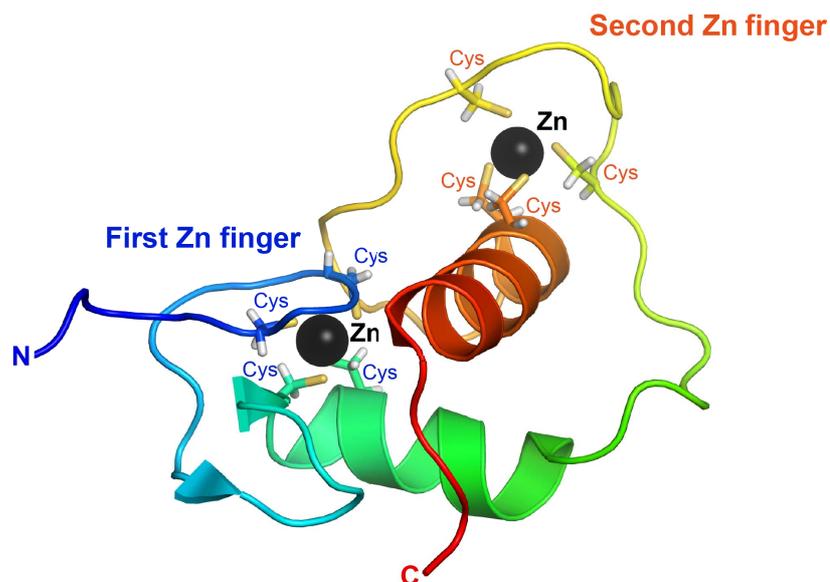


Figure 1.2: ER DBD structure as determined by NMR. Pdb code: 1HCP (Schwabe et al. 1990).

1.1.4 - The response elements recognition by the DBD

The crystallographic determination of the DBD structure bound to the DNA revealed how NRs discriminate between their response elements (Luisi et al. 1991; Schwabe et al. 1993). After the last two cysteines of the N-terminal zinc finger motif there is a recognition helix (**Figure 1.2**) which inserts directly into the major groove of DNA and

confers DNA half-site binding. Different NRs use distinct sets of amino acids on the exposed face of their DNA recognition α -helix. Within the second zinc finger, the loop around the zinc dictates the half-site spacing (**Figure 1.3**).

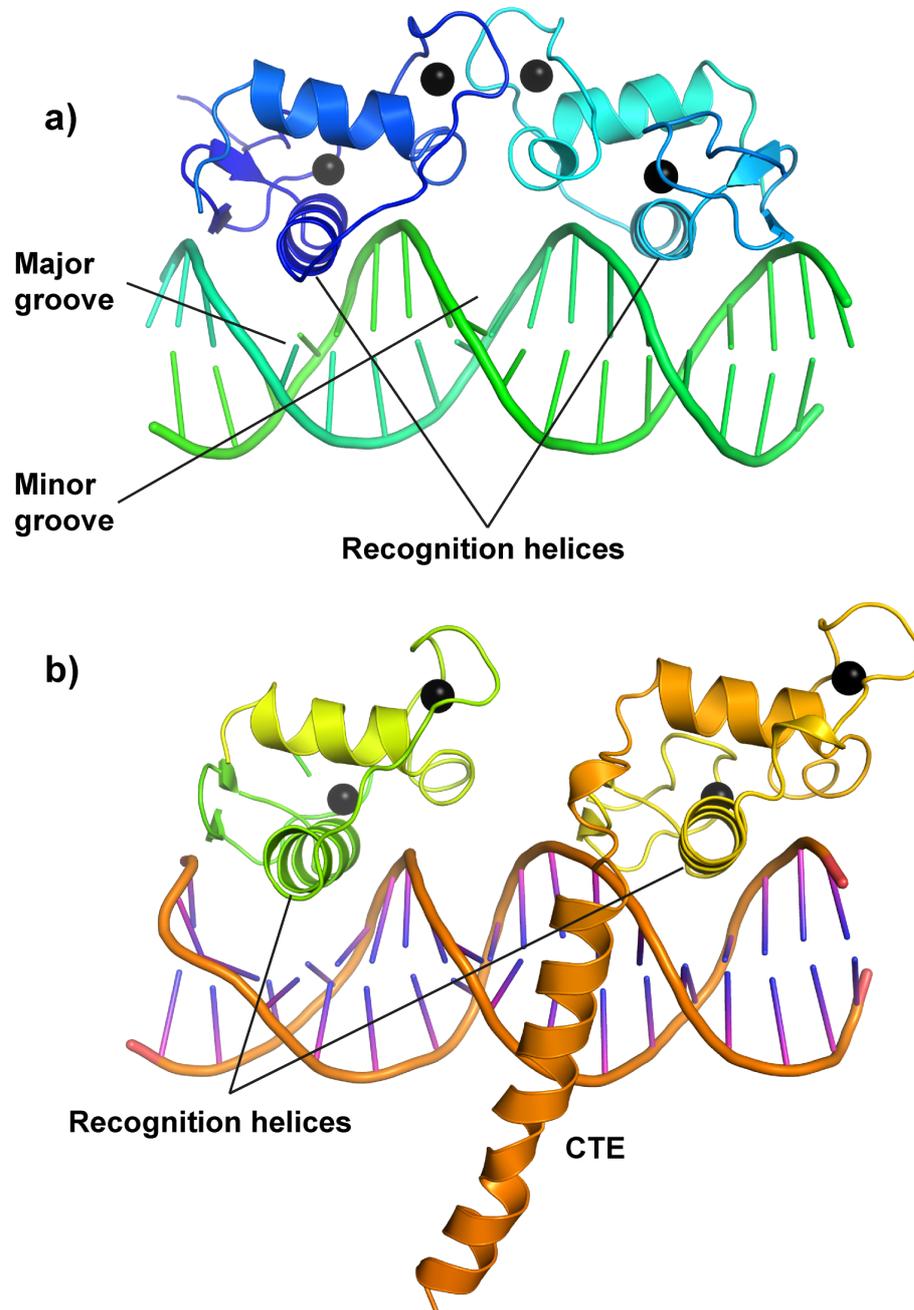


Figure 1.3: Structural organization of two DBDs bound to DNA. a) Structural elements of two ER α DBDs bound as a homodimer to a palindromic sequence of DNA separated by 3 nucleotides (pdb code 1HCQ) (Schwabe et al. 1993) b) Structural elements of two DBDs corresponding to the RXR and TR heterodimer bound to a direct repeat separated by 4 nucleotides (DR4) (pdb code 2NLL) (Rastinejad et al. 1995).

Typically, HRE are hexa-nucleotide sequences derived of the canonical sequence 5'-RGGTCA-3' (in which R is a purine), called half-sites, which are spatially oriented as direct or inverted repeats or simply as monomeric sequences that are functionally associated with their target genes (Beato 1989). The DBDs bind as homo- or heterodimers to their palindromic binding site consisting of two consensus half sites with specific base pairs between the sites. DNA target selectivity relies on the recognition of the geometry associated with the arrangement of the two half sites (Umesono et al. 1991) and, in order to recognize this geometry, two DBDs are required to interact with each other and make contacts with the minor groove of the DNA (Luisi et al. 1991). The minor groove of the DNA between the two half sites is required for the dimerization process, which implies that the DBD contributes to appropriate dimerization pairing of the receptors (**Figure 1.3**).

Steroid receptors bind to palindromic repeats of a hexameric half-site separated by 3 base pairs of spacer (IR3) as homodimers. Except ER, the steroid receptors (GR, MR, AR, and PR) recognize the consensus sequence 5'-AGAACA-3'. ER binds similar symmetric sites but with consensus 5'-AGGTCA-3' half sites (Schwabe et al. 1993; Beato et al. 1995). Therefore, the first order level of specificity is derived from the sequence and type of repetition showed in the HRE.

Monomeric NRs such as NGFI-B, Rev-Erb, ROR and SF-1 bind to extended single half-site elements and rely on the immediate flanking sequence upstream of their half site for response element discrimination (Wilson et al. 1993; Giguere et al. 1995; Harding & Lazar 1995; Charles et al. 1999).

RXR heterodimers recognize direct repeats (DR) where there is a second-order level of specificity determined by the unique spacing of the half-site repeats (Koenig et al. 1987; Näär et al. 1991; Umesono et al. 1991). This spacing between repeats comprises a small number of nucleotides, typically 1-5, and the direct repeats are referred to according to this number: DR1-DR5 (Mangelsdorf & Evans 1995). RXR forms heterodimers with a variety of non-steroid receptors such as RAR, VDR, TR, PPAR, and several others (Yu et al. 1991; Bugge et al. 1992; Hallenbeck et al. 1992; Kliewer et al. 1992; Leid et al. 1992; Marks et al. 1992; Zechel et al. 1994; Forman et al. 1995). Since the different RXR heterodimers recognize the same DRs of hexameric half-sites (5'-AGGTCA-3'),

the half-site spacing is essential for the different RXR heterodimers to recognize and discriminate their specific response element. For instance, the response element for VDR is a direct repeat spaced by 3 nucleotides, for TR the spacing is 4 nucleotides, and for RAR the spacing is 5 nucleotides (Umesono et al. 1991; Perlmann et al. 1993).

In RXR heterodimers, RXR can be situated at the upstream or downstream half-site in relation to its partner. For most heterodimers, RXR binds at the 5' response element and the heterodimer partner binds at the 3' response element (Rastinejad et al. 1995). In contrast, for PPAR-RXR heterodimers, which bind to DR1 elements, the RXR molecule binds at the 3' response element (Chandra et al. 2008). Thus, the spacing also dictates RXR heterodimer pairing orientation.

Depending on the type of receptor, the C-terminal extension plays a role in sequence recognition and/or dimerization. The TR CTE contacts the DNA minor groove and projects across the minor groove of the half site (Rastinejad et al. 1995). The minor groove binding of the CTE prevents co-occupancy of receptors on the wrong response elements (Rastinejad et al. 1995; Zhao et al. 1998).

1.1.5 - The ligand-binding domain

The LBD is less conserved than the DBD among the NRs, however numerous crystallographic structures have shown that all the LBDs share a similar overall globular fold. This comprises twelve α -helices and two or three β -sheets organized in a three-layer anti-parallel α -helical “sandwich”. The lower half of the domain is characterized by a hydrophobic cavity or pocket in which lipophilic ligands bind (**Figure 1.4**) (Wurtz et al. 1996).

The architecture of the LBD is formed by H1-H3 constituting one face; H4, H5, S1, S2, H6, H8, H9 corresponding to the central layer of the domain and H7 and H10 forming the second face (Bourguet et al. 1995; Renaud et al. 1995). The overall similarity among all the available LBD structures is particularly significant in the top half of the LBD, which includes H1, H4, H5, and H7-H10, and defines a structurally conserved region (**Figure 1.4**). The ligand binding pocket (LBP) is contained in the lower part of

the LBD, which is more variable in size and amino acid composition, to specifically accommodate the diverse ligands (Wurtz et al. 1996).

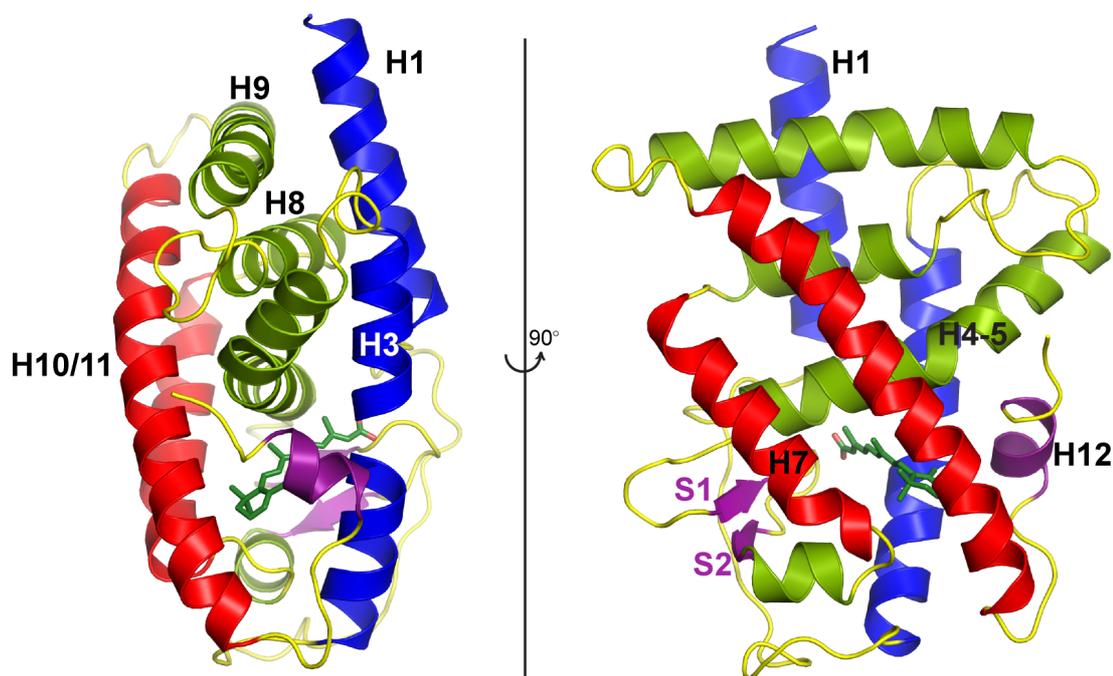


Figure 1.4: Structural organization of the NR LBD. The LBD is formed from a three-layer α -helical sandwich. The layers are shown in blue, green and red. In the lower region of the structure the central helical layer is absent, creating a mainly hydrophobic cavity in which the ligand (dark green) binds. Helix 12 and β -strands (purple) close the front and back of the ligand-binding cavity, as can be seen on the left. The structure corresponds to the RAR γ LBD bound to all-trans retinoic acid (pdb code 2LBD) (Renaud et al. 1995).

The LBP is generally located behind H4-5 and in the front of H7 and H10 (**Figure 1.4**) and has a volume that varies from almost absent in the NR4A orphan receptor ((Li et al. 2003), to more than 1400 Å³ in the subtypes of PPARs (Nolte et al. 1998). The LBP is formed predominantly by hydrophobic residues, with one or more polar residues also present that allow effective hydrogen bonding with hydrophilic groups that may be present on some NR ligands. Specific ligands are discriminated by van der Waals forces that detect the surface, volume and shape features of the ligands, as well as a few polar interactions that play an essential role in the correct positioning and enforce the selectivity of the pocket.

The LBD is functionally very complex. It mediates four distinct but related functions that can be localized within its structure: a dimerization surface, which mediates

interaction with partner LBDs; the LBP which specifically interacts with diverse and small lipophilic molecules; a coregulator-binding surface, which interacts with regulatory protein complexes such as corepressor and coactivator complexes and an activation function helix, called AF-2, which is responsible for ligand-dependent transactivation, corepressor release and coactivator recruitment. AF-2 is formed by an amphipathic α -helix located in the C-terminal which, together with helices 3 and 4, forms the coactivator-binding site.

To date, many crystal structures of LBDs forming monomers, homodimers, and heterodimers LBDs have been described, allowing comparison of the homo and heterodimerization interfaces. Comparison of the PPAR γ -RXR α heterodimer (Gampe et al. 2000) and the ER α -ER α homodimer (Brzozowski et al. 1997), revealed that the global heterodimeric structure is very similar to that of the homodimer; what is more, the structural elements producing the dimerization interface are identical. The core of the dimer interface is mainly composed of interactions mediated through the N-terminal of H9 and H10/11 composed of complementary hydrophobic residues. The residues from H7, and the L9/10 are involved in the formation of additional hydrogen bonds that further increase the specificity and stability of the dimer interface (Gampe et al. 2000).

The main differences are the symmetry and the surface of the dimer interface. Notably heterodimer interfaces are slightly asymmetric, in contrast with the nearly perfect symmetry of the homodimer interface of PPAR γ , ER α and apo-RXR α LBDs (Bourguet et al. 1995; Brzozowski et al. 1998). In steroid receptors, which preferentially form homodimers, H7, H9, H10/11 and L9/10 are longer allowing additional contacts, making the homodimeric surfaces larger (around 1700 \AA^2 for the ER α homodimer) than the heterodimeric surfaces (around 915 \AA^2 for PPAR γ -RXR α heterodimer) (**Figure 1.5**) (Brzozowski et al. 2000).

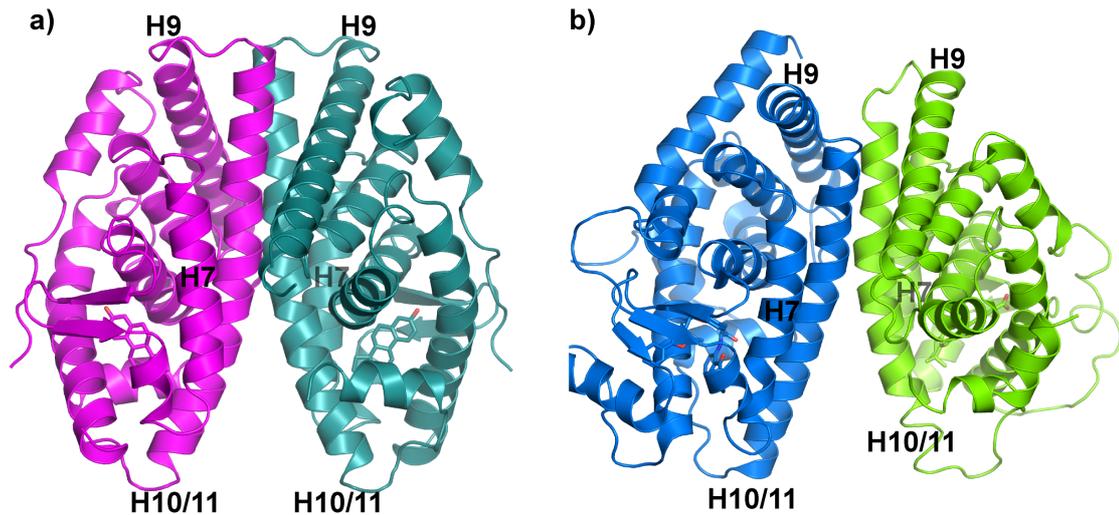


Figure 1.5: Typical dimerization surfaces that form between receptor LBDs. a) ER α -ER α homodimer structure with the helices involved in dimerization labeled (pdb code 1ERE) (Brzozowski et al. 1997). b) PPAR γ -RXR α heterodimer structure with the helices forming the dimerization interface labeled (pdb code 1FM9) (Gampe et al. 2000). Surfaces from H7, H9, and H10/11 participate in forming the dimeric interfaces.

Heterodimers can be regulated depending on the level of permissiveness exhibited by the complex (Mangelsdorf & Evans 1995). For instance, some of the RXR heterodimers, including those formed between the PPARs, LXRs, FXR, PXR and CAR, are called “permissive” since they can be activated by the specific ligand of the RXR α , 9-*cis*-retinoic acid (9cRA), as well as the cognate ligand of the partner. The non-permissive RXR partners such as RAR, TR and VDR, generally do not respond to RXR ligands (Forman et al. 1995).

Ligand binding affects the stability and propagation of signals across the heterodimerization interface, allowing integration of ligand-dependent signals across the dimer interface (Cheskis & Freedman 1996; Thompson et al. 1998; Shulman et al. 2004). The structural basis of this heterodimer-specific signaling is poorly understood, but some researchers have suggested an allosteric mechanism for the silencing of RXR signaling by a non-permissive dimer partner (DiRenzo et al. 1997; Li et al. 2004; Pérez et al. 2012; Venäläinen et al. 2009).

1.2 - Nuclear Hormone Receptor Mechanism of Action

The physiological status of the human body is highly regulated as a whole and is able to respond to environmental changes by altering endocrine signaling or metabolic parameters that lead to dramatic changes in endogenous ligand concentrations. These small lipophilic molecules serve as molecular messengers to enable communication among tissues and respond to the environment. NRs are the molecules responsible for transducing these chemical signals from their ligands into changes in gene expression in order to respond to the environmental changes.

NR transcriptional activity depends on sets of associated proteins, called coactivators and corepressors, interacting with them. Therefore, from the response elements (HRE) NRs recruit other complexes that bring repression or activation to their target genes by modifying histone tails and remodeling chromatin structure to ultimately, promote or repress transcription. In the liganded state, NRs recruit coactivator proteins that act as scaffolds to recruit complexes that harbor histone acetyltransferase activity (HAT). Usually, in the absence of ligand, NRs recruit corepressors, such as nuclear receptor corepressor 1 (NCoR 1) or silencing mediator for retinoid and thyroid hormone receptor (SMRT), which interact with histone deacetylases (HDAC) (**Figure 1.6**) (Nagy 2004).

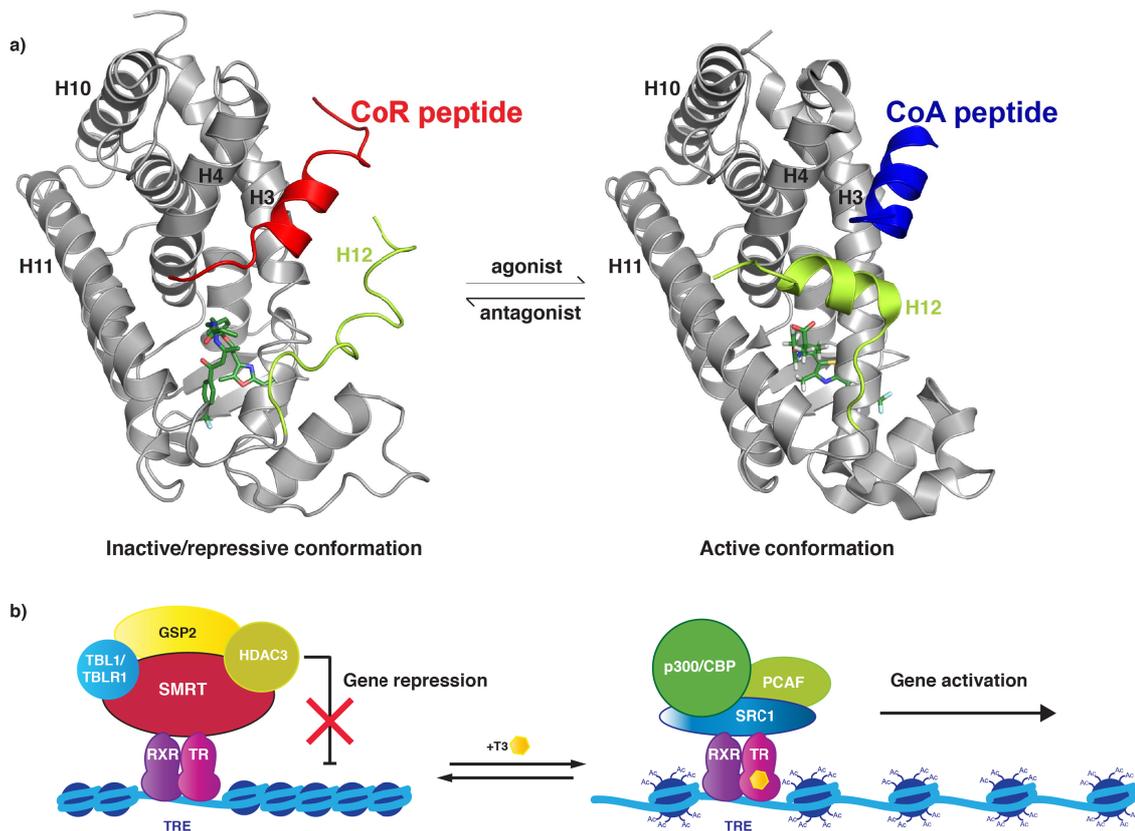


Figure 1.6: Ligand-induced NR activation. a) Structure of the PPAR α LBD bound to a SMRT corepressor peptide (red) and a SRC-1 coactivator peptide (blue) (pdb codes: 1KKQ and 2P54, respectively) (Xu et al. 2002; Sierra et al. 2007). The molecular switch involves many subtle conformational changes that globally stabilize the protein, and a dramatic conformational change of H12 which becomes ordered and caps the ligand cavity. b) Schematic representation of the NR mechanism of action showing how ligand binding modulates the interaction of the LBD with a number of coregulator complexes such as corepressor complexes with deacetylase activity and coactivator complexes with acetylase activity.

1.2.1 - Ligand-induced NR activation

NRs respond to ligand binding through distinct mechanisms depending on, firstly, the location of the receptors in the absence of ligand. In contrast with other NRs, NRs in group 3, which includes the GR, AR, ER, PR, and MR, in the absence of ligand are predominantly localized in the cytoplasm associated with molecular chaperones such as heat-shock proteins (Hsp90 and Hsp56) that restrain them in the inactive state (reviewed in (Sever & Glass 2013)). These cytoplasmic complexes are still able to recognize and respond to specific ligands, which after binding promote NR activation by dissociation of the chaperone complex, translocation to the nucleus, homodimerization and association with the response elements.

In contrast, all other NRs constitutively bind to DNA independently of ligand binding. In fact, it has been demonstrated that in the absence of ligand, a number of NRs including TR, LXR and RAR, repress target gene transcription below a basal level (Baniahmad et al. 1992; Horlein et al. 1996; Hu et al. 2003). Ligand association promotes a dramatic molecular switch in the function of these NRs that causes both relief of repression as well as an increase in target gene transcription by inducing a conformational change in the LBD (**Figure 1.6**) (Nagy 2004). In the unliganded state, these repressive transcription factors are bound to a set of transcriptional corepressors (NCoR1, SMRT), which recruit transcriptional complexes that contain specific HDACs. By deacetylation of chromatin, these enzymes generate a condensed chromatin structure over the target promoter that results in gene repression (**Figure 1.6**).

Corepressors are large scaffold proteins that include one or several regions called the NR interaction domain (CoRNR box), each of which contains an isoleucine-rich sequence (LxxxIxxx[I/L]) that specifically docks to a hydrophobic groove in the surface of the LBD comprising helices H3 and H4. Mutational assays followed by two hybrid analysis demonstrated that this motif is necessary and sufficient to mediate the binding of the corepressor to unliganded NRs (Hu & Lazar 1999; Nagy et al. 1999).

In most cases, ligand binding promotes subtle conformational arrangements in the LBD that place H12 in a stable position, causing disruption of the corepressor interaction surface and corepressor complex dissociation. This ligand-induced conformational move of H12 creates a new hydrophobic groove in the surface of the LBD formed by H3, H4 and H12 that is no longer recognized by corepressors, and allows coactivator recruitment. Molecular studies established that a specific leucine-rich sequence (LxxLL) is used by coactivators to bind NRs; one or more of these motifs are present in the NR interaction domains of the coactivators, called NR boxes (Heery et al. 1997). The determination of the PPAR γ and TR β structures bound to coactivator peptides revealed that there is a particularly important charge clamp, formed by a lysine and glutamate in H3 and H12, respectively, which interacts with the helix dipole formed by the interaction motif (Darimont et al. 1998; Nolte et al. 1998). The active position of H12 is then, essential to support coactivator binding.

Subsequent determination of the PPAR α LBD structure in complex with both a corepressor peptide and a coactivator peptide allowed the comparison of the binding sites for both peptides within the LBD (Xu et al. 2002; Sierra et al. 2007). Supported by mutational assays, the results revealed that corepressors and coactivators interact with overlapping surfaces on the LBD and implied that the binding of the two coregulators is mutually exclusive. The longer binding motif of corepressor proteins forms an extra helical turn so that it does not require H12 to be in the active position (Hu & Lazar 1999; Nagy et al. 1999). This larger corepressor interaction surface that extends underneath H12 may account for the preferential binding of corepressor over coactivator to the unliganded LBD. Therefore, despite overlap, there are various differences in their binding interfaces. For instance, the SMRT corepressor motif covers 736 Å of the PPAR α surface making additional hydrogen bond contacts with the LBD, whereas the SRC-1 motif covers only 478 Å (Xu et al. 2002; Sierra et al. 2007).

i) Mechanism of the molecular switch

Since the first determination of the three-dimensional structure of the rat TR α LBD by Wagner and co-workers in 1995 (Wagner et al. 1995) and the prediction of a common fold for all NR LBD (Wurtz et al. 1996), there has been significant progress in understanding the nuclear receptor mechanism of action at the molecular level. However, despite the biochemical and biophysical information available, the mechanism of the molecular switch carried out by the NRs in response to ligand remains unclear.

Comparison of the liganded LBD structures with the unliganded LBD structures represents a very useful approach to evaluate the changes that take place between the two conformations and discern the specific adjustments in position or interaction that occur upon ligand binding. Initially, comparison of the ligand-bound RAR γ (Renaud et al. 1995) with the apo-RXR (Bourguet et al. 1995) suggested that NRs undergo a very specific switch between two conformations and a 'mouse trap-like' mechanism was proposed where the ligand induced a structural transition to position H12 in the active conformation sealing the ligand-binding cavity.

However, the determination of the structures of other receptors in the inactive state including PPAR γ , PXR and LRH-1 revealed that there are not significant differences between the structures in the active and in the inactive state, suggesting a mechanism that would involve more subtle structural changes (Nolte et al. 1998; Watkins et al. 2001; Sablin et al. 2003). Instead of a change in the conformation of H12, the moderate rearrangements upon ligand binding would result in a global stabilization of the receptors, lowering the degree of conformational dynamics and ultimately causing the establishment of H12 in the active conformation.

The determination of the crystal structure of the apo-PPAR γ LBD also showed H12 in a disordered state supporting the idea of the apo-LBDs adopting a mobile globular state, especially in the lower region of the protein, that is stabilized after the binding of the ligand which promotes coactivator binding (Nolte et al. 1998). Further biochemical and biophysical studies also support these results. For instance, fluorescence studies showed that H12 is more dynamic in the apo-state of PPAR γ LBD compared with the liganded state of the LBD (Kallenberger et al. 2003). NMR spectroscopy indicated that the apo-state of PPAR γ LBD is in a conformationally mobile state, and that ligand binding is associated with a marked stabilization of the conformation. Proteolytic sensitivity assays and secondary structure melting studies in ER α showed that H12 is intrinsically dynamic, and that ligand binding leads to a more rigid and well ordered conformation (Leng et al. 1993; Keidel et al. 1994; Pissios et al. 2000; Nagy 2004; Raghuram et al. 2007).

The differences in the dynamic properties of the apo- and holo-receptors suggested a “dynamic stabilization model” in which in the apo-state, H12 is not fixed in any single position, but rather mobile along with other portions of the LBD. Therefore, the NRs in the inactive or apo-state resemble a molten globule.

1.2.2 - Coactivators and transcriptional activation

Ligand binding switches NRs from an inactive to an active state by inducing stabilization and conformational changes. The active conformation recruits coactivator

complexes, which contain chromatin-modifying enzymes required for transcription and, ultimately, enhance the recruitment and/or function of the transcription machinery. Proteomic co-precipitation and expression cloning approaches have been used to identify a large number of factors that interact with NRs in a ligand-dependent manner. These factors function as components of large, multiprotein complexes. Transcriptional processes are believed to be regulated through the sequential and combinatorial interactions of this large number of modulatory multiprotein complexes to reorganize chromatin (Pollard & Peterson 1998).

The p160/SRC family has been reported to be among the first complexes recruited by activated NRs, both in solution and on DNA (Cavailles et al. 1994; Halachmi et al. 1994). The p160 family includes three related genes that encode the p160 factors, called SRC1/NcoA1, TIF2/GRIP1/NcoA2 and p/CIF/AIB1/ACTR/RAC/TRAM1 (Onate et al. 1995; Kamei et al. 1996; Anzick et al. 1997; Chen et al. 1997; Hong et al. 1997; Li et al. 1997; Takeshita et al. 1997). Biochemical and structural data showed that p160 proteins could directly interact with LBDs in the active conformation through the highly conserved α -helical LxxLL motif (NR box) (Heery et al. 1997). The NR interaction domain of these factors contains three repeated motifs of the consensus sequence (Voegel et al. 1998). Although the short sequence encompassing the LxxLL motif is sufficient for NR-coactivator interaction, additional amino acids surrounding the motif appear to make additional contact with the LBD (Darimont et al. 1998; Nolte et al. 1998; McInerney et al. 1998). These residues are not conserved among different coactivators suggesting that they can play a role in specifying coactivator and NR interaction (**Figure 1.7**).

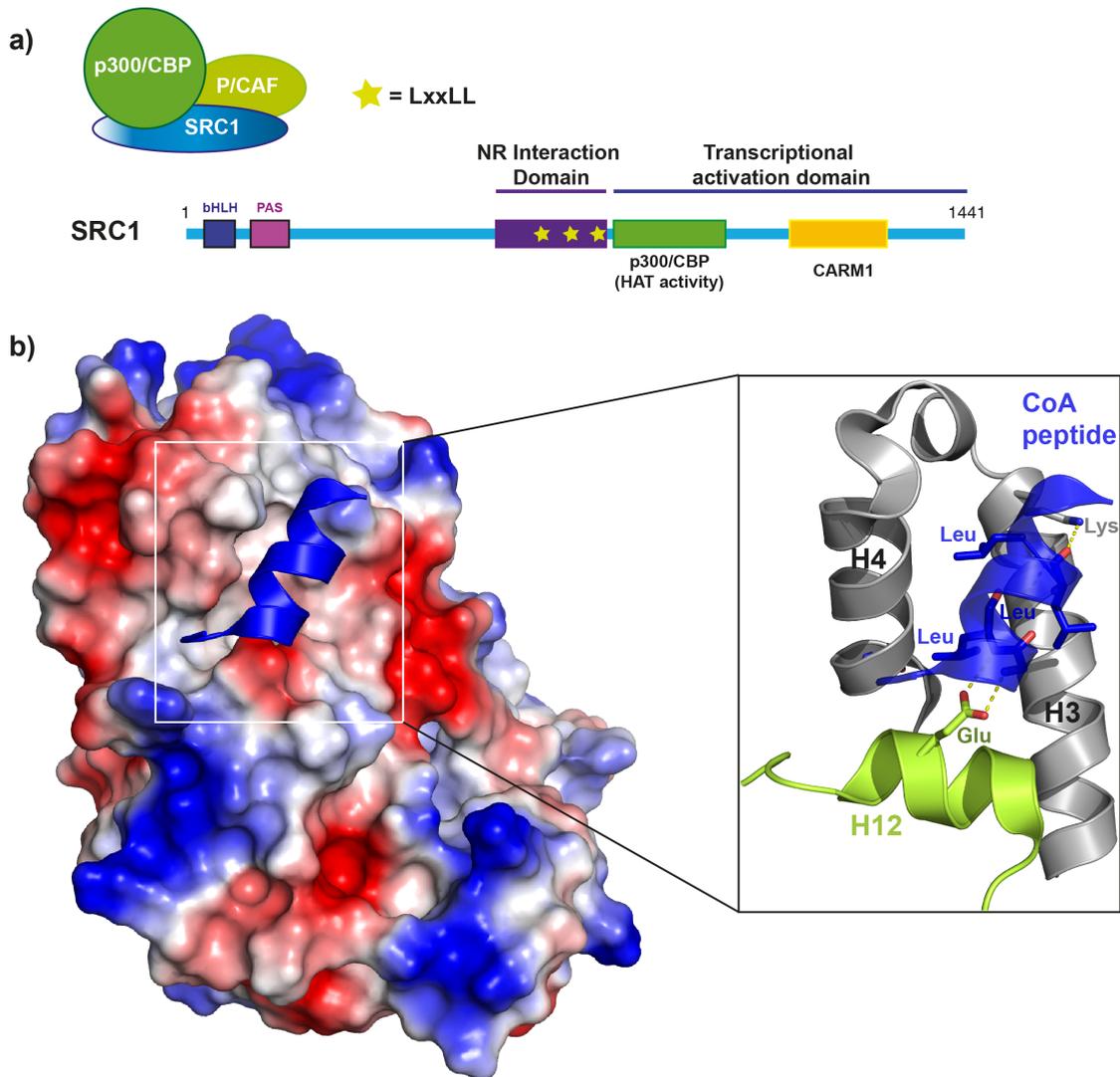


Figure 1.7: Functional domains of SRC/p160 family members and LBD:CoA interaction interface. a) Schematic organization of members of the SRC/p160 family. SRC1 acts as a scaffold to recruit histone acetyltransferase enzymes (HAT, CBP/p300, P/CAF) and methyltransferases (CARM1) to the promoters targeted by the NR. b) Crystal structure of PPAR α LBD:GW735:CoA complex in a surface view colored by charge. The enlarged view shows the leucine residues of the LxxLL motif packed within the hydrophobic cavity formed by H3 and H4 and the electrostatic interactions between the Glu from H12 in the active conformation and the Lys from H3 with the backbone of the coactivator peptide forming the charge clamp essential for the binding of coactivator proteins (pdb code: 2P54, (Sierra et al. 2007)).

SRC/p160 coactivators act as scaffold proteins to recruit three classes of chromatin remodeling factors that have been reported to play critical roles in transcriptional activation by NRs: ATP-dependent nucleosome remodeling complexes; factors that contain histone acetyltransferase activity (HAT); and coactivator-associated methyltransferases (CARM1) (Table 1.2).

Chromatin remodeling factor	Function
SWI/SNF complex	ATP-dependent nucleosome remodeling complex that causes local changes in chromatin structure (Owen-Hughes & Workman 1996)
P/CAF (Xiang-Jiao Yang et al. 1996), CBP (CREB binding protein) (Bannister & Kouzarides 1995BC), p300 (adenovirus E1A binding protein) (Ogryzko et al. 1996)	HAT activity: lysine acetylation of histones, weakening the interaction between histones and nucleosome and decondensing the chromatin. Molecular scaffolds: P/CAF interacts with CBP and p300; CBP recruits RNA polymerase II (Nakajima et al. 1996).
CARM1 (Coactivator-associated arginine methyltransferase) (Cheng 2000)	Histone 3 methylation Interaction with p160

Table 1.2: Chromatin remodeling factors recruited by SRC/p160 coactivator proteins. The table shows different factors recruited to the chromatin by coactivator proteins and their function in transcription activation.

Various coactivator multiprotein complexes, called mediators, have been reported to bind to liganded receptors and RNA polymerase II. The main function of mediator complexes is to transmit signals from the transcription factors to the RNA polymerase II. Mediator complexes include TRAP (TRs associated protein) (Fondell et al. 1996), DRIP (VDR-interacting protein) (Rachez et al. 1999) and ARC (activator-recruited cofactor) (Naar et al. 1999). These complexes can bind to NRs in the active state mainly through mediator of RNA polymerase II transcription subunit 1 (MED1), which is a common component of various mediator complexes and harbors functional LxxLL NR box motifs (reviewed in (Allen & Taatjes 2015)).

Considering the different protein complexes involved in transcription activation by NRs, it is plausible that they have a sequential role. The p160/SRC family proteins and chromatin remodeling complexes containing HAT activity are recruited to the promoter initially. Histone modification will lead to the reorganization of chromatin to a more accessible state. Then additional complexes that include TRAP/DRIP/ARC are required

to recruit the RNA polymerase II and its factors to initiate transcription (Rosenfeld et al. 2006).

Additional transcription factors, called “pioneer factors” were identified by Chromatin Immunoprecipitation (ChIP)-CHIP and ChIP-SEQ approaches (Carroll et al. 2005). These factors are involved in initiating chromatin remodeling and subsequent recruitment of additional chromatin modifiers to induce transcription. For NRs including ER α , the binding of these factors appears to be required for robust activation of a large number of target genes (reviewed in Zaret & Carroll 2011).

1.2.3 - Corepressors and transcriptional repression

Many NRs such as TRs and RARs, are constitutively bound to target promoters and exhibit transcriptional repression in the absence of ligand (Baniahmad et al. 1995). Repression is mediated by interaction with transcriptional corepressors such as NCoR and SMRT, which were originally identified as components of the repressive complex associated with unliganded RAR and TR (Chen & Evans 1995; Horlein et al. 1996).

SMRT and NCoR are large homologous proteins that can be purified from HeLa cell extracts as part of a complex with an apparent molecular weight of between one and two mega Daltons (Guenther et al. 2000; Li et al. 2000; Wen et al. 2000). Both corepressors are predicted to be mostly unstructured, with only a few regions possessing an inherently folded structure. These structured regions are believed to form the domains through which SMRT and NCoR act as scaffolds to interact with other proteins (**Figure 1.8**).

The N-terminal of the corepressors, which is the most conserved region between SMRT and NCoR1 (83 % identity), has a strong transcriptional repression activity region called repression domain 1 (RD1) (Chen & Evans 1995; Horlein et al. 1996). The amino terminal portion of RD1 recruits both GPS2 and TBL1/TBLR1 which also interact directly with each other. After these two protein-protein interaction domains, there are two other structured regions that fold into SANT-like domains (Aasland 1996). The SANT1, which is responsible for both recruiting and activating HDAC3, was called

deacetylase activation domain (DAD) (Codina et al. 2005; Guenther et al. 2001; Zhang et al. 2002). SANT2 was reported to interact directly with histone tails and was called the histone interaction domain (HID) (Yu et al. 2003; Hartman et al. 2005).

The C-terminal part of both NCoR and SMRT contains a region that specifically recognizes and binds to a hydrophobic groove in the surface of the LBD of unliganded NRs. This interaction is mediated by two conserved receptor interaction motifs called ID1 and ID2 located in the CoRNR box of the corepressors and containing the LxxxIxxx[I/L] specific sequence (Hu & Lazar 1999; Nagy et al. 1999; Perissi et al. 1999). These two motifs (17 and 19 amino acids, respectively) are conserved in both position and sequence between NCoR and SMRT and adopt an amphipathic α -helical conformation that makes additional flanking interactions with the LBD so that the binding surface for corepressors is more extensive than that for the coactivators (**Figure 1.8**). The crystal structure of the PPAR α LBD bound to the antagonist GW6471 and a peptide corresponding to the SMRT ID2 motif established that the corepressor peptide adopts a three-turn α -helix which displaces H12 from the active position since the corepressor-binding site overlaps the H12 active position (Xu et al. 2002; Wang 2006; Madauss et al. 2007).

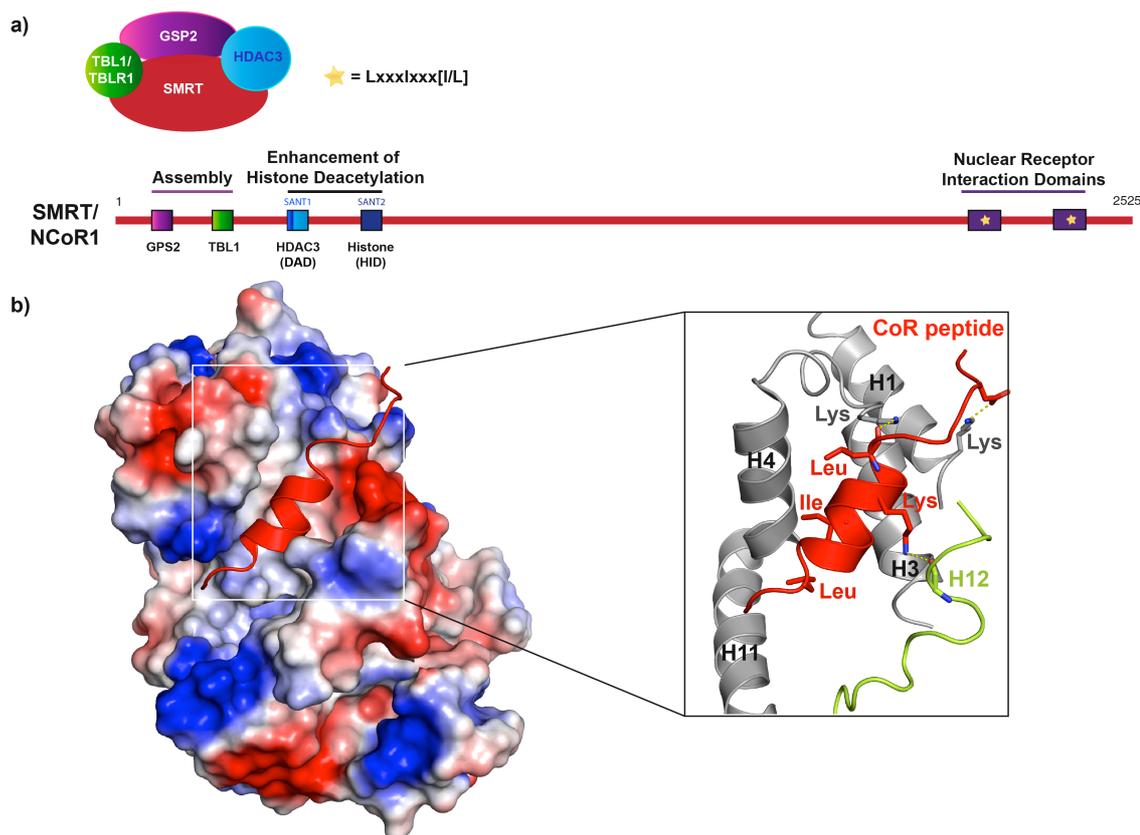


Figure 1.8: Corepressor functional domains and LBD interaction. a) Schematic view of NCoR and SMRT. Carboxy-terminal domains I and II harbor the leucine-rich domains (LxxxIxxx[I/L]) which bind the LBDs. The N-terminal, which is called the repression domain 1, contains the interaction domains that bind GPS2, TBL1, HDAC3, and the histone tails (HID). b) Crystal structure of the PPAR α LBD:GW6471:CoR in a surface view colored by charge. The enlarged view shows the leucine residues of the LxxxIxxx[I/L] packed within the hydrophobic groove formed by H3 and H4 and interactions of the peptide with H1 and H11. H12 is required to be in the inactive state for the corepressor to bind (pdb code: 1KKQ) (Xu et al. 2002).

The core of the repression complex is formed by the interaction of HDAC3, TBL1/TBLR1, GPS2, and SMRT or NCoR. Using a variety of structural and functional approaches, Oberoi and colleagues elegantly demonstrated that this core-corepressor complex contains four TBL1 or TBLR1 molecules as a scaffold for two GPS2 molecules, two SMRT or NCoR molecules and two HDAC3 to form the huge 1-2 MDa complex. GPS2 and TBL1 interact with the N-terminal region of SMRT RD1 but also directly with each other. The NMR structure of the interacting regions of SMRT and GPS2 shows that the interaction is mediated by an antiparallel coiled-coil. One end of this coiled-coil contains the TBL1 interaction domains of both proteins so that both proteins interact with the same domain of TBL1 (Oberoi et al. 2011). At the same time, the DAD domain of SMRT specifically recruits HDAC3 catalytic domain through the interaction with both its C- and N-terminal (Guenther et al. 2001; Zhang

et al. 2002; Yang et al. 2002). The TBL1 tetramer, located in the center of the complex, mediates the assembly of two SMRT/GSP2 heterodimer bound to the HDACs in a ternary complex (**Figure 1.9**) (Oberoi et al. 2011).

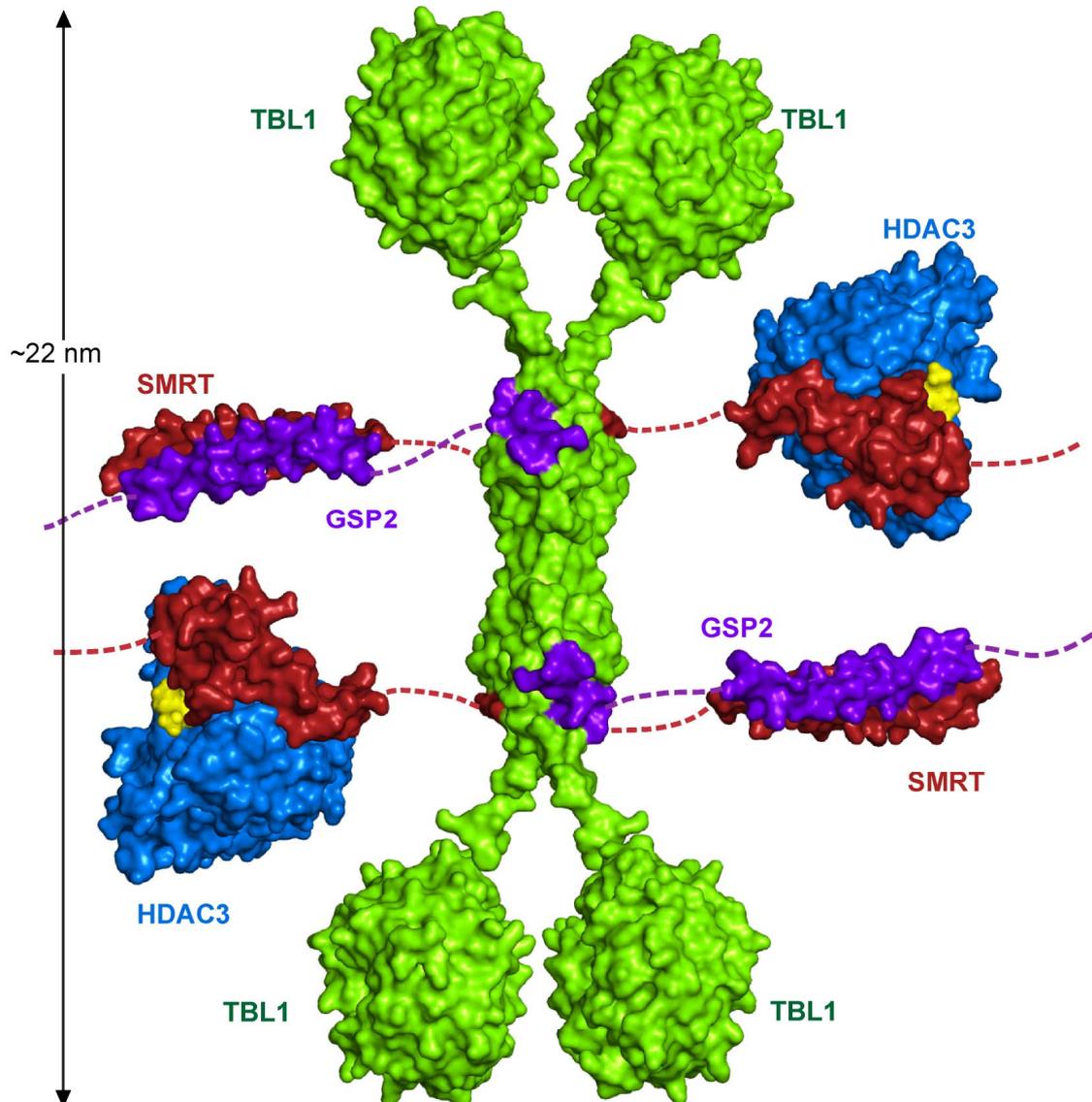


Figure 1.9: Model of the assembly of the core SMRT/NCoR repression complex. The TBL1 tetramer is shown in green in the center of the figure. The coiled-coil (pdb code 2L5G) and TBL1 interacting regions from SMRT and GSP2 (pdb code 2XTC) are shown in red and purple respectively. The DAD domain from SMRT (pdb code 1XC5) is also shown in red and an HDAC catalytic domain (pdb code 3HGQ) is shown in blue.

1.2.4 - Full-length NR complexes behaviour on the DNA

Several structures of full-length NRs and associated proteins have demonstrated the diversity of the organization of these complexes when bound to DNA (Chandra et al. 2008; Chandra et al. 2013; Lou et al. 2014). Despite the strong similarities in NR secondary structural organization, the fact that different NRs recognize different response element configurations and have different dimerization preferences suggest that each dimer might be organized in a different quaternary structure upon DNA binding.

In the first crystal structure of full-length NRs, the PPAR γ -RXR α heterodimeric complex in the active conformation bound to coactivator peptides, the LBD and DBD portions align perfectly with the previously solved structures of the isolated domains (Chandra et al. 2008; Lee & Mahdavi 1993; Nolte et al. 1998; Gampe, Montana, et al. 2000; Egea et al. 2000). The two receptor proteins form a very compact asymmetric complex, with PPAR γ upstream of RXR α . The LBD and the DBD are intimately coupled and coordinated forming a central zone that has been suggested to be a path of communication between the domains (Chandra et al. 2008). The PPAR γ CTE appeared to determine the polarity of the interaction between the two receptors, largely interacting with the 5' upstream sequence of the PPAR γ receptor element of the DNA (PPRE). In contrast, the RXR α CTE, which interacts with the PPAR γ DBD, lacks secondary structure and is more flexible. This flexibility may account for its promiscuity as a heterodimer partner for various receptors, which allows the RXR heterodimer complexes to interact with diverse direct repeats with multiple spacer sizes and half-site geometries.

Interestingly, in addition to the DBD-DBD and the LBD-LBD interfaces, there was a third heterodimerization interface between the PPAR γ LBD and the DBD CTE region of the RXR α (**Figure 1.10**). Mutational assays of key residues that form this region showed that PPAR γ LBD/ RXR α DBD interaction is DNA-dependent and can strongly influence DNA binding. This was already suggested by affinity assays that showed that intact nuclear receptors bind more strongly to the DNA than their isolated DBDs (Ozers et al. 1997).

A different structural organization was found for the HNF4 α homodimer bound to coactivator peptides and the same DR1 DNA target element (Chandra et al. 2013). This suggests that different quaternary structures might be adopted by different members of the NR superfamily. Despite the different overall organization, there are some common features between the two structures. First, the asymmetric HNF4 α homodimer structure has a convergence zone that lies at the center of the complex formed by the LBDs, the upstream-positioned DBD, and the hinge portion of the downstream-positioned receptor. This closely coupled region provides a path for signal communication from one end of the complex to the other. Another common feature is the creation of a third DNA-dependent interacting surface formed between the DBD of the upstream subunit and the hinge region of the downstream subunit (**Figure 1.10**). The physical connection between the LBDs and the upstream DBD suggests that the integration of diverse signals from the LBD and the DBD is essential to ensure high-affinity DNA binding (Chandra et al. 2013).

The crystal structure of the full-length RXR α -LXR β heterodimer on a DR4 element revealed yet a different structural organization (Lou et al. 2014). Again, the individual domains of both receptors in the complex adopt the predicted canonical organization. Both LBDs are in the active conformation, bound to coactivator peptides and making the usual heterodimer interactions. The structure of the DBDs closely correspond to that seen in previous structures of RXR DBD heterodimers such as RXR DBD-TR β DBD which also binds to DR4 elements (Rastinejad et al. 1995). This highlights the importance of the spacing nucleotides in the overall organization of the DBDs upon the DNA. As in the TR β DBD, the LXR β DBD contains a short α -helical C-terminal extension (CTE) that is not seen in RXR α DBD which contacts the DNA minor groove.

The main difference with the structures described in this section is the overall quaternary organization. The RXR α -LXR β complex overall structure exhibits an open asymmetric organization with both receptors crossed over and parallel to the DNA. The RXR α DBD occupies the 5' half-site with the LXR β DBD situated on the 3' half-site. Whereas, the RXR α LBD occupies the 3' position with the LXR β LBD situated in the 5' position (**Figure 1.10**). In contrast with the closed conformations of the previous structures, this open conformation of the RXR α -LXR β may be a consequence of the increased spacing between DR4 half-sites versus the DR1 elements. Superimposing the

RXR α -LXR β LBDs with the PPAR γ -RXR α LBDs reveals a significant twist of the DBDs. Since there are no obvious differences in the RXR α DBD or LBD structure, this structural alteration could be a consequence of adjustments in the RXR α hinge which exhibits high flexibility and can adapt to permit different NR partners to bind to their cognate DNA elements (Lou et al. 2014).

The quaternary structural organization seems to depend on hinge, CTE or LBD residues that are not conserved among NRs. These residues as well as the response element configuration are key determinants for the diverse quaternary organization that can be adopted by the NRs.

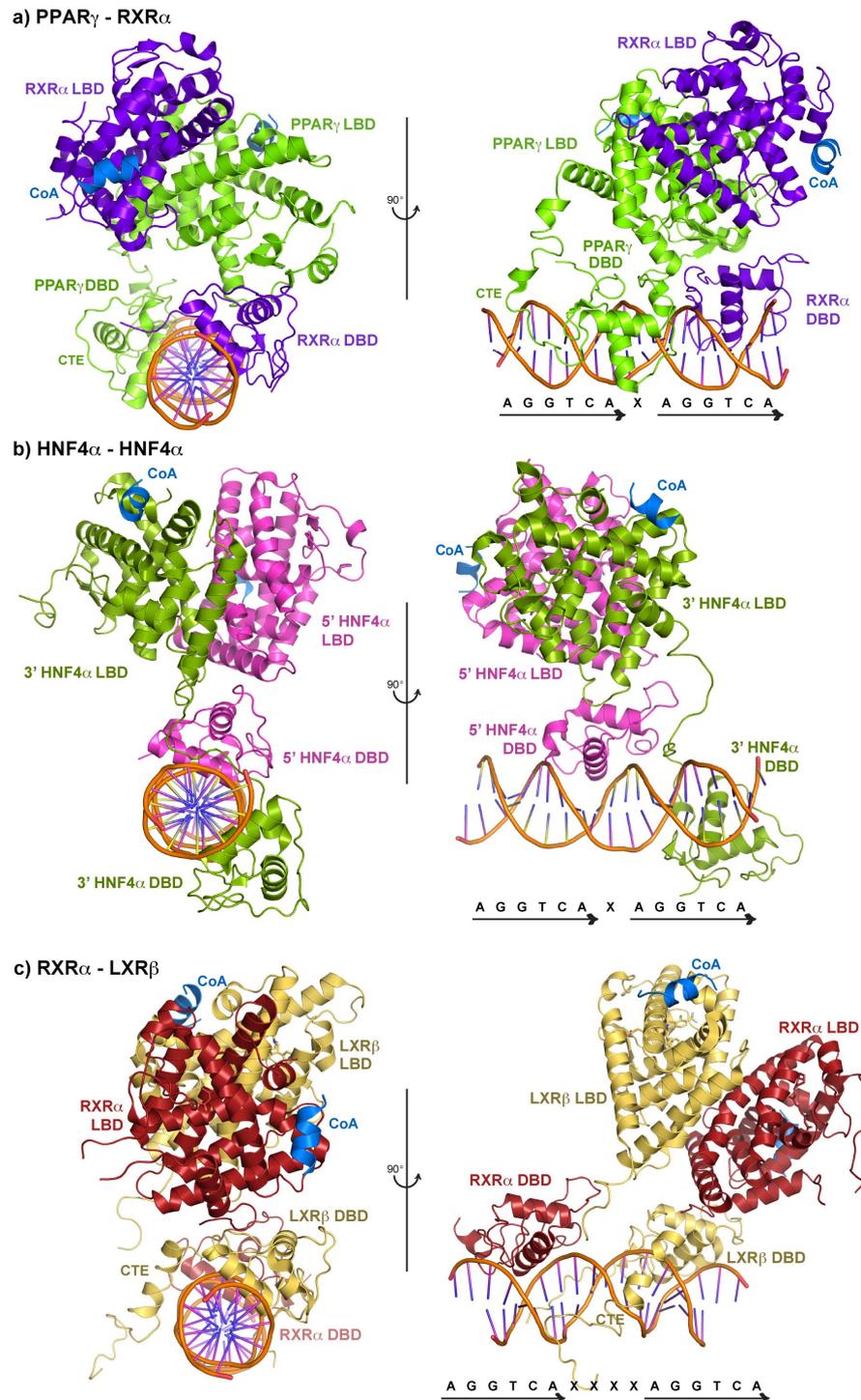


Figure 1.10: Overall structural organization of full-length NRs in the DNA. a) Liganded PPAR γ -RXR α heterodimer bound to coactivator peptides co-crystallized with DR1 target DNA. PPAR γ (green) is located upstream of RXR α (purple) (pdb 4DZY). b) Liganded HNF4 α homodimer bound to coactivator peptides and co-crystallized with the DR1 target DNA (pdb 4IQR). c) Liganded RXR α -LXR β structure bound to coactivator peptides and co-crystallized with the DR4 target DNA. RXR α (red) is located upstream of LXR β (yellow) (pdb 4NQA).

