ENGINEERING THE PREGNANE X RECEPTOR AND ESTROGEN RECEPTOR ALPHA TO BIND NOVEL SMALL MOLECULES USING NEGATIVE CHEMICAL COMPLEMENTATION

A Thesis
Presented to
The Academic Faculty

by

Hally A. Shaffer

In Partial Fulfillment
of the Requirements for the Degree
Doctorate of Philosophy in the
School of Chemistry and Biochemistry

Georgia Institute of Technology
May 2011

ENGINEERING THE PREGNANE X RECEPTOR AND ESTROGEN RECEPTOR ALPHA TO BIND NOVEL SMALL MOLECULES USING NEGATIVE CHEMICAL COMPLEMENTATION

Approved by:

Dr. Donald Doyle, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Bahareh Azizi, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Andreas Bommarius, Coadvisor
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Nick Hud
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Sheldon May
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Adegboyega Oyelere
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Loren Williams, Coadvisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Date Approved: April 4, 2011

To my Lord and Savior whose Spirit guided me through this work.

ACKNOWLEDGEMENTS

I have been privileged with many mentors throughout my educational career, without whom I may not have pursued a graduate degree. Likewise, the process of obtaining my Ph. D. has been possible through the support and encouragement of many different people. I would like to give special thanks to Dr. Bahareh Azizi for her continual dedication toward teaching me to love and tackle science. Throughout my graduate career, she has been my mentor, teacher, and friend as she pushed me beyond what I thought was capable and then helped lead the way toward the goal. This work is an outpouring of that effort on her part. I would like to thank Dr. Donald Doyle for developing the various nuclear receptor *and* yeast projects on which I worked. Through the challenges of the different projects, I was able to develop as a scientist.

I would like to thank Dr. Andreas Bommarius, Dr. Nick Hud, Dr. Sheldon May, Dr. Adegboyega Oyelere, and Dr. Loren Williams, who served on my committee the past several years. I would also like to thank Dr. Wendy Kelly, who served on my committee during my first several years at Tech. I would like to give thanks to Dr. Andreas Bommarius and Dr. Loren Williams. Andy has provided insight into my projects from a viewpoint I would not have otherwise seen. The past couple years, Dr. Loren Williams has supported my work and was always available to answer questions when I stopped by his office. I would also like to thank Dr. Charles Liotta and the Chemistry & Biochemistry department for their support offered the past couple years.

The work done in any lab is never alone, and therefore, there are many lab members that deserve thanks. I would like to thank previous members Dr. Terry Watt and

Dr. Kenyetta Johnson. Terry provided scientific advice and help with some of my initial work. Even though she has graduated, Kenyetta still receives emails from me and continues to answer my questions. I would also like to thank James Kratzer who was always there listening to us ladies go on and on and then how he continued to stop by to say “hi.” I would like to thank Amanda Ousley for her service to me during our time here. Without her rides, I literally would not have been able to get any research done. It was good to share an office and lab bench with her. I would like to thank Jennifer Taylor for training me in cell culture and, more importantly, for the many late night discussions throughout the past couple years that has developed into a strong friendship. I would like to thank Hilda Castillo for her friendship and consistent gentle nature. I’m glad there was another lab member who paid attention to detail as much as (or more than) I did. I would also like to thank Anna Duraj-Thatte for training me to do the modeling and the many different scientific discussions about ER. Through that process, it was fun helping you say “yeast” and not “east.” I would like to thank Michael Rood for his help the past several years. You were always there to discuss science, math, and food. I’ll always laugh thinking about the power of genetic selection.

Many other scientists have had their hand in helping me through this work. I would like to thank Heba Gill for continuing to setup quant after quant and did hundreds of noncolumn minipreps. Through her work, she also took the time to understand negative chemical complementation. I would also like to thank Lena James for her work this past summer. I think she won the award for the most noncolumn minipreps in one day. Janna Blum was also always there, especially when she moved her bench closer to mine. Her expertise obviously goes beyond the bench-work when you look at the many

experiences of her life that she has shared. Additional 3D wing friends and supporters have been Pritha Bagchi whose calm, yet determined nature amazes me. I would also like to thank Dr. Karl Huettinger and Dr. Raegan McRae for their help at different times throughout my work. I also thank Prabuddha Bransal for his help with some statistical work. It may go without saying, but I also thank the significant others of all those I've mentioned who, by indirect relation, were also a support through all of it. I would also like to thank the Student Teacher Enhancement Partnership (STEP) program for teaching me more about the field of education and the many possibilities it holds.

I would like to thank Smyrna Christian Church (I wish I could list all of you) as well as Blake and Leslie Stephens for their support and prayers throughout my time at Tech. That is a huge unseen role involved in this work. In addition, I would like to thank Berry College. The entire Chemistry department was always there to help me through my first semester of teaching as a faculty member while I was finishing writing.

My family has certainly been there the longest. All of them have been there during my time before undergrad as well as throughout my graduate career. I wholeheartedly thank my mom, Peter, my dad, and Joan. All of you played a role, and are continuing to do so, in this process with me. You were always there to listen and then give me courage to get back on the horse. I would like to thank Sara, Annie, Jessie, and Emily. We had such fun times. Keep pushing Sara and Annie. It is well worth it. Thank you for your support. And, to my one and only twin, I wish I could express it all. Marci, you have been my biggest role model. Thank you.

Oh my beloved Zack. Thank you for everything: your prayers, hugs, for sleepless nights with me, for doing all the chores over and over and over. You are the most patient

and loving man I know. Thank you for your solid, continual support to motivate me to run this race. I thank my God every time I remember you (Phil 1:3).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

3-AT	3-amino-1,2,4-triazole
OHT	4-hydroxytamoxifen
5'FOA	5'Floroorotic acid
5'FU	5'Fluorouracil
9cRA	9-*cis* retinoic acid
E2	17-β estradiol
17αE2	17-α estradiol
AF-2	Activation Function 2
ACTR	Activator for Thyroid and Retinoid Receptors
AR	Androgen Receptor
DBD	DNA-binding Domain
EC_{50}	Half maximal effective concentration
ER	Estrogen Receptor
GAD	Gal4 Activation Domain
GBD	Gal4 DNA-binding Domain
GR	Glucocorticoid Receptor
HDAC	Histone Deacetylase
IC_{50}	Half maximal inhibitory concentration
LBD	Ligand-binding Domain
LXR	Liver X Receptor
MR	Mineralocorticoid Receptor
ONPG	*o*-nitrophenyl β-D-galactopyranoside
PDB	Protein Data Bank

PXR	Pregnane X Receptor
RE	Response Element
RNAP	RNA Polymerase
RXR	Retinoid X Receptor
SERM	Selective Estrogen Receptor Modulator
SRC-1	Steroid Receptor Coactivator-1
Y2H	Yeast Two Hybrid

SUMMARY

Nuclear receptors are ligand-activated transcription factors that play significant roles in various biological processes within the body, such as cell development, hormone metabolism, reproduction, and cardiac function. As transcription factors, nuclear receptors are involved in many diseases, such as diabetes, cancer, and arthritis, resulting in approximately 10-15% of the pharmaceutical drugs presently on the market being targeted toward nuclear receptors. Structurally, nuclear receptors consist of a DNA-binding domain (DBD), responsible for binding specific sequences of DNA called response elements, fused to a ligand-binding domain (LBD) through a hinge region. The LBD binds a small molecule ligand. Upon ligand binding, the LBD changes to an active conformation leading to the recruitment of coactivator (CoAC) proteins and initiation of transcription. As a result of their involvement in disease, there is an emphasis on engineering nuclear receptors for applications in gene therapy, drug discovery and metabolic engineering.

Chemical complementation uses an engineered yeast strain, PJ69-4A, which contains Gal4 response elements (Gal4 REs) controling expression of the *ADE2* and *HIS3* genetic selection markers, along with the *lacZ* gene, which can be used for a colorimetric screening assay. In chemical complementation, the interaction between a nuclear receptor and a small molecule is linked to the survival of the yeast cells in selective media. Chemical complementation uses two fusion proteins: a fusion of the Gal4 DBD with the nuclear receptor LBD and a fusion of the nuclear receptor CoAc with the Gal4 activation

domain (GAD). Chemical complementation can be used for drug discovery as well as protein engineering applications.

The original focus of this work was to use chemical complementation to engineer a nuclear receptor to bind and activate transcription in response to a novel small molecule, specifically the following β-lactam antibiotics: penicillin G, oxacillin, cloxacillin, ampicillin, amoxicillin, or nafcillin. The first target nuclear receptor to engineer to activate transcription in response to the β-lactam antibiotics was the pregnane X receptor (PXR), a nuclear receptor involved in xenobiotic metabolism. A PXR library toward the antibiotics was designed based on crystal structure analysis as well as amino acid sequence alignments. The PXR library was assessed in the chemical complementation system using selective plates containing the antibiotics. All of the variants obtained were constitutively active, meaning they grow without the presence of an exogenous small molecule. To overcome the constitutive activity, three different approaches were taken. First, a knockout yeast strain was developed that could remove any potential endogenous ligands in the yeast that could be activating PXR in chemical complementation. Second, the nuclear receptor was changed to the human estrogen receptor alpha (hERα). Lastly, a selection system called negative chemical complementation was utilized to remove the constitutively active variants from a library of nuclear receptor variants.

Yeast steroid metabolism is different from mammalian cholesterol metabolism. Therefore, to develop a knockout strain, the *ERG24* gene, involved in the yeast steroid biosynthetic pathway, was replaced by a *Kan* selection marker, which renders yeast resistant to the antibiotic geneticin. Removing the *ERG24* gene removes steroid

metabolites that could be causing the constitutive activity observed with PXR. The knockout strain was engineered; however, a pure knockout strain was not obtained. The engineered strain contained both the parent genome and the *ERG24* knockout. This strain was therefore not used to engineer nuclear receptor variants.

The second approach was to design a library of hERα variants to be activated by the β-lactam antibiotics. The residues chosen to mutate in the hERα library was based on modeling the antibiotics into the pocket of the crystal structure with the natural ligand 17-β estratiol (E2). The mutations were chosen to introduce hydrogen bonding potentials into the ligand binding pocket (LBP). When put through selection, the hERα library produced functional variants that were activated by E2, however none were activated by the β-lactam antibiotics. The rational library produced three variants that contain a point mutation at residue L346: L346M, L346V, and L346E, which displayed varying levels of sensitivity toward E2. Therefore, this position was saturated to investigate the role this residue has in ligand binding and activity and to determine if changing the chemical properties at position 346 would cause the variant to bind a different steroid molecule.

The last approach taken to overcome constitutive activity in chemical complementation was to use negative chemical complementation, a selection system that uses the *URA3* selection marker. In negative chemical complementation when a ligand activates the nuclear receptor and the *URA3* gene is expressed in media containing 5'fluoorotic acid (5'FOA), the toxin 5-fluorouracil (5'FU) is produced, leading to cell death. To engineer a nuclear receptor to be activated by the β-lactam antibiotics, a random mutagenic library of hERα was put through selection in media containing 5'FOA to remove the constitutively active variants from the library. This method removed most

of the constitutively active variants and produced functional hERα variants, but none of the variants were activated by the antibiotics.

Negative chemical complementation can also be used in a high-throughput method toward drug discovery applications. Various ligands function as nuclear receptor activators or repressors, known as agonists or antagonists. hERα was therefore tested with the current drugs 4-hydroxy tamoxifen (OHT) and the potent antagonist Fulvestrant to assess the use of this system for drug discovery of potential antagonists. Growth above basal was not observed in the negative selection system using uracil selective media with 5'FOA with an antagonist. Initial work toward developing a positive selection method using nuclear receptor corepressors to discover nuclear receptor antagonists was performed. Modifying this selection system could not only engineer nuclear receptors to bind novel small molecules, but also make nuclear receptor antagonist drug discovery possible.

CHAPTER 1

NUCLEAR RECEPTORS

Nuclear receptors are a diverse superfamily of ligand-activated transcription factors that play significant roles in various biological processes within the body, such as cell development, hormone metabolism, reproduction, cardiac function, and immune responses [1-3]. There are 48 known human nuclear receptors that bind various ligands, including hormones, retinoids, fatty acids, and xenobiotics [1, 4, 5]. Commonly known nuclear receptors include the estrogen receptor (ER), the retinoid X receptor (RXR), and the liver X receptor (LXR) [1, 2, 4-6]. These nuclear receptors are divided into several subfamilies based on the type of ligand each receptor binds [4, 7]. Nuclear receptors that are referred to as orphan receptors do not have an identified physiological ligand [1, 3, 4, 6]. Nuclear receptors contain several isotypes, such as α, β, and γ, as a result of different promoter usage as well as RNA-splicing. These isotypes are all involved in individual biological pathways, thus adding an additional level of complexity to this family of proteins [8].

Before nuclear receptors were cloned, the biological effect of these receptors was known. Studies had showed that steroids played a role in inflammation, yet the mechanism of action was not identified [4, 7]. The first human nuclear receptor that was cloned was the glucocorticoid receptor (GR), soon followed by other steroid receptors, such as the estrogen receptor [4, 7, 9]. In comparing the different sequences, conservation in the various domains was observed, leading to the current classifications of nuclear receptors within the superfamily [4, 7]. The classifications are also based on the types of ligands that are known to interact with these receptors [4, 7].

Their role as transcription factors and regulators of gene expression results in the involvement of these receptors in various metabolic and physiological related diseases, such as diabetes, obesity, arthritis, and cancer [1, 10]. As a result, approximately 10-15% of the pharmaceutical drugs presently on the market are targeted toward nuclear receptors, as ligands for these receptors [11]. In fact, in 2003, thirty-four out of the two-hundred most commonly prescribed drugs were targeted toward nuclear receptors [7]. Examples of current drugs include thiazolidinediones for the peroxisome proliferator-activated receptor gamma (PPARγ) and dexamethasone for the glucocorticoid receptor (GR), targeting type II diabetes and inflammatory diseases, respectively [1, 10, 12, 13]. One common drug used to inhibit expression of genes regulated by the human estrogen receptor alpha (hERα) is 4-hydroxytamoxifen (OHT), which is currently prescribed to treat breast cancer. Mifepristone (RU486) is also used as an anti-progestin and anti-glucocorticoid for the progesterone receptor (PR) and glucocorticoid receptor (GR) [14, 15]. However, there is still a vast interest in the pharmaceutical industry to discover novel nuclear receptor ligands for additional therapeutic applications [3, 7].

1.1 Nuclear Receptor Structure

Nuclear receptors have a modular structure consisting of five to six common domains, known as domains A through F (Figure 1.1) [1, 2, 4, 8, 16, 17]. The A/B domain has the highest variability in the nuclear receptor superfamily, containing the ligand-independent activation function 1 (AF-1) with a poorly defined structure [4, 18, 19]. The A/B domain has been thought to be involved in mediating gene activation and contains phosphorylation sites for providing interactions with other domains or proteins [18-20]. Domains C and E, the DNA-binding domain (DBD) and ligand-binding domain

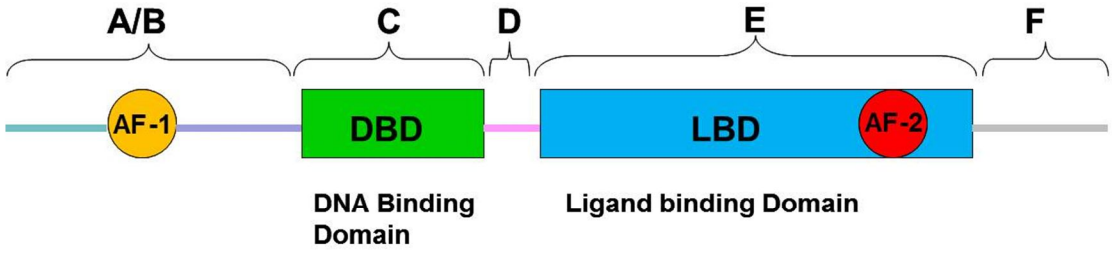


Figure 1.1 Scheme of nuclear receptor domains.

(LBD), respectively will be discussed below. These two domains are connected through the D domain, a hinge region, providing rotational flexibility between domains C and E. The F domain, located at the C-terminus, also does not have a defined function, although this domain has been implicated in regulating nuclear receptor interactions with other proteins [16, 21, 22].

The DNA-binding domain (DBD) and ligand-binding domain (LBD) of nuclear receptors, domains C and E, are evolutionarily conserved [1, 12, 23, 24]. Between the DBD and LBD, the DBD is the most conserved domain among all of the nuclear receptor domains [3, 4]. The DBD consists of zinc fingers, two zinc ions coordinated by two cysteine or histidine residues, which bind specific DNA sequences called response elements (REs) [2, 12, 23, 25, 26]. Nuclear receptors bind a variety of response elements, that consist of multiple copies of a consensus sequence either as direct repeat (DR) or inverted repeat (IR) of the sequence, often separated by a spacer sequence; thus providing the specificity for nuclear receptor binding [23, 26, 27]. Steroid receptors mainly recognize the sequence AGAACA, whereas non-steroid receptors recognize the sequence AGGTCA [23]. Most nuclear receptors also function as homo or heterodimers, a function that is regulated by the recognition of the DBD with the response element [3, 4]. Figure 1.2 shows the DBDs of a crystal structure of two nuclear receptors in a heterodimer bound to a DNA response element. The response element consisting of a direct repeat of two consensus sequences for the two different receptors involved in the heterodimer can be seen (Figure 1.2).

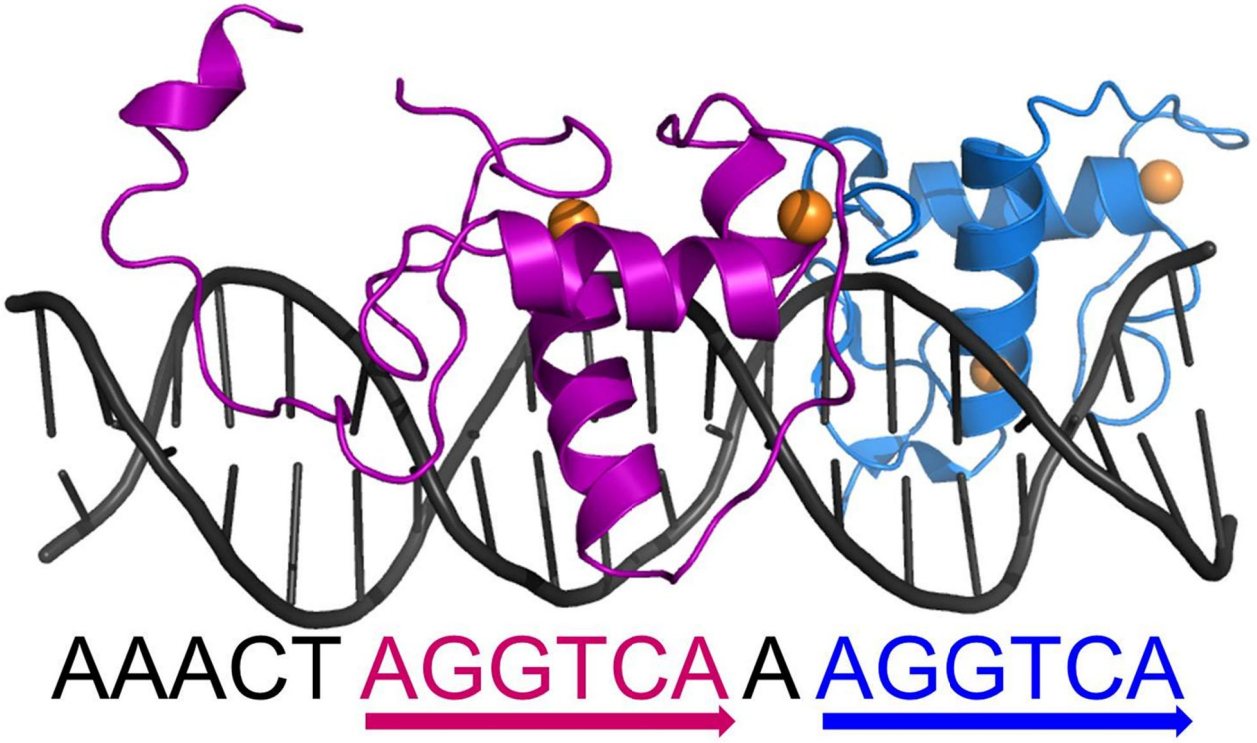


Figure 1.2. Crystal structure of the DBDs of PPARγ with RXRα bound to DNA (PDB 3DZY). PPARγ is shown in purple and RXRα is shown in marine. The DNA is grey and the zinc ions are shown as spheres and colored orange. The consensus sequence response element is shown below the structure. The sequence colors correlate to the respective nuclear receptor DBDs. An upstream extension site and a spacer nucleotide are shown in black. Arrows indicate the 5' to 3' regions of the response element repeat.

In most nuclear receptors, the three-dimensional structure of the LBD consists of primarily an α-helical structure with two β-sheets that fold into an α-helical sandwich with one main β-turn [24]. The helices within the LBD are numbered one through twelve [8, 17, 28]. This numbering system is based on the original crystal structure of the *apo* RXR LBD, which contained twelve defined α-helices [28]. Some nuclear receptors, such as the pregnane X receptor (PXR) or the vitamin D receptor (VDR), however, have additional helices creating a larger LBD [17]. The carboxy-terminal of the LBD includes the ligand-dependent activation function 2 (AF-2), located within Helix 12. The AF-2 is crucial for the functionality of activating these receptors. When the ligand binds to the LBD, the AF-2 domain is repositioned into an active conformation, making contacts with helices 3, 4/5, and 10/11 [1, 3]. The conformational change is clearly seen in the difference between the *apo* and *holo* structures of the retinoid X receptor alpha (RXRα), where Helix 12 is repositioned over the ligand binding pocket in the *holo* structure (Figure 1.3) [7, 28, 29]. Although Helix 12 makes the largest change, helices 3 and 10/11 also shift slightly to form the α-helical sandwich around the ligand.

The ligand binding pocket (LBP) is formed from helices 3, 7, and 12, although residues from additional helices make contacts with the ligand [3, 7, 17]. The LBPs can vary in size and shape, having pocket volumes ranging from less than 400 $Å^3$ to greater than 1500 $Å^3$ [3, 7, 30]. Despite the size and shape differences, nuclear receptor LBP's all have similar characteristics. The core of the pocket contains primarily hydrophobic residues, making key hydrophobic contacts for ligand binding [3]. The pocket often also contains important hydrogen bonding residues, used to anchor the ligand into the pocket [5, 7, 17]. For example, 17-β estradiol (E2) forms two hydrogen bonds with residues in

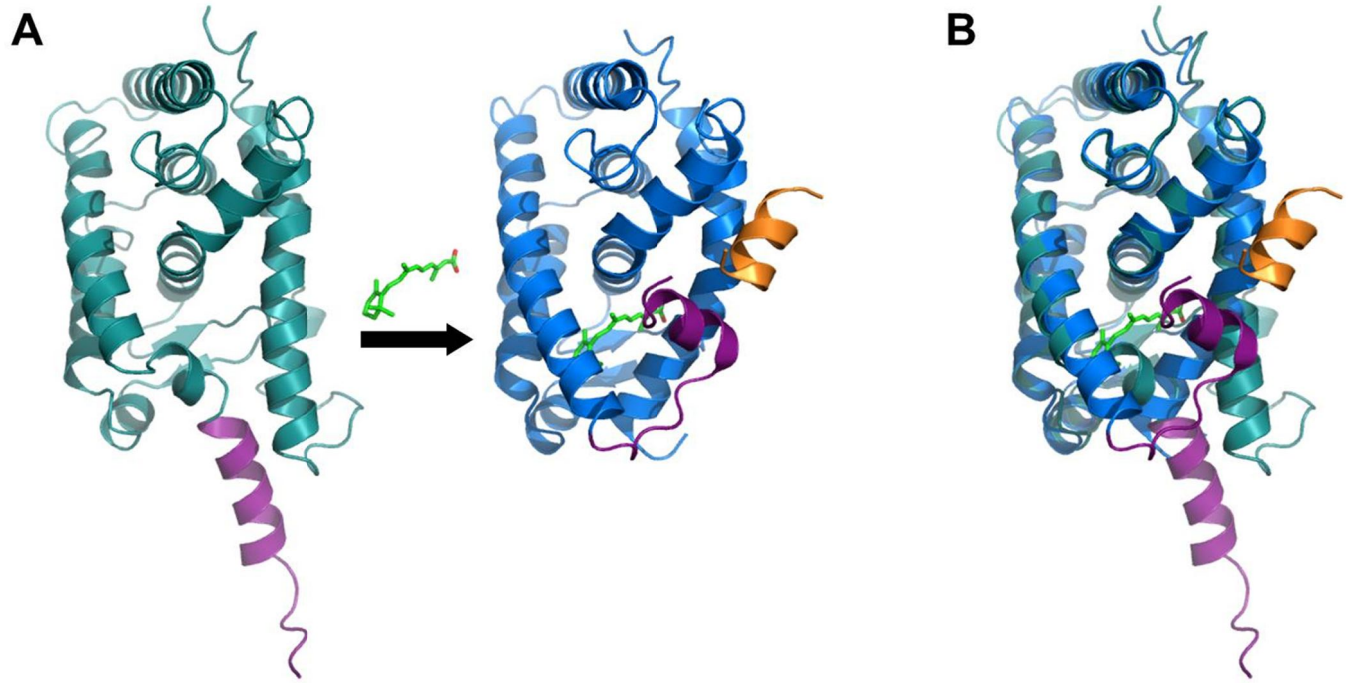


Figure 1.3. The ligand binding domain (LBD) of RXRα. (A) The *apo* and *holo* RXRα LBD (PDB files 1LBD and 1XDK). The *apo* structure is shown in dark green. The *holo* structure is shown in marine with the coactivator peptide shown in orange. In both structures, Helix 12 is colored in purple. The ligand, 9-*cis* retinoic acid is displayed as sticks and colored green (B) A superimposition of the *apo* and *holo* LBDs of RXRα. The color scheme is the same as in (A).

the LBP, located on opposite ends of the pocket, of the estrogen receptor [31]. A glutamate in helix 3, E353, an arginine in helix 6, R394, and a water molecule form a network of hydrogen bonds with the 3-hydroxy group of E2 at one end of the pocket. Histidine 524, located in helix 11 on the opposite side of the pocket, also forms a hydrogen bond with the 17-hydroxy group of E2 [3, 31].

Although the structural elements of the DBD and LBD have been studied extensively and the general functions of these domains are identified, the first full-length crystal structure containing all the domains of the nuclear receptors was only recently published (Figure 1.4) [19]. The complete structure is a heterodimer between RXRα and PPARγ, each bound to a ligand, a coactivator peptide, and the DNA [19]. Chandra *et al.* show that the PPARγ LBD plays a significant role in directing the DBDs of both nuclear receptors to bind the DNA [19]. In the crystal structure, the A/B domains of both receptors could not be visualized due to the highly unstructured nature of that domain. In addition, the structure highlights the interface of multiple domains, providing evidence that a larger network of interactions among the nuclear receptor domains and the DNA exists [19].

1.2 Nuclear Receptor Function

The mechanism of nuclear receptor activation and gene regulation involves multiple proteins and requires that several important processes take place. Nuclear receptors are either localized in the cytoplasm or in the nucleus, depending on the type of receptor. When ligand binds, nuclear receptors in the cytoplasm translocate into the nucleus, interacting with their respective response elements and activating gene expression [1, 3, 4, 32].