1.3 - Genetic disorders of nuclear receptors

Several human genetic disorders are caused by mutations in NR genes. These disorders are associated with pathogenic genetic variants of NRs which cause diverse phenotypes of different severity depending on the affected receptor and the mutation. Recently, NR gene mutations have been identified by whole exome sequencing, correlating human phenotypes with pathogenic genetic variants. To date, genetic defects in 20 of the 48 known human NRs have been related to human disorders (reviewed in Achermann et al. 2017)). Normally, the phenotype of the disorder is associated with the roles in endocrine pathways and physiological processes in which the affected receptor is involved.

Pathogenic mutations include nonsense, frame shift, and missense mutations. Nonsense and frame shift mutations often lead to more severe phenotypes, depending on the position of the mutation. In general, shorter proteins result in more severe phenotypes. Missense mutations often affect key amino acids in the DBD or LBD that impair the essential functions of DNA binding, ligand binding and/or ligand-binding transactivation.

Inheritance patterns also differ for the different NR genes. Mutations in NRs without repressive activity in the absence of ligand such as the VDR, MR and ER α generally exhibit recessive inheritance. Mutations located in the DBD lead to a complete loss of function in homozygotes and mutations in the LBD normally affect ligand binding or transactivation. However, in heterozygotes, the wild type allele function is not affected and is able to respond to the specific ligand, so that in most cases, patients with heterozygous changes tend to have a milder phenotype. The few cases of dominant inheritance reported are due to an inhibition or delay in nuclear translocation of the mutant receptor (Kino et al. 2001; Charmandari et al. 2004) or a ligand-independent nuclear localization of the inactive receptor (Inoue et al. 2000).

In contrast, NRs with transcriptional repressive activity in the absence of ligand such as TR and RAR, exhibit autosomal dominant inheritance. Heterozygous mutations normally disrupt the LBD function and inhibit wild-type receptor action in a dominant-

negative manner. The dominant negative activity involves enhanced corepressor recruitment and gene repression by the mutant receptor. Mutations causing dominant-negative activity are located in the LBD and affect ligand binding and/or ligand transactivation. Intact DNA binding and heterodimerization function of mutant receptors is essential for them to exert the dominant negative activity (Reviewed in (Gurnell et al. 2016)).

Naturally occurring human mutations in the PPAR γ LBD, associated with severe insulin resistance, diabetes mellitus and hypertension, also exhibit dominant negative activity. Functional studies showed that the two unrelated missense PPAR γ mutations (P467L, V290M) retain DNA binding but exhibit significant impairment of transcriptional activation and coactivator recruitment in response to ligands (Barroso et al. 1999). In the wild type PPAR γ LBD crystal structure, the residues Pro 467 and Val 290 are involved in H12 packing in the active state (**Figure 1.7**) (Nolte et al. 1998). The P467L and V290M mutations change the dynamics of H12, which favors interaction with corepressors (Xu et al. 2002; Kallenberger et al. 2003; Agostini et al. 2004).

1.4 - Mechanism of thyroid hormone action

Thyroid hormones (TH) are essential for normal development and growth, as well as for regulating metabolism in the adult of all vertebrates (Magnus-Levy 1895; Gudernatsch 1912; Cohen 1970; Morreale de Escobar et al. 1987; Nunez et al. 1991; Porterfield & Hendrich 1993; Oppenheimer & Schwartz 1997; Klein & Ojamaa 2001; Silva 2006; Fliers et al. 2010).

TH is produced by the thyroid gland in response to thyroid stimulating hormone (TSH) and is secreted by the anterior pituitary (Salter 1940; Stanley & Astwood 1949; Hoskins 1949). The thyroid gland consists of follicles, where TH is synthesized through iodination of tyrosine residues (Etkin & Gona 1974). The most common species secreted is 3,5,3',5'-tetraiodothyronine (thyroxine or T4) and to a lesser extent 3,5,3'-triiodothyronine (triiodothyronine or T3) (Kaplan 1984; Berne & Levy 1990). TH synthesis and secretion are tightly regulated by the hypothalamus-pituitary axis in order to maintain a constant cellular level of T3, especially in the brain, and a constant plasma

concentration of T4. The hypothalamus releases thyroid-releasing hormone (TRH) that subsequently stimulates the secretion of TSH by the pituitary gland. Circulating T4 and T3 downregulate the synthesis of TRH and TSH at hypothalamus and pituitary level, indirectly decreasing the production and release of TH by the thyroid gland (**Figure 1.11**) (Rondeel et al. 1988; Dahl et al. 1994).

THs, which are hydrophobic small molecules, were believed to cross membranes through passive diffusion. However, there are TH specific transporters in the membrane such as the MTC family (MTC8 and MTC10), the organic-anion-transporting polypeptide 1c1 (OATP 1c1) and the nonspecific L-type amino acid transporters 1 and 2 (LAT1, LAT2), that mediate the uptake of the circulating TH into peripheral tissues and entry into target cells (Christensen et al. 1954; Sorimachi & Robbins 1978; Rao et al. 1976; Krenning et al. 1978; Everts et al. 1996; Docter et al. 1997; Abe et al. 1998).

T4 is derived from thyroid gland secretion, whereas the majority of circulating T3 is produced by deiodination of T4 in peripheral tissues (**Figure 1.11**) (Chanoine et al. 1993). Although the circulating T4 concentration is 4-fold greater than circulating T3, the TR-binding affinity for T3 is 15-fold higher than its affinity for T4 ((Lin et al. 1990)). Thus T4 must be converted to T3 for mediation of genomic thyroid hormone action ((Bianco & Kim 2006)). Local production of the active form, T3, from the circulating prohormone T4 at the tissue level is carried out primarily by the action of deiodinases. Type 1 deiodinase (1 5'-deiodinase, D1) catalyzes removal of inner or outer ring iodine atoms to generate T3, rT3 or 3,3'-diiodothyronine (T2) from T4 and T3, respectively. D1 is present in peripheral tissues, especially in liver, kidney, thyroid tissue and CNS (Hennemann 1986; Kohrle et al. 1987; Leonard & Koehrle 1996; St Germain & Galton 1997). Type 2 deiodinase (5'-deiodinase, D2) is expressed in many T3-target tissues such as skeleton, brown adipose tissue, anterior pituitary, placenta and CNS (hypothalamus and pituitary) and catalyzes outer ring deiodination to generate T3 from T4 ((Friesema et al. 2006; Fonseca et al. 2013; Luongo et al. 2015)). There is a third type of deiodinase (D3) that irreversibly inactivates T3 or prevents T4 being activated by inner ring deiodination, converting T3 or T4 into inactive metabolites (T2 or rT3, respectively). D3 is mainly found in the CNS and the placenta and its physiological role is to prevent or limit access of TH to specific tissues at critical times

during development and in tissue repair ((Toyoda et al. 1997; Visser et al. 1998; Bianco & B. W. Kim 2006)).

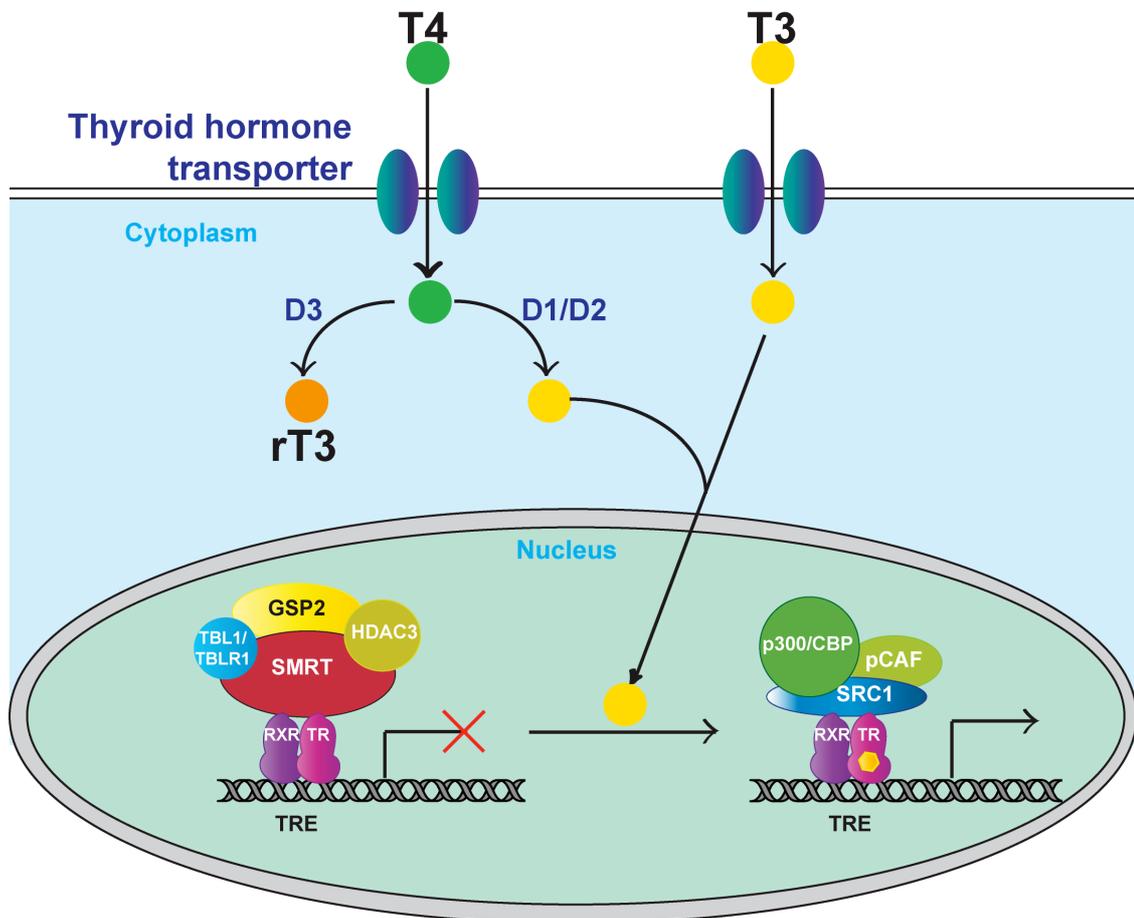


Figure 1.11: Schematic of thyroid hormone action. In specific tissues, such as the brain, transporters such as MCT8 transport T4 and T3 into the cell. Circulating T4 is converted locally in some tissues by D1/D2 enzymes to the active form, T3. D3 converts T3 and T4 to the inactive forms T2 and rT3, respectively. Unliganded TRs heterodimerize with RXR and bind to a TRE in order to repress gene expression by recruiting corepressors such as SMRT. T3 binding results in corepressor release and coactivator recruitment, which ultimately leads to transcription activation.

1.4.1 - Human Thyroid Hormone Receptor

TRs are T3-inducible transcription factors that bind to thyroid hormone-response elements (TREs) in the regulatory region of T3-target genes in order to mediate the biological effects of the TH.

Tata and Widnell were the first to demonstrate that thyroid hormone activates nuclear RNA synthesis by stimulating the RNA polymerase activity and subsequent cytoplasmic protein synthesis in isolated rat nuclei (Tata & Widnell 1966). Successive studies suggested the presence of molecules with high-affinity, low-capacity T3 binding sites in the nuclei of rat tissues and cultured GH cells (Oppenheimer et al. 1974; Samuels et al. 1974). The cloning of these molecules in 1986 allowed their biochemical characterization and the study of their transcriptional activity on T3 target genes (Sap et al. 1986; Weinberger et al. 1986).

Two major subtypes of TRs (TR α and TR β) are encoded by two paralog genes located on two different chromosomes (Weinberger et al. 1986; C. C. Thompson et al. 1987; Laudet et al. 1992). The *THRB* gene, located on chromosome 3, encodes two major T3-binding isoforms (β 1 and β 2) generated through differential promoter usage. The TR β 1 and β 2 isoforms share high sequence homology in the DBD and LBD but differ in length and sequence in the amino terminal A/B domain (**Figure 1.12**) (Lazar 1993). The TR α gene, located on chromosome 17, encodes two different proteins generated by alternate splicing (TR α 1 and TR α 2). The TR α 2 splice variant, which is longer than TR α 1 and has a different amino acid sequence in the C-terminal region, shows no T3-binding activity (Mitsuhashi et al. 1988).

Different TR isoforms are expressed in a tissue-dependent and developmentally regulated fashion (Cheng 2000). TR α 1 is ubiquitously expressed in all tissues to some extent, with particular abundance in the CNS, myocardium, gastrointestinal tract, bone, cartilage and skeletal muscle. TR α 2 is expressed in a variety of tissues such as brain and testis however its biological function is poorly understood, although it is thought to act as a weak dominant negative inhibitor of thyroid hormone action (Lazar 1993; Xiang-Jiao Yang et al. 1996). TR β 1 is widely expressed and is the principal isoform in liver and kidney, while TR β 2 expression is restricted to the hypothalamus, pituitary, inner ear, and retina (Lazar 1993).

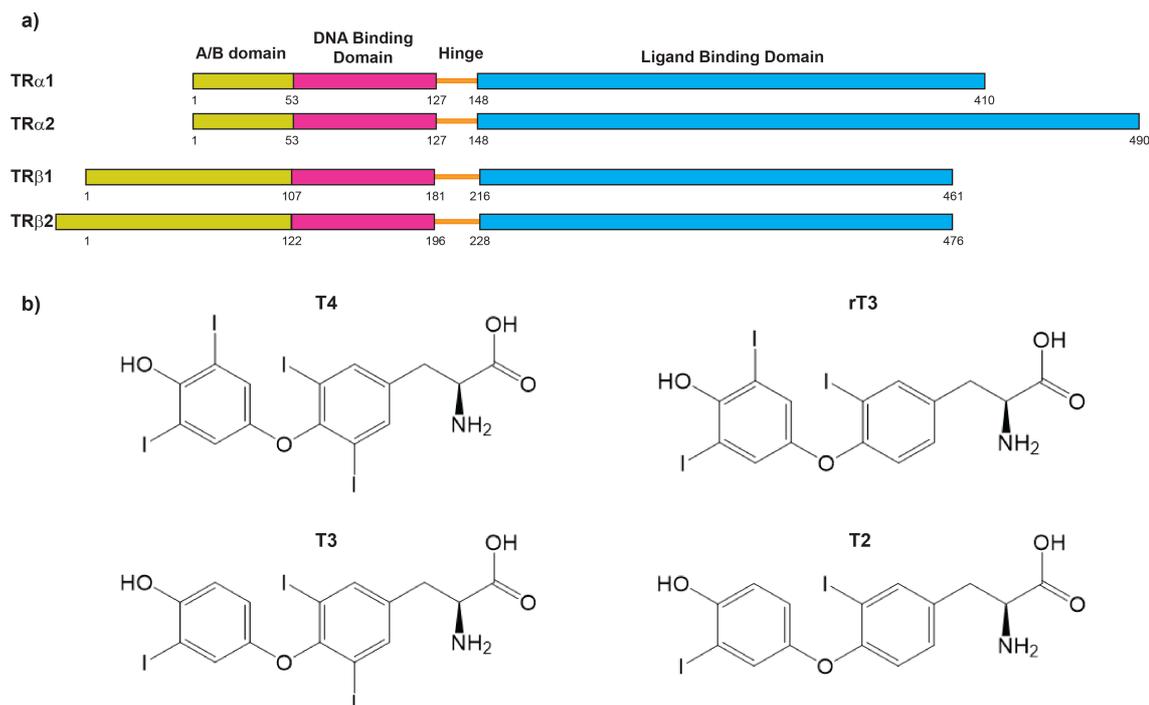


Figure 1.12: Scheme of TR isoforms and T3. a) Schematic view of the major isoforms product of the *THR* genes: TR α and TR β showing the modular organization and domains of TR α 1, TR α 2, TR β 1, and TR β 2. b) The thyroid hormone forms found in the human body: T4, T3 and rT3.

Although partial redundancy may exist for a limited number of functions of TR α and TR β proteins, in general these proteins mediate specific functions in a time- and tissue-specific manner. Both isoforms bind T3 with high affinity and specificity (dissociation constant TR α K_d = 0.058 nM and TR β K_d = 0.081 nM determined by an ¹²⁵I-T3 competitive binding assay) (Chiellini et al. 1998). Sequence analysis shows that there is a high degree of amino acid sequence homology in the DBD (90 % identity and 93 % similarity) and in the LBD (85 % identity and 93 % similarity) among the major forms of the receptors (Altschul et al. 1990). However, the A/B and hinge regions show little sequence similarity and TR β has a longer N-terminal domain (**Figure 1.12**) (Oppenheimer & Schwartz 1997).

1.4.2 - Ligand binding pocket of Thyroid Hormone Receptor

The structural fold of TR α and TR β LBDs consists of twelve α -helices (H1-H12) and two short β -strands (S1, S2). The LBP is deeply buried within the hydrophobic core of

the LBD and is almost completely filled by the ligand. The LBP is predominantly composed of hydrophobic amino acids but has two polar regions. The thyronine rings are stabilized by a series of non-polar interactions with the residues composing the cavity while hydrogen bonds stabilize the ends of the ligand. A single histidine located in H11 (His 381) forms a hydrogen bond with the phenolic hydroxyl of the outer thyronine ring. Other polar interactions are located at the other end of the hormone, formed by one asparagine located between H1 and H2 (Asn 179), three arginines located on H3 (Arg 228), H6 (Arg 262) and on the second β -strand (Arg 266), and by a serine (Ser 277) located on the third β -strand. The polar interactions of the ligand with this region of the protein form an intricate hydrogen bond network which involves surrounding water molecules (**Figure 1.13**) (Wagner et al. 1995).

The only difference between the LBP of TR α and TR β is that Ser 277 in TR α is substituted by an asparagine (Asn 331) in TR β . Ser 277 (TR α) and Asn 331 (TR β) form part of the binding cavity and make hydrogen bonds with the Arg 228 (TR α) and Arg 282 (TR β), respectively. Despite this difference, all the residues in the polar pocket of both receptors adjust to adopt the same conformations and make the same interactions with the ligand (**Figure 1.13**). Mutation studies, substituting an asparagine residue for a serine 277 (S277N) in TR α , and a reciprocal variant in TR β demonstrated that the single amino acid substitution is responsible for the different ligand affinities showed by the two receptors (Wagner et al. 2001).

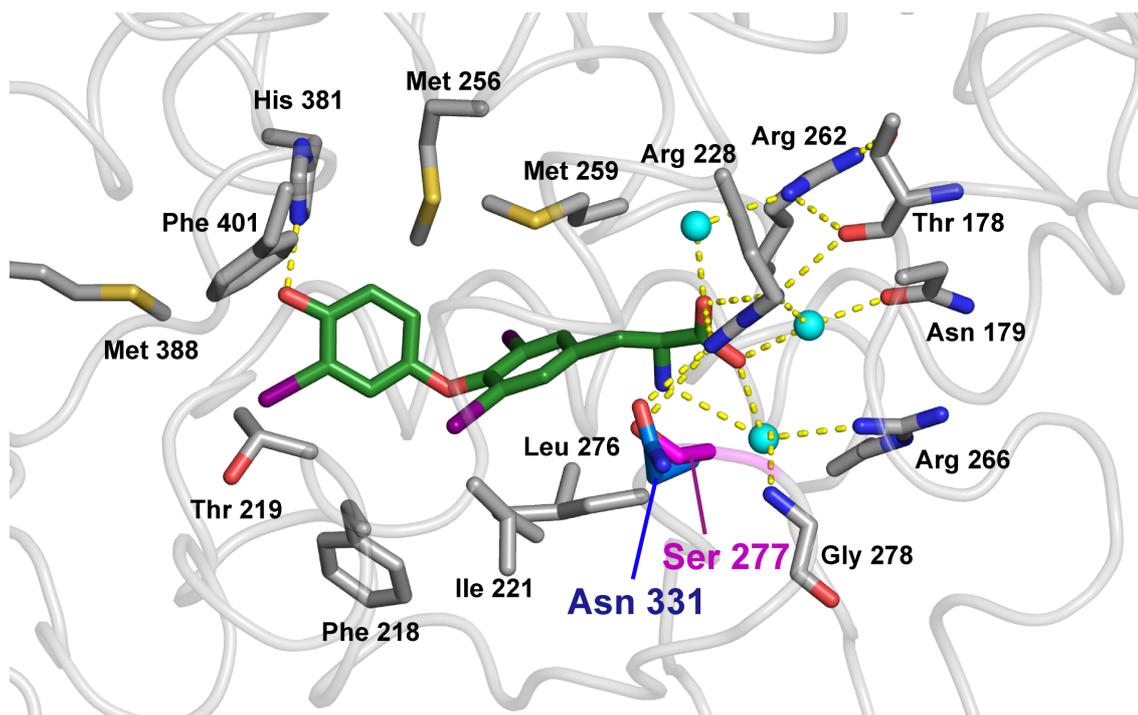


Figure 1.13: Superimposition of the LBP structure of TR α and TR β . LBP of TR α and TR β showing the residues involved labeled, the main difference in sequence highlighted (Asn 277 in TR β in blue and Ser 277 in TR α in magenta), and the intricate hydrogen bond network formed in the polar region of the LBP with the surrounding water molecules (pdb codes: 2H79 for TR α and 3GWS for TR β).

1.4.3 - Human Thyroid Receptor: mutations and disease

Several mutations have been identified in both TR α and TR β isoforms involved in a disease termed Resistance to Thyroid Hormone (RTH). Mutations in any of the isoforms have the same main consequence: an impaired TH response. However, due to the different tissue distribution of the two receptors, the symptoms of the disorder are different depending on which TR is affected.

Resistance to Thyroid Hormone beta (RTH β) is a dominant-inherited disorder caused by mutations in *THRB*. RTH β is biochemically recognizable by elevated circulating levels of T4 and T3 with non-suppressed pituitary TSH levels, reflecting the main role of TR β in the hypothalamus-pituitary axis regulation. Approximately 160 heterozygous TR β mutations have been identified. These mutations are mainly localized in three clusters within the TR β LBD, and show different levels of impairment depending upon the specific function of the residues affected by the mutation (reviewed in Refetoff & Dumitrescu 2007). The syndrome, which affects around 1 in 40,000 individuals, is also associated with variable resistance to hormone action in peripheral tissues where RTH β is expressed, especially in liver and kidney (Refetoff et al. 1993).

A homologous human disorder with defective TR α was expected given the significant amino acid sequence similarity exhibited between TR β and TR α . However, it was not until 2012 that the first patient carrying a mutation in TR α was identified. To date, 14 different mutations of human *THRA* have been identified in 28 different patients from 12 different families (**Figure 1.14**). These cases share hypothyroid features and RTH in TR α target tissues (**Figure 1.15**) (Bochukova et al. 2012; van Mullem et al. 2012; van Mullem et al. 2013; Moran et al. 2013; Moran et al. 2014; Espiard et al. 2015; Yuen et al. 2015; Tylki-Szymańska et al. 2015; Demir et al. 2016; Moran et al. 2016).

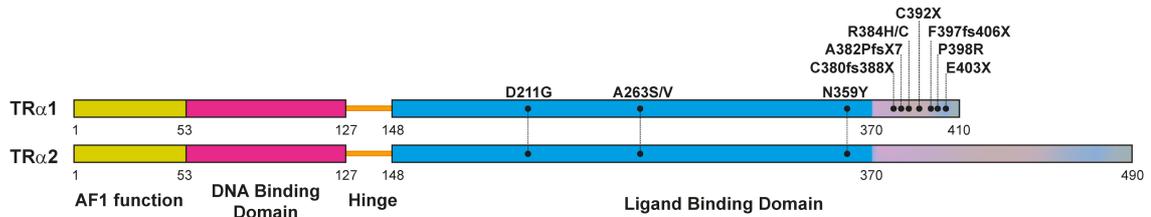


Figure 1.14: Schematic representation of the TR α 1 and TR α 2. Schematic view of the proteins derived from the *THRA1* locus, showing the location of the human mutations identified so far. The LBD is showed in blue, with the non-homologous areas in light blue. Three mutations D211G, A263V and N359Y affect both TR α 1 and TR α 2 transcripts.

Both TR α and TR β resistance disorders are dominantly inherited which means that the mutant receptor inhibits the wild type receptor in a dominant negative manner (Reviewed in Gurnell et al. 2016).

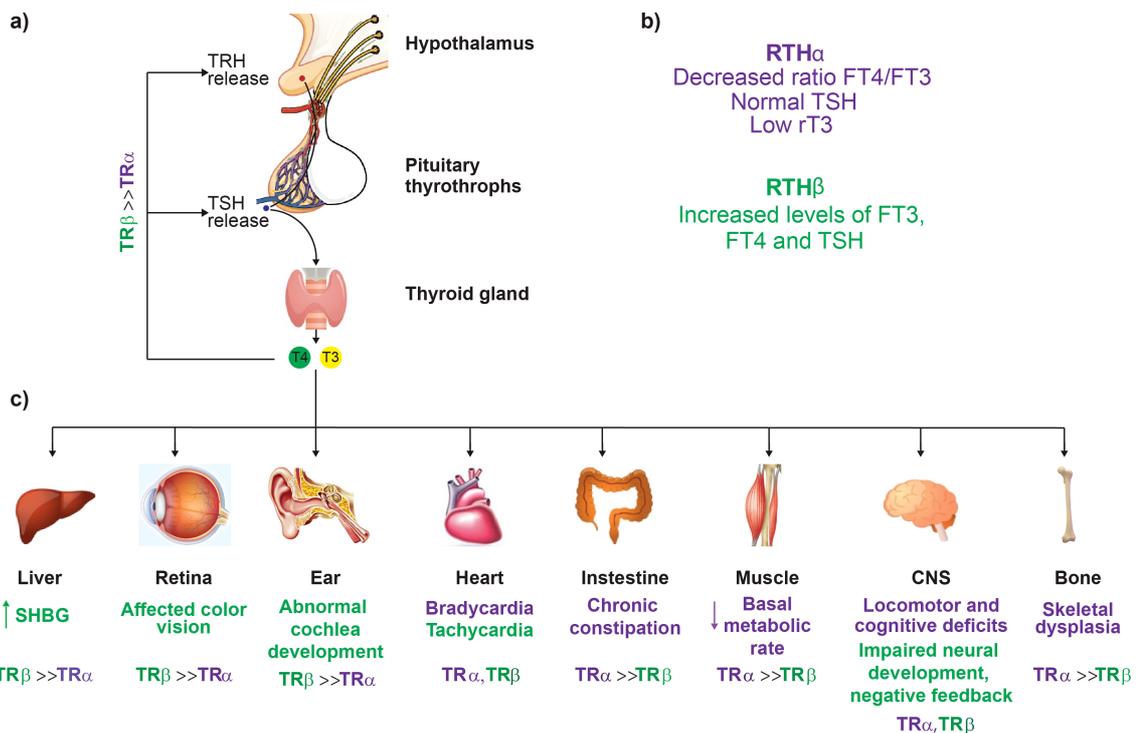


Figure 1.15: TR α and TR β tissue distribution. The diagram shows the tissues where the different isoforms are mainly expressed and the different symptoms of RTH α and β due to the tissue specificity of expression. a) Hypothalamus-pituitary axis regulating the TH secretion through the release of TRH and TSH. The feedback is mainly carried out by the TR β . b) Biochemical parameters shown by patients affected by RTH α and RTH β . c) Symptoms of the RTH α (purple) and RTH β (green) distributed among the tissues affected.

1.5 - Aims and objectives

This thesis is focused on the biophysical and biochemical characterization of the proteins resultant from the first three mutations identified in the *THRA1* gene. These three mutations have been recently identified in four patients with symptoms of tissue-specific hypothyroidism (growth and developmental retardation, skeletal dysplasia, and constipation) associated with biochemical irregularities that included low to low-normal T4 and high to high-normal T3 concentrations, a subnormal T4/T3 ratio, low levels of rT3, and normal levels of thyroid stimulating hormone (TSH). Whole exome sequencing indicated that these patients were affected by Resistance to Thyroid Hormone alpha (RTH α) (Bochukova et al. 2012; van Mullem et al. 2012; Moran et al. 2013).

The RTH α disorder is mediated by heterozygous, loss-of-function, mutations involving *THRA1*. The mutations include a non-sense mutation (E403X) (Bochukova et al. 2012), a single nucleotide insertion that results in a frame-shift from codon 397 (F397fs406X) (van Mullem et al. 2012) and a single nucleotide insertion that also shifts the reading frame at codon 382 (A382PfsX7) (Moran et al. 2013) (**Figure 1.16**).

All these mutations result in prematurely truncated TR α proteins, without the C-terminal region (H12). Collaborators at the University of Cambridge showed that the TR α mutant proteins have minimal T3-dependent transcriptional activation. However, these TR α mutants constitutively repress the transcription of T3-target genes. Hence, when coexpressed, these TR α mutants inhibit transcriptional activation of their wild type receptor counterparts in a dominant negative manner (**Figure 1.16**) (Bochukova et al. 2012; van Mullem et al. 2012; Moran et al. 2013).

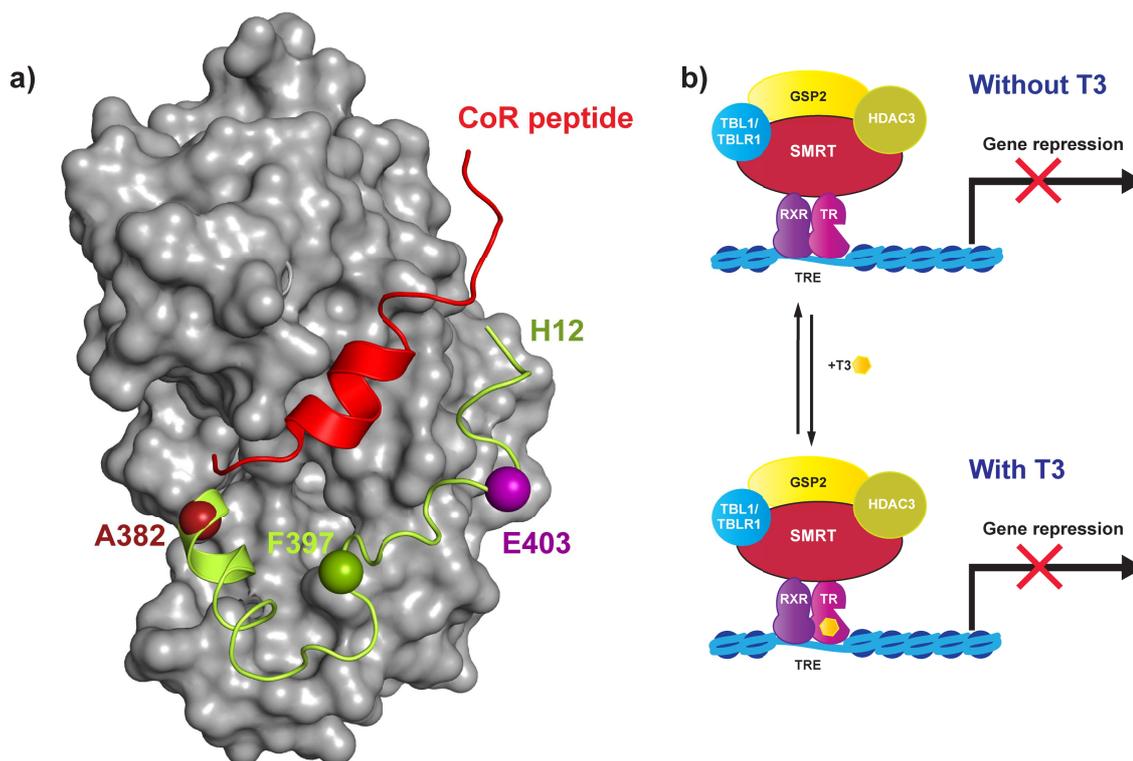


Figure 1.16: Position of the TR α LBD mutations and the mechanism of action. a) Crystallographic model corresponding to the superimposition of TR α LBD (pdb code 2H79) and PPAR α LBD (pdb code 1KKQ). The model shows the TR α LBD in the inactive state with H12 in the disorder conformation and the position of the different mutations. b) The mutant TR α can recruit the corepressor complex and inhibit basal gene transcription but are unable to respond to T3, resulting in persistent inhibition of gene transcription, even in the presence of hormone.

The aim of this thesis is to investigate the RTH α at the molecular level in order to better understand the behavior of the mutant TR α in key aspect such as interaction with coregulator proteins, the mechanisms underlying repression and response to ligands by structural and biophysical approaches.

Chapter 3 describes the purification and biophysical characterization of the WT and the mutant LBDs. Fluorescence anisotropy and circular dichroism approaches are used to investigate the interaction of the receptors with coregulator proteins in the presence and in the absence of T3. The results demonstrate that the mutant LBDs constitutively interact with corepressors and are not able to recruit coactivators even in the presence of T3. However, the mutant LBDs are able to bind T3, so the absence of T3 response is a consequence of the failure of the mutant receptors to recruit coactivators. The most likely reason for this impaired T3 response is the lack of H12 exhibited by the mutant

LBDs. The active conformation of H12 adopted by the WT LBD in response to T3 is absolutely essential to bind coactivators and ultimately, activate the T3 target genes transcription.

The ability of the mutant LBDs to accommodate ligands in their ligand binding pocket provided an opportunity to design and test T3 analogues as potential agents to disrupt the interaction between corepressors and the mutant LBDs. Disruption of this interaction would alleviate the constitutive repression and the dominant negative activity carried out by the mutant LBDs. The WT receptor, also present in the cells, would be able to respond to T3 by activating transcription. Two types of T3 analogues were designed with the purpose to displace corepressor peptide and/or recruit coactivator peptide.

An additional alternative approach is investigated as a potential therapy to treat the disease: stapled peptides. Stapled peptides are synthetic peptides with a chemical modification called the staple that keeps the peptide in the helical bioactive conformation even in solution. Stapled peptides were designed based on the native corepressor peptide and synthesized to compete with native corepressor for binding to the mutant LBDs.

These alternative strategies are described in Chapter 4 and are designed to avoid the ineffectiveness and adverse effects of T4 treatment. *In vitro* biophysical studies are performed to investigate the ability of these agents to prevent corepressor and mutant LBDs interaction and, consequently to relieve the dominant negative activity of the mutant LBDs.

Chapter 5 describes the numerous crystallization attempts performed in order to obtain the crystal structure of the mutant LBDs. Different strategies were tried to determine the crystallographic structure of the interface between mutant LBDs and corepressor. The strategy consists of trying to increase the stability of the mutant proteins by optimizing the boundaries, adding different ligands (T3 and T3 analogues), and different versions of the corepressor peptide (native SMRT and stapled SMRT). The determination of the structure of a construct based on the mutant LBDs confirms that mutant LBDs are able to bind T3 and reveals the position of the hormone and the size and shape of the ligand-

binding cavity. This knowledge will aid the design of T3 analogues with improved properties to treat the disease.

The disorder derived from mutations in the *THR1* gene, RTH α , is difficult to diagnose due to the lack of clear biochemical markers exhibited by the patients affected. In order to identify and anticipate potential pathogenic variants of the *THR1* gene, an exome database was examined. 15 potential polymorphisms were selected as possibly disease causing and were analyzed by functional assays in mammalian cells. Chapter 6 describes the identification of three probably mutations confirmed from these 15 potentially pathogenic variants.

Chapter 2 - Materials and Methods

2.1 - Materials

2.1.1 - Plasmid constructs

Vectors for expressing recombinant proteins in *Escherichia coli* or mammalian cells (JEG-3) were provided by the PROTEX cloning service at the University of Leicester. The plasmids used for expression in *E. coli* are based on the pGEX vector and are all regulated a lac operon inducible using IPTG (pLEICS14). For mammalian system assays, mammalian expression vectors based on pcDNA3 and pCMV were used (pLEICS12). Amino-terminal affinity tags are spaced from the inserts by a TEV protease recognition sequence and in addition to a selective antibiotic resistance gene, the vectors allow selecting for positive clones thanks to the presence of a *sacB* gene in the cloning region, which is replaced by the gene of interest. *SacB* gene expression is lethal for cells growing in sucrose agar plates, which allows the positive selection of colonies containing correct inserts. The vectors are also modified to allow ligase-free ligation using a BD In-Fusion kit. All vectors used in the project are detailed in **Table 2.1** and **Figure 2.1**.

Vector name	Backbone	AR	Tag	Expression Host
pLEICS12	pcDNA3	Amp	His ₁₀ / 3 x Flag	Mammalian
pLEICS14	pGEX-4T-1	Amp	GST	<i>E. coli</i>

Table 2.1: An overview of the plasmids designed by the PROTEX cloning facility indicating the backbone vector from which they were derived, antibiotic resistance, affinity tag and expression host.

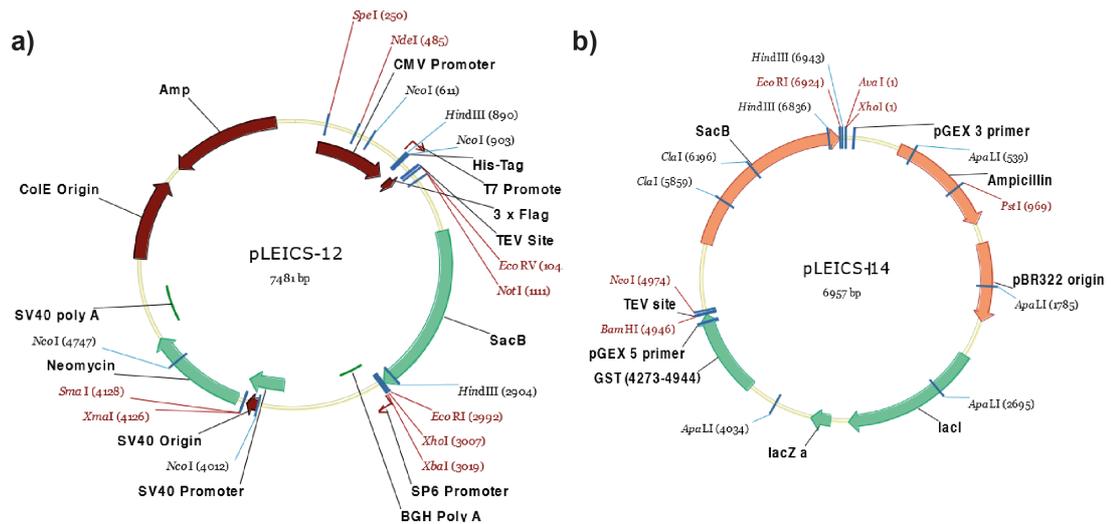


Figure 2.1: Plasmid maps. Schematic representation of the plasmids pLEICS12 (a) and pLEICS14 (b) indicating size, restriction sites, genes and promoters.

2.1.2 - Primers

All primers were purchased from Eurofins MWG Operon and designed using the MacVector and the QuickChange Primer Design Program from Agilent Technologies. Some of them were created to clone the gene of interest in the different vectors, while the others were designed to introduce specific point mutations or nonsense mutations. A list of the oligonucleotides used can be found in Appendix 1.

2.1.3 - Bacterial cell lines

Competent DH5 α and Rosetta 2 (DE3) *E. coli* cells were used, both lines purchased from Invitrogen. Rosetta 2 host strains are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*. These strains supply tRNAs for 7 rare codons (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) on a compatible chloramphenicol-resistant plasmid (**Figure 2.2**).

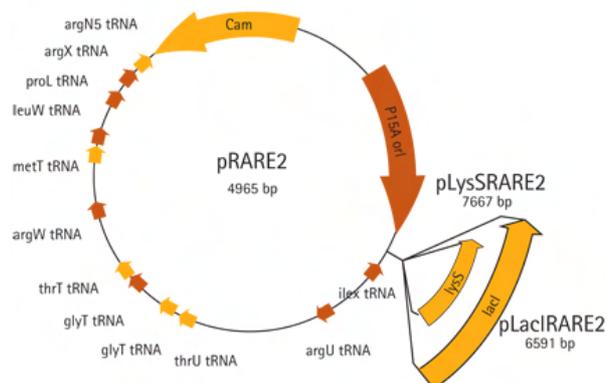


Figure 2.2: Rosetta 2 *E. coli* cells plasmid. Rosetta 2 strains express 7 rare codons on this chloramphenicol-resistant plasmid, pRARE2.

2.1.4 - Mammalian cells

The human choriocarcinoma-derived JEG-3 cells were provided kindly by the University of Cambridge and maintained in house. JEG-3 cells were cultured in Opti-MEM (Gibco™ Opti-MEM™ reduced serum media) enriched with 10 % FCS (Fetal Calf Serum) and 1 % PSF (100x Penicillin 10,000 Units/ml-Streptomycin 10,000 Units/ml-Amphotericin B 250 Units/ml).

2.1.5 - Standard Chemicals and Reagents

All chemicals and reagents were purchased from Sigma-Aldrich, Melford or Fisher Scientific unless otherwise specified. DNA restriction enzymes were purchased from NEB, KOD Hot Start polymerase was supplied by Sigma Aldrich and Tobacco ETCH Virus Protease (TEV) was expressed and purified in house by Ms. Jacquie Greenwood. Complete EDTA Free Protease Inhibitor Cocktail Tablet was acquired from Roche Diagnostics.

Pre-cast Novex® NuPAGE® 4-12 % gradient bis-tris SDS PAGE gels and NuPAGE® MES-SDS running buffer were purchased from Life Technologies (UK).

3,3',5-Triiodo-L-thyronine (T3) was purchased from Sigma-Aldrich and was dissolved in 100 % ethanol (stock solution 10 mM). T3 analogues were synthesized and kindly donated for analysis by Nick Tomkinson, University of Strathclyde.

Tissue culture reagents were purchase from Sigma-Aldrich, except from Opti-MEM reduced serum media that was purchased from Gibco.

Protein Crystallography sparse matrix screens used in this project such as NR-LBD, JCSG-plus, ProPlex™, Stura and MacroSol, MiDAS™, Morpheus®, NR LBD™ and PACT *premier* were obtained from Molecular Dimensions and distributed by hand to deep well blocks and stored at 4 °C. Chemicals to make optimized screen were also purchased from Molecular Dimensions or as specified.

Chemicals needed for NMR measurements were kindly provided by the department of Chemistry at the University of Leicester.

2.2 - Generating protein constructs and mutagenesis: cloning, DNA sequencing and DNA purification methods

2.2.1 - Primer design

Primers were designed as described in paragraph 2.1.2. Genes of interest were amplified from plasmid templates previously designed and purified by Dr. Maura Agostini (University of Cambridge) and Dr. Louise Fairall. Primers used for cloning contained 12-21 nucleotides complementary to the required insert and a further 15-18 nucleotides complementary to the homology tag of the vector. Primers were designed to have a melting temperature between 68 °C and 72 °C.

Primers for mutagenesis were designed to be 25-45 nucleotides in length containing a central mismatch corresponding to the required mutation, with a melting temperature of ≥ 78 °C. Both primers containing the desired mutation anneal to the same sequence on opposite strands of the plasmid. The two additional nucleotides were needed to clone the new mutant constructs in the pLEICS12 (mammalian expression vector). These

primers were designed to contain 12-21 nucleotides complementary to the insert and further 15-18 nucleotides complementary to the homology tag of the vector. The melting temperature was between 68 and 72 °C.

2.2.2 - Cloning procedure

All cloning was conducted by the PROTEX service at the University of Leicester using a BD in-fusion system. PCR products were cloned into the required bacterial expression or mammalian expression vectors by Dr. Xiaowen Yang and Ms. Dipti Vashi. The target gene with homology tags is added to linearized plasmid and incubated with a BD in-fusion enzyme to fuse PCR product to the vector. The resulting plasmid is transformed into DH5 α cells and plated onto sucrose containing agar plates. Inoculating onto sucrose plates enables selection for clones in which the target gene has replaced the *sacB* gene. Positive colonies were verified by colony PCR.

2.2.3 - Site Directed Mutagenesis

Site directed mutagenesis was performed using overhanging PCR. Mutagenesis primers were designed to be 25-30 nucleotides with a minimum of 9 bases preceding and following the site of mutation, as described in paragraph 2.2.1. A PCR reaction is performed with the 5' ORF primer, the reverse mutagenesis primer and a second PCR reaction with the forward mutagenesis and the reverse ORF primers. PCR products were purified and then used as template in a final PCR reaction using forward and reverse ORF primers. This resulted in two PCR products annealing and the product from this PCR reaction was used as the insert for cloning (**Figure 2.3**).

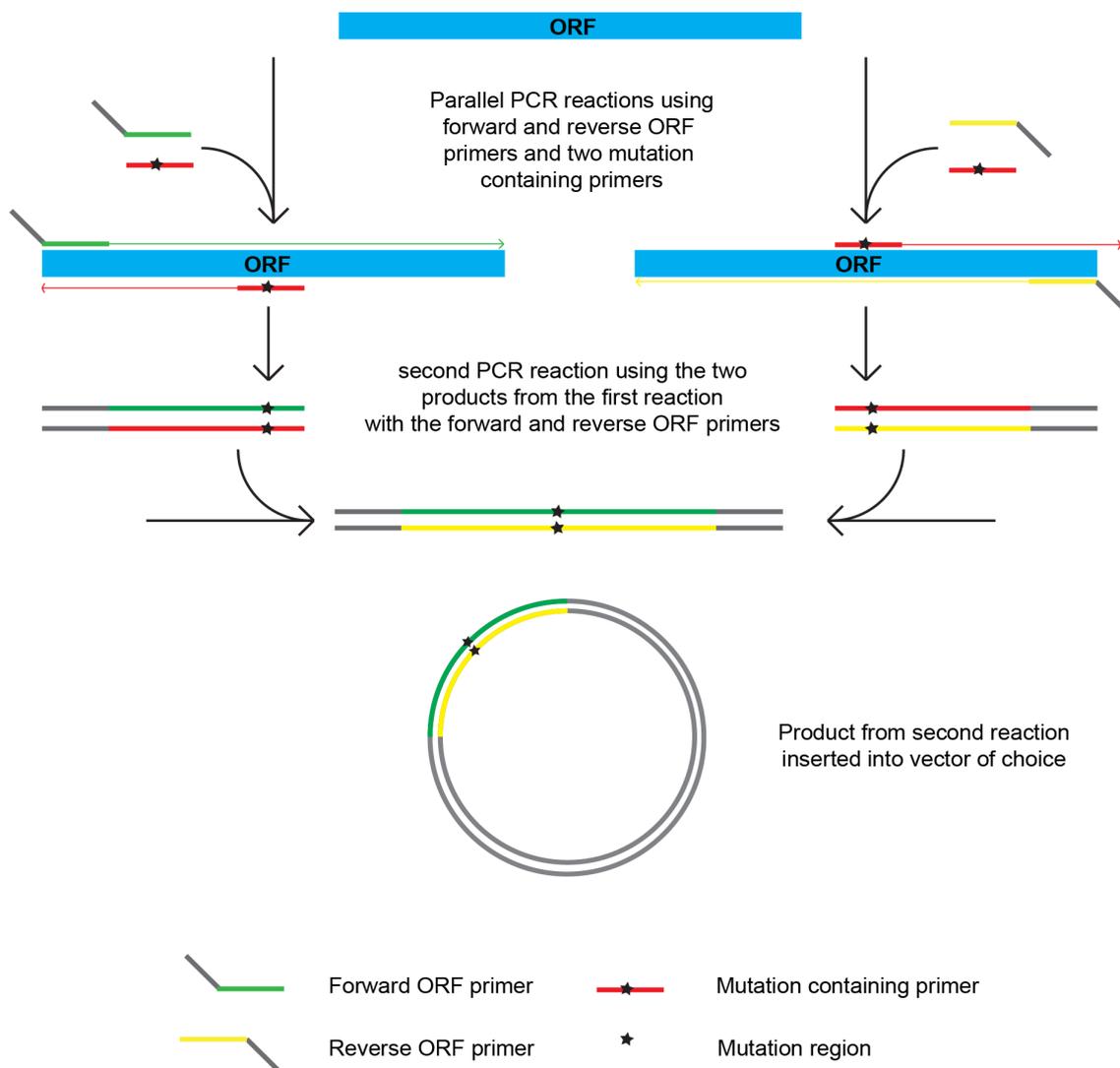


Figure 2.3: Overview of overhanging PCR used in the production of mutant genes. Two PCR reactions were performed with one primer to the 5' or 3' end of the gene and a mismatch primer to the location where the mutation is required. The two PCR products are used as a template with the 5' and 3' primers to produce a full-length product with the expected mutation. The product of the final PCR was gel purified and inserted into vectors by PROTEX.

2.2.4 - Small-scale plasmid DNA purification

Small amounts of plasmid DNA were purified using commercially available plasmid DNA purification kits based on a modified alkaline lysis method followed by an ion-exchange purification step using resin to bind DNA under appropriate low-salt and pH conditions.

As specified by the Quiagen Miniprep protocol, 5-8 ml cultures of transformed *E. coli* were grown and the DNA was purified using the Quiagen Miniprep kit following the manufacturer's instructions.

DNA from larger cultures (up to 100 ml) was purified with the Quiagen Midiprep plasmid purification kit following the manufacturer's instructions.

2.2.5 - Large-scale plasmid DNA purification

Large-scale purification was carried out with a LiCl and PEG based method optimized by Dr. Louise Fairall.

750 ml of 2xYT media (15 g/L BactoTryptone, 10 g/L Bacto Yeast Extract, and 5 g/L NaCl) containing 200 µg/ml of ampicillin were inoculated with a single colony of transformed *E. coli* DH5α cells containing the plasmid of interest and incubated for 20 hours in an orbital shaker incubator at 37 °C and 160 rpm. Cells were harvested by centrifugation at 3503 x g (or 4000 rpm on a Sorvall Evolution RC with a F8S-6x1000y rotor programmed with the SLC 6000 setting) for 10 minutes, re-suspended in 15 ml of suspension buffer (25 mM Tris pH 8.0, 10 mM EDTA) and lysed by adding 50 mg of lysozyme previously dissolved in 5 ml of alkaline buffer (0.2 M NaOH, 1 % SDS). Genomic DNA and SDS were precipitated by adding 22.5 ml of neutralization buffer (3 M phosphate acetate pH 4.8) and, after incubating 5 minutes at room temperature, centrifugation at 3220 x g (or 4000 rpm on a Eppendorf 5810 R centrifuge with a Eppendorf A-4-81 swing bucket rotor) for 10 minutes at 4 °C. Then, the supernatant was filtered through miracloth. Nucleic acids were precipitated out of solution by adding 0.6 volumes of isopropanol, incubate 5 minutes, then pelleted by centrifugation at 3220 x g for 10 minutes at 4 °C and re-dissolved into 15 ml of suspension buffer. RNA and protein impurities were taken out of solution by adding 1 volume of 5 M LiCl kept at -20 °C, incubating on ice for 5 minutes and precipitated by centrifugation at 3220 x g for 10 minutes at 4 °C. Nucleic acids dissolved in the supernatant were pulled out of the LiCl solution by adding 0.6 volumes of isopropanol, incubating 5 minutes and centrifugation at 3220 x g for 10 minutes at 4 °C. The pellet was re-dissolved in 10 ml of suspension buffer 2 (20 mM Tris-HCl pH 8.0, 1 mM EDTA). Remaining RNA

contaminants were digested by adding 50 µl of 10 mg/ml heat treated RNase A to the sample tube followed by an incubation of 15 minutes at room temperature. Small RNA fragments resulting from the RNase A digestion were discarded by the addition of 0.25 volumes of PEG solution followed by a 30 minutes incubation on ice. PEG only precipitates large nucleic acid molecules leaving the small RNA contaminants in solution. The precipitated plasmid was then collected at the bottom of the centrifuge tube by centrifuging at 3220 x g for 10 minutes at 4 °C and the pellet was re-dissolved into 10 ml of suspension buffer 2. Remaining PEG impurities were extracted by adding 2 ml of chloroform to the sample, vortexing briefly and centrifuging for 1 minute at 4 °C. The aqueous layer was retained and the purified plasmid was precipitated out of solution by adding 0.1 volumes of 5 M NaCl and 3 volumes of absolute ethanol, followed by centrifugation step at 3220 x g for 10 minutes at 4 °C. The pellet was washed in 70 % ethanol and let to dry upside down overnight. The plasmid was finally dissolved into an appropriate volume of filter sterilized Milli Q H₂O and stored at -20 °C.

2.2.6 - Determination of DNA concentration

DNA concentration was determined by nanodrop using an IMPLEN nanophotometer in accordance with the manufacturer's instructions.

2.2.7 - Plasmid sequencing

Protein expression vectors were sequenced by the Protein and Nucleic Acid Chemistry Lab (PNAACL) at the University of Leicester. Sequencing results were analyzed using 4 peaks, EnzymeX, MacVector and Standard Nucleotide BLAST®.

2.3 - Mass spectrometry analysis

Protein bands from Coomassie stained SDS PAGE gels were excised and submitted to the PNAACL at the University of Leicester where they were analyzed either by MALDI-ToF or LC-MS/MS.

2.4 - Protein expression and purification

2.4.1 - Cell transformation and growth

Rosetta 2 cells were transformed with the appropriate bacterial expression vector carrying the different gene constructs. Transformation was carried out by incubating the competent cells on ice for 30 minutes with 1 µg of plasmid DNA. The transformed cells were then rescued in 400 µl of LB media (10 g/L bactotryptone, 10 g/L NaCl, 5 g/L bacto yeast extract) for about one hour at 37 °C, plated out onto 2YT agar containing 30 µg/ml chloramphenicol and 30 µg/ml ampicillin and incubated at 37 °C overnight.

Starter cultures were prepared by inoculating a single colony resulting from the transformation plates into 10 ml of 2TY containing 30 µg/ml chloramphenicol and 30 µg/ml ampicillin. Cultures were incubated at 37 °C in an orbital shaker for 6 hours and then transferred into 750 ml of 2TY with 30 µg/ml of chloramphenicol and 30 µg/ml of ampicillin. These flasks were incubated at 37 °C in a shaking incubator. Once the optical density of an absorbance of 600 nm had reached 0.1, protein expression was induced adding 40 µM isopropyl-1-thio-β-D-galactopyranoside (IPTG) and the temperature was decreased to 20 °C, the optimal temperature for protein expression. Finally, cells were harvested by centrifugation at 3503 x g for 10 minutes at 4 °C after 18 hours of incubation. Cell pellets were sometimes frozen at -20 °C for later use.

2.4.2 - GST-tagged protein purification

Frozen cell pellets were defrosted and re-suspended in lysis buffer containing 1xPBS, 0.5 mM DTT, 1 % triton X-100 and a CompleteTM EDTA-free protease-inhibitor cocktail tablet. Then, cells were lysed by sonication for 3 minutes with 30-second on/off intervals using a medium probe 13-mm sonotrode in a Soniprep 150 Ultrasonic Disintegrator (Sanyo Gallen Kamp PLC). Cell debris was eliminated by centrifugation for 20 minutes at 108,669 x g (or 30,000 rpm on Avanti J-30I centrifuge with a JA-30.50 Ti fixed angle rotor), and the supernatants were isolated and then mixed with 1/10 volume of GST-resin slurry (pre-equilibrated in wash buffer) per liter of culture and

incubated for 30 minutes at 4 °C in continuous agitation. The resin was then washed three times with wash buffer (1xPBS, 0.5 mM DTT, 1 % Triton X-100). Bound proteins were eluted using the cleavage buffer (20 mM Tris-HCl pH 8, 100 mM NaCl, 0.5 mM DTT) and finally, the GST tag was removed by incubation with TEV (100:1 molar ratio) overnight at 4 °C.

2.4.3 - Ionic exchange chromatography

Eluted protein samples were filtered through a 0.22 µm filter and loaded onto 5-ml HiTrap Q HP Ion Exchange column (HiTrap Q HP IEX), mounted on an ÄKTA purifier, previously equilibrated in low salt buffer (20 mM Tris-HCl pH 7.4, 50 mM NaCl, 1 mM DTT). Proteins were separated by elution with high salt buffer (20 mM Tris-HCl pH 7.4, 500 mM NaCl, 1 mM DTT) at a flow of 1.5 ml/min. 1 ml fractions were collected into a fraction collector.

2.4.4 - Size exclusion chromatography

Partially purified diluted samples from the ionic exchange chromatography were transferring into a 15 ml Amicon® ultracentrifugal filter (Millipore) with a 10.0 kDa molecular weight cut-off and centrifuged until the sample was concentrated down to a volume of approximately 500 µl. The concentrated protein complex was filtered through a 0.22 µm filter and then loaded onto a size exclusion chromatography (Superdex 75 10/300 GL, GE Healthcare Bio-Science) mounted on an ÄKTA purifier, already equilibrated with Gel Filtration buffer (30 mM Tris-HCl pH 8.0, 50 mM NaCl, 5 % glycerol v/v, 1 mM DTT, and 0.5 mM EDTA). An isocratic run was carried out at flow rate of 0.5 ml/min in order to elute the different proteins and collect them in 500 µl fractions of the sample.

2.4.5 - Protein concentration and quantification

Relevant fractions containing the protein of interest were merged together and transferred into a 15 ml Amicon® ultracentrifugal filter with a 10 kDa molecular weight

cut-off and centrifuged at 3220 x g until the sample was concentrated down to a volume of 500 μ l. The samples were desalted and the buffer replaced following multiple rounds of concentration and dilution using the Amicon® ultracentrifugal filter. Samples used for fluorescence anisotropy (FA) and circular dichroism (CD) experiments were kept in a buffer consisting of 1x PBS and 1 mM DTT and samples for crystallization trials were maintained in 20 mM Tris pH 8.0, 50 mM NaCl and 1 mM of DTT.

Protein concentration was determined using the Bio-Rad Protein Assay following the manufacturer's guidelines: 2 μ l of pure protein was added to 1 ml mixture containing 20 % Bio-Rad reagent diluted in ultrapure 18 M Ω water. Absorbance was measured at 595 nm, against a diluted reagent blank. The protein concentration in mg/ml was then calculated using a factor which had been previously determined from a BSA standard curve.

2.4.6 - Protein analysis by SDS-PAGE

Proteins from different fractions were analyzed by SDS-PAGE. Electrophoresis was performed using NuPAGE pre-cast 4-12 % gels (Invitrogen). Samples were prepared by mixing with 2x SDS sample buffer (20% glycerol, 70 mM Tris pH 6.8, 0.54 mg/ml bromophenol blue, 2.0 % SDS, 200 mM DTT). Electrophoresis was performed using the manufacturer recommended running buffer at 125 V for 35 minutes.

Proteins were visualized by staining with Instant Blue (Expedeon) for 1 hour before de-staining with water. To examine protein size, SeeBlue Plus2 protein marker from Invitrogen was included as an electrophoretic sample.

2.5 - Peptide synthesis

The peptides used in this project were designed for biochemical studies of protein interaction and for crystal trials. Two different types of peptides were used, native peptides and stapled peptides. Stapled α -helical peptides are synthetic peptides locked into their bioactive α -helical fold through the site-specific introduction of a chemical

brace, a hydrocarbon staple located between two unnatural amino acids; this modification shows improvements in the helicity and stability of the peptides in solution, therefore stapling can greatly refine the pharmacologic performance of the peptides. Peptides were synthesized, purified and labeled in house by Dr. Naomi Robertson in the Department of Chemistry (**Table 2.2**).

Peptide	Sequence	Structure
Native SMRT	Ac-STNMGLEAIIRKALMG-NH ₂	
Stapled SMRT	Ac-STNMGLES ₅ IIRS ₅ ALKG-NH ₂	
TAT Native SMRT	H ₂ N-YGRKKRRQRRR-STNMGLEAIIRKALMG-NH ₂	
TAT Stapled SMRT	H ₂ N-YGRKKRRQRRR-STNMGLES ₅ IIRS ₅ ALMG-NH ₂	
GRIP1	Ac-KHKILHRLLQDSSC-NH ₂	

Table 2.2: Peptides used in this project. Peptides included in this table are the native and the stapled SMRT corepressor peptides based on the first nuclear receptor interaction domain of the SMRT corepressor protein and GRIP1 coactivator peptide based on the second nuclear receptor interaction domain of the GRIP1 coactivator protein. A TAT sequence was added to the corepressor peptides in order to study the ability of this sequence to improve the entrance of the peptides in the cell.