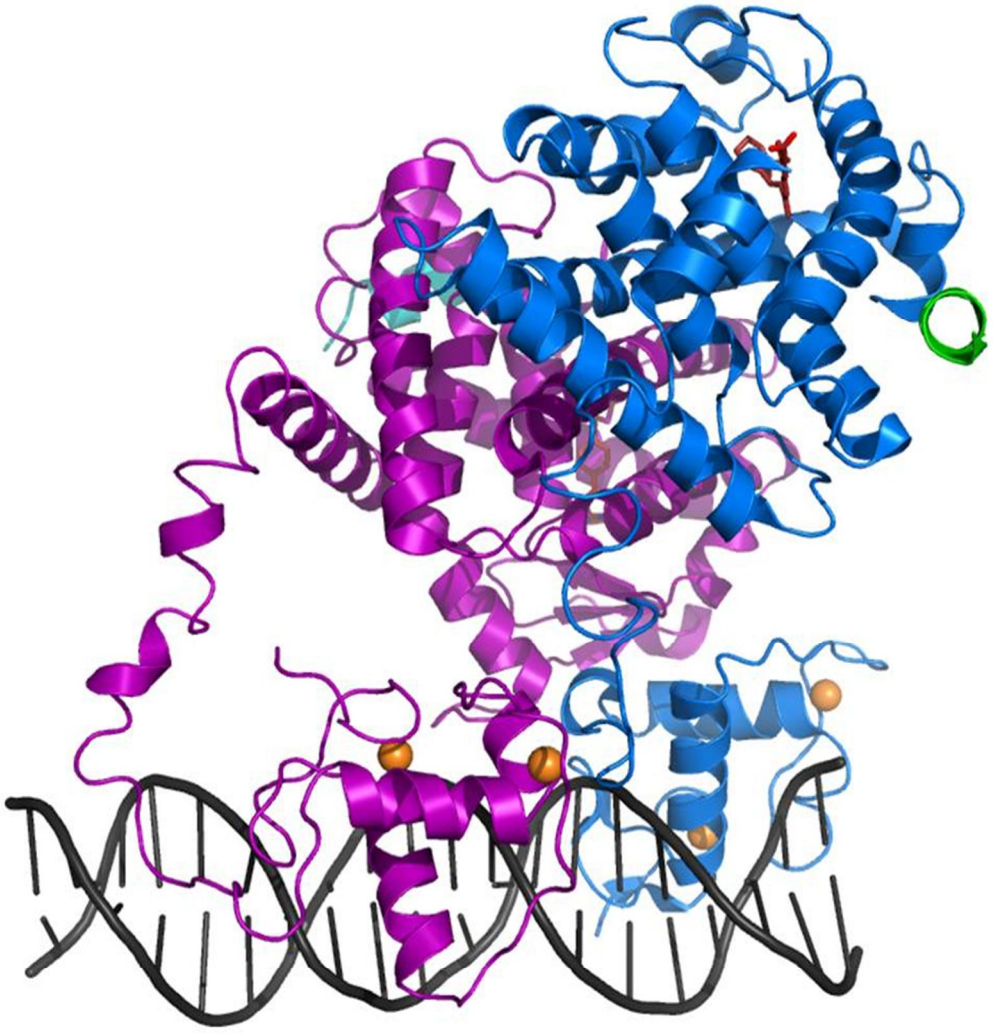


Figure 1.4. Full length crystal structure of PPARγ with RXRα bound to DNA (PDB 3DZY). PPARγ is shown in purple with the coactivator peptide in cyan. RXRα is shown in marine with the coactivator peptide in green. The ligands are displayed as sticks and colored dark red. The DNA is grey and the zinc ions are shown as spheres and colored orange.

Nuclear receptors also associate with coregulator proteins, known as coactivators (CoAc) and corepressors (CoR) [4, 33]. Both interact with specific domains on the LBD, known as interaction domains (ID), through amino acid motifs made up of LXXLL or LXXXIXXXI/L, where X is any amino acid, for coactivators and corepressors, respectively [4, 33, 34]. These motifs are receptor specific and can be repeated in combinations of one, two, or more, depending on the specific coregulator protein [35-37]. The IDs on the LBD are hydrophobic clefts formed by helices 3, 4, and 12, thus creating the region for coregulator interaction. Coactivators and corepressors associate with similar regions in this hydrophobic cleft on the LBD, but not identical locations due to the differences in the structural fold of the active versus inactive receptor [1].

As previously mentioned, nuclear receptors are transcription factors that regulate gene expression through small molecule ligands [4, 8, 24, 38]. Functionally, in the absence of ligand, most nuclear receptors are bound to corepressor proteins, such as the silencing mediator for retinoid and thyroid hormone receptors (SMRT), and the receptor is in an inactive conformation. Corepressors are then involved in recruiting other proteins, such as histone deacetylases (HDACS) that deacetylate histones, condensing the chromatin, inhibiting RNA polymerase from binding to the DNA and therefore transcription cannot occur [1, 34, 39, 40]. However, upon ligand binding, the LBD undergoes a conformational change, positioning Helix 12 in the active conformation [1, 3, 24]. Corepressor proteins dissociate from the nuclear receptor, allowing coactivators, such as the steroid receptor coactivator 1 (SRC-1) and the activator for thyroid and retinoid receptors (ACTR) to associate with the active LBD (Figure 1.5) [33, 34, 40-44]. These coactivators then recruit histone acetyl transferases (HATs) that acetylate histones

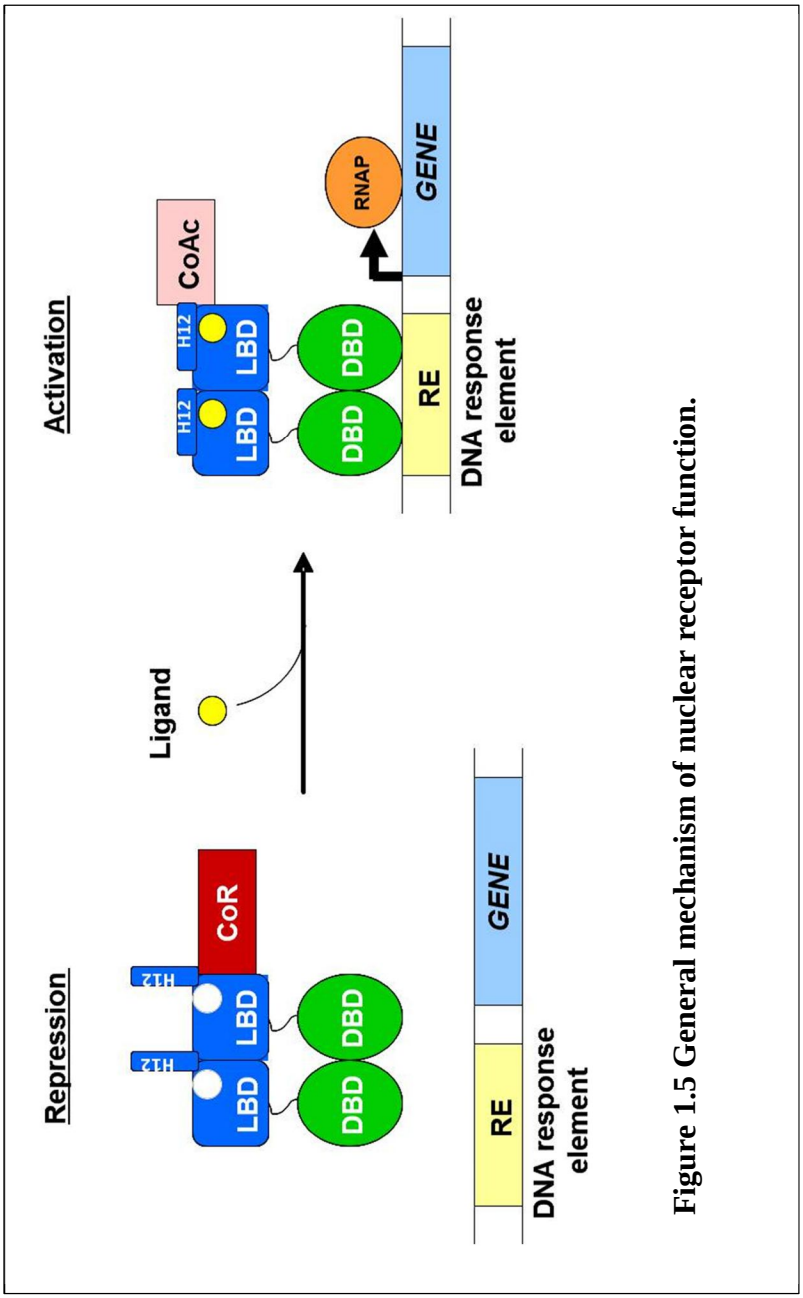


Figure 1.5 General mechanism of nuclear receptor function.

causing the destabilization of the chromatin, leading to the recruitment of RNA polymerase (RNAP) to activate transcription [1, 3, 14, 24, 38, 39].

Nuclear receptors interact with two main types of ligands, agonists and antagonists. Agonists are small molecule ligands that bind and activate the nuclear receptor, through inducing the proper conformational change necessary in the AF-2 domain of the LBD. Antagonists, on the other hand, are small molecule ligands that bind to the nuclear receptor, but do not cause the proper conformational change to occur. Therefore corepressors associate with the nuclear receptor and transcription is repressed [14, 45]. Due to the variety of ligands and large network of interactions with which nuclear receptors are associated, investigation towards understanding these interactions are continually pursued [24, 46]. One tool that can be used toward further understanding nuclear receptor-ligand interactions is chemical complementation.

1.3 Chemical Complementation

Genetic selection systems have been demonstrated to be powerful tools for discovering and deciphering protein-protein interactions and small molecule-protein interactions [47-51]. Various genetic selection systems have been developed in microorganisms, such as bacteria and yeast. As a simple eukaryote, yeast have proven to be quick, cheap, and accessible hosts for selection, As shown by the success of the yeast two-hybrid system developed by Fields and Song [47, 52-54]. Advances in these genetic selection techniques and the availability of numerous microbial strains allows for the possibility of developing additional *in vivo* genetic selection systems, increasing the

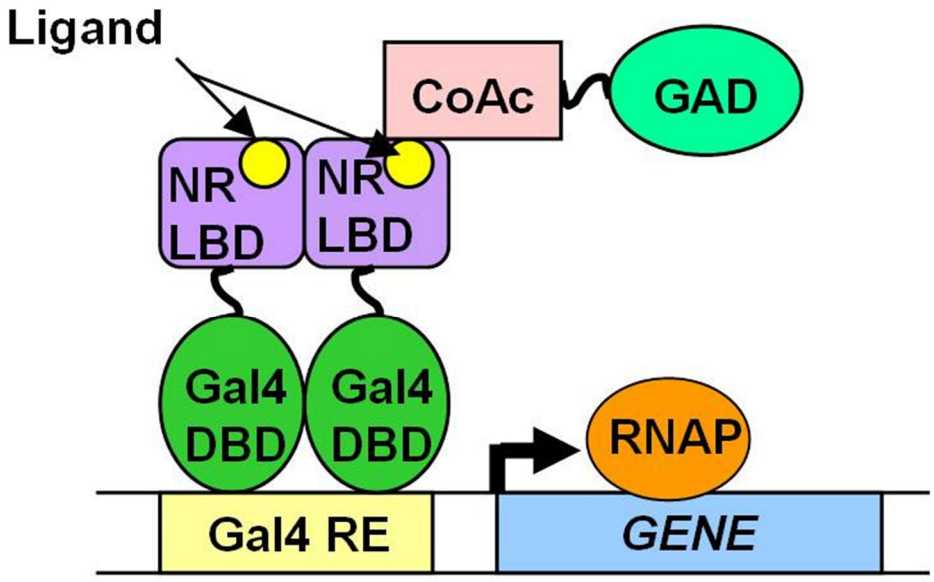


Figure 1.6. Scheme of the chemical complementation selection system. The nuclear receptor (NR) ligand binding domain (LBD) fused to the Gal4 DNA-binding domain (DBD), is introduced into a yeast strain containing Gal4 response elements (Gal4REs) controlling expression of genetic selection genes. Upon ligand binding, a conformational change occurs; recruiting a fusion protein of the NR coactivator (CoAc) and the Gal4 activation domain (GAD) that binds to the LBD, leading to the initiation of transcription of a selection marker. Gene expression occurs and cells survive on media with ligand.

feasibility and efficiency of alternative detection systems for protein-protein, protein-RNA or DNA, and protein-small molecule interactions [55-57].

Our lab developed a genetic selection system in *Saccharomyces cerevisiae*, called chemical complementation, for detecting the interaction of a small molecule with a nuclear receptor [58]. Chemical complementation uses an engineered yeast strain, PJ69-4A, which contains Gal4 response elements (Gal4 REs) controling expression of the *ADE2* and *HIS3* genetic selection markers, along with the *lacZ* gene, which can be used for a colorimetric screening assay [59].

Chemical complementation uses two fusion proteins. The Gal4 DNA-binding domain is fused to the nuclear receptor LBD and a human nuclear receptor coactivator is fused to the Gal4 activation domain (GAD). Gal4 is a ligand-independent transcription factor [60]. Ligand binding to the nuclear receptor LBD causes the recruitment of the CoAc:GAD fusion protein, initiating transcription of the *ADE2* or *HIS3* selection marker required for yeast survival (Figure 1.6). The absence of a ligand, or the inability of the small molecule to activate the nuclear receptor LBD, causes cell death [14].

The advantage of using genetic selection systems in yeast, such as chemical complementation, provide versatile, feasible, and fast methods for deciphering protein-protein and small molecule-protein interactions in comparison to mammalian cell assays [47, 49, 61-63]. Even though yeast are simple eukaryotes, the advantage is they do not possess any endogenous nuclear receptors present in mammalian cells. For example, libraries of synthetic or natural compounds can be assessed in a semi high-throughput manner using a 96-well plate format, where ligands able to bind and activate the nuclear receptor lead to cell growth. In terms of protein engineering, functional protein variants

can be discovered through the survival of the yeast as well. Ultimately, molecular interactions involved in cellular processes can be analyzed through high-throughput selection methods like chemical complementation.

1.4 Applications of Chemical Complementation

1.4.1 Drug Discovery

Nuclear receptors are excellent targets for drug discovery because their essential role as transcriptional regulators leads to their involvement in several diseases [11, 48, 63]. Consequently, a vast number of pharmaceutical companies are currently targeting and developing nuclear receptor ligands as potential therapeutics, either as agonists or antagonists [7, 11]. Chemical complementation has the ability to rapidly and reproducibly analyze large libraries for protein-ligand interactions, and therefore provides a tool for discovering potential, novel, nuclear receptor ligands [64]. Through the high-throughput assay, novel nuclear receptor ligands can be developed.

1.4.2 Engineering nuclear receptors

Chemical complementation is also a powerful tool for directed protein evolution for engineering and discovering functional protein variants within large protein libraries [51, 55, 63]. The ability to influence protein-protein interactions as well as protein-ligand interactions can be used for applications such as gene therapy and metabolic engineering through the ability to control gene expression using a ligand-mutant receptor pair [65]. Examples of this have been shown, where nuclear receptors have been engineered to activate transcription in response to a novel small molecule using various yeast-three hybrid genetic selection systems [58, 64-67].

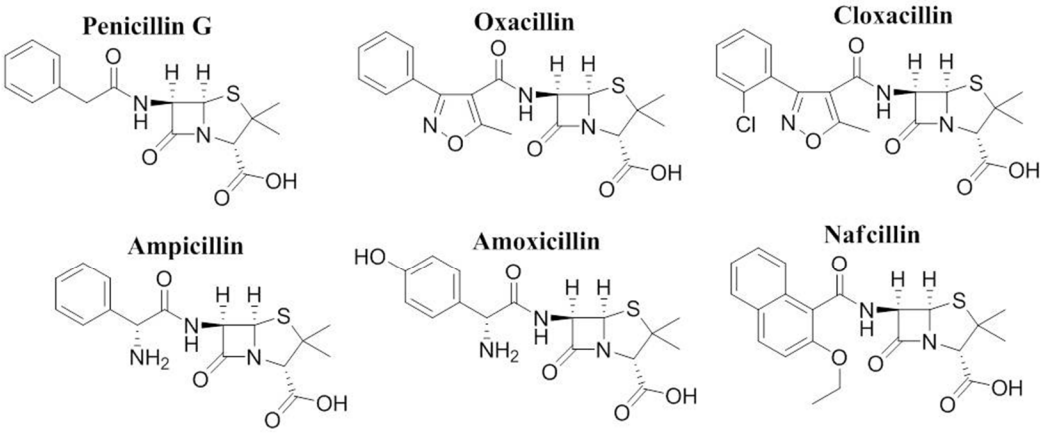


Figure 1.7. Structures of various β-lactam antibiotics.

Previous research has utilized chemical complementation to engineer RXRα variants that are activated by the synthetic ligand LG335 and not by the wild-type ligand 9-*cis* retinoic acid (9cRA) [64]. Directed evolution of hERα has also been successfully performed making new variants with different specificity [65-67]. For example, Zhao *et al.* increased the affinity and specificity of hERα toward testosterone from E2 through a yeast two-hybrid system that screened for testosterone activity [13].

1.5 Using Nuclear Receptors to Develop a Novel Biocatalyst

Nuclear receptor research can also be used to design and engineer novel biocatalysts [24, 46]. Consequently, the original focus of this work was to engineer a nuclear receptor to bind and activate transcription in response to a novel small molecule, specifically any of the following β-lactam antibiotics: penicillin G, oxacillin, cloxacillin, ampicillin, amoxicillin, or nafcillin (Figure 1.7) [68]. In particular, oxcillin, cloxacillin, and nafcillin, which are β-lactamase resistant antibiotics and currently are not synthesized enzymatically. The β-lactam antibiotics have been in clinical use for over 60 years, however the industrial production generates a significant amount of waste [68-70]. To overcome excess waste production, a biocatalytic process where an enzyme is engineered to make the antibiotic would prove extremely valuable. To create a system that can serve as an alternative process for mass producing the β-lactam antibiotics, two steps need to be accomplished to achieve enzyme activated growth. First, a nuclear receptor needs to be engineered to be activated by an antibiotic. Secondly, a library of engineered enzymes will be created to make the antibiotic from precursor molecules (Figure 1.8). To do this, yeast cells containing the desired enzyme, along with the nuclear receptor variant, and the precursors will produce the antibiotic to bind the nuclear receptor and allow the yeast to

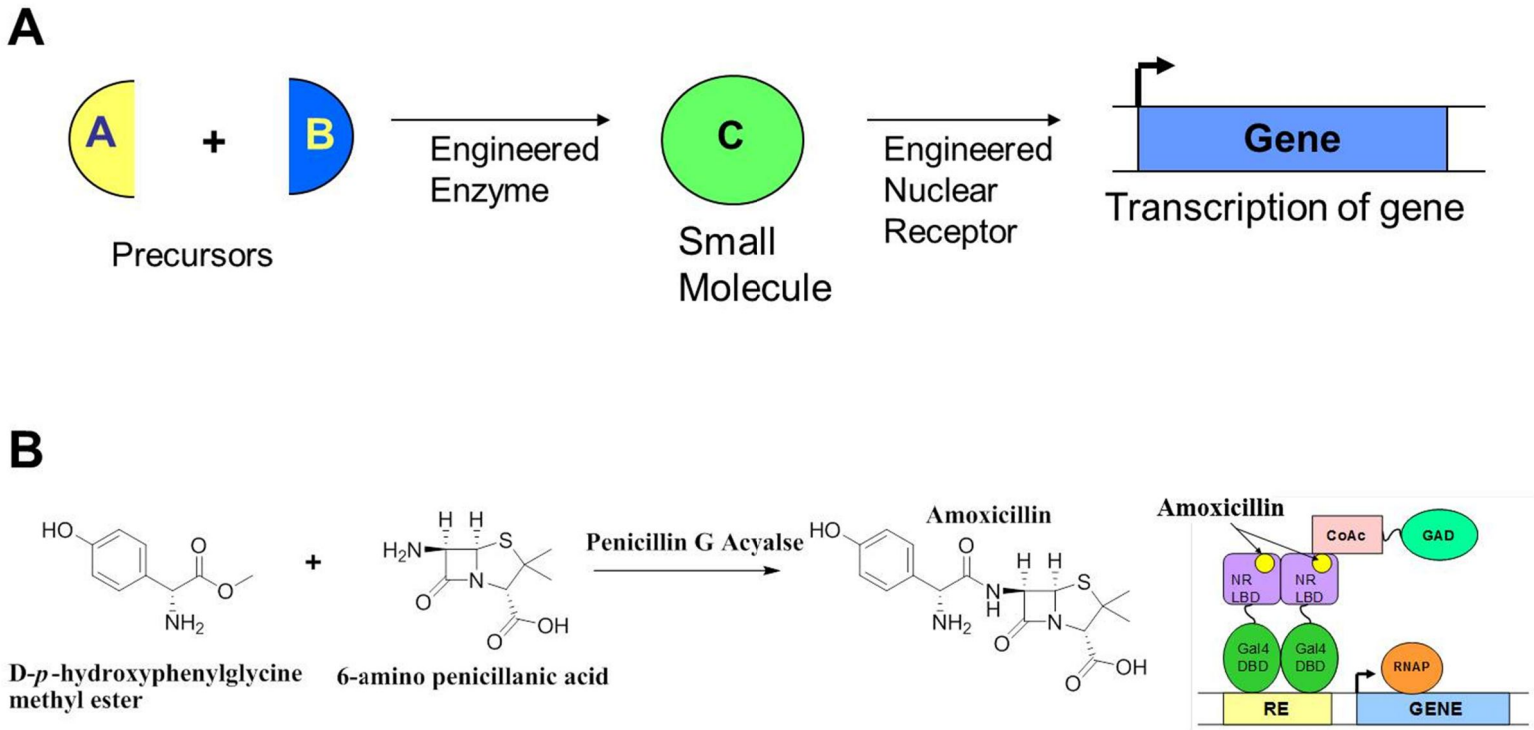


Figure 1.8. Enzyme Activated Growth. (A) General scheme of enzyme activated growth using an engineered nuclear receptor in order to engineer an enzyme to make a small molecule antibiotic from the precursors. (B) Enzyme activated growth in yeast through an antibiotic activating a nuclear receptor. The precursor molecules are added to the media with yeast cells containing the desired enzyme along with the nuclear receptor variant. The enzyme converts the precursors to the antibiotic to bind the nuclear receptor, lead to the initiation of transcription, and allow the yeast to survive.

survive (Figure 1.8). The first target nuclear receptor to engineer to be activated by an antibiotic was PXR, which will be discussed further in Chapter 2.

1.6 Summary

In summary, nuclear receptors control many biological functions through gene expression, and as a result, are involved in many different diseases. The mechanism of nuclear receptor activation through a small molecule ligand and coregulator proteins is well studied, but more research is being conducted to develop additional drugs for therapeutic purposes [46]. The focus of this research involves engineering a nuclear receptor to bind and activate transcription in response to a small molecule ligand, specifically the β-lactam antibiotics penicillin G, oxacillin, cloxacillin, ampicillin, amoxicillin, and nafcillin.

1.7 References

1. Gronemeyer, H., J.A. Gustafsson, and V. Laudet, *Principles for modulation of the nuclear receptor superfamily.* Nature Reviews Drug Discovery, 2004. **3**(11): p. 950-964.

2. Sonoda, J., L.M. Pei, and R.M. Evans, *Nuclear receptors: Decoding metabolic disease.* Febs Letters, 2008. **582**(1): p. 2-9.

3. Huang, P.X., V. Chandra, and F. Rastinejad, *Structural Overview of the Nuclear Receptor Superfamily: Insights into Physiology and Therapeutics.* Annual Review of Physiology, 2010. **72**: p. 247-272.

4. Germain, P., et al., *Overview of nomenclature of nuclear receptors.* Pharmacological Reviews, 2006. **58**(4): p. 685-704.

5. Greschik, I. and D. Moras, *Structure-activity relationship of nuclear receptor-ligand interactions.* Current Topics in Medicinal Chemistry, 2003. **3**(14): p. 1573-1599.

6. Noy, N., *Ligand specificity of nuclear hormone receptors: Sifting through promiscuity.* Biochemistry, 2007. **46**(47): p. 13461-13467.

7. Moore, J.T., J.L. Collins, and K.H. Pearce, *The nuclear receptor superfamily and drug discovery.* ChemMedChem, 2006. **1**(5): p. 504.

8. Steinmetz, A.C.U., J.P. Renaud, and D. Moras, *Binding of ligands and activation of transcription by nuclear receptors.* Annual Review of Biophysics and Biomolecular Structure, 2001. **30**: p. 329-359.

9. Hollenberg, S.M., et al., *Primary Structure and Expression of a Functional Human Glucocorticoid Receptor cDNA.* Nature, 1985. **318**(6047): p. 635-641.

10. Bai, Z.L. and R. Gust, *Breast Cancer, Estrogen Receptor and Ligands.* Archiv Der Pharmazie, 2009. **342**(3): p. 133-149.

11. Chen, T.S., et al., *Coactivators in assay design for nuclear hormone receptor drug discovery.* Assay and Drug Development Technologies, 2003. **1**(6): p. 835-842.

12. Laudet, V. and H. Gronemeyer, *The Nuclear Receptor FactsBook.* London: Academic Press, 2002.

13. Chen, Q., et al., *A yeast two-hybrid technology-based system for the discovery of PPAR gamma agonist and antagonist.* Analytical Biochemistry, 2004. **335**(2): p. 253-259.

14. Smith, C.L. and B.W. O'Malley, *Coregulator function: A key to understanding tissue specificity of selective receptor modulators.* Endocrine Reviews, 2004. **25**(1): p. 45-71.

15. Zhang, J.H., F.T.F. Tsai, and D.S. Geller, *Differential interaction of RU486 with the progesterone and glucocorticoid receptors.* Journal of Molecular Endocrinology, 2006. **37**(1): p. 163-173.

16. de Lera, A.R., et al., *Design of selective nuclear receptor modulators: RAR and RXR as a case study.* Nature Reviews Drug Discovery, 2007. **6**(10): p. 811-820.

17. Folkertsma, S., et al., *A family-based approach reveals the function of residues in the nuclear receptor ligand-binding domain.* Journal of Molecular Biology, 2004. **341**(2): p. 321-335.

18. Warnmark, A., et al., *Activation functions 1 and 2 of nuclear receptors: Molecular strategies for transcriptional activation.* Molecular Endocrinology, 2003. **17**(10): p. 1901-1909.

19. Chandra, V., et al., *Structure of the intact PPAR-gamma-RXR-alpha nuclear receptor complex on DNA.* Nature, 2008. **456**(7220): p. 350-356.

20. Bain, D.L., et al., *Nuclear receptor structure: Implications for function.* Annual Review of Physiology, 2007. **69**: p. 201-220.

21. Farboud, B. and M.L. Privalsky, *Retinoic acid receptor-alpha is stabilized in a repressive state by its C-terminal, isotype-specific F domain.* Molecular Endocrinology, 2004. **18**(12): p. 2839-2853.

22. Yang, J., et al., *The F-domain of estrogen receptor-alpha inhibits ligand induced receptor dimerization.* Molecular and Cellular Endocrinology, 2008. **295**(1-2): p. 94-100.

23. Rastinejad, F., *Retinoid X receptor and its partners in the nuclear receptor family.* Current Opinion in Structural Biology, 2001. **11**(1): p. 33-38.

24. Chambon, P., *The nuclear receptor superfamily: A personal retrospect on the first two decades.* Molecular Endocrinology, 2005. **19**(6): p. 1418-1428.

25. Jakob, M., et al., *Novel DNA-binding element within the C-terminal extension of the nuclear receptor DNA-binding domain.* Nucleic Acids Research, 2007. **35**(8): p. 2705-2718.

26. Khorasanizadeh, S. and F. Rastinejad, *Nuclear-receptor interactions on DNA-response elements.* Trends in Biochemical Sciences, 2001. **26**(6): p. 384-390.

27. Lehmann, J.M., et al., *The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions.* Journal of Clinical Investigation, 1998. **102**(5): p. 1016-1023.

28. Bourguet, W., et al., *Crystal Structure of the Ligand Binding Domain of the Human Nuclear Receptor RXR alpha.* Nature, 1995. **375**(6530): p. 377-382.

29. Pogenberg, V., et al., *Characterization of the interaction between retinoic acid receptor/retinoid X receptor (RAR/RXR) heterodimers and transcriptional coactivators through structural and fluorescence anisotropy studies.* Journal of Biological Chemistry, 2005. **280**(2): p. 1625-1633.

30. Goyanka, R., et al., *Nuclear Receptor Engineering Based on Novel Structure Activity Relationships Revealed by Farnesyl Pyrophosphate.* Protein Engineering Design & Selection, 2010.

31. Brzozowski, A.M., et al., *Molecular basis of agonism and antagonism in the oestrogen receptor.* Nature, 1997. **389**(6652): p. 753-758.

32. Stanisic, V., D. Lonard, and B. O'Malley, *Modulation of Steroid Hormone Receptor Activity.* Progress in Brain Research, 2010. **181**: p. 153-176.

33. Rosenfeld, M.G., V.V. Lunyak, and C.K. Glass, *Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptionaf response.* Genes & Development, 2006. **20**(11): p. 1405-1428.

34. Privalsky, M.L., *The role of corepressors in transcriptional regulation by nuclear hormone receptors.* Annual Review of Physiology, 2004. **66**: p. 315-360.

35. McInerney, E.M., et al., *Determinants of coactivator LXXLL motif specificity in nuclear receptor transcriptional activation.* Genes & Development, 1998. **12**(21): p. 3357-3368.

36. Watkins, R.E., et al., *2.1 angstrom crystal structure of human PXR in complex with the St. John's wort compound hyperforin.* Biochemistry, 2003. **42**(6): p. 1430-1438.

37. Heldring, N., et al., *Structural insights into corepressor recognition by antagonist-bound estrogen receptors.* Journal of Biological Chemistry, 2007. **282**(14): p. 10449-10455.

38. Mangelsdorf, D.J., et al., *The nuclear receptor superfamily-the 2nd decade.* Cell, 1995. **83**(6): p. 835-839.

39. Guenther, M.G., O. Barak, and M.A. Lazar, *The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3.* Molecular and Cellular Biology, 2001. **21**(18): p. 6091-6101.

40. Nagy, L., et al., *Mechanism of corepressor binding and release from nuclear hormone receptors.* Genes & Development, 1999. **13**(24): p. 3209-3216.

41. Leo, C. and J.D. Chen, *The SRC family of nuclear receptor coactivators.* Gene, 2000. **245**(1): p. 1-11.

42. Leo, C., et al., *Role of retinoid receptor coactivator pockets in cofactor recruitment and transcriptional regulation.* Journal of Biological Chemistry, 2001. **276**(25): p. 23127-23134.

43. Love, J.D., et al., *Transcriptional repression by nuclear receptors: mechanisms and role in disease.* Biochemical Society Transactions, 2000. **28**: p. 390-396.

44. McKenna, N.J. and B.W. O'Malley, *Combinatorial control of gene expression by nuclear receptors and coregulators.* Cell, 2002. **108**(4): p. 465-474.

45. Shiau, A.K., et al., *The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen.* Cell, 1998. **95**(7): p. 927-937.

46. O'Malley, B., *The Year in Basic Science: Nuclear Receptors and Coregulators.* Molecular Endocrinology, 2008. **22**(12): p. 2751-2758.

47. Fields, S. and O.K. Song, *A novel genetic system to detect protein-protein interactions.* Nature, 1989. **340**(6230): p. 245-246.

48. McDonnell, D.P., E. Vegeto, and M.A.G. Gleeson, *Nuclear hormone receptors as targets for new drug discovery.* Bio-Technology, 1993. **11**(11): p. 1256-1261.

49. Woycechowsky, K.J. and D. Hilvert, *Deciphering enzymes - Genetic selection as a probe of structure and mechanism.* European Journal of Biochemistry, 2004. **271**(9): p. 1630-1637.

50. Lin, H.N., H.Y. Tao, and V.W. Cornish, *Directed evolution of a glycosynthase via chemical complementation.* Journal of the American Chemical Society, 2004. **126**(46): p. 15051-15059.

51. Taylor, S.V., P. Kast, and D. Hilvert, *Investigating and engineering enzymes by genetic selection.* Angewandte Chemie International Edition, 2001. **40**(18): p. 3310-3335.

52. Feilotter, H., et al., *Construction of an improved host strain for two hybrid screening.* Nucleic Acids Research, 1994. **22**(8): p. 1502-1503.

53. Cribb, P. and E. Serra, *One- and two-hybrid analysis of the interactions between components of the Trypanosoma cruzi spliced leader RNA gene promoter binding complex.* International Journal for Parasitology, 2009. **39**(5): p. 525-532.

54. Rajmohan, R., et al., *Characterization of Wiskott-Aldrich syndrome (WAS) mutants using Saccharomyces cerevisiae.* FEMS Yeast Research, 2009. **9**(8): p. 1226-1235.

55. Drees, B.L., *Progress and variations in two-hybrid and three-hybrid technologies.* Current Opinion in Chemical Biology, 1999. **3**(1): p. 64-70.

56. Uetz, P., et al., *A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.* Nature, 2000. **403**(6770): p. 623-627.

57. Licitra, E.J. and J.O. Liu, *A three-hybrid system for detecting small ligand-protein receptor interactions.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(23): p. 12817-12821.

58. Azizi, B., E.I. Chang, and D.F. Doyle, *Chemical complementation: small-molecule-based genetic selection in yeast.* Biochemical and Biophysical Research Communications, 2003. **306**(3): p. 774-780.

59. James, P., J. Halladay, and E.A. Craig, *Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast.* Genetics, 1996. **144**(4): p. 1425-1436.

60. Lohr, D., P. Venkov, and J. Zlatanova, *Transcriptional regulation in the yeast GAL gene family: A complex genetic network.* Faseb Journal, 1995. **9**(9): p. 777-787.

61. Zeng, J., et al., *Genome wide screens in yeast to identify potential binding sites and target genes of DNA-binding proteins.* Nucleic Acids Research, 2008. **36**(1): p. 8.

62. Pollock, R. and T. Clackson, *Dimerizer-regulated gene expression.* Current Opinion in Biotechnology, 2002. **13**(5): p. 459-467.

63. Gillies, A.R., G. Skretas, and D.W. Wood, *Engineered systems for detection and discovery of nuclear hormone-like compounds.* Biotechnology Progress, 2008. **24**: p. 8-16.

64. Schwimmer, L.J., et al., *Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(41): p. 14707-14712.

65. Chockalingam, K., et al., *Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(16): p. 5691-5696.

66. Chen, Z.L., et al., *Directed evolution of human estrogen receptor variants with significantly enhanced androgen specificity and affinity.* Journal of Biological Chemistry, 2004. **279**(32): p. 33855-33864.

67. Islam, K.M.D., et al., *Directed evolution of estrogen receptor proteins with altered ligand-binding specificities.* Protein Engineering Design & Selection, 2009. **22**(1): p. 45-52.

68. Elander, R.P., *Industrial production of beta-lactam antibiotics.* Applied Microbiology and Biotechnology, 2003. **61**(5-6): p. 385-392.

69. Demain, A.L. and R.P. Elander, *The beta-lactam antibiotics: past, present, and future.* Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology, 1999. **75**(1-2): p. 5-19.

70. Wegman, M.A., et al., *Towards biocatalytic synthesis of beta-lactam antibiotics.* Advanced Synthesis & Catalysis, 2001. **343**(6-7): p. 559-576.

CHAPTER 2

ENGINEERING NUCLEAR RECEPTORS TOWARD ANTIBIOTICS: PREGNANE X RECEPTOR

2.1 Engineering Nuclear Receptors to Bind Antibiotics

Nuclear receptors play a vital role in gene regulation, and consequently are involved in many diseases. As a result, there is an emphasis on engineering nuclear receptors for applications in gene therapy, drug discovery and metabolic engineering. As mentioned in Chapter 1, the goal of this work is to engineer a nuclear receptor toward the β-lactam antibiotics via chemical complementation. The β-lactam antibiotics are widely used in the pharmaceutical industry, however the cost and production of waste is high [1]. Therefore, new synthetic or biosynthetic pathways are desirable to overcome the production and cost concerns. One mechanism to do this is to engineer an enzyme to biosynthetically mass produce the antibiotic, a method that is currently unavailable. Once this enzyme is able to create the desired product, a selection mechanism in the microorganism is needed to produce continual biosynthesis of the product. This work, therefore, focuses on engineering a nuclear receptor variant that is able to bind and activate in response to the β-lactam antibiotics and then use that receptor in chemical complementation to engineer an enzyme to make the antibiotics. The pregnane X receptor (PXR) was chosen to engineer to bind and activate in response to β-lactam antibiotics.

2.2 Promiscuous Activity of PXR

PXR, a member of the nuclear receptor superfamily, was originally defined as an orphan receptor due to the lack of a known "wild-type" ligand [2]. Prior to the discovery of PXR, certain compounds, such as the synthetic glucocorticoid dexamethasone, the antiglucocorticoid pregnenolone 16α-carbonitrile (PCN), the macrolide antibiotic rifampicin, and the herbal antidepressant hyperforin were known to turn on expression of a cytochrome P450 (CYP450) gene, specifically the cytochrome P-450 monooxygenase 3A4 (*CYP3A4*) gene [2-5]. The family of CYP450 enzymes regulate endobiotic and xenobiotic metabolism, however the mechanism leading to the activation of *CYP3A4* was unknown [3, 6, 7]. Initially, the glucocorticoid receptor (GR) was thought to be involved in this mechanism because the ligands leading to the activation of the *CYP3A4* gene were glucocorticoids [2, 3, 6]. Once isolated, however, PXR was found to be the protein responsible for the expression of the *CYP3A4* gene [2, 3, 6, 8, 9]. Additionally, PXR was also found to regulate expression of the *CYP2C9* and *CYP2B6* genes, both CYP450s involved in drug metabolism [7, 10-12]. PXR was also observed to be activated by naturally occurring pregnane steroids and synthetic steroids such as dexamethasone, as well as involved in regulating bile acid homeostasis through the detoxification of cholesterol and oxysterols [2, 5, 6, 8].

Evolutionarily, PXR is related to the constitutive androstane receptor (CAR) and the vitamin D receptor (VDR) due to the similar sequence identities among the DBDs and LBDs of these receptors [6, 13-15]. CAR, like PXR, is involved in xenobiotic metabolism, whereas VDR plays a role in bile acid homeostasis [6, 7, 16]. Functionally, PXR forms a heterodimer with the retinoid X receptor (RXR) and, through regulation of *CYP3A4* gene expression, is responsible for the metabolism of more than 50 % of human

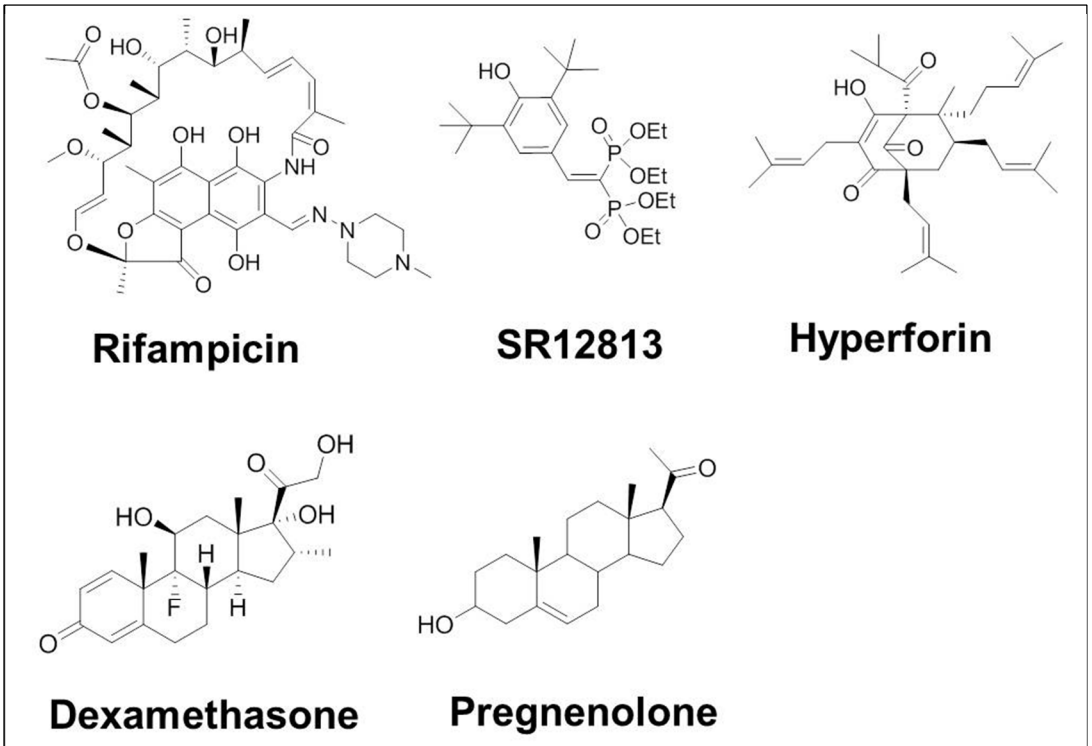


Figure 2.1. Structures of some ligands that activate PXR.

drugs [2]. PXR ligands vary in size and shape, ranging from the large antibiotic rifampicin (MW: 823 g/mol) to the cholesterol lowering drug SR12813 (MW: 505 g/mol), and hyperforin (MW: 537 g/mol). (Figure 2.1) [2, 17-21]. These ligand binding capabilities make PXR a promiscuous nuclear receptor.

In addition to being activated by a broad range of compounds, another unique characteristic of PXR is the differences in ligand specificities among different species [5, 22-25]. For example, mouse and rat PXR are strongly activated by PCN, but human and rabbit are not, whereas rifampicin activates human PXR well and has little to no activity with mouse and rat PXR [4]. On average, the mammalian PXR LBDs only share approximately 75-80 % sequence identity and among the non-mammalian orthologs only share about 50 % sequence identity [4, 13, 21]. Many nuclear receptors share at least 60-70 % sequence identities across all species [21].

Structurally, PXR resembles most nuclear receptors, consisting of 12 α-helices, but contains two additional β-strands, making a total of five β-strands whereas most nuclear receptors contain two or three. In addition, the 1,200 $Å^3$ binding pocket is significantly larger than most nuclear receptors, which have pocket sizes ranging from approximately 400 $Å^3$ to 700 $Å^3$ [26-28]. Evidence for the ability of PXR to accommodate large ligands is seen in the crystal structure of PXR with rifampicin, where partial unwinding of helix 7 expands the volume of PXR's binding pocket to almost 1,600 $Å^3$ (Figure 2.2) [6, 18, 29]. In addition, the *apo* crystal structure of PXR LBD shows the receptor in the active conformation (PDB 1ILG) (Figure 2.3). Helix 12 is bent towards the rest of the protein forming the α-helical sandwich normally seen in the presence of a ligand [19]. The promiscuous activity of PXR is displayed through the

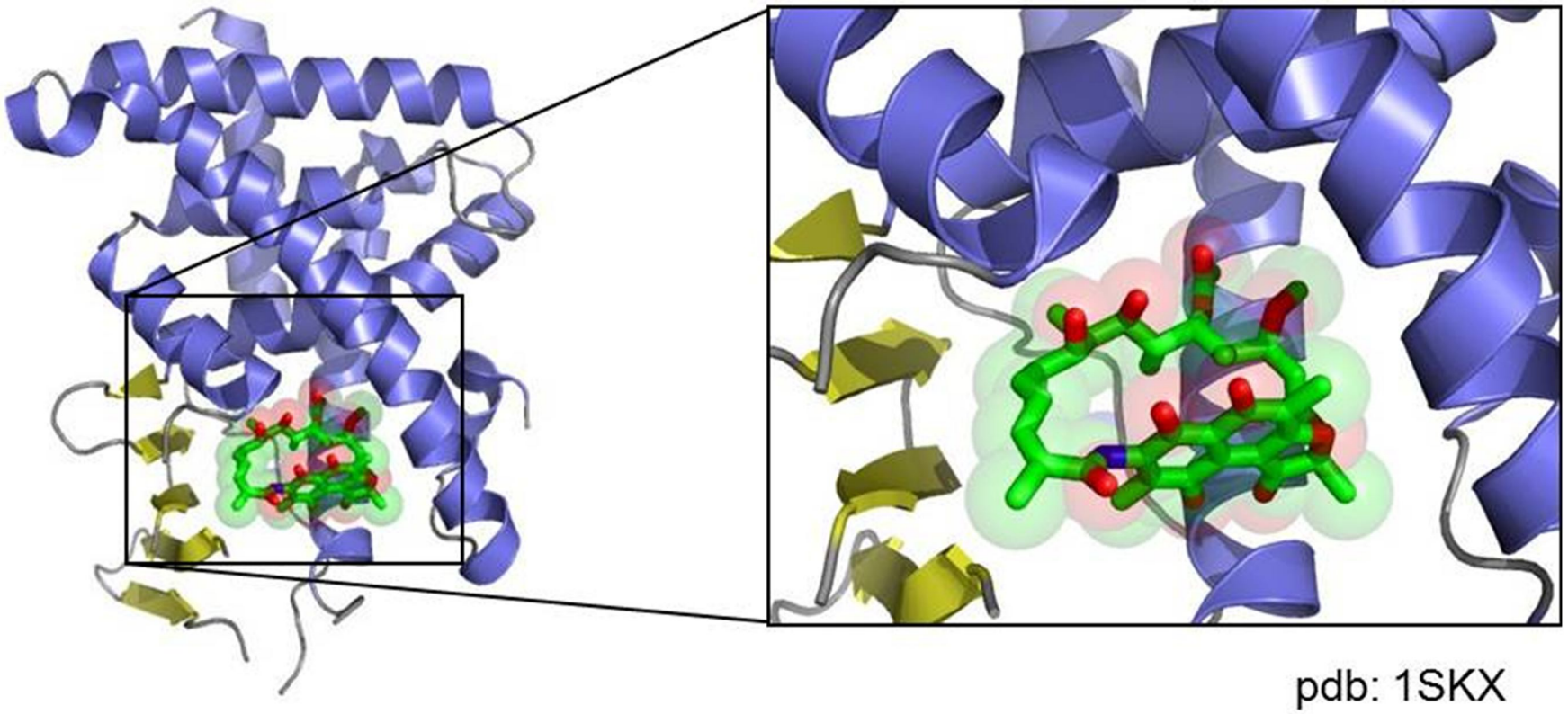


Figure 2.2. Crystal structure of the LBD of PXR with rifampicin (PDB 1SKX). All α-helices are shown in blue and β-strands are shown in yellow. The ligand is shown in green.

enhanced volume of the binding pocket, the additional structural elements, and the ability to be activated by ligands of various sizes, making this receptor an attractive candidate to engineer to bind the β-lactam antibiotics [2, 3, 19, 22, 30].

2.3 PXR in Chemical Complementation

As mentioned in Chapter 1, the development of chemical complementation, a genetic selection system in yeast for analyzing nuclear receptor interactions, was useful for many applications, such as protein and enzyme engineering. Prior to engineering PXR to activate transcription in response to a β-lactam antibiotic, this receptor was first tested in chemical complementation to determine if the receptor was functional in our yeast system. Overall, if the ligand is able to bind the nuclear receptor and transcription is activated, the yeast will survive on selective media.

To test PXR in chemical complementation, a yeast expression plasmid containing the Gal4 DNA binding domain (Gal4DBD) fused to the PXR LBD (pGBDPXR), containing a tryptophan marker, was transformed into the yeast strain PJ69-4A [31]. Alongside that, a fusion of the nuclear receptor coactivator, the activator for thyroid hormone and retinoid receptors (ACTR), fused to the Gal4 activation domain (GAD) (pGAD10BAACTR), containing a leucine marker, was transformed into PJ69-4A. As a positive control, a plasmid with the *holo* Gal4 gene (pGBT9Gal4), containing a tryptophan marker, was also transformed into PJ69-4A with pGAD10BAACTR. Gal4 is a ligand-independent transcription factor [32]. In addition, the plasmid with the Gal4DBD fused to the RXRLBD, pGBDRXR, and pGAD10BAACTR were co-transformed in PJ69-4A, also as controls.

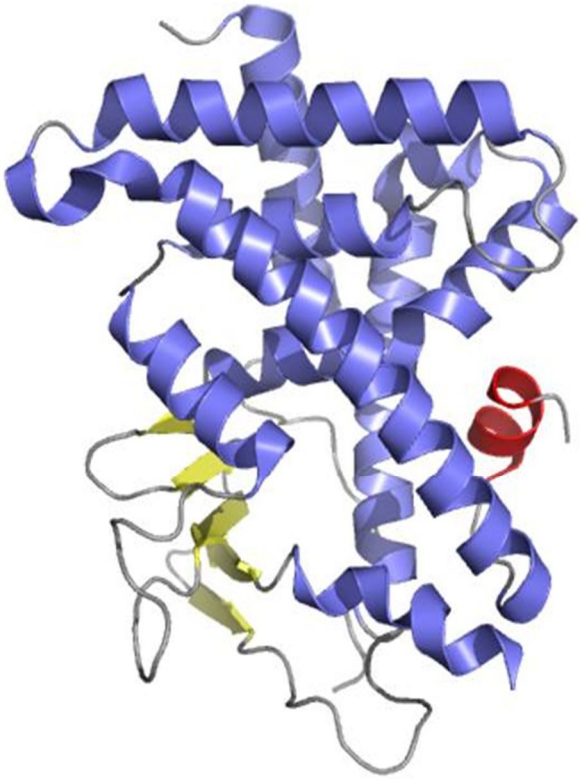


Figure 2.3. Crystal structure of the LBD of *apo* PXR (PDB 1ILG). All α-helices are shown in blue and β-strands are shown in yellow. Helix 12 is shown in red.

Yeast transformants containing both pGBDPXR and pGAD10BAACTR were then analyzed on selective agar media. The selective media used was synthetic complete media lacking an essential nutrient. As a control, the transformants were streaked onto synthetic complete media lacking leucine and tryptophan (SC-LW), selecting for both fusion plasmids. As previously mentioned, the chemical complementation selection system uses either the *ADE2* or *HIS3* selective genes, involved in the adenine and histidine biosynthetic pathways, respectively. In adenine or histidine selective media, if the small molecule ligand activates the nuclear receptor and turns on transcription of the selective gene, the yeast will survive.

PXR as well as the Gal4 and RXR controls were tested on synthetic complete media lacking adenine, leucine, and tryptophan (SC-ALW) with or without the ligand 9-*cis* retinoic acid (9cRA), the wild-type ligand for RXR (Figure 2.4). As a ligand independent transcription factor, the sectors of the adenine selective plate with Gal4 display growth with and without 9cRA as well as the nonselective SC-LW plate, as expected (Figure 2.4A-C). RXR did not grow on the adenine selective plate (SC-ALW) and shows ligand activated growth on the adenine selective plate with 9cRA (Figure 2.4). PXR showed growth on all three plates, and is therefore constitutively active, meaning the yeast grow on selective media without the presence of an exogenous small molecule (Figure 2.4A-C). Due to the observed growth on the SC-ALW plate without ligand, these streaking results confirm that wild-type PXR is constitutively active in chemical complementation (Figure 2.4B).

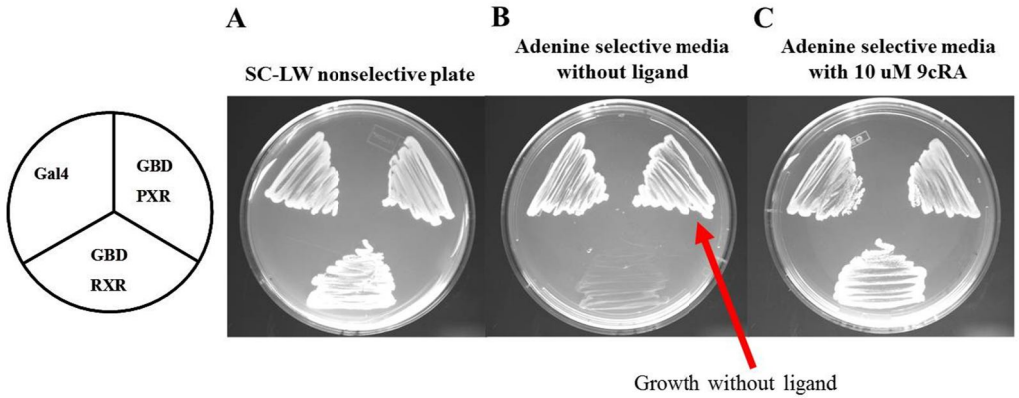


Figure 2.4. Chemical complementation assay with GBDPXR, GBDRXR, and Gal4 with ACTR on selective agar plates. (A) Synthetic complete media lacking leucine and tryptophan (SC-LW) show growth on all sectors of the plate, as expected, selecting for the presence of the plasmids containing the fusion proteins (B) Adenine selective media (SC-ALW) without ligand. RXR does not show growth and Gal4 grows, as expected. PXR is constitutively active, displaying growth without the presence of ligand as indicated by the arrow. (C) Growth is observed on all sectors of the plate in the adenine selective media containing 9-*cis* retinoic acid (9cRA).

2.4 PXR Library for Antibiotics

2.4.1 Library Design

Despite the constitutive activity of PXR in chemical complementation, engineering a PXR variant toward the antibiotics became a two-fold issue. The first goal was to create a variant that was not constitutively active, and the second goal was to rationally design a library to find a variant able to bind and activate in response to any of the β-lactam antibiotics: ampicillin, amoxicillin, penicillin G, oxacillin, cloxacillin, and nafcillin. As mentioned in Chapter 1, the nuclear receptor variant can be used in chemical complementation to engineer an enzyme to make the antibiotics.

To design a library of PXR variants able to be activated by the antibiotics, two main steps were taken. First, the residues in the receptor's ligand binding pocket were chosen for mutagenesis. These residues are selected based on structural analysis as well as on known protein-ligand interactions. Once a list of residues was generated, mutations at each of those positions were determined based on sequence alignments focusing on evolutionarily conserved residues and the amino acid properties within the wild-type protein.