2.5.1 - Measurement of peptide concentration by NMR

All the chemicals were obtained commercially (Sigma-Aldrich) and provided by the Department of Chemistry. ^1H NMR spectra were obtained using a Bruker DPX300 spectrometer equipped with a 60-position automatic sample changer (BACS-60) operating at 400 MHz for proton NMR. Both the p-nitrophenol and the peptides were previously dissolved in D_2O to avoid water interferences. Tetramethylsilane (TMS) was used as internal standard and all the measurements were run at room temperature. Additionally to the chemical shift that indicates how many different types of hydrogens are found in a molecule, integrations reveal the number of hydrogens of each type. So integration reveals the ratio of one type of hydrogen to another within the same spectrum.

Therefore, the measurement of the peptide concentration was achieved comparing the integration signal from ^1H chemical shift of p-nitrophenol protons (**Figure 2.4**) with the integration of the ^1H chemical shift from some protons carefully selected from the different peptides. For that reason, prior to performing the measurements, peptide sequences were analyzed and the peptides could be divided in three types depending on the ^1H chemical shift measured. Native SMRT and GRIP 1 peptides concentration was determined by measuring the integration of the acetyl group located in the N terminus of the peptide, stapled peptide concentration was determined by the integration of the signal from the hydrogens of the stapled double bond and integration from the methionine acetyl groups established TAT peptides concentration.

First of all, accurate concentration of p-nitrophenol in D_2O was measured by UV spectroscopy. The average of three independent measurements was used to calculate the p-nitrophenol concentration according to the Beer-Lambert law, knowing the p-nitrophenol extinction coefficient at 320 nm. Secondly, 200 μl of this solution was mixed with 300 μl of the peptide solution in a 35 mm NMR tube. After 10 seconds of sonication, NMR tubes were analyzed and the spectra obtained were used to calculate the concentration of the peptide present in the tube.

The whole procedure was performed with the assistance of Dr. Yana K. Rennie (Department of Chemistry, University of Leicester).

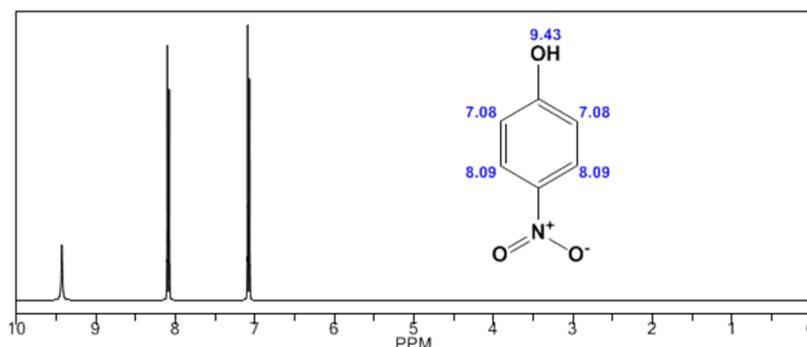


Figure 2.4: ¹H NMR spectrum of p-nitrophenol.

2.5.2 - Peptide coupling

Two different fluorophores were coupled to the peptides in order to use them for FA experiments.

FITC-labeling peptides were obtained by coupling the FITC fluorophore to the amino terminal of the peptide still attached to the resin. This procedure was performed by Dr Naomi Robertson.

BODIPY-TMR C₅ maleimide (Invitrogen) was coupled to peptides through an N-terminal cysteine residue, which was introduced to the peptides during synthesis. 90 μM peptide was incubated with a 5 fold molar excess of BODIPY in a 1 ml reaction with constant stirring for 2 hours in darkness at room temperature. The purification of the labeled peptide from free dye was performed using a PD-10 column (GE Healthcare) pre-equilibrated with 1xPBS containing 0.5 % TCEP. Eluted fractions were concentrated to 50 μl using an Amicon centrifugal concentrator.

2.6 - Fluorescence anisotropy assay

Using 96 well black plates (Corning Life Sciences), a constant amount of fluorescent peptide (FITC-labeled or BODIPY-labeled peptides) was titrated with increasing

concentrations of TR α LBD. The fluorescence anisotropy (FA) value was measured at each receptor concentration in a final volume of 50 μ l after incubating at room temperature for 5 minutes with slow shaking and centrifuging the multiwell plates. The FA reaction was performed in a buffer containing 1xPBS, 0.01 % Triton X-100, and 0.1 mg/ml Bovine Serum Albumin (BSA). FA values were measured with a Victor X5 multilabel plate reader (Perkin Elmer, Singapore) using a 480 nm excitation filter and 535 nm emission filters to measure FITC emission and 542 nm excitation filter and 572 nm emission filters to measure BODIPY fluorescence. Blank fluorescence values were subtracted in each polarization plane.

FA data were used to determine the equilibrium dissociation constants (Kd), using the Prism software (Graphpad) and the nonlinear regression analysis (saturation binding equation, one-site specific binding: $FA = B_{max} \times [ligand] / (Kd + [ligand])$). B_{max} (dimensionless) represents the maximum recordable FA value, which is indicative of maximum specific peptide binding. And the Kd equals the concentration of receptor at which the FA value is half of the maximum FA value ($FA = B_{max}/2$). Or, in other words, Kd represents the concentration of receptor at which the free concentration of fluorescent ligand is half the total concentration of the ligand. In any case, Kd values indicate the affinity of the interaction between the ligand and the protein.

2.7 - Freeze-drying

T3 analogues, previously dissolved in ethanol to perform the CD experiments, were dissolved in 100 % DMSO after removing the solvent by the freeze-drying method. T3 analogues samples were transferred to a 15 ml falcon tube and rapidly frozen by direct submersion in a liquid nitrogen batch (-195.79 $^{\circ}$ C, 77 K). The prefrozen tubes were quickly introduced in a vial and attached to the drying chamber or manifold to prevent warming. Prior to that, a vacuum pump had been started to create low pressure (1 mbar) in the drying chamber cooled to -50 $^{\circ}$ C. Vacuum was created in the product container quickly, and the sublimation of the solvent started; the collector condensed out all condensable gases, and the vacuum pump removed all non-condensable gases. The samples were left overnight.

2.8 - Circular dichroism

Thermal unfolding of proteins was monitored by CD spectroscopy, over a wavelength range of 200-250 nm, using a Chirascan Spectrometer (Applied Photophysics) equipped with a temperature controller (Quantum Northwest TC125). CD spectra were measured from samples in 1 mm path length quartz cuvettes, using a scanning speed of 100 nm/min, a spectral bandwidth of 1 nm, and a response time of 1 second.

The folding or secondary structure of the proteins was assessed by visual inspection of CD spectra from 200-250 nm (**Figure 2.5**). Helical regions are indicated by two troughs at 210 and 222 nm whereas β sheet is indicated by a single trough at approximately 216 nm. Unstructured material is characterized by a low absorption between 190 and 200 nm with a low flat peak between 210 and 230 nm often seen. Then, the thermal denaturation of the proteins was characterized by measuring the ellipticity changes at 222 nm induced by a temperature increase from 20 to 90 °C at steps of 1 degree. Samples of 1 mg/ml protein in 1xPBS and 1 mM DTT were analyzed by CD in order to obtain the melting temperature of the proteins alone or in the presence of different ligands.

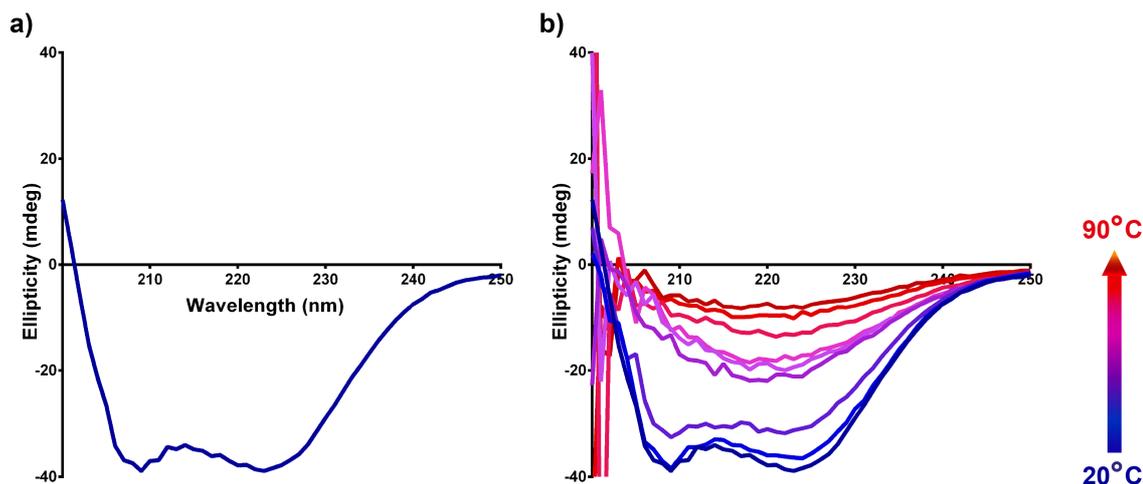


Figure 2.5: CD spectrum of the purified E403X LBD mutant protein. a) CD spectrum at 20 °C that typically corresponds to proteins containing elements of α -helical structure characterized by two negative bands at 208 and 222 nm. b) Change in the CD spectrum as the temperature increases. The helical content of the protein is decreasing until the helical signal at 222 nm is almost 0, which means that the protein is unfolded.

2.9 - Crystallization trials

Stock solutions for crystallization containing protein, peptide and ligand were prepared prior to dispensing on to the crystallization plates. Crystallization trials were performed into MRC 96 well sitting drop crystallization plates that were set up using robotics. Each reservoir well was manually filled with 80 μ l of mother liquor from commercially available crystallization screens (Section 2.1.5). Protein sample and precipitant were dispensed with equal volumes to produce a 100 nl drop. The plates were finally sealed with a transparent sheet that allowed checking for the formation of protein crystals with the aid of an optical microscope.

2.9.1 - Protein sample preparation

The different proteins of interest were expressed and purified as described in section 2.4. Proteins were kept in a Tris-base buffer (20 mM Tris pH 8, 50 mM NaCl and 1 mM DTT) at about 3 mg/ml and prior to crystallization, were combined with the other components of the experiment if necessary and the mix was concentrated up to

approximately 8 mg/ml. Protein concentration was measured following the Bio-Rad method previously described in section 2.4.

Peptides were synthesized as described in section 2.5 and diluted in ultrapure 18 M Ω water to a concentration of 2 mg/ml. The concentration of the peptides was defined by solubilizing a previously determined mass.

Ligands were obtained either by purchasing from Sigma-Aldrich such as T3 or kindly donated by Professor Nick Tomkinson. T3 was dissolved in 100 % ethanol up to a concentration of 10 mM. T3 analogues were dissolved in 100 % DMSO to the same concentration.

2.9.2 - Cryoprotection, collection and freezing of crystals

Prior to analysis, crystals were frozen and kept in liquid nitrogen (-195 °C) in order to take them safely to the synchrotron and reduce the rate of radiation damage. Rapid freezing could cause the formation of small ice crystals from the aqueous solution surrounding the protein crystals which can diffract X-rays producing a ring-like in the diffraction pattern that would affect the quality of the data collected. In addition, the cooling process increases the amount of disorder within the crystal making them weaker and reducing the quality of the diffraction data by increasing the mosaicity. In order to prevent this, crystals were frozen using a cryoprotectant solution consisting of mother liquor and the minimum effective amount of a cryoprotectant such as glycerol. Cryoprotectants lower the freezing point of the solvent increasing the rate of cooling and disrupting the nucleation and formation of ice. Cryoprotectants also make the aqueous solution to freeze as an amorphous glass without significant diffraction.

A nylon loop is used to pick single crystals from the drop, once the well containing them had been opened and some cryoprotectant had been added to it. Nylon loops accommodating the crystals are quickly put into a puck submerged in liquid nitrogen.

2.9.3 - Data collection

Crystals were taken to the Diamond Light Source synchrotron (UK) where they were analyzed at the microfocal X-ray beamline I-24.

2.10 - JEG-3 mammalian cells: tissue culture and transient transfection

2.10.1 - Preparation and culture of JEG-3 cells

In order to avoid contamination and work the whole time in sterile conditions, a class II laminar flow hood was used for growing, maintaining, transfection and inducing JEG-3 cells. Before starting to work, both the hood and the required materials were carefully sterilized with 70 % ethanol.

A vial of frozen JEG-3 cells were removed from liquid nitrogen and thawed quickly in a 37 °C water-bath. Immediately, cells were transferred to a sterile, conical 50-ml tube containing 30 ml of pre-warmed culture medium, and mixed gently. Then, 2 ml of the cell suspension was placed into each of three 35 mm glass culture plates containing 14 ml of fresh pre-warmed medium, labeled and grown in a 37 °C incubator supplied with 5 % CO₂. The media was replaced after 24 hours and the process repeated 3 times before cells were used for transfection to ensure cells were fully recovered from freezing.

2.10.2 - Seeding and passaging of JEG-3 cells

In order to maintain and expand the cell stock, JEG-3 adherent cells were sub-cultured in Opti-MEM media supplemented with 10 % FCS and 1 % PSF when they had reached an approximate 85-90 % of confluence. The media was removed from the plates by aspiration and the cells were washed once with 10 ml of sterile PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 2 mM KH₂PO₄, pH 7.4) before detaching using 1 ml of trypsinization solution (0.25 % trypsin, 1 mM EDTA) pre-warmed at 37 °C. After 1 minute of incubation at 37 °C, trypsin was neutralized by re-suspending the cells into 5 ml of media. Finally, 500 µl of cells were placed into one or more 35 mm plates

containing 10 ml of fresh pre-warmed media and then incubated in a humidified incubator at 37 °C and 5 % CO₂.

2.10.3 - Transient transfection and cotransfection of JEG-3 cells

Eighteen hours prior to transfection, the media was replaced with Opti-MEM containing 10 % resin-stripped FCS and 1 % PSF and the cells were seeded into 96-well plates. The plates were incubated for 18 hours into a humidified incubator at 37 °C and 5 % CO₂.

Once cells had reached about 80 % confluence after incubation, were ready for transfection. As all the experiments were performed in triplicate, for every three wells the following solutions were prepared:

Solution A:

30 µl plain Opti-MEM + 1 µl Lipofectamine 2000

Solution B:

1 µl mix of plasmid DNA used for the transfection (purified from DH5α using a QIAprep midiprep Kit from Quiagen).

Solution A was incubated at room temperature for 5 min then added to solution B and incubated at room temperature for a further 20 minutes. After incubation the mixture was added to the cells. Transfected cells were incubated for 4 hours in the humidified incubator so that they could incorporate the exogenous DNA and then were exposed to increasing concentrations of T3 (from 0 to 10,000 nM). Following additional 36 hours of incubation at 37 °C after T3 exposure, cells were harvested and assayed for reporter activity.

2.11 - Luciferase and β-galactosidase dual reporter assay

Transfection procedure was performed to measure the ability of a nuclear receptor to induce the transcription of a luciferase reporter gene in response to a ligand, T3 in this

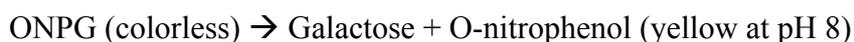
case. Therefore, in order to measure the luciferase activity produced by the cells, transfected and induced cells were harvested and lysed in glycyl-glycine buffer (0.5 M glycyl-glycine pH 7.8, 1 M MgSO₄, 180 mM EGTA) containing 1 % (v/v) Triton X-100 and 1 mM DTT. Extracts were analyzed using an Auto Lumat LB 960 Microplate Luminometer (Berthold Technologies) to measure luciferase activity and an EMAX Plus Microplate Reader (Molecular Devices) to determine β-galactosidase activity.

50 μl of luciferase assay buffer containing 1x glycyl-glycine buffer, 100 mM K phosphate, 200 mM ATP and 100 nM DTT was added to every well before adding 100 μl of 100 mM luciferin, in 1x glycyl-glycine buffer and 1 M DTT. The reaction catalyzed by luciferase produced by the cells in response to a ligand is the following **(Figure 2.6)**:



The emitted light is measured by the luminometer.

In order to determine the efficiency of the transfection, β-galactosidase activity was also assayed and used to normalize luciferase values for transfection efficiency. A β-galactosidase solution was required and consisted of 100x Mg solution (1 M MgCl₂, 4.5 M β-MeSH), 1xONPG (O-nitrophenyl-β-D-galactopyranoside) and 0.1 M Na phosphate pH 7.5. 150 μl of this solution was added to every well to perform the following reaction which is catalyzed by the β-galactosidase produced by the transfected cells **(Figure 2.6)**:



Absorbance at 420 nm from the o-nitrophenol was measured by the microplate reader.

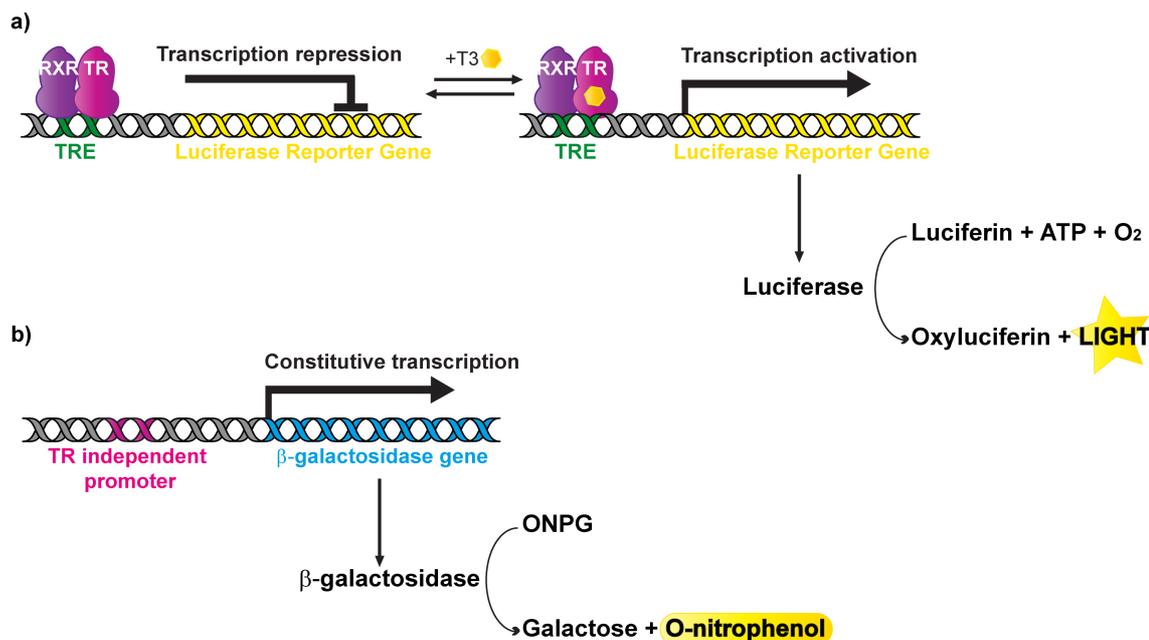


Figure 2.6: Schematic representation of the luciferase and β -galactosidase reporter assay.

a) In response to T3, TRs activate the transcription of luciferase enzyme which produces light. The luminescence is proportional to the activity of the receptor. b) β -galactosidase activity measured by the amount of o-nitrophenol produced by the cells.

2.12 - Western blot analysis

Cells were harvested after 48 hours post-transfection and 24 hours of T3 exposure (section 2.10.3). After washing with PBS and aspirating the media, cells were lysed into 1.0 ml of lysis buffer and scraped off the plate into 1 ml of fresh PBS using a plastic cell scraper. Cell were transferred into centrifuge tubes and spun down at 4000 rpm at 4 °C for 10 minutes. The pellets were discarded and the supernatants were run on SDS-PAGE gels for 40 minutes at 200 V using Pre-cast 4-12 % gels (Invitrogen) and commercially available MES-SDS running buffer (section 2.4.6).

Proteins were transferred onto a nitrocellulose membrane using a semi-dry method (ThermoFisher) following the manufacturer's instructions. After the transfer, the membrane was blocked for 1 hour with milk blocking buffer on a rocker. The membranes were then washed 3 times for 10 minutes in PBS and incubated over night at 4 °C in falcon tubes containing the primary antibodies diluted (1/500) in milk blocking buffer. Following the incubation, the antibody solutions were discarded and the membranes washed 3 times for 10 minutes with PBS.

Next, the membrane was incubated for 2 hours with horseradish peroxidase conjugated secondary antibodies previously diluted (1/10,000) into milk blocking buffer. The membrane was washed again 3 times for 10 minutes in PBS. Bound antibodies were visualized by ECL, using the Amersham ECL detection kit (GE Healthcare) used in accordance with manufacturer`s instructions.

Chapter 3 - Biophysical characterization of the interaction between WT and mutant TR α LBDs with their coregulator proteins

3.1 - Introduction

This project is focused on the biophysical and biochemical characterization of three mutations identified in the *THRA1* gene. These mutations were identified in four patients with symptoms of tissue-specific hypothyroidism, associated with low to low-normal T4 levels and high to high-normal T3 levels, a sub-normal T4/T3 ratio, low levels of rT3, and normal levels of TSH. The patients were affected by RTH α (Bochukova et al. 2012; van Mullem et al. 2012; Moran et al. 2013).

i) Molecular genetics and functional properties of the TR α mutants

The RTH α disorder is mediated by heterozygous, loss-of-function, mutations involving *THRA1*. The first mutation characterized is a non-sense mutation c1207 G \rightarrow T, p. E403X that results in an abnormal receptor lacking the last 8 amino acid residues (Bochukova et al. 2012). The second mutation identified is a single nucleotide insertion (c1190 ins. T) that results in a frame-shift from codon 397 of TR α and a premature truncation of the protein five residues before its carboxy-terminus (F397fs406X) (van Mullem et al. 2012). The third mutation is a single nucleotide deletion (c1144 del. G) that shifts the reading frame at codon 382 which alters the six subsequent residues and then introduces a premature stop codon (A382PfsX7), deleting 22 residues in the C-terminal (Moran et al. 2013) (**Figure 3.1**).

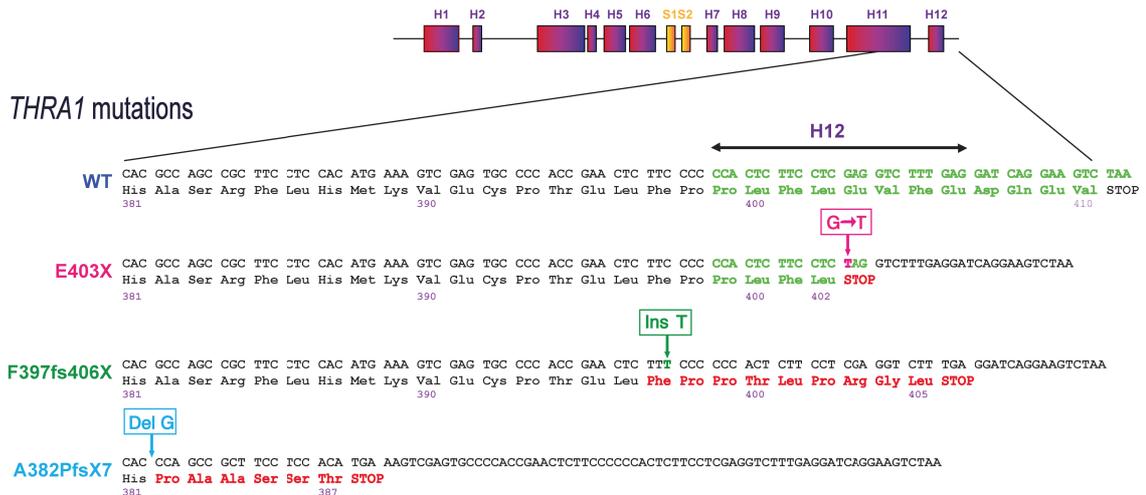


Figure 3.1: DNA and amino acid sequence of C-terminal WT and mutant TR α . 3' end DNA sequence of the *THRA1* wild type and the three mutants in study, indicating the position and type of mutation, and the consequences of the mutation in the amino acid sequence of the proteins.

WT TR α associates with corepressor complexes to bring about transcriptional repression in the absence of T3. T3 binding causes some conformational changes in the structure of the receptors, especially in H12 which adopts a highly ordered helical conformation. The active conformation of the receptor produces the release of corepressor complexes and leads to the recruitment of coactivator complexes responsible for the activation of T3 target gene transcription.

However, all the mutations studied in this thesis result in prematurely truncated TR α , without the C-terminal region (H12). Previous results indicated that the TR α mutants have minimal T3-dependent transcriptional activation and constitutively repress the transcription of T3 target genes. Hence, when coexpressed, these TR α mutants inhibit transcriptional activation of their WT receptor counterparts in a dominant negative manner (Bochukova et al. 2012; van Mullem et al. 2012; Moran et al. 2013).

ii) Clinical features

The first four patients identified had some common clinical features due to the lack of T3 response in the specific tissues where mutant *THRA1* is expressed. However, the level of impairment depends on the position and severity of the mutation. The symptoms of the disease include short stature due to growth retardation that mainly

affects the lower segment of the body, delayed tooth eruption, and severe constipation, in addition to mild cognitive deficits, decreased muscle tone and delayed motor development that leads to impaired gross and fine motor skills. The delayed bone development in these patients confirmed that TR α 1 plays a major role in bone development (Bassett & Williams 2008). The delay in motor development and the mild cognitive deficits are in accordance with the important role of TR α 1 in brain development (Venero 2005). Similarly, as the TR α 1 is the predominant isoform in the heart, defects in the cardiovascular system were expected and included decreased heart rate and blood pressure. Interestingly, the patient carrying the A382PfsX7 frame shift mutation has epilepsy and severe cognitive impairment in addition to the other symptoms (**Table 3.1**).

In contrast to patients with RTH β , the hypothalamus-pituitary axis in patients carrying *THRA1* mutations is minimally affected, highlighting that TR β is the main isoform expressed in the CNS. Therefore, the patients showed normal circulating TSH levels, normal or low-normal levels of total T4 and free T4, high-normal or elevated levels of total T3 and free T3, and low levels of rT3, resulting in markedly subnormal ratios of free T4 to T3 (FT4/FT3) (**Table 3.1**).

	TRα1- A382PfsX7	TRα1- F397fs406X	TRα1- E403X
Genotype			
Mutation	Frame shift	Frame shift	Nonsense
Zygoty	Heterozygous	Heterozygous	Heterozygous
Inheritance	<i>De novo</i>	<i>De novo</i> and inherited	<i>De novo</i>
Phenotype			
Appearance	Flattened nasal bridge Broad face, thickened lips Macroglossia Coarse facies, skin tags and moles	Flattened nasal bridge Broad face, thickened lips Macroglossia	Flattened nasal bridge Broad face, thickened lips
Bone development	Disproportionate short stature (growth deficit in the lower segment of the body) Macrocephaly Delayed tooth eruption	Disproportionate short stature (growth deficit in the lower segment of the body) Delay bone development Macrocephaly Delayed tooth eruption	Disproportionate short stature (growth deficit in the lower segment of the body) Delay bone development Macrocephaly Delayed tooth eruption
Constipation	Severe	Mild	Severe
Mental development	Cognitive impairment Epilepsy	Mild cognitive deficits Delayed motor development	Cognitive impairment Reduced muscle tone Slow initiation of movement Fine and gross motor incoordination (dyspraxia)
Cardiovascular	Low heart rate		Low heart rate and blood pressure
Metabolic	Low metabolic rate		Low metabolic rate
Hematological	Mild anemia	Mild anemia	Mild anemia
Biochemical markers			
TSH	Elevated-normal	Normal	Normal
fT4	Low-normal	Low-normal	Low-normal
fT3	Normal	Elevated	Elevated-normal
T3/T4 ratio	Elevated	Elevated	Elevated
rT3	Low	Low	Low
SHBG	Elevated		Elevated

Table 3.1: Summary and comparison of the clinical features, genotype and phenotype of RTHα patients. The patients are carrying three specific mutations in the LBD region of the *THRA* gene (A382PfsX7, F397fs406X and E403X) (Bochukova et al. 2012; van Mullem et al. 2012; Moran et al. 2013).

Little structural and biophysical work has been carried out upon the mutant TR α since the majority of the analyses have been performed in cells. The biophysical understanding of the mechanism of the mutant LBDs action at molecular level is essential to design strategies to treat the disease. Mutant TR α activity is modulated through the action of T3 and the functional interaction with coregulator proteins such as SMRT corepressor and GRIP1 coactivator. Therefore, the aim of this chapter is to characterize the interaction between the WT and the mutant A382PfsX7, F397fs406X and E403X LBDs with their coregulator proteins by interaction assays and spectroscopy techniques in the absence and the presence of T3.

3.2 - Design of WT and LBD TR α constructs

WT and mutant A382PfsX7, F397fs406X and E403X LBD constructs were prepared in order to investigate the effect of the mutations in the LBD ability to interact with coregulator proteins such as corepressors and coactivators. The constructs were designed to contain an N-terminal GST tag followed by a TEV protease site for expression and purification in *E. coli*.

The template used to create the different constructs was the full-length WT TR α which was kindly provided by Dr. Maura Agostini (University of Cambridge). The objective was to create shorter constructs that only contained the LBD region of the receptor. Therefore, several combinations of primers were designed to specifically produce the WT and the mutant LBDs and tested in their ability to produce protein (section 2.2.1, materials and methods).

The 5' primers were designed based on predicted secondary structure elements, disorder prediction parameters, and previous structures solved of the TR α LBD (**Figure 3.2**). The first one, based on the structure of the TR α LBD, started at residue 152 (Nascimento et al. 2006). Other two primers were chosen based on the predicted secondary structure of the hinge domain, the region between the DBD and the LBD, and trying not to truncate the protein at low entropy regions or regions of high conservation. These two additional primers were designed to produce proteins from residues 130 and 148.

Regarding the 3' primers, the WT construct was designed to extend to the natural C-terminal sequence without any modification (residue 410). The other oligonucleotides were designed to introduce the appropriate mutation in the *THRA1* gene: a single nucleotide change at position 1207 (G to T) to make the E403X mutant receptor, a T single nucleotide insertion after position 1190 that results in a the frame-shift mutation that introduces six different residues and ends at 406 (F397fs406X), and a G single nucleotide deletion at position 1144 that results in other frame-shift mutation that introduces six different residues before ending prematurely (A382PfsX7).

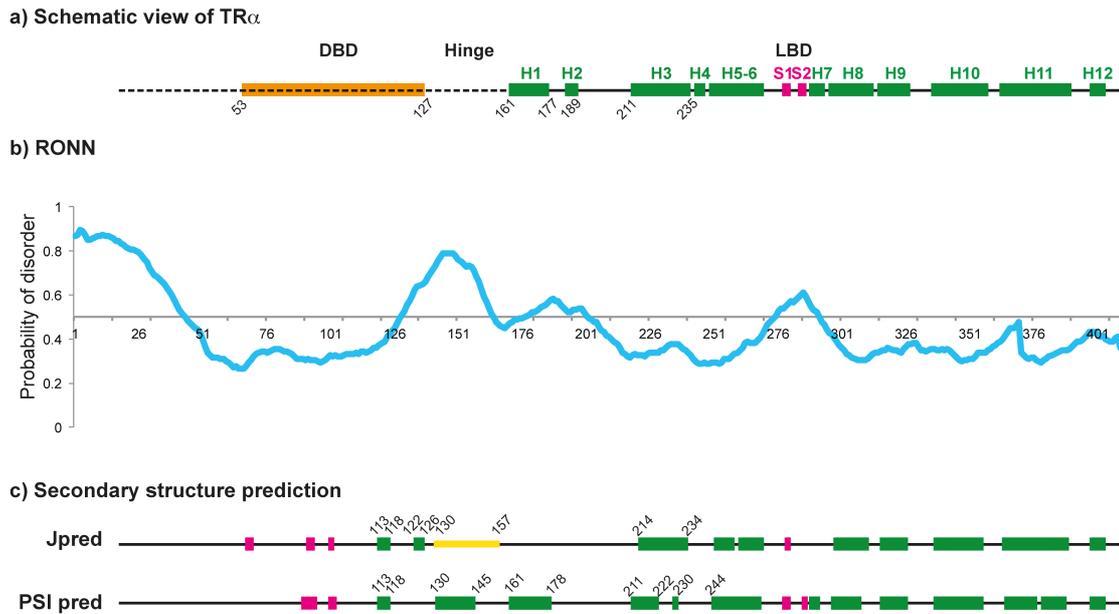


Figure 3.2: Disorder and secondary structure predictions of full-length TR α . a) Schematic view of the TR α structure, showing a dotted line for the regions with unknown structure (DBD and hinge) and a continuous line with the different secondary elements for the LBD region (TR α LBD structure based on the 2H79) b) RONN (Yang et al. 2005), c) Jpred, and PSIPred (Cole et al. 2008). Green box: α -helix. Pink box: β -sheets. Yellow line: coiled-coil. Black line: disordered.

3.3 - Optimization of protein expression and purification

Three WT LBD constructs were cloned into pLEICS14 *E. coli* expression vectors. The best expressing protein was selected prior to preparing the mutant constructs. Purification trials were performed to select the best N-terminal boundary in terms of expression level and purification yield. The cloning was performed by Dr. Xiaowen Yang and Ms. Dipti Vashi from the Protex service (University of Leicester).

A GST affinity purification step was performed to every sample. Samples were lysed and purified using 3 ml affinity resin by the batch method before expression was verified by SDS-PAGE. Detection of protein at the expected mass by SDS-PAGE indicated that all the WT LBD constructs were expressed and the expression level was sufficient for further applications such as structural and functional studies. However, there were some contaminants in the samples containing the target protein and further purification was required (**Figure 3.3a**).

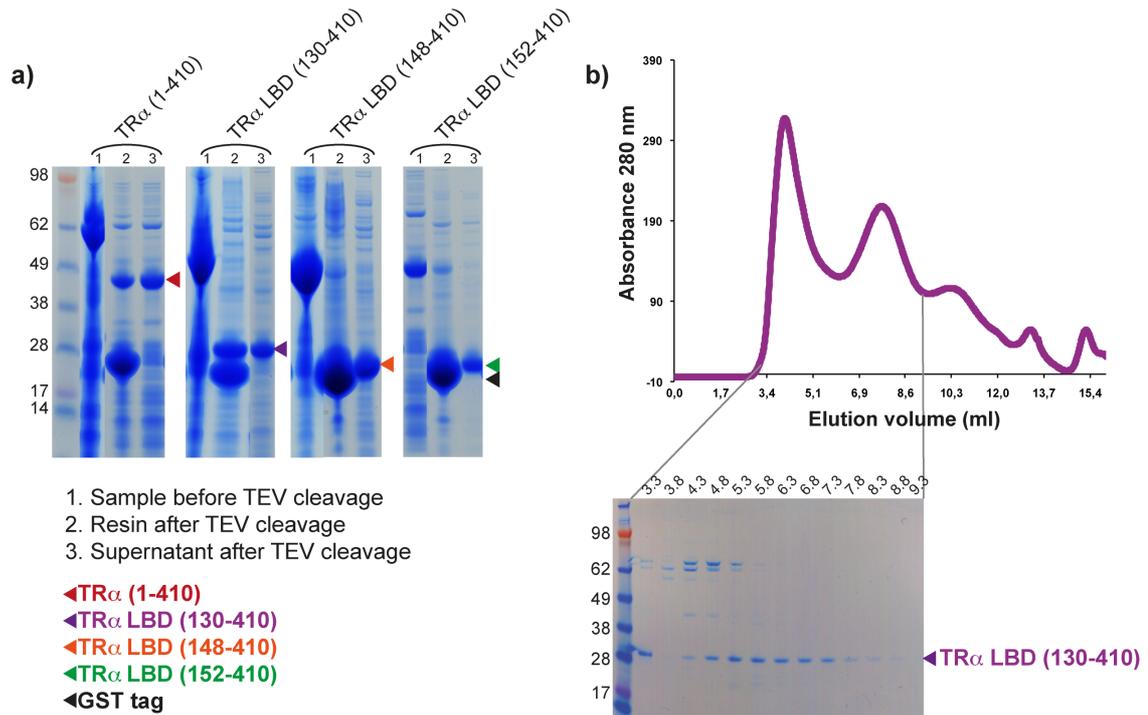


Figure 3.3: Expression and purification of TR α constructs. a) SDS-PAGE analysis of expression and GST purification of the different TR α constructs purified using glutathione sepharose 4B. b) S75 GF chromatography of the 130-402 TR α construct.

TR α LBD 130-410 (Mw 31.3 kDa, pI 7.26) was purified by gel filtration chromatography (S75 GL). The results showed that all the fractions obtained contained the expected protein and many others contaminants, meaning that the purification was not completely efficient (**Figure 3.3b**).

TR α LBD 148-410 (Mw 28.8 kDa, pI 5.74), and TR α LBD 152-410 (Mw 28.3 kDa, pI 5.87) were expressed and purified on GST resin followed by HiTrap Q HP IEX column and the purification was successful. High levels of pure protein were obtained after running through the column. The greatest yield of protein was observed with the TR α LBD 148-410 construct. The mutant constructs (A382PfsX7 LBD, F397fs406X LBD, and E403X LBD) were produced with the same 5' primer.

After sequencing the constructs and transforming *E. coli*, the LBDs were purified following the same strategy, a GST affinity purification step followed by an IEX chromatography. The GST elution was relatively pure with a few contaminating

proteins accounting for less than 20 % of the material observed (**Figure 3.4, 3.5**). A high level of expression was achieved for all the different constructs and the IEX chromatography removed the majority of the contaminating bands after TEV cleavage removed the affinity tag.

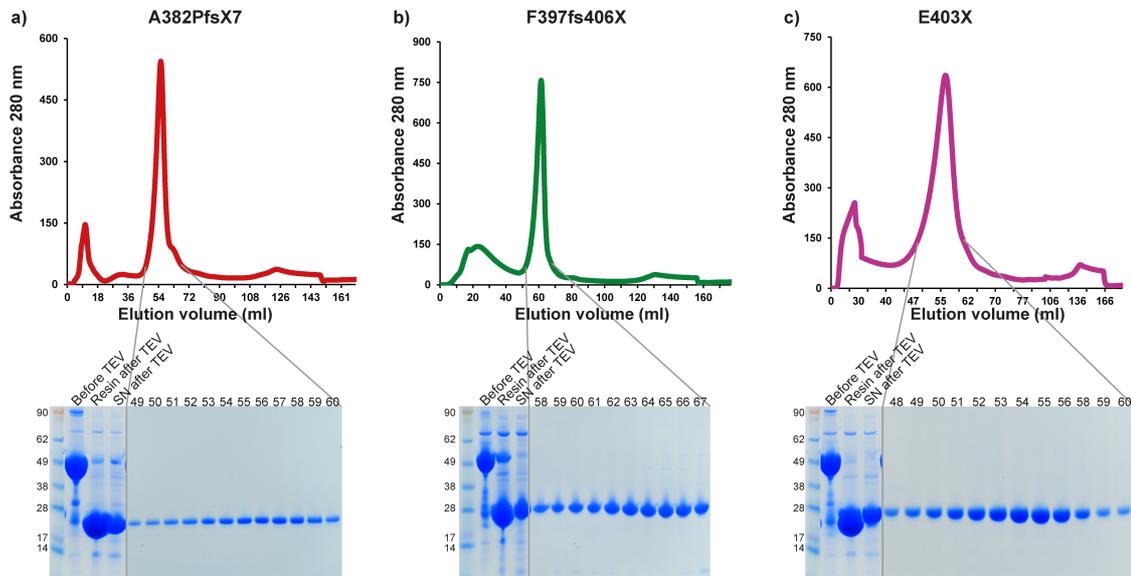


Figure 3.4: Purification of mutant TR α LBD constructs by HiTrap Q IEX. Purification of mutant TR α LBDs constructs using GST-tag affinity chromatography followed by HiTrap Q IEX chromatography. The purity and yield was assessed by SDS-PAGE analysis after every purification step. a) A382PfsX7 LBD, b) F397fs406X LBD, and c) E403X LBD.

The SDS-PAGE of the WT LBD revealed some contaminating bands after the IEX chromatography indicating that a further purification step was required. Therefore, after concentration of the fractions that contained the protein, partially purified WT LBD was loaded onto the S75 GL chromatography to remove the copurified contamination (**Figure 3.5**).

Protein purity (> 99 %) was assessed by SDS-PAGE, and protein concentration was determined using the Bradford dye assay (Bio-Rad). The average yield of the proteins was 2-3 mg per liter of culture.

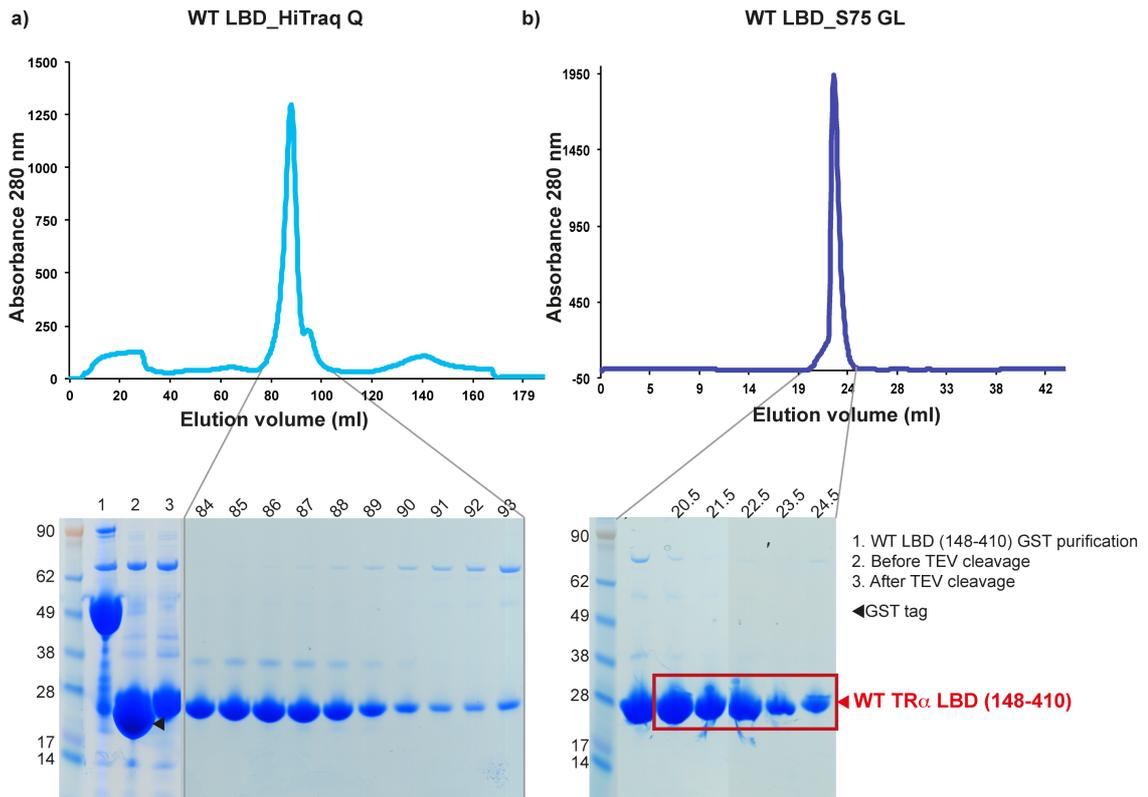


Figure 3.5: WT purification process. a) HiTrap Q IEX of the WT LBD and SDS-PAGE analysis of the protein expression and purification. b) Purification of the WT LBD by S75 GL chromatography and SDS-PAGE analysis of the fractions obtained.

3.4 - Fluorescence anisotropy assays to study the interaction between the TR α LBD and coregulator peptides

The binding affinity between the different LBDs and their coregulator proteins were measured by FA. FA is a powerful tool for studying molecular interactions between a large molecule such as a protein and small ligands or peptides previously coupled to a fluorophore. Due to the rapid tumble of small fluorescent molecules in solution during its fluorescence lifetime (the time between excitation and emission), when the molecule is excited with plane-polarized light, the emitted light is largely depolarized. However, if the fluorophore is bound to a larger molecule its effective molecular volume increases, the fluorophore rotation slows and the emitted light becomes polarized. As a consequence, the bound and free states of the fluorophore have an intrinsic polarization value: high values for the bound state and low values for the free state. Thus, the

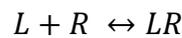
measured polarization or anisotropy provides a direct measure of the fraction of fluorophore or ligand bound to the receptor.

The relationship between the intensities of the detected light can be expressed in fluorescence polarization (FP) or FA units. The FP measures the light emitted in two planes, parallel (I_{\parallel}) and perpendicular (I_{\perp}). The FA measures the light detected in three planes, the same plane as that of the exciting light (parallel I_{\parallel}) and the two other perpendicular planes (I_{\perp}) according to the following formula:

$$P = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + I_{\perp}}$$

$$A = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}}$$

In order to quantitatively analyze the binding of a small fluorescent molecule to a larger protein, saturation binding curves are generated by varying protein concentration and measuring the fraction of the ligand bound to the protein at every concentration by FA. These FA values are used to calculate the equilibrium K_d of the interactions. The K_d provides a measure of the strength of the interaction between the protein and the ligand and corresponds to the concentration of ligand (L) at which half of concentration of a particular protein (R) is bound to ligand and the other half is ligand free.



$$K_d = \frac{[L][R]}{[LR]}$$

$$\alpha = \frac{[LR]}{Rt} = \frac{[L]/K_d}{1 + [L]/K_d}$$

When α equals 0.5 (i.e., 50 % of R have bound L) then $[L] = K_d$. Therefore, the K_d can be calculated by measuring the saturable relationship between $[L]$ and $[LR]$ at equilibrium.

Two peptides were designed for using in the FA assay. An N-terminal FITC labeled 14-aa length peptide with sequence based on the interaction domain 1 of the SMRT corepressor protein (CoRNR box 1: Ac-STNMGLEAIRKALMG-NH₂), containing the corepressor NR recognition motif LxxxIxxx[I/L], and a C-terminal BODIPY-TMR labeled 16-aa length peptide with sequence based on the second NR interaction box of GRIP1 coactivator protein (NID 2: Ac-KHKILHRLQLDSSC-NH₂) containing the coactivator NR recognition motif LxxLL (Li et al. 1997; Darimont et al. 1998; Fen Ding et al. 1998; Nagy et al. 1999; Webb et al. 2000). Both peptides were synthesized and purified by Dr. Naomi Robertson (Department of Chemistry, University of Leicester).

3.4.1 - Optimization of fluorescence anisotropy assays

FA experiments require thorough optimization processes in order to obtain reliable K_d values. The assay requires protein concentration to be much higher than labeled peptide concentration in order to assure that protein concentration does not change significantly when the peptide binds to the protein. Therefore, the aim is to reduce the concentration of labeled peptide to the lowest value compatible with keeping adequate fluorescence signal. An adequate fluorescence signal is provided when there is a measurable and significant change in FA, and a high signal/noise ratio, in response to the binding of the labeled peptide to the protein. It is recommendable to start with a concentration of labeled peptide of approximately ten times lower than the expected K_d and confirm whether there is an adequate fluorescence signal.

Previous studies have estimated that the interaction affinity between SMRT and WT TR α LBD is approximately 0.7 μ M in the absence of T3 and 2.7 μ M in the presence of T3 (Levy-Bimbot et al. 2012). Therefore a range of FITC-SMRT peptide concentration from 50 nM to 5 nM was used to be titrated against increasing concentration of LBD (0-50 μ M). Finally, a fixed concentration of 5 nM FITC-SMRT peptide was chosen to be the lowest fluorescent peptide concentration necessary to produce a significant difference in FA. The starting point of protein concentration was decided to be 5 μ M after finding out the saturation concentration.

Regarding the GRIP1 peptide and the WT TR α LBD interaction, it has been reported to be strongly T3-dependent (Levy-Bimbot et al. 2012). The affinity is 2.2 μ M in the presence of T3 and 20 μ M in the absence of T3. So, again a range of GRIP1-BODIPY peptide concentrations from 200 to 5 nM was tested in order to finally choose a fix concentration of 5 nM labeled peptide to perform the experiments.

3.4.2 - WT and mutant LBDs bind to SMRT peptides with similar affinity

In the absence of T3, unliganded WT LBDs interact with corepressor proteins, repressing gene transcription below the basal level. Mutant LBDs are believed to function as constitutive repressors as they never activate gene transcription. A study of the interaction affinity between the WT and the mutant LBDs and corepressor peptides was performed by FA in order to compare the affinity of the corepressor binding to the WT and to the mutant LBDs. The results will help to discern whether the constitutive repression is a consequence of a higher affinity of the interaction between the corepressor and the mutant LBDs that makes the corepressor remain bound to the mutant LBDs.

The characterization of the interaction between the WT and mutant LBDs and corepressor peptides was performed by FA using FITC-SMRT corepressor peptide. Increasing amount of the WT and mutant LBDs, expressed and purified as described above (section 3.3), were incubated with a fixed amount of FITC-labeled SMRT peptide. The K_d values were determined to be 158.5 nM for WT LBD, 91.1 nM for F397fs406X LBD, 90.8 nM for A382PfsX7 LBD and 212.2 nM for E403X LBD (**Figure 3.6**).

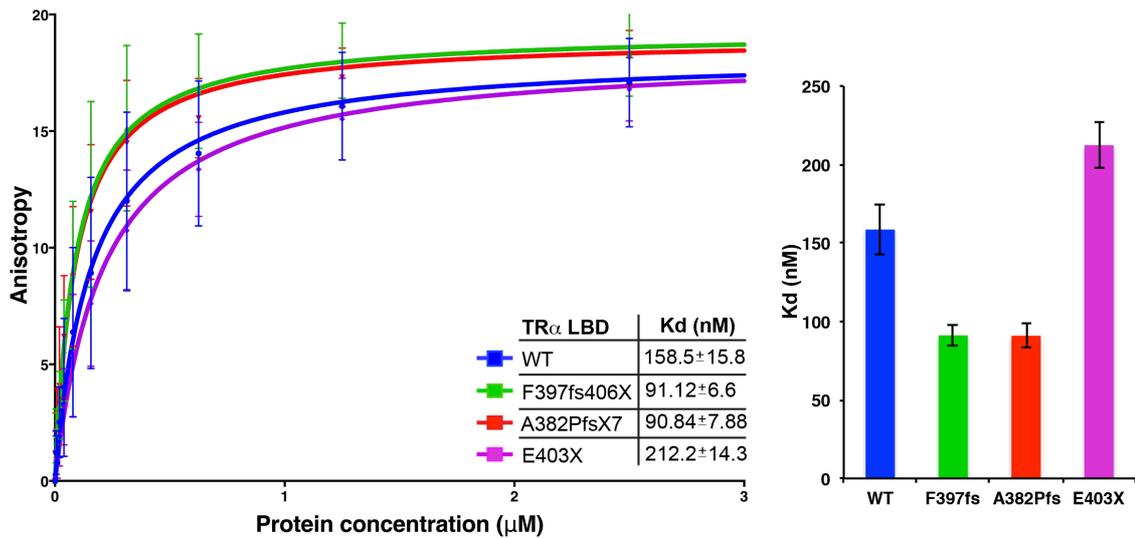


Figure 3.6: Binding assays of native SMRT peptide to TR α . Saturation-binding curves were generated for the WT LBD and the mutant LBDs in the absence of T3. The FA was plotted against an increasing concentration of the LBDs and the Kd was calculated using Graphpad prism software. FA values are the mean \pm standard error of the mean (SEM) of the measurements obtained from at least 5 independent experiments.

The results showed that increasing the concentration of LBDs promoted a dose-dependent increase in anisotropy of labeled SMRT peptide, reflecting an increase of the binding of this peptide to the LBD in all the cases. The interaction between the corepressor and the LBDs appear to be strong because all the Kd corresponding to the different LBDs are in the nanomolar range. The strongest interactions were found between the SMRT corepressor and the LBD mutations F397fs406X and A382PfsX7. The LBD mutant E403X and SMRT corepressor interaction was the weakest one (**Figure 3.6**).

The differences in affinity are probably related to the different C-terminal that shows the different LBDs. The WT LBD has a complete H12 which, in the inactive state, is highly mobile. Due to its dynamic properties, H12 moves independently of the rest of the protein and occupies different positions. This movement could be interfering in corepressor binding by partially occluding the corepressor-binding surface of the receptor. Since the LBD mutation E403X still has most of H12, the remaining C-terminal residues of the protein could also be partially occluding the corepressor-binding site even to a greater extent. The LBD mutations F397fs406X and A382PfsX7

generate the shortest proteins, without any residue of H12. Therefore, the C-terminal part of these proteins would not interfere in corepressor binding.

Nevertheless, these small differences in corepressor affinity between the WT and the mutant LBDs do not account fully for the constitutive binding of corepressor to the mutant LBDs reported *in vivo* ((Bochukova et al. 2012; van Mullem et al. 2012; Moran et al. 2013).

3.4.3 - Neither WT nor mutant LBDs are able to interact with coactivator peptides in the absence of T3

The characterization of the interaction between the WT and mutant LBDs with the coactivator peptides was performed in order to confirm that coactivator peptides are not able to interact with LBDs in the absence of T3. FA assays were performed using a fixed amount of BODIPY-labeled GRIP1 peptide incubated with increasing concentration of the LBDs.

The interaction of the BODIPY-GRIP1 peptide with the WT and E403X mutant LBD were not detectable. The K_d values of GRIP1 peptide binding to the LBD mutants F397fs406X and A382PfsX7 were significantly high (190.4 μ M and 23 μ M respectively). This suggests that none of the proteins is able to interact with coactivator proteins in the absence of T3 (**Figure 3.7**).

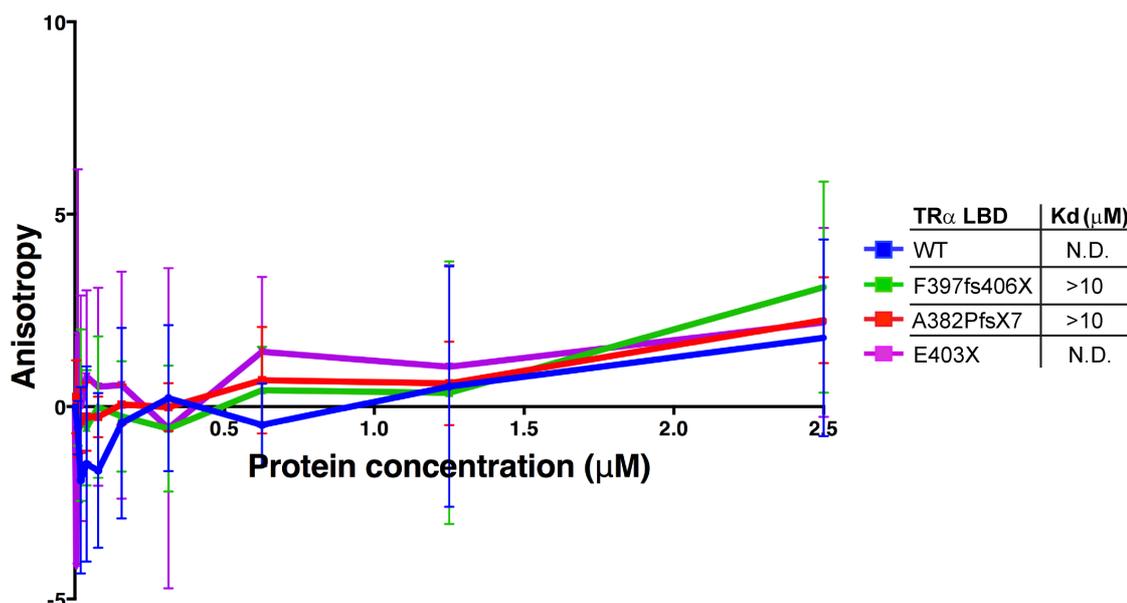


Figure 3.7: Binding assays of BODIPY-labeled GRIP1 coactivator peptides to TR α LBDs. Saturation-binding curves were generated in triplicate for the WT and the mutant LBDs in the absence of T3. The FA values were plotted against the increasing concentration of the LBDs and the Kd was calculated using Graphpad prism software. FA values are the mean \pm SEM of the measurements obtained from at least 5 independent experiments.

3.4.4 - Mutant LBDs do not release corepressor peptides in response to T3

Additional FA experiments were performed to further elucidate the molecular mechanism of coregulator binding to the different mutant LBDs in the presence of T3. T3 binding causes diverse conformational changes in the structure of the WT LBD that lead to the active state of the receptor which no longer supports corepressor binding, so a decrease in the affinity of the interaction between corepressors and the WT LBD is expected. On this section the effect of T3 in the affinity of the interaction between the mutant LBDs and the corepressor peptides is investigated.

The same concentration of LBD and labeled peptides previously determined were used to study the T3 potency to prevent the formation of the LBD:SMRT corepressor complex and to induce the LBD:GRIP1 coactivator complex. A molar ratio 1 to 10 (LBD:T3) was decided to assure that all the LBD molecules were bound to T3. An apparent Kd will be obtained showing the affinity of the interaction between the coregulators and the LBD in the presence of T3. A comparison between this apparent

K_d and the one obtained in the previous experiments will show the effect of the ligand in the affinity of the interaction.

The results showed that the affinity of the interaction between the WT LBD and the SMRT corepressor decreased almost 10 fold in the presence of T3 (**Figure 3.8a**). This suggests that SMRT peptide binding is T3-dependent in the case of WT LBD. The decrease in the affinity of the interaction between the WT LBD and SMRT corepressor in response to T3 indicates that the binding of T3 causes conformational changes in the WT LBD that lead to the release of the corepressor from the LBD. The active conformation of the WT LBD induced by T3 no longer permit the corepressor interaction.

The binding of SMRT corepressor to the liganded WT LBD (WT LBD:T3) observed in the graph at high concentration of the protein suggests that SMRT corepressor is able to bind to the active conformation of the LBD (**Figure 3.8a**). This apparent interaction is probably due to the mimetic activity of the corepressor. The coactivator and corepressor NR interaction motifs are reasonably similar to allow the corepressor to interact with the liganded state of the WT LBD since there is no coactivator to compete for the corepressor/coactivator binding site in the LBD.

On the other hand, the affinity of the interaction between mutant LBDs and SMRT peptide did not decrease in the presence of T3. It indicates that mutant LBDs do not release corepressor in response to T3 and therefore, SMRT peptide remains bound to the mutant LBDs in the presence of T3. Actually, mutants A382PfsX7 and F397fs406X bind SMRT peptide even tighter in the presence of T3, suggesting that T3 stabilizes the interaction between these mutant LBDs and the corepressor (**Figure 3.8**).

These results confirmed that mutant LBDs are indeed constitutively bound to corepressor proteins with high affinity independently of the presence of T3.