Amino acid residues from the crystal structures of three different PXRLBD-ligand structures, specifically with rifampicin, SR12813, and hyperforin (PDB files 1SKX, 1ILH, and 1M13, respectively) were analyzed [17-19]. In each structure, residues within 4.0 Å of the ligand were considered for mutagenesis (Figure 2.5). Four angstroms is an optimal distance between the ligand and receptor to generate a list of amino acids that contact the ligand. In Figure 2.5, a Venn diagram of the residues common to all three structures and within 4 Å of the ligand is shown, useful for determining differences

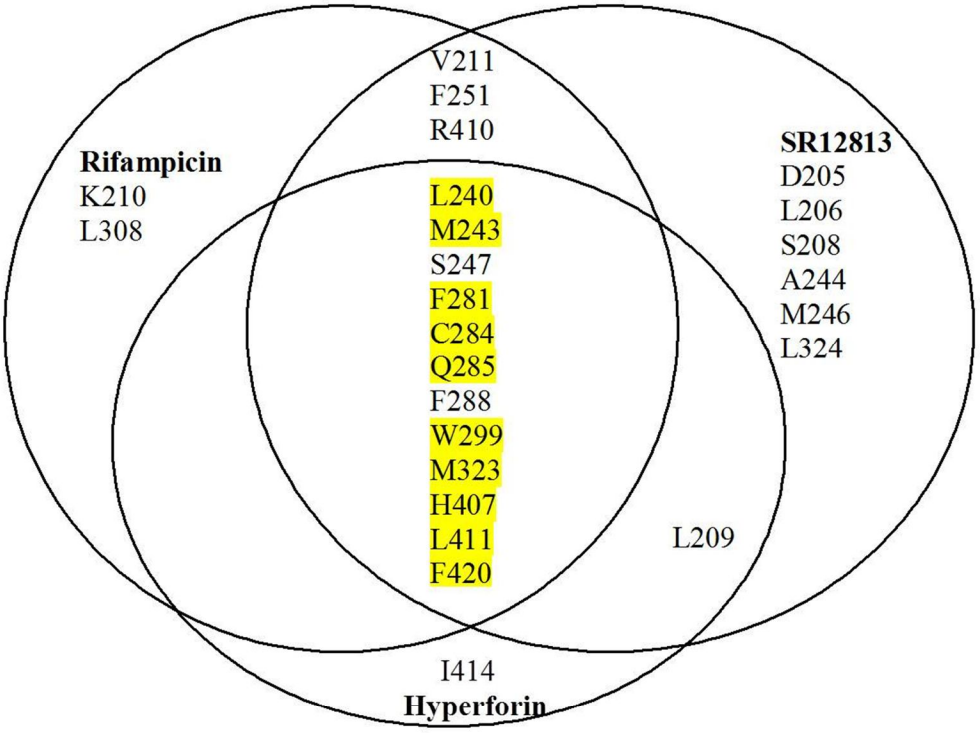


Figure 2.5. Venn diagram of common residues within 4 Å of the ligand in three different crystal structures of the PXRLBD with the ligands rifampicin, SR12813, and hyperforin (PDB files 1SKX, 1M13, and 1ILH, respectively). All residues highlighted in yellow were chosen to mutate in the PXR library. Phenylalanine 288 and serine 247 were not chosen to minimize the library size.

among the residues that contact the ligand in the various structures. For example, L209 is common between the structures with SR12813 and hyperforin, but not the structure with rifampicin, whereas V211, F251 and R410 are common between the structures with SR12813 and rifampicin, but not the structure with hyperforin. All of the common residues among the three structures were considered for mutagenesis in the library. Phenylalanine 288 and serine 247 were not chosen to minimize the library size, resulting in L240, M243, F281, C284, Q285, W299, M323, H407, L411, and F420 as the residues for the library.

The types of mutations made at each position were determined based on sequence alignments of PXR, VDR, and CAR since these three receptors belong to the same subgroup as PXR within the nuclear receptor superfamily (Figure 2.6) [13, 16]. Five of the ten positions, L240, M243, F281, M323, and L411, were chosen to be mutated to the hydrophobic amino acids leucine, isoleucine, valine, phenylalanine, and methionine to maintain the hydrophobic residues in the ligand binding pocket (LBP). Previously, Chrenik *et al.* showed that a mutation of Q285 to isoleucine did not significantly affect activity with rifampicin, despite the presence of two hydrogen bonds between Q285 and rifampicin [18]. Q285, therefore, was also mutated to a leucine, isoleucine, valine, phenylalanine, and methionine to be consistent with maintaining the hydrophobic nature of the pocket. Tryptophan 299 was mutated to a phenylalanine based on the presence of a phenylalanine at that position in CAR. Due to the nature of the genetic code, a cysteine was also introduced into the design at position 299, producing a complete list of changes consisting of tryptophan, phenylalanine, leucine, and cysteine at position 299. At residue C284, PXR orthologs and VDR contain a conserved isoleucine, therefore an isoleucine

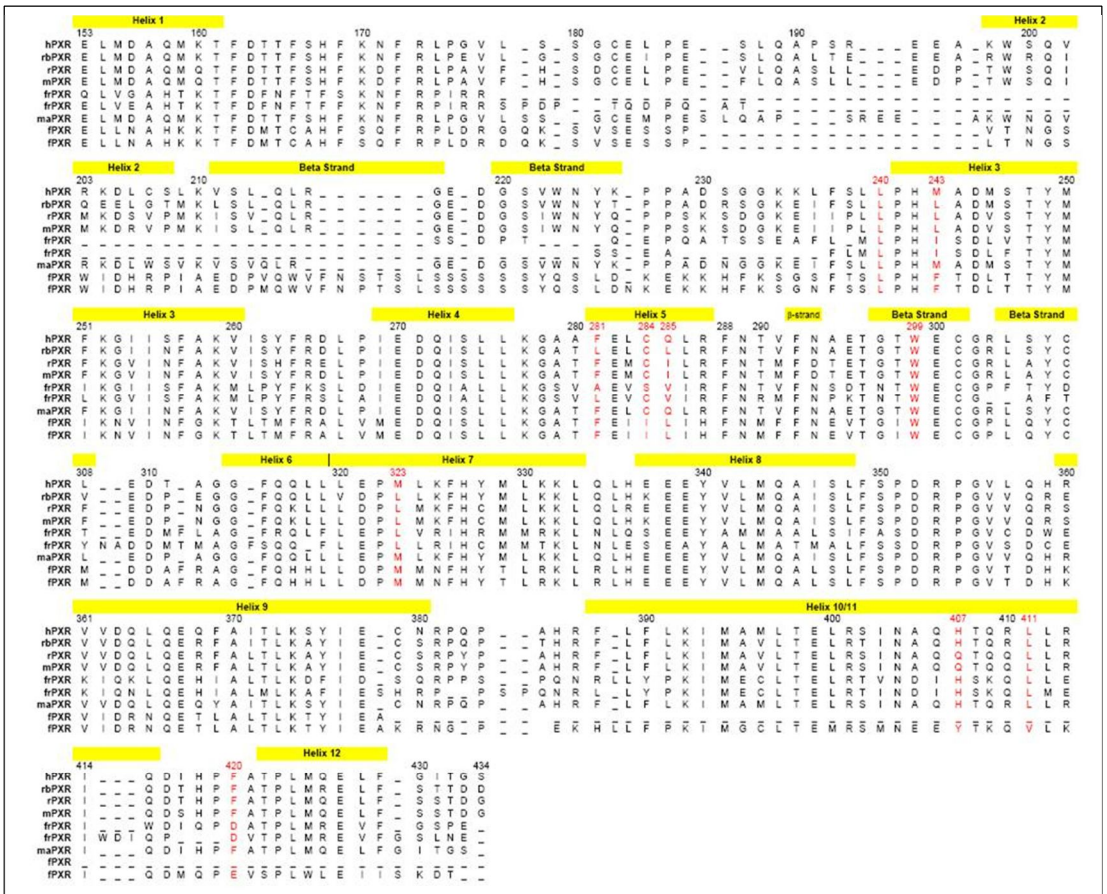
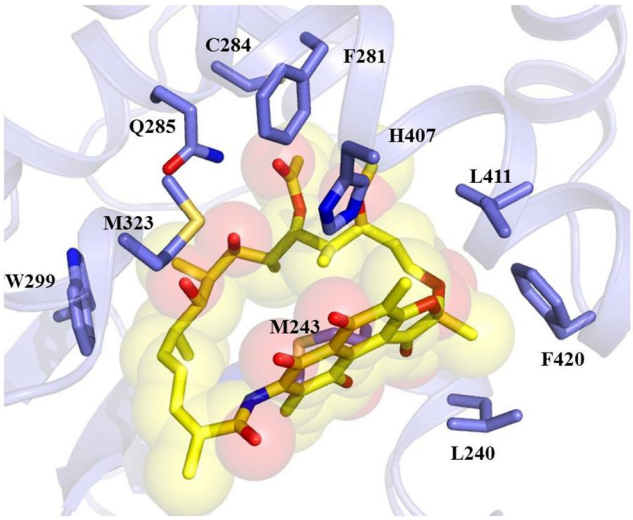


Figure 2.6. Alignment of PXR orthologs. Residues for the library are shown in red. Secondary structural elements are shown in yellow above the corresponding residues. The Residue numbers are shown every ten residues. The species abbreviations are h: human, rb: rabbit, r: rat, m: mouse, fr:frog, ma: Macaca mulatta (Rhesus macaque), and f: fish.

A



B

Residue	Chosen Mutations
L240	LIV_2FM
M243	LIV_2FM
F281	LIV_2FM
C284	CISF
Q285	LIV_2FM
W299	F_2WLC_2
M323	LIV_2FM
H407	HQ
L411	LIV_2FM
F420	FDYV

Figure 2.7. PXR library design. (A) Ligand binding pocket of LBD with rifampicin showing residues selected for mutagenesis. (B) Residues chosen to mutate in the PXR library and mutations at each position. The subscripts indicate amino acids that appear with more than one codon at the randomized amino acid position and are represented twice in the design.

was incorporated at that position, as well as, serine due to the similarity in size and polarity to cysteine. Therefore, the genetic code introduce a phenylalanine, resulting in four possible substitutions: cysteine, isoleucine, serine, and phenylalanine. As seen in Figure 2.6, position 407 is a glutamine in the PXR orthologs therefore H407 was chosen to be mutated to a glutamine (Figure 2.6). In order to maintain the wild-type polarity at H407, that position was not mutated to any hydrophobic amino acids. Also based on the sequence alignments, position 420 was substituted with phenylalanine, tyrosine, valine, and aspartic acid. The theoretical library size at the amino acid level for these designed mutations is 2.0×10^6 variants (Figure 2.7). The list of oligonucleotides containing the designed changes for the PXR library is shown in Table 2.1.

2.4.2 PXR Library Construction and Selection in Chemical Complementation

To create the library of PXR variants, oligonucleotides containing the randomized mutations were combined using overlap extension PCR to produce a complete cassette with the designed mutations (Figure 2.8) [33]. Each oligonucleotide contains overlap regions at the 5' and 3' ends that are complementary to the previous and following oligonucleotides (Figure 2.8). The fragments hybridize at the complementary ends to generate larger cassettes containing the randomized mutations and are amplified via PCR. These larger cassettes are also combined through PCR to make the complete insert cassette.

The ends of the complete insert cassette are complementary to regions in the background plasmid. The background plasmid was created to eliminate any possible contamination from the wild-type receptor when the library is transformed into yeast. To make the background plasmid, *SacII* and *KpnI* restriction sites were inserted into the

Table 2.1. PXR library oligos showing sequences in black and overlap regions are colored and underlined. Matching overlap regions are the same color. Melting temperatures based on the Vector NTI DNA program are indicated below the respective overlap region. Randomized positions are highlighted in yellow.

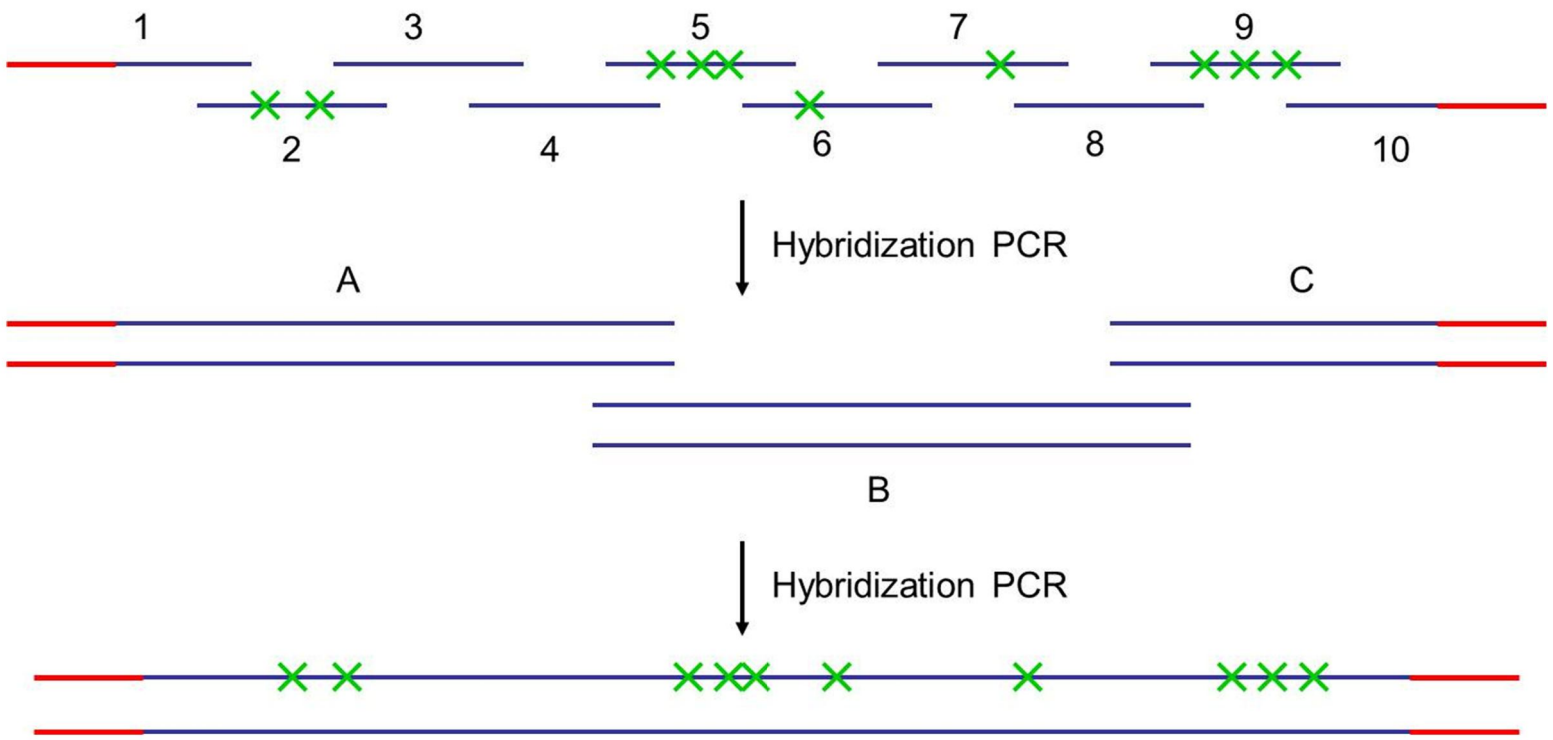


Figure 2.8. Scheme of PCR reaction steps used to generate library of variants. Oligonucleotides (oligos) containing the randomized codons are combined into larger cassettes through overlap extension PCR. These larger cassettes are combined through PCR to make the full cassette containing the designed library mutations. The oligos are shown in blue and mutations are shown as green. Red areas are regions for homologous recombination.

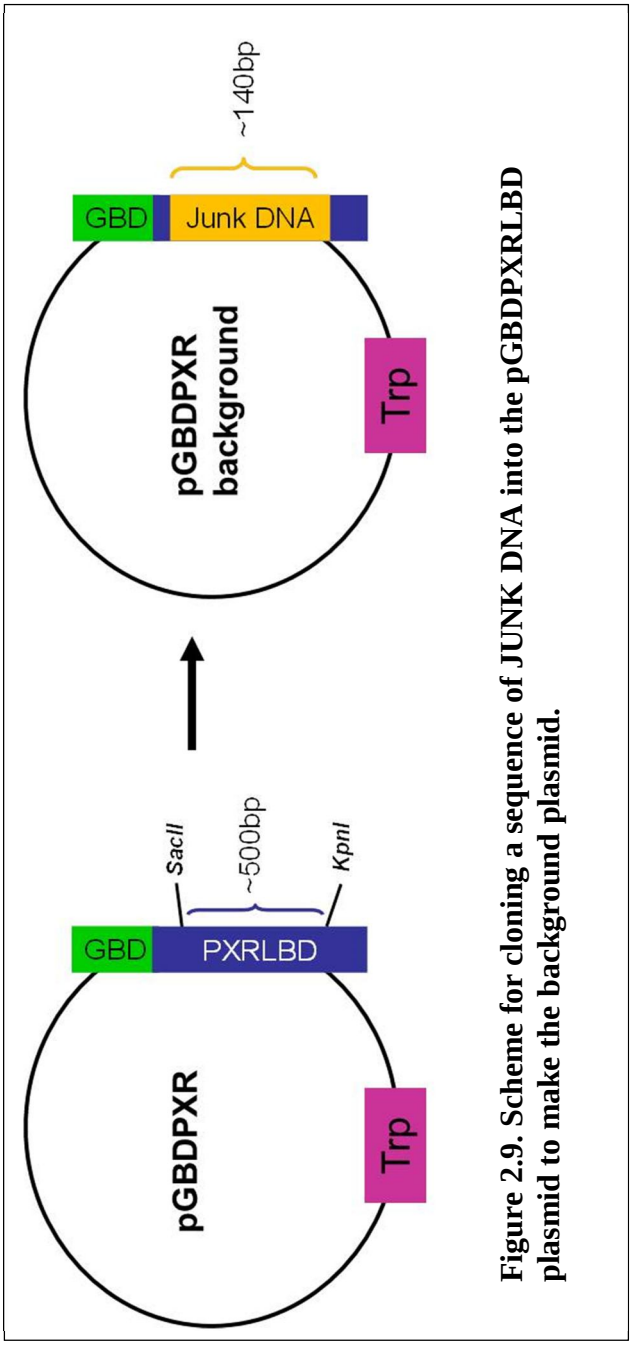


Figure 2.9. Scheme for cloning a sequence of JUNK DNA into the pGBDPXRLBD plasmid to make the background plasmid.

wild-type PXRLBD gene in the pGBDPXR plasmid. The section of the gene between these two sites was then digested and replaced with a random DNA sequence (Figure 2.9). When transcription of this gene occurs, multiple STOP codons are produced, leading to a nonfunctional protein, thus eliminating background due to the wild-type receptor.

Prior to transforming the PXR library cassette into yeast, the yeast strain containing the GAD:ACTR fusion protein was grown to competency. The background plasmid is then digested with *SacII* and *KpnI* and the library cassette along with the digested background vector were added to the yeast cells (Figure 2.10). Through homologous recombination of the library cassette with the background plasmid, yeast cells containing a library of mutants were plated onto adenine selective media containing ligand sets with the β-lactam antibiotics combined with several additional small molecule ligands (Figure 2.10). The first ligand set contained the antibiotics amoxicillin, cloxacillin, penicillin G as well as the ligand 17-β estradiol. The second ligand set contained the antibiotics ampicillin, oxacillin, nafcillin, as well as 9cRA and resveratrol. Ligand sets were designed based on similarities in the structures of the ligands. The library of mutants was also plated onto SC-LW nonselective media to determine whether homologous recombination occurred as well as to determine the library size and diversity of mutations in the library.

The transformation produced a library size of 1.42×10^5 colonies and a transformation efficiency of 9.5×10^5 colonies/μg vector casette DNA. Approximately 100 colonies from the adenine selective plates containing the various ligands were

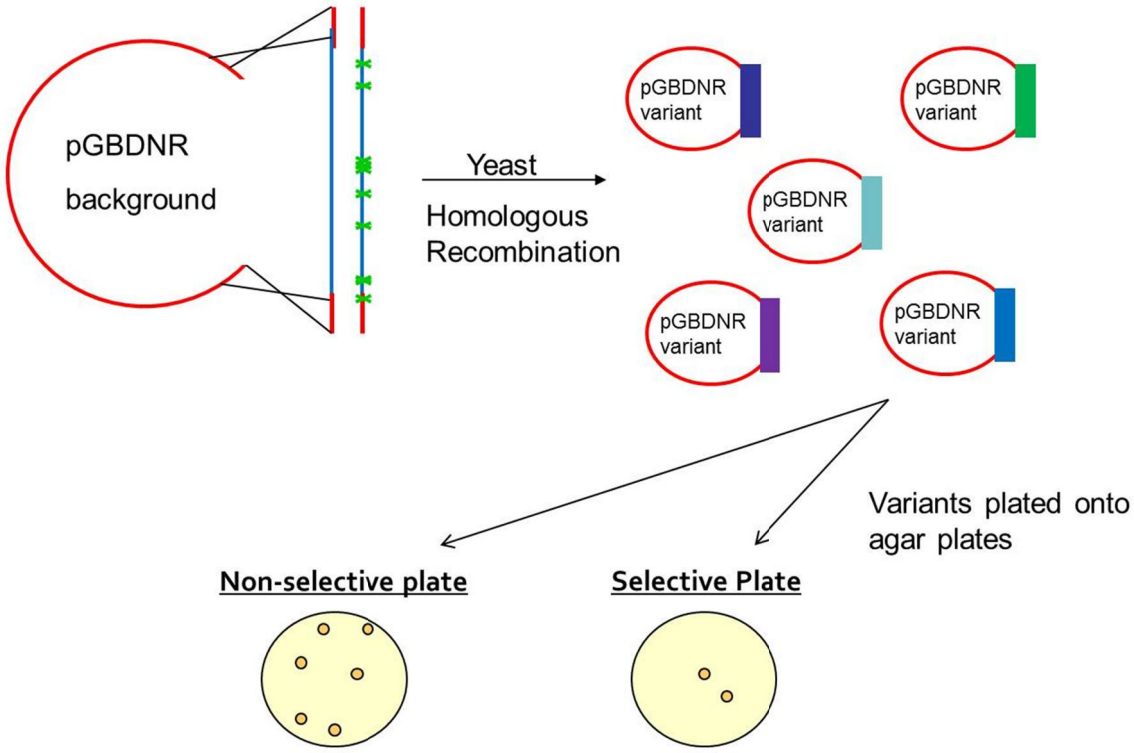


Figure 2.10. Scheme for generating libraries of nuclear receptor mutants in yeast. The library cassette is transformed with the cut background plasmid and, through homologous recombination, a library of nuclear receptor mutants is made. The library of mutants is plated onto nonselective and selective plates, selecting for ligand-activated growth.

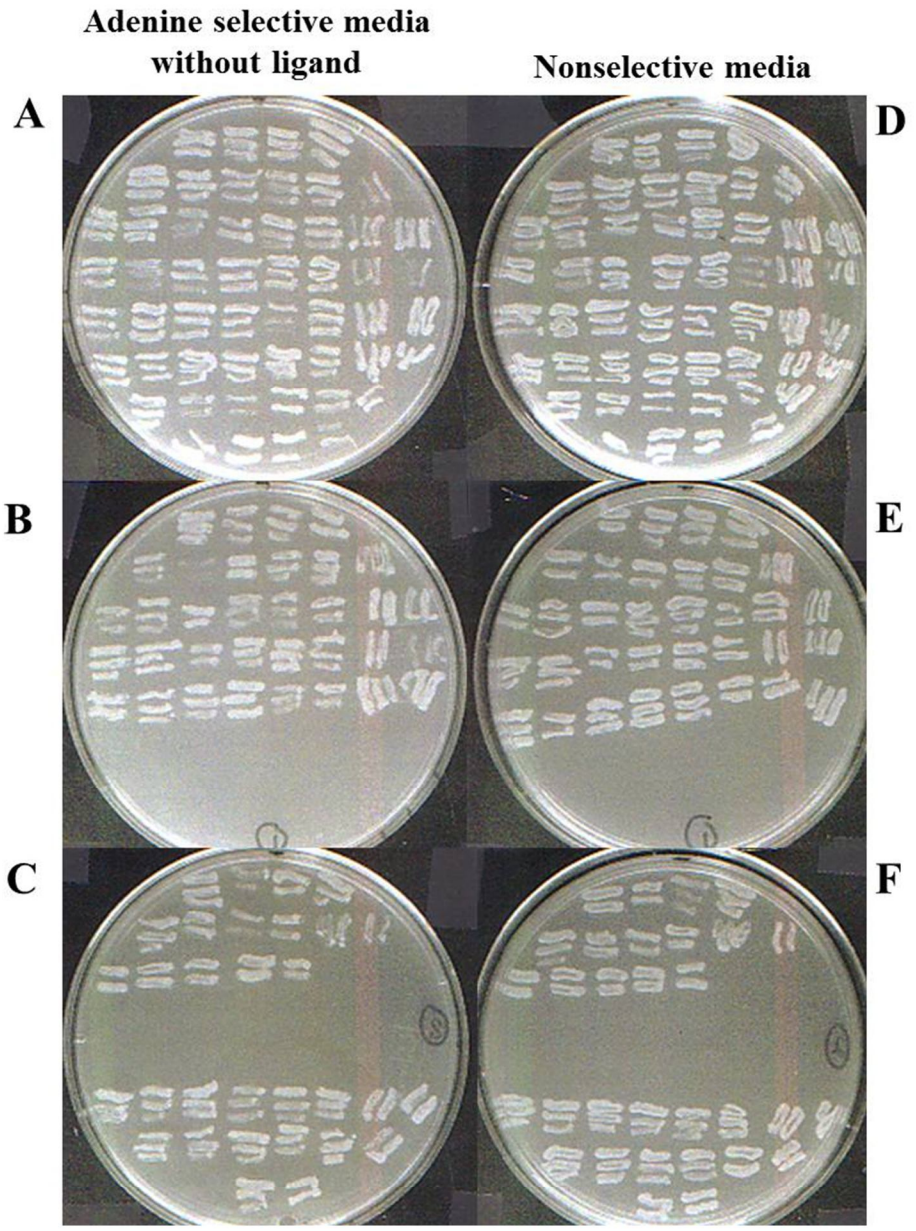


Figure 2.11. PXR Library Streaking. Mutants from the selective plates from the PXR library are streaked onto selective plates without ligand and nonselective plates to determine if the mutants are constitutively active. A-C shows mutants streaked onto an adenine selective plate without ligand. All colonies display growth on plates A-C, and therefore, all variants are constitutively active. Plates D-F show the same mutants streaked onto synthetic complete media lacking leucine and tryptophan (SC-LW) nonselective plates.

streaked onto adenine selective plates without ligand (SC-ALW) as well as nonselective plates (SC-LW). The expected result was to obtain a variant that was not constitutively active and ligand-activated by one of the ligands on the plates. Streaking potential candidates showed that all variants were constitutively active, showing growth on the adenine selective plate without ligand (Figure 2.11).

Colonies were obtained and sequenced alongside variants from the nonselective plates, providing sequencing for approximately 0.02 % of the total library (Table 2.2). Sequencing confirmed mutational diversity at each of the amino acid positions in the library, however all variants from the nonselective plates contained frame shifts. The diversity in the sequencing can be seen, for example, at C284, which had multiple variants with a mutation to a serine in the selective variants in comparison to those seen in the nonselective variants. Additionally, phenylalanine 420 maintained the wild-type phenylalanine at that position for most of the selective variants and had greater diversity in the mutations for the nonselective variants, demonstrating a lack of tolerance for change at that position. Despite the overall sequence diversity, definitive conclusions about the tolerance at each position cannot be made due to the presence of the frame shift mutations found from sequencing of variants from the nonselective plates.

Although the results did not produce any variants that were not constitutively active, this was not surprising due to the activity of wild-type PXR in chemical complementation. However, these results display the power of genetic selection through ability to select functional versus nonfunctional protein variants as well as obtain sequence diversity. Wild-type PXR containing its own DBD is not constitutively active in

Table 2.2 PXR Library Sequencing

	LIV_2FM	LIV_2FM	LIV_2FM	CISF	LIV_2FM	F_2WLC_2	LIV_2FM	HQ	LIV_2FM	FDYV
Residue	**L240**	**M243**	**F281**	**C284**	**Q285**	**W299**	**M323**	**H407**	**L411**	**F420**
					Nonselective Variants					
Variant										
1*	F	V	V	I	I	C	I	H	V	Y
2*	M	V	I	C	V	W	F	Q	na	na
3*	I	I	L	I	M	W	F	Q	I	V
4*	I	I	V	I	F	F	F	H	V	F
5*	na	F	I	S	I	L	V	Q	V	Y
6*	M	V	F	C	L	W	F	H	I	D
7*	V	V	M	I	M	F	M	H	F	V
					Selective Variants					
1	M	I	I	I	V	F	F	Q	F	F
2	L	V	I	S	V	F	L	H	L	Y
3	L	V	V	I	I	F	L	Q	F	Y
4	L	V	V	S	V	W	M	Q	F	F
5	L	V	M	S	I	F	I	H	L	F
6	L	L	F	S	I	F	L	H	V	F
7	L	F	V	I	V	L	V	H	I	F
8	V	F	M	S	I	F	L	Q	V	F
9	L	F	F	S	V	L	L	H	L	F
10	L	I	F	S	V	L	L	H	L	F
11	I	F	I	F	F	F	V	Q	I	F
12	L	F	V	S	I	F	F	Q	F	Y
13	L	F	V	I	F	F	V	Q	V	F
14	L	F	L	I	F	F	V	Q	I	F
15	V	L	I	I	L	F	V	Q	I	F
16	L	L	F	S	V	W	L	H	V	F
17	L	I	M	I	I	F	F	H	L	Y

*: contains frame shifts
na: unclear sequencing readout

mammalian cells. However, the Gal4DBD:PXRLBD construct that was used in chemical complementation had not been tested previously to determine whether in cell culture this fusion protein was ligand-activated. To confirm that PXR was only constitutively active in the chemical complementation selection system, the Gal4DBD:PXRLBD was tested for its activity in mammalin cell culture.

2.5 PXR in HEK293T Mammalian Cells

Full length PXR containing the DBD and LBD is shown to be ligand activated in mammalian cells by ligands ranging from the large ligand rifampicin to smaller steroid-like ligands, such as pregnenolone [18, 21, 34]. In chemical complementation, the Gal4 DNA binding domain (GBD) is fused to the PXR LBD, however, this fusion had not been tested in mammalian cells. The GBDPXRLBD gene was therefore cloned into the mammalian expression vector pCMX, which contains a cytomegalovirus (CMV) promoter, using *SacII* and *NotI* restriction sites. This plasmid, pCMXGBDPXRLBD, was transfected into human embryonic kidney (HEK293T) cells along with a reporter plasmid containing four Gal4 response elements controlling a luciferase gene (p17*4TATALuc), as previously described. When ligand binds the LBD and the Gal4 DBD binds the Gal4 response elements, the luciferase gene should be expressed and luciferase activity can be measured. As a control, a plasmid with a β-galactosidase gene (pCMXβgal) was transfected for use in a β-galactosidase colormetric assay to normalize the GBDPXRLBD activation levels [35].

A dose response using rifampicin was performed using the HEK293T mammalian cell line. PXR was shown to be activated by rifampicin with a fold activation level of 22±5 and an EC_{50} of approximately 3 μM, which is comparable to published values

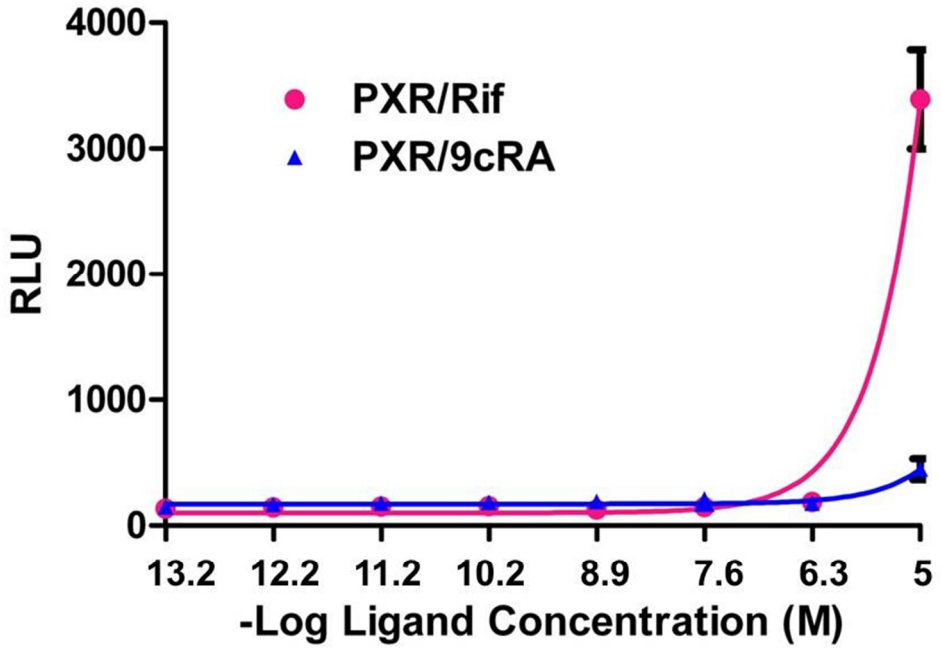


Figure 2.12. Activation profiles in HEK293T cells for Gal4DBD-PXRLBD (GBDPXR) with the Gal4RE reporter controlling *renilla* luciferase (RLU) in response to rifampicin and 9cRA.

around 5 μM (Figure 2.12) [17, 18]. As a control, PXR was also tested in a dose response with the RXR ligand 9cRA. No activation was observed in response to 9cRA, as expected, since this is the ligand for RXR. An interesting fact determined from this set of experiments is that the PXR functions as a homodimer when fused to the Gal4 DBD in mammalian cells. Overall, the results have shown that PXR is ligand activated in mammalian cells and confirms that PXR is constitutively active in yeast. One possible reason for the constitutive activity observed in yeast is an endogenous yeast metabolite activating PXR not present in mammalian cells.

2.6 Summary and Future Work

The goal of this work was to engineer a PXR variant able to bind novel small molecules, specifically the β-lactam antibiotics, through analysis of a PXR library in chemical complementation. PXR was found to be constitutively active; therefore, the first step was to engineer a PXR variant to no longer be constitutively active. Sequencing results of the library showed variants with sequence diversity at the randomized amino acid positions. Streaking results, however, showed that all mutants were constitutively active, showing growth on the adenine selective plate without ligand. The constitutive activity with PXR was only observed in yeast. When tested in mammalian cell culture, PXR showed ligand activation with rifampicin.

Due to the constitutive activity seen in the PXR library, and to continue pursuing a receptor with the ability to bind and activate in response to the β-lactam antibiotics, several different approaches were taken. First, a knockout yeast strain was developed to possibly eliminate metabolites that could be activating PXR and leading to PXR's constitutive activity. Secondly, due to PXR's constitutive nature in yeast, a new receptor,

the human estrogen receptor alpha (hERα), was chosen as the target for protein engineering with chemical complementation.

2.7 Materials and Methods

2.7.1 Strains

PJ69-4A (*MATa trp1-901 leu2-3,112 ura3-52 his3-200 gal4*Δ *gal80*Δ *LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*), was a kind gift from Dr. Philip James (University of Wisconsin, Madison). The PJ69-4A strain contains three reporter genes, *HIS3*, *ADE2*, and *lacZ*, which are under the control of different Gal4 inducible promoters.

2.7.2 Plasmids and Primers

pCMX-hRXRα was a kind gift from Dr. Ronald Evans (Salk Institute for Biological Studies, La Jolla, CA) [103]. pGBDRXRα (containing residues 1-151 of the Gal4 gene and residues 44-462 of the RXR gene) was a kind gift of Dr. Kenji Miyata (McMaster University; Ontario, Canada) [104].

The plasmid pRESTA-PXR (containing residues 130-434) was obtained from Dr. Mathew Redinbo (University of North Carolina, Chapel Hill, NC) and was previously cloned into pGBD [36].

To make the pGBDPXR background plasmid, QuikChange™ (Stratagene) site-directed mutagenesis was used to insert *SacII* and *KpnI* sites into the PXR gene in pGBDPXR, producing the pGBDPXRSacIIKpnI plasmid. The plasmid was digested with *SacII* and *KpnI*, removing 500 bp of the PXR gene. A 145 bp random sequence of DNA was amplified from the plasmid pMSCVeGFP. The random DNA sequence was digested

with *SacII* and *KpnI* and ligated into the pGBDPXRSacIIKpnI plasmid between the two restriction sites, generating three STOP codons. The resulting plasmid was confirmed by sequencing and named pGBDPXRbackground (Operon, Huntsville, AL)

To make pCMXGBDPXRLBD, the GBDPXRLBD fusion was amplified from pGBDPXRLBD using PCR with the following primers: GBDSacIIFor 5'tccccgcggatgaagctactgtcttctatcgaacaag3' and PXRNotIRev 5'aaggaaaaaagcggccgctcagctacctgtgatgccgaac3'. The plasmid pCMXrERβ as well as the GBDPXRLBD cassette were digested with *SacII* and *NotI*, and ligated to create the pCMXGBDPXRLBD plasmid.

2.7.3 Ligands

9-*cis* retinoic acid (MW=300.44 g/mol) was purchased from ICN Biomedicals (Aurora, OH). Amoxicillin (MW=365.41 g/mol) and Oxacillin (MW = 423.40 g/mol) were purchased from Fluka (Milwaukee, MI). Ampicillin (MW=371.4), Penicillin G (MW=372.48 g/mol), Cloxacillin (MW=475.9 g/mol), and Nafcillin (MW=454.5 g/mol) were purchased from Sigma (St. Louis, MO). Rifampicin (MW=823 g/mol) was purchased from Sigma (St. Louis, MO). 10 mM stocks of each ligand were dissolved in 80% ethanol:20% DMSO and stored at 4 °C.

2.7.4 PCR reactions

Oligonucleotides were ordered from Operon (Huntsville, AL) with overlapping ends shown in Table 2.2. To create the cassette containing the randomized codons at the ten positions, overlap extension PCR was utilized. Through PCR, the fragments can be were amplified in a PCR reaction containing 100 ng of each oligo, 125 ng of each primer, 100 mM dNTPs, and 1 μL *Pfu* polymerase. The PCR program was 95 °C for 2 minutes,

95 °C for 1 minute, 55 °C for 1 minute, 72 °C for 1 minute, repeated 30 times followed by 72 °C for 2 minutes. The resulting cassette was 788 bp and was gel purified using the Zymo Gel DNA Recovery Kit (Zymo Research, Orange, CA).

2.7.5 Yeast Transformations

Synthetic complete (SC) media were made as previously described [37]. Selection plates were made of synthetic complete media minus any of the following nutrients: histidine, leucine, tryptophan, adenine, and uracil. The pGBDPXR background plasmid was digested with *SacII* and *KpnI* and transformed with the PXR library cassette into PJ69-4A containing pGAD10BAACTR using the lithium acetate transformation method as described previously [38, 39]. Cells were plated onto synthetic complete media lacking leucine and tryptophan (SC-LW) and two sets of synthetic complete media lacking adenine, leucine, and tryptophan (SC-ALW) containing the antibiotics and incubated at 30 °C for 3 days. The first ligand set contained the antibiotics amoxicillin, cloxacillin, penicillin G as well as the ligand 17-β estradiol. The second ligand set contained the antibiotics ampicillin, oxacillin, nafcillin, as well as 9cRA and resveratrol. Transformation efficiency was determined from the number of colonies on the plate per μg of vector DNA added to each plate.

2.7.6 Transfection in HEK293T Cells

The transfection assay was performed as previously described [35]. Briefly, the pCMXGBDPXRLBD plasmid was transfected into human embryonic kidney (HEK293T) cells (ATCC, USA) with a reporter plasmid p17*4TATALuc, containing four Gal4 response elements controlling a *Renilla* luciferase gene as well as the

pCMXβgal plasmid, containing a β-galactosidase gene under the control of a CMV promotor. The cationic lipid used was lipofectamine 2000 (Invitrogen, USA).

2.8 References

1. Elander, R.P., *Industrial production of beta-lactam antibiotics.* Applied Microbiology and Biotechnology, 2003. **61**(5-6): p. 385-392.

2. Kliewer, S.A., et al., *An orphan nuclear receptor activated by pregnanes defines a novel steroid signaling pathway.* Cell, 1998. **92**(1): p. 73-82.

3. Lehmann, J.M., et al., *The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions.* Journal of Clinical Investigation, 1998. **102**(5): p. 1016-1023.

4. Kliewer, S.A. and T.M. Willson, *Regulation of xenobiotic and bile acid metabolism by the nuclear pregnane X receptor.* Journal of Lipid Research, 2002. **43**(3): p. 359-364.

5. Handschin, C. and U.A. Meyer, *Induction of drug metabolism: The role of nuclear receptors.* Pharmacological Reviews, 2003. **55**(4): p. 649-673.

6. Timsit, Y.E. and M. Negishi, *CAR and PXR: The xenobiotic-sensing receptors.* Steroids, 2007. **72**(3): p. 231-246.

7. Ekins, S., L. Mirny, and E.G. Schuetz, *A ligand-based approach to understanding selectivity of nuclear hormone receptors PXR, CAR, FXR, LXR alpha, and LXR beta.* Pharmaceutical Research, 2002. **19**(12): p. 1788-1800.

8. Bertilsson, G., et al., *Identification of a human nuclear receptor defines a new signaling pathway for CYP3A induction.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(21): p. 12208-12213.

9. Blumberg, B., et al., *SXR, a novel steroid and xenobiotic-sensing nuclear receptor.* Genes & Development, 1998. **12**(20): p. 3195-3205.

10. Goodwin, B., et al., *Regulation of the human CYP2B6 gene by the nuclear pregnane X receptor.* Molecular Pharmacology, 2001. **60**(3): p. 427-431.

11. Cheng, X.G. and C.D. Klaassen, *Regulation of mRNA expression of xenobiotic transporters by the pregnane X receptor in mouse liver, kidney, and intestine.* Drug Metabolism and Disposition, 2006. **34**(11): p. 1863-1867.

12. Chen, Y.P., et al., *Induction of human CYP2C9 by rifampicin, hyperforin, and phenobarbital is mediated by the pregnane X receptor.* Journal of Pharmacology and Experimental Therapeutics, 2004. **308**(2): p. 495-501.

13. Moore, L.B., et al., *Pregnane X receptor (PXR), constitutive androstane receptor (CAR), and benzoate X receptor (BXR) define three pharmacologically distinct classes of nuclear receptors.* Molecular Endocrinology, 2002. **16**(5): p. 977-986.

14. Reschly, E.J., et al., *Functional evolution of the vitamin D and pregnane X receptors.* BMC Evolutionary Biology, 2007. **7**.

15. Auwerx, J., et al., *A unified nomenclature system for the nuclear receptor superfamily.* Cell, 1999. **97**(2): p. 161-163.

16. Krasowski, M.D., et al., *Evolutionary selection across the nuclear hormone receptor superfamily with a focus on the NR1I subfamily (vitamin D, pregnane X, and constitutive androstane receptors).* Nucl Recept, 2005. **3**: p. 2.

17. Watkins, R.E., et al., *2.1 angstrom crystal structure of human PXR in complex with the St. John's wort compound hyperforin.* Biochemistry, 2003. **42**(6): p. 1430-1438.

18. Chrencik, J.E., et al., *Structural disorder in the complex of human pregnane X receptor and the macrolide antibiotic rifampicin.* Molecular Endocrinology, 2005. **19**(5): p. 1125-1134.

19. Watkins, R.E., et al., *The human nuclear xenobiotic receptor PXR: Structural determinants of directed promiscuity.* Science, 2001. **292**(5525): p. 2329-2333.

20. Xue, Y., et al., *Crystal structure of the pregnane X receptor-estradiol complex provides insights into endobiotic recognition.* Molecular Endocrinology, 2007. **21**(5): p. 1028-1038.

21. Ekins, S., et al., *Evolution of pharmacologic specificity in the pregnane X receptor.* BMC Evolutionary Biology, 2008. **8**: p. 21.

22. Jones, S.A., et al., *The pregnane x receptor: A promiscuous xenobiotic receptor that has diverged during evolution.* Molecular Endocrinology, 2000. **14**(1): p. 27-39.

23. Ekins, S., et al., *Evolution of pharmacologic specificity in the pregnane X receptor.* BMC Evolutionary Biology, 2008. **8**.

24. Krasowski, M.D., Ni, A., Hagey, L. R., Ekins, S., *Evolution of promiscuous nuclear hormone receptors: LXR, FXR, VDR, PXR, and CAR.* Molecular and Cellular Endocrinology, 2010. **Article In Press**.

25. LeCluyse, E.L., *Pregnane X receptor: molecular basis for species differences in CYP3A induction by xenobiotics.* Chemico-Biological Interactions, 2001. **134**(3): p. 283-289.

26. Huang, P.X., V. Chandra, and F. Rastinejad, *Structural Overview of the Nuclear Receptor Superfamily: Insights into Physiology and Therapeutics.* Annual Review of Physiology, 2010. **72**: p. 247-272.

27. Moore, J.T., J.L. Collins, and K.H. Pearce, *The nuclear receptor superfamily and drug discovery.* ChemMedChem, 2006. **1**(5): p. 504.

28. Egea, P.F., et al., *Crystal structure of the human RXR alpha ligand-binding domain bound to its natural ligand: 9-cis retinoic acid.* Embo Journal, 2000. **19**(11): p. 2592-2601.

29. Xue, Y., et al., *Crystal structure of the PXR-T1317 complex provides a scaffold to examine the potential for receptor antagonism.* Bioorganic & Medicinal Chemistry, 2007. **15**(5): p. 2156-2166.

30. Bloom, J.D. and F.H. Arnold, *In the light of directed evolution: Pathways of adaptive protein evolution.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**: p. 9995-10000.

31. James, P., J. Halladay, and E.A. Craig, *Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast.* Genetics, 1996. **144**(4): p. 1425-1436.

32. Lohr, D., P. Venkov, and J. Zlatanova, *Transcriptional regulation in the yeast GAL gene family: A complex genetic network.* Faseb Journal, 1995. **9**(9): p. 777-787.

33. Higuchi, R., B. Krummel, and R.K. Saiki, *A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions.* Nucleic Acids Res, 1988. **16**(15): p. 7351-67.

34. Lemaire, G., et al., *Discovery of a highly active ligand of human pregnane X receptor: a case study from pharmacophore Modeling and virtual screening to " in vivo" biological activity.* Molecular Pharmacology, 2007. **72**(3): p. 572-581.

35. Taylor, J.L., et al., *Characterization of a molecular switch system that regulates gene expression in mammalian cells through a small molecule.* BMC Biotechnology, 2010. **10**: p. 12.

36. Azizi, B., *Chemical complementation: A genetic selection system for drug discovery, protein engineering, and deciphering biosynthetic pathways (PhD)*. 2005, Georgia Institute of Technology: Atlanta, GA.

37. Frederick M., A.R., Robert E. Kingston, David D. Moore, J. G. Seidman, John A. Smith, Kevin Struhl, ed. *Short Protocols in Molecular Biology*. 4th ed. 1999, Wiley John & Sons Inc.

38. Schwimmer, L.J., et al., *Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(41): p. 14707-14712.

39. Gietz, R.D. and R.A. Woods, *Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method*, in *Guide to Yeast Genetics and Molecular and Cell Biology, Pt B*. 2002, Academic Press Inc: San Diego. p. 87-96.

CHAPTER 3

KNOCKOUT YEAST STRAIN AND PROTEIN EXPRESSION

3.1 PXR is Activated by Natural and Synthetic Steroids

Although the primary function of PXR is the removal of a large variety of xenobiotics from the body, PXR is also known as the steroid xenobiotic receptor (SXR) because of this receptor's ability to be activated in response to naturally occurring steroids such as pregnenolone, progesterone, 17α-hydroxyprogesterone, cortisol, and other intermediates in the steroid biosynthetic pathway in mammalian cells [1-4]. This receptor is also activated by synthetic steroids such as pregnenolone16α-carbonitrile (PCN), dexamethasone, and RU486 [1, 3, 4]. As mentioned in Chapter 2, PXR, through interaction with steroids, mediates regulation of bile acid homeostasis [5, 6]. Thus, this promiscuous receptor not only functions as a xenobiotic receptor, but also plays a significant role in steroid metabolism [5, 6].

In Chapter 2, when PXR was tested in the chemical complementation yeast selection system, this receptor was found to be constitutively active. One hypothesis for the observed constitutive activity of PXR in the chemical complementation is the presence of an endogenous yeast metabolite activating the receptor. Metabolites in yeast and mammalian cells vary due to the different metabolic pathways of these two oraganisms. Cholesterol is the primary sterol in mammalian cells, whereas ergosterol serves this role in *Saccharomyces cerevisiae* (Figure 3.1) [7-13].

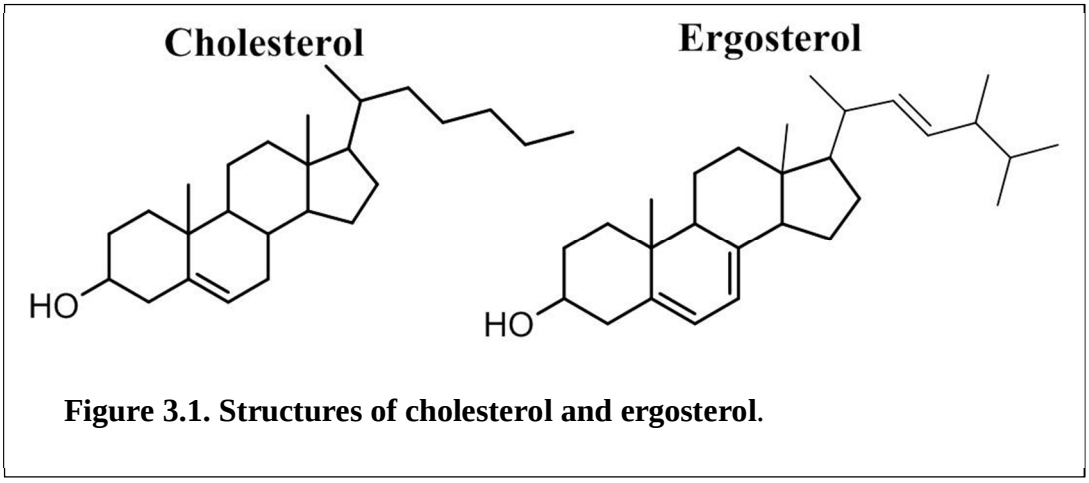


Figure 3.1. Structures of cholesterol and ergosterol.

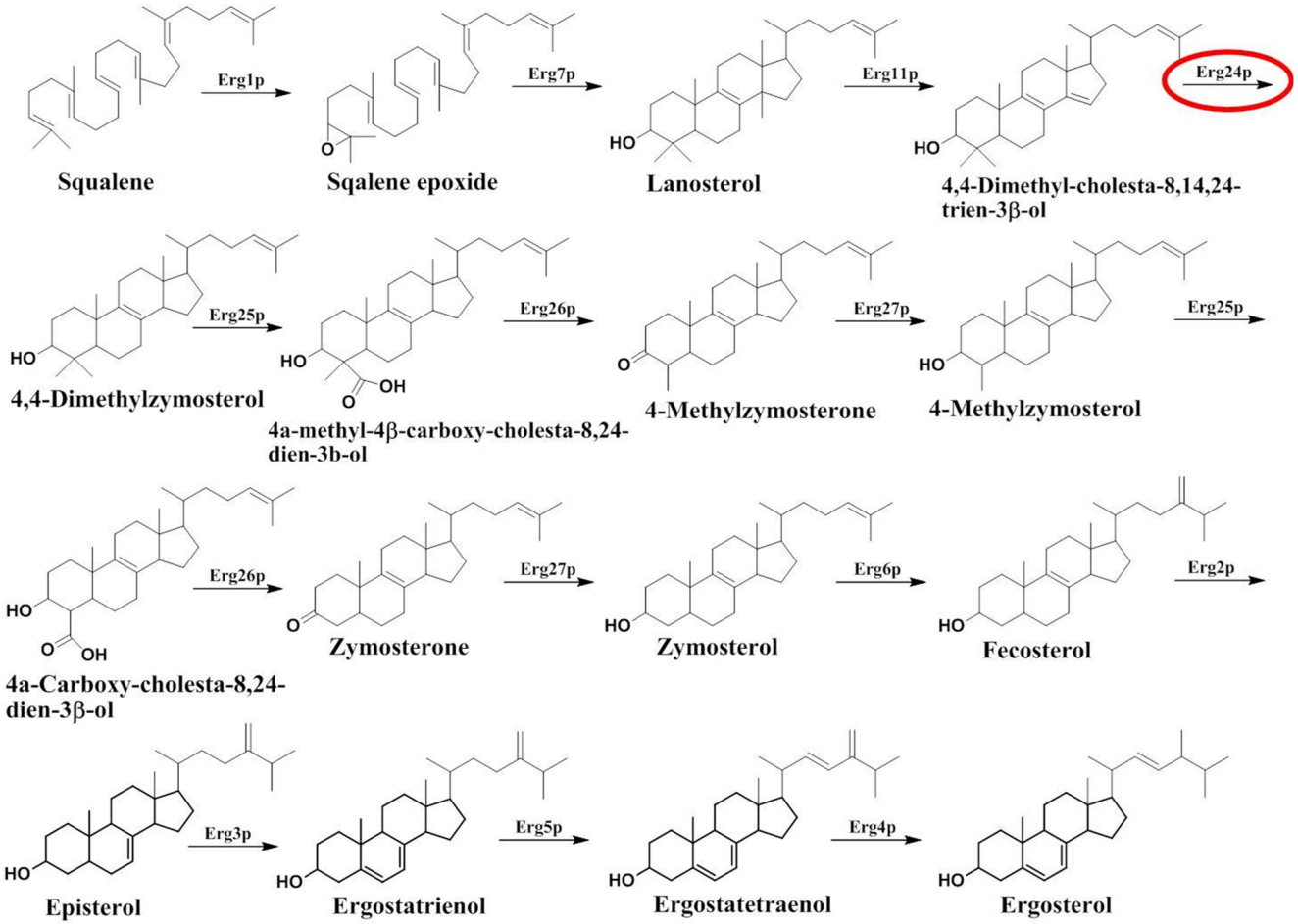


Figure 3.2 The ergosterol biosynthetic pathway. The enzyme for each step is shown above the arrow. The names of each steroid in the pathway are indicated.

3.2 The Ergosterol Biosynthetic Pathway in Yeast

The yeast biosynthetic pathway from squalene to ergosterol consists of fifteen steps involving twelve of different enzymes (Figure 3.2) [8, 9, 11-15]. The enzymes for each step in this pathway, designated ERG, have been characterized [8, 11, 13]. Several of the ERG genes, ERG24, ERG6, ERG2, ERG3, ERG5, and ERG4, are considered nonessential, such that when the gene is deleted or the enzyme is inhibited, the yeast are still able to survive despite various side effects, such as disruption of membrane permeability, mating and other cellular functions [11, 13, 15, 16]. In the biosynthetic pathway, the *ERG24* gene, encoding the sterol C-14 reductase, is the most upstream nonessential gene in the synthesis of ergosterol [13, 17-20]. Previous work has shown that the *ERG24* null yeast produce an accumulation of the steroid ignosterol, a steroid not toxic to the yeast [13, 21].Thus, eliminating the *ERG24* gene and consequently the sterol C-14 reductase enzyme from the biosynthetic pathway could lead to the disruption of a number of the steroid metabolites, perhaps removing the metabolite binding PXR. The hypothesis was made that by creating a new strain containing an *ERG24* knockout, downstream metabolites would not be produced and eliminating the endogenous ligand activating PXR.

3.2.1 ERG24 Knockout Design

As the first nonessential gene in the yeast ergosterol biosynthetic pathway, *ERG24* was chosen to be deleted from the yeast strain BAPJ69-4A [8, 22, 23]. BAPJ69-4A, a derivative of the strain PJ69-4A, was created to contain three selective markers *ADE2*, *HIS3*, and *URA3* as well as the *lacZ* gene controlled by Gal4 promoters [22-24]. In order to determine the effectiveness of the gene knockout, the gene is replaced with a

selectable marker [25, 26]. Wach *et al.* developed a system where the kanamycin (*Kan*) gene is used as the selectable marker to disrupt the gene of interest, rendering yeast resistant to geneticin [26, 27]. Geneticin is an antibiotic, that inhibits the elongation step of protein synthesis [28]. In order for yeast to survive in the presence of geneticin, a resistant gene, such as kanamycin needs to be present. If the gene of interest has been replaced with the *Kan* gene, the yeast should be able to grow in the presence of geneticin. Therefore, to develop the knockout strain, the *ERG24* gene was replaced by a *Kan* selection marker [26, 27, 29].