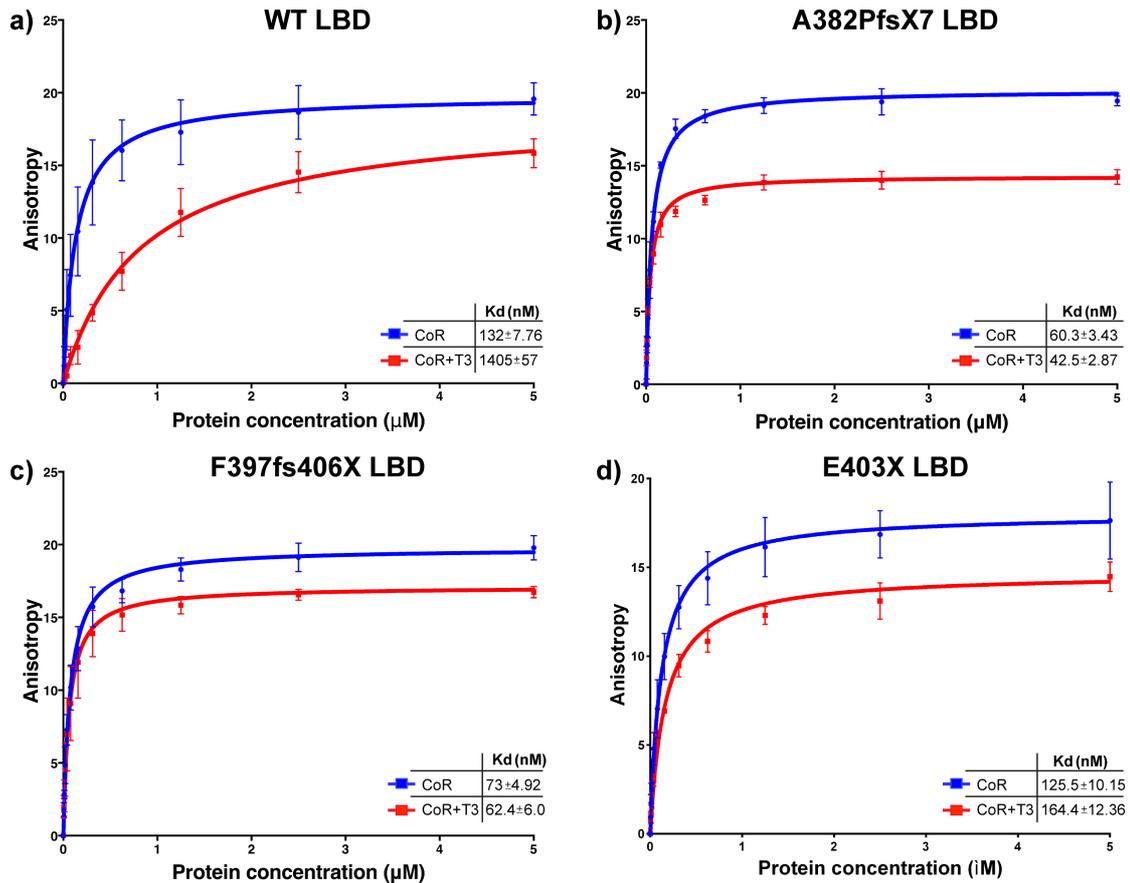


Figure 3.8: Binding assays of SMRT peptide to TR α LBDs. Increasing concentration of the mixture LBD:T3 (ratio 1:10) was incubated with 5 nM FITC-labeled corepressor peptide. FA values are the mean \pm SEM of measurements obtained from triplicate experiments. a) WT LBD, b) A382PfsX7 LBD, c) F397fs406X LBD, and d) E403X LBD.

3.4.5 - Mutant LBDs are not able to interact with coactivator peptides even in the presence of T3

T3 binding causes conformational changes of the WT LBD, which adopts an active state that enables coactivator interaction and activates the transcription of target genes. The ability of T3 to recruit coactivator peptides was studied by FA in order to address the effect of T3 on the mutant LBDs. The affinity of the interaction between the WT and mutant LBDs with the GRIP1 peptide in the presence of T3 was obtained in order to examine the change in binding affinity caused by T3.

The results showed that the binding of T3 to the WT LBD causes a notable increase in the affinity of the interaction between the WT LBD and the coactivator peptide,

indicated by the K_d value which became 67 nM (**Figure 3.9a**). This suggests that GRIP1 coactivator peptide binding is T3-dependent in the case of the WT LBD indicating that coactivator binding requires the WT LBD to be in the active or bound state.

On the other hand, as it is shown in **Figure 3.9**, the binding of T3 to the mutant LBDs barely change the affinity of the interaction between the mutant LBDs and the coactivator peptides, which suggests that coactivator proteins are not able to bind the coactivator binding site of the mutant LBDs in the presence of T3 either.

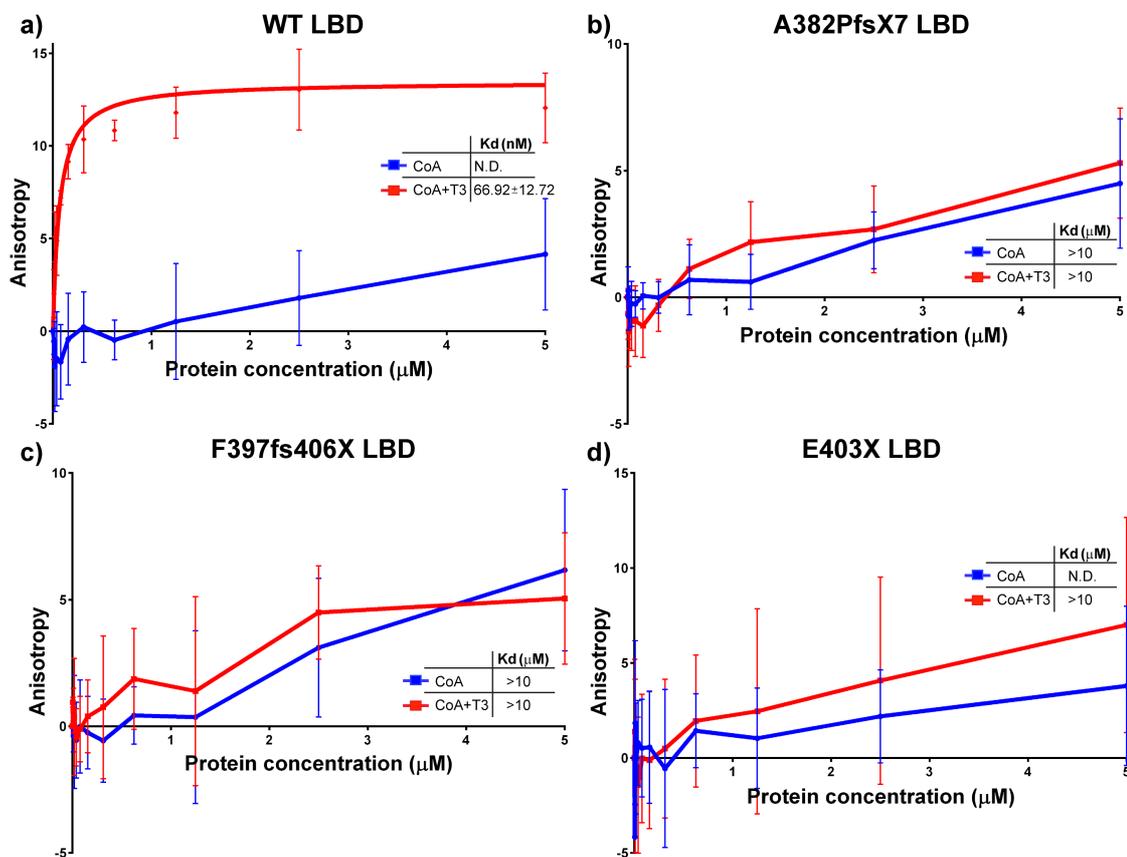


Figure 3.9: Binding assays of GRIP1 peptide to TR α LBDs. Increasing concentration of the mixture LBD:T3 (ratio 1:10) was incubated with 5 nM BODIPY-labeled coactivator peptide. FA values are the mean \pm SEM of measurements obtained from triplicate experiments. a) WT LBD, b) A382PfsX7 LBD, c) F397fs406X LBD, and d) E403X LBD.

3.5 - Circular dichroism study of protein folding and peptide binding interactions

CD experiments were set up in order to investigate the ability of the coactivator and corepressor peptides and T3 to improve the thermal stability of the WT and the mutant LBDs.

CD is a spectroscopic technique for obtaining information about the folding of proteins and is also used to follow their unfolding as a function of temperature. When proteins are folded they often have highly asymmetric secondary structural elements, such as α -helices and β -sheets, which have characteristic CD spectra. When proteins unfold they lose these highly ordered structures and the CD spectra change. Therefore, analysis of CD spectra is useful to characterize the secondary structure of a protein of interest and the changes in CD spectrum as a function of temperature are used to determine the melting temperature. The melting temperature (T_m) is the midpoint of the unfolding transition and can be used as an indicative of the thermal stability of a protein. A change in the T_m of unfolding after adding specific ligands or peptides to the protein sample reveals whether the ligand or the peptide under study stabilizes the protein.

WT LBD and mutant LBDs were expressed using the constructs designed in section 3.2.2 and purified following the protocol described in section 2.4. An initial spectrum of each protein alone (apo-LBDs) was taken to determine the main secondary elements that form the protein (**Figure 3.10**). The CD spectra at 20 °C of purified LBDs were typical of folded proteins containing elements of α -helical secondary structure, characterized by two negative peaks at 208 and 222 nm (no major trough at approximately 200 nm that would indicate random coil) (**Figure 3.10**).

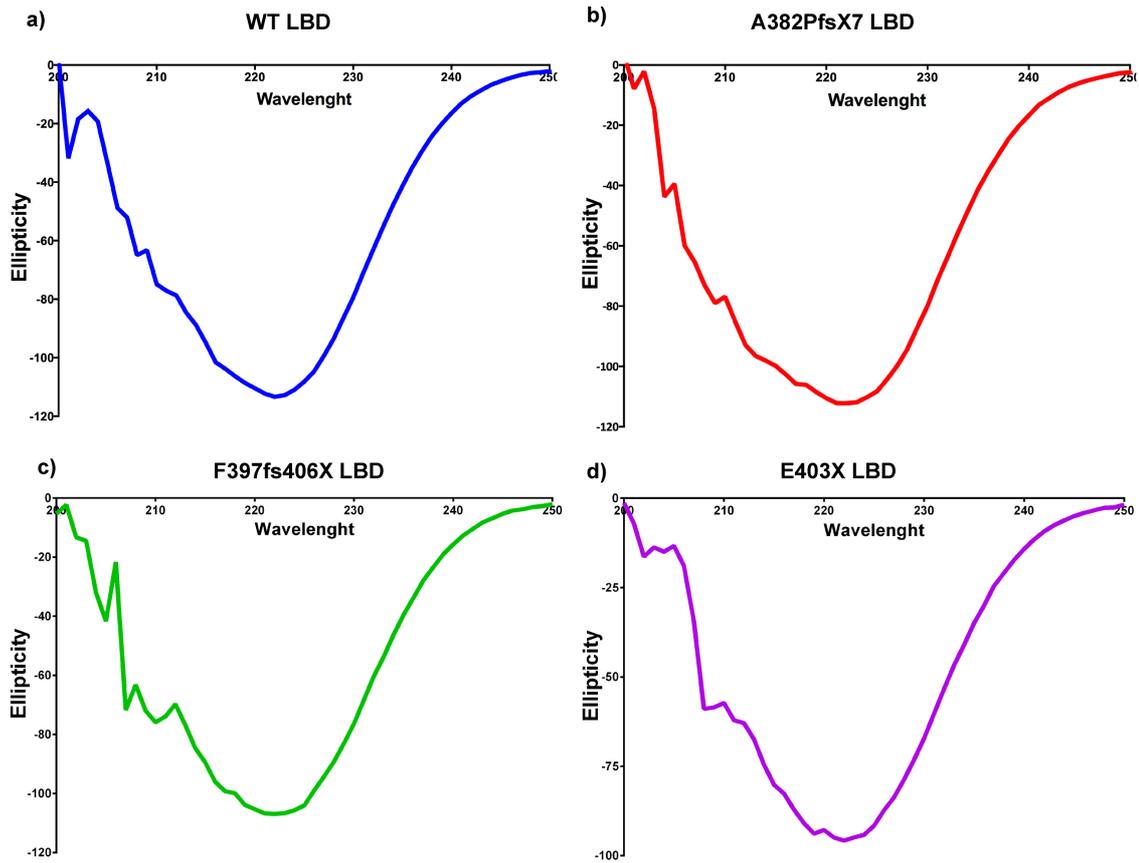


Figure 3.10: CD spectra of the WT and mutant LBDs. Data represent the average ellipticity values obtained from triplicate experiments. a) WT LBD, b) A382PfsX7 LBD, c) F397fs406X LBD, and d) E403X LBD.

3.5.1 - T3 increases the thermal stability of the mutant TR α LBDs

FA results shown in the previous section indicated that T3 presence does not lead to coactivator recruitment in the mutant LBDs, which means that the mutant LBDs do not respond to T3. Subsequently, the question that arose is whether the mutant LBDs are impaired in binding to T3. In order to investigate the ability of the mutant receptors to bind T3, thermal melt CD spectroscopy was performed using the WT and the mutant LBDs in the presence of T3.

The results in **Figure 3.11** show the T_m values of the unfolding transition of the different TR α LBDs alone and in the presence of T3. The presence of T3 stabilized all the TR α LBDs, including the mutants, as indicated by the increase in melting

temperature after adding T3. This suggests that all the TR α LBDs are able to bind T3, including the mutant ones (**Figure 3.11**).

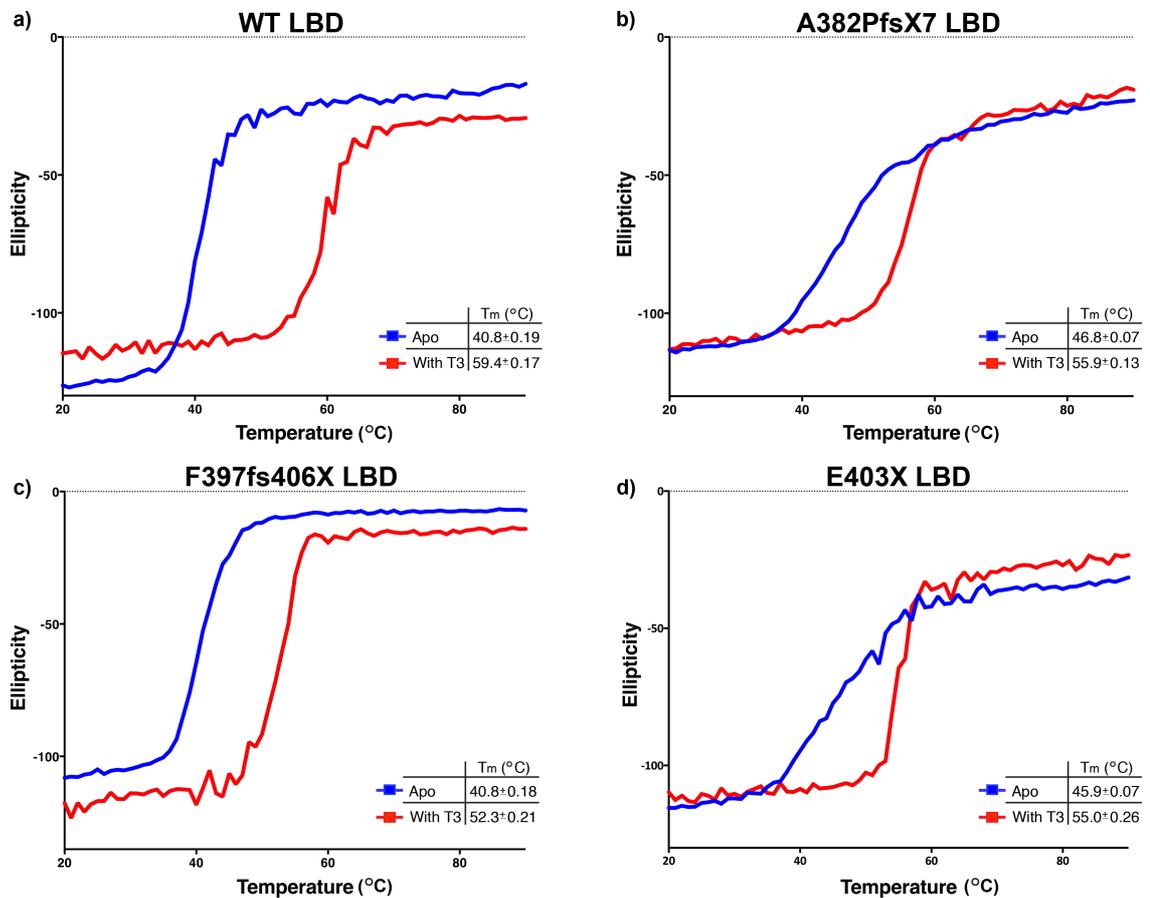


Figure 3.11: Temperature induced denaturation experiments of the TR α LBDs as measured by CD at 222 nm. Apo form of the proteins (blue) and TR α LBDs in solution with T3 (red): a) WT LBD b) A382PfsX7 LBD, c) F397fs406X LBD, and d) E403X LBD.

The stabilizing effect of T3 on the mutant LBDs is quite significant, as shown by the substantial increase of the T_m in approximately 10 degrees (**Figure 3.11**). This large stabilization effect indicates that T3 causes several conformational changes in the mutant LBDs. As a result, the liganded mutant LBDs are more stable and, in the case of the LBD mutations A382PfsX7 and F397fs406X, this leads to a stronger interaction with SMRT corepressor (**Figure 3.8**).

However, the greatest stabilization effect is promoted by the binding of T3 to the WT LBD with an increase in the T_m of almost 20 °C indicating a WT-preference of T3. Since the WT LBD has a functional and complete H12, it also maintains the ability to

further interact with T3 by some residues belonging to this helix such as Phe 401. This interaction further stabilizes the protein as a result of T3 binding.

3.5.2 - Study of the ability of mutant LBDs to bind coregulator peptides

FA results described previously showed that mutant LBDs interact tightly with corepressor peptides even in the presence of T3. In contrast, mutant LBDs are not able to recruit coactivator peptides in response to T3. CD experiments were performed using coactivator and corepressor peptides in order to confirm WT and mutant LBDs properties to interact with the coregulator peptides. Chemically modified peptides were also tested for their ability to interact with the WT and the mutant LBDs.

These chemical modifications were added in order to improve the potential abilities of the peptides as therapeutic agents. Firstly, a site-specific chemical hydrocarbon arm, called a staple, was introduced to reinforce the native α -helical structure of the peptides. Forcing a helical structure of a peptide improves its stability and reduces the entropic penalty of having to fold before binding any substrate. Thus, this modification confers the peptide an improved affinity to its target. Secondly, a TAT (trans-activator of transcription) sequence (YGRKKRRQRRR) was attached to the N-terminus of the SMRT corepressor peptide in order to increase the cell permeability. The TAT peptide is derived from the transcriptional activator of human immunodeficiency virus (HIV) and belongs to the cell-penetrating peptides (CPPs) superfamily (Green & Loewenstein 1988; Frankel & Pabo 1988; Mann & Frankel 1991). CPPs are short peptides that facilitate cellular intake of various molecules (peptides, proteins, antisense oligonucleotides, large iron beads and liposomes) in a receptor-independent and concentration-dependent fashion (Schwarze et al. 1999; Lindsay 2002; Cao et al. 2002; Wadia et al. 2004). Further investigation of the therapeutic potency of the stapled peptides will be described in Chapter 4.

i) Mutant LBDs interact with corepressors but do not interact with coactivators

The thermal stability of the apo-LBDs, and T3 liganded-LBDs was tested on the addition of native SMRT, stapled SMRT, TAT SMRT, and GRIP1 coactivator as a function of the temperature by measuring the ellipticity at 222 nm.

The presence of both SMRT corepressor peptides, the native and the stapled ones, stabilized all the TR α LBDs, as determined by the increase in the melting temperature after adding them. This suggests that SMRT corepressor binds to all the TR α LBDs. On the contrary, both TAT peptides (the native and the stapled ones) did not stabilize any of the LBDs. This absence of stabilization is probably due to the highly positive charged sequence attached to the C-terminal of the SMRT peptides that interferes in the interaction between the corepressor binding motif and the hydrophobic LBD surface, preventing the binding (**Figure 3.12, Table 3.2**).

In accordance with the FA results that showed negligible interaction between coactivators and the mutant LBDs, GRIP1 peptide did not stabilize the mutant LBDs (**Figure 3.12, Table 3.2**).

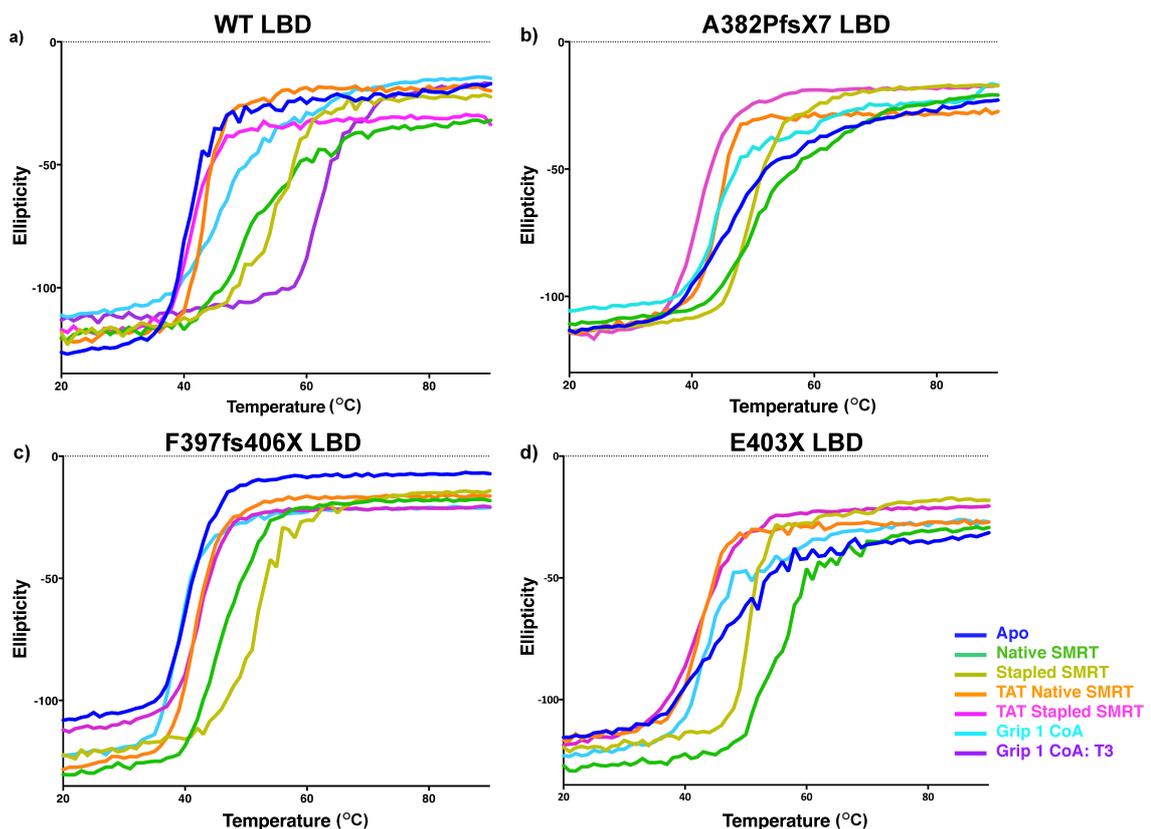


Figure 3.12: Temperature induced denaturation experiments of the TR α LBDs as measured by CD at 222 nm. Apo- (dark blue), native SMRT (green), stapled SMRT (gold yellow), TAT peptides (orange and pink), and GRIP1 (light blue) in solution with TR α LBDs: a) WT LBD b) A382fsX7 LBD, c) F397fs406X LBD, and d) E403X LBD. In the case of the WT it is also indicated the thermal stabilization of GRIP1 coactivator in the presence of T3 (violet).

T_m (°C)						
	Apo	Native SMRT	Stapled SMRT	TAT native SMRT	TAT stapled SMRT	GRIP1
WT	40.8 ± 0.11	51.6 ± 0.28	54.6 ± 0.37	43.4 ± 0.05	41.6 ± 0.11	47.4 ± 0.17 62.6 ± 0.21
A382PfsX7	46.8 ± 0.4	52.3 ± 0.28	50.1 ± 0.24	43.8 ± 0.17	41.4 ± 0.10	45.3 ± 0.33
F397fs406X	40.8 ± 0.08	46.4 ± 0.31	51.5 ± 0.31	41.8 ± 0.15	42.3 ± 0.18	39.3 ± 0.09
E403X	45.9 ± 0.20	55.7 ± 0.38	50.7 ± 0.11	42.1 ± 0.07	42.4 ± 0.40	44.4 ± 0.23

Table 3.2: Melting temperature values of the different LBDs alone or with SMRT peptides, TAT peptides or GRIP1 peptide (measured in degree Celsius). Melting temperature values are the mean ± SEM of measurements obtained from triplicate experiments. In the case of the WT LBD, the melting temperature of the receptor in the presence of both GRIP1 coactivator and T3 is also indicated.

3.5.3 - Liganded mutant LBDs are able to interact with corepressor peptides

CD experiments were performed to analyze the thermal stability of the WT and mutant LBDs in the presence of T3 with SMRT or GRIP1 peptides in order to investigate the formation of ternary complexes. As illustrated in **Figure 3.13**, SMRT corepressor peptide is not able to bind WT LBD when the receptor is bound to T3 since there is no significant change in the T_m of the WT LBD:T3:CoR in comparison with the T_m of the WT LBD:T3. This means that the active conformation of H12 does not allow the corepressor to bind. In contrast, GRIP1 coactivator peptide is able to bind to the WT LBD even in the absence of T3, which means that coactivators are able to bind to the inactive state of the LBD, with H12 in the disordered state. This result supports the “dynamic stabilization” model that describes H12 in the apo-state as a rather mobile structure not fixed in any single position (Nagy 2004). The stabilizing effect of the GRIP1 binding is considerable higher when the WT protein is bound to T3 and the WT LBD:T3:CoA ternary complex forms. This interaction between the coactivator and the active state of the WT LBD causes a T_m increment of 21.8 °C (**Figure 3.13**, **Table 3.2**).

On the other hand, CD results suggest that the mutant LBDs are able to bind T3 and the corepressor at the same time forming a ternary complex that showed the highest level of thermal stabilization. Binding of coactivator to the mutant LBDs was really impaired even in the presence of T3, supporting the FA data. In **Figure 3.13** is shown that the thermal stabilization carried out by the coactivator in the presence of T3 is the same as the one carried out by T3 alone, demonstrating that the observed stabilization effect is due to T3. In the case of the E403X LBD, there is some stabilization upon adding coactivator to the E403X LBD:T3. However, corepressor binding showed more stabilizing effect and, since both proteins compete for binding to the same hydrophobic surface of the receptor, the corepressor binding would be favored.

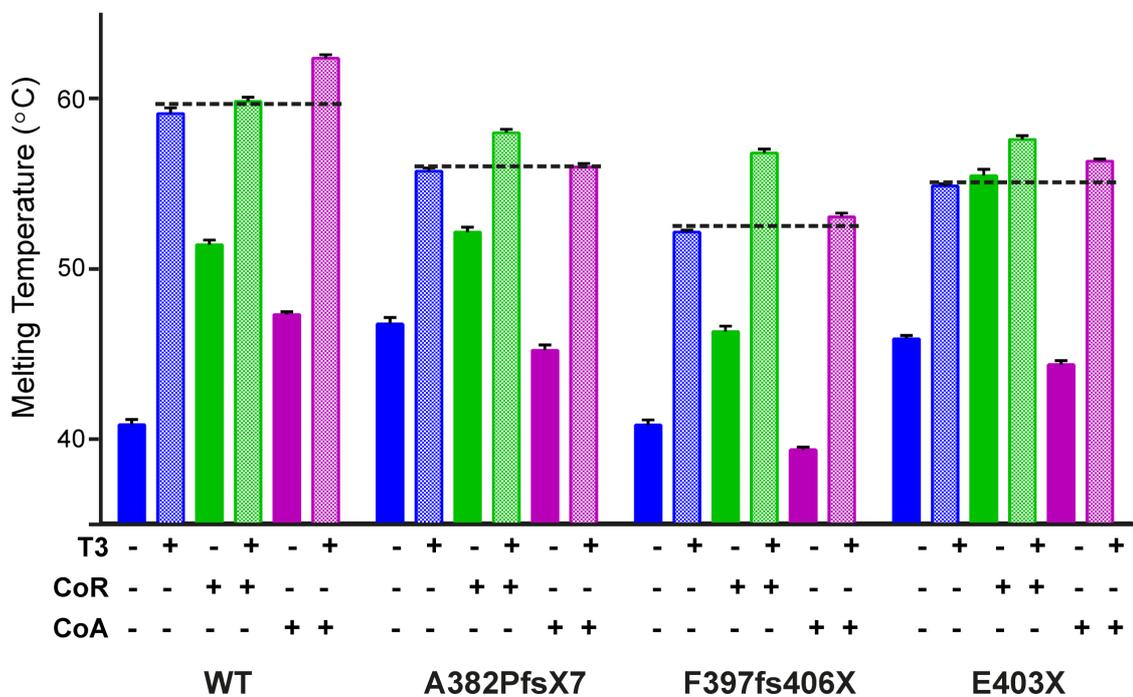


Figure 3.13: Melting temperature values of the different LBDs alone (blue) or in solution with T3 (dotted blue), SMRT (green) or both corepressor and T3 (dotted green), GRIP1 coactivator (pink) or both coactivator and T3 (dotted pink) as measured by CD at 222 nm and presented in degree Celsius.

3.6 - Discussion

In this study the WT and three mutant (A382PfsX7, F397fs406X and E403X) TR α LBDs were successfully expressed, purified, and characterized in their ability to interact with coregulator proteins. Biochemical and spectroscopic techniques such as FA and CD were used to investigate the affinity of the interaction between the WT and mutant LBDs with coactivator and corepressor peptides in response to T3. The results obtained from FA and CD were reassuringly in agreement and reveal some insights in H12 behavior in the active and the inactive state.

i) Corepressor recruitment and binding by mutant LBDs

Data from the SMRT corepressor peptide binding assay and CD suggest that mutant LBDs bind more strongly to SMRT corepressor than WT LBD, with the exception of the E403X mutant LBD. In the inactive conformation, H12, as a dynamic and mobile structure, could occupy different positions in solution, and one of these possibilities is a position where the corepressor/coactivator binding surface will be occluded or partially occluded. Therefore, the complete loss of H12, as in the LBD mutations A382PfsX7 and the F397fs406X, would expose the hydrophobic corepressor-binding site on the receptor surface facilitating their recruitment and resulting in a higher affinity between the corepressor and the mutant receptors. The additional helical turn in corepressor motif requires a larger interaction interface with the LBD surface than coactivator motif, so corepressor binding overlaps the coactivator binding surface and extends underneath H12 (Xu et al. 2002). Therefore, the exposition of the corepressor-binding site should be beneficial for the preferential binding of corepressor to these mutants.

FA results also indicate that T3 binding to the LBD mutations A382PfsX7 and F397fs406X modestly increased the affinity between corepressors and the mutant LBDs suggesting that H12 is not the only feature responsible for corepressor binding. These results are consistent with previous reports that showed that there is a moderate increase in the affinity between the TR β LBD and corepressors after the deletion of H12. The authors also demonstrated that when H12 is deleted corepressor binding was stimulated by 3,5,3'-triiodothyroacetic acid (Triac), an agonist similar to T3, supporting the FA results reported here (Marimuthu et al. 2002).

In the case of the E403X LBD mutant, the remaining H12 residues could interfere with corepressor recruitment. The last amino acids (Leu 400, Phe 401 and Leu 402) are hydrophobic residues which are likely looking for hydrophobic interactions within the protein to avoid solvent contact, for example with the hydrophobic groove of the corepressor-binding site (H3 and H5 groove). Since Leu 401 residue interacts with T3, forming part of the T3 hydrophobic pocket, it is also probable that this end of the mutant LBD interacts closer with T3 when the hormone is on the cavity. As a result, the C-terminal of the E403X LBD mutation which would interact with T3 and H3 residues, partly occupying the corepressor-binding site, would need to be displaced in order for the corepressor to fit on the LBD surface. So, this interference of the C-terminal residues with the corepressor-binding surface could be the reason for the weaker binding of the peptide to this mutant.

On the other hand, the affinity of the interaction is similar for all the LBDs studied, with K_d within the nanomolar range (from 90 to 210 nM). This suggests that the constitutive binding of corepressor to the mutant LBDs is not a consequence of the higher affinity showed by the mutant LBDs to the corepressor since there is no significant increase in the binding affinity, but a consequence of the mutant LBDs failure to bind coactivators in response to T3.

ii) Response to T3 and coactivator recruitment by mutant LBDs

All the mutants LBDs are able to bind T3, however the binding of T3 does not cause either the displacement of the SMRT corepressor from the mutant LBD surface or the recruitment of coactivators. This is probably due to the lack of H12 function in the mutant receptors, since it has been demonstrated that the active position of H12 and, especially, the Glu 403 residue are absolutely necessary for coactivator binding (Darimont et al. 1998).

Destabilization of the active conformation of H12 by mutations in the PPAR γ LBD was reported to be associated with severe insulin resistance, diabetes mellitus and hypertension (Barroso et al. 1999). Two independent mutations, P467L in the N-terminal and V290M in H3, were discovered affecting H12 and the coactivator-binding

site of the LBD, respectively. A mutation in the beginning of H12 of TR β LBD (P448H) was also found to lead to resistance to the RTH β disease (Chatterjee et al. 1991). All these mutations led to the destabilization of H12 in the active conformation and showed dominant negative activity. Supporting the results presented here, this lack of H12 function also led to an impaired transactivation in response to ligand and a constitutive repression of target genes (Chatterjee et al. 1991; Barroso et al. 1999).

Interestingly, CD results showed that the WT and mutant LBDs have rather low melting temperatures in the absence of T3 suggesting that the LBDs show a great range of dynamic behavior that correlates well with their strong basal repressive activity and justify the absence of apo-LBD crystallographic structures. The melting temperature shows a marked increase on ligand binding supporting the “dynamic stabilization model” that claims that ligand binding stabilizes the LBD, resulting in a more compact and rigid structure.

To summarize, **Figure 3.14** shows the binding affinity results in terms of K_d. It is clear that the mutant LBDs not only failed to release corepressor, but also failed to recruit coactivator in response to T3. This suggests that the active position of H12 and Glu 403 are essential for coactivator binding (**Figure 3.14**). These results clarify previously reported data that showed a clear absence of T3 response by cells transfected with the different mutant TR α (Bochukova et al. 2012; van Mullem et al. 2012; Moran et al. 2013).

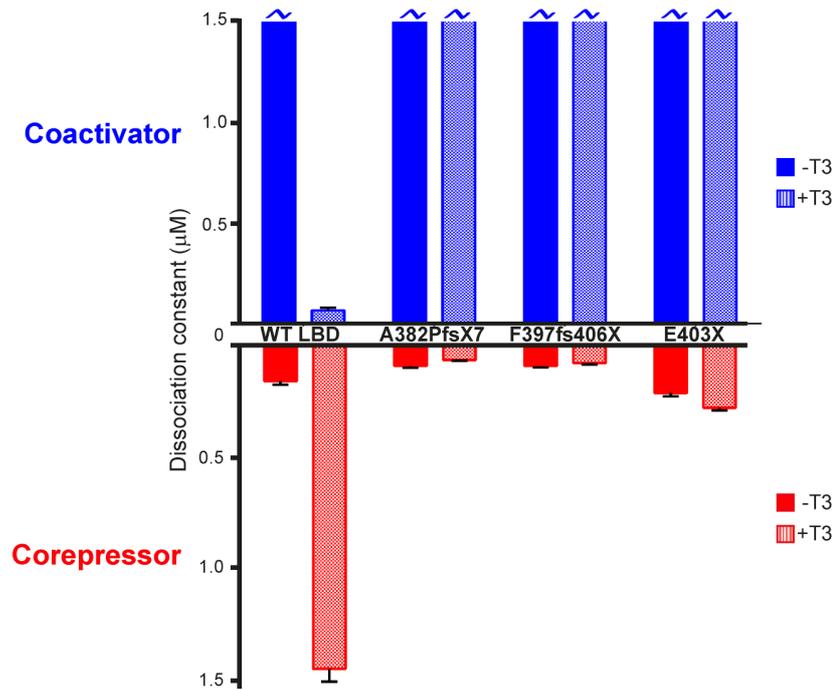


Figure 3.14: Summary of coactivator and corepressor binding assay results showing the K_d in the presence and absence of T3. Binding affinity of the different LBDs for coactivator in the absence of T3 (blue) and in the presence of T3 (dotted blue); affinity of the LBDs for corepressor in the absence of T3 (red) and in the presence of T3 (dotted red).

Chapter 4 - Action of stapled peptides and T3 analogues in corepressor and coactivator interaction

4.1 - Introduction

Mutant LBDs inhibit the WT transcriptional activity by interacting constitutively with corepressor complexes. This leads to a constitutive repression of T3 target gene transcription even in the presence of T3. WT LBD is expressed at the same level as mutant receptors and, as it is present in the cell, responds to T3 but its transcription activation of T3 target genes is not enough to trigger the proper T3 signal in the cell.

Some of the patients have been treated with T4 after the identification of several symptoms of hypothyroidism. Overall, 11 patients have been T4 treated either from early childhood or in adult life (Bochukova et al. 2012; Moran et al. 2013; Moran et al. 2014; van Mullem et al. 2012; Espiard et al. 2015; Demir et al. 2016; van Gucht et al. 2016). In the majority of cases, the treatment resulted in decreased TSH levels, and a normalization of FT4 and rT3 levels, together with a marked increase of SHBG levels (Sex Hormone Binding Globulin, a hepatic marker of thyroid hormone action), suggesting that TR β remains active and responsive in the hypothalamic-pituitary axis and liver to the increased levels of thyroid hormones following treatment with T4. T4 therapy also raised the basal metabolic rate of the patients and improved their constipation. However, the heart rate, the blood pressure and the growth rate remained abnormally low (Moran et al. 2013; van Mullem et al. 2013). In general, initiation of T4 therapy at a young age had beneficial developmental effects in childhood cases; however, patients with frame-shift and nonsense mutations, who exhibit the most severe clinical phenotype, did not ameliorate their growth retardation, and cardiac and renal problems (Moran et al. 2013; van Mullem et al. 2012; Demir et al. 2016). Therefore, even though increasing the levels of T3 as a treatment of the RTH α disease shows some relief in specific symptoms, it is not a suitable therapy and has undesirable secondary effects due to the cross-activation of TR β .

Preventing the interaction between mutant LBDs and corepressor proteins would eliminate the dominant activity exerted by these receptors, allowing the WT LBD, also present in the cells, to respond to T3. As suggested by the biophysical data showed in Chapter 3 and confirmed by the structural data discussed in Chapter 5, mutant LBDs are able to recruit and bind T3 in their hydrophobic cavity in the same position as in the WT LBD. This knowledge provided an opportunity to design several T3 analogues trying to promote corepressor release or coactivator recruitment.

In this project, two different strategies have been investigated in order to promote the dissociation of corepressor proteins from mutant LBDs. T3 analogues and stapled peptides have been designed and tested to use them as agents to displace SMRT corepressors from the mutant LBDs. In a third strategy, another type of T3 analogue was designed in order to promote coactivator recruitment.

The main aim of this chapter is to study the ability of T3 analogues and stapled peptides to block the interaction between mutant LBDs and corepressors and/or recruit coactivators.

4.2 - Receptor binding and corepressor displacement properties of T3 analogues

Two types of T3 analogues were designed based on available liganded TR α LBD crystal structures with two different purposes. Both types of analogues were designed to resemble T3 at least in terms of overall molecular shape in order to fit buried in the binding cavity of the LBD, so they only differ in the extension length and chemical structure.

The first type of T3 analogue (**Figure 4.1a**) was designed and tested as a potential therapeutic agent able to remove SMRT corepressor from the corepressor-binding surface of the mutant LBDs. A library of extensions were added to the 4'OH position of the outer thyroxine ring of T3 in order to modify the corepressor-binding site structure of the LBDs, decreasing the affinity of the interaction between the corepressor and mutant LBDs and thus, releasing the corepressor from the mutant LBDs.

The second type of T3 analogue (**Figure 4.1b**) was designed to recruit coactivator proteins by mimicking the side chain of Glu 403 of the LBD. This amino acid belongs to H12 and is essential to recruit coactivator proteins in response to T3. The conformational changes that take place after ligand binding place H12 in an active conformation. The active conformation of H12 places the Glu 403 in a new position as well, allowing it to interact with coactivator proteins. The interaction between the Glu 403 and the backbone of the NR interaction domain of the coactivator forms a charge clamp that is critical for the interaction between these two proteins (LBD and coactivators) (Nolte et al. 1998; Darimont et al. 1998). Therefore, these second group of analogues were modified to have different length hydrocarbon extensions coupled to the 4'OH and end with a carboxylic acid group. Different extensions were tested with the aim of finding out the most suitable one that places the carboxylic acid of the T3 analogue in the proper position to interact with coactivator peptides. All the compounds used in this study were kindly provided by Professor Nick Tomkinson (University of Strathclyde).

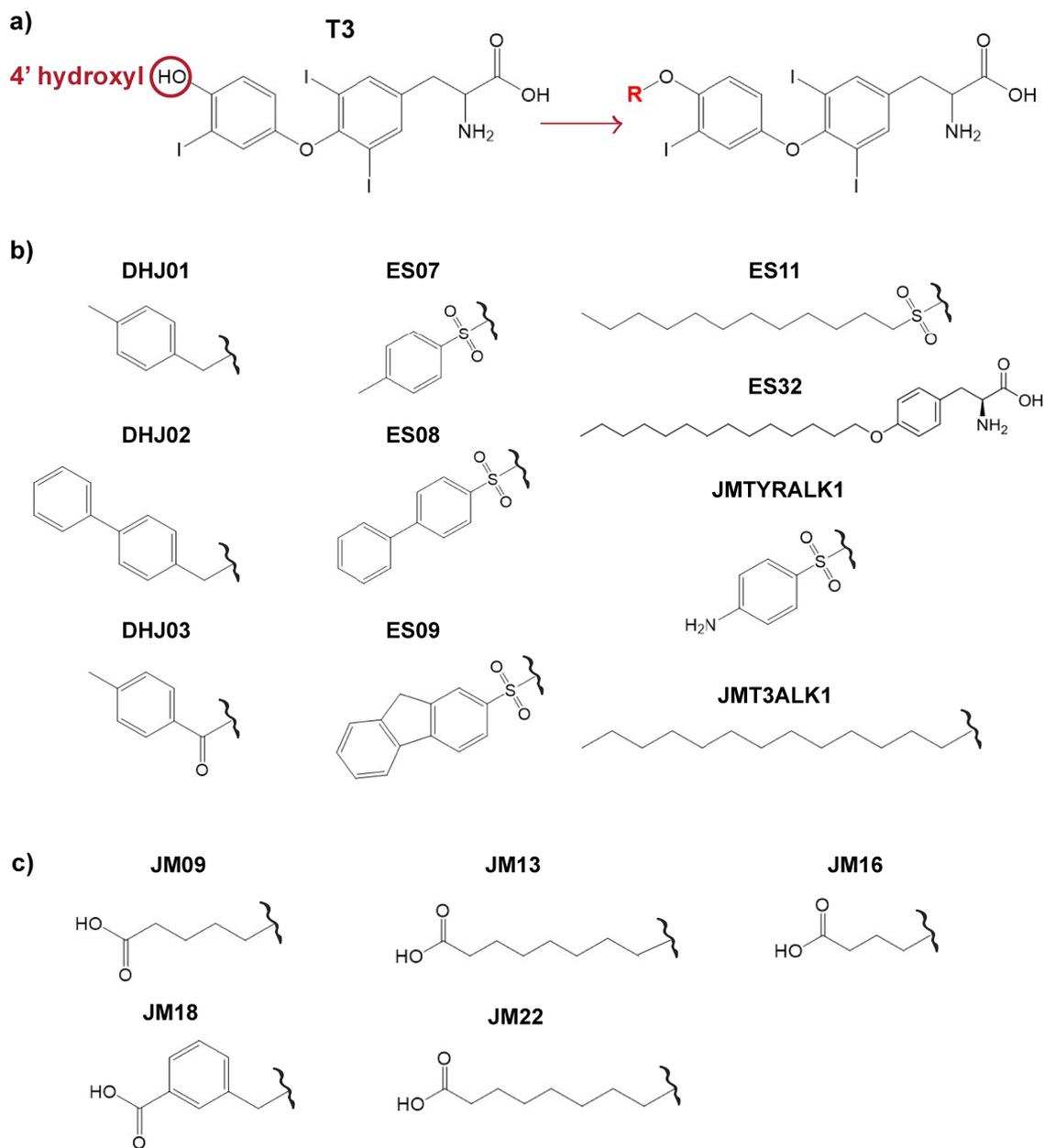


Figure 4.1: T3 analogues. a) T3 formula highlighting the position where the modifications were introduced. b) First type of T3 analogue synthesised to displace corepressor proteins from the corepressor-binding surface of mutant LBDs. c) Second type of T3 analogue that includes those thyroidmimetics created to recruit coactivators. Different length extensions and chemical functionalities were designed in order to test their effectiveness in corepressor dissociation and/or coactivator recruitment.

4.2.1 - Most of T3 analogues increase the thermal stability of the WT and mutant LBDs

The strategy to characterize the interaction between the WT and mutant LBDs with the different T3 analogues consisted of firstly, CD experiments designed to investigate whether the T3 analogues are able to stabilize to the different LBDs, and secondly, FA assays performed to find out the effect of the interaction in corepressor dissociation or coactivator recruitment.

CD is a spectroscopic technique sensitive to the conformation of molecules, specifically to asymmetric or chiral molecules such as DMSO or nitrate. T3 analogues were perfectly soluble in DMSO, however DMSO containing samples would interfere in the CD measurements because of the solvent absorption. Therefore T3 analogues were dissolved in 100 % ethanol at 1 mM stock.

The binding of the T3 analogues was investigated by CD. As it is shown in **Figure 4.2**, the melting temperature of the proteins increased by adding T3 or T3 analogues suggesting that all the T3 analogues were able to bind to all the mutant LBDs. The successful stabilization effect of the analogues was expected as they were designed to fit in the hydrophobic binding cavity of the receptor.

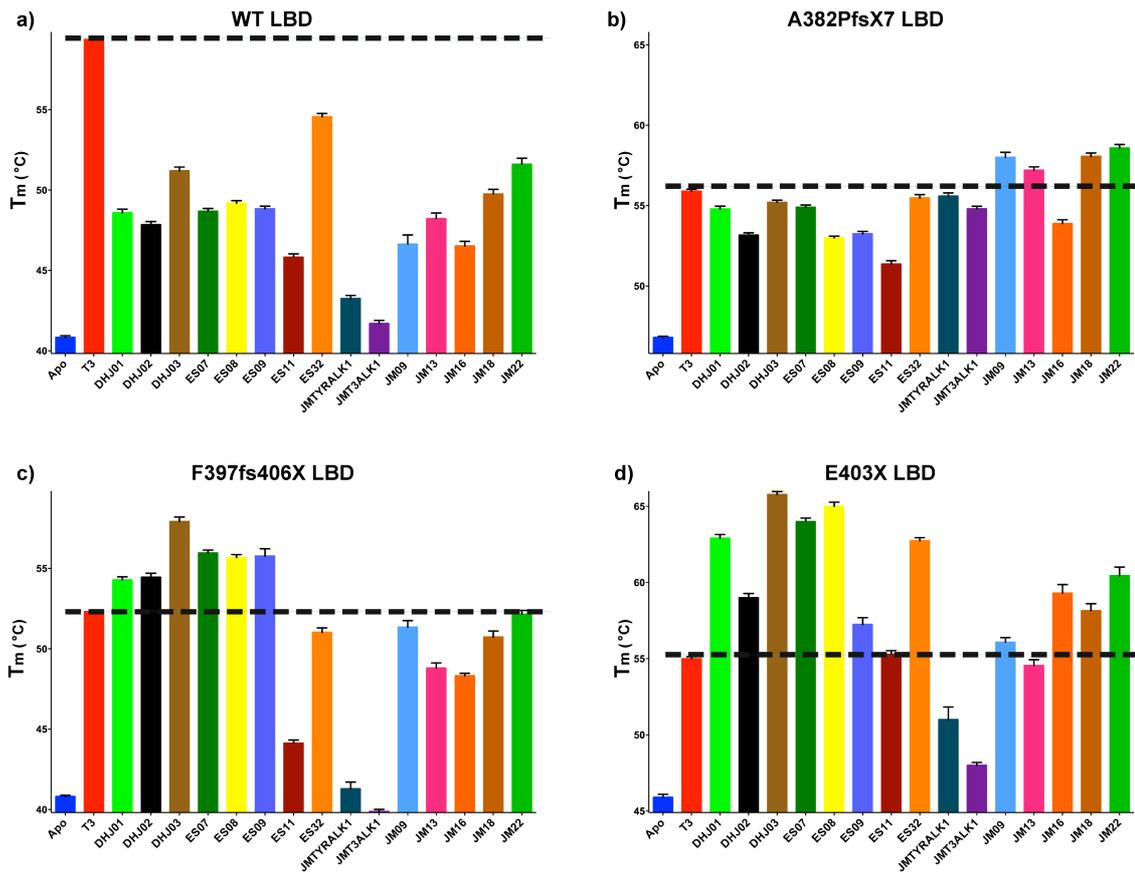


Figure 4.2: Temperature induced denaturation experiments of the TR α LBDs as measured by CD at 222 nm. T_m values are the mean \pm SEM from triplicate experiments. Apo- and ligand-bound TR α LBDs: a) WT, b) A382PfsX7, c) F397fs406X and d) E403X.

In the case of the WT LBD, all the T3 analogues increased the thermal stability of the receptor to a certain extent, especially ES32 which showed the greatest stabilizing effect. However, none of them matched the stabilizing effect of T3 binding to the WT LBD (almost 20 °C) (**Figure 4.2**).

Regarding the A382PfsX7 LBD, all the T3 analogues caused an increase in the thermal stability of the receptor, especially JM09, JM16, JM18 and JM22 which showed more stabilizing effect than T3. Similarly, all the analogues were able to bind E403X LBD, most of them causing a higher stabilization effect than T3 except for ES11, JMTYRALK1, JMT3ALK1, JM09 and JM13. In the case of F397fs406X LBD, the results showed that JMTYRALK1 and JMT3ALK1 did not stabilize the protein, since their presence in the assay did not change its thermal stability. The binding of DHJ01, DHJ02, DHJ03, ES07, ES08 and ES09 to the F397fs406X LBD caused higher thermal stabilization than T3. The high level of thermal stabilization accomplished by some of

the analogues reveals different, and possibly more, interactions between the mutant LBDs and the analogues than between the mutant LBDs and T3 (**Figure 4.2**). Therefore, the results suggest that the extensions introduced in some analogues are accommodated within the protein and further interact with specific residues in such a favorable way that increases the thermal stability of the protein.

Overall, the T3 analogues that exhibited the least stabilization effect were JMT3ALK1 and JMTYRALK1. Both compounds did not produce any significant effect in the thermal stability of F397fs406X, which suggests that they are not able to bind this particular mutant LBD. These results suggest that the second or outer ring of the hormone is essential for the compound to fit in the hydrophobic cavity and the lack of it as in the JMTYRALK1 prevents the compound from binding. In the case of the JMT3ALK1, it is possible that the hydrocarbon elongation is too long to fit in the binding cavity or to access to it as well as the others T3 analogues. However, A382PfsX7 LBD is able to recognize and bind these compounds perfectly well; A382PfsX7 LBD is the shortest mutant and the lack of 22 residues might make it more permissive to bind different ligands than the others.

4.2.2 - T3 analogues show moderate potency to release corepressor peptides

FA assays were performed in order to study the effect of the T3 analogues in corepressor binding to the mutant LBDs with the aim of finding out if any of the compounds are able to displace corepressor from the mutant LBDs.

After CD analysis, a freeze-drying method was followed as described in 2.7 of materials and methods to remove the ethanol from the T3 analogue vials. As discussed before, T3 analogues are more soluble in DMSO than in ethanol, so 10 mM concentration stocks were made in DMSO and a 1 to 10 molar ratio (LBD:T3 analogue) was achieved to perform the FA experiments. Using more than 5 % of DMSO in the FA assays had some effects in protein and peptide stability, so in order to keep them soluble, a more concentrated T3 analogues stock was required.

FA assays were performed using FITC-labeled SMRT corepressor peptide at 5 nM and serially diluted complex of the LBDs with the different T3 analogues (1 to 10 molar ratio) from 10 μ M of protein and 100 μ M compound. The FA results were plotted to obtain corepressor binding curves from which an apparent K_d was calculated for each T3 analogue and each mutant LBD (**Figure 4.3**).

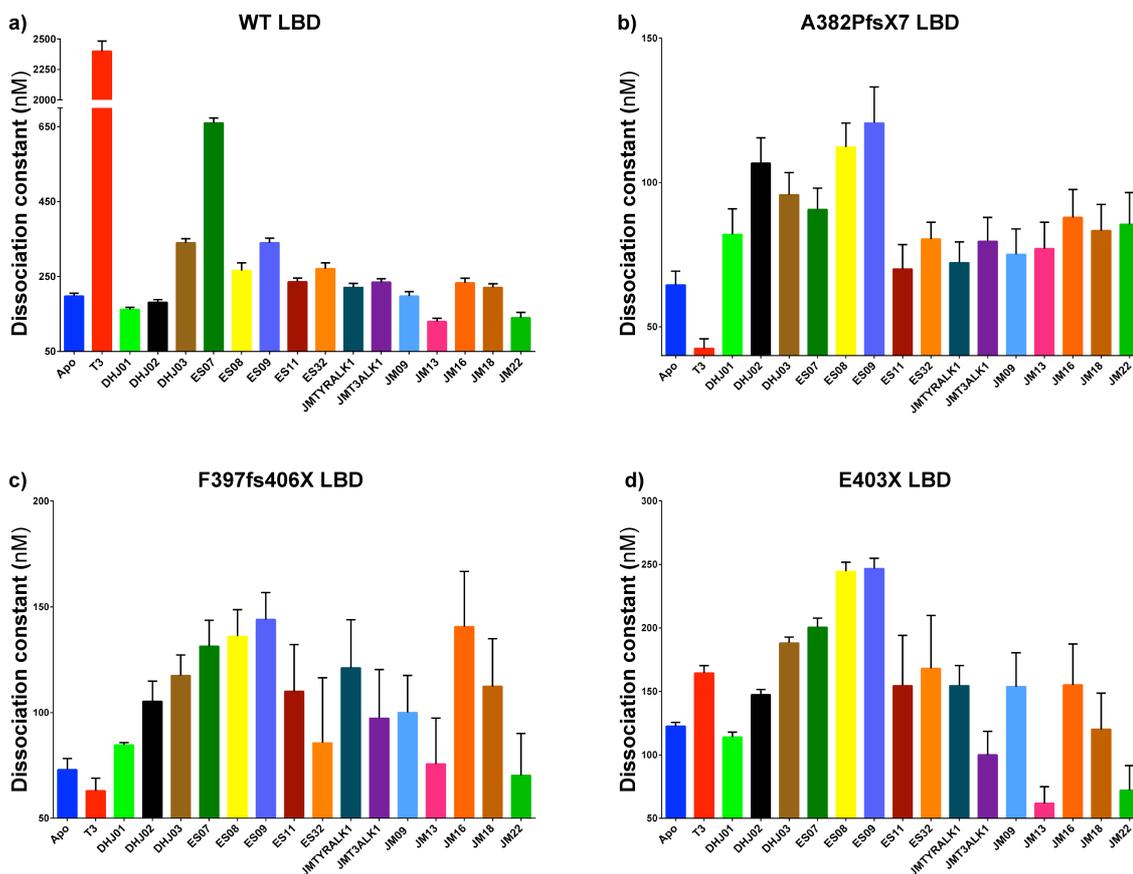


Figure 4.3: Apparent dissociation constant of T3 analogues. The apparent dissociation constant was determined by FA and indicates the T3 analogues ability to displace the labeled SMRT peptide from the different LBDs. Increasing concentration of the mixture LBD:T3 analogue (ratio 1:10) was incubated with 5 nM FITC-SMRT peptide. K_d values are the mean \pm SEM of measurements obtained from triplicate experiments. a) WT LBD, b) A382PfsX7 LBD, c) F397fs406X LBD, and d) E403X LBD.

As shown by the results, all these compounds had relatively low potency in corepressor displacement that limits their usefulness, in spite of the significant LBD stabilization that most of them showed by CD. A382PfsX7 and F397fs406X mutant LBDs respond relatively well to some of the compounds by releasing corepressor peptides, especially compared to T3. Since these mutant LBDs completely lack H12, they absolutely fail to displace corepressor and in fact, the stabilization effect of T3 upon binding increases the

affinity of the interaction between A382PfsX7 and F397fs406X mutant LBDs and corepressor peptides as discussed in Chapter 3. Therefore, any extension somewhat able to interfere with the corepressor-binding surface seems to cause a moderate effect, especially the extensions of ES08 and ES09 analogues. ES08 and ES09 compounds also displayed moderately potency for reducing SMRT peptide binding affinity to the E403X mutant LBD compared to T3. However, the other T3 analogues tested showed little or no effect on the interaction between the mutant LBDs and the corepressor peptide (**Figure 4.3**).

4.3 - Corepressor displacement properties of stapled corepressor peptides

Stapled peptides that specifically bind the TR α LBDs offer an alternative approach to T3 analogues for the modulation of the receptor interaction to coregulator proteins. Stapled peptides are synthetic peptides locked into their bioactive α -helical fold through the site-specific introduction of a chemical arm, a hydrocarbon staple. Stapling can greatly improve the pharmacologic activity of peptides, increasing their target affinity, and their levels of cell penetration (Walensky et al. 2004). In addition to reinforcing the biologically active secondary structure of the native SMRT peptide, helix stabilization buries the amide backbone in the interior of the helix core, protecting the peptide from proteolysis and making them poor substrates for enzymatic hydrolysis (Schafmeister et al. 2000).

In this project, stapled peptides were designed based on the nuclear receptor interaction domain 1 (NID 1) of SMRT corepressor protein to compete with native SMRT corepressor complex in the interaction with LBD proteins. Stapled peptides were synthesized by Dr. Naomi Robertson (Department of Chemistry, University of Leicester) and tested in their ability to displace native SMRT corepressor proteins from the mutant LBDs.

First of all, CD experiments were performed in order to find out the effect of stapled peptides in the stability of the WT and mutant LBDs. The results shown in Chapter 3 section 3.5.3 indicated that stapled peptides are able to significantly stabilize all the

LBDs suggesting that they are able to bind to all of them. Secondly, FA experiments were performed using FITC-stapled SMRT corepressor peptide in order to find out the affinity of the interaction between the stapled peptide and the LBDs. Comparing the dissociation constant of the stapled corepressor with the dissociation constant of the native corepressor will provide an indication of the potency of stapled peptides as selective peptide inhibitors that compete at the corepressor-binding site of the LBDs.

4.3.1 - Design of stapled corepressor peptides

The α -helical peptide represents a common structural motif to mediate interaction between signaling proteins. An effective peptide-protein interaction requires the peptide to be in a rigid, well-defined structure in order to make the specific hydrogen bond contacts with the protein and pack in the protein-peptide interface. However, in isolation a short peptide adopts an unstructured, flexible conformation and is less likely to adopt its bioactive secondary structure. To reinforce the α -helical structure and overcome the entropic penalty of folding, conformational constraints can be employed. The constraint allows the peptide to stabilize a more helical conformation. In this work, stapled peptides containing an all-hydrocarbon link (the “staple”) as the conformational constraint were designed based on the native SMRT corepressor peptide.

The hydrocarbon stapling technique allows the incorporation of the constraint at different positions along the peptide sequence. In order to design a functional stapled peptide that maintains the bioactive and physicochemical characteristics of the native peptide, a study of the location where the staple is placed within the peptide is generally required to give the best staple positions. In this case, structural information about the SMRT peptide binding to NRs is available and the details of the binding are known, such as the sequence of the motif that specifically interacts with the LBD as well as its position and orientation (Xu et al. 2002).