To disrupt the *ERG24* gene, a cassette containing the Kan gene was constructed. The Kan insert cassette (KanIC) should hypothetically contain approximately 350 base pairs of upstream (UF) and downstream (DF) regions of the target *ERG24* gene. Approximately 30 base pairs of the beginning and end of the *ERG24* gene, and a *Kan* gene inserted in place of the remainder of the *ERG24* gene (Figure 3.3). Experimentally, to create the cassette containing the *Kan* gene for the knockout strain several different PCR reactions were required.

First, the genomic DNA of the BAPJ69-4A yeast strain was isolated using a Zymo Yeast Genomic DNA Prep Kit (Zymo Research, Orange, CA). The genomic DNA was used as the template in a PCR reaction to isolate the UF and DF regions of approximately 300-400 base pairs upstream and downstream of the *ERG24* gene (Figure 3.3A). These UF and DF PCR fragments contain 20-40 base pairs of the N-terminal and C-terminal ends of the *ERG24* gene, respectively, as well as a 20-30 base pair sequence of the *Kan* gene (Figure 3.3A). The terminal sequences of the *ERG24* gene are required

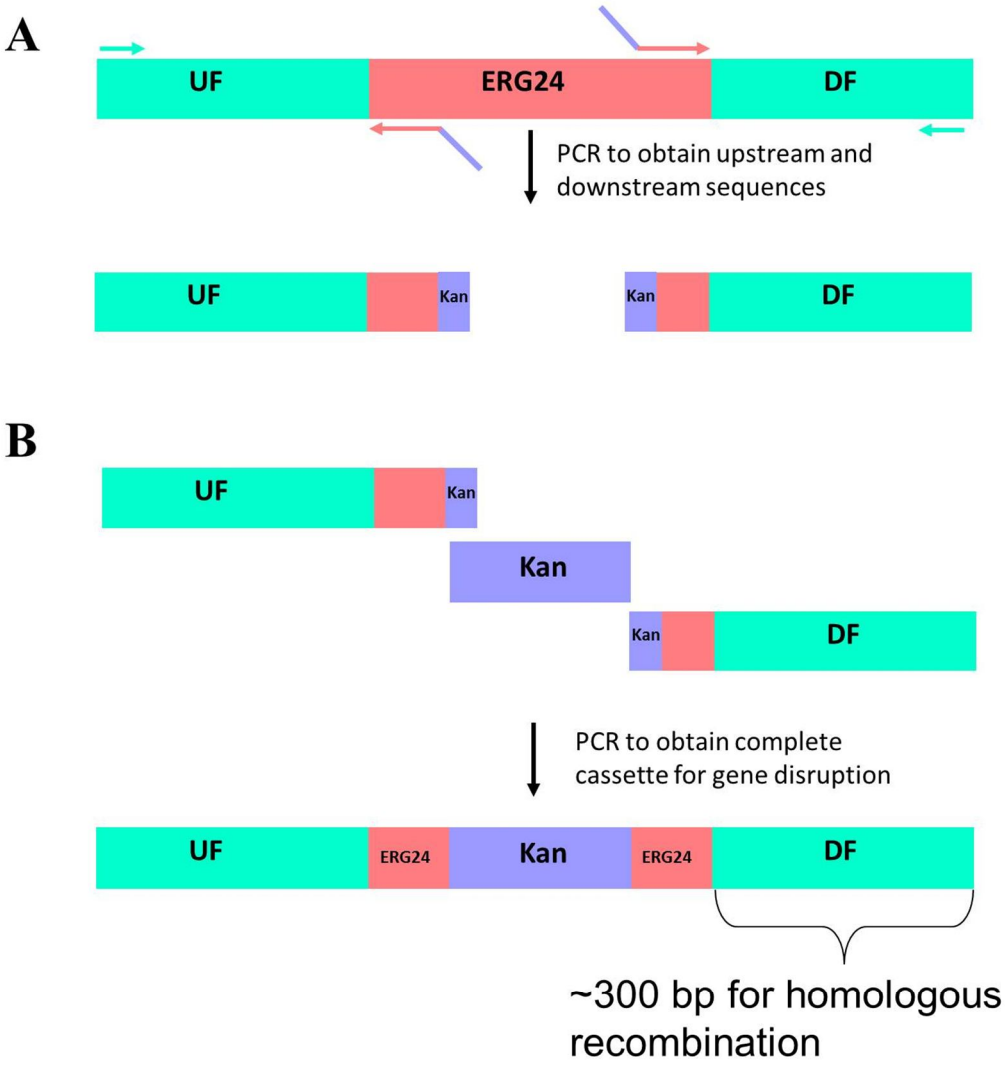


Figure 3.3. Scheme for making the cassette containing the disrupted *ERG24* gene. (A) Upstream and Downstream flanking regions (UF and DF) from the *ERG24* gene are made through PCR. (B) The UF and DF products are used in a PCR reaction with the kanamycin (*Kan*) gene to generate the complete cassette to use in the transformation to make the knockout strain.

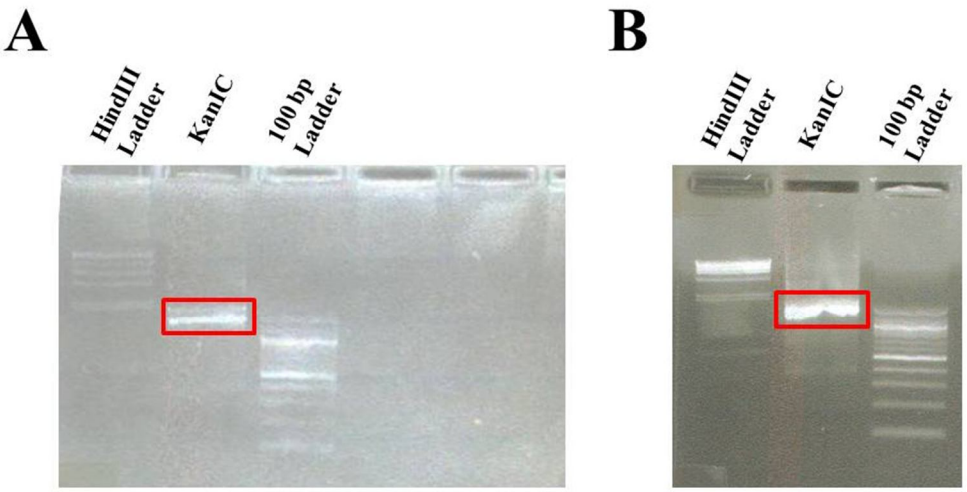


Figure 3.4. KanIC PCR products. (A) Gel showing concentration of KanIC PCR product. Band size should be approximately 1950 base pairs. (B) Gel showing concentration of KanIC from secondary PCR of product from gel in A.

so enough of the *ERG24* sequence is present for homologous recombination to occur with the gene targeted for disruption [26, 27]. As seen in Figure 3.3B, once the UF and DF fragments were created, the fragments were used as primers in a PCR reaction with a plasmid containing the *Kan* gene (Figure 3.3B). The PCR reaction did not produce the expected KanIC cassette and therefore a different PCR reaction was designed.

The goal therefore was to generate the full KanIC as a "full pot" synthesis through a PCR containing all the necessary components. This reaction contained the genomic yeast DNA, the pET28a plasmid containing the *Kan* gene, all primers needed to make the UF and DF fragments, and two the primers needed to amplify the *Kan* gene. Through PCR, the fragments can be amplified and hybridize with the complementary ends, producing a fragment that is approximately 1,950 base pairs. As seen in Figure 3.4A, an insert cassette containing the correct band size was made (Figure 3.4A). To increase the concentration of the KanIC, a secondary PCR reaction was performed using the KanIC as the template in a reaction containing the outermost primers for the full cassette and confirmed via sequencing (Figure 3.4B).

3.2.2 Strain Construction and Analysis

Prior to transforming the KanIC into the BAPJ69-4A yeast strain, the optimal concentration of geneticin to use for the knockout strain was determined. Previously, Wach *et al.* had assessed the optimal concentration of geneticin to be 0.2 mg/ml with the yeast strain [27]. To determine the optimal concentration of geneticin for BAPJ69-4A, the strain was grown on a rich yeast media containing yeast extract, peptone, and dextrose (YPD) with a range of geneticin concentrations above and below 0.2 mg/mL. A

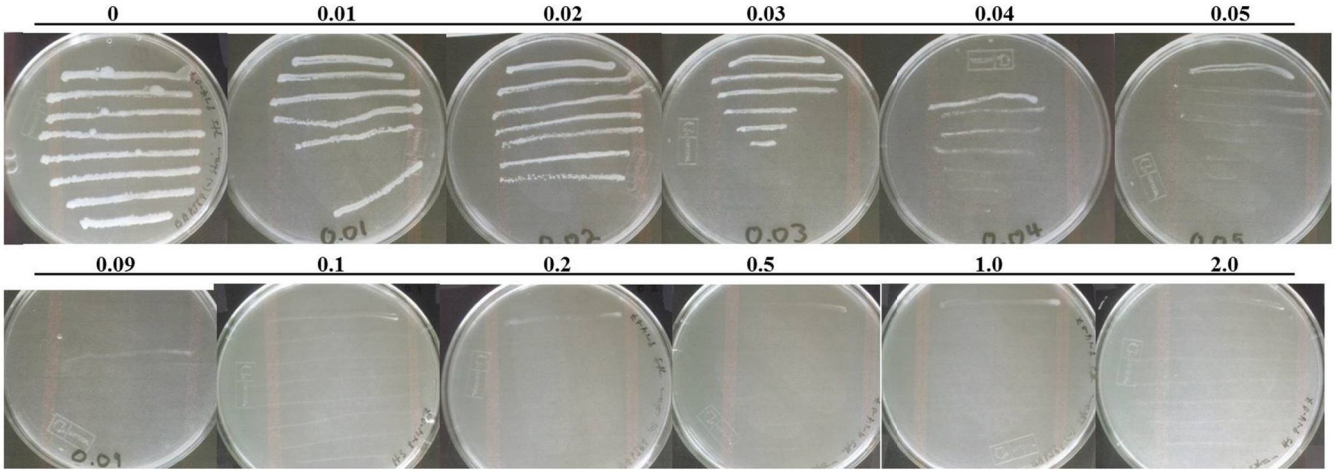


Figure 3.5. Streaking BAPJ69-4A. YPD plates containing geneticin concentrations of 0.01, 0.02, 0.03, 0.04, 0.05, 0.09, 0.1, 0.2, 0.5, 1.0, and 2.0 mg/mL. The geneticin concentration for each plate is indicated above the plate (mg/mL).

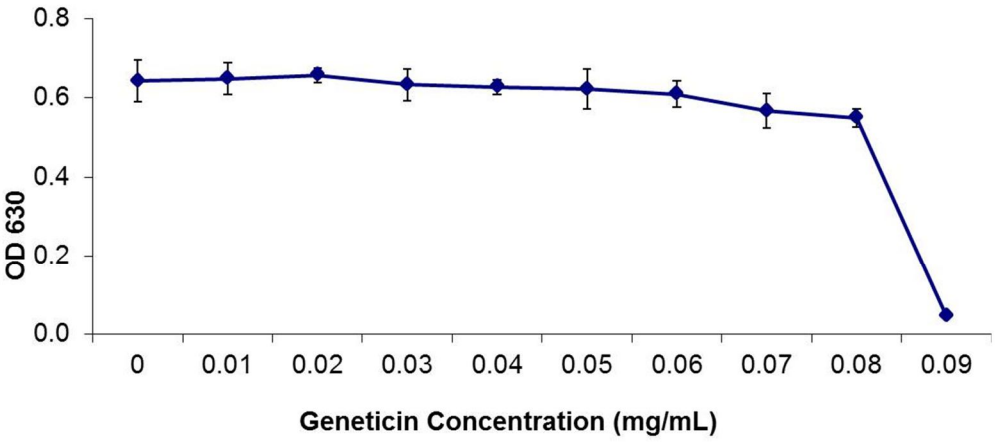


Figure 3.6. Growth assay with BAPJ69-4A yeast strain in YPD media with increasing concentrations of geneticin.

YPD plate without geneticin was used as a control (Figure 3.5). The yeast should survive on the YPD plate and should not grow on the media with the appropriate concentration of geneticin required to kill the parent BAPJ69-4A strain. No growth was observed for BAPJ69-4A on the YPD plates containing geneticin at concentrations 0.08 mg/mL and above (Figure 3.5). A liquid growth assay was also performed with varying concentrations to confirm the optimal geneticin concentration needed to make the knockout strain. No growth was observed at 0.09 mg/mL geneticin, therefore this concentration was used in the yeast transformation with the KanIC (Figure 3.6). If integration of the KacIC occurs, the new strain should be able to grow on media containing 0.09 mg/mL geneticin.

The KanIC containing the disrupted gene was transformed into BAPJ69-4A using the lithium acetate transformation protocol as previously described [30, 31]. Yeast transformants were grown on YPD media containing 0.09 mg/mL geneticin to determine if integration of the *Kan* gene at the *ERG24* gene in the genome occured via homologous recombination. Seventy colonies grew on the YPD plate containing geneticin and were further analyzed for resistance to the antibiotic through confirming growth on a YPD plate containing 0.09 mg/ml geneticin. All colonies displayed growth on the YPD plate and approximately 25 showed growth on the YPD plate containing geneticin.

Ten of the potential knockout strain mutants were tested for integration. A genomic PCR was performed on these 10 mutant strains followed by a PCR to determine if the KanIC was integrated based on the size of the PCR fragment. Secondary PCRs for the best four mutants from this result was performed. One variant, called I1, showed a band with the correct size, and the strain was further characterized alongside BAPJ69-4A

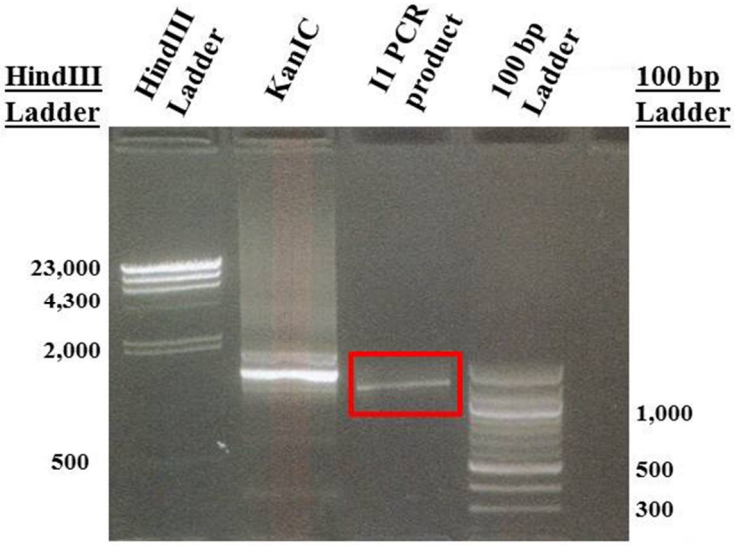


Figure 3.7 Gel of original KanIC PCR product and PCR cassette from genome prep of I1 potential knockout strain.

by streaking onto YPD plates with and without geneticin to confirm resistance to geneticin (Figure 3.7). Both strains grow on the YPD plate, as expected. BAPJ69-4A did not show growth on the YPD plate with geneticin, whereas I1 shows growth on the plate containing geneticin (Figure 3.8). However, sequencing results of the I1 fragment showed a hybrid of the knockout strain and the parent BAPJ69-4A strain. This can occur when two haploid cells mate, resulting in diploid yeast, containing two genomes [32]. In this case, the hypothesis is that the two strains, the parent and the knockout, mated, resulting in a hybrid strain containing both genomes. The complete experiment with the transformation of the KanIC was repeated to obtain a complete knockout strain, but no positive colonies were obtained.

Despite sequencing showing I1 as a hybrid strain, the GBDPXRLBD construct was tested for constitutive activity in the hybrid strain. The GBDPXRLBD fusion plasmid (pGBDPXRLBD), containing a tryptophan marker, and the GAD:ACTR fusion (pGAD10BAACTR) plasmid, containing a leucine marker, were transformed into the hybrid strain as performed in Chapter 2 and grown on synthetic complete media lacking leucine and tryptophan with 0.09 mg/mL geneticin (SC-LW+geneticin), to maintain the *ERG24* knockout containing the *Kan* gene. As a control, PXR was also transformed into the parent BAPJ69-4A strain. All colonies from the SC-LW+geneticin plate were streaked for constitutive activity onto a synthetic complete media lacking adenine, leucine, and tryptophan with 0.09 mg/mL geneticin (SC-ALW+geneticin) plate. Streaking showed growth for all colonies, confirming that PXR was still constitutively active in the hybrid strain (Figure 3.9).

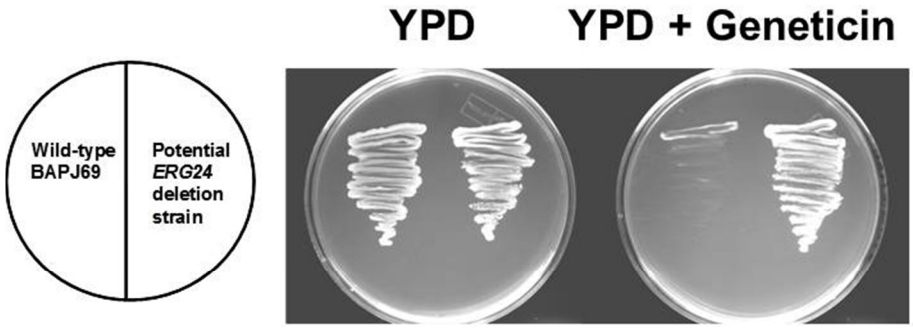


Figure 3.8 Streaking BAPJ69-4A and I1 potential knockout strain on a YPD nonelective plate and a YPD plate containing 0.09 mg/mL geneticin. Both yeast strains grow on the YPD plate. BAPJ69-4A does not show growth on the plate containing geneticin, whereas the potential knockout strain displays growth.

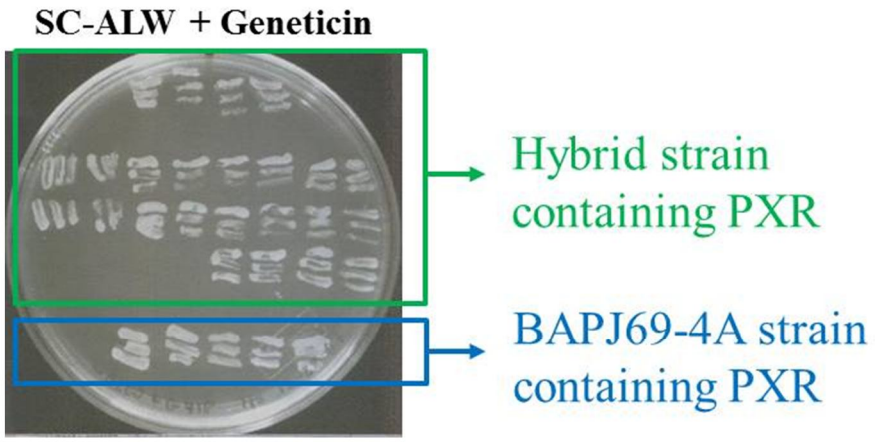


Figure 3.9. Streaking GBDPXRLBD in the hybrid strain and the BAPJ69-4A parent strain on synthetic complete media lacking adenine, leucine, and tryptophan containing 0.09 mg/mL geneticin (SC-ALW+Geneticin). All PXR transformants in the BAPJ69-4A and the hybrid strain show growth.

In summary, we constructed a cassette designed to knockout the *ERG24* gene in the strain BAPJ69-4A and replace the gene with a *Kan* gene to allow the yeast to grow on plates containing the antibiotic geneticin. A strain was engineered that grew on geneticin selective plates, however sequencing revealed a hybrid strain between the parent strain and one with the *Kan* gene integrated into the *ERG24* gene. In addition, PXR was still determined to be constitutively activity in the hybrid strain. Thus, we could not determine if an endogenous steroid is activating PXR in chemical complementation. The next step was identifying the ligand binding and activating PXR. Therefore, PXR was overexpressed and purified from yeast to identify the ligand bound to the LBD.

3.3 Expression and purification in yeast

3.3.1 Yeast Expression using pESCtrp

To express PXR in yeast, the PXRLBD alone and the GBDPXRLBD fusion containing N-terminal His tags were cloned into the yeast expression plasmid pESCtrp using restriction enzymes *NotI* and *SpeI*, placing the PXR gene under the control of a Gal10 promoter. The Gal10 promoter in the plasmid can be induced with galactose [33]. The resulting plasmid, pESCtrpGBDPXRLBD, containing a tryptophan marker, was transformed into PJ69-4A with pGAD10BAACTR, containing the GAD:ACTR fusion protein, as before. As a control, an empty pESCtrp vector was transformed with pGAD10BAACTR.

Transformants were grown to competency in synthetic complete media lacking leucine and tryptophan (CM-LW) media, selecting for both plasmids, and induced in media containing 2 % galactose overexpressing the PXR gene. After cell lysis, the GBDPXRLBD protein was purified using Clontech Talon ® Co^{2+} resin (Mountain View,

CA). To obtain pure protein, the soluble fraction containing all the proteins is subjected to a column containing the Co^{2+} resin. The GBDPXRLBD protein contains a histidine sequence, which is able to bind the cobalt resin, and pure protein is eluted. A buffer containing imidazole is used, where the desired 52 kDa protein is observed on a SDS-PAGE protein gel. As shown in Figure 3.10, the protein was present in both the initial soluble and insoluble fractions. However, the purified and dialyzed fractions only show a band approximately 30 kDa. The same band was observed in the control sample containing the empty pESCtrp vector (Figure 3.10). Several tactics were taken to troubleshoot and enhance expression of PXR. Varying galactose concentrations during the induction time or varying growth times did not produce a protein band with the correct size (Data not shown).

Yeast are not optimal organisms for overexpression of a protein, however RXR has successfully been expressed, purified in yeast, and tested for activation with 9cRA [34]. One hypothesis for the lack of a purified protein on the gel could be due to expression levels. Expression is strain dependent, and using galactose induction does not seem to suit the BAPJ69-4A strain. To troubleshoot the low expression levels, a new plasmid lacking galactose-dependent induction was used.

3.3.2 Yeast Expression using pUGPD

The pUGPD vector, containing a glyceraldehyde 3-phosphate dehydrogenase (GPD) promoter, was obtained. This promoter is continually expressed, and does not require induction [35, 36]. The GBDPXRLBD gene was cloned into pUGPD using the *NotI* restriction site and transformed with the pGAD10BAACTR plasmid into BAPJ69-

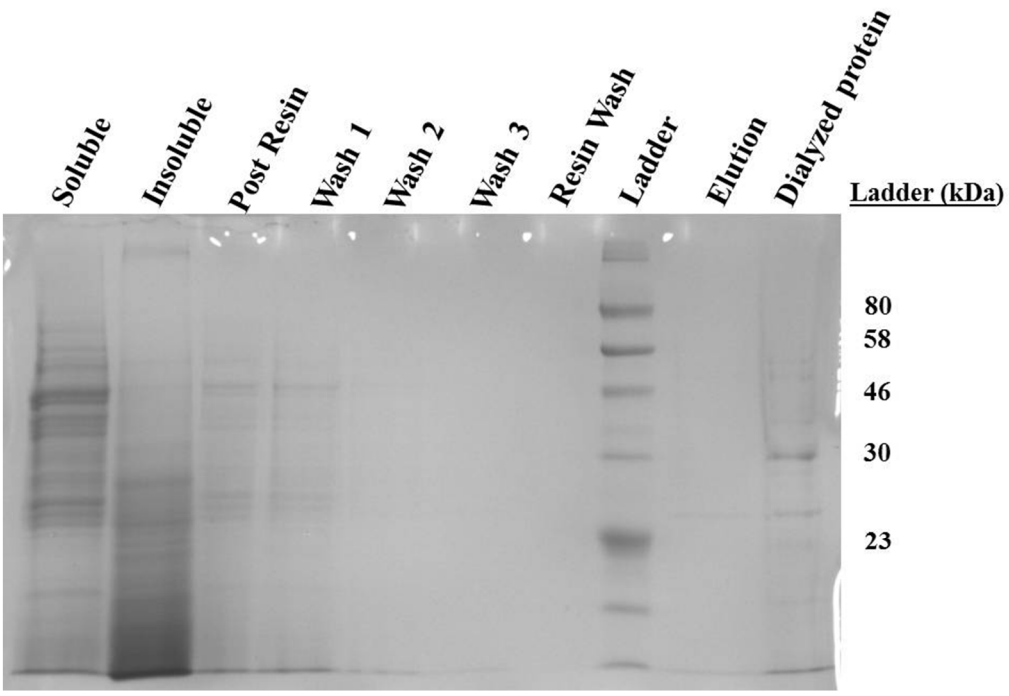


Figure 3.10. Gel of protein expression using pESCtrpGBDPXRLBD vector with induction using galactose.

4A. The pUGPD empty vector with pGAD10BAACTR was used as a negative control. Yeast cells were plated onto synthetic complete media lacking leucine and uracil (SC-LU), selecting for the GAD:ACTR and pUGPDGBDPXRLBD vectors.

To express GBDPXRLBD in the pUGPD vector, transformants were grown in 25 mL of SC-LU media at 30 °C overnight. Various lysis mechanisms were used to determine which method would produce a higher overall protein yield in the soluble fraction. These lysis methods included different lysis buffers, varying concentration of zymolase, incubation times, and the use of glass beads. The protein gel with the soluble and insoluble fractions for each method showed similar levels of protein concentration for five different lysis methods (data not shown). The soluble fraction from the first lysis method was purified as done previously. However, low levels of protein concentration were observed in all fractions, and as seen in Figure 3.11, the same incorrect 30 kDa band is observed in the elution fraction (Figure 3.11).

3.4 Summary and Future work

In summary, the design of an *ERG24* knockout strain used to remove the constitutive activity seen with PXR produced a hybrid strain. To identify the potential small molecule binding and activating PXR in yeast, PXR was overexpressed and purified from the yeast strain. Two different yeast expression plasmids were used, one containing a galactose inducible promoter and one with a constitutively expressed promoter, as well as multiple lysis methods. However, the protein concentrations remained low and troubleshooting did not enhance protein levels. Therefore, since PXR was constitutively active in chemical complementation, this receptor is not an ideal

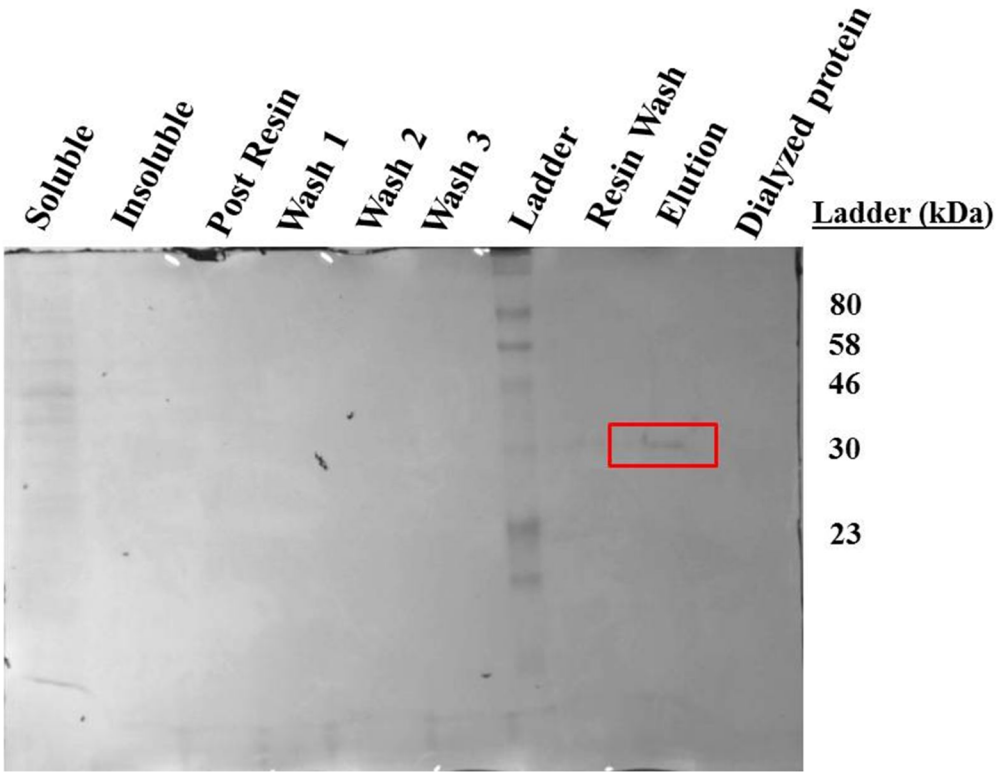


Figure 3.11. Gel of purified protein from expression in yeast with pUGPDGBDPXRLBD. A 30 kDa artifact is indicated with the red box.

candidate for protein engineering. Thus, the target nuclear receptor was switched to the human estrogen receptor alpha (hERα).

3.5 Materials and methods

3.5.1 Yeast Strain

The yeast strain used was the BAPJ69-4A strain developed in the lab with the genotype *MATa trp1-901 leu2-3, 112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ ho::GAL1-URA3* [22, 23].

3.5.2 Plasmids

The pGBDPXRLBD and pGAD10BAACTR plasmids are described in Chapter 2. The pET28a plasmid was used to obtain the kanamycin gene used in to make the knockout strain. Initial protein expression experiments in yeast were performed using the pESCTrp plasmid (Stratagene). The pUGPD vector was a kind gift from Dr. Yuri Chernoff's lab (Georgia Institute of Technology, Atlanta, GA). All sequencing was performed at Operon (Huntsville, AL).

3.5.3 Reagents

CelLytic-Y yeast cell lysis/extraction reagent and the galacose were purchased from Sigma (St. Louis, MO). The Talon ® affinity cobalt resin was purchased from Clontech (Mountain View, CA). Geneticin (G418) was purchased from Sigma (St. Louis, MO).

3.5.4 Primers and PCR Reactions

The PCR reactions to obtain the upstream flanking (UF) and downstream flanking (DF) regions of the *ERG24* gene contained 10x *Pfu* polymerase buffer, 100 mM dNTPs, 100 ng of genomic DNA, 125 ng of each primer, and 1 μL *Pfu* polymerase. The PCR

program was 95 °C for 2 minutes, 95 °C for 1 minute, 61 °C and 55 °C for the UF and DF regions, respectively for 1 minute, 72 °C for 1 minute, repeated 25 times followed by 72 °C for 2 minutes. These cassettes were gel purified using the Zymo Gel DNA Recovery Kit (Zymo Research, Orange, CA).

The "full pot" PCR reaction contained 10x *Pfu* polymerase buffer, 100 mM dNTPs, 100 ng of genomic DNA, 50 ng of pET28a, 125 ng of each primer, and 1 µL *Pfu* polymerase. The PCR program was 95 °C for 2 minutes, 95 °C for 1 minute, 55 °C for 1 minute, 72 °C for 2 minute, repeated 25 times followed by 72 °C for 3 minutes.

To create pESCtrpGBDPXRLBD, the GBDPXRLBD fusion was amplified from pGBDPXRLBD using PCR with the following primers: GLBDNotI_HisFor 5' aaggaaaaaa gcggccgc atgatg caccaccaccaccaccac3' and PXRSpeIRev 5' ctagactagtctcagctacctgtgatgccgaac3'. The pESCtrp plasmid as well as the GBDPXRLBD cassette were digested with *NotI*, and *SpeI* and ligated to create the pESCtrpGBDPXRLBD plasmid.

To create pUGPDGBDPXRLBD, the GBDPXRLBD fusion was amplified from pGBDPXRLBD using PCR with the following primers: GLBDNotI_HisFor 5' aaggaaaaaa gcggccgc atgatg caccaccaccaccaccac3' and PXRNotIRev 5'aaggaaaaaagcggccgctcagctacctgtgatgccgaac3'. The plasmid pUGPD plasmid as well as the GBDPXRLBD cassette were digested with *NotI*, and ligated to create the pUGPDGBDPXRLBD plasmid.

3.5.5. Transformation of KanIC

The KanIC (6 µg) was transformed into BAPJ69-4A using the lithium acetate protocol [31]. Yeast cells containing the transformed KanIC were added to YPD media

and incubated for 3 hours at 30 °C with shaking at 300 rpm. A suspension containing 1.5 mL of cells was transferred into a microcentrifuge tube and pelleted at 5,000 rpm for 1 minute**.** One milliliter of the supernatant was removed and the cells were resuspended in the remaining volume. Cells were plated onto YPD plates containing 0.09 mg/mL geneticin and grown at 30 °C for 5 days. Colonies containing the potential *ERG24* knockout strain were streaked onto a YPD plate and a YPD plate containing 0.09 mg/mL geneticin and grown at 30 °C for 2 days.

3.5.6 Expression and Purification

Transformants containing pGBDPXRLBD and pGAD10BAACTR plasmids were inoculated into 5 mL of synthetic complete media lacking leucine and tryptophan SC-LW) and grown at 30 °C overnight with shaking at 300 rpm. This overnight culture was used to inoculate a 200 mL culture of SC-LW media and the culture was grown to competency. The larger culture was pelleted by centrifugation at 4,000 rpm for 7 minutes and washed with water to remove all media. Cells were added to 300 mL of SC-LW media containing 2 % galactose as the carbon source to induce expression of the GBDPXRLBD gene and grown at 30 °C for 90 hours with shaking at 300 rpm. After 70 hours of growth, an additional 1 % galactose was added to the culture. The culture was pelleted by centrifugation at 8,000 rpm for 15 minutes, and the cells were lysed through resuspension in a CelLytic-Y yeast cell lysis/extraction reagent from Sigma (St. Louis, MO) followed by incubation with zymolase. The cell lysate was purified using the Talon ® affinity cobalt resin (Mountain View, CA).

To express GBDPXRLBD in the pUGPD vector, transformants were grown in 25 mL of SC-LU media at 30 °C overnight with shaking at 300 rpm. The 25 mL culture was

separated into 5 mL volumes; each lysed using a different lysis method to determine which would produce a higher overall protein yield in the soluble fraction.

SDS-PAGE gels of the protein fractions used 0.05 % SDS in trisglycine buffer through a 12 % polyacrylamide gel with a 4 % stacking gel.

3.6 References

1. Kliewer, S.A., et al., *An orphan nuclear receptor activated by pregnanes defines a novel steroid signaling pathway.* Cell, 1998. **92**(1): p. 73-82.

2. Cheng, X. and C.D. Klaassen, *Regulation of mRNA expression of xenobiotic transporters by the pregnane x receptor in mouse liver, kidney, and intestine.* Drug Metab Dispos, 2006. **34**(11): p. 1863-7.

3. Bertilsson, G., et al., *Identification of a human nuclear receptor defines a new signaling pathway for CYP3A induction.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(21): p. 12208-12213.

4. Kliewer, S.A. and T.M. Willson, *Regulation of xenobiotic and bile acid metabolism by the nuclear pregnane X receptor.* Journal of Lipid Research, 2002. **43**(3): p. 359-364.

5. Echchgadda, I., et al., *The xenobiotic-sensing nuclear receptors pregnane x receptor, constitutive androstane receptor, and orphan nuclear receptor hepatocyte nuclear factor 4 alpha in the regulation of human Steroid- Bile acid-sulfotransferase.* Molecular Endocrinology, 2007. **21**(9): p. 2099-2111.

6. Timsit, Y.E. and M. Negishi, *CAR and PXR: The xenobiotic-sensing receptors.* Steroids, 2007. **72**(3): p. 231-246.

7. Maczek, J., et al., *Metabolic flux analysis of the sterol pathway in the yeast Saccharomyces cerevisiae.* Bioprocess Biosyst Eng, 2006. **29**(4): p. 241-52.

8. Daum, G., et al., *Biochemistry, cell biology and molecular biology of lipids of Saccharomyces cerevisiae.* Yeast, 1998. **14**(16): p. 1471-1510.

9. Mo, C.Q. and M. Bard, *Erg28p is a key protein in the yeast sterol biosynthetic enzyme complex.* Journal of Lipid Research, 2005. **46**(9): p. 1991-1998.

10. Zinser, E., F. Paltauf, and G. Daum, *Sterol composition of yeast organelle membranes and subcellular disruption of enzymes involved in sterol metabolism.* Journal of Bacteriology, 1993. **175**(10): p. 2853-2858.

11. Davies, B.S.J., H.S. Wang, and J. Rine, *Dual activators of the sterol biosynthetic pathway of Saccharomyces cerevisiae: Similar activation/regulatory domains but different response mechanisms.* Molecular and Cellular Biology, 2005. **25**(16): p. 7375-7385.

12. Maczek, J., et al., *Metabolic flux analysis of the sterol pathway in the yeast Saccharomyces cerevisiae.* Bioprocess and Biosystems Engineering, 2006. **29**(4): p. 241-252.

13. Parks, L.W., et al., *Use of Sterol Mutants as Probes for Sterol Functions in the Saccharomyces cerevisiae.* Critical Reviews in Biochemistry and Molecular Biology, 1999. **34**(6): p. 399-404.

14. Athenstaedt, K., et al., *Identification and characterization of major lipid particle proteins of the yeast Saccharomyces cerevisiae.* Journal of Bacteriology, 1999. **181**(20): p. 6441-6448.

15. Gaber, R.F., et al., *The yeast gene ERG6 is required for normal membrane function but is not essential for biosynthesis of the cell-cycle-sparking sterol.* Molecular and Cellular Biology, 1989. **9**(8): p. 3447-3456.

16. Bhuiyan, M.S.A., et al., *Synthetically lethal interactions involving loss of the yeast ERG24: The sterol C-14 reductase gene.* Lipids, 2007. **42**(1): p. 69-76.

17. Lai, M.H., et al., *The identification of a gene family in the saccharomyces cerevisiae ergosterol biosynthesis pathway.* Gene, 1994. **140**(1): p. 41-49.

18. Lorenz, R.T. and L.W. Parks, *Cloning, sequencing, and disruption of the gene encoding sterol C-14 reductase in saccharomyces cerevisiae.* DNA and Cell Biology, 1992. **11**(9): p. 685-692.

19. Crowley, J.H., et al., *Aerobic isolation of an ERG24 null mutant of Saccharomyces cerevisiae.* Journal of Bacteriology, 1996. **178**(10): p. 2991-2993.

20. Crowley, J.H., S. Tove, and L.W. Parks, *A calcium-dependent ergosterol mutant of Saccharomyces cerevisiae.* Current Genetics, 1998. **34**(2): p. 93-99.

21. Marcireau, C., D. Guyonnet, and F. Karst, *CONSTRUCTION AND GROWTH-PROPERTIES OF A YEAST-STRAIN DEFECTIVE IN STEROL 14-REDUCTASE.* Current Genetics, 1992. **22**(4): p. 267-272.

22. Azizi, B., *Chemical complementation: A genetic selection system for drug discovery, protein engineering, and deciphering biosynthetic pathways (PhD)*. 2005, Georgia Institute of Technology: Atlanta, GA.

23. Shaffer, H.A., B. Azizi, and D.F. Doyle, *BAPJ69-4A: A Yeast Two-Hybrid Strain for Both Positive and Negative Genetic Selection.* To be submitted.

24. James, P., J. Halladay, and E.A. Craig, *Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast.* Genetics, 1996. **144**(4): p. 1425-1436.

25. Brachmann, C.B., et al., *Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications.* Yeast, 1998. **14**(2): p. 115-132.

26. Wach, A., et al., *New heterologous modules for classical or PCR-based gene disruptions in saccharomyces cerevisiae.* Yeast, 1994. **10**(13): p. 1793-1808.

27. Wach, A., *PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in S-cerevisiae.* Yeast, 1996. **12**(3): p. 259-265.

28. Jimenez, A. and J. Davies, *Expression of a transposable antibiotic resistance element in Saccharomyces.* Nature, 1980. **287**(5785): p. 869-71.

29. Walker, M., et al., *PCR-based gene disruption and recombinatory marker excision to produce modified industrial Saccharomyces cerevisiae without added sequences.* Journal of Microbiological Methods, 2005. **63**(2): p. 193-204.

30. Voth, W.P., et al., *Yeast vectors for integration at the HO locus.* Nucleic Acids Research, 2001. **29**(12).

31. Gietz, R.D. and R.A. Woods, *Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method*, in *Guide to Yeast Genetics and Molecular and Cell Biology, Pt B*. 2002, Academic Press Inc: San Diego. p. 87-96.

32. Herskowitz, I., *Life cycle of the budding yeast Saccharomyces cerevisiae.* Microbiological Reviews, 1988. **52**(4): p. 536-553.

33. Budde, C., et al., *Purification and characterization of recombinant protein acyltransferases.* Methods, 2006. **40**(2): p. 143-150.

34. Allegretto, E.A., et al., *Transactivation properties of retinoic acid and retinoid X receptors in mammalin cells and yeast: correlation with hormone binding and effects of metabolism* Journal of Biological Chemistry, 1993. **268**(35): p. 26625-26633.

35. Chang, H.C.J., D.F. Nathan, and S. Lindquist, *In vivo analysis of the Hsp90 cochaperone Sti1 (p60).* Molecular and Cellular Biology, 1997. **17**(1): p. 318-325.

36. Musti, A.M., et al., *Transcriptional mapping of two yeast genes coding for glyceraldehyde 3-phosphate dehydrogenase isolated by sequence homology with the chicken gene.* Gene, 1983. **25**(1): p. 133-143.

CHAPTER 4

ENGINEERING NUCLEAR RECEPTORS TOWARD ANTIBIOTICS: HUMAN ESTROGEN RECEPTOR ALPHA

4.1 Human Estrogen Receptor alpha

The estrogen receptor (ER) is a member of the nuclear receptor superfamily of ligand-dependent transcription factors [1-4]. The two ER isotypes, ERα and ERβ, each play different roles in metabolic processes [1-10]. Both human ERα (hERα) and human ERβ (hERβ) are expressed in the breast, bone, cardiovascular system, and brain, however hERα is mainly expressed in the breast, uterus, and bone and therefore is involved in mammary gland development and bone density [1, 2]. As transcription factors, these receptors are also involved in many diseases, playing significant roles in osteoporosis, cardiovascular disease and the development of breast cancer [1, 11, 12].

The wild-type ligand for hERα and hERβ is 17-β estradiol (E2), which mediates growth and differentiation. In addition to E2, hERα has been shown to activate transcription through a broad range of compounds, including the isoflavone genistein, resveratrol, bisphenol A, and the synthetic estrogens diethylstilbestrol (DES) and propylpyrazole (PPT) (Figure 4.1) [1, 12-17]. Both hERα and hERβ, however, have varying ligand specificities, despite the fact that these two estrogen receptor subtypes share common residues within their binding pockets [1, 14, 16, 18-21]. For example, E2 has a higher affinity for hERα whereas the phytoestrogen genistein preferentially activates hERβ over hERα [7, 14, 22, 23]. The ligand binding domains have 58 % amino acid sequence identity, but only two residues differ within the pocket [24]. L384 in hERα is a methionine in hERβ and M421 in hERα is an isoleucine in hERβ. Both receptors

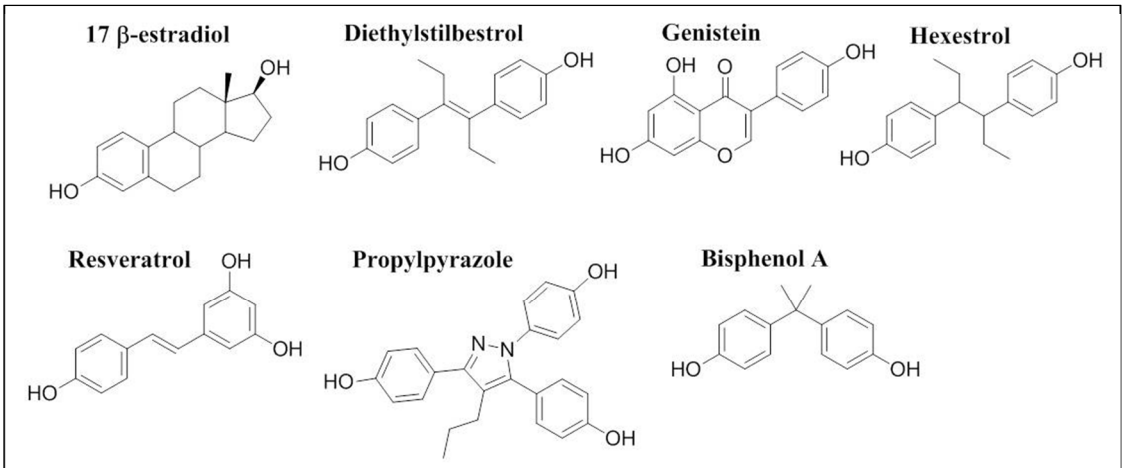


Figure 4.1. Various ligands for the human estrogen receptor alpha (hERα).

maintain a methionine in the pocket, suggesting the important role this amino acid plays in ligand activation of these two isotypes of the estrogen receptor [3]. In addition, hERa is known to localize in the cell membrane, leading to activation of different signaling pathways [25].

Some ligands modulate ER activity differently in varying tissues and are known as selective estrogen receptor modulators (SERMs) [1, 12, 26]. For example, 4-hydroxy tamoxifen (OHT), currently used to treat ER-positive breast cancer cells, acts as an antagonist in breast cancer cells, but as an agonist in bone, uterus, liver, and the cardiovascular system [1, 19, 27, 28]. The dual nature of this ligand adds a level of complexity for drug discovery and development.

Like other nuclear receptors, hERα consists of a DNA-binding domain (DBD) fused to a ligand-binding domain (LBD) through a hinge region [3, 4, 7, 29]. The LBD folds into an α-helical sandwich containing twelve α-helices, forming a pocket of approximately 450 $Å^3$ [3, 8, 30]. Upon ligand binding, the terminal helix, helix 12, is repositioned into an active conformation, making contacts with helices 3, 4/5, and 10/11, closing the ligand binding pocket (Figure 4.2) [3, 22, 31]. The ligand binding pocket (LBP) is formed from helices 3, 7, and 12, although residues from additional helices make contacts with the ligand [3, 29, 32, 33]. This conformational change allows coactivator proteins to bind to the LBD and recruit transcriptional machinery to turn on transcription [4, 34-39].

Ligand interaction with hERα occurs through specific hydrogen bonds as well as hydrophobic contacts among residues within the LBP [3, 8]. Three residues form the

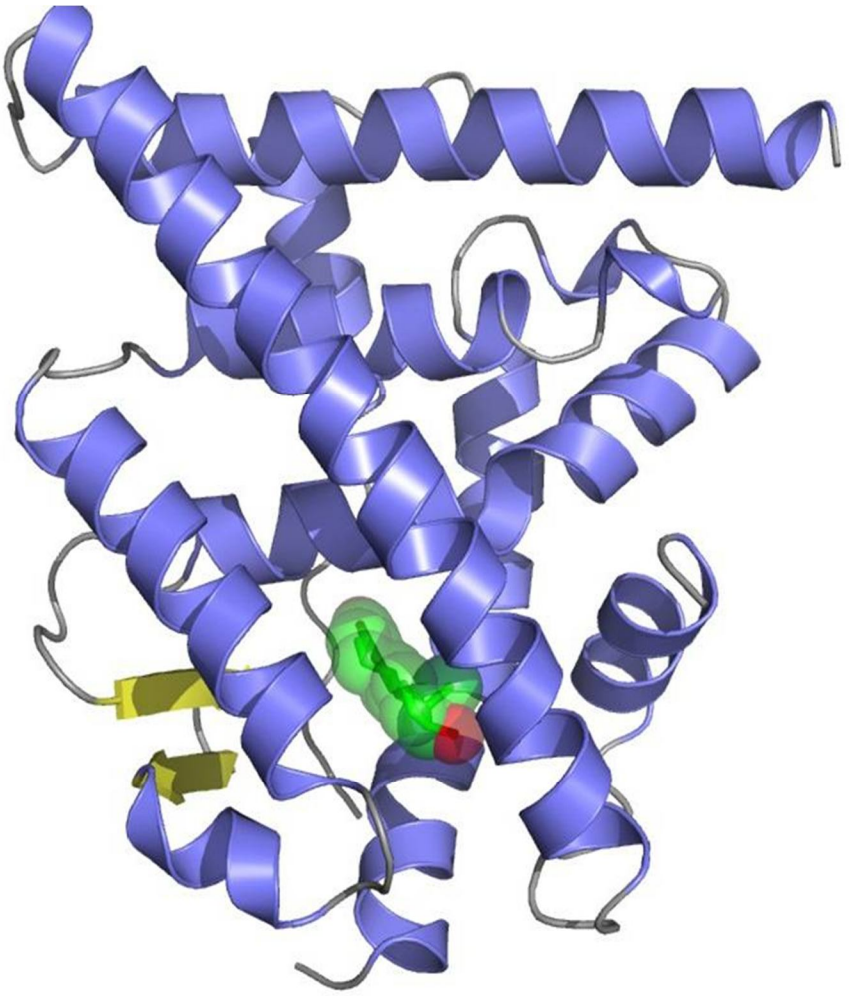


Figure 4.2. Crystal structure of the hERα LBD with 17 β-estradiol (E2) (PDB 1ERE). Alpha helices are shown in blue and the β-sheet is shown in yellow. The ligand is shown in green sticks with spheres.

hydrogen bond contacts between hERα and E2. A glutamate in helix 3, E353, an arginine in helix 6, R394, and a water molecule form a network of hydrogen bonds with the 3-hydroxy group of E2 at one end of the pocket. Histidine 524, located in helix 11 on the opposite side of the pocket, also forms a hydrogen bond with the 17-hydroxy group of E2 [3, 8, 10, 29, 31]. Additional residues in the pocket that surround the steroid ring system of E2 make conserved hydrophobic contacts with the ligand [8, 29]. One characteristic of hERα is the high sensitivity this receptor has toward E2, having a subnanomolar affinity for the wild-type ligand [11, 40]. Comparison of crystal structures of the LBD with a variety of ligands, mutational analysis, and modeling of hERα have identified important residues involved in ligand interaction [8, 11, 26, 29, 40-51]. These studies also demonstrate different residues responsible for ligand specificity and sensitivity [5, 47, 52-57]. Understanding the ligand-receptor interactions is essential for elucidating the mechanism of ER activation.

To continue toward the goal of engineering a nuclear receptor to bind and activate in response to the β-lactam antibiotics, the target nuclear receptor was switched from the pregnane X receptor (PXR) to hERα. A rational designed library of hERα was created and tested in chemical complementation (Figure 4.3). As mentioned previously, chemical complementation uses two fusion proteins in yeast expression plasmids: the yeast Gal4 DNA binding domain (GBD) fused to the hERα LBD, containing a tryptophan marker, and a plasmid containing a human coactivator fused to the Gal4 activation domain (GAD), containing a leucine marker (Figure 4.3). The coactivator in this system contains peptide motif repeats of LXXLL of the steroid receptor coactivator-1 (SRC-1). The full length SRC-1 coactivator is not necessary for activation to occur, therefore the smaller

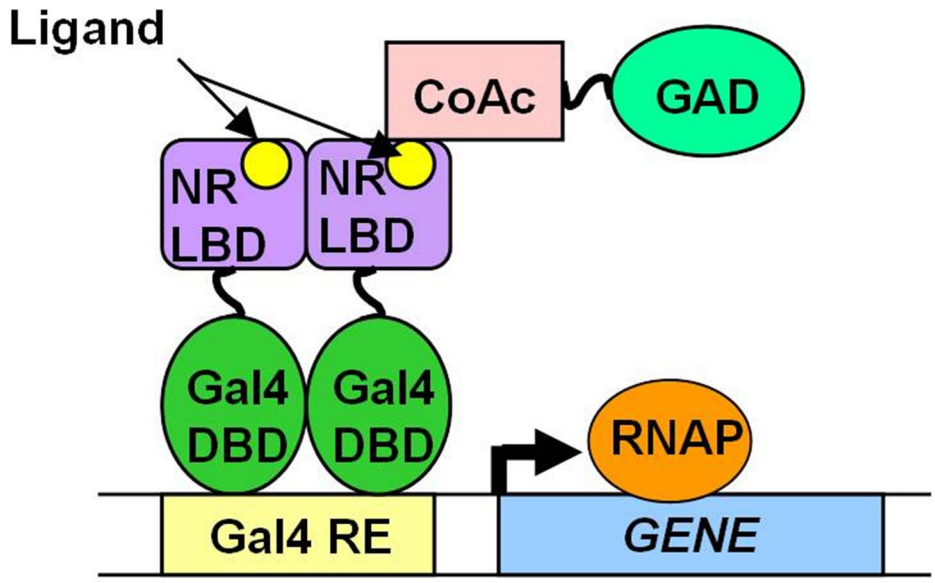


Figure 4.3. Chemical complementation selection system. The nuclear receptor (NR) ligand binding domain (LBD) fused to the Gal4 DNA-binding domain (DBD), is introduced into a yeast strain containing Gal4 response elements (Gal4REs) controlling expression of genetic selection genes. Upon ligand binding, a conformational change occurs; recruiting a fusion protein of the NR coactivator (CoAc) and the Gal4 activation domain (GAD) that binds to the LBD, leading to the initiation of transcription of a selection marker. Gene expression occurs and cells survive on media with ligand.

coactivator peptide was used [36]. In this work, chemical complementation is used to test a rationally designed hERα library, as well as to test random mutagenic libraries of hERα.

4.2 Human Estrogen Receptor Alpha Library for Antibiotics

4.2.1 Library Design

To engineer a hERα variant to activate transcription in response to the β-lactam antibiotics ampicillin, amoxicillin, penicillin G, oxacillin, cloxacillin, and nafcillin, a rational library was designed toward the β-lactam antibiotics. As performed with the library for the PXR in Chapter 2, residues in the ligand binding pocket were chosen for mutagenesis based on the known receptor-ligand interactions, as well as previously assessed knowledge about various residues in the LBP.

Hydrophobic residues, including M343, L346, L349, A350, L384, L387, M388, L391, F404, M421, I424, F425, L428, and G521, that make contacts with the ligand were considered for mutagenesis [7, 16, 26, 29, 46]. The three known hydrogen bonding residues E353, R394, and H524 were not considered for mutagenesis. Mutating all interacting residues would result in a large library size and a complete analysis of mutations would not be obtainable. Therefore, within this list of contacting residues, a subset was chosen based on proximity to the ligand as well as to have residues located in the different α-helices that form the LBP and surround the ligand. The complete list of residues chosen to mutate were L346, L384, M388, F404, M421, I424, and L428 (Figure 4.4). Leucine 346 is located on helix 3, L384 and M388 are located on helix 5, F404 is located on a β-strand, and M421, I424, and L428 are located on helix 7, thus providing residues from several surrounding α-helices as well as the β-sheet.