The staple constraint is required across the interaction motif or the corepressor peptide (LxxxIxxx[I/L]) between successive turns of the α -helix in order to specifically stabilize the conformation of this motif. The staple constraint also needs to be positioned on the

solvent-exposed side of the helix, so as not to interfere with any residues necessary for the binding to the LBD surface.

The 16-residue native SMRT peptide consisted of 3 α -helical turns. This could be conformationally constrained by an $i, i+4$ staple, an $i, i+7$ staple or two $i, i+4$ staples. Since many peptide-protein interactions require some flexibility of the peptide for binding to its protein target, an $i, i+4$ staple was chosen to constrain the SMRT peptide across one turn of the helix. The position of the staple was determined by carrying out a staple scan which provided a helical wheel projection. The residues Ala 687 and Met 691 were mutated to the unnatural S₅ amino acids and coupled by the hydrocarbon link (**Figure 4.4**).

The solubility of the stapled peptide is also important to produce high micromolar concentration stocks required to perform the CD experiments, the FA assays, and the crystallization trials. Therefore, it was essential to determine and optimize the solubility of the stapled peptides. The first stapled peptide synthesized was completely insoluble in water, so it was dissolved in 100% DMSO prior to stepwise dilution into aqueous buffers, varying the pH and salt concentration. As it is essential for the peptide to be fully dissolved in the assay buffer to be completely active, the insoluble peptide was redesigned to incorporate an extra charged residue, a lysine instead of the methionine 693. This one was completely soluble in water and other aqueous solvents, including assay buffer and tissue culture medium.

Two stapled peptides were synthesized by Dr. Naomi Robertson with the same sequence, but the second one had a FITC fluorophore attached to the N-terminus of the peptide. Unlabeled stapled SMRT corepressor peptides were used for crystallization trials and CD and labeled stapled SMRT were used to perform the FA experiments.

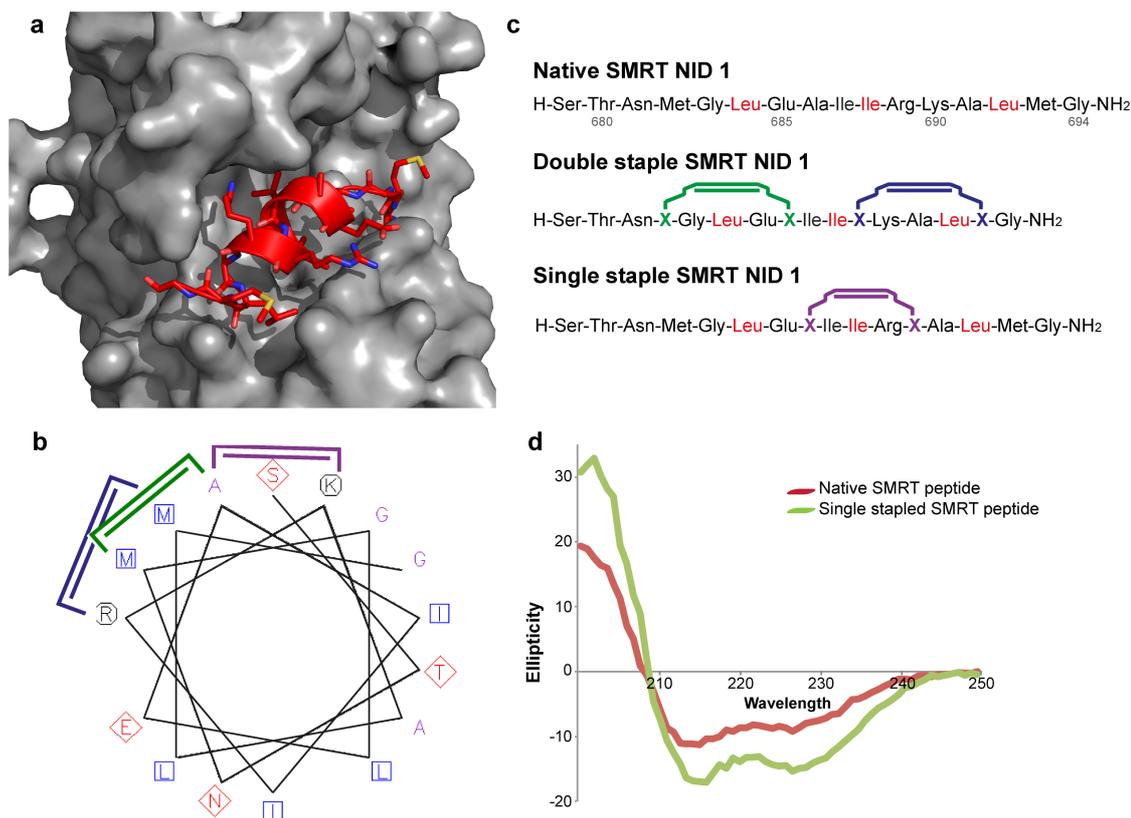


Figure 4.4: Design and CD study of stapled peptides. a) Crystal structure of SMRT corepressor peptide bound to the surface of PPAR α LBD (PDB code: 1KKQ) (Xu et al, 2002). b) Helical wheel projection of the staple scan of SMRT corepressor peptide, where the positions of the staples are shown (green, blue and purple). All of them show the binding motif of the SMRT corepressor peptide to the PPAR α LBD at the back face of the helix free from any interactions. c) Sequences of the two peptides hypothesized from the staple scan. d) Far-UV CD spectra showing enhanced helical content in the selected single SMRT stapled peptide.

4.3.2 - WT and mutant LBDs bind stapled corepressor peptides with great affinity

Having confirmed that stapled SMRT peptides are able to stabilize the different LBDs, a FA assay using FITC-labeled stapled SMRT peptides was performed in order to determine the affinity of this interaction. The K_d of the interaction between the stapled peptides and the different LBDs was determined using increasing concentration of the WT and the mutant LBDs from 10 μ M and a fixed concentration of 5 nM labeled stapled SMRT peptide (**Figure 4.5**).

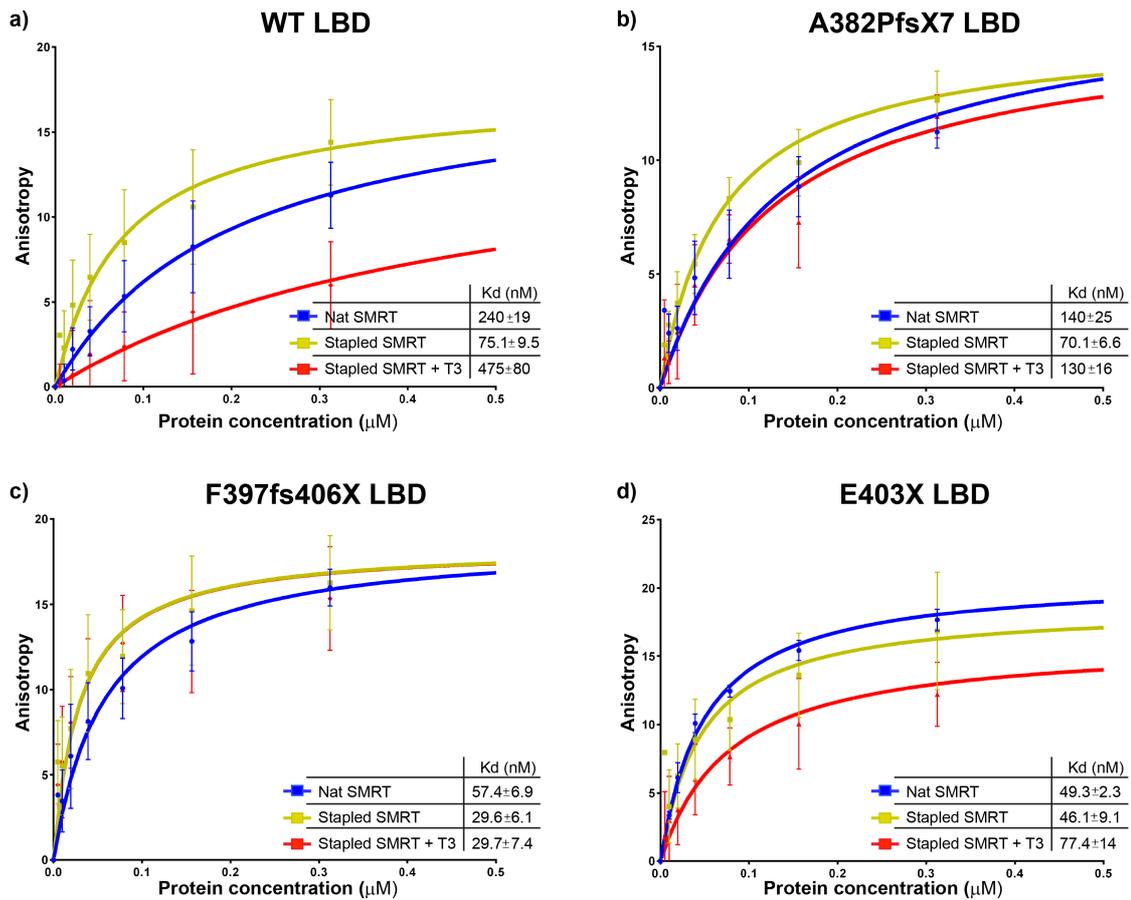


Figure 4.5: Binding assays of stapled SMRT peptide to TR α LBDs showing a comparison between the binding affinity of the native and the stapled SMRT corepressor peptides. Increasing concentration of the mixture LBD:T3 (ratio 1:10) was incubated with 5 nM FITC-stapled SMRT or 5 nM FITC-native SMRT. FA values are the mean \pm SEM of measurements obtained from triplicate experiments. Binding assay LBD:stapled SMRT peptide: a) WT LBD, b) A382PfsX7 LBD, c) F397fs406X LBD, and d) E403X LBD.

The results showed that stapled SMRT peptides are able to bind more strongly to the different LBDs than the native SMRT peptide which means that the improvements in helicity and stability that show these modified peptides actually benefited them to bind tighter to target proteins (**Figure 4.5**). The binding affinity increased by 2 fold in the case of A382PfsX7 LBD and F397fs406X LBD. There is an exception for the E403X LBD. The binding affinity is similar between the native and the stapled SMRT corepressor and the E403X LBD (**Figure 4.5d**), which means that the stapled SMRT does not bind more strongly to this mutant. The explanation for this fact remains unknown. As previously mentioned, it is possible that the remaining H12 residues present in this mutant partially occlude the corepressor-binding site of the LBD,

hindering the peptide binding. This difficulty affects both peptides, the native and the stapled one, and appears to be determinant to limit the binding of the peptides.

Figure 4.5 also shows that, by the same token as in the native SMRT interaction, the stapled SMRT peptide binding was significantly T3-dependent in the case of WT LBD, since the K_d increased in the presence of T3 (**Figure 4.5a**). This 6-fold increment in the dissociation constant is significant and implies that the WT LBD is able to release the stapled corepressor in response to T3. However, the effect of T3 over the stapled peptide is slightly lower than that caused over the native peptide which was almost 10-fold (Chapter 3, Section 3.4.5, **Figure 3.9**).

However, in the case of mutant LBDs, stapled SMRT binding was T3-independent or partially T3-independent since the K_d hardly changed even at high concentrations of T3. That means that T3 does not cause the release of the stapled corepressor from the mutant LBDs.

4.4 - None of the T3 analogues show considerable effect in coactivator recruitment

The third strategy involved using a second type of T3 analogue designed to mimic the Glu 403 side chain. This residue, which belongs to H12, makes an essential salt bridge with the backbone of coactivator specific NR interaction sequence (LxxLL) to recruit coactivator proteins to the DNA (**Figure 4.6c**). Since the mutant LBDs lack H12, they also lack Glu 403 and cannot recruit coactivators in response to T3. The carboxyl was introduced in these T3 analogues to substitute the Glu 403 side chain and recruit coactivators. Different length aliphatic extensions were analyzed in order to find out the one that locates the ending carboxyl in the proper position.

2-fold serial dilutions of purified LBDs and T3 analogues were prepared from a stock of 1 to 10 molar ratio protein: compound (10 μ M protein and 100 μ M compound) in 96-well plates using a fix concentration of 5 nM BODIPY-labeled GRIP1 coactivator peptides. The apparent K_d of the interaction between the coactivator and the different LBDs were determined by FA (**Figure 4.6**)

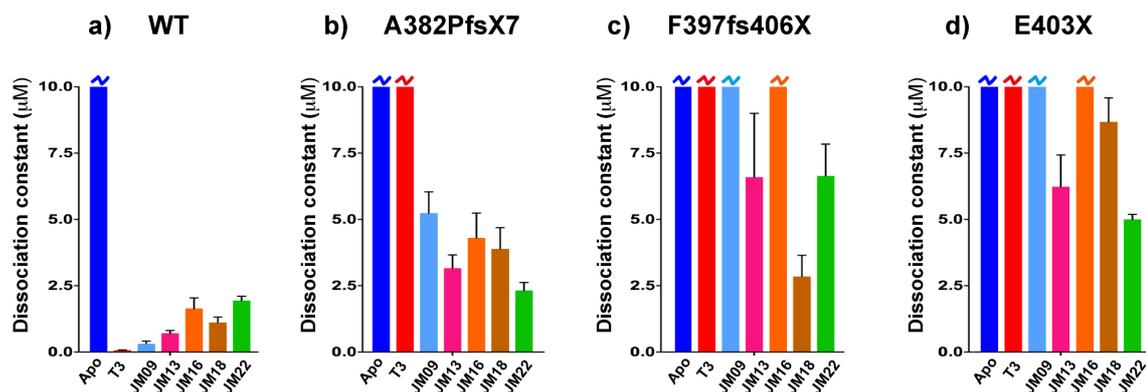


Figure 4.6: Apparent dissociation constant of T3 analogues. The apparent dissociation constant was determined by FA and indicates the T3 analogues ability to displace the labeled SMRT peptide from the different LBDs. Increasing concentration of the mixture LBD:T3 analogue (ratio 1:10) was incubated with 5 nM BODIPY-GRIP1 peptide. K_d values are the mean \pm SEM of measurements obtained from triplicate experiment. a) WT LBD, b) A382PfsX7 LBD, c) F397fs406X LBD, and d) E403X LBD.

The coactivator recruitment properties of these T3 analogues are presented in **Table 4.1**. The results appeared to be slightly different for each mutant LBD. A382PfsX7 mutant LBD seemed to respond considerably well to all the T3 analogues as the apparent affinity between the protein and the coactivator peptide increased in response to them, especially to JM22 which is the one with the longest aliphatic extension before the carboxylic group. The affinity between F397fs406X and coactivator peptides increased more in response to JM18 which is the compound with a ring insertion between the hormone-like structure and the carboxylic acid. None of the T3 analogues had a considerable effect in coactivator recruitment in the case of E403X; the best results were achieved by using JM13 and JM22, the compounds with the longer hydrocarbon extension, but the dissociation constants were still very high.

In comparison to the WT LBD ability to recruit coactivator peptides in response to T3, none of the compounds were strong coactivator recruiters since the affinity of the interaction between the coactivator peptide and the mutant LBDs did not achieved the nanomolar range as in the case of the WT LBD.

Nevertheless, since the mutant LBDs did not recruit coactivators at all in response to T3, shown by the non-detected interaction between them in the presence of T3, coactivator recruitment was considerably better in response to these T3 analogues than in response to T3 (**Figure 4.6**).

4.5 - Discussion

In this chapter T3 analogues and stapled peptides were designed and tested for their ability to promote the release of corepressor proteins from the mutant LBDs or to recruit coactivator proteins.

T3 analogues contain extensions at the 4'OH position of the outer thyronine ring of different chemical composition. First type of T3 analogues were designed to disrupt the corepressor-binding surface of the mutant LBDs and to release the corepressor from that surface, while the second type of T3 analogues were designed to recruit coactivator proteins in response to the analogues. One or two cyclohexanes were added by an ether bond to the 4'OH position of T3 to form the DHJ compounds. Different combinations of cycles were forming the extension coupled to the 4'OH position by a sulfonyl bond in the ES compounds. And some linear or cycle aliphatic extensions were added by an ether bond as well to form the JM family. Thus, different chemical functionalities were tested in affinity and potency at the same time by CD and FA.

According to the results, all the compounds were able to bind all the LBDs to a certain extent, but they exhibited relatively low potency to promote either release of corepressors or recruit of coactivators. JMT3ALK1 and JMTYRALK1 showed comparatively weak LBD binding revealing that the lack of one ring in the structure of the analogue or a very long linear hydrocarbon extension (14 Cs) do not permit a proper fit or access of the compound in the binding cavity of the proteins.

Among all the analogues, compounds ES08 and ES09 displayed the best ability to promote corepressor dissociation in all of the mutant LBDs. These two compounds share a sulfonate ester bond between the 4'OH position of the outer ring and the extension made of two and three hydrocarbon cycles. The coupling of the extension could be the reason of the analogues activity; it is plausible that the bend induced by the sulfonate ester link in the structure of the compound and the relatively rigidity of the extension sterically hinders the corepressor-binding surface of the LBD and lowers the affinity for corepressor binding (**Figure 4.7**).

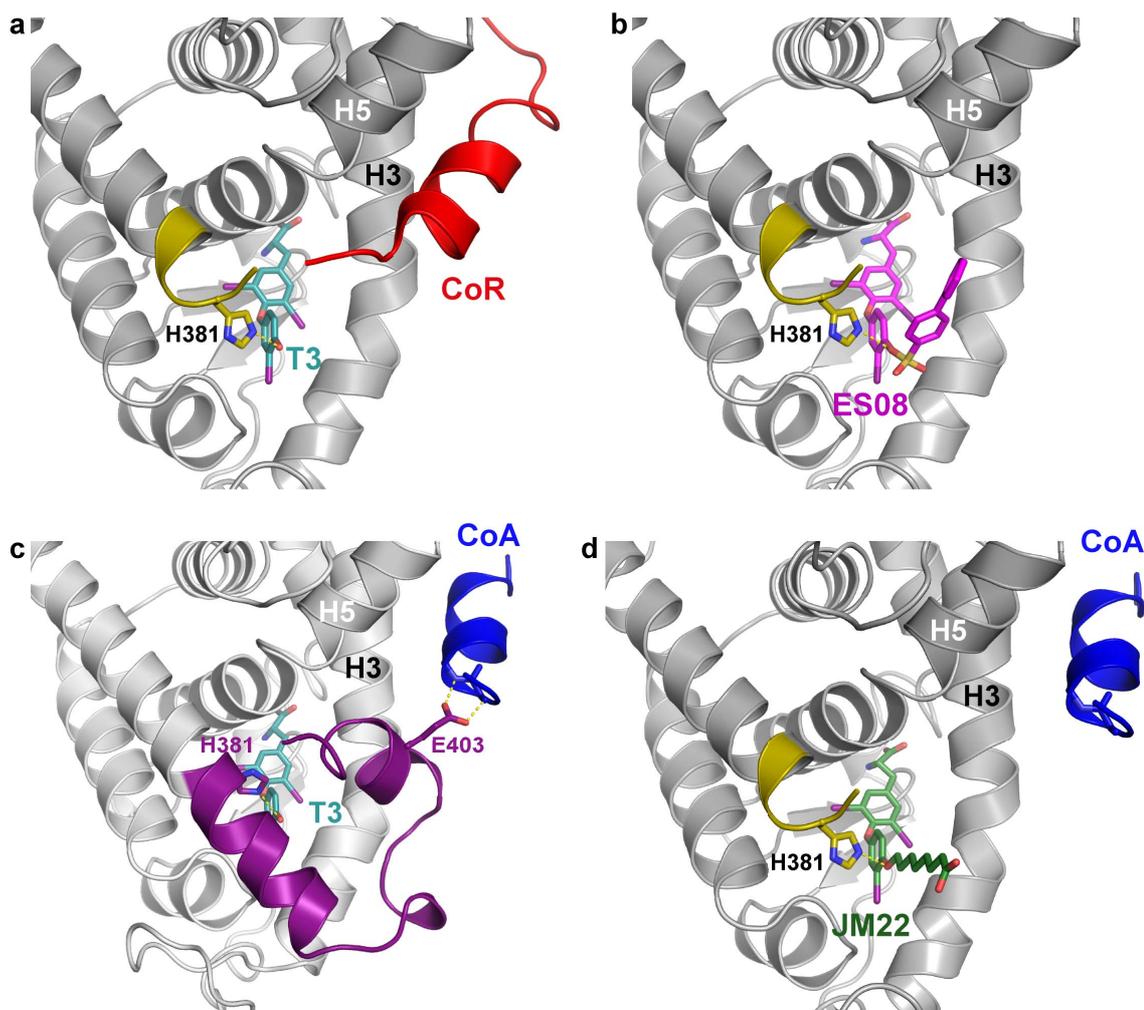


Figure 4.7: Fitting of T3 analogues in the LBP as if they were occupying the same position as T3. In yellow and purple are highlighted the different C-terminal of the mutant LBDs and the WT respectively. a) Model of the T3-liganded mutant LBD in interaction with a corepressor peptide. The corepressor is bound to the CoR binding surface of the LBD, consisting of a hydrophobic groove formed by H3 and H5. b) ES08 fitting into the mutant LBD binding pocket with the 4'OH extension pointing towards the CoR binding surface (H3 and H5) theoretically disrupting it in order to release the corepressor. c) T3-liganded WT LBD showing the Glu 403 making polar contacts with the backbone of the coactivator peptide in order to stabilize its binding. d) JM22 fitting into the mutant LBD binding cavity showing the extension as a rigid rod; however, the flexible composition of the extension would not allow the prediction of its position.

On the other hand, among second type of T3 analogues, JM22 appeared to be the best compound to promote coactivator recruitment. Different length linear hydrocarbon extensions have been tried (3C-7C) to find out the best one in coactivator recruitment. The longest compound was found to be the best, although it is still quite far from the coactivator potency showed by the Glu 403 that it tried to mimic (**Figure 4.6**). Due to the flexibility of linear hydrocarbon chains, it is possible that the carboxylic terminus of

the compound would be able to move and interact with the coactivator peptide in order to encourage its binding.

Corepressor peptides were enhanced by the hydrocarbon staple. The helicity of the stapled SMRT improved as shown in the CD studies that determined that the solution conformation of the stapled SMRT was highly α -helical compared to the linear native SMRT peptide that was weakly α -helical (**Figure 4.4**). FA results showed that stapled SMRT gave an almost 2-fold increase in binding potency relative to native SMRT for the WT, A382PfsX7 and FS397fs406X interaction highlighting the potential of the hydrocarbon staple to further enhance binding affinity while preserving target selectivity.

The comparison of the native and stapled SMRT peptides ability to bind LBDs was achieved by comparing the dissociation constants obtained using the same experimental set up, except for the labeled peptide. Competition assays would have been an option to find out the stapled peptides ability to competitively displace FITC-native corepressor from the LBD. However, since the concentration of fluorescent peptide would have been much lower than the concentration of protein in the assay, only a very small fraction of protein would have been interacting with it and the competition assay results would not have been reliable.

Ideally, stapled SMRT peptides would directly target LBDs corepressor-binding surface, neutralize its interaction with native corepressor proteins and relieve the dominant activity of the mutant receptors. However, the usefulness of stapled peptides *in vivo* raises some pharmacological questions including efficient cell and nucleus penetration, specific high-affinity binding in cells, and proper stability *in vivo* due to their proteolytic degradation resistance.

Stapled peptides capacity for cellular entrance depends on a combination of factors that include charge, hydrophobicity, and α -helical structure, with a high α -helical content and a charge from 0 to +2 can often enhance the cell permeability of stapled peptides (Bird et al. 2008; Bird et al. 2011). The high α -helical content and the +2 charge of stapled SMRT corepressor peptide should promote cellular uptake but experiments

should be performed in order to confirm the cellular delivery and nucleus entrance of the stapled peptides.

4.5.1 - Design of TR α improved analogues

In order to design a perfect T3 analogue useful as a treatment for the RTH α , there are some previous considerations to take into account. First at all, it should be specific for the alpha isoform of the TR and unable to bind to TR β . Secondly, it should bind the mutant LBDs strongly enough (subnanomolar range) to displace T3 from the binding cavity, therefore knowing the dissociation constants of both, T3 and the T3 analogue, and being able to regulate the analogue concentration, the equilibrium would be displaced in favor of the analogue binding. And finally, the T3 analogue binding should release the corepressor proteins present in the cell and this could be obtained by preparing T3 derivatives with bulky extensions at key position that would perturb the corepressor-binding surface of the receptor. These analogues would share characteristics with T3 and therefore dock in the LBP, but would prevent corepressor binding by disrupting their binding surface.

There are, however, inherent synthetic difficulties in further developing T3 analogues. The three iodine atoms of T3 are sensitive to deiodination and limit the substitution at certain positions in the aromatic rings. In addition, due to its enhanced acidity, the hydrogen atom on the hydroxyl group is easily replaced by other substituents. This is why all the compounds tested in this study have extension coupled to the 4' OH position of the T3 outer ring.

However, examination of the P393G T394X LBD (Chapter 5) and the liganded WT LBD structure (pdb code 2H79) reveals that the 5' position of the outer ring points to the space formed by H5, H11 and H12, so an extension in this position might disrupt the formation of the corepressor-binding surface. On the other hand, too flexible extensions might not be efficient enough to disrupt this part of the protein and probably a more rigid linkage such as ethynyl would act as a rigid rod to secure the extension in place. Following these ideas, Nguyen and collaborators designed a potent TR antagonist with nanomolar affinity that also inhibits TR action in an animal model (Nguyen et al. 2002).

This T3 analogue called NH-3 (**Figure 4.8**) also promotes corepressor release in mammalian two hybrid assays and in GST pull-down assays (Nguyen et al. 2002), although it failed to promote coactivator recruitment. NH-3 mechanism of action suggests that it is disrupting somehow the binding surface of the LBD where corepressors and coactivators bind interfering and modifying H3 and/or H5 structure.

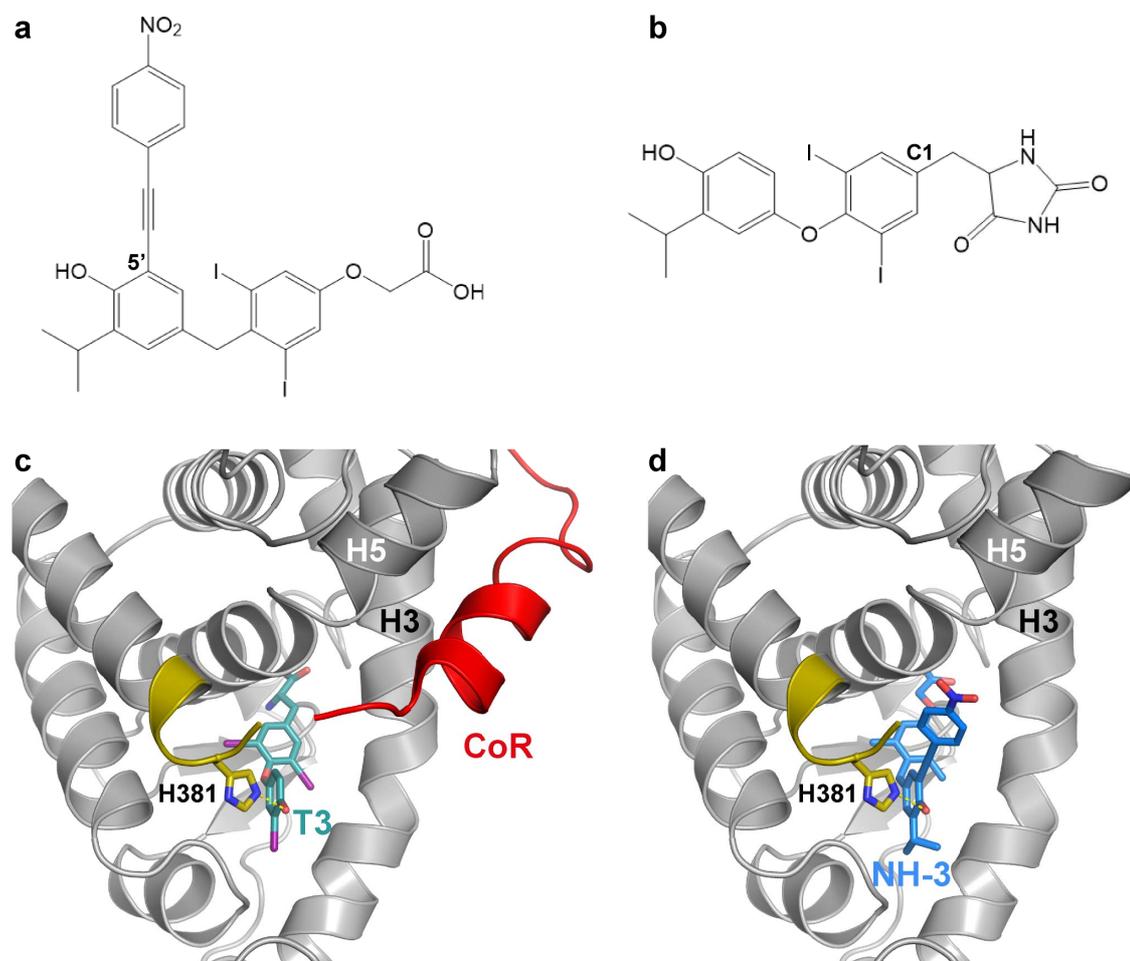


Figure 4.8: NH-3 & CO23 chemical formula and NH-3 fitting in the LBP of mutant LBD. a) NH-3 is a T3 derived analogue that has several modifications in its structure: an extension in the 5' outer ring of T3, two methyl substitutions in the 3 and 5 first ring positions, a dimethyl link between the two rings and an isopropyl substitution in the 3' outer ring that would solve the chemical restrictions. All these extra modifications showed TR β selectivity (Nguyen et al. 2002). b) CO23 compound showing a thiazolidinedione modification in the C1 that conferred TR α -specificity. The two iodines kept at 3 and 5 position of the inner-ring that also improved TR α -specificity (Ocasio & Scanlan 2006). c) Model of the T3-liganded mutant LBD bound to a CoR peptide in the CoR-binding surface of the LBD. d) NH-3 fitting into the mutant LBD binding pocket with the 5' extension pointing towards the CoR binding surface (H3 and H5) disrupting it in order to release the CoR.

The only difference between the LBP of TR α and TR β is that Ser 277 in TR α is substituted by an asparagine (Asn 331) in TR β . Ser 277 (TR α) and Asn 331 (TR β) form part of the binding cavity and make hydrogen bonds with the Arg 228 (TR α) and Arg 282 (TR β), respectively. This substitution induces a change in the position of a structural element of the LBD, the β -hairpin between S1 and S2 (residues Ser 277 to Glu 279 of TR α , residues Asn 331 to Glu 333 of TR β). Nevertheless, all the residues in the polar pocket of both receptors adjust to adopt the same conformations and make the same interactions with the ligand (Wagner et al. 2001). Mutational studies, substituting an asparagine residue for a serine 277 (S277N) in TR α , and a reciprocal variant in TR β , demonstrated that the single amino acid substitution is responsible for the different ligand affinities showed by the two receptors (Wagner et al. 2001). Many TR β specific agonists have been made based on this difference such as Triac (Schueler et al. 1990; Takeda et al. 1995) and GC-1 (Chiellini et al. 1998). Wagner and colleagues compared the structures of TR α and TR β LBDs in complex with Triac and concluded that, despite the few differences that distinguish the two receptors, in the polar part of the binding pocket, the Arg 228 (H3, TR α) forms a hydrogen bond to Ser 277 while Arg 282 (TR β) points away from the ligand. The alternate conformation of Arg 228/282 results from both structural and sequence differences between the isoforms and leads to the binding preferences of Triac for TR β (Wagner et al, 2001).

A comparison of the structures of both receptors revealed another difference between them. Despite having the same sequence, the residues in the loop between H1 and H3 are ordered in the TR α , forming a reverse turn, but are disordered in the TR β . These residues form van der Waals contacts with hydrophobic residues in the loop between H11 and H12. It has been proposed that this structural difference between the TR subtypes could influence the position of H12 and produce differences in hormone affinity (Wagner et al. 2001).

These differences could be exploited to design synthetic isoform-selective ligands that specifically interact with one of the receptors as agonists or antagonists. In fact, Ocasio and Scanlan tried to design a TR α selective agonist and found that thyromimetics that incorporate a thiazolidinedione in the C1 region conferred TR α -specificity *in vitro* and especially in cultured human U2OS osteosarcoma cells (**Figure 4.8**) (Ocasio & Scanlan 2006).

Chapter 5 - Investigating corepressor and mutant LBDs interaction: a structural approach

5.1 - Introduction

The first structures determined of the LBD were solved in 1995. These were the rat TR α (Wagner et al. 1995), the human RXR α (Bourguet et al. 1995), and the human RAR γ (Renaud et al. 1995). These structures led to the prediction of a common fold for all NR LBDs (Wurtz et al. 1996).

Wurtz *et al.* found that the overall fold of the RXR α apo- and RAR γ holo- LBDs were very similar, creating a common primarily helical structural scaffold that forms a single protein domain. This domain was described as a three-layer anti-parallel α -helical sandwich and can be divided in two halves. Half of the domain is occupied by a non-polar cavity in which the ligands bind (Wurtz et al. 1996). This common structure was unexpected since RXRs and RARs belong to two evolutionarily distinct branches of the NR superfamily (Gronemeyer & Laudet 1995).

Despite the similarities and the structure conservation, there appears to be some discrepancies in the mechanism of the molecular switch carried out by the LBDs in response to ligand. Initially, a comparison of the ligand-bound RAR γ (Renaud et al. 1995) with the apo-RXR (Bourguet et al. 1995) suggested that receptors undergo a very specific switch between two conformations that involves the arrangement of H12 from the inactive position to close the binding cavity in the active conformation. The structures suggested that H12 can be located in two positions, but it is always in an ordered helical structure that is displaced to the active position after ligand binding (Renaud et al. 1995). Later, a comparison of several other structures in the inactive and the active state of the receptors (Nolte et al. 1998; Watkins et al. 2001; Xu et al. 2002; Sablin et al. 2003) together with a number of biophysical and dynamic studies of the apo-receptors (Kallenberger et al. 2003; Yan et al. 2004) suggested that ligand binding causes a global stabilization of the LBD, lowering the overall conformational dynamics and involving subtle conformational changes that ultimately lead to the stabilization of

H12 in the active conformation. This “dynamic stabilization” model implies that the unliganded LBD behaves as a molten globule and H12 is not either ordered or fixed in any single position, but rather mobile along with other parts of the protein.

A number of TR structures have been solved providing mechanistic insights into how the protein is folded, how ligands are specifically recognized, how DNA recognition and receptor dimerization are achieved, and how coactivators are recruited. These structures include the DIMIT-liganded rat TR α LBD (Wagner et al. 1995), the T3-liganded human TR α LBD (Nascimento et al. 2006), and structures of the TR α and TR β LBDs bound to synthetic ligands such as GC-1 and Triac (Wagner et al. 2001; Martínez et al. 2009; Bleicher et al. 2008). The determination of these structures defined the structural differences between the TR α and TR β isoforms and their selectivity in ligand recognition. The dimerization features of TR α with RXR α were structurally determined, including the shape and size of the heterodimerization interface (Putchá et al. 2012). The structure of the T3-liganded TR β LBD in complex with a coactivator peptide provided essential information about coactivator recognition and binding (Darimont et al. 1998).

However, the structure of the ligand-free TR remains undetermined because the absence of ligand typically destabilizes the LBD, making it challenging to crystallize. As demonstrated in Chapter 3 (section 3.2.4), the TR α LBD shows rather low melting temperatures in the absence of ligand indicating that TRs have a large range of dynamic behavior. To determine the structure of the TR α LBD in the inactive state in complex with corepressor would be essential in order to elucidate both the mechanism of the molecular switch and the molecular interactions between the corepressor and the LBD.

The aim of the research described in this chapter is to determine the structure of the WT or mutant LBDs in the inactive state with and without corepressor peptides in order to establish the molecular details of the interaction between the corepressor and the TR α LBDs and the structural basis of repression. The determination of the structure of a construct based on the mutant LBD is shown revealing some characteristics that contribute to understand the molecular pathology of the RTH α disease.

5.2 - Crystallization trials

5.2.1 - Initial crystallization studies

Crystallization trials were set up by sitting drop vapor diffusion with the WT and the mutant LBDs. The proteins were purified as described in section 2.9. Protein purity was > 99 % as assessed by SDS-PAGE, and protein concentration was determined using the Bradford dye assay (Bio-Rad).

Drops were set up with a 1:1 ratio of protein solution to screen condition (100 nl sample + 100 nl mother liquor) using six commercial sparse matrices. Protein samples for initial crystallization trials were prepared by mixing the WT or mutant LBDs with either the native or stapled corepressor peptide. The samples were concentrated by centrifugation using a 10 kDa molecular weight cut off filter. The temperature was also considered an important variable, and two conditions were tested; duplicate plates were set up at room temperature and in the cold room. A summary of the optimization process and the initial crystallization attempts can be seen in **Table 5.1**.

Unfortunately, crystallization trials without ligand only produced amorphous precipitation and no hits were observed. For the WT LBD, the flexibility of H12 in the absence of ligand could be hindering crystal formation (Kallenberger et al. 2003). For the LBD frame-shift mutations, it is likely that the C-terminal end consists of a disordered region formed by the loop between H11 and H12. The E403X LBD mutation results in an incomplete H12. This could be as dynamic as H12 in the absence of ligand and adopt different conformations. Since ligand binding not only stabilizes H12 but also globally the whole LBD, adding the ligand to the sample will probably enhance crystal formation.

This was initially avoided in these crystallization trials because the goal of the experiments was to determine the structure of the apo-state of the receptor interacting with the corepressor peptide. However, as the biophysical and biochemical characterization of the mutant LBDs progressed, the experiments described in Chapters 3 and 4 demonstrated that neither T3 nor T3 analogues prevent the interaction of the

mutant LBDs with the corepressor. Therefore, crystallization experiments were repeated with the addition of a ligand.

Commercial screens	Type of protein
NR LBD	WT LBD
Stura and MacroSol	A382PfsX7 LBD
JCSG-plus	F397fs406X LBD
MiDAS	E403X LBD
Morpheus	
PACT <i>premier</i>	

Peptide	Ligand
None	None
Native SMRT corepressor	T3
Stapled SMRT corepressor	ES family T3 analogues JM family T3 analogues

Protein concentration	Molar ratio (protein:peptide:ligand)
15 mg/ml	1:1
10 mg/ml	1:2
5 mg/ml	1:5 1:1:2 1:1:5 1:1:10

Temperature
Room temperature (20 °C)
Cold room (4 °C)

Table 5.1: Initial crystallization experiments. The table shows the variations carried out to optimize the protein concentration and the molar ratios, using a number of commercial screens (obtained from Molecular Dimensions).

The best results in terms of crystallization plate appearance were obtained using 8 mg/ml of protein concentration (60 % of the wells with granular, amorphous or heavy precipitation and 40 % mostly clear drops), with a molar ratio 1 to 1 with either of the peptides, and a molar ratio 1 to 5 with T3 or T3 analogues. Higher protein to peptide or protein to ligand ratios produced high levels of sample precipitation.

The majority of the crystallization plates containing the ternary complex (protein: peptide: ligand) still produced amorphous precipitation as well as clear drops. However, in one of the plates containing the F397fs406X LBD plus either native or stapled corepressor peptide and T3 (molar ratio 1:1:5) there were some oily areas of phase separation. This phase separation consists of highly concentrated protein aggregates that are packed together but not yet in a crystal lattice. Further incubation led to the formation of spherulites or round-shaped small crystals. These quasi crystals were obtained in three related conditions and the condition with the best spherulites was optimized (**Table 5.2**).

5.2.2 - Optimization of initial crystallization conditions

Several different approaches were used to optimize the initial crystallization hits and increase the size and quality of the crystals. These approaches included altering the components of the condition (pH, precipitant and salt concentration), adding further compounds to the original condition which may alter the stability or packing of the protein, altering the ratio of protein to precipitant, and increasing the drop size. A summary of the optimization process used for the F397fs406X LBD can be found in **Table 5.2**.

Crystals obtained were mounted and frozen in a loop. However, when the crystals were X-rayed at the Diamond Light Synchrotron, on the Microfocus beamline, they did not diffract.

Random microseed matrix screening (rMMS) was performed as a last attempt to produce extra hits using the commercial screens available, and a seed stock made from the F397fs406X:CoR:T3 crystals. The results are summarized in **Table 5.3**.

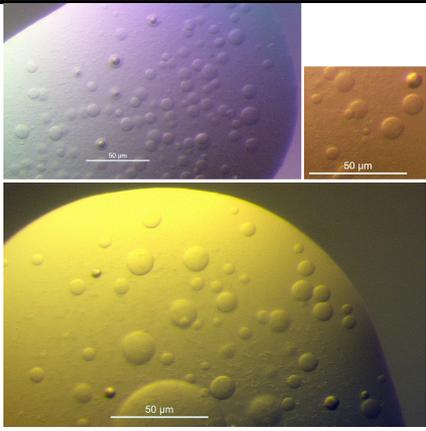
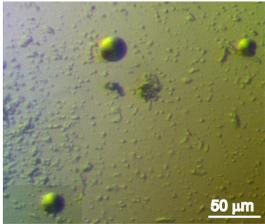
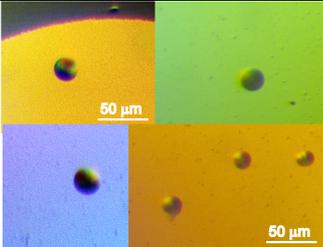
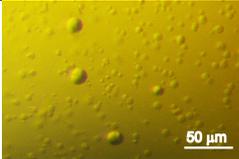
Protein complex	Optimization screen and condition	Picture	Outcome
F397fs406X LBD: Native SMRT: T3 (1:1:5) and F397fs406X LBD: Stapled SMRT: T3 (1:1:5)	First hits in 1.0 M lithium sulphate, 0.1 M imidazole malate pH 6.5, 2 % w/v PEG 8K.		Small round spherulites within oily areas of phase separation Optimization (improve size, shape and quality)
	Optimization plate: variation of salt and PEG concentration, and pH		No change in size or shape observed.
	Increase sample concentration (14 mg/ml), increase protein: precipitant ratio (2:1 and 3:2), and drop size		Bigger crystals, still rather round. Did not diffract.
	Using 10–30 % JCSG+ as an additive to the optimization plate		Slightly bigger single crystals. Did not diffract.
	Adding a range of 1–15 % glycerol		Unsuccessful
	Matrix micro-seeding using a seed stock made of same crystals		Unsuccessful

Table 5.2: Optimization of initial crystallization conditions. The table shows the optimization process performed in order to improve the size, shape and quality of the spherulites obtained from the first crystallization trials.

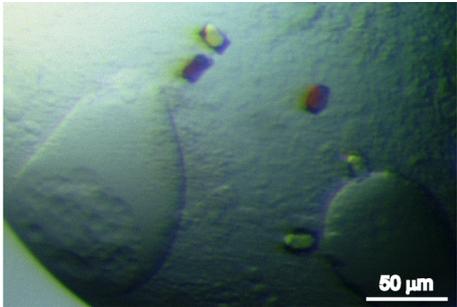
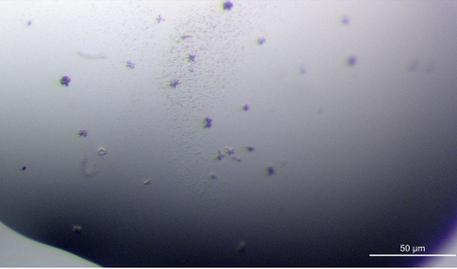
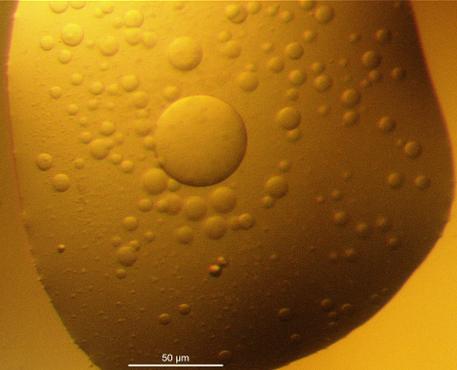
Protein complex	rMMS commercial screen and condition	Picture	Outcome
F397fs406X LBD: Native SMRT: T3 (1:1:5) and F397fs406X LBD: Stapled SMRT: T3 (1:1:5)	2 M potassium/sodium phosphate pH 7		Did not diffract No further hits were observed in the optimization plate
	0.1 M HEPES Na pH 8.2, 50 % v/v PEG 500 MME		No hits were found in the optimization plate
	0.2 M MgCl ₂ , 0.1 M HEPES Na pH 7.5 and 30 % v/v 2-propanol		No hits were found in the optimization plate
	0.1 M ammonium acetate, 0.1 M HEPES Na pH 7.5 and 0.8 M lithium sulphate		No better hits were found in the optimization plate

Table 5.3: Crystals formed in the Stura & MacroSol and NR LBD screens using a rMMS experimental set up. A seed stock from the F397fs406X:CoR:T3 previous crystals was used to seed two commercial screens obtaining the four main hits shown in the table. The best hits were optimized by varying pH and salt and precipitant concentration of the conditions shown in the table.

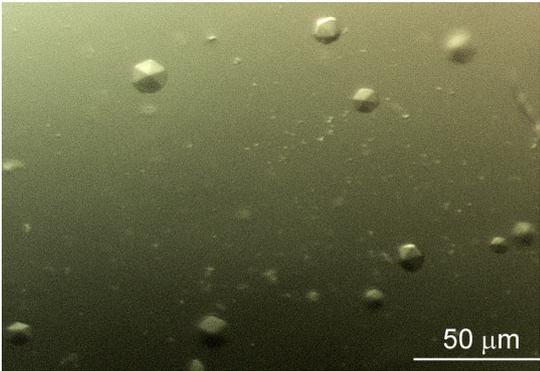
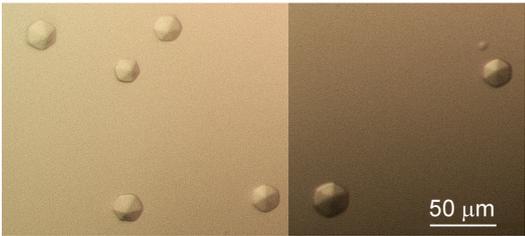
5.2.3 - Modification of LBD constructs

Optimization of the LBD boundaries was attempted as an alternative approach to obtain new and better crystals. The C-terminal of all the mutant LBDs is probably unstructured. In E403X, the loop Glu 391 to Leu 402 between H11 and H12 is probably disordered. In the frame shift mutations, A382PfsX7 and F397fs406X, there are six residues C-terminal to the frame shift that are not TR α LBD and are expected to be unstructured. These extensions may potentially disrupt crystal packing.

The presence of unstructured regions may have been unfavorable in the initial crystallization experiments and therefore four new LBD constructs were produced to remove these amino acids. The termini were removed from the original constructs to produce truncated proteins that ended in the proline 398 (P399X LBD) or in the proline 393 (T394X LBD). A further modification was included to change the last proline to a glycine (P398G P399X LBD and P393G T394X) so that the end will have a small side chain to avoid structural and hydrophobic interference.

The new constructs were expressed and purified following the same protocol as the original constructs. Crystallization experiments were performed with each new construct complexed with corepressor peptide and ligand (molar ratio 1:1:5), using six sparse matrix screens.

Two different crystal forms were observed when using the two different length constructs ending in Gly (P398G P399X LBD and P393G T394X LBD) (summarized in **Table 5.4**). These crystals were taken to the Diamond Light Source and X-rayed on the Microfocus beamline where some of them formed by P393G T394X:T3 diffracted up to 4 Å, allowing data collection.

Protein complex	Optimization trial	Picture	Outcome
P398G P399X LBD: Native SMRT: T3 and P398G P399X LBD: Stapled SMRT: T3 (9.5 mg/ml)	Initial hits: 0.1 M sodium thiocyanate, 0.1 M Tris pH 8.5 and 9 % w/v PEG 10 K		Many small 6-sided bi-pyramidal crystals (5-10 μm) per drop. Optimization
	Varying salt and PEG concentration, change pH		Similar size and number to the previous ones.
	Adding 10–30 % JCSG+ as an additive screen to optimization screen.		Bigger and single crystals suitable for collection. 4 °C: smaller crystals. Did not diffract.
	TCEP as reducing agent in the protein solution. Stura and MacroSol, NR LBD and optimization screen		Most of the drops with granular or full precipitate; clear drops for the remaining conditions.

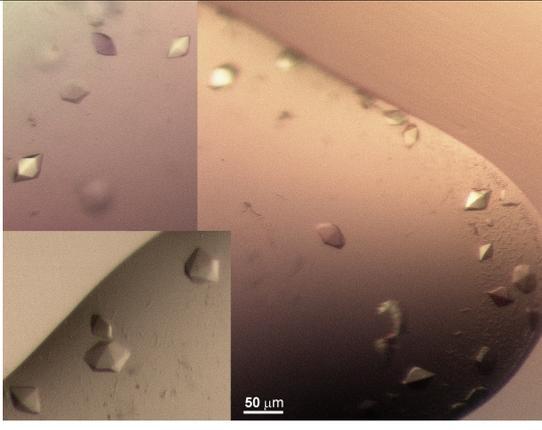
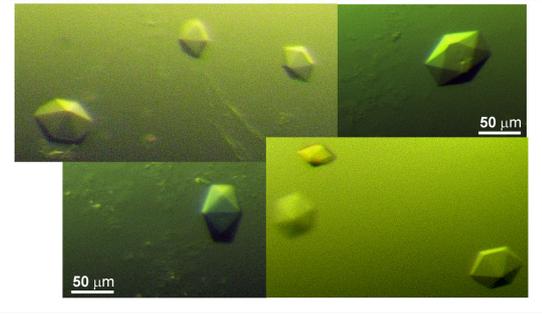
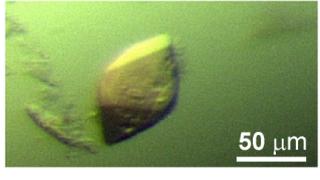
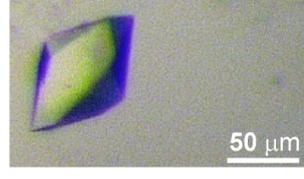
P393G T394X LBD:T3 (1:5) (9.5 mg/ml)	Initial hits observed in many conditions of the NR LBD commercial screen		Large (ranged from 50 to 100 μm), diverse shape yet well-defined sharp edges, single crystals suitable for data collection.
	0.1 M Tris pH 8, 3.5 M sodium formate		Crystals suitable for X-ray diffraction.
	0.2 M sodium chloride, 0.1 M Tris pH 8.5, 10 % w/v PEG 8K		Crystals suitable for X-ray diffraction.
	0.2 M sodium chloride, 0.1 M Tris pH 8.5, 1.0 M lithium sulphate		Crystals suitable for X-ray diffraction. Diffraction up to 3 Å, data collection
	All the previous optimization screens at 4°C		Many smaller crystals, often in clumps.

Table 5.4: Optimization process of initial hits conditions. The table shows the crystallization trials conducted using the new constructs that resulted in P398G P399X LBD and P393G T394X LBD crystals and the optimization process performed afterwards.

5.2.4 - Structure determination

In order to collect a complete data set from single crystals, especially using intense synchrotron radiation, crystals are rapidly frozen at 100 K using liquid nitrogen. The addition of a cryoprotectant is crucial to prevent small ice crystals forming that would affect the quality of the data collection. The cryoprotectant buffer was based on the mother liquor from the crystallization condition with the incorporation of various concentrations of glycerol (0, 5, 10, 15, 20, 25 %) (Section 2.9.2, materials and methods). The buffers were tested by flash cooling at 100 K using the in-house X-ray source. The best cryoprotectant buffer contained 20 % glycerol included in the respective mother liquor solution.

i) Data collection

A selection of 19 crystals of the P393G T394X LBD:T3 complex were taken to the Diamond Light Source synchrotron (UK) where they were analyzed at the microfocal X-ray beam line I-24. Most of the crystals diffracted at very low resolution (6-8 Å), however it was possible to collect complete data sets for four crystals containing T3.

First, three images were collected at 0, 45 and 90° to determine the resolution of the diffraction data derived from that crystal and the unit cell symmetry and dimension (crystal system). Then, if the diffraction data resolution was at least 4 Å, a complete data set of 1800 images was collected. At Diamond, strategy calculations are automatically carried out using Mosflm (data harvesting program to integrate images) (Leslie et al. 2011; Powell et al. 2017) and EDNA (sample characterization and data collection strategy program) (Incardona et al. 2009). For each crystal 180° of diffraction data were collected starting from the angle recommended by Mosflm using an angle of oscillation of 0.1°/frame. The wavelength was 0.969 Å, the beam size of 50 x 50 µm, the time exposure of 0.1 s/frame, and the transmission of 50 %.

The best diffraction data were collected on a single P393GP394X LBD: T3 crystal from the H3 optimization screen and were processed automatically using Xia2. Xia2 is an automatic data processing system which includes software such as Mosflm, Labelit

(Sauter et al. 2004), Pointless (Evans 2005), CCP4 (Collaborative Computational Project 1994) and XDS (Kabsch 1993).

ii) Data processing

The data collected already processed by Xia2 were reanalyzed using Pointless, Aimless and Truncate:

- Pointless determines the scores for all the possible Laue groups consistent with the crystal class based on the cell dimension restraints and suggests a space group by checking sets of reflections which may be systematically absent
- Aimless (Evans & Murshudov 2013) merges partial reflections together and each set of symmetry equivalent reflections into a single observation.
- Truncate (French & Wilson 1978) converts intensities to structure factors from merged data. In addition, Truncate calculates a number of statistics from the intensity data which can be used to assess data quality.

Originally, data were processed in P6₂22. Molecular replacement was performed using WT TR α LBD (PDB code 2H79; Nascimento et al. 2006) as a search model in Phaser (McCoy et al. 2007). However subsequent refinement in Refmac failed to reduce the R_{free} indicating that the structure did not refine. This suggested that something was incorrect, probably the space group. It was then realized that Phaser had changed the space group to P6₄22.

The data obtained from Diamond were then reprocessed using Mosflm. Mosflm is a data processing program part of to the CCP4 platform and can be used to process diffraction images in an interactive mode. A comparison of the first images with the last ones suggested a deterioration of the crystal; the last images showed fewer and weaker spots. After spotfinding and indexing, a list of solutions sorted by the penalty score was generated. The space group was manually chosen as P6₄22, the one with the highest symmetry and lowest penalty. After the cell refinement and the integration of the images, the output file was processed using Aimless and then Phaser. This confirmed that the correct space group was P6₄22.

After checking the Aimless summary statistics, it became apparent that the crystal was slightly anisotropic and had been damaged by radiation, so there were some bad batches at the end of the collection. The R_{merge} was relatively stable across all batches until number 1600, where it increased suggesting radiation damage. An examination of the accumulative intensity of the data showed that the data were 100 % complete using the first 680 images, therefore an exclusion of batches from 1601 to 1800 would not eliminate unique reflections. The maximum resolution, indicated by the parameters $CC_{1/2}$ and $I/\sigma(I)$ ($CC_{1/2} \geq 0.3$ and $I/\sigma(I) \approx 1.5$), was 2.7 Å along the H and K axis and 2.92 Å along the L axis. Therefore, the data was reprocessed in Aimless using a resolution cutoff of 2.7 Å and excluding batches from 1601 to 1800. The resulting statistics improved, showing that the overall R_{meas} value was acceptable (0.186), the data was 100 % complete with a 16.8 redundancy, and the crystal was untwined (**Table 5.5**).

The number of molecules in the crystallographic asymmetric unit was estimated by running the Matthews Cell Content Analysis (Matthews 1968; Kantardjieff & Rupp 2003). The results revealed that there was one molecule in the asymmetric unit if the solvent content is about 50 % in the crystal.

The structure was solved by molecular replacement using WT TR α LBD (PDB code 2H79) as a search model in Phaser. Once a preliminary model was defined, it was refined using Refmac5 (Murshudov et al. 1999) against the data to improve the phases that resulted in noticeable clearer maps. Consequently, the LBD sequences were rebuilt iteratively using multiple rounds of refinement and building using Refmac5 and Coot (Emsley et al. 2010).

Data collection	P393G T394X: T3
Space group	P6 ₄ 22
Cell dimensions	
<i>a, b, c</i> (Å)	143.33 143.33 88.50
α, β, γ (°)	90.00 90.00 120.00
Resolution (Å)	72.06–2.7
R_{sym} or R_{meas}	0.186
$I/\sigma I$	16.7
Completeness (%)	100.0
Redundancy	16.8
Refinement	
Space group	P6 ₄ 22
Cell dimensions	
<i>a, b, c</i> (Å)	143.33 143.33 88.50
α, β, γ (°)	90.00 90.00 120.00
Resolution (Å)	124.126–2.7
Unique no. reflections	15,235
Total no. reflections	255,225
$R_{\text{work}}/ R_{\text{free}}$	19.54/ 22.04
No. atoms	
Protein	3824
Ligand	23
Water	9
B-factors	
Protein	61.78
Ligand	60.87
Water	52.55
R.m.s deviations	
Bond lengths (Å)	0.0196
Bond angles (°)	2.1217

Table 5.5: Data collection and refinement statistics. The table indicates the data collection and refinement statistics found for the P393G T394X LBD:T3 structure.

iii) Model building and structure refinement

The electron density corresponding to the iodines from T3 was very intense so that the hormone could be unambiguously placed within the binding pocket of the LBD (**Figure 5.1**). The C-terminal of the WT LBD was clearly different from the P393G T394X LBD so this part of the template was deleted from the model. C-terminal residues from Met 376 were carefully added, making sure that they fitted the observed electron density.

After every round of manual building by Coot, a Refmac5 run was made in order to refine the model against the electron density. Refmac5 refines the atomic model by adjusting the model parameters to the experimental data and compares the calculated intensities with the experimental intensities in order to determine the free R-factor (R_{free}) and the work R-factor (R_{work}). These factors are used as indicators of the refinement progress; the better the model fits in the electron density, the more similar the model becomes to the experimental data, and the lower the R-factor values. Therefore, after every round of refinement, a decrease in the R-factors is expected until reasonable values are obtained. If the free R-factor increases it is an indicator of over fitting.

The final model contained three molecules in the asymmetric unit each with amino acids 156–391 from the WT TR α LBD and one molecule of T3. The final model with $R_{\text{work}} = 19.54\%$ and $R_{\text{free}} = 22.05\%$, had 95.73% residues in the favored region, 4.27% in the allowed region and none in the outlier region of the Ramachandran plot (**Table 5.5**).

The B-factor values of the protein and the ligand were slightly high, consistent with the values obtained for other structures of the TR LBD with T3 (Darimont et al. 1998; Nascimento et al. 2006; Putcha et al. 2012). The stabilization obtained by ligand binding does not stabilize the C-terminal which has higher B-factors.

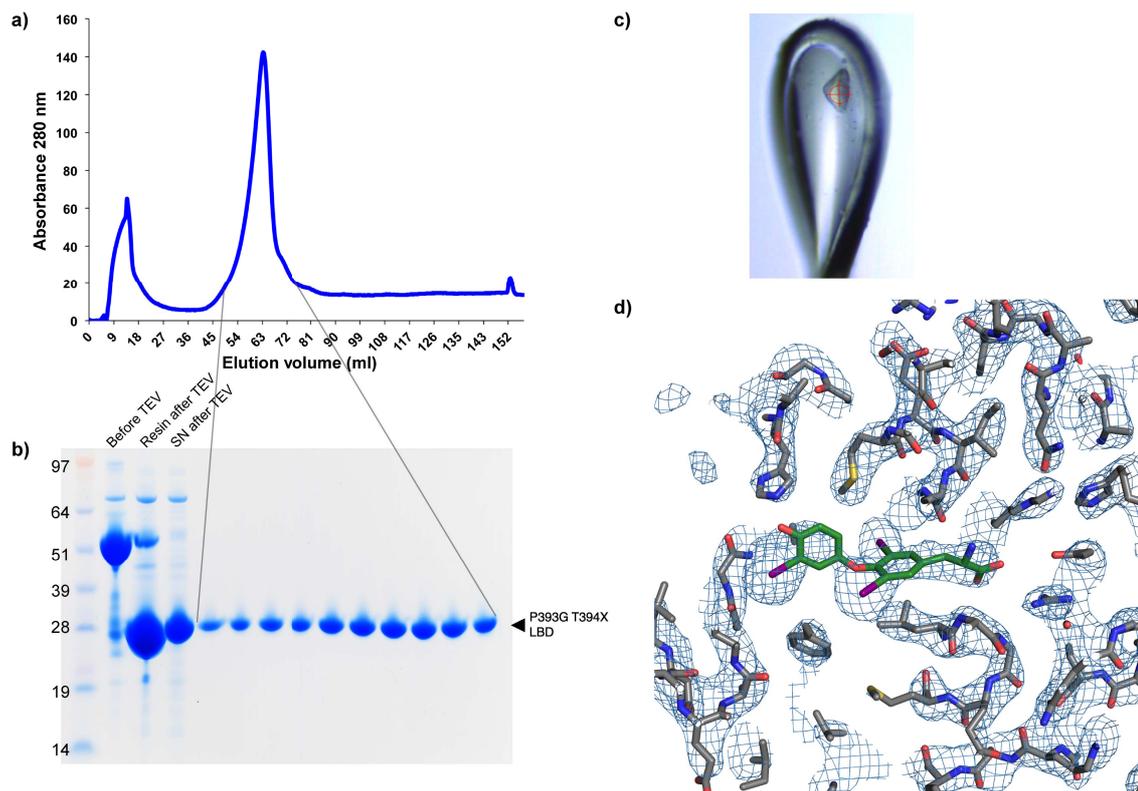


Figure 5.1: Purification and crystallization of the P393G T394X LBD: T3 (1: 5) complex.
 a/b) HiTrap Q IEX chromatogram and SDS-PAGE gel showing the results of the purification process of the protein by GST affinity chromatography and IEX. c) Image of a P393G T394X LBD: T3 crystal mounted in a loop at the Diamond Synchrotron beam line I24. d) Electron density ($2F_0 - F_c$) contoured at 1.0σ around the binding pocket of the LBD and showing the residues interacting with T3.

5.2.5 - Asymmetric unit and crystal packing

The lattice parameters of the unit cell ($\mathbf{a} = \mathbf{b}$; $\alpha = \beta = 90^\circ$; $\gamma = 120^\circ$) define a hexagonal crystal system with a primitive distribution of the lattice points. The minimum symmetry is one 6-fold-rotation axis along \mathbf{c} . Two of the axes (\mathbf{a} and \mathbf{b}) are of equal length, separated by equal angles (90°), and lie in the same plane. The third axis (\mathbf{c}) is perpendicular to the plane of the other three axes. (**Figure 5.2a**).

The asymmetric unit of the model consists of one molecule (**Figure 5.3**). This molecule within the unit cell is organized in an arrangement that is described by the symmetry operations defined by the space group (P6₄22). The capital letters define the lattice type and the numbers define the symmetry operations (rotation and/or screw axes) that are carried out to create the contents of the unit cell. The P6₄22 space group defines a hexagonal crystal system with the lattice points located only at the vertices of the unit cell forming a primitive lattice (P). The hexagonal crystal system typically has one 6-fold-rotation axis along the c axis, which defines two of the axes of the unit cell to be equal in length ($\mathbf{a} = \mathbf{b}$) and separated by 90° ($\alpha = \beta$), and the other angle (γ) to be 120° . The point group, as a description of the rotational symmetry, is 622 which signifies that there are two 2-fold rotation axes along x and y and a 6-fold rotation axis along z. The combination of the rotational symmetry (rotation and screw axes) defined by the point group and the translational symmetry defined by the crystal system results in 12 symmetry operators. These 12 symmetry operators (crystallographic symmetry) carried out by the asymmetric unit define the coordinates of the 12 asymmetric unit equivalent positions within the unit cell (**Figure 5.2**) [*International Tables for X-ray Crystallography* (Space group 181, pp. 574-575)] (Aroyo et al. 2006). Knowing the space group (P6₄22) and the content of the asymmetric unit (1 molecule), the position of every atom in the crystal can be defined.

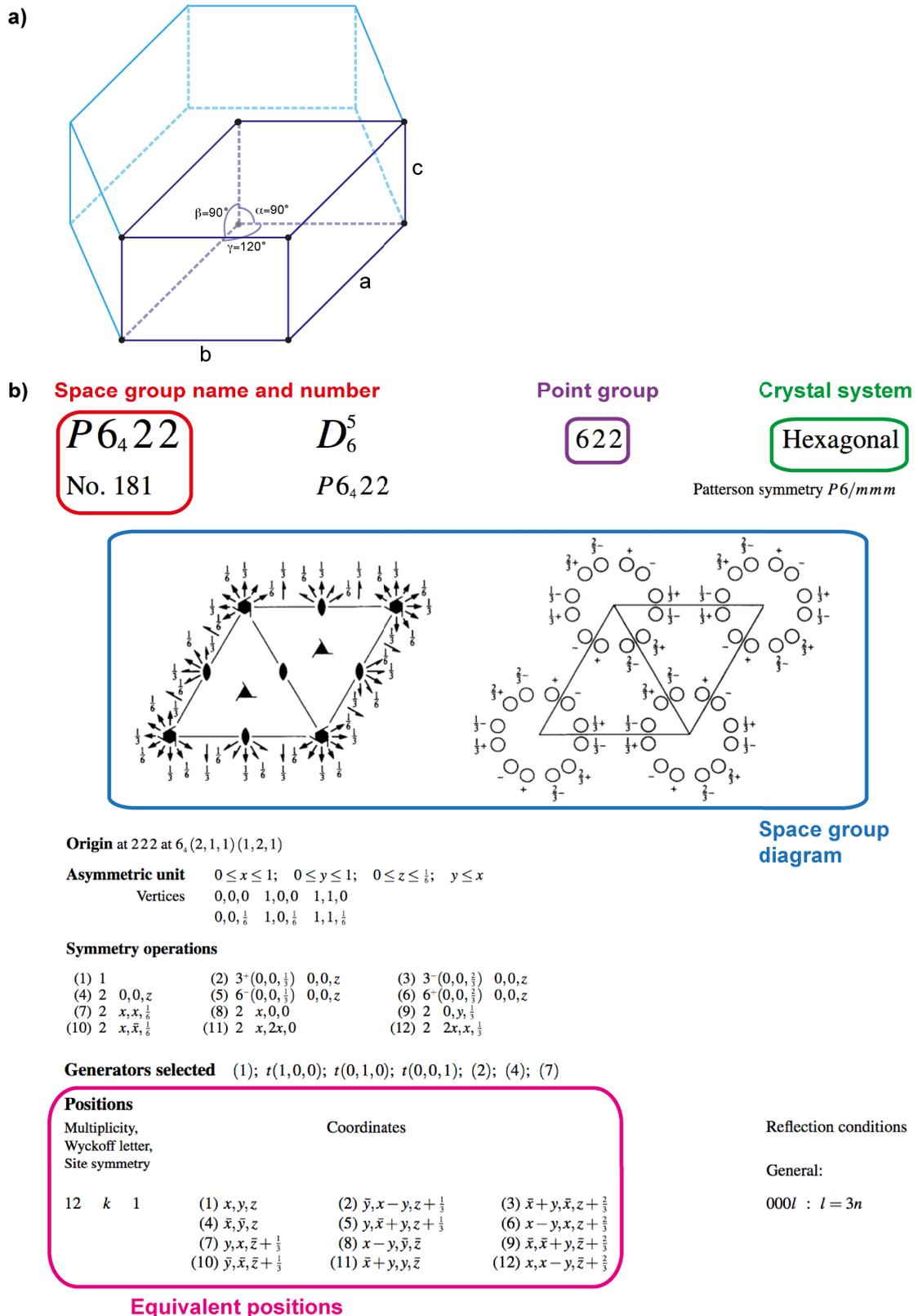


Figure 5.2: P6₄22 space group description. a) Picture of the hexagonal crystal system with primitive distribution of the lattice points. b) Description of the P6₄22 space group adapted from the International Tables for X-ray Crystallography highlighting the most useful information (Aroyo et al. 2006).

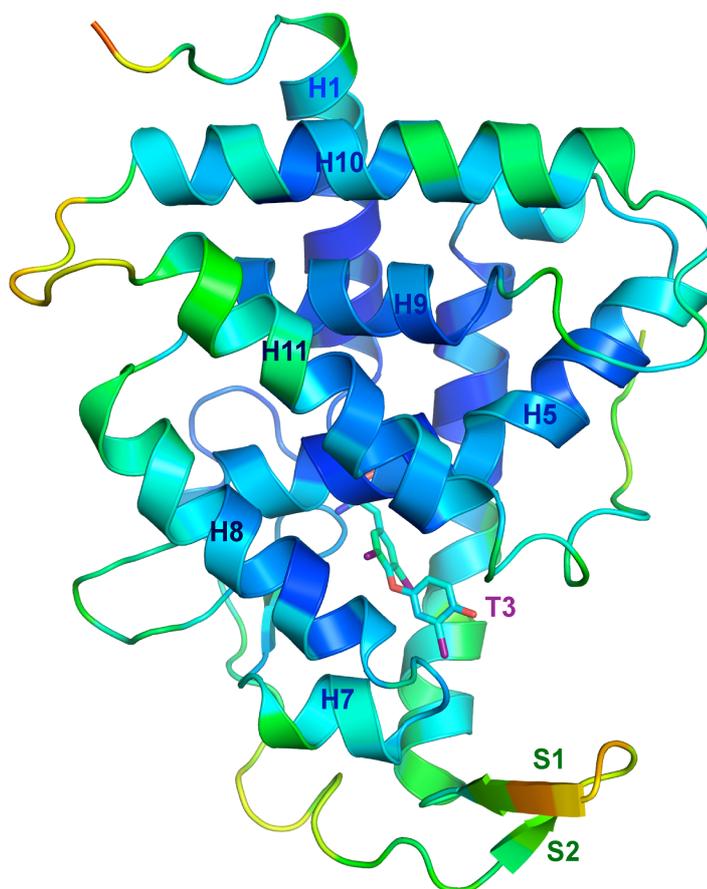


Figure 5.3: Crystallographic asymmetric unit. The picture shows the molecule that forms the asymmetric unit with the elements of secondary structure labeled.

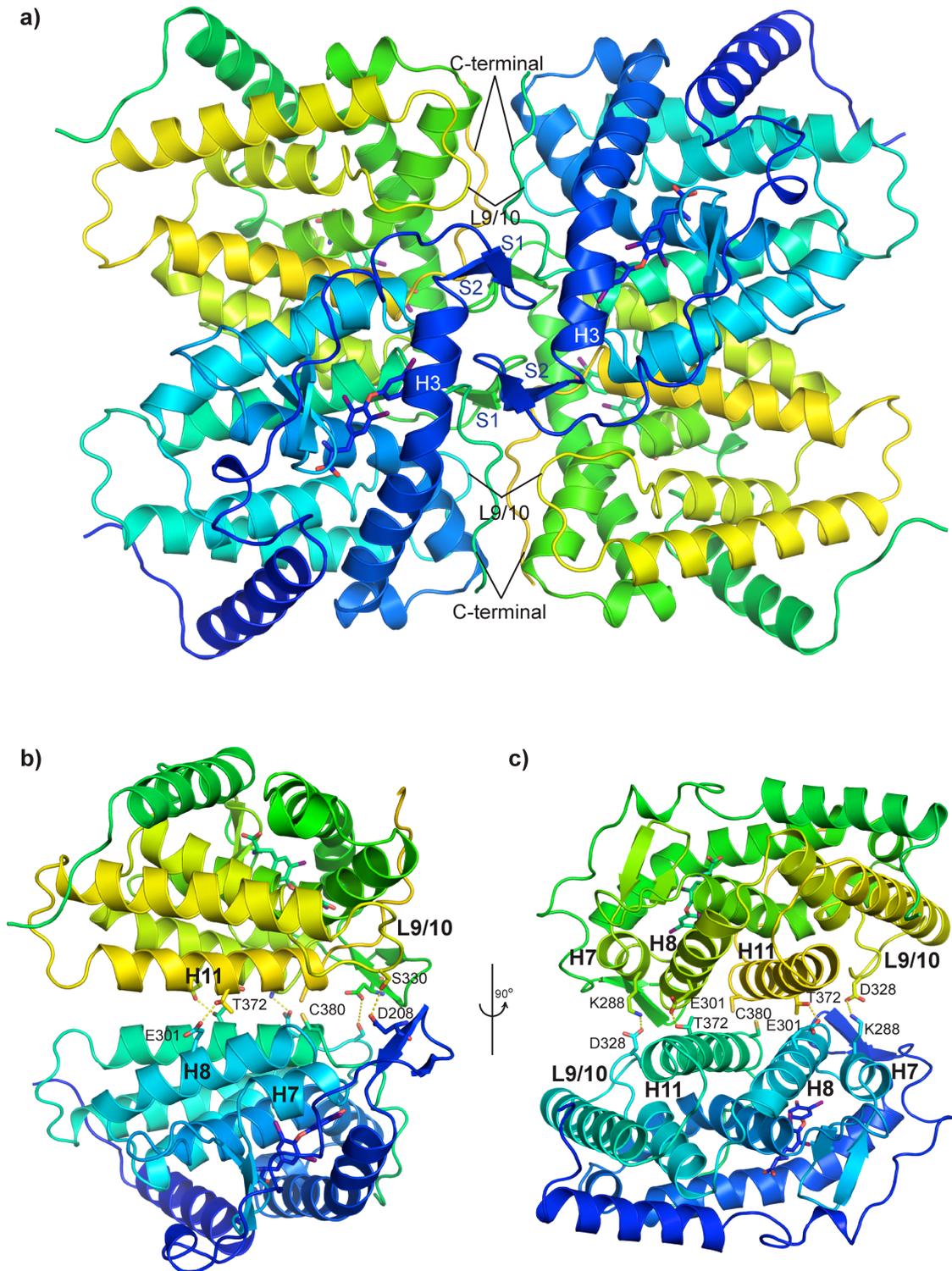
A closer examination of the crystal packing of the model revealed that the proteins are forming a tetramer (**Figure 5.4**). The proteins forming the tetramer are tightly packed, interacting with each other through close intermolecular contacts that could introduce distortions in the structure, making it different from the protein in solution. This possible disruption of the real structure will be discussed later.

Each tetramer is made up of two dimers related by a two fold rotational symmetry. The interaction between the two dimers to form the tetramer is mainly through van der Waals contacts between the two first β -strands (S1 and S2) and the beginning of H3 as well as the C-terminal tail and the loop between helices 9 and 10 (L9/10) of the other molecule. The interface area between them is about 630 \AA^3 (**Figure 5.4a**).

The two molecules that form the dimer are related by rotational symmetry through a 2-fold axis. The dimer interface is 714.6 \AA^3 and involves interactions through the H8, the

L9/10 and the C-terminal end of H11. Interaction between these two symmetry-related molecules results in the formation of a disulphide bond between Cys 380 of each protein. This covalent link appears to stabilize the dimerization and favour nucleation and crystal growth (**Figure 5.4b**).

The interactions among the molecules within the crystal is a consequence of the crystal packing, so it does not mean that the mutant TR α LBDs are forming dimers or tetramers in solution or the dimerization and/or tetramerization play a physiological role.



5.2.6 - Overall architecture of the P393G T394X LBD

The P393G T394X LBD protein consists of a single structural domain packed in three layers, composed of 11 α -helices, H1-H11, and four short β -strands, S1-S4. The overall fold of the protein is essentially the same as WT TR α LBD. H11 is shorter than WT TR α LBD and H12 is not present in the P393G T394X LBD (Figure 5.5, 5.6).

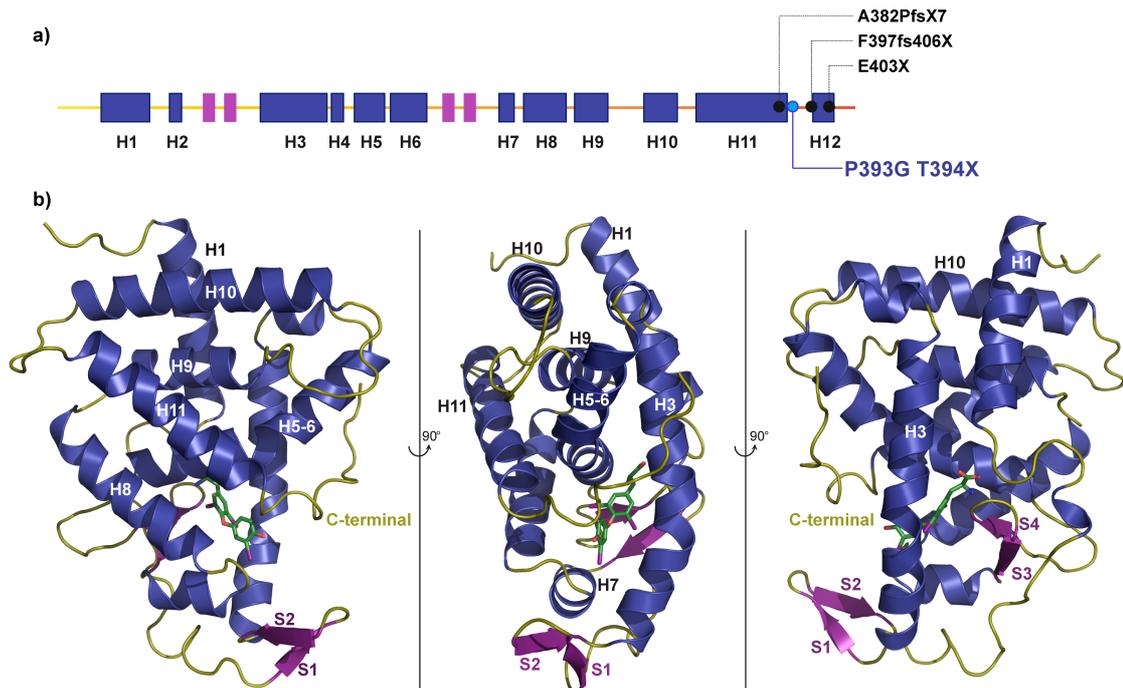


Figure 5.5: Structure of P393G T394X LBD construct. a) Schematic view of the TR α sequence, showing the relative position of the mutant receptors in study and the construct made in order to improve the crystal growing. b) Overall structure of the P393G T394X LBD with secondary structure elements labeled. T3 is depicted as green sticks, β -strands are purple, coil conformations are yellow and α -helices are blue.

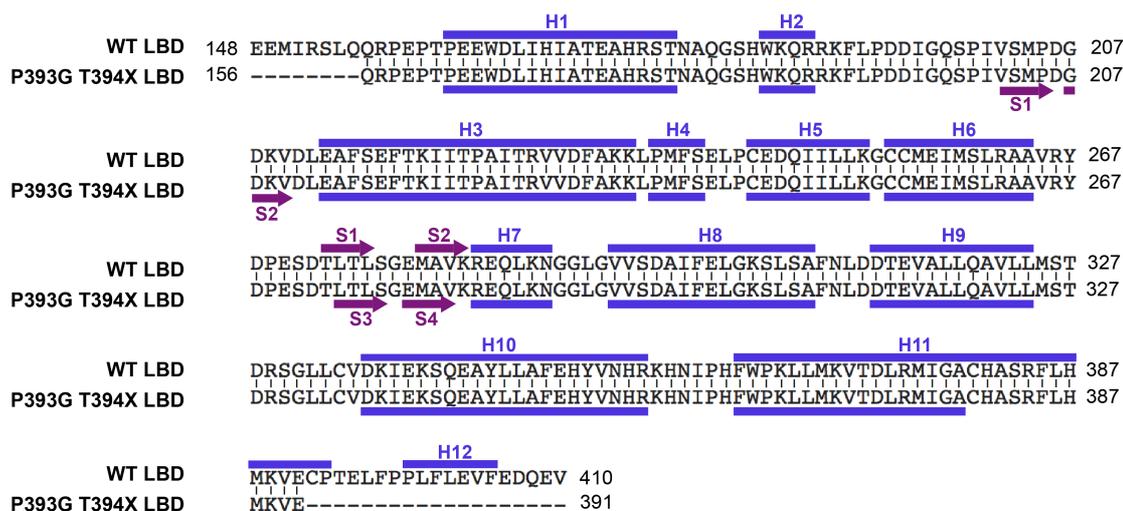


Figure 5.6: Alignment of the sequences for the P393G T394X LBD and WT LBD (2H79) showing the differences in secondary structural elements (α -helices and β -sheets).

In the P393G T394X LBD there are two β -sheets (S1: Val 202-Met 204; S2: Asp 208-Val 210) running between H2 and H3 instead of the long loop that runs between H2 and H3 in the WT LBD. The two β -strands are occupying space available because the C-terminal of H11 and H12 are missing. The S1 and S2 β -strands form part of the LBP (**Figure 5.7**).

In the WT TR α LBD H11 extends diagonally across the full-length of the molecule, but in the P393G T394X TR α LBD it abruptly turns at Ala 379 to form an unstructured region composed mostly of polar residues. The C-terminal of H11 runs antiparallel to the adjacent H5 until the putative corepressor/coactivator-binding site where it interacts with H3 and H5. Some polar and Van der Waals interactions stabilize the end of the H11 and keep it in position next to H5. Arg 384 makes a hydrogen bond with the 4' hydroxyl of T3 and electrostatic contacts with Thr 219 contributing to place the C-terminal tail. In contrast, in the WT TR α LBD Arg 384 is situated on the other side of the protein, exposed to the solvent (**Figure 5.7**).

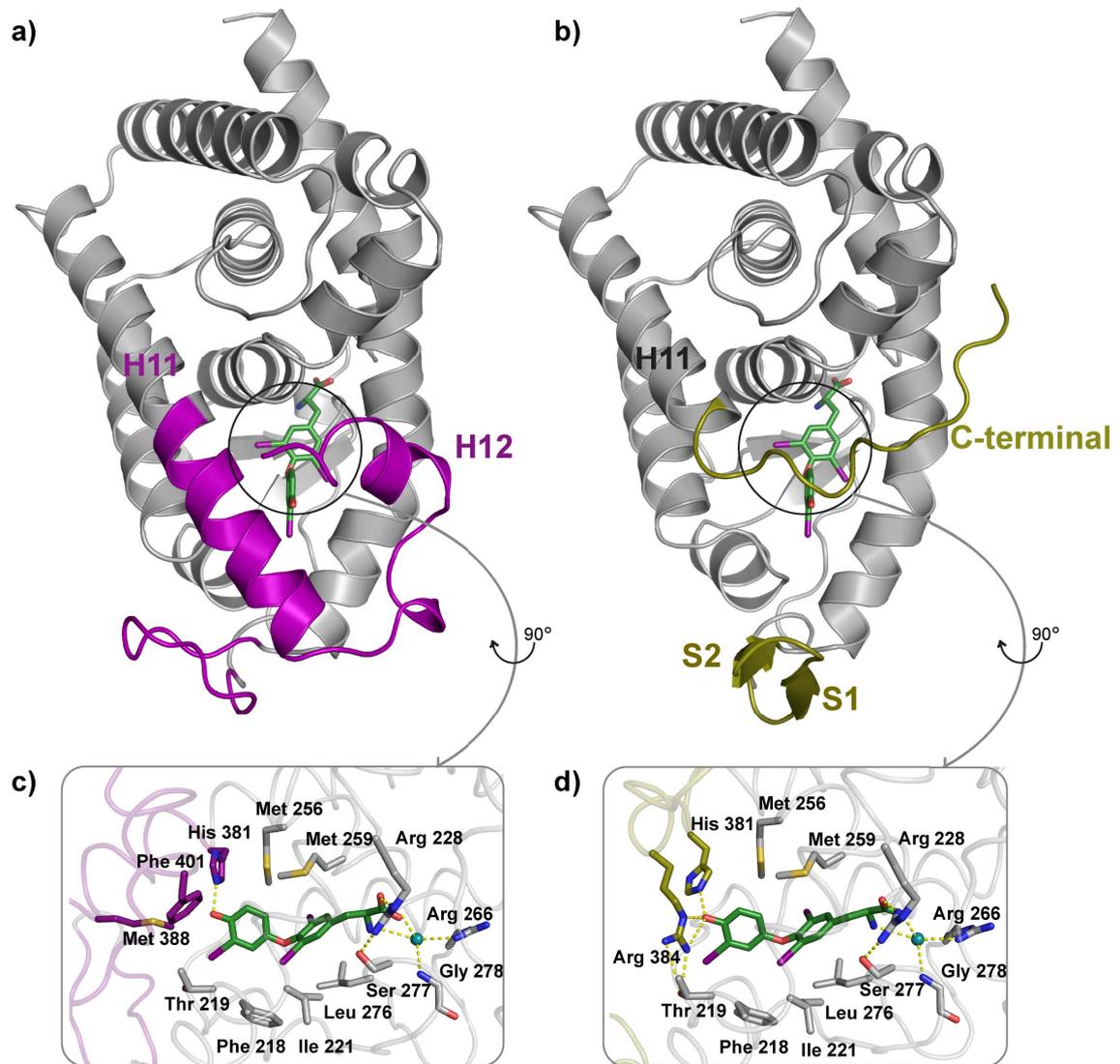


Figure 5.7: Significant differences between P393G T394X LBD and the WT LBD (2H79). a/b) Overall structures of P393G T394X LBD and WT LBD (2H79) with the regions of significant difference colored in purple (2H79) and yellow (P393G T394X). c/d) Comparison of the WT (c) and P393G T394X (d) LBP showing that residues His 381 and Arg 384 are located in different positions and make different interactions.

ii) The hormone-binding cavity

The T3 hormone binds very similarly to both P393G T394X LBD and the WT LBD. However, the interactions between T3 and the residues from H11 and the volume of the LBP are slightly different. The turn of H11 at Ala 379 means that the next residues are not a helix and are in a different position. This affects the orientation of His 381, which still forms a hydrogen bond with the 4' hydroxyl, and the position of Arg 384, which also interacts with the 4' hydroxyl of T3. Arg 384 in P393G T394X LBD is occupying the position of the Phe 401 of the WT LBD and contributes to the LBP in this side of the protein (**Figure 5.7**).

The volume of the LBP corresponding to the P393G T394X LBD and the WT LBD was calculated using POCASA (Yu et al. 2009). The results showed that the volume of the P393G T394X LBP is about 200 Å³ which is larger than the volume of the WT LBP (160 Å³). In the WT LBD the presence of a complete H11 and a H12 means that the LBP is tightly closed with H11 interacting with T3 through His 381 and Met 388, and H12 interacting with T3 through Phe 401. In contrast, in the P393G T394X LBD Met 388 is not interacting with T3 and its side chain is pointing to the solvent, and Arg 384 substitutes Phe 401. As a consequence of the different structure of H11 and the lack of H12, the P393G T394X LBP is bigger and more accessible to solvent than the WT LBP (**Figure 5.8**).

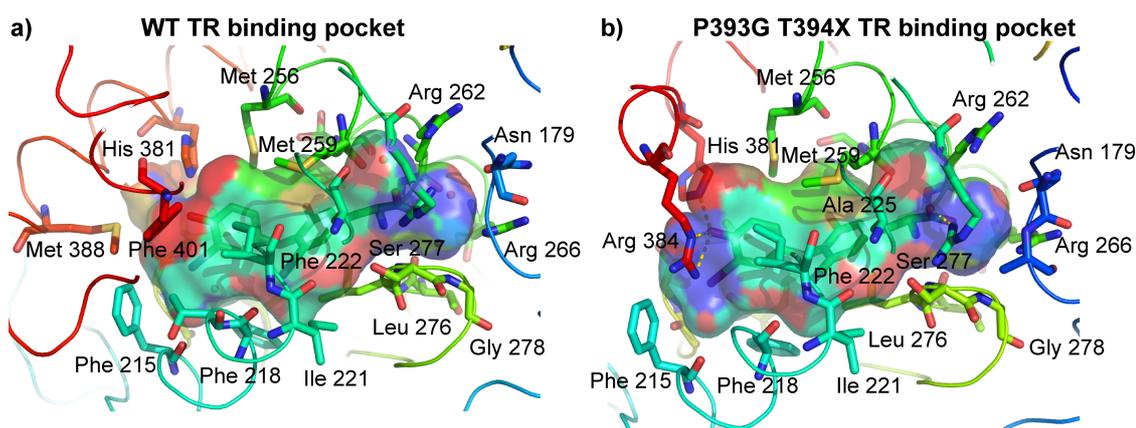


Figure 5.8: Surface view of the binding cavity in a) WT LBD (PDB code: 2H79) and b) P393G T394X LBD showing the larger LBP of P393G T394X. Residues which are forming the ligand binding cavity are shown as sticks and labeled.

iii) Heterodimerization interface

TR α heterodimerization interface between TR α and RXR α comprises of L9/10, H10, L10/11 and the beginning of H11 of TR α LBD. The buried surface area is 961.4 Å² and involves the interaction of 28 residues from each protein. The interactions that keep the proteins together are salt bridges between the Asp 328, Glu 339, Glu 343 and Lys 366 from TR α LBD and Lys 356, Glu 394, Lys 417 and Arg 421 from the RXR α LBD and hydrophobic Van der Waals contacts involving specific residues from Val 320 to Met 376 (Putcha et al. 2012).

In P393G T394X LBD the heterodimerization interface is conserved, so the heterodimerization capability is assumed for this construct and, by extension, also for the other mutant proteins. This heterodimerization interface contributes to the dimer interface between chains A and B (**Figure 5.4**).

iv) P393G T394X LBD structure suggest that the H11 C-terminal can mimic the corepressor/coactivator interaction motif

As discussed earlier, in P393G T394X LBD H11 is an ordered regular α -helix from Phe 363 until Ala 379; the following residue has a 90° turn and the remaining C-terminal tail of the protein runs unstructured and antiparallel to the adjacent H5 until the end of the molecule.

Superposition of the corepressor and coactivator peptides from PPAR α LBD complexed with corepressor and TR β LBD complexed with coactivator (PDB codes: 1KKQ and 1BSX) with the P393G T394X LBD structure reveals that the C-terminal amino acids are overlapping at the coactivator and corepressor-binding site (**Figure 5.9**).

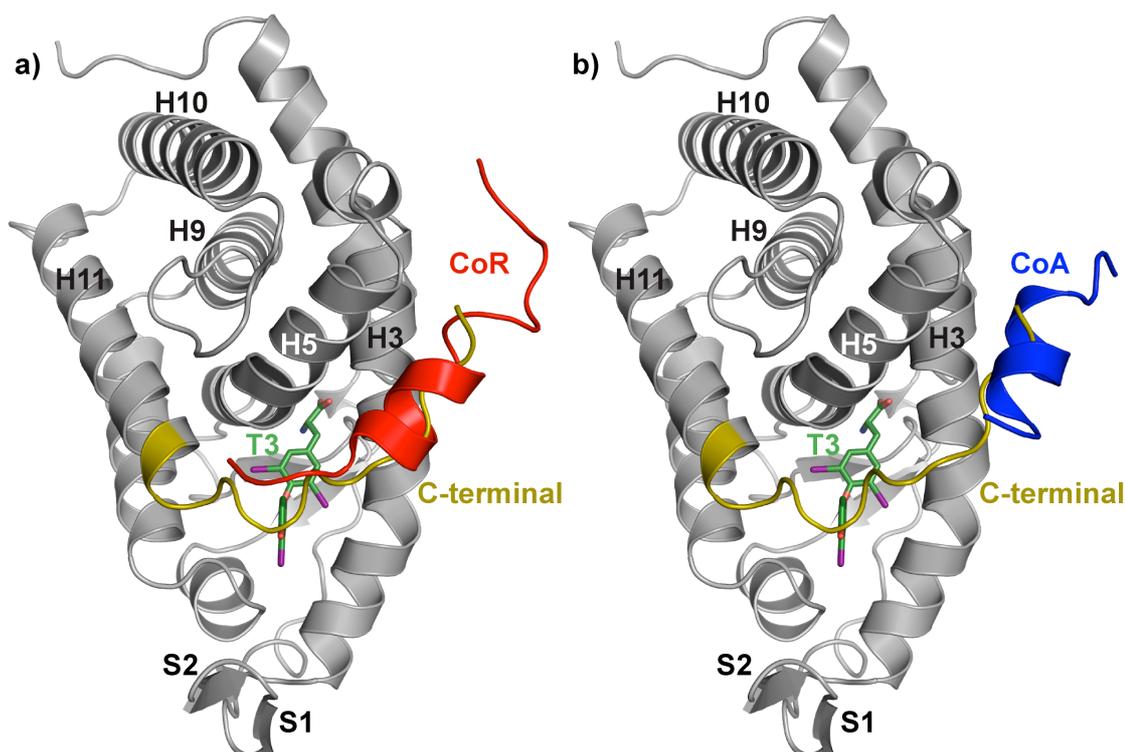


Figure 5.9: Structural model of P393G T394X LBD in interaction with corepressor and coactivator peptides. The corepressor SMRT peptide in (a) and coactivator GRIP1 peptide in (b) are superimposed on the P393G T394X LBD structure from the PPAR α LBD:SMRT complex (pdb code: 1KKQ) and the TR β LBD:GRIP1 complex (pdb code: 1BSX).

However, the biophysical characterization of the construct showed that it could bind corepressor peptide and T3 at the same time, forming a ternary complex. The FA results confirmed the interaction between the LBD and the corepressor was of high affinity (102 nM) and was enhanced by T3 binding up to 37 nM (**Figure 5.10**). These results suggest that the C-terminal position in the crystallized protein is a consequence of the crystal packing of the molecules within the crystal and that the C-terminal unstructured tail is probably a dynamic segment of the protein.

As described earlier in **Figure 5.4**, chains A and B in P393G T394X LBD interact with each other via residues from the C-terminal tail of the proteins. Unfortunately, the location of the C-terminal of the protein in the crystal means that this crystal packing cannot accommodate a corepressor peptide.

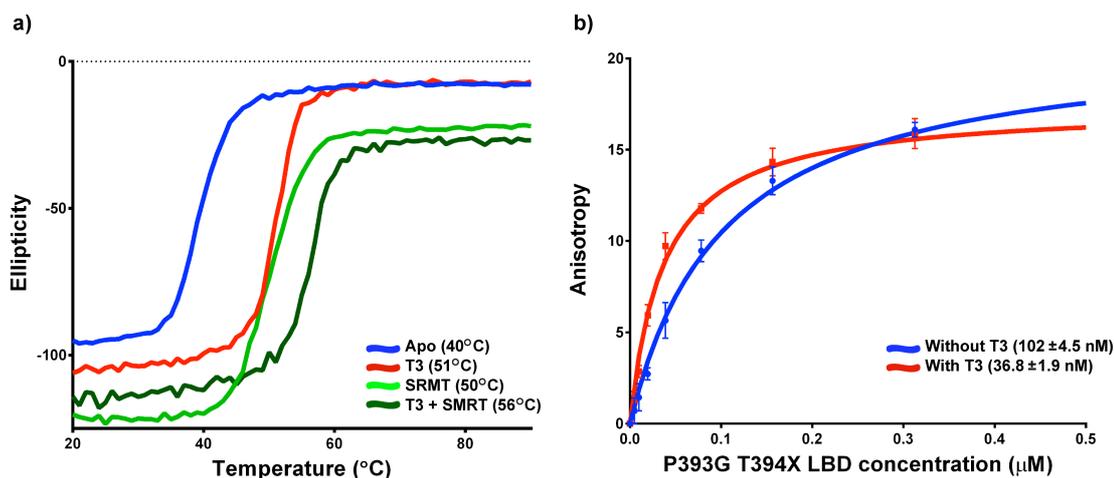


Figure 5.10: Biophysical characterization of P393G T394X LBD. a) CD results showing the melting temperature of the P393G T394X LBD alone (blue) or in the presence of T3 (red), corepressor peptide (green) and both T3 and corepressor (dark green). b) FA results showing the affinity values of the interaction between P393G T394X LBD with corepressor peptide, in the liganded and unliganded state. FA values are the mean \pm SEM of measurement obtained from triplicate experiments.

5.2.7 - Crystallization trials of T3 analogues liganded LBDs

Crystallization experiments were also performed using the P393G T394X LBD in complex with T3 analogues and with or without corepressor peptide. For the first round of experiments, two ratios of LBD: T3 analogue were tested (1:2 and 1:5) using the NR LBD Molecular Dimensions screen and the four optimization plates already prepared in the previous assays based on previous hits from the NR LBD screen (A4, B9, D10 and H3). No crystals were found in the plates that contained corepressor peptide.

The optimization screens based on the NR LBD screen generated a large number of crystals, with similar 6-sided bi-pyramidal shape and approximately 50 μm size to the crystals previously obtained. ES32, JM13 and JM18 containing samples produced crystals which required no further optimization as the crystals were large enough and had defined edges (**Table 5.6**).

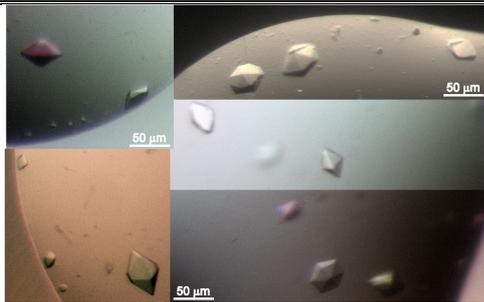
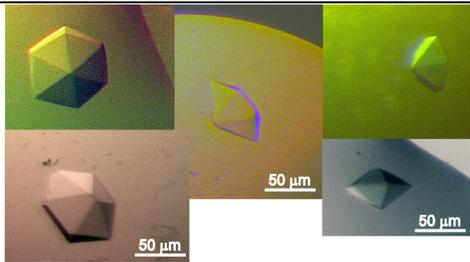
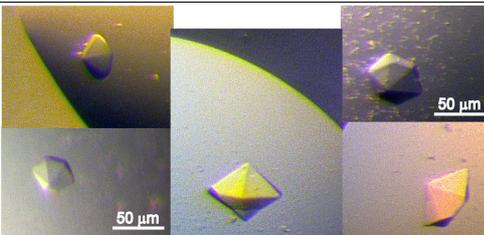
Protein complex	Condition	Picture	Outcome
P393G T394X LBD:ES32 (1:5)	0.2 M sodium chloride, 0.1 M Tris pH 8.5, 10 % w/v PEG 8K (NR LBD screen, condition B9)		Crystals suitable for collection that diffracted to 4.0 Å
P393G T394X LBD:JM13 (1:5)	0.2 M sodium chloride, 0.1 M Tris pH 8.5, 1.0 M lithium sulphate (NR LBD screen, condition H3)		Crystals suitable for collection that diffracted to 3.0 Å
P393G T394X LBD:JM18 (1:5)	0.2 M sodium chloride, 0.1 M Tris pH 8.5, 1.0 M lithium sulphate (NR LBD screen, condition H3)		Crystals suitable for collection that diffracted to 3.4 Å

Table 5.6: Pictures of the P393G T394X LBD crystals containing T3 analogues (ES32, JM13, JM18). The table shows the sample composition and the crystals growing in the different optimization plates, B9 NR LBD screen for ES32 and H3 NR LBD screen for JM compounds.

The data collected from the P393G T394X LBD: ES32 complex was not good enough to solve the structure as the resolution was too low. In the case of the P393G T394X LBD interacting with the JM compounds, the structures were solved by molecular replacement using the previous model (P393G T394X LBD without the T3) in Phaser. The asymmetric unit contains one molecule, as in the previous model, which is packed within the crystal forming tetramers. The data was processed using the same space group as the previous structure (P6₄22). The statistics for data collection and refinement are shown in **Table 5.7**.

Data collection	P393G T394X: JM13
Space group	P6 ₄ 22
Cell dimensions	
<i>a, b, c</i> (Å)	142.81 142.81 90.32
α, β, γ (°)	90.00 90.00 120.00
Resolution (Å)	123.7–3.00
R_{sym} or R_{meas}	0.182
$I/\sigma I$	13.3
Completeness (%)	99.9
Redundancy	12.4
Refinement	
Space group	P6 ₄ 22
Cell dimensions	
<i>a, b, c</i> (Å)	142.81 142.81 90.32
α, β, γ (°)	90.00 90.00 120.00
Resolution (Å)	123.7–3.00
Total no. reflections	140,988
Total no. unique	11,339
$R_{\text{work}}/ R_{\text{free}}$	14.99/ 22.85
No. atoms	
Protein	3824
Ligand	41
Water	-
B-factors	
Protein	60.8
Ligand	63.6
Water	-
R.m.s deviations	
Bond lengths (Å)	0.0188
Bond angles (°)	2.2747

Table 5.7: Data collection and refinement statistics. The table indicates the data collection and refinement statistics found after solving the P393G T394X LBD: JM13 structure.

However, while the T3 part of the compound was easy to position due to the electron density from the iodines, there was not any extra electron density for the modification of the 4' hydroxyl (**Figure 5.11c**). A possible conclusion is that the compound was not stable during the crystallization experiments and turned out to be degraded leaving only T3 in the hydrophobic pocket. In support of this idea, it took longer for these crystals to grow, approximately 1 month.

Another possibility is that the compound extension, as a mobile and flexible portion of the compound, adopted multiple conformations and occupied different positions during the crystallization process so that it cannot be seen in the electron density. In support of this idea, some extra density was revealed and the compound extension could be fitted in the position showed in **Figure 5.11**. However, the density disappeared upon refinement in Refmac5.

Composite simulated annealing omit maps were calculated using Phenix (Terwilliger et al. 2008). These maps also had no density for the aliphatic tail.

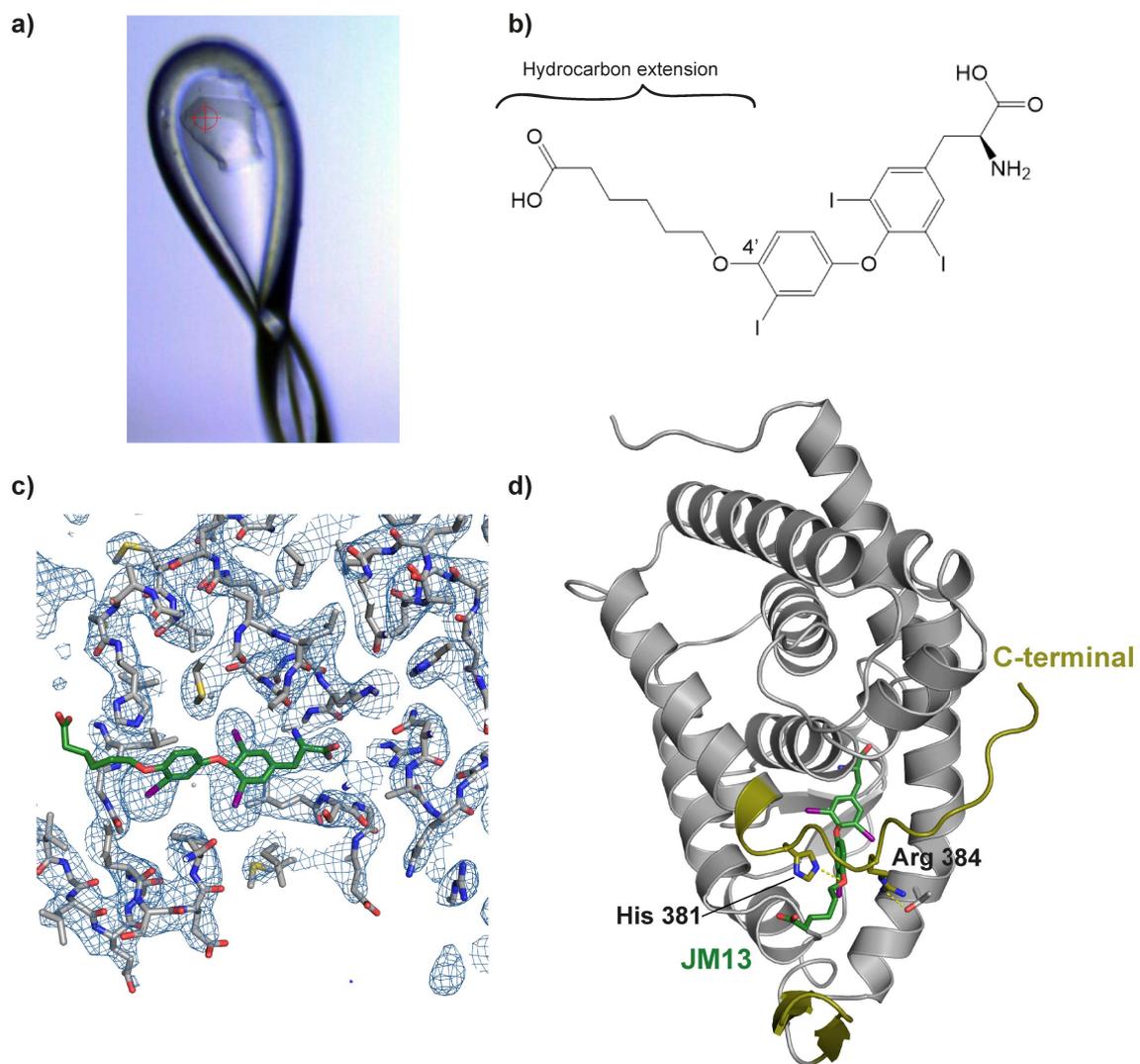


Figure 5.11: P393G T394X LBD:JM13 (1:5) complex. a) Image of a P393G T394X LBD:JM13 crystal mounted in a loop at the Diamond Light Synchrotron beam line I24. b) Chemical formula of JM13. c) Electron density ($2F_0-F_c$) contoured at 1σ around the LBP, showing the JM13 ligand and the residues surrounding it. d) Overall structure of the P393G T394X LBD with the JM13 compound in the LBP.

5.3 - Discussion

The principal aim of the experiments described in this chapter was to investigate the structure of the WT and mutant TR α LBDs in complex with corepressor peptides. However, it was only possible to solve the structure of a designed construct (P393G T394X LBD) based upon the mutant LBD sequence that removes the potentially disordered amino acids after the frame-shift mutation. It was possible to crystallize this construct in the presence of T3 as a ligand, but not in the presence of corepressor peptide. Nevertheless, solving this structure has confirmed that mutant LBDs are able to accommodate T3 within the hydrophobic pocket in the same position as in the WT.

Superimposition of the WT LBD structure and the P393G T394X LBD model revealed two large differences. The first one is the formation of two extra β -strands (S1 and S2) between H2 and H3. Secondly, the structure revealed a fundamental unexpected difference in the C-terminal of the LBD. H11 of P393G T394X LBD is much shorter than in the WT LBD because a 90° turn at Ala 379 disrupts the helix. The C-terminal amino acids run antiparallel to the adjacent H5 and beneath it, until the corepressor/coactivator-binding site where the tail packs loosely against H3 and H5. A new hydrogen bond is formed between the Arg 384 and the 4' hydroxyl of T3 stabilizing the position of the C-terminal tail. The C-terminal tail helps to enclose the hydrophobic cavity on that side of the protein and contributes to the globular shape. However, the position of the C-terminal tail in the structure is probably a consequence of the packing of the molecules within the crystal. In solution the C-terminal tail is most likely dynamic and behaves as a mobile structure independent of the rest of the protein that is not situated in any particular position.

Examination of the binding pocket structure revealed that all the structural and chemical characteristics of the cavity remained the same in the shorter LBD. Residues that form the hydrophobic region of the binding pocket are situated in the same position as in the WT protein and make the same van der Waals interactions with the hydrophobic rings and iodines. The polar region of the binding pocket, formed by side chains from H3, H4 and S3 interacting with the T3 amino propionic and water molecules, also remains the same. Despite the 90° turn of H11 at Ala 379, His 381 is in a similar position but

different orientation from His 381 in the WT and still forms a hydrogen bond with the 4' hydroxyl of the T3 outer ring. Since there is a high degree of similarity between the binding cavity of the construct and the WT, it is reasonable to expect that mutant LBDs are also capable of accommodate T3 in the same position as in the WT. However, although binding of T3 to the mutant LBDs changes the dynamic behavior and stabilizes the LBDs, it does not release the corepressor from its binding site as described in Chapter 3.

The hydrophobic pocket of the P393G T394X LBD is bigger than for the WT LBD. This suggests that mutant LBDs might accommodate bigger compounds than T3 in the LBP. Compounds with extensions of the outer ring of T3 designed to modify the corepressor-binding surface would be able to access and fit in the cavity. The extension should be pointing to the corepressor-binding site in order to interfere with the corepressor binding, therefore, extensions in the 5' position of the outer ring would be recommended. In addition, it is clear that for maximal effect the extensions should be rigid in order to limit the freedom of movement. These conclusions will be taken into account to design new T3 analogues and further investigate their potential as therapeutic agents.

Regarding the crystal packing, P393G T394X LBD molecules form tetramers within the crystal. Some NRs structures have been already reported to contain a tetrameric organization of the molecules within the crystal. For instance, the apo-RXR α LBD, the RXR α LBD:SMRT complex, and the RXR α LBD:agonist:SMRT complex crystallize as tetramers (Gampe et al. 2000; Zhang et al. 2011; Zhang et al. 2011). However, most of the NRs structures solved are homodimers or heterodimers upon crystallization such as the TR β LBD in the active state co-crystallized with GRIP1 coactivator peptide (Darimont et al. 1998), the PPAR γ LBD co-crystallized with SRC1 coactivator peptide (Nolte et al. 1998), and the structure of the heterodimer TR α LBD and RXR α LBD complexed with T3 and 9-cis retinoic acid (Putcha et al. 2012). These LBDs that crystallized as homo and heterodimers shared the same dimerization interface, mainly formed by residues from L9/10, H10, L10/11 and H11.

Chapter 6 - Functional characterization of potentially pathogenic TR α variants

6.1 - Introduction

RTH α is a dominant negative disorder characterized by tissue specific hypothyroidism associated with mutations in the *THRA1* gene. There are a wide range of clinical features related to the disease that could differ depending on the type (missense, nonsense or frame-shift), the position, and the severity of the mutation. The properties of the mutant protein generated by the mutant gene correlate with the severity of the phenotype.

RTH α could be difficult to diagnose from the medical point of view due to the phenotype variability and the absence of specific markers which only include tissue-specific symptoms of hypothyroidism and near-normal circulating levels of T4, T3 and TSH. However, since the identification and characterization of the first TR α LBD mutation in 2012, many more have been identified although the prevalence is still lower than expected (Moran & Chatterjee 2015). The current prevalence of the RTH β , a disorder caused by mutations in the TR β homologous receptor, is approximately 1 in 40,000 (Gurnell et al. 2016), therefore more patients affected by RTH α showing different mutations are anticipated.

All the TR α mutations have been identified by exome sequencing from patients that exhibited tissue specific symptoms of hypothyroidism. Clinical exome sequencing is used routinely to identify single nucleotide variants (synonymous, missense and nonsense variations) and single insertions or deletions that lead to frame-shift variations in disease-related genes. However, a proper functional characterization of the variants identified is required to avoid diagnostic uncertainty or misclassification. An alternative approach to confirm the pathogenicity of the variants involves the comparison of the new variant identified to the “reference standard” sequence of the gene of interest. Due to the high diversity of human genome which contains an average of one variant every eight nucleotides of the exome, there are many different sequences of the same gene

that would generate a functional protein so the “reference standard” sequence is difficult to define. Consequently, this approach would require the comparison of the new variants identified with adequate number of sequences from unaffected people which should include all the polymorphic functional variants found in the population. An extensive knowledge of genetic variation and a database that includes a significant number of genetic sequences are essential to provide references that enable the comparison of the variants observed in patients carrying rare Mendelian diseases (Bamshad et al. 2011; MacArthur et al. 2012).

i) Broad Institute Exome Aggregation Consortium (ExAC)

In an effort to provide these references, MacArthur and colleagues have generated the largest catalogue so far of human variation in protein-coding regions called the Exome Aggregation Consortium (ExAC) by collecting high-throughput DNA sequencing results (Lek et al. 2016). This directory of human genetic diversity aggregates high-quality sequence data from 60,706 human exomes and is publicly accessible in the database <http://exac.broadinstitute.org>. Each genome introduced in the ExAC has been confirmed to be of high enough quality to be confidently considered in clinical assays (Lek et al. 2016).

The study of the sequences found in the ExAC promotes the discovery of genes involved in rare diseases and the possible mutations or changes in amino acids involved. Pathogenic variants are expected to have very low frequency in the population and therefore in the ExAC database due to negative selection. Since the RTH α has been under diagnosed, the database could be used to find new very low frequency variants of the *THRA1* gene in the population that could be potentially pathogenic but have remained unnoticed.

By doing a search in ExAC, it is possible to generate a list that includes all the variants of a gene of interest present in the database. Then, from a structural point of view, new potentially pathogenic mutations can be selected from this list of variation by analyzing the position and the function of the residues affected as well as the type of variation.

As discussed before, many NR structures have been determined, including the human TR β LBD in complex with coactivator peptides (Darimont et al. 1998), the human TR α LBD in the liganded state (Nascimento et al. 2006) and as a heterodimer with the RXR α LBD (Putchá et al. 2012). There are other NR structures solved with corepressor peptides such as PPAR α LBD in complex with an antagonist and a SMRT corepressor peptide (Xu et al. 2002). Due to the high degree of structural similarity exhibited by the majority of the NRs, the function and importance of specific residues in the TR α LBD could be anticipated by comparing the TR α LBD sequence to the other LBDs sequences. Consequently, the TR α residues involved in T3 binding, coactivator and corepressor interaction, and dimerization can be predicted. Alteration of these residues could lead to RTH α . This way, an ExAC search coupled with structural analysis can be used to understand why individuals carrying mutations have an impaired response to T3 and may help to diagnose RTH α .

The main aim of the research described here is to functionally characterize potentially pathogenic variants of TR α selected by structural modeling. First at all, the *THRA1* gene was introduced in the ExAC database in order to identify all the different variations of the gene found in the population. Next, potentially pathogenic variants were selected from a structural point of view, and finally, a functional characterization of every variant was carried out by transactivation and dominant negative assays in mammalian cells in order to discern whether these variants were really pathogenic.

6.2 - Identification of novel pathogenic human TR α LBD variants by searching the ExAC

First of all, a search in the ExAC for the *THRA1* gene was performed with the aim of finding low frequency non-previously identified variants in the *THRA1* gene that can potentially cause the RTH α disorder. The searching results show that the ExAC database contains 834 different *THRA1* sequences which is a significant high number compared to the variability found in other NR genes such as *THRB* that had 410 variants, *PPARA* with 529 variants and *RXRA* with 551 variants. The genetic variation corresponding to the *THRA1* sequences includes synonymous, nonsense, missense and loss of function (LoF) variants as well as copy number variants (CNVs). The LoF variation includes nonsense, splice acceptor, and splice donor variants caused by single nucleotide changes. The results of the search showed 76 synonymous variants, 97 non-synonymous variants, 5 LoF variants and 2 CNVs. From the 97 non-synonymous variants found in the *THRA1*, including nonsense and missense variants, 63 belonged to the DBD, 38 were localized in the LBD, and 52 variants were common to TR α 1 and α 2 (**Table 6.1**).

Position	Frequency	N ^o of alleles	TR β equivalent mutation
E148K	0.000008248	1	
R152Q	0.000008248	1	
Q156R	0.000008248	1	
P158S	0.000008250	1	
E159Q	0.000008250	1	
D166G	0.00003302	4	
I170V	0.0004874	59	
A171S	0.000008248	1	
R188K	0.000008248	1	
P193S	0.000008248	1	P274L
M204V	0.000008248	1	
I222V	0.00001648	2	I276L/N
A225V	0.000008248	1	
M238L	0.00002473	3	
S240T	0.000008248	1	
E241K	0.000008248	1	
I249V	0.000008248	1	
A264G	0.000008248	1	
V265A	0.000008248	1	
K283R	0.000008248	1	
V294I	0.000008248	1	V348E
T314P	0.00001652	2	
T327A	0.00001681	2	
S330L	0.000008248	1	
S330X	0.000008248	1	
D336G	0.000008248	1	
E339Q	0.000008248	1	
A344V	0.00006591	8	
V353I	0.00002472	3	
R356C	0.000008238	1	
H358P	0.000008248	1	
L367P	0.000008248	1	
M369L	0.000008248	1	
E395K	0.000008248	1	E449X
P399Q	0.000008248	1	P453A/H/S/T/L

Table 6.1: LBD TR α variants found in the ExAC database. Table indicates the position of the variation within the protein (residue number), frequency of the mutated allele (number of people found carrying the variation divided by the total number of individuals sequenced), number of alleles found in the database carrying the variation and, if is applicable, the equivalent mutation in TR β already recognized to mediate RTH β .

The ExAC searching results also show the number of observed variants with respect to the number of expected variants and calculate a Z score for synonymous, missense variations, and CNVs. The Z score value, which represents the deviation of observed counts from the expected number, is positive for both synonymous and missense variations indicating that the gene has fewer variants than expected and suggesting that the *THRA1* gene is intolerant to this type of variation. CNV represents deletions and duplications and the Z score value is also positive showing a *THRA1* intolerance to permit these types of variations. The pLI is the probability that a given gene is extremely intolerant to LoF variation and is calculated using the observed and the expected variants counts. The *THRA1* gene is intolerant to LoF variation as well, as indicated by a pLI higher than 0.1 (Table 6.2).

Constraint from ExAC	Expected n° variants	Observed n° variants	Constraint metric
Synonymous	104.4	76	Z = 1.72
Missense	221.3	90	Z = 4.32
LoF	19	5	pLI = 0.17
CNV	5.5	2	Z = 0.62

Table 6.2: Output of the *THRA1* gene searching in ExAC database. Z-score indicates the deviation of observed number of variants from the expected number. Positive Z scores indicate increased constraint (intolerance to variation) and therefore that the gene has fewer variants than expected. The probability of intolerance to LoF variation (pLI) is calculated by the number of observed and expected LoF variants.

Some interesting conclusions could be drawn from this initial search. Firstly, *THRA1* gene exhibits a great deal of variability within its sequence compared to *THRB* or *PPARA*. However, non-synonymous variations are not well tolerated, and almost twice the number of the missense changes are located in the DBD (63 missense variants found in the DBD, while only 38 are found in the LBD) suggesting that this domain more readily accepts changes in its sequence. In contrast, no mutations in the DBD have been reported so far involved in RTH, either in RTH α or RTH β . This could be due to the fact that DNA binding ability of the receptors is essential to repress T3 target gene transcription and, therefore, to have dominant negative activity over the WT receptor also present in the cell. Lack of dominant negative activity allows variation

accumulation in the DBD sequence of the population hidden in heterozygous individuals.

Interestingly, none of the mutations previously reported as pathogenic in *TRHAI* gene were found in this database. As stated before, deleterious mutations are expected to have low allele frequencies due to negative selection; it is therefore not particularly surprising that mutated *THRAI* sequences have such low frequency that do not appear in the database.

Potentially pathogenic variants were selected from the structural point of view taking into account the following characteristics:

- Position and function of the residue within the WT receptor: degree of involvement of the residue in ligand binding, coregulator proteins binding and recruitment, interactions with other residues within the protein, and interaction with residues from other NRs.
- Type of non-synonymous mutation: missense, nonsense.
- Presence of already reported homologous mutation in TR β
- Biochemical features of the new residue compared to the WT residue (degree of similarity between the WT residue and the new residue).

Based on a structural and biochemical characterization of all the variants located in the TR α LBD, the 15 potentially disease causing variations that follow were selected (**Table 6.3, Figure 6.1**).

Position of the variant	Possible consequences of the change
D166G	Significantly different residues
P193S	Homologous mutation in TR β (P247L)
A214V	Homologous mutation in TR β (A268D/G) Valine's side chain is significantly more bulky than alanine
I222V	Homologous mutation in TR β (I276L/N) I222 forms part of the binding cavity and the more bulky side chain of valine could interfere sterically in the cavity structure
V294I	Homologous mutation in TR β (V384E)
T314P	Significantly different residues: proline could introduce a turn and change the structure of the protein.
S330L	Significantly different residues: leucine residues are not usually in solvent accessible parts of the protein.
S330X	Truncated protein without H12 function
A334V	Significantly different residues: valine's side chain much bulky than alanine's
V353I	Significantly different residues: isoleucine's side chain is more bulky and could interfere sterically in the inner of the protein displacing other residues.
R356C	Significantly different residues Arginine 356 makes many polar interactions and contribute to stabilize the protein joining the loop between H9 and H10 with the loop between H8 and H7
H358P	Significantly different residues Histidine 358 makes many polar interactions and contribute to stabilize the protein joining the loop between H9 and H10 with the loop between H8 and H7
L367P	Significantly different residues Involved in the dimerization domain
M369L	Involved in the dimerization domain
P399Q	Significantly different residues Located at the beginning of the H12 possibly contributing to the active conformation of it in response to T3

Table 6.3: TR α variants selected for the functional analysis. Table indicates a short summary of reasons to justify the decision.

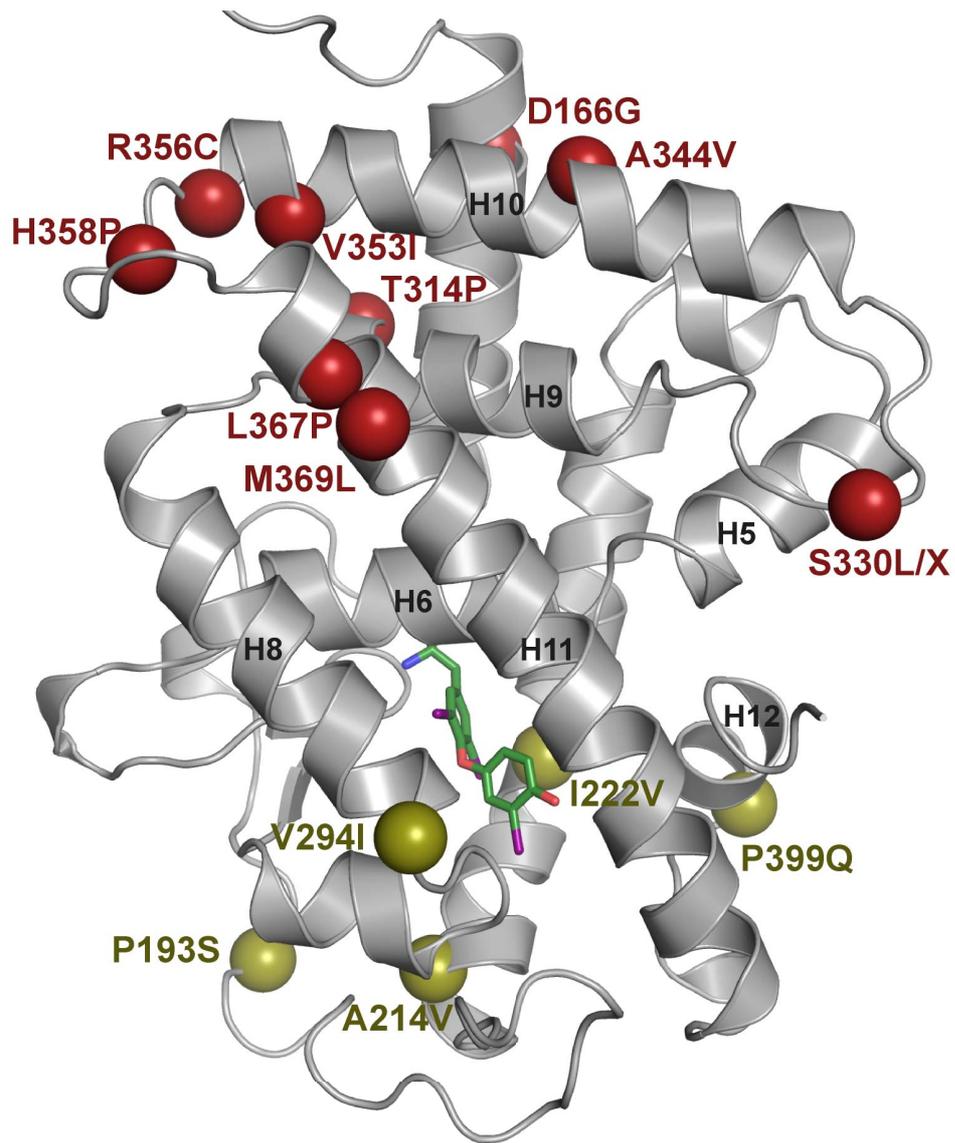


Figure 6.1: Crystallographic structure of TR α LBD showing the position of selected variants in red (new ones identified in the database) and yellow (the ones with an equivalent mutation in TR β) spheres.

6.2.1 - Potentially pathogenic human TR α LBD variants functional analysis

Functional assays were performed in mammalian cells in order to assess the effects of the variations on the transcriptional activity of the TR α .

Site directed mutagenesis was performed following the instructions presented in 2.2.3 of materials and methods and using full-length WT human TR α as a template. Afterwards, the constructs were cloned into the pLEICS12 mammalian expression vector and taken to the Metabolic Research Laboratories (University of Cambridge) to perform the functional assays under the supervision of Dr. Maura Agostini and Prof. Krishna Chatterjee.

Two different types of functional assays were performed: T3 transactivation and dominant negative assays. The T3 transactivation assays were accomplished by transient transfection of the different variants together with a reporter gene (luciferase gene in the construct MAL-TKLUC) located after a TRE. Dominant activity potency of each variant was analyzed by transient cotransfection of the WT and every variant in JEG-3 cells using the same reporter gene. In this case, the measured fluorescence corresponds to the T3 transactivation properties of both receptors, the WT and the variant cotransfected with it (**Figure 6.2**). These experiments were performed following the instructions described in 2.10 and 2.11 materials and methods.

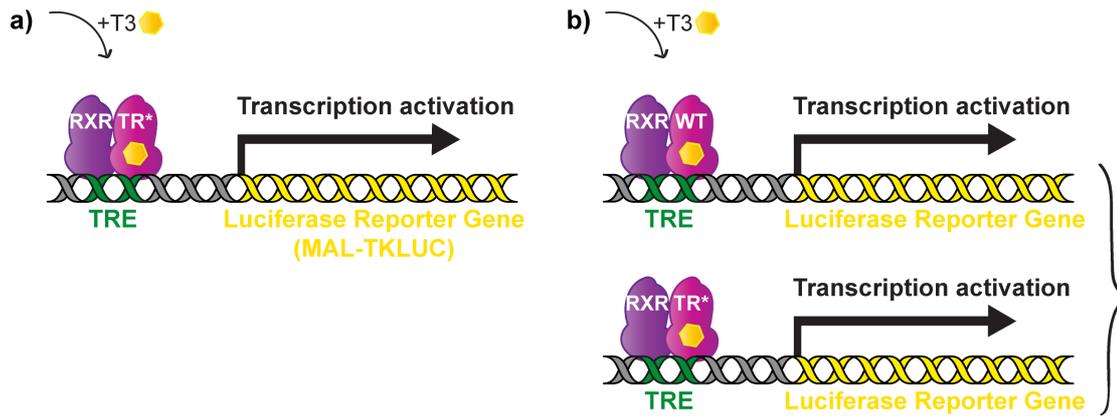


Figure 6.2: Scheme of the functional assays performed in mammalian cells. a) T3-dependent transcriptional activation by WT and TR α variants. JEG-3 cells were cotransfected with empty vector, WT or mutant TR α expression vectors together with the reporter construct MAL-TKLUC. b) Dominant negative assay to find out the inhibition of WT activity by TR α variants. JEG-3 cells were cotransfected with the WT expression vector together with the empty vector (pcDNA3) or one of the TR α variants and MAL-TKLUC.

The different variants were classified according to the results in potentially pathogenic, non-pathogenic and possibly pathogenic variants.

i) Potentially pathogenic variants

According to the results, the S330X, L367P and P399Q variants are most likely disease causing variants.

S330X variant showed negligible activation of the reporter gene in response to any of the T3 concentrations studied. When coexpressed, the mutant S330X suppressed between 35-40 % of WT TR α function in a dominant negative manner at T3 concentrations between 10 to 1,000 nM (**Figure 6.3**). Therefore, S330X protein was transcriptionally inactive even at the highest concentration of T3 and inhibited the action of the WT counterpart in a dominant negative manner.

The L367P and P399Q variants showed deleterious effects on the function of TR α (**Figure 6.3**). The transcriptional response of these variants to T3 showed a right-shifted transcriptional activation profile, meaning that they required between 10 and 100 times more T3 to activate the reporter gene transcription. However, these variants never reached the maximal transcriptional response achieved by the WT. L367P variant

showed a maximal transcriptional response of approximately 65 % compared to the WT. P399Q variant maximal transcriptional response was lower, reaching approximately the 58 % of the WT maximal activity. These variants also exhibited dominant negative activity, inhibiting approximately between 35-40 % of WT transcriptional activity even at the highest T3 concentrations studied (**Figure 6.3**).

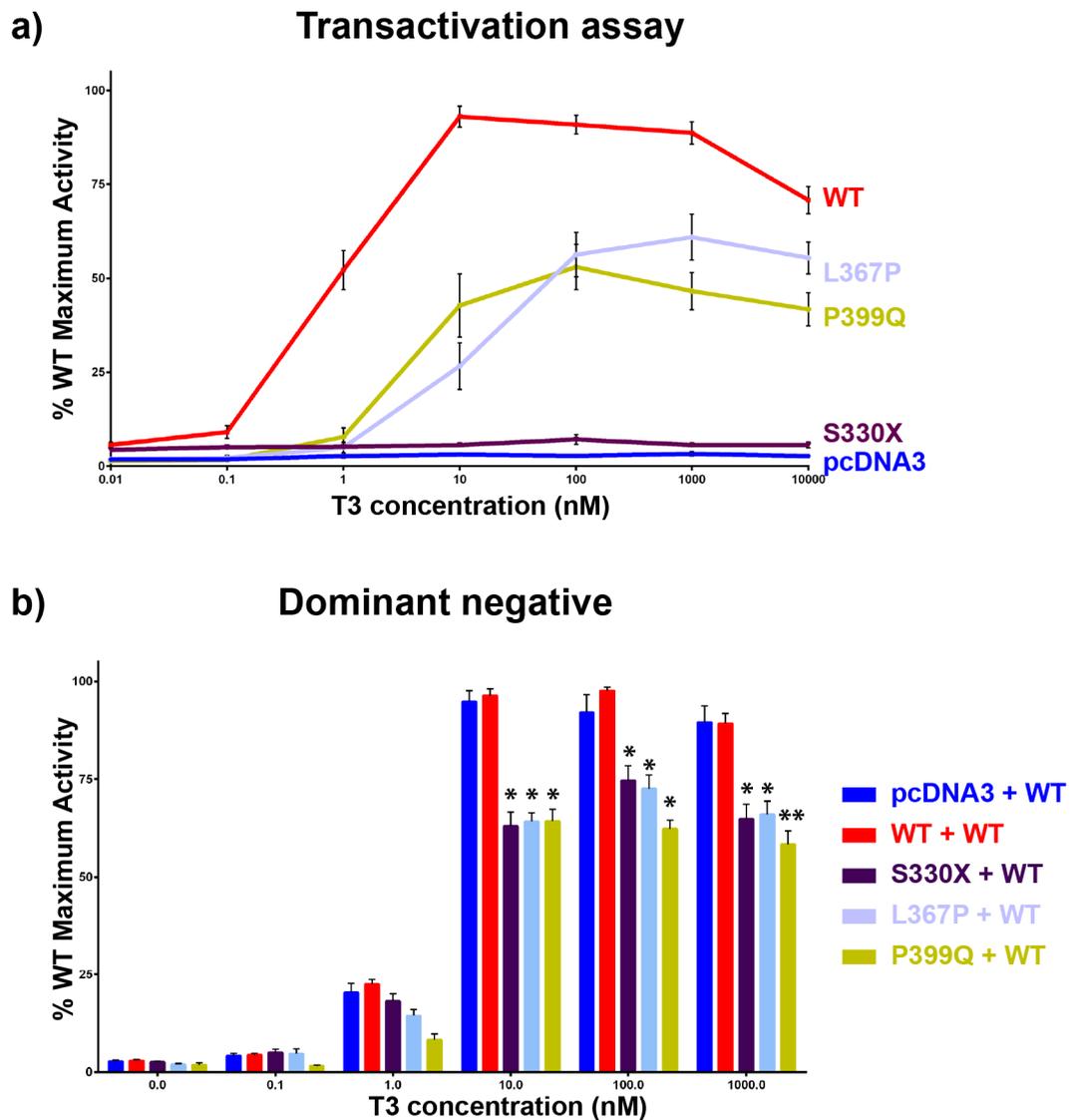


Figure 6.3: T3-dependent transcriptional activation and dominant negative results of the S330X, L367P and P399Q variants. Transcriptional activation in response to increasing amounts of T3 was normalized against the Bos- β gal assay and expressed as a percentage of the maximum WT receptor response. The data shown represent the mean \pm SEM of at least five independent experiments. Significance (T-student): * $p < 0.05$, ** $p < 0.01$.

ii) Non-pathogenic variants

In marked contrast, the biological activity of the missense variants D166G, P193S, A214V, I222V, V394I, S330L, V353I and T314P was comparable to that of WT TR α at most T3 concentrations. Coexpression of WT and these variants did not significantly change the function of WT TR α (Figure 6.4). Therefore, these variants could be considered as benign polymorphisms.

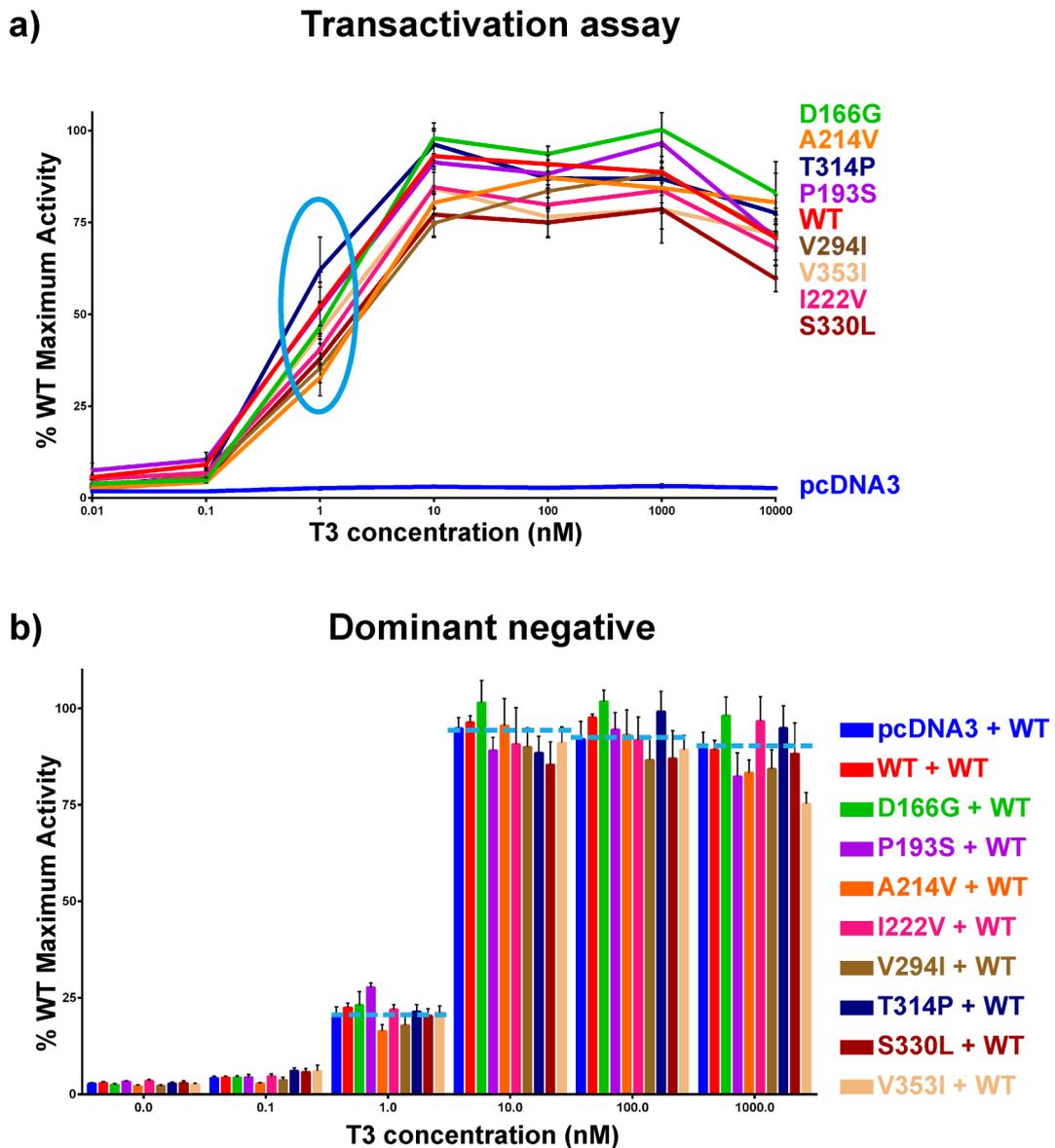


Figure 6.4: T3-dependent transcriptional activation and dominant negative activity of the D166G, P193S, A214V, I222V, V394I, S330L, V353I and T314P variants. Transcriptional activation in response to increasing amounts of T3 was normalized against the Bos- β gal assay and expressed as a percentage of the maximum WT receptor response. The data shown represent the mean \pm SEM of at least five independent experiments.

Apart from the variants D166G, P193S and T314P, the remaining showed between 5-20 % lower transcriptional activation values than the WT TR α at every T3 concentration investigated. Nevertheless, the activation profile was the same as the WT TR α and no significant dominant negative effect was found (**Figure 6.4**).

iii) Possibly non-pathogenic variants

Those with the same profile of transcriptional activation as the WT TR α , showing similar transcriptional response at every concentration of T3, but with a noticeable level of dominant negative activity were classified as probably benign variants. Variants A344V, R356C, H358P and M369L belong to this group and are not probably causing a disease even though they did not reach the maximal WT transcriptional activity at any T3 concentration studied and exerted moderate dominant negative activity (**Figure 6.5**).

The variants R356C and H358P showed normal transcriptional activation profiles and non-significant dominant negative activity at any concentration of T3 studied. In contrast, variants A344V and M369L showed significant dominant negative activity of approximately 25 % at 10 nM and 10,000 nM, respectively. This dominant negative activity is non-significant at different values of T3 (**Figure 6.5**).

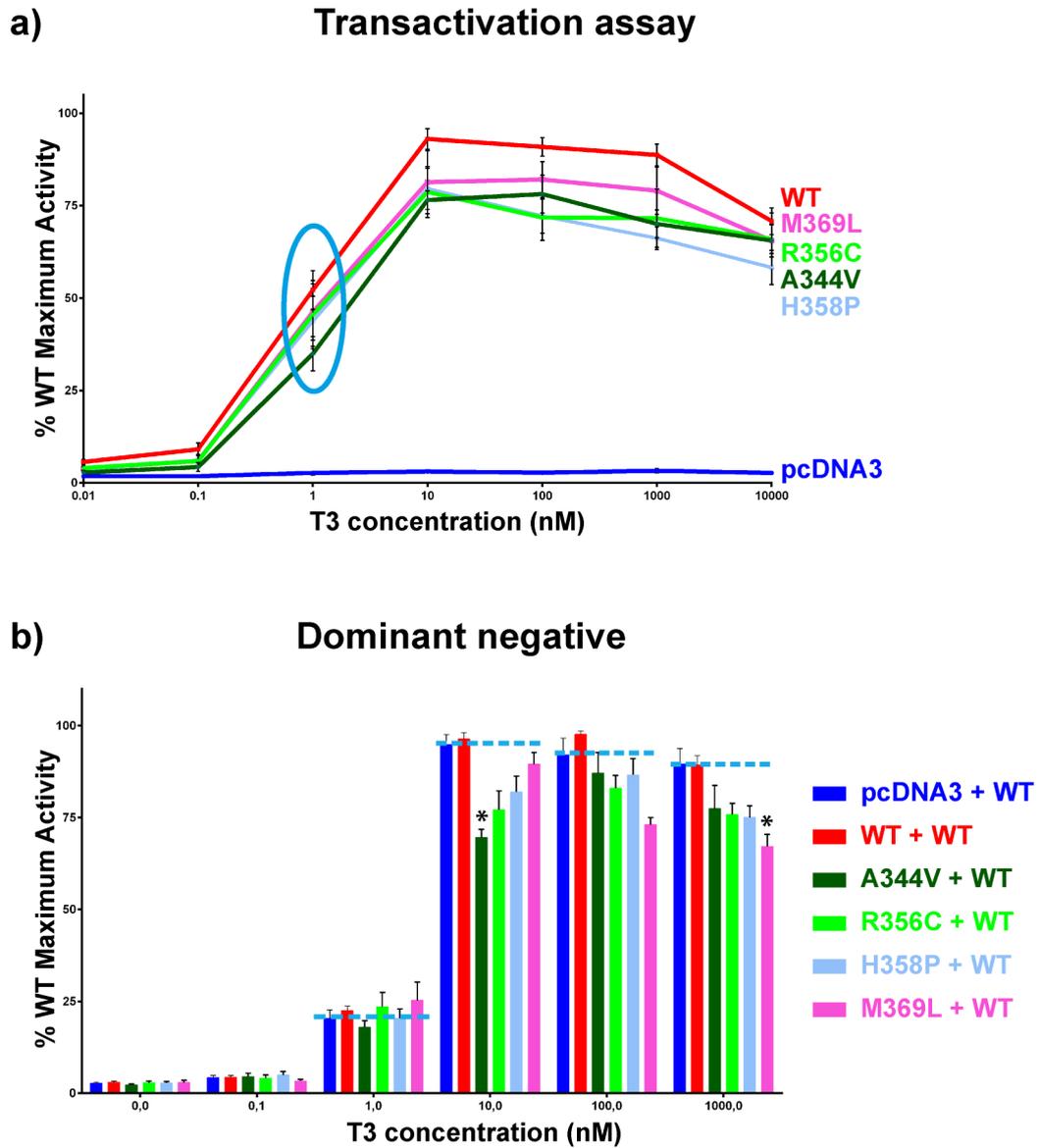


Figure 6.5: T3-dependent transcriptional activity and dominant negative results of the A344V, R356C, H358P and M369L variants. Transcriptional activation in response to increasing amounts of T3 was normalized against the Bos- β gal assay and expressed as a percentage of the maximum WT receptor response. The data shown represent the mean \pm SEM of at least five independent experiments. Significance (T-student): * $p < 0.05$.

6.2.2 - Expression of Flag-tagged TR α variants

Western blot analysis were carried out in order to confirm that the potentially disease causing variants were expressed correctly in the cells following the procedure described in 2.12 of materials and methods.

JEG-3 cells were transiently cotransfected with the WT TR α and the potentially disease causing TR α variants S330X, L397P and P399Q. All the constructs were made to contain a FLAG tag attached to the N-terminal of the protein. After 48 hours of incubation, cells were lysed and the presence of TR α was detected by FLAG antibodies.

The results indicated that all the constructs corresponding to the potentially disease causing variants were appropriately expressed in the JEG-3 mammalian cells. The expression of the TR α (Mw: 49.6 kDa) in the cotransfected cells (WT+WT, WT+L367P and WT+P399Q) was approximately double of the expression of the WT TR α in cells transfected by the construct alone indicating that the L367P and the P399Q constructs were expressed correctly. The S330X (Mw: 40 kDa) also was expressed by the cells as indicated by the presence of a lower band in the gel (**Figure 6.6**).

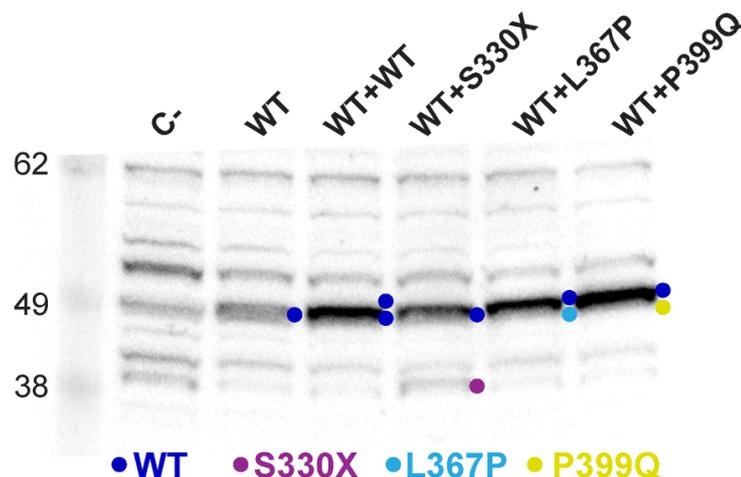


Figure 6.6: Expression of N-terminal flag-tagged TR α variants. The figure shows an anti-flag western blot representing the expression of the N-terminal flag-tagged WT, S330X, L367P and P399Q TR α .

6.3 - Discussion

In this chapter it has been demonstrated that determining the pathogenicity of non-synonymous variants incidentally found in the population by functional assays is a necessary step after identifying them by exome sequencing and before diagnosing them as mutations involved in a genetic disease. Since *THRA1 LBD* exome contains 38 non-synonymous rare variants and only 3 of them are in fact pathogenic validated by single-variant assays, functional characterization of every non-synonymous variant becomes essential to avoid diagnostic uncertainty and misdiagnosis (when a benign variant is presumed pathogenic).

An initial selection of the variants with more probability to cause a perturbation in the protein structure and/or function was made from the *THRA1* variants found in the ExAC database. Based on this first structural examination, 15 variants were chosen to perform transcriptional activation and dominant negative assays.

i) S330X TR α as mutation probably associated with RTH α

S330X variant should be considered a disease causing mutation. The transcriptional activation profile and the dominant negative activity coincide with the ones showed by the initial patients discovered who had frame-shift mutations with a prematurely introduced stop codon (A382PfsX7 and F397fs406X) or a nonsense mutation (E403X), all of them generating prematurely truncated proteins (Bochukova et al. 2012; van Mullem et al. 2012; Moran et al. 2013). The phenotype of the disorder caused by S330X is probably as severe as in these patients.

The results imply that the S330X variant generates a functional protein that is able to inhibit the WT transcriptional activity in a dominant negative manner. The truncated protein generated ends at residue Arg 329, consequently this protein completely lacks H10, H11 and H12 which means that the protein lacks the heterodimerization domain that expand along the L9/10, H10, L10/11 and the beginning of H11 in TR α LBD (Putchá et al. 2012). Dimerization domain appears to be required for the mutant TR β receptors to exert dominant negative activity (Nagaya & Jameson 1993; Kitajima et al. 1995) (reviewed in (Gurnell et al. 2016)). The truncated protein also lacks the activation

function 2 (AF-2) mediated by H12 which explains the total absence of response to T3. The results also imply that the truncated S330X protein is able to bind corepressor proteins with sufficient stability to repress T3 target gene transcription. It is probably that the truncated protein is able to recruit corepressor peptides since the corepressor interface of the LBD formed by H3 and H5 is still intact.

ii) L367P TR α as probably mutation associated with RTH α

L367P variant was also found to be potentially pathogenic. L367 residue in TR α belongs to the beginning of H11 and is pointing to the inside of the protein making several van der Waals contacts with others hydrophobic residues from L8/9 (Phe 309 and Leu 311), H10 (Phe 349) and H11 (Phe 363, Trp 364, and Val 371) (**Figure 6.7**).

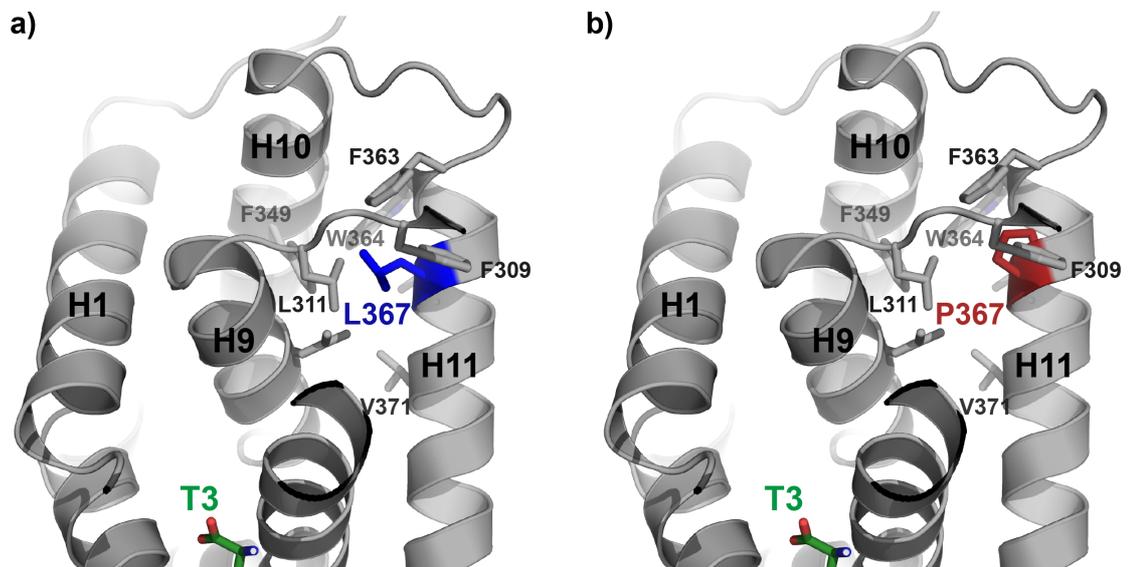


Figure 6.7: Position and function of the L369. a) Location of L367 in the WT LBD TR α and the surrounding residues. b) L367P mutation made by PYMOL to show the position of the new residue in the WT LBD TR α . No structural distortion is generated by the program, but the proline would introduce a turn in that position and disrupt the structure of H11.

Proline residues do not naturally form α -helices or β -sheets and normally introduces a turn in the structure of the protein. The transcriptional activity profile showed that the L367P variant required 100-fold higher concentration of T3 than that by WT to activate the transcription of the reporter gene, suggesting that L367P has lower affinity to T3. The change from a leucine to a proline might lead to major conformational disruptions that modify the binding cavity and thus, the interaction with T3. In addition, the protein

does not reach the maximum transcriptional activity of the WT receptor suggesting further structural modifications that may prevent the protein from recruiting coactivators. The dominant negative activity suggests that the L367P receptor still has the corepressor and dimerization domain completely functional.

iii) P399Q TR α as probably mutation associated with RTH α

The last variant that is probably pathogenic is P399Q. Proline 399 is located in a key position of the protein and is responsible for the turn introduced in the protein in response to T3 that places the H12 in the active conformation. The TR β P453H mutant, which is homologous to the P399Q, exhibited about 10-fold reduced affinity to T3, partial T3 responsiveness (approximately 30 % of WT activity after incubation with 5 nM T3) at mediating activation and repression of the reporter genes and a dominant negative effect of 30 % when cotransfecting with the WT (Chatterjee et al. 1991). These results are consistent with the results reported here which showed an approximately 39 % of transcriptional activation and 60 % transcriptional activity when cotransfecting with WT at 10 nM T3. The authors attributed the functional impairment of P453H receptor to its partial binding to T3 (Chatterjee et al. 1991). Since the WT phenotype is not completely recovered even at significantly high concentrations of T3 and there is significant inhibition of the WT counterpart activity, few structural perturbations might also be responsible for the functional impairment in addition to the attenuated binding to T3.

Potential reasons for the marked impairment in the P399Q variant functional properties could be investigated by modeling the effect of the amino acid change. There are two prolines at the beginning of H12 that play a structural role forming two successive turns that bend the molecule almost 90 ° making H12 change to the active conformation in response to T3. Proline is a residue which cannot make the hydrogen bonds needed to fold an α -helix or a β -sheet conformation and the proline-proline sequence at the C-terminus of the receptor may result in a unique conformation. The predicted secondary structure of P399Q LBD is more likely to form a stretch of α -helices from amino acid positions His 358 to Leu 402 (from the beginning of H11 to the end of H12) compared to the WT receptor with a proline-proline sequence at codon 398 and 399 (**Figure 6.8**).

Although P399 is notably required to place H12 in the active conformation, functional results show that the impairment could be partially rescued with higher concentrations of T3 because, even damaged, H12 remains there.

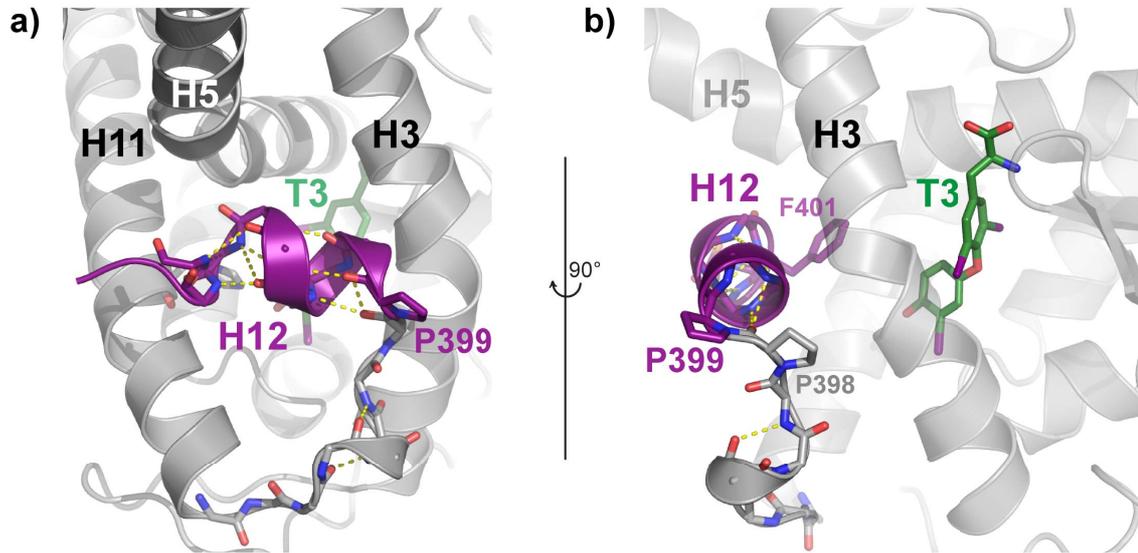


Figure 6.8: Position and function of the P399. a) Front view of the WT LBD TR α (pdb code: 2H79) structure highlighting the P399 position and the 90 ° turn made by the two prolines located at the beginning of H12 (purple). It is also shown the hydrogen bonds made by the residues forming H12 to assemble the α -helical structure in response to T3. b) Side view of the WT LBD TR α showing the essential structural turns made by the P398 and P399 in order to place H12 in the active conformation.

Except for these three probably disease causing variants, the majority of the non-synonymous variants found in ExAC database are probably only polymorphisms even though most of them are uncommon in the population. Considering that these variants are present in the population as heterozygous and, therefore, are sharing their function with a totally functional WT receptor, it can be concluded that the non-pathogenic variants are benign polymorphic substitutions that could explain the variability found in the population involved in metabolism, development or response to stimuli. Since their prevalence appears to be low or very low among the sequences found in the ExAC database, it could be interesting to study if they are associated to other disorders or the propensity to suffer other disorders. It is also noticeable that four of them (P193S, A214V, I222V and V294I) were studied because homologous mutations in TR β had been already reported as being involved in RTH β . This discrepancy raises the question whether the activity of these two homologous receptors is as similar as thought and

remarks the different function that the TR isoforms carry out in the specific tissues where they are expressed.

The variants classified as probably non-pathogenic are most likely completely functional at T3 cellular levels even though they did not reach the maximal WT activity and exerted some dominant negative activity at the highest concentration of T3. That is why they have not been selected against and consequently, have been transmitted through generations up to now and are contributing to the pool of variation found among the population in metabolic and developmental rate.

Interestingly, the extent of thyroid dysfunction *in vivo* seems consistent with the magnitude of receptor impairment *in vitro* (Gurnell et al. 1999). However, due to the marked divergence in phenotypes found in the RTH α and RTH β affected individuals, it is difficult to predict the possible phenotype associated with the functional properties found in the different variants. It is not possible either to classify the non-conclusive variants as benign polymorphism since they could lead to middle phenotypes of the RTH α disorder.

Finally, it is necessary to highlight the importance of a very early identification and diagnosis of the RTH α to prevent neurodevelopmental defects. Since patients often display residual sensitivity to thyroid hormones due to heterozygote status and the presence of WT TR α , an early T3 treatment could thus be beneficial.

Chapter 7 - Discussion and future perspectives

The work in this thesis was towards three major goals. Firstly, to investigate the molecular mechanism that underlies the pathology of three independent mutant TR α LBDs as well as to examine the structural basis for repression by the mutant receptors. Secondly, to explore diverse approaches as potential therapies to treat the disease caused by the mutant TR α . And thirdly, to carry out functional studies on new potentially pathogenic TR α variants found in the population.

7.1 - Molecular pathology of the mutant TR α LBDs

Previous functional studies in mammalian cells indicated that mutant receptors do not activate transcription in response to T3 and, in a cotransfection experiment, they inhibit the activity of the WT equivalent in a dominant negative manner (Bochukova et al. 2012; van Mullem et al. 2012; Moran et al. 2013). These results suggested that the mutant receptors are constitutively bound to corepressor complexes and therefore, constitutively repress T3 target gene transcription.

This study is focused on exploring the ability of WT and mutant LBDs to interact with their coactivator and corepressor in the presence and in the absence of T3. These experiments showed that mutant LBDs interact with corepressor peptides with significant affinity (nanomolar range) in the presence and in the absence of T3 confirming that mutant LBDs are constitutively bound to corepressor peptides. However, mutant LBDs are not able to interact with coactivator peptides even in the presence of T3. Interestingly, the strength of the interaction of the mutant LBDs to the corepressor peptides is similar to that of the WT LBD, suggesting that the constitutive binding of corepressors to the mutant LBDs is not a consequence of a higher binding affinity, but a consequence of the mutant LBDs failure to recruit coactivators in response to T3. The most likely reason is the lack of H12 function in the mutant receptors, which is absolutely necessary for coactivator binding (Nolte et al. 1998; Darimont et al. 1998).

Most mutations in the TR β LBD showed impaired ability to bind T3 and because of that, impaired transactivation (reviewed in (Gurnell et al. 2016)). Circular dichroism experiments were performed in order to examine the mutant TR α LBDs ability to bind T3. The results suggest that all the mutant TR α LBDs are able to interact with T3. This T3 binding capacity of the mutant LBDs was confirmed at structural level through the determination of the structure of a construct based on the mutant LBDs, P393GT394X LBD. The structure clearly revealed a molecule of T3 accommodated in the binding pocket of the LBD. The structural and chemical characteristics of the cavity remained the same in the shorter LBD as in the WT LBD. There was, however, one remarkable difference between the P393GT394X LBD and the WT LBD: a sudden 90° turn of H11 at Ala 379 that disrupts the helix and becomes a disordered coil. The position of this tail in the crystal structure is probably due to the packing of the molecules within the crystal. In solution, it is more probably that the C-terminal tail behaves as a dynamic and mobile structure independent of the rest of the protein. Hence, these experiments provide novel insight into the structural basis of the molecular pathology of mutant TR α .

7.2 - Structural basis of TR α mechanism of repression

Many structures of NRs have been determined so far providing insight in the molecular mechanisms of NRs action, including full-length receptors bound to their cognate DNA target element (Chandra et al. 2008; Chandra et al. 2013; Lou et al. 2014). Understanding the interaction between NRs and their coregulator proteins is essential to identify the mechanisms involved in activation or repression of the receptor itself, or in other words, how the binding of ligand translates into regulation of transcription. The interaction between NRs and coactivators is well understood since there are numerous structural studies of NR LBDs in the active conformation interacting with coactivator peptides, including the TR β LBD in complex with GRIP1 coactivator peptide (Darimont et al. 1998). So far, unfortunately, only two structures of complexes between ligand-free receptors and corepressor peptides have been solved: the apo-Rev-erbA- α bound to NCoR corepressor peptide and the apo-RXR α interacting with SMRT corepressor peptide (Phelan et al. 2010; Zhang et al. 2011). The structure of apo-Rev-erbA- α is unique and the mode of corepressor binding differs from the previous

determined by other NR structures bound to corepressors. Therefore, the outstanding question is whether these mechanisms, established for RXR α and Rev-erb- α , applies to other NRs, especially since NRs are rather diverse in their mechanism of action.

To better understand the molecular features of repression by TR α or the mutant TR α , which function as constitutive repressors, attempts were made to try to co-crystallize apo-TR α LBDs with SMRT peptides. Unfortunately, no crystals were obtained. Co-crystallization using a ternary complex formed by mutant LBDs, T3 or T3 analogues and native or stapled corepressor peptides produced crystals. But these crystals did not diffract. Determination of the crystal structure of a shorter construct bound to T3 revealed that the crystal packing of the molecules was too tight to permit corepressor binding, frustrating all attempts to obtain crystals from the ternary complex. Co-crystallization trials with even shorter constructs might overcome this difficulty.

Other approaches could contribute to an understanding of the dynamics of H12 in apo-TR α LBD or the C-terminal tail of mutant TR α LBDs such as hydrogen/deuterium exchange mass spectrometry (H/D ex MS) studies or fluorescence spectroscopic analysis. These studies have been successfully used to characterize the dynamic properties of H12 in the inactive state of RXR α LBD (Yan et al. 2004), PPAR γ LBD (Kallenberger et al. 2003; Hamuro et al. 2006), and ER α and ER β LBD (Dai et al. 2009).

7.3 - Functional characterization of potentially pathogenic TR α variants

The ExAC database provided an opportunity to find low frequency potentially pathogenic TR α variants present in the population. Since the RTH α disorder has been under diagnosed due to the lack of clear biochemical indicators in patients affected by the disease, these new variants might indeed be disease causing variants. Identification of mutations that actually lead to the RTH α disorder could accelerate the process of diagnosis and provide the opportunity to rapidly treat affected individuals.

From the 38 non-synonymous variations found in the TR α LBD, 15 were selected as potentially pathogenic, taking into account the position and function of the residue affected. Functional analysis using luciferase reporter assays determined that only three variants are non-functional: S330X, L367P, and P399Q. It is likely that these mutations in TR α would be pathogenic and cause the RTH α disorder; however the severity of the phenotype would be difficult to predict.

Several assays could be performed in order to better understand the molecular pathology of S330X, L367P and P399Q variants. Biophysical characterization of these variants could be performed by interaction assays and spectroscopy techniques such as fluorescence anisotropy and circular dichroism. In addition, electrophoretic mobility shift assay could be used to analyze the hetero- and homo- dimerization properties of the different variants.

The approach of analyzing the exome has been recently reported for PPAR γ (Majithia et al. 2016). Since *PPARG* exhibits great variability in the population (approximately 0.2 % of the population carrying a rare variant (Majithia et al. 2014)), classification of newly identified variants requires functional assays to avoid misdiagnosis or diagnosis uncertainty. Majithia and colleagues analyzed 55 new low-frequency missense variants identified by exome sequencing to find that 6 of them resulted in defective proteins confirmed by classical transactivation assays.

The TR α and PPAR γ results highlight the importance of functional characterization assays of the missense variants found routinely by exome sequencing in disease-related genes. Due to the great variability exhibited by the human genome, which shows one variant for every 8 base pairs (Lek et al. 2016), many rare missense variants produce functional proteins and are carried by unaffected people.

7.4 - Stapled peptides and T3 analogues investigation in corepressor and coactivator interaction

The biophysical and structural characterizations demonstrated that mutant TR α LBDs are able to bind T3 and provided an opportunity to design several T3 analogues in order

to promote corepressor release or coactivator recruitment. In addition, stapled peptides were designed and tested for their ability to competitively displace native corepressor from the corepressor-binding site of the mutant TR α LBDs. These two approaches were investigated as potential strategies to treat the disorder caused by the mutant TR α LBDs.

The biophysical results suggest that most of the T3 analogues have relatively weak ability to disrupt corepressor or coactivator interaction with the mutant LBDs. Only the ES08 and ES09 analogues displayed significant ability to promote corepressor dissociation. The structural basis for the action of the T3 analogues was explored by X-ray crystallography in order to design new improved T3 analogues based on the structural information. Several crystals with the construct P393G T394X LBD containing different compounds were successfully obtained. Unfortunately, only the P393G T394X LBD:JM13 crystals diffracted well enough to obtain a data set. The determination of the structure was challenging since the electron density of the compound extension disappeared upon refinement. An explanation is that the 4'OH extension of the compound is mobile and adopts multiple conformations, which means that it is not seen in the electron density.

The biophysical results using the stapled peptides were much more encouraging. The stapled corepressor is able to bind to A382PfsX7 and F397fs406X mutant LBDs with improved affinity compared to the native corepressor and to E403X mutant LBD with similar affinity.

7.5 - TR α as pharmacological target

NRs constitute important therapeutic targets for endocrine, metabolic and circadian disorders. However, the varying roles of even the same receptor in multiple pathways, target tissues, and on individual genes pose limitations in developing effective pharmacological agents that lack unwanted side effects. The efficacy of a treatment relies on the extent of biochemical and structural understanding of the target and the molecular mechanism of the pathology. This information is essential to design a proper therapeutic strategy and to predict the possible side effects.

i) T3 analogues as potential therapeutic agents

As discussed in Chapter 4, new improved T3 analogues could be designed and tested for their ability to bind mutant LBDs and promote corepressor release now that the biophysical and structural basis of the pathology are better understood. The strategy to analyze T3 analogues *in vitro* presented in Chapter 4 constitutes a good first approach to identify the best chemical design of the T3 analogues tested. However, *in vitro* assays have several limitations that compromise the correlation of the results *in vivo*. For instance, in a cell nucleus, the availability of corepressor and coactivator proteins are limited and the NR transcriptional activity can be influenced by both coregulators as well as the spectrum of relative coactivator and corepressor affinities. A T3 analogue might have improved potency in promoting corepressor release *in vivo* if there are coactivator proteins competing to bind the LBD surface. Therefore, extending the T3 analogue studies to cellular gene expression assays in mammalian cells is required to determine the ability of these T3 analogues to modify the transcriptional activity of the mutant LBDs.

Many molecules have been successfully developed for clinical treatment of a variety of disorders involving other nuclear receptors. A commonly prescribed drug for breast cancer treatment is raloxifene, a selective ER antagonist. Raloxifene and estradiol (the cognate ligand for ER) are chemically very similar and directly compete to bind to the ER LBP. However, raloxifene contains a bulky substituent that makes different contacts with the ER LBD. Its side chain makes extensive hydrophobic contacts with H3, H5/6, H11 and the loop between H11 and H12. As a consequence, H12 is displaced by the bulky substituent preventing coactivator binding (Brzozowski et al. 1997; Shiau et al. 1998). PPAR γ is generally known as the target for two thiazolines, pioglitazone and rosiglitazone, which are used clinically to increase insulin sensitivity in type 2 diabetes mellitus (Cariou et al. 2012). Other compounds, called tyrosine agonists, are agonists of PPAR γ function and have been reported to reverse the dominant negative activity associated with mutations in PPAR γ (Agostini et al. 2004)). These compounds are able to stabilize H12 in the active conformation of mutant PPAR γ to displace corepressor and relieve the dominant negative activity of the mutant receptors (Agostini et al. 2004).

In most individuals affected by RTH β , administration of supraphysiological doses of L-T4 or L-T3 compensates the receptor defect in some tissues (Refetoff et al. 1993).

However, this therapy needs careful monitoring to avoid adverse cardiac effects or the excess catabolism associated with TR α cross-activation. In order to overcome the side effects and the poor efficiency, a T3 analogue was developed. Triac has greater affinity, potency, and activity than T3 for TR β and exhibits a higher affinity for TR β than TR α *in vitro* (Schueler et al. 1990; Takeda et al. 1995). Accordingly, Triac activates thyroid response predominantly in pituitary and liver, where TR β is mainly expressed (Bracco et al. 1993). L-T4 or L-T3 treatment in RTH α has been beneficial for improving some symptoms such as constipation (Moran et al. 2013; van Mullem et al. 2013). Other drugs have been reported to ameliorate some phenotypic abnormalities (growth, bone development), such as suberoylanilide hydroxamic acid (SAHA), an inhibitor of histone deacetylase (Kim et al. 2014).

Future therapies include TR α -selective thyromimetics with higher affinity and selective activity for mutant TR α than for normal TR α or TR β . Identification of T3 analogues that preferentially stimulate TR α activity and overcome resistance in TR α -expressing tissues would represent a major therapeutic advance.

ii) Development of stapled peptides to potentially prevent mutant LBDs: corepressor interaction

In order to progress the studies with stapled peptides, *in vivo* studies are required not only to confirm their effectiveness to promote corepressor release from mutant LBDs, but also to study their cellular and nuclear uptake.

Many approaches have been reported to investigate the stapled peptides capacity for cellular penetrance. However, the explicit mechanisms of uptake remain an active area of research. Using a variety of techniques that involved FITC-labeled stapled peptides such as live confocal microscopy, FACS analysis, and fluorescence scan, Walensky and colleagues have evaluated cellular uptake. Cellular penetrance appears to be time- and energy-dependent, consistent with pinocytosis (Walensky et al. 2004). Active transport from the pinosomes to intracellular sites of biological activity such as nucleus also has been observed by microfocal microscopy (Walensky et al. 2004; LaBelle et al. 2012; Edwards et al. 2013). A combination of factors influences cell penetrance ability of stapled peptides such as overall charge, hydrophobicity, and α -helical structure. It has

been reported that stapled peptides with overall charge to 0 to +2 and greater α -helical content can enhance cell permeability (Bird et al. 2008; Bird et al. 2011). Therefore, stapled peptides can be modulated in sequence and structure in order to achieve cellular uptake.

A wide range of cellular and *in vivo* studies using a great variety of stapled peptides have been achieved in the fields of cancer, infectious disease, metabolism, and neuroscience (reviewed in (Walensky & Bird 2014)). The use of stapled peptides as specific regulators of NR biological function constitutes an active area of investigation. Stapled coactivators have been designed to target ER and VDR and inhibit NR-coactivator interaction (Phillips et al. 2011; Demizu et al. 2013). The biophysical characterization of the stapled coactivators showed a marked increase in helicity and in binding potency for the NR binding site of NRs, agreeing with the data reported here. A careful design of the peptides would be essential to assure a specific binding of the peptides to the target NR as well as to promote cell penetrance. Adjusting the flanking residues of the corepressor or coactivator NR interacting motif might provide NR specificity.

Despite its biological implications, little is known about the effects of cellular environment on the interaction and dynamics of proteins. In the nucleus, the interaction between the stapled peptide and the hydrophobic surface of NRs will be affected by a variety of factors, including competition with the whole corepressor complex (**Figure 7.1**).

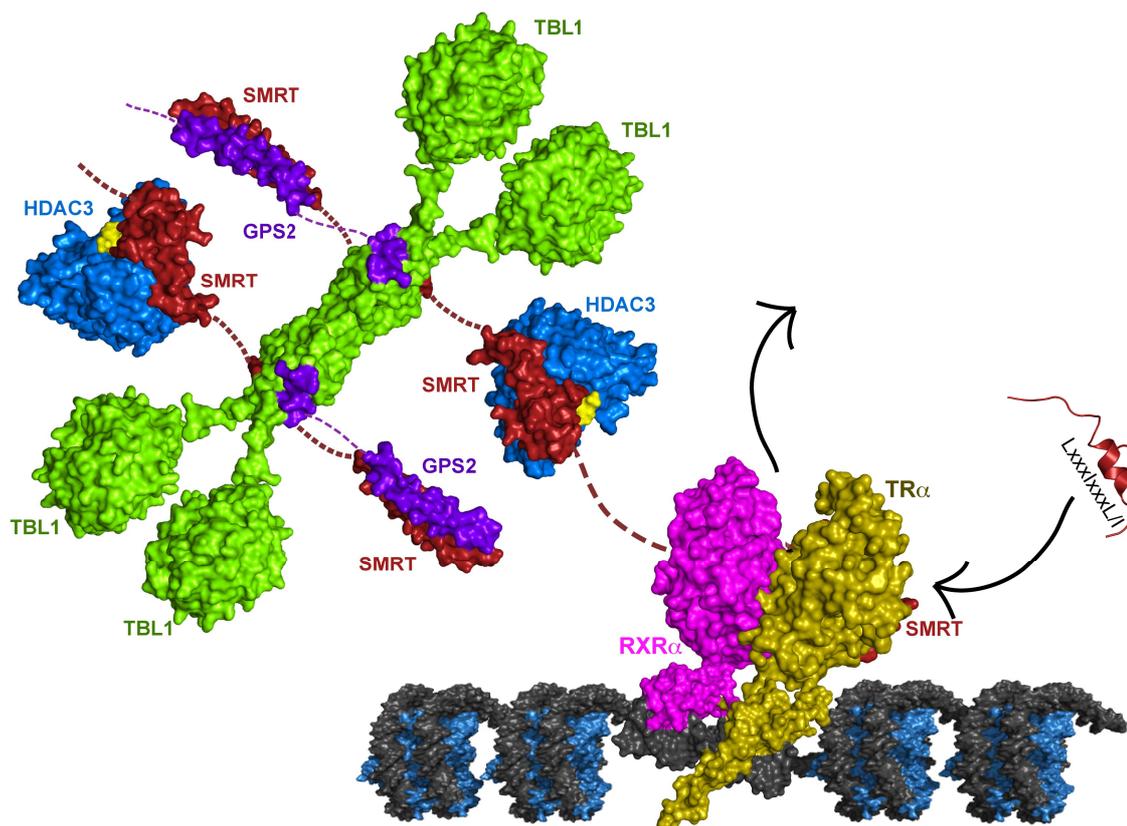


Figure 7.1: Representation of the SMRT corepressor complex interacting with RXR α - TR α heterodimer bound to chromatin. The SMRT corepressor interacts with RXR α -TR α heterodimer in the inactive state and mediates the transcriptional repression of T3 target genes. SMRT corepressor is a component of the SMRT corepressor complex which is formed by TBL1 (green), GPS2 (purple), HDAC3 (blue) and SMRT (red) (Oberoi et al, 2011). The stapled SMRT peptide is expected to interact with the hydrophobic surface of NRs with greater affinity and displace the native SMRT from the NR surface. Pdb codes: SMRT/GPS2 interaction surface 2L5G, TBL1 WD40 4LG9, TBL1 tetramerization domain 2XTC, HDAC3/SMRT interaction surface 4A69, nucleosome 1ZBB, RXR α LBD-TR α LBD heterodimer 3UVV, RXR α DBD-TR α DBD heterodimer 2NLL.

SMRT corepressor functions as a platform protein that coordinates the assembly of the corepressor complex core. The active complex consists of SMRT and two further core proteins TBL1 and GSP2, which together target the HDAC activity to chromatin (Oberoi et al., 2011). It has a modular structure that ensures flexibility to complete diverse tasks that range from targeting specific NRs in the inactive state to remodeling chromatin. These multiple functionalities are coordinated to bring about the required transcriptional response. The SMRT corepressor complex is able to interact with chromatin through chromatin-binding modules with apparent lack of specificity. The interaction of the corepressor complex with chromatin is carried out through the SANT domain of SMRT (Yu et al. 2003) and through TBL1 (Yoon et al. 2003). These

interactions further ensure that corepressor complexes associate stably with specific promoters once the target NR heterodimer has been recognized. It remains to be seen whether full-length coregulators engage in more extensive contacts with DNA-bound NRs. Therefore, corepressor complex displacement seems to be challenging work due to the large number of interactions that take place. On the other hand, the CoRNR sequence motifs are located within regions of the corepressor protein that are intrinsically disordered. The helical structure appears to be formed only on interaction with the NR LBD (Ahmad et al. 2003; Liu et al. 2007). The entropic cost of forming a fixed structure results in a relative low binding affinity that would favour the binding of the stapled corepressor. In addition, the interaction between SMRT protein and NRs appears to be essential to target the complex to chromatin since only a few subtle conformational changes caused by the ligand binding are enough to displace the corepressor complex from the NR and chromatin. Therefore, it is reasonable to assume that the disruption of this interaction by stapled peptides would be sufficient to mediate corepressor dissociation.

In conclusion, stapled peptides that contain corepressor or coactivator specific binding motifs (LxxLL or LxxxIxxxI/L) constrained by a hydrocarbon link represent a very useful tool to study peptide-protein interactions and a potential agent to combat a number of diseases. The biophysical data reported here represents an encouraging start to further investigate the potency of the stapled SMRT to competitively displace native SMRT *in vivo*.

Appendix 1: Table of Oligonucleotides

Oligo ID	Sequence (5'→3')	Purpose
THRA1_LF_1	gtattttcagggcgcccaggagatgatccgatcactg	Produce A382PfsX7, F397fs406X and E403X TR α LBDs in the pLEICS 12 from the full-length mutant receptors previously placed in the pcDNA3 (N-terminal).
THRA1_LF_6	gacggagctcgaattttagacttctgatcctcaaagac	Produce A382PfsX7, F397fs406X and E403X TR α LBDs in the pLEICS 12 from the full-length mutant receptors previously placed in the pcDNA3 (C-terminal). Produce the full-length TR α variants for the polymorphic study (C-terminal) in the pLEICS14.
THRA1_LF_2	gtattttcagggcgcccgatcactgcagcagcgacca	Shorten the TR α LBD constructs for crystallization trials (N-terminal).
THRA1_P_398	gacggagctcgaattttaggggaagagttcggtggggcac	Produce the P399X TR α LBD construct for crystallization trials (C-terminal).
THRA1_G_398	gacggagctcgaatttatccgaagagttcggtggggcac	Produce the P398G P399X TR α LBD construct for crystallization trials (C-terminal).
THRA1_P_393	gacggagctcgaattttaggggcactcgactttcatgtgg	Produce the P394X TR α LBD construct for crystallization trials (C-terminal).
THRA1_G_393	gacggagctcgaatttatccgcactcgactttcatgtgg	Produce the P393G P394X TR α LBD construct for

		crystallization trials (C-terminal).
THRA1_LF_4	gtatthtcagggcgccatggaacagaagccagcaag	Produce the full-length TR α variants for the polymorphic study (N-terminal).
THRA1_D166G_F	actcctgaagagtgggtctgatccacattgcc	Produce the D166G mutant in full-length TR α
THRA1_D166G_R	ggcaatgtggatcagaccccactcttcaggagt	Produce the D166G mutant in full-length TR α
THRA1_P193S_F	gaggcggaaattcctgagcgtacattggccag	Produce the P193S mutant in full-length TR α
THRA1_P193S_R	ctggccaatgcatcgctcaggaattccgcctc	Produce the P193S mutant in full-length TR α
THRA1_A214V_F	aaggtggacctggaagtcttcagcgagttacc	Produce the A214V mutant in full-length TR α
THRA1_A214V_R	ggtaaactcgctgaagactccaggtccacctt	Produce the A214V mutant in full-length TR α
THRA1_I222V_F	gcgagttaccaagatcgtcacccggcca	Produce the I222V mutant in full-length TR α
THRA1_I222V_R	tggccggggtgacgatcttggtaaactcgc	Produce the I222V mutant in full-length TR α
THRA1_A263V_F	gtccctgcgggtggctgtccgct	Produce the A263V mutant in full-length TR α
THRA1_A263V_R	agcggacagccaccgcagggac	Produce the A263V mutant in full-length TR α
THRA1_V294I_F	ggcggcctgggcatagtctccgacg	Produce the V294I mutant in full-length TR α
THRA1_V294I_R	cgtcggagactatgccaggccgcc	Produce the V294I mutant in full-length TR α
THRA1_T314P_F	ctttaacctggatgaccggaagtggctctgct	Produce the T314P mutant in full-length TR α
THRA1_T314P_R	agcagagccacttccgggtcatccaggttaaag	Produce the T314P mutant in full-length TR α
THRA1_S330X_F	ctaattgcaacagaccgctagggcctgctgtg	Produce the S330X mutant in full-length TR α
THRA1_S330X_R	cacagcaggccctagcggctctgttgacattag	Produce the S330X mutant in full-length TR α

THRA1_S330L_F	ctaagtcaacagaccgcttgggctgctgtg	Produce the S330L mutant in full-length TR α
THRA1_S330L_R	cacagcaggccaagcggctgttgacattag	Produce the S330L mutant in full-length TR α
THRA1_A344V_F	gaagagtcaggaggtgtacctgctggcgt	Produce the A344V mutant in full-length TR α
THRA1_A344V_R	acgccagcaggtacacctctgactcttc	Produce the A344V mutant in full-length TR α
THRA1_V353I_F	gcgttcgagcactacatcaaccaccgcaaac	Produce the V353I mutant in full-length TR α
THRA1_V353I_R	gtttgcggtggttgatgtagtgctcgaacgc	Produce the V353I mutant in full-length TR α
THRA1_R356C_F	agcactacgtcaaccactgcaaacacaacattccg	Produce the R356C mutant in full-length TR α
THRA1_R356C_R	cggaatggtgtgttgcagtggtgacgtagtgct	Produce the R356C mutant in full-length TR α
THRA1_H358P_F	gtcaaccaccgcaaaccaacattccgcacttc	Produce the H358P mutant in full-length TR α
THRA1_H358P_R	gaagtgcggaatgttgggttgcggtggttgac	Produce the H358P mutant in full-length TR α
THRA1_L367P_F	acttctggccaagccgctgatgaaggtgac	Produce the L367P mutant in full-length TR α
THRA1_L369P_R	gtcacctcatcagcggcttgggccagaagt	Produce the L367P mutant in full-length TR α
THRA1_M369L_F	ggccaagctgctgttgaaggtgactgac	Produce the M369L mutant in full-length TR α
THRA1_M369L_R	gtcagtcaccttcaacagcagcttgggcc	Produce the M369L mutant in full-length TR α
THRA1_P399Q_F	cgaactctcccccaactcttctcgagg	Produce the P399Q mutant in full-length TR α
THRA1_P399Q_R	cctcgaggaagagttgggggaagagttcg	Produce the P399Q mutant in full-length TR α

Appendix 2: Buffers and Solutions**Bacterial Expression Media**LB Medium (1 L)

10 g	Bactotryptone
10 g	Bactoyeast extract
5 g	NaCl

2YT Medium (1 L)

16 g	Bactotryptone
10 g	Bactoyeast extract
5 g	NaCl

2YT plates (1 L)

15 g	Agar
10 g	Bactotryptone
5 g	Bactoyeast extract
8 g	NaCl

DNA Purification BuffersSuspension buffer

25 mM	Tris-HCl pH 8
10 mM	EDTA

Suspension buffer 2

20 mM	Tris-HCl pH 8
1 mM	EDTA

Lysis buffer

50 mg	Lysozime
0.2 M	NaOH
1 %	SDS

Neutralization buffer

3 M	Phosphate acetate pH 4.8
-----	--------------------------

Trial Protein BuffersGST Lysis/Suspension Buffer

1x	PBS
1 M	NaCl
1 % (v/v)	Triton X-100
0.5 mM	DTT
#1	Complete EDTA-free protease-inhibitor cocktail tablet (Roche)

Glutathione Sepharose 4B binding buffer

1x	PBS
1 % (v/v)	Triton X-100
0.5 mM	DTT

Glutathione Sepharose 4B cleavage buffer

20 mM	Tris-HCl pH 8
100 mM	NaCl
0.5 mM	DTT

Low salt HiTrap Q IEX buffer

20 mM	Tris-HCl pH 7.4
50 mM	NaCl
1 mM	DTT

High salt HiTrap Q IEX buffer

20 mM	Tris-HCl pH 7.4
500 mM	NaCl
1 mM	DTT

Gel Filtration running buffer

30 mM	Tris-HCl pH 8
50 mM	NaCl
5 % (v/v)	Glycerol
1 mM	DTT
0.5 mM	EDTA

Crystallization buffer

20 mM	Tris-HCl pH 8
50 mM	NaCl
1 mM	DTT

SDS sample buffer

20 %	Glycerol
70 mM	Tris-HCl pH 6.8
0.54 mg/ml	Bromophenol blue
2 %	SDS
200 mM	DTT

PBS buffer

137 mM	NaCl
2.7 mM	KCl
10 mM	Na ₂ HPO ₄
2 mM	KH ₂ PO ₄ pH 7.4

Biophysical characterization buffersFP reaction buffer

1x	PBS
0.01 % (v/v)	Triton X-100
0.1 mg/ml	BSA

CD reaction buffer

1x	PBS
0.1 mg/ml	BSA

JEG-3 cells cultureOpti-MEM grown media

1x	Opti-MEM reduced serum media
10 %	FCS
1 %	PSF

Opti-MEM transfection media

1x	Opti-MEM reduced serum media
10 %	Resin-stripped FCS
1 %	PFS

Trypsinization Solution

0.25 %	Trypsin
1 mM	EDTA

Luciferase and β -galactosidase dual reporter assay buffers

Cell Lysis/Wash buffer

1x	Glycyl-glycine pH 7.8
1 % (v/v)	Triton X-100
1 mM	DTT

Luciferase assay buffer

1x	Glycyl-glycine pH 7.8
100 mM	Potassium phosphate monobasic
200 mM	ATP
100 mM	DTT

Luciferin assay buffer

1x	Glycyl-glycine pH 7.8
1 mM	Luciferin
100 mM	DTT

Glycyl-glycine pH 7.8 buffer

0.5 mM	Glycyl-glycine pH 7.8
1 M	MgSO ₄
180 mM	EGTA

β -galactosidase assay solution

100x	Mg Solution
1x	ONPG
0.1 M	Sodium Phosphate buffer pH 7.5

100x Mg Solution

1 M	MgCl ₂
4.5 M	β-Mercaptoethanol

Sodium Phosphate buffer pH 7.5

0.2 M	Na ₂ HPO ₄ ·2H ₂ O
0.2 M	NaH ₂ PO ₄ ·2H ₂ O

Western blot

Milk blocking buffer

3 %	Dry and fat-free milk made in PBS
-----	-----------------------------------

Appendix 3: Fluorescence anisotropy results of T3 analogues

	Kd (nM)			
	WT LBD	A382PfsX7	F397fs406X	E403X
Apo	198 ± 7.6	64.5 ± 4.8	73 ± 12.28	123 ± 3.1
T3	2400 ± 84	42.5 ± 3.3	63 ± 9.2	164 ± 5.9
DHJ01	162 ± 5.9	82 ± 8.9	85 ± 11.5	114 ± 3.9
DHJ02	181 ± 7.3	106 ± 8.8	105 ± 9.5	147 ± 4.1
DHJ03	340 ± 11.2	95 ± 7.7	118 ± 9.7	188 ± 4.6
ES07	660 ± 13.1	90 ± 7.4	131 ± 12.3	200 ± 7.2
ES08	266 ± 20.3	112 ± 8.3	136 ± 12.6	245 ± 7.2
ES09	340 ± 12.6	121 ± 12.5	144 ± 12.7	246 ± 8.0
ES11	236 ± 10.0	70 ± 8.5	110 ± 22.16	154 ± 39.7
ES32	271 ± 15.3	80.4 ± 5.8	86 ± 30.8	168 ± 41.7
JMT3ALK1	221 ± 10.9	72.2 ± 7.2	121 ± 22.8	154 ± 16.0
JMTYRALK1	235 ± 9.2	79.6 ± 8.3	97 ± 26.6	100 ± 18.6
JM09	198 ± 11.7	75.1 ± 8.8	100 ± 17.5	154 ± 26.8
JM13	130 ± 8.6	77.1 ± 9.1	76 ± 21.7	62 ± 13.2
JM16	233 ± 12.6	87.9 ± 9.7	141 ± 46.2	155 ± 32.3
JM18	221 ± 9.8	83.3 ± 9.1	112 ± 22.6	120 ± 28.5
JM22	140 ± 14.7	85.5 ± 11.0	70.3 ± 19.8	72 ± 19.4

Appendix 3: FA results of T3 analogues using FITC-SMRT corepressor peptide. Table indicates the apparent Kd of the corepressor binding to the different LBDs in the presence of T3 analogues calculated using Graphpad Prism. SEM indicates error introduced through pipetting.

	Kd (μ M)			
	WT	A382PfsX7	F397fs406X	E403X
Apo	N.D.	23.8 \pm 23.4	N.D.	N.D.
T3	0.07 \pm 0.01	19.90 \pm 12.6	N.D.	N.D.
JM09	0.31 \pm 0.1	5.24 \pm 0.8	N.D.	N.D.
JM13	0.71 \pm 0.1	3.16 \pm 0.5	6.6 \pm 2.4	6.23 \pm 1.2
JM16	1.64 \pm 0.4	4.3 \pm 0.9	15.5 \pm 8.1	28.4 \pm 17.9
JM18	1.12 \pm 0.2	3.89 \pm 0.8	2.85 \pm 0.8	8.68 \pm 0.9
JM22	1.94 \pm 0.2	2.32 \pm 0.3	6.64 \pm 3.2	5.00 \pm 0.2

Appendix 4: Summary of the binding affinity of GRIP1 coactivator peptide to the WT and mutant LBDs. Table includes the FA results obtained from saturation binding curves using the WT and mutant LBDs in the presence of T3 analogue tested in coactivator recruitment. The FA was plotted against an increasing concentration of LBD:T3 analogue complex and Kd was calculated using Graphpad prism software. Kd values are the mean \pm SEM of measurements obtained from triplicate experiments.

Chapter 8 - References

- Aasland, R., 1996. The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional corepressor N-CoR and TFIIIB. *Trends in Biochemical Sciences*, 21, pp.87–88.
- Abe, T. et al., 1998. Molecular characterization and tissue distribution of a new organic anion transporter subtype (oatp3) that transports thyroid hormones and taurocholate and comparison with oatp2. *The journal of biological chemistry*, 273, pp.22395–22401.
- Achermann, J.C. et al., 2017. Genetic disorders of nuclear receptors. *Journal of Clinical Investigation*, 127(4), pp.1181–1192.
- Agostini, M. et al., 2004. Tyrosine agonists reverse the molecular defects associated with dominant-negative mutations in human peroxisome proliferator-activated receptor gamma. *Endocrinology*, 145(4), pp.1527–1538.
- Ahmad, K.F. et al., 2003. Mechanism of SMRT corepressor recruitment by the BCL6 BTB domain. *Molecular Cell*, 12(6), pp.1551–1564.
- Allen, B.L. & Taatjes, D.J., 2015. The Mediator complex: a central integrator of transcription. *Nature reviews. Molecular cell biology*, 16(3), pp.155–166.
- Altschul, S.F. et al., 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215, pp.403–410.
- Anzick, S.L. et al., 1997. AIB1, a Steroid Receptor Coactivator Amplified in Breast and Ovarian Cancer. *Science*, 277, pp.965–968.
- Aroyo, M.I., Muller, U. & Wondratschek, H., 2006. Historical introduction. In *International Tables for Crystallography Volume A1: Symmetry relations between space groups*. Symmetry relations between space groups. Chester, England: Springer Netherlands, pp. 2–5.
- Bamshad, M.J. et al., 2011. Exome sequencing as a tool for Mendelian disease gene discovery. *Nature Reviews Genetics*, 12(11), pp.745–755.
- Baniahmad, A. et al., 1995. The Tau4 Activation Domain of the Thyroid Hormone Receptor Is Required for Release of a Putative Corepressor(s) Necessary for Transcriptional Silencing. *Molecular and Cellular Biology*, 15, pp.76–86.
- Baniahmad, A., Kohne, A.C. & Renkawitz, R., 1992. A transferable silencing domain is present in the thyroid hormone receptor, in the v-erbA oncogene product and in the retinoic acid receptor. *The EMBO Journal*, 11, pp.1015–1023.
- Bannister, A.J. & Kouzarides, T., 1995. The CBP co-activator is a histone acetyltransferase. *Nature Clinical Practice Endocrinology & Metabolism*, 384, pp.641–643.

- Barroso, I. et al., 1999. Dominant negative mutations in human PPAR gamma associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature Clinical Practice Endocrinology & Metabolism*, 402(6764), pp.880–883.
- Bassett, J.H.D. & Williams, G.R., 2008. The skeletal phenotypes of TR and TR mutant mice. *Journal of Molecular Endocrinology*, 42(4), pp.269–282.
- Battye, T.G.G. et al., 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Cryst (2011). D67*, 271–281, pp.1–11.
- Beato, M., 1989. Gene Regulation by Steroid Hormones. *Cell*, 56, pp.335–344.
- Beato, M., Herrlich, P. & Schutz, G., 1995. Steroid Hormone Receptors: many actors in search of a plot. *Cell*, 83, pp.851–857.
- Berne, R.M. & Levy, M.N., 1990. *Principles of Physiology*, Mosby Company, St. Louis.
- Bianco, A.C. & Kim, B.W., 2006. Deiodinases: implications of the local control of thyroid hormone action. *Journal of Clinical Investigation*, 116(10), pp.2571–2579.
- Bird, G.H. et al., 2008. Synthesis and Biophysical Characterization of Stabilized α -Helices of BCL-2 Domains. In *Programmed Cell Death, The Biology and Therapeutic Implications of Cell Death, Part B*. Methods in Enzymology. Elsevier, pp. 369–386.
- Bird, G.H., Christian Crannell, W. & Walensky, L.D., 2011. *Chemical Synthesis of Hydrocarbon-Stapled Peptides for Protein Interaction Research and Therapeutic Targeting*, Hoboken, NJ, USA: John Wiley & Sons, Inc.
- Bleicher, L. et al., 2008. Structural basis of GC-1 selectivity for thyroid hormone receptor isoforms. *BMC Structural Biology*, 8(1), p.8.
- Bochukova, E. et al., 2012. A Mutation in the Thyroid Hormone Receptor Alpha Gene. *New England Journal of Medicine*, 366(3), pp.243–249.
- Bourguet, W. et al., 1995. Crystal structure of the ligand-binding domain of the human nuclear receptor RXR alpha. *Nature Clinical Practice Endocrinology & Metabolism*, 375, pp.377–382.
- Bracco, D. et al., 1993. Comparison of the metabolic and endocrine effects of 3,5,3'-triiodothyroacetic acid and thyroxine. *The Journal of Clinical Endocrinology & Metabolism*, 77(1), pp.221–228.
- Brzozowski, A.M., Pike, A.C., Dauter, Z., Hubbard, R.E., Bonn, T., Engström, O., et al., 1997. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature Clinical Practice Endocrinology & Metabolism*, 389(6652), pp.753–758.
- Bugge, T.H. et al., 1992. RXRalpha, a promiscuous partner of retinoic acid and thyroid hormone receptors. *The EMBO Journal*, 11, pp.1418–1428.

- Burris, T.P. et al., 2013. Nuclear receptors and their selective pharmacologic modulators. *Pharmacological Reviews*, 65(2), pp.710–778.
- Cao, G. et al., 2002. *In vivo* delivery of a Bcl-xL fusion protein containing the TAT protein transduction domain protects against ischemic brain injury and neuronal apoptosis. *The Journal of Neuroscience*, 22, pp.5423–5431.
- Cariou, B., Charbonnel, B. & Staels, B., 2012. Thiazolidinediones and PPAR γ agonists: time for a reassessment. *Trends in endocrinology and metabolism: TEM*, 23(5), pp.205–215.
- Carroll, J.S. et al., 2005. Chromosome-Wide Mapping of Estrogen Receptor Binding Reveals Long-Range Regulation Requiring the Forkhead Protein FoxA1. *Cell*, 122(1), pp.33–43.
- Cavaillès, V. et al., 1994. Interaction of proteins with transcriptionally active estrogen receptors. *Proceedings of the National Academy of Sciences*, 91, pp.10009–10013.
- Chandra, V. et al., 2013. Multidomain integration in the structure of the HNF-4 α nuclear receptor complex. *Nature Clinical Practice Endocrinology & Metabolism*, 495(7441), pp.394–398.
- Chandra, V. et al., 2008. Structure of the intact PPAR γ –RXR α nuclear receptor complex on DNA. *Nature Clinical Practice Endocrinology & Metabolism*, pp.350–356.
- Chanoine, J.P. et al., 1993. The thyroid gland is a major source of circulating T3 in the rat. *Journal of Clinical Investigation*, 91, pp.1–5.
- Darwin, C., 1860. *On the Origin of Species by Means of Natural Selection, Or the Preservation of Favoured Races in the Struggle for Life*, London: Murray.
- Charles, J.P., Shinoda, T. & Chinzei, Y., 1999. Characterization and DNA-binding properties of GRF, a novel monomeric binding orphan receptor related to GCNF and FTZ-F1. *European Journal of Biochemistry*, 266, pp.181–190.
- Charmandari, E. et al., 2004. Natural glucocorticoid receptor mutants causing generalized glucocorticoid resistance: molecular genotype, genetic transmission, and clinical phenotype. *The Journal of Clinical Endocrinology & Metabolism*, 89(4), pp.1939–1949.
- Chatterjee, V.K. et al., 1991. Thyroid hormone resistance syndrome. Inhibition of normal receptor function by mutant thyroid hormone receptors. *Journal of Clinical Investigation*, 87(6), pp.1977–1984.
- Chen, H. et al., 1997. Nuclear Receptor Coactivator ACTR Is a Novel Histone Acetyltransferase and Forms a Multimeric Activation Complex with P/CAF and CBP/p300. *Cell*, 90, pp.569–580.
- Chen, J.D. & Evans, R.M., 1995. A transcriptional corepressor that interacts with nuclear hormone receptors. *Nature Clinical Practice Endocrinology & Metabolism*, 377, pp.454–459.

- Cheng, S.Y., 2000. Multiple Mechanisms for Regulation of the Transcriptional Activity of Thyroid Hormone Receptors. *Endocrine and Metabolic Disorders*, 1, pp.9–18.
- Cheskis, B. & Freedman, L.P., 1996. Modulation of Nuclear Receptor Interactions by Ligands: Kinetic Analysis Using Surface Plasmon Resonance. *Biochemistry*, 35, pp.3309–3318.
- Chiellini, G. et al., 1998. A high-affinity subtype-selective agonist ligand for the thyroid hormone receptor. *Chemistry and Biology*, pp.299–306.
- Christensen, H.N., Hess, B. & Riggs, T.R., 1954. Concentration of Taurine, beta-Alanine, and Triiodothyronine by Ascites Carcinoma Cells. *American Association for Cancer Research*, 14(2), pp.124–127.
- Codina, A. et al., 2005. Structural insights into the interaction and activation of histone deacetylase 3 by nuclear receptor corepressors. *PNAS*, 102, pp.6009–6014.
- Cohen, P.P., 1970. Biochemical differentiation during Amphibian Metamorphosis. *Science*, 168(3931), pp.533–543.
- Cole, C., Barber, J.D. & Barton, G.J., 2008. The Jpred 3 secondary structure prediction server. *Nucleic Acids Research*, 36(Web Server), pp.W197–W201.
- Collaborative Computational Project, N.4., 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallography D Biological Crystallography*, pp.760–763.
- Dahl, G.E. et al., 1994. A central negative feedback action of thyroid hormones on thyrotropin-releasing hormone secretion. *Endocrinology*, 135, pp.2392–2397.
- Dai, S.Y. et al., 2009. Unique ligand binding patterns between estrogen receptor alpha and beta revealed by hydrogen-deuterium exchange. *Biochemistry*, 48(40), pp.9668–9676.
- Darimont, B.D. et al., 1998. Structure and specificity of nuclear receptor-coactivator interactions. *Genes & Development*, 12(21), pp.3343–3356.
- Demir, K. et al., 2016. Diverse Genotypes and Phenotypes of Three Novel Thyroid Hormone Receptor- α Mutations. *The Journal of Clinical Endocrinology & Metabolism*, 101(8), pp.2945–2954.
- Demizu, Y. et al., 2013. Development of stapled short helical peptides capable of inhibiting vitamin D receptor (VDR)-coactivator interactions. *Bioorganic & Medicinal Chemistry Letters*, 23(15), pp.4292–4296.
- DiRenzo, J. et al., 1997. Peroxisome Proliferator-Activated Receptors and Retinoic Acid Receptors Differentially Control the Interactions of Retinoid X Receptor Heterodimers with Ligands, Coactivators, and Corepressors. *Molecular and Cellular Biology*, 17, pp.2166–2176.
- Docter, R. et al., 1997. Expression of Rat Liver Cell Membrane Transporters for Thyroid Hormone in *Xenopus laevis* Oocytes. *Endocrinology*, 138, pp.1841–1847.

- Edwards, A.L. et al., 2013. Multimodal interaction with BCL-2 family proteins underlies the proapoptotic activity of PUMA BH3. *Chemistry and Biology*, 20(7), pp.888–902.
- Egea, P.F. et al., 2000. Crystal structure of the human RXR ligand-binding domain bound to its natural ligand: 9-cis retinoic acid. *The EMBO Journal*, 19(11), pp.2592–2601.
- Emsley, P. et al., 2010. Features and development of Coot. *Acta Cryst (2010). D66, 486-501*, pp.1–16.
- Escriva, H., Delaunay, F. & Laudet, V., 2000. Ligand binding and nuclear receptor evolution. *BioEssays*, 22, pp.717–727.
- Espiard, S. et al., 2015. A Novel Mutation in THRA Gene Associated With an Atypical Phenotype of Resistance to Thyroid Hormone. *The Journal of Clinical Endocrinology & Metabolism*, 100(8), pp.2841–2848.
- Etkin, W. & Gona, A.G., 1974. The evolution of thyroid function in poikilothermic vertebrates. *Handbook of Physiology: The Thyroid*, American Physiological Society, Washington, DC.
- Evans, P., 2005. Scaling and assessment of data quality. *Acta Cryst (2006). D62, 72-82*, pp.1–11.
- Evans, P.R. & Murshudov, G.N., 2013. How good are my data and what is the resolution? *Acta Cryst (2013). D69, 1204-1214*, pp.1–11.
- Everts, M.E. et al., 1996. Different regulation of thyroid hormone transport in liver and pituitary: its possible role in the maintenance of low T3 production during nonthyroidal illness and fasting in man. *Thyroid*, 6, pp.359–368.
- Fen Ding, X. et al., 1998. Nuclear Receptor-Binding Sites of Coactivators Glucocorticoid Receptor Interacting Protein 1 (GRIP1) and Steroid Receptor Coactivator 1 (SRC-1): Multiple Motifs with Different Binding Specificities. *Molecular Endocrinology*, pp.302–313.
- Fliers, E., Klieverik, L.P. & Kalsbeek, A., 2010. Novel neural pathways for metabolic effects of thyroid hormone. *Trends in Endocrinology & Metabolism*, 21(4), pp.230–236.
- Fondell, J.D., Ge, H. & Roeder, R.G., 1996. Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proceedings of the National Academy of Sciences*, 93, pp.8329–8333.
- Fonseca, T.L. et al., 2013. Coordination of hypothalamic and pituitary T3 production regulates TSH expression. *Journal of Clinical Investigation*, 123(4), pp.1492–1500.
- Forman, B.M. et al., 1995. Unique Response Pathways Are Established by Allosteric Interactions among Nuclear Hormone Receptors. *Cell*, 81, pp.541–550.
- Frankel, A.D. & Pabo, C.O., 1988. Cellular uptake of the Tat protein from Human

- Immuno deficiency Virus. *Cell*, 55, pp.1189–1193.
- French, S. & Wilson, K., 1978. On the treatment of negative intensity observations. *Acta Crystallographica*, A34, pp.517–525.
- Friesema, E.C. et al., 2006. Thyroid hormone transport by the human monocarboxylate transporter 8 and its rate-limiting role in intracellular metabolism. *Molecular endocrinology (Baltimore, Md.)*, 20(11), pp.2761–2772.
- Gallastegui, N. et al., 2015. Advances in our structural understanding of orphan nuclear receptors. *Trends in Biochemical Sciences*, 40(1), pp.25–35.
- Gampe, R.T., Montana, V.G., et al., 2000. Asymmetry in the PPAR γ /RXR α crystal structure reveals the molecular basis of heterodimerization among nuclear receptors. *Molecular Cell*, 5, pp.545–555.
- Gampe, R.T., Mountana, V.G., et al., 2000. Structural basis for autorepression of retinoid X receptor by tetramer formation and the AF-2 helix. *Genes & Development*, 14, pp.2229–2241.
- Giguere, V. et al., 1986. Functional Domains of the Human Glucocorticoid Receptor. *Cell*, 46, pp.645–652.
- Giguere, V., McBroom, L.D.B. & Flock, G., 1995. Determinants of Target Gene Specificity for ROR1: Monomeric DNA Binding by an Orphan Nuclear Receptor. *Molecular and Cellular Biology*, 15, pp.2517–2526.
- Green, M. & Loewenstein, P.M., 1988. Autonomous Functional Domains of Chemically Synthesized Human Immunodeficiency Virus Tat *Trans*-Activator Protein. *Cell*, 55, pp.1179–1188.
- Green, S. & Chambon, P., 1987. Oestradiol induction of a glucocorticoid-responsive gene by a chimaeric receptor. *Nature Clinical Practice Endocrinology & Metabolism*, 325, pp.75–78.
- Green, S. et al., 1986. Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. *Nature Clinical Practice Endocrinology & Metabolism*, 320, pp.134–139.
- Gronemeyer, H. & Laudet, V., 1995. Transcription factors 3: nuclear receptors. *Protein Profile*, 2, pp.1173–1308.
- Gronemeyer, H., Gustafsson, J.-A. & Laudet, V., 2004. Principles for modulation of the nuclear receptor superfamily. *Nature Reviews Drug Discovery*, 3(11), pp.950–964.
- Gudernatsch, J.F., 1912. Feeding Experiments on tadpoles. *Archiv fur Entwicklungsmechanik der Organismen*, 35(3), pp.357–483.
- Guenther, M.G. et al., 2000. A core SMRT corepressor complex containing HDAC3 and TBL1, a WD40-repeat protein linked to deafness. *Genes & Development*, 14, pp.1048–1057.

- Guenther, M.G., Barak, O. & Lazar, M.A., 2001. The SMRT and N-CoR Corepressors Are Activating Cofactors for Histone Deacetylase 3. *Molecular and Cellular Biology*, 21(18), pp.6091–6101.
- Gurnell, M. et al., 2016. *Resistance to Thyroid Hormone* Seventh Edition, Elsevier.
- Gurnell, M. et al., 1999. Three Novel Mutations at Serine 314 in the Thyroid Hormone beta Receptor Differentially Impair Ligand Binding in the Syndrome of Resistance to Thyroid Hormone. *Endocrine*, 140(12), pp.5901–5906.
- Halachmi, S. et al., 1994. Estrogen Receptor-associated proteins: possible mediators of hormone-induced transcription. *Science*, 264, pp.1455–1460.
- Hallenbeck, P. et al., 1992. Heterodimerization of thyroid hormone (TH) receptor with H-2RIIBP (RXR beta) enhances DNA binding and TR-dependent transcriptional activation. *Proceedings of the National Academy of Sciences*, 89, pp.1–5.
- Hamuro, Y. et al., 2006. Hydrogen/deuterium-exchange (H/D-Ex) of PPAR gamma LBD in the presence of various modulators. *Protein Science*, 15(8), pp.1883–1892.
- Harding, H. & Lazar, M.A., 1995. The monomer-binding orphan receptor Rev-Erb represses transcription as a dimer on a novel direct repeat. *Molecular and Cellular Biology*, 15(9), pp.4791–4802.
- Hartman, H.B. et al., 2005. The histone-binding code of nuclear receptor co-repressors matches the substrate specificity of histone deacetylase 3. *EMBO reports*, 6(5), pp.445–451.
- Hård, T. et al., 1990. ¹H NMR studies of the glucocorticoid receptor DNA-binding domain: sequential assignments and identification of secondary structure elements. *Biochemistry*, 29(38), pp.9015–9023.
- Heery, D.M. et al., 1997. A signature motif in transcriptional coactivators mediates binding to nuclear receptors. *Nature Clinical Practice Endocrinology & Metabolism*, 387, pp.733–737.
- Hennemann, G., 1986. *Thyroid hormone deiodination in healthy man*, Thyroid Hormone Metabolism. Marcel Dekker, New York.
- Hollenberg, A.N., Monden, T. & Wondisford, F.E., 1995. Ligand-independent and -dependent functions of Thyroid Hormone Receptor Isoforms depend upon their distinct amino termini. *The journal of Biological Chemistry*, 270, pp.14274–14280.
- Hong, H. et al., 1997. GRIP1, a Transcriptional Coactivator for the AF-2 Transactivation Domain of Steroid, Thyroid, Retinoid, and Vitamin D Receptors. *Molecular and Cellular Biology*, 17(5), pp.2735–2744.
- Horlein, A., Heinzl, T. & Rosenfeld, M.G., 1996. Gene regulation by thyroid hormone receptors. *Current Opinion in Endocrinology and Diabetes*, 3(5), pp.412–416.
- Hoskins, R.G., 1949. The thyroid-pituitary apparatus as a servo (feed-back) mechanism. *Journal of Clinical Investigation*, 9, pp.1429–1431.

- Hu, E. et al., 1996. Inhibition of Adipogenesis Through MAP Kinase-Mediated Phosphorylation of PPAR gamma. *Science*, 274, pp.1–5.
- Hu, X. & Lazar, M.A., 1999. The CoRNR motif controls the recruitment of corepressors by nuclear hormone receptors. *Nature Clinical Practice Endocrinology & Metabolism*, 402, pp.93–96.
- Hu, X. et al., 2003. Liver X Receptors interact with corepressors to regulate gene expression. *Molecular Endocrinology*, 17, pp.1019–1026.
- Incardona, M.F. et al., 2009. EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. *Journal of synchrotron radiation*, 16(Pt 6), pp.872–879.
- Inoue, S. et al., 2000. An Estrogen Receptor β Isoform That Lacks Exon 5 Has Dominant Negative Activity on both ER α and ER β . *Biochemical and Biophysical Research Communications*, 279(3), pp.814–819.
- Kabsch, W., 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst (1993)*. 26, 795–800, pp.1–6.
- Kallenberger, B.C. et al., 2003. A dynamic mechanism of nuclear receptor activation and its perturbation in a human disease. *Nature Structural Biology*, 10(2), pp.136–140.
- Kamei, Y. et al., 1996. A CBP Integrator Complex Mediates Transcriptional Activation and AP-1 Inhibition by Nuclear Receptors. *Cell*, 85, pp.403–414.
- Kantardjieff, K.A. & Rupp, B., 2003. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Science*, 12(9), pp.1865–1871.
- Kaplan, M.M., 1984. The role of thyroid hormone deiodination in the regulation of hypothalamo-pituitary function. *Neuro Endocrinology*, 38, pp.254–260.
- Keidel, S., Lemotte, P. & Apfel, C., 1994. Different agonist- and antagonist-induced conformational changes in retinoic acid receptors analyzed by protease mapping. *Molecular and Cellular Biology*, pp.287–298.
- Kim, D.W. et al., 2014. A histone deacetylase inhibitor improves hypothyroidism caused by a TR α 1 mutant. *Human molecular genetics*, 23(10), pp.2651–2664.
- Kino, T. et al., 2001. Pathologic human GR mutant has a transdominant negative effect on the wild-type GR by inhibiting its translocation into the nucleus: importance of the ligand-binding domain for intracellular GR trafficking. *The Journal of Clinical Endocrinology & Metabolism*, 86(11), pp.5600–5608.
- Kitajima, K., Nagaya, T. & Jameson, J.L., 1995. Dominant negative and DNA-binding properties of mutant thyroid hormone receptors that are defective in homodimerization but not heterodimerization. *Thyroid*, 5(5), pp.343–353.

- Klein, I. & Ojamaa, K., 2001. Thyroid Hormone and the Cardiovascular System. *New England Journal of Medicine*, 344, pp.501–510.
- Kliwer, S.A. et al., 1992. Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D₃ signalling. *Nature Clinical Practice Endocrinology & Metabolism*, 355, pp.446–449.
- Koenig, R.J. et al., 1987. Thyroid hormone receptor binds to a site in the rat growth hormone promoter required for induction by thyroid hormone. *Proceedings of the National Academy of Sciences of the United States of America*, 84(16), pp.5670–5674.
- Kohrle, J., Brabant, G. & Hesch, R.D., 1987. Metabolism of the Thyroid Hormones. *Homone Research*, 26, pp.58–78.
- Krenning, E.P. et al., 1978. Active transport of triiodothyronine (T₃) into isolated rat liver cells. *FEBS Journal*, 91, pp.113–116.
- Kumar, V. et al., 1987. Functional Domains of the Human Estrogen Receptor. *Cell*, 51, pp.941–951.
- LaBelle, J.L. et al., 2012. A stapled BIM peptide overcomes apoptotic resistance in hematologic cancers. *Journal of Clinical Investigation*, 122(6), pp.2018–2031.
- Laudet, V. et al., 1992. Evolution of the nuclear receptor gene superfamily. *The EMBO Journal*, 11, pp.1003–1013.
- Lazar, M.A., 1993. Thyroid Hormone Receptors: Multiple Forms, Multiple Possibilities. *Endocrine Reviews*, 14, pp.1–10.
- Lee, Y. & Mahdavi, V., 1993. The D Domain of the Thyroid Hormone Receptor alpha 1 Specifies Positive and Negative Transcriptional Regulation Functions. *The journal of biological chemistry*, 268, pp.2021–2028.
- Leid, M. et al., 1992. Purification, Cloning, and RXR Identity of the HeLa Cell Factor with Which RAR or TR Heterodimerizes to Bind Target Sequences Efficiently. *Cell*, 68, pp.377–395.
- Lek, M. et al., 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature Publishing Group*, 536(7616), pp.285–291.
- Leng, X. et al., 1993. Ligand-dependent conformational changes in thyroid hormone and retinoic acid receptors are potentially enhanced by heterodimerization with retinoic X receptor. *The Journal of Steroid Biochemistry and Molecular Biology*, 46, pp.643–661.
- Leonard, J.L. & Koehrle, J., 1996. Intracellular pathways of iodothyronine metabolism. *L.R. Publishers, ed., National Library of Medicine*.
- Leslie, A.G.W., Recent changes to the MOSFLM package for processing film and image plate data. *Newsletter on Protein Crystallography*.

- Levy-Bimbot, M. et al., 2012. Tetrabromobisphenol-A disrupts thyroid hormone receptor alpha function in vitro: Use of fluorescence polarization to assay corepressor and coactivator peptide binding. *Chemosphere*, 87(7), pp.782–788.
- Li, D. et al., 2004. Novel Roles of Retinoid X Receptor (RXR) and RXR Ligand in Dynamically Modulating the Activity of the Thyroid Hormone Receptor/RXR Heterodimer. *Journal of Biological Chemistry*, 279(9), pp.7427–7437.
- Li, H. et al., 1997. Characterization of Receptor Interaction and Transcriptional Repression by the Corepressor SMRT. *Molecular Endocrinology*, 11, pp.2025–2037.
- Li, J. et al., 2000. Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. *The EMBO Journal*, 19, pp.4342–4350.
- Li, Y., Lambert, M.H. & Xu, H.E., 2003. Activation of nuclear receptors: a perspective from structural genomics. *Structure*, 11(7), pp.741–746.
- Lin, K.H., Fukuda, T. & Cheng, S.Y., 1990. Hormone and DNA binding activity of a purified human thyroid hormone nuclear receptor expressed in Escherichia coli. *Journal of Biological Chemistry*, 265(9), pp.5161–5165.
- Lindsay, M.A., 2002. Peptide-mediated cell delivery: application in protein target validation. *Current Opinion in Pharmacology*, pp.587–594.
- Liu, Y. et al., 2007. Structural basis for recognition of SMRT/N-CoR by the MYND domain and its contribution to AML1/ETO's activity. *Cancer cell*, 11(6), pp.483–497.
- Lou, X. et al., 2014. Structure of the retinoid X receptor alpha-liver X receptor beta (RXR alpha-LXR beta) heterodimer on DNA. *Nature Publishing Group*, 21(3), pp.277–281.
- Luisi, B.F. et al., 1991. Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature Clinical Practice Endocrinology & Metabolism*, 352, pp.497–505.
- Luongo, C. et al., 2015. The selective loss of the type 2 iodothyronine deiodinase in mouse thyrotrophs increases basal TSH but blunts the thyrotropin response to hypothyroidism. *Endocrinology*, 156(2), pp.745–754.
- MacArthur, D.G. et al., 2012. A systematic survey of loss-of-function variants in human protein-coding genes. *Science*, 335(6070), pp.823–828.
- Madauss, K.P. et al., 2007. A Structural and in Vitro Characterization of Asoprisnil: A Selective Progesterone Receptor Modulator. *Molecular Endocrinology*, 21(5), pp.1066–1081.
- Magnus-Levy, A., 1895. Uber den respiratorischen gaswechsel unter dem einfluss der thyroidea sowie unter verschiedenen pathologischen zustanden. *Berliner Klinische Wochenschrift*, 34, pp.650–652.

- Majithia, A.R. et al., 2016. Prospective functional classification of all possible missense variants in PPARG. *Nature Genetics*, pp.1–8.
- Majithia, A.R. et al., 2014. Rare variants in PPARG with decreased activity in adipocyte differentiation are associated with increased risk of type 2 diabetes. *Proceedings of the National Academy of Sciences*, 111(36), pp.13127–13132.
- Mangelsdorf, D.J. & Evans, R.M., 1995. The RXR Heterodimers and Orphan Receptors. *Cell*, 83, pp.841–850.
- Mangelsdorf, D.J. et al., 1995. The nuclear receptor superfamily: The second decade. *Cell*, 83(6), pp.835–839.
- Mann, D.A. & Frankel, A.D., 1991. Endocytosis and targeting of exogenous HIV-1 Tat protein. *The EMBO Journal*, 10, pp.1733–1739.
- Margulis, L., 1993. *Symbiosis in Cell Evolution*, W H Freeman & Company.
- Marimuthu, A. et al., 2002. TR surfaces and conformations required to bind nuclear receptor corepressor. *Molecular endocrinology (Baltimore, Md.)*, 16(2), pp.271–286.
- Marks, M.S. et al., 1992. H-2RIIBP (RXRbeta) heterodimerization provides a mechanism for combinatorial diversity in the regulation of retinoic acid and thyroid hormone responsive genes. *The EMBO Journal*, 11, pp.1419–1435.
- Martínez, L. et al., 2009. Gaining ligand selectivity in thyroid hormone receptors via entropy. *PNAS*, 106(49), pp.20717–20722.
- Matthews, B.W., 1968. Solvent content of protein crystals. *Journal of Molecular Biology*, 33, pp.491–497.
- McCoy, A.J. et al., 2007. Phaser crystallographic software. *J. Appl. Cryst (2007)*. 40, 658-674, pp.1–17.
- McEwan, I.J., 2004. Molecular mechanisms of androgen receptor-mediated gene regulation: structure-function analysis of the AF-1 domain. *Endocrine-Related Cancer*, 11, pp.281–293.
- McInerney, E.M. et al., 1998. Determinants of coactivator LxxLL motif specificity in nuclear receptor transcriptional activation. *Genes & Development*, 12, pp.3357–3368.
- Mitsubishi, T., Tennyson, G.E. & Nikodem, V.M., 1988. Alternative splicing generates messages encoding rat c-erbA proteins that do not bind thyroid hormone. *Proceedings of the National Academy of Sciences*, 85, pp.5804–5808.
- Moran, C. & Chatterjee, K., 2015. Resistance to thyroid hormone due to defective thyroid receptor alpha. *Best Practice & Research Clinical Endocrinology & Metabolism*, 29(4), pp.647–657.
- Moran, C. et al., 2013. An Adult Female With Resistance to Thyroid Hormone

- Mediated by Defective Thyroid Hormone Receptor α . *The Journal of Clinical Endocrinology & Metabolism*, 98(11), pp.4254–4261.
- Moran, C. et al., 2014. Resistance to Thyroid Hormone due to a mutation in thyroid hormone receptor alpha 1 and alpha 2 variant protein. pp.1–57.
- Moran, C. et al., 2016. Resistance to thyroid hormone caused by a mutation in thyroid hormone receptor TR α 1 and TR α 2: clinical, biochemical, and genetic analyses of three related patients. *Diabetes & Endocrinology*, 2(8), pp.619–626.
- Morreale de Escobar, G., Obregon, M.J. & Escobar del Rey, F., 1987. Fetal and maternal thyroid hormones. *Hormone Research*, 26, pp.12–27.
- Murshudov, G.N. et al., 1999. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Cryst (1999)*. D55, 247-255, pp.1–9.
- Naar, A.M. et al., 1999. Composite co-activator ARC mediates chromatin-directed transcriptional activation. *Nature Clinical Practice Endocrinology & Metabolism*, 398, pp.828–832.
- Nagaya, T. & Jameson, J.L., 1993. Thyroid Hormone Receptor Dimerization Is Required for Dominant Negative Inhibition by Mutations That Cause Thyroid Hormone Resistance. *The journal of biological chemistry*, 268(21), pp.15766–15771.
- Nagy, L., 2004. Mechanism of the nuclear receptor molecular switch. *Trends in Biochemical Sciences*, 29(6), pp.317–324.
- Nagy, L. et al., 1999. Mechanism of corepressor binding and release from nuclear hormone receptors. *Genes & Development*, 13, pp.3209–3216.
- Nakajima, T. et al., 1996. The Signal-Dependent Coactivator CBP Is a Nuclear Target for pp90RSK. *Cell*, 86, pp.465–474.
- Nascimento, A.S. et al., 2006. Structural Rearrangements in the Thyroid Hormone Receptor Hinge Domain and Their Putative Role in the Receptor Function. *Journal of Molecular Biology*, 360(3), pp.586–598.
- Näär, A.M. et al., 1991. The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. *Cell*, 65(7), pp.1267–1279.
- Nguyen, N.H. et al., 2002. Rational Design and Synthesis of a Novel Thyroid Hormone Antagonist That Blocks Coactivator Recruitment. *Journal of Medicinal Chemistry*, 45(15), pp.3310–3320.
- Nolte, R.T. et al., 1998. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature Clinical Practice Endocrinology & Metabolism*, 395, pp.137–142.
- Nuclear Receptors Nomenclature Committee, 1999. A unified nomenclature system for the nuclear receptor superfamily. *Cell*, 97, pp.161–163.

- Nunez, J. et al., 1991. Regulation by Thyroid Hormone of Microtubule Assembly and Neuronal Differentiation. *Neurochemical Research*, 16, pp.975–982.
- Oberoi, J. et al., 2011. Structural basis for the assembly of the SMRT/NCoR core transcriptional repression machinery. *Nature Structural and Molecular Biology*, 18(2), pp.177–184.
- Ocasio, C.A. & Scanlan, T.S., 2006. Design and Characterization of a Thyroid Hormone Receptor α (TR α)-Specific Agonist. *ACS Chemical Biology*, 1(9), pp.585–593.
- Ogryzko, V.V. et al., 1996. The Transcriptional Coactivators p300 and CBP Are Histone Acetyltransferases. *Cell*, 87, pp.953–959.
- Onate, S.A. et al., 1995. Sequence and characterization of a coactivator for the Steroid Hormone Receptor superfamily. *Science*, 270, pp.1–5.
- Oppenheimer, J.H. & Schwartz, H.L., 1997. Molecular Basis of Thyroid Hormone-Dependent Brain Development. *Endocrine Reviews*, 18, pp.462–475.
- Oppenheimer, J.H., Schwartz, H.D. & Surks, M.I., 1974. Tissue differences in the concentration of triiodothyronine nuclear binding sites in the rat: liver, kidney, pituitary, heart, brain, spleen, and testis. *Endocrinology*, 95, pp.897–903.
- Owen-Hughes, T. & Workman, J.L., 1996. Remodeling the chromatin structure of a nucleosome array by transcription factor-targeted *trans*-displacement of histones. *The EMBO Journal*, 15, pp.4702–4712.
- Ozers, M.S. et al., 1997. Equilibrium Binding of Estrogen Receptor with DNA Using Fluorescence Anisotropy. *The journal of biological chemistry*, 272, pp.30405–30411.
- Perissi, V. et al., 1999. Molecular determinants of nuclear receptor–corepressor interaction. *Genes & Development*, 13, pp.3198–3208.
- Perlmann, T. et al., 1993. Determinants for selective RAR and TR recognition of direct repeat HREs. *Genes & Development*, 7, pp.1411–1422.
- Pérez, E. et al., 2012. Modulation of RXR function through ligand design. *Molecular and Cell Biology of Lipids*, 1821(1), pp.57–69.
- Phelan, C.A. et al., 2010. Structure of Rev-erb α bound to N-CoR reveals a unique mechanism of nuclear receptor-co-repressor interaction. *Nature Structural and Molecular Biology*, 17(7), pp.808–814.
- Phillips, C. et al., 2011. Design and Structure of Stapled Peptides Binding to Estrogen Receptors. *Journal of the American Chemical Society*, 133(25), pp.9696–9699.
- Pissios, P. et al., 2000. Dynamic Stabilization of Nuclear Receptor Ligand Binding Domains by Hormone or Corepressor Binding. *Molecular Cell*, 6, pp.245–253.
- Pollard, K.J. & Peterson, C.L., 1998. Chromatin remodeling: a marriage between two

- families? *BioEssays*, 20, pp.771–780.
- Porterfield, S.P. & Hendrich, C.E., 1993. The Role of Thyroid Hormones in Prenatal and Neonatal Neurological Development—Current Perspectives. *Endocrine Reviews*, 14, pp.1–13.
- Powell, H.R. et al., 2017. Integrating macromolecular X-ray diffraction data with the graphical user interface iMosflm. *Nature Protocols*, 12(7), pp.1310–1325.
- Putchá, B.D.K. et al., 2012. Structural basis for negative cooperativity within agonist-bound TR:RXR heterodimers. *PNAS*, 109, pp.6084–6087.
- Rachez, C. et al., 1999. Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. *Nature Clinical Practice Endocrinology & Metabolism*, 398, pp.824–829.
- Raghuram, S. et al., 2007. Identification of heme as the ligand for the orphan nuclear receptors REV-ERBa and REV-ERBβ. *Nature Structural and Molecular Biology*, 14(12), pp.1207–1213.
- Rao, G.S. et al., 1976. Uptake of thyroid hormone by isolated rat liver cells. *Biochemical and Biophysical Research Communications*, 73(1), pp.98–104.
- Rastinejad, F. et al., 1995. Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature Clinical Practice Endocrinology & Metabolism*, 375, pp.203–211.
- Refetoff, S. & Dumitrescu, A.M., 2007. Syndromes of reduced sensitivity to thyroid hormone: genetic defects in hormone receptors, cell transporters and deiodination. *Best Practice & Research Clinical Endocrinology & Metabolism*, 21(2), pp.277–305.
- Refetoff, S., Weiss, R.E. & Usala, S.J., 1993. The syndromes of resistance to thyroid hormone. *Endocrinology Reviews*, 14, pp.348–399.
- Renaud, J.-P. et al., 1995. Crystal structure of the RARγ ligand-binding domain bound to all-trans retinoic acid. *Nature Clinical Practice Endocrinology & Metabolism*, 378(6558), pp.681–689.
- Roeder, R.G., 1996. The role of general initiation factors in transcription by RNA polymerase II. *Trends in Biochemical Sciences*, pp.327–335.
- Rondeel, J.M. et al., 1988. Effect of Thyroid Status and Paraventricular Area Lesions on the Release of Thyrotropin-Releasing Hormone and Catecholamines into Hypophysial Portal Blood. *Endocrinology*, 123, pp.1–5.
- Rosenfeld, M.G., Lunyak, V.V. & Glass, C.K., 2006. Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes & Development*, 20(11), pp.1405–1428.
- Sablin, E. et al., 2003. Structural Basis for Ligand-Independent Activation of the Orphan Nuclear Receptor LRH-1. *Molecular Cell*, 11, pp.1575–1585.

- Salter, W.T., 1940. Endocrine function of iodine. *Harvard University press*.
- Samuels, H.H. et al., 1974. In vitro characterization of solubilized nuclear receptors from rat liver and cultured GH1 cells. *The Journal of Clinical Investigation*, 54, pp.853–865.
- Sap, J. et al., 1986. The c-erb-A protein is a high-affinity receptor for thyroid hormone. *Nature Clinical Practice Endocrinology & Metabolism*, 324(6098), pp.635–640.
- Sauter, N.K., Grosse-Kunstleve, R.W. & Adams, P.D., 2004. Robust indexing for automatic data collection. *Journal of Applied Crystallography*, 37(3), pp.399–409.
- Schafmeister, C.E., Po, J. & Verdine, G.L., 2000. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *Journal of the American Chemical Society*, 122(24), pp.5891–5892.
- Schueler, P.A. et al., 1990. Binding of 3,5,3'-triiodothyronine (T3) and its analogs to the in vitro translational products of c-erbA protooncogenes: differences in the affinity of the alpha- and beta-forms for the acetic acid analog and failure of the human testis and kidney alpha-2 products to bind T3. *Molecular endocrinology*, 4(2), pp.227–234.
- Schwabe, J.W., Neuhaus, D. & Rhodes, D., 1990. Solution structure of the DNA-binding domain of the oestrogen receptor. *Nature Clinical Practice Endocrinology & Metabolism*, 348(6300), pp.458–461.
- Schwabe, J.W.R. et al., 1993. The Crystal Structure of the Estrogen Receptor DNA-Binding Domain Bound to DNA: How Receptors Discriminate between Their Response Elements. *Cell*, 75, pp.567–578.
- Schwarze, S.R. et al., 1999. *In Vivo* protein transduction: delivery of a biologically active protein into the Mouse. *Science*, pp.1569–1572.
- Sever, R. & Glass, C.K., 2013. Signaling by nuclear receptors. *Cold Spring Harbor Perspectives in Biology*, 5(3), pp. 16709–16709.
- Shiau, A.K. et al., 1998. The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen. *Cell*, 95, pp.927–937.
- Shulman, A.I. et al., 2004. Structural determinants of allosteric ligand activation in RXR heterodimers. *Cell*, 116, pp.417–429.
- Sierra, M.L. et al., 2007. Substituted 2-[(4-Aminomethyl)phenoxy]-2-methylpropionic Acid PPAR α Agonists. 1. Discovery of a Novel Series of Potent HDLc Raising Agents. *Journal of Medicinal Chemistry*, 50(4), pp.685–695.
- Silva, J.E., 2006. Thermogenic Mechanisms and Their Hormonal Regulation. *Physiological Reviews*, 86(2), pp.435–464.
- Sorimachi, K. & Robbins, J., 1978. Uptake and metabolism of Thyroid Hormones by cultured monkey hepatocarcinoma cells. *Biochemical and Biophysical Acta*, 542,

pp.515–526.

- St Germain, D.L. & Galton, V.A., 1997. The deiodinase family of selenoproteins. *Thyroid*, 7, pp.655–668.
- Stanley, M.M. & Astwood, E.B., 1949. The response of the thyroid gland in normal human subjects to the administration of thyrotropin, as shown by studies with IL31. *Endocrinology*, 44, pp.49–60.
- Takeda, T. et al., 1995. Triiodothyroacetic acid has unique potential for therapy of resistance to thyroid hormone. *The Journal of Clinical Endocrinology & Metabolism*, 80(7), pp.2033–2040.
- Takeshita, A. et al., 1997. TRAM-1, A Novel 160-kDa Thyroid Hormone Receptor Activator Molecule, Exhibits Distinct Properties from Steroid Receptor Coactivator-1. *The journal of biological chemistry*, 272, pp.27629–27634.
- Taneja, R. et al., 1997. Phosphorylation of activation functions AF-1 and AF-2 of RAR alpha and RAR gamma is indispensable for differentiation of F9 cells upon retinoic acid and cAMP treatment. *The EMBO Journal*, 16, pp.6452–6465.
- Tata, J.R. & Widnell, C.C., 1966. Ribonucleic Acid synthesis during the early action of thyroid hormones. *Journal of Biological Chemistry*, 98, pp.604–619.
- Terwilliger, T.C. et al., 2008. Iterative-build OMIT maps: map improvement by iterative model building and refinement without model bias. *Acta Crystallographica Section D Biological Crystallography*, 64(Pt 5), pp.515–524.
- Thompson, C.C. et al., 1987. Identification of a novel Thyroid Hormone Receptor expressed in the mammalian Central Nervous System. *Science*, 237, pp.1610–1614.
- Thompson, P.D. et al., 1998. Heterodimeric DNA Binding by the Vitamin D Receptor and Retinoid X Receptors Is Enhanced by 1,25-Dihydroxyvitamin D₃ and Inhibited by 9-*cis*-Retinoic Acid. *The journal of biological chemistry*, 273, pp.8483–8491.
- Thornton, J.W. & Desalle, R., 2000. A new method to localize and test the significance of incongruence: detecting domain shuffling in the nuclear receptor superfamily. *Systematic Biology*, 49, pp.183–201.
- Toyoda, N. et al., 1997. Structure-Activity Relationships for Thyroid Hormone Deiodination by Mammalian Type I Iodothyronine Deiodinases. *Endocrinology*, 138, pp.213–218.
- Tylki-Szymańska, A. et al., 2015. Thyroid hormone resistance syndrome due to mutations in the thyroid hormone receptor α gene (*THRA*). *Journal of Medical Genetics*, 52(5), pp.312–316.
- Umesono, K. & Evans, R.M., 1989. Determinants of Target Gene Specificity for Steroid/Thyroid Hormone Receptors. *Cell*, 57, pp.1139–1146.
- Umesono, K. et al., 1991. Direct repeats as selective response elements for the thyroid hormone retinoic acid, and vitamin D₃ receptors. *Cell*, 65, pp.1255–1266.

- van Gucht, A.L.M. et al., 2016. Resistance to Thyroid Hormone Alpha in an 18-Month-Old Girl: Clinical, Therapeutic, and Molecular Characteristics. *Thyroid*, 26(3), pp.338–346.
- van Mullem, A.A. et al., 2012. Clinical Phenotype and Mutant TR alpha 1. *New England Journal of Medicine*, 366(15), pp.1–3.
- van Mullem, A.A. et al., 2013. Clinical Phenotype of a New Type of Thyroid Hormone Resistance Caused by a Mutation of the TR α 1 Receptor: Consequences of LT4 Treatment. *The Journal of Clinical Endocrinology & Metabolism*, 98(7), pp.3029–3038.
- Venäläinen, T. et al., 2009. Molecular mechanism of allosteric communication in the human PPAR α -RXR α heterodimer. *Proteins: Structure, Function, and Bioinformatics*, 78(4), pp.873–887.
- Venero, C., 2005. Anxiety, memory impairment, and locomotor dysfunction caused by a mutant thyroid hormone receptor 1 can be ameliorated by T3 treatment. *Genes & Development*, 19(18), pp.2152–2163.
- Visser, T.J. et al., 1998. Characterization of thyroid hormone sulfotransferases. *Chemico-Biological Interactions*, 109, pp.279–291.
- Voegel, J.J. et al., 1998. The coactivator TIF2 contains three nuclear receptor-binding motifs and mediates transactivation through CBP binding-dependent and -independent pathways. *The EMBO Journal*, 17(2), pp.507–519.
- Wadia, J.S., Stan, R.V. & Dowdy, S.F., 2004. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nature Medicine*, 10(3), pp.310–315.
- Wagner, R.L. et al., 1995. A structural role for hormone in the thyroid hormone receptor. *Nature Clinical Practice Endocrinology & Metabolism*, 378, pp.690–697.
- Wagner, R.L. et al., 2001. Hormone Selectivity in Thyroid Hormone Receptors. *Molecular Endocrinology*, 15(3), pp.398–410.
- Walensky, L.D. & Bird, G.H., 2014. Hydrocarbon-Stapled Peptides: Principles, Practice, and Progress. *Journal of Medicinal Chemistry*, 57(15), pp.6275–6288.
- Walensky, L.D. et al., 2004. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science*, 305, pp.1466–1470.
- Wang, C.Y., 2006. Structural Model Reveals Key Interactions in the Assembly of the Pregnane X Receptor/Corepressor Complex. *Molecular Pharmacology*, 69(5), pp.1513–1517.
- Watkins, R.E. et al., 2001. The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. *Science*, 292, pp.2329–2333.
- Webb, P. et al., 2000. The nuclear receptor corepressor (N-CoR) contains three isoleucine motifs (I/LXXII) that serve as receptor interaction domains (IDs).

- Molecular Endocrinology*, 14, pp.1976–1985.
- Weinberger, C. et al., 1986. The c-erb-A gene encodes a thyroid hormone receptor. *Nature Clinical Practice Endocrinology & Metabolism*, 324(6098), pp.641–646.
- Wen, Y.D. et al., 2000. The histone deacetylase-3 complex contains nuclear receptor corepressors. *PNAS*, 97, pp.7202–7207.
- Wilson, T.E., Fahrner, T.J. & Milbrandt, J., 1993. The orphan receptors NGF1-B and Steroidogenic Factor 1 establish monomer binding as a third paradigm of Nuclear Receptor-DNA interaction. *Molecular and Cellular Biology*, 13, pp.5794–5804.
- Wurtz, J.M. et al., 1996. A canonical structure for the ligand-binding domain of nuclear receptors. *Nature Structural Biology*, 3(1), pp.87–94.
- Xu, H.E. et al., 2002. Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPAR α . *Nature Clinical Practice Endocrinology & Metabolism*, 415, pp.813–817.
- Yan, X. et al., 2004. Dynamics and ligand-induced solvent accessibility changes in human retinoid X receptor homodimer determined by hydrogen deuterium exchange and mass spectrometry. *Biochemistry*, 43(4), pp.909–917.
- Yang, X. et al., 1996. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature Clinical Practice Endocrinology & Metabolism*, 382, pp.319–325.
- Yang, X. et al., 2006. Nuclear Receptor Expression Links the Circadian Clock to Metabolism. *Cell*, 126(4), pp.801–810.
- Yang, X., Zhang, F. & Kudlow, J.E., 2002. Recruitment of O-Glc-N-ac-transferase to promoters by corepressor mSin3A: Coupling protein O-Glc-N-Acylation to transcriptional repression. *Cell*, 110, pp.69–80.
- Yang, Z R et al., 2005. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics*, 21(16), pp.3369–3376.
- Yoon, H.G. et al., 2003. Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. *The EMBO Journal*, 22(6), pp.1336–1346.
- Yu, J. et al., 2003. A SANT motif in the SMRT corepressor interprets the histone code and promotes histone deacetylation. *The EMBO Journal*, 22, pp.3403–3410.
- Yu, J. et al., 2009. Roll: a new algorithm for the detection of protein pockets and cavities with a rolling probe sphere. *Bioinformatics*, 26(1), pp.46–52.
- Yu, V.C. et al., 1991. RXR beta: A Coregulator That Enhances Binding of Retinoic Acid, Thyroid Hormone, and Vitamin D Receptors to Their Cognate Response Elements. *Cell*, pp.1251–1266.