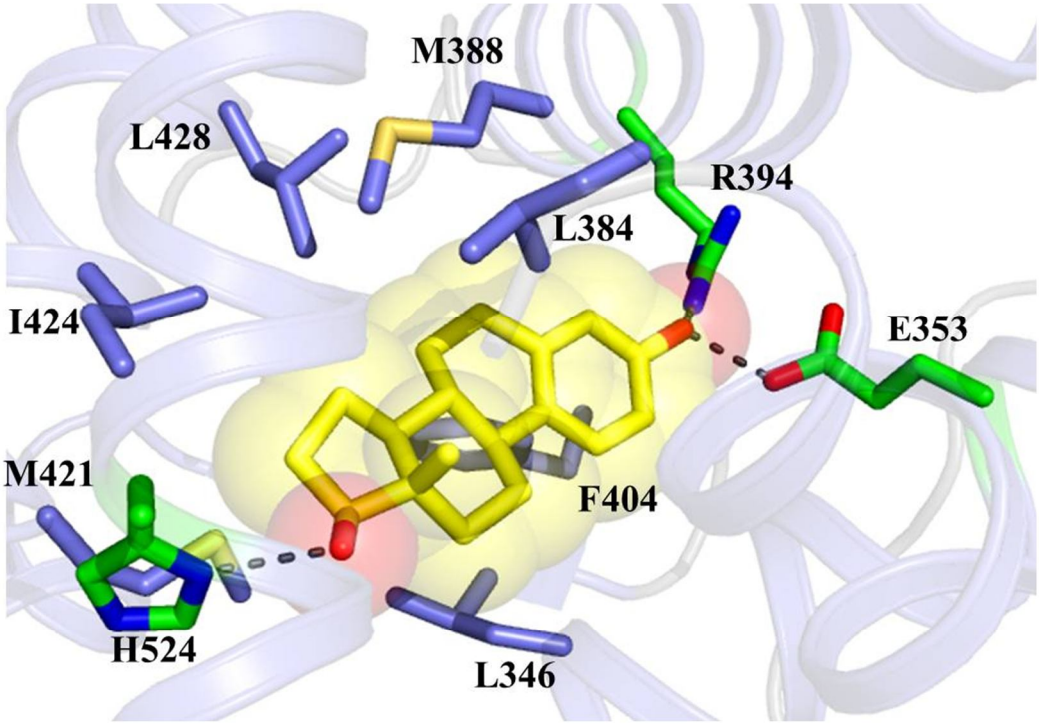


Figure 4.4. hERα library residues. Ligand binding pocket of the LBD with E2 showing residues selected for mutagenesis (PDB 1ERE). All residues chosen for mutagenesis in the library are shown in blue and the hydrogen bonding residues are shown in green. E2 is shown in yellow sticks with

The mutations that were substituted at each position were based on molecular modeling of the β-lactam antibiotics into the LBP of hERα. Modeling was performed with AutoDock Vina, using the crystal structure of the ligand binding domain of hERα with E2 (PDB 1ERE) [58]. As a control, E2 was modeled into the pocket (PDB 1ERE), producing a binding energy of -10.7 kcal/mol. When superimposed onto the crystal structure, the modeled E2 overlays in the same position as E2 in the crystal structure (Figure 4.5). Ampicillin, amoxicillin, penicillin G, and oxacillin were also modeled into the LBP of hERα using AutoDock Vina, producing binding energies of -7.5, -7.2, -8.5, and -8.1 kcal/mol, respectively (Figure 4.6). As seen in Figure 4.6, oxacillin is located in the opposite orientation as seen with ampicillin, amoxicillin, and penicillin G (Figure 4.6). This is not surprising since oxacillin has the additional isoxazole ring not present in the other antibiotics. Moreover, the second lowest binding energy for oxacillin was consistent with the position of the other antibiotics (Figure 4.6). These data provided a theoretical orientation of the lactam ring as well as the "side chain" of each antibiotic in the LBP of hERα, and therefore aided in determining distances and potential interactions between these groups and the residues in the library.

Figure 4.7A shows the amino acids chosen for mutagenesis in the hERα library (Figure 4.7A). Amino acid changes were chosen based on size and chemical characteristics of the substituted residue. Due to the polarity of the antibiotics, these residues were chosen to be mutated to amino acids with hydrogen bonding potentials, such as glutamine, asparagine, aspartic acid, glutamic acid, serine, histidine, and arginine (Figure 4.7B). In Figure 4.7B, the desired mutations at each position are indicated. The degeneracy of the genetic code also produces additional mutations at each residue (Figure

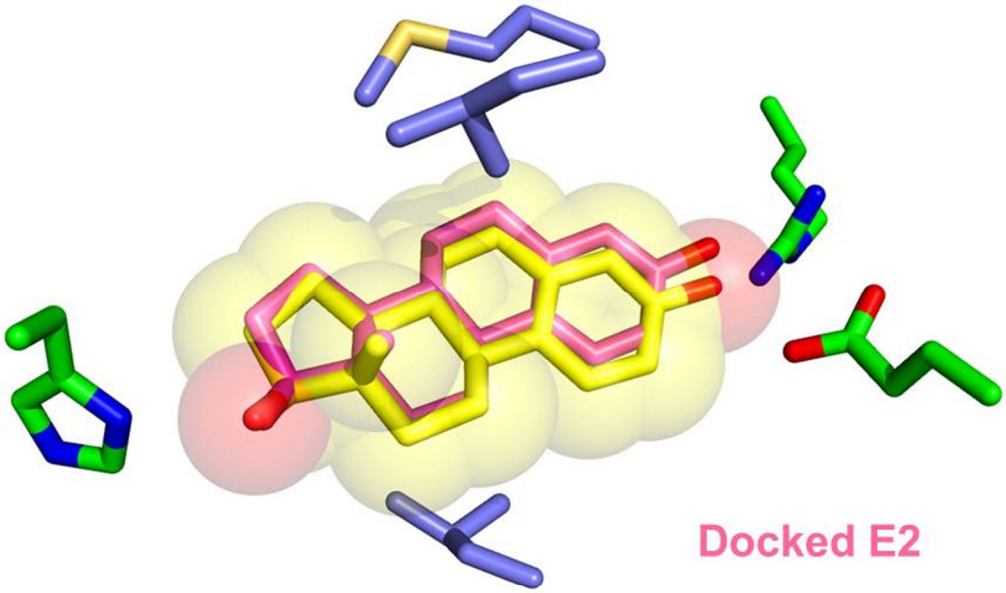


Figure 4.5 Modeling of E2 into the crystal structure of the hERα LBD with E2 (PDB 1ERE). All residues chosen for mutagenesis in the library are shown in blue and the hydrogen bonding residues are shown in green. E2 is shown in yellow sticks with spheres. The docked E2 is shown in pink sticks. For clarity only the three hydrogen bonding residues E353, R394, and H524 and the library residues L346, L384, and M388 are shown.

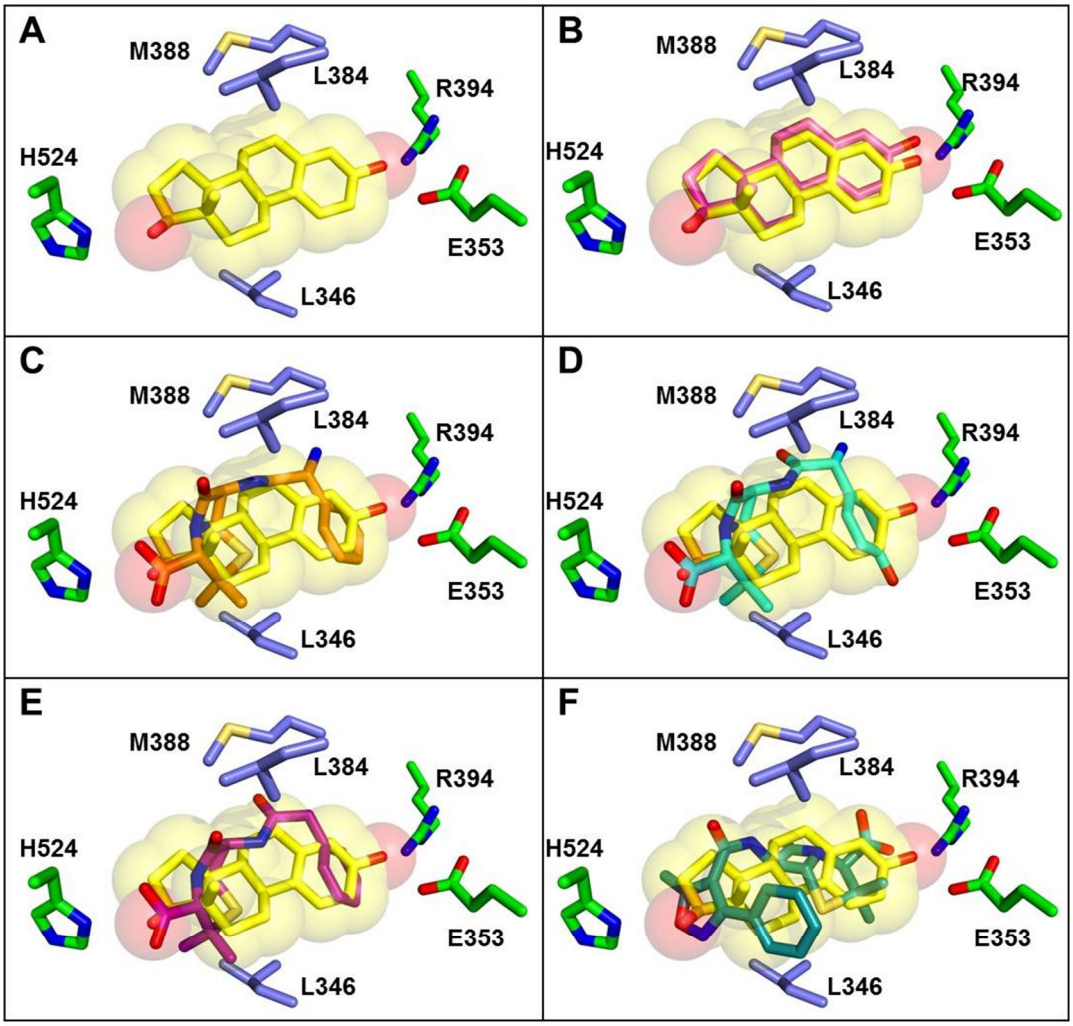
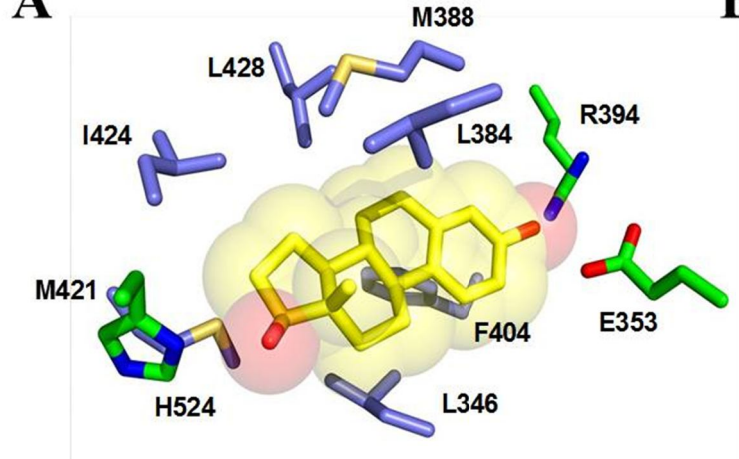


Figure 4.6. Modeling of hERα crystal structure with the β-lactam antibiotics (PDB 1ERE). (A) The wild-type crystal structure of hERα with E2. Modeling with (B) E2 shown in pink C) Ampicillin shown in orange, (D) Amoxicillin shown in cyan, (E) Penicillin G shown in purple, and (F) Oxacillin shown in dark green. The crystal structure coloring is the same as in Figure 4.5.

A



B

Residue	Desired Mutations	Complete Mutations
L346	LDENQS	LQHVEDMIKNRGS
L384	LDEQ	LQHVED
M388	MQ	LQMK
F404	FQE	LQHVEDFYstop
M421	MNY	LMIKNFYstop
I424	IQS	LQHMIKNRS
L428	LQHR	LQHR

Figure 4.7. hERα library design. (A) Ligand binding pocket of the LBD with E2 showing residues chosen for mutagenesis and the three hydrogen bonding residues E353, R394, and H524 (PDB 1ERE). (B) Residues chosen to mutate in the hERα library are shown. The desired mutations are the amino acids chosen in the rational design at each position. Due to the degeneracy of the genetic code, the desired mutations produce the complete list of amino acid mutations for each residue in the library. These mutations produce library sizes of 8.09×10^5 and 3.98×10^6 at the amino acid level and nucleotide level, respectively.

4.7B). All positions in the library, except M421, had a designed mutation that included glutamine due to the larger volume of this amino acid, having a volume of 143.8 $Å^3$, and to provide several hydrogen bonding atoms [59]. M421 was located near the lactam ring in the modeling results, and, more importantly, is one of the two residues in the pocket that differ between hERα and hERβ. M421, therefore, was only mutated to asparagine and tyrosine to introduce polar amino acids having the average or larger volumes of 114.1 $Å^3$ and 193.6 $Å^3$. As charged amino acids, aspartic acid and or glutamic acid were chosen as possible mutations for L346, L384, and F404, to also introduce hydrogen bonding potentials into the LBP. Therefore, at L346, aspartic acid, glutamic acid, asparagine, glutamine, and serine were introduced and at L384, an aspartic acid, glutamic acid, and glutamine were introduced (Figure 4.7B). A glutamic acid and glutamine were introduced at F404 (Figure 4.7B). Previous work by Chockalingam and coworkers, proposed that M388Q could donate a hydrogen bond with the synthetic ligand 4,4'-dihydroxybenzil (DHB) [45]. Therefore, M388 was chosen to be mutated to a glutamine. Serine was included at L346 and I424 to incorporate a smaller polar residue that could increase the volume of the pocket to accommodate the larger antibiotics. The combination of mutations at all the positions produces a theoretical library sizes of 8.09 x $10^5$and 3.98 x 10^6 at the amino acid level and nucleotide level, respectively (Figure 4.7B).

4.3 Yeast Transformations

4.3.1 Library Construction and Selection in Chemical Complementation

As performed in Chapter 2 with the PXR library, to evaluate the hERα library, chemical complementation was used. A background plasmid for hERα, designated pGBDhERαbackground, was made to remove background contamination from the wild-

type receptor in the library. Expression of the gene in the background plasmid produces three stop codons, eliminating contamination of the library with the wild-type hERα receptor. Next, to create the library of hERα variants, oligonucleotides containing the randomized mutations were combined using overlap extension PCR to generate a cassette with the designed mutations [60].

The library cassette containing the randomized codons was transformed into the yeast with the linearized background plasmid and, via homologous recombination, yeast cells containing a library of variants were plated onto adenine selective media containing ligand sets with the β-lactam antibiotics combined. The first ligand set contained the antibiotics amoxicillin, cloxacillin, and penicillin G. The second ligand set contained the antibiotics ampicillin, oxacillin, and nafcillin. The library of variants was also plated onto SC-LW nonselective media to determine whether homologous recombination occurred as well as to determine the library size and diversity of mutations in the library.

The transformation of the hERα library produced a library size of 2.94×10^3 colonies and a transformation efficiency of 5.88×10^4 colonies/μg vector cassette DNA. Eighty-five colonies from the adenine selective plates with ligands were streaked onto adenine selective plates without ligand (SC-ALW), adenine selective plates with E2, and nonselective plates (SC-LW). Most of the variants were constitutively active, meaning the cells show growth on the SC-ALW plate without the presence of the exogenous small molecule (Figure 4.8). Streaking results, however, also showed that approximately 10 % of the library was not constitutively active (Figure 4.8).

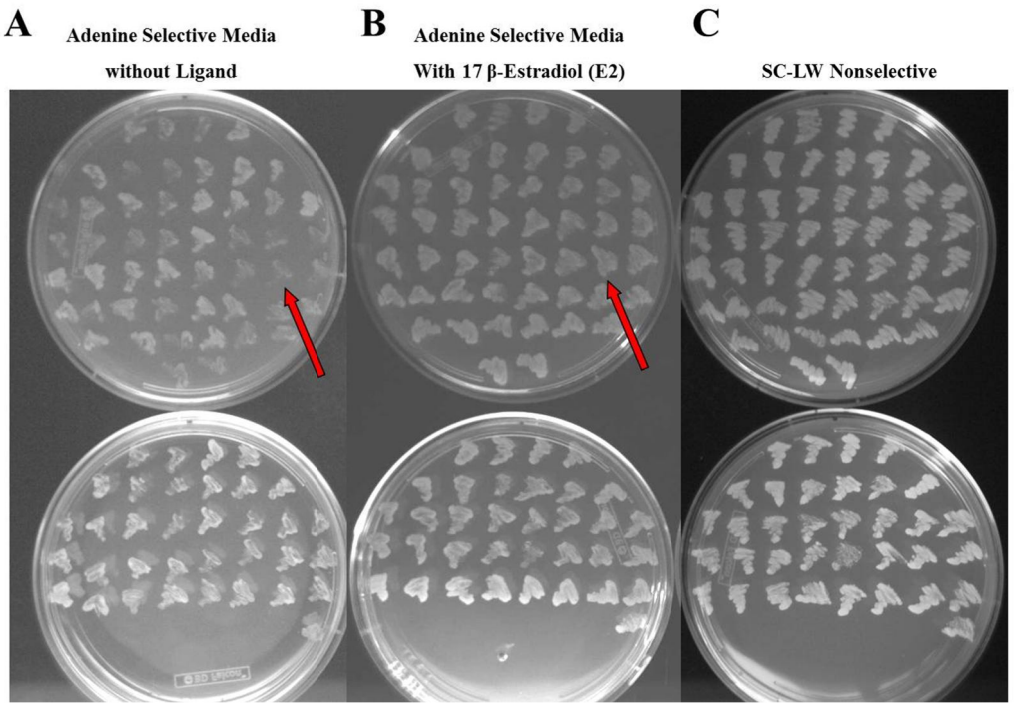


Figure 4.8 Streaking mutants from selective plates containing the antibiotics. (A) Adenine selective media without ligand. (B) Adenine selective media with E2. (C) A nonselective plate containing synthetic complete media lacking leucine and tryptophan (SC-LW), selecting for both fusion plasmids. The arrows indicate an example of a mutant that is not growing on the adenine selective plate and displays growth on the adenine selective plate with E2.

Table 4.1 Summary of hERα library sequencing

Variant	Library Mutations	Undesigned Mutations	Fold Activations
1	L346M		5.4
2	L346V		5.2
3	L346E		5.4
4	L346M; M388Q	S450F	5.4
5	L346M; L384L	D538G	5.6
6	L346I	D374E	5.6
7	L346N	R363R	4.6
8	L346L; L384L; M388Q		5.8
9	L384E; M388L		5.0
10	M388Q	D351E	5.0
11	L384Q	R363R	5.4
12	L384L; M388Q	H356L	5.5

Colonies from the hERα library that were sequenced showed that only three of the seven positions mutated in the library contained mutations (Table 4.1). These mutations occur at L346, L384, and M388, but not at F404, M421, I424, and L428, indicating that the first three positions are more tolerant to change than the other four positions. In addition, sequencing results indicate that seven out of the twelve variants contain mutations that were not in the library design (Table 4.1). The absence of mutation at the majority of the residues in the library produced poor overall sequence diversity in the library.

These data suggest, however, that targeting a large group of amino acids and introducing multiple changes at those positions may be too ambitious to discover many functional protein variants. In contrast, introducing fewer changes, such as a variant with one to three mutations, could produce more functional receptors. The rescued and sequenced hERα variants from the library transformation were retransformed into yeast and tested for functionality in a quantitative liquid growth assay with E2 as well as the antibiotics.

4.3.2 hERα Variants in Chemical Complementation

To confirm if the variants from this library are constitutively active, dead, or ligand activated, a growth assay in liquid media was performed. In this liquid assay, yeast containing the variants are grown in a 96-well plate in selective media without ligand as well as with increasing concentrations of ligand (Figure 4.9). If the ligand able to bind and activate the nuclear receptor fusion protein (GBD:hERαLBD), the coactivator fusion protein (SRC-1:GAD) is recruited and growth is observed in adenine or histidine selective media. The hERα variants from the library were tested in a dose response with

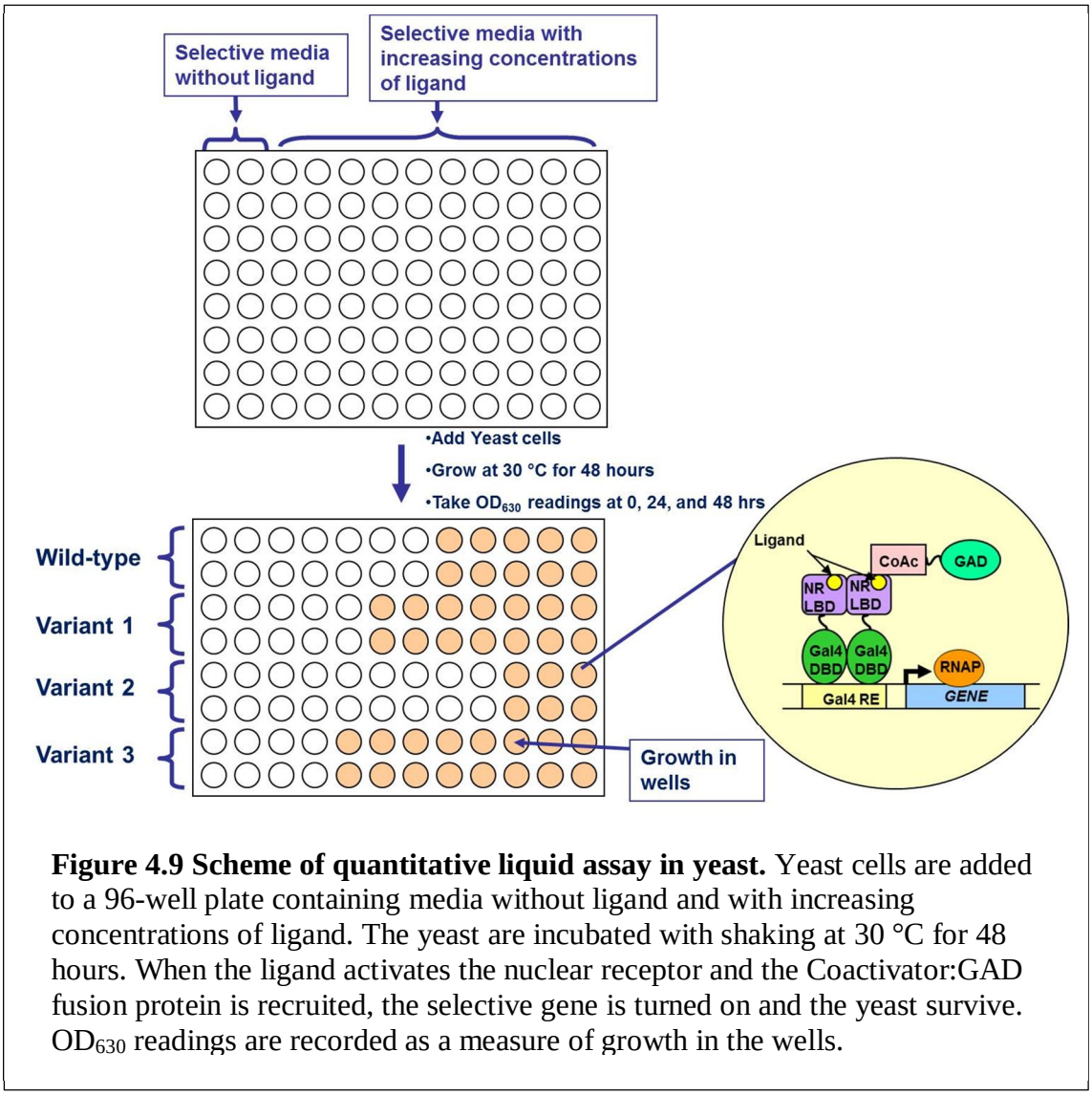


Figure 4.9 Scheme of quantitative liquid assay in yeast. Yeast cells are added to a 96-well plate containing media without ligand and with increasing concentrations of ligand. The yeast are incubated with shaking at 30 °C for 48 hours. When the ligand activates the nuclear receptor and the Coactivator:GAD fusion protein is recruited, the selective gene is turned on and the yeast survive. OD_{630} readings are recorded as a measure of growth in the wells.

E2 ranging from 10 nM to 10 μM in histidine selective media, containing synthetic complete media lacking histidine, leucine, and tryptophan with 10 mM 3-amino-1,2,4-triazole 3-AT (SC-HLW+3-AT). Due to the known leaky of expression of the *HIS3* gene, 3-AT was used to reduce basal background [61]. Variants were also tested with the target β-lactam antibiotic molecules. As performed in Chapter 2, Gal4, a ligand-independent transcription factor, was used as the positive control.

As expected, Gal4 displays growth in the histidine selective media with and without ligand (Figure 4.10, green circle). Seven of the twelve variants tested displayed the same growth trend as wild-type hERα (pink circle), although the variants were not tested with E2 concentrations below 10 nM, indicating that these variants may differ from wild-type hERα if tested at lower concentrations of E2 (Figure 4.10). All the variants display approximately five fold activation (Table 4.1).Variants 3, 7, and 9 display decreased sensitivity toward E2, having EC_{50} values, based on growth, at approximately 3 μM greater than 10 μM, and 1 μM, respectively (Figure 4.10, orange square, lavender diamond, and blue square, respectively).Variants 3 and 7 contain the mutation L346E and L346N, respectively (Table 4.1). In both variants, a hydrophobic amino acid, leucine, has been replaced with a polar amino acid. Variant 9 contains mutations L384E and M388L, indicating that the introduction of a charged amino acid into the pocket causes decreased sensitivity toward the wild-type ligand E2. When tested in a liquid assay with the antibiotics, none of the variants displayed ligand activated growth (data not shown). Therefore, to continue pursuing the goal of discovering a variant that is activated by the antibiotics, random mutagenesis approach was taken. Random mutagenesis provides the

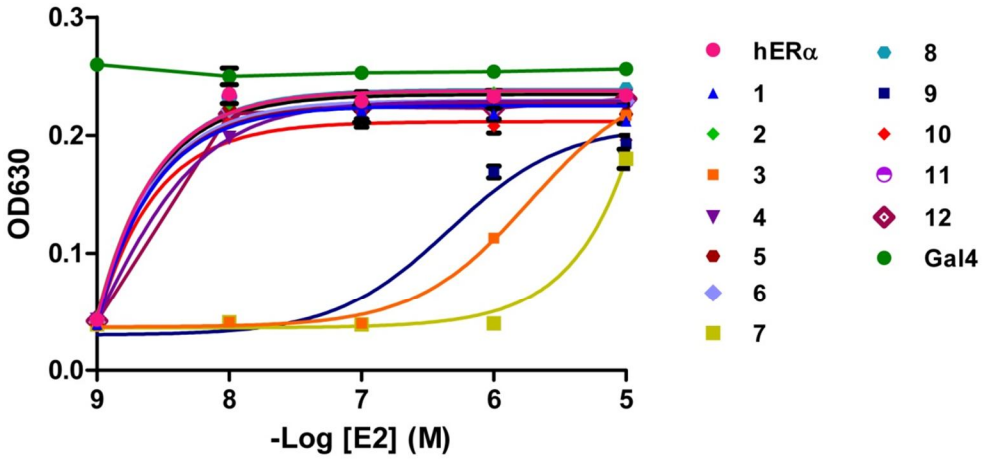


Figure 4.10 Growth profiles in yeast cells with E2 for wild type hERα variants obtained from designed library in histidine selective media containing 10 mM 3-AT. Sequencing for each variant is found in Table 4.1

possibility of many different mutations because amino acids outside of the LBP can be mutated, allowing for more diverse variants.

4.3.3 Random Mutagenesis on hERα

Random mutagenesis methods have been developed and are currently used to create and discover protein variants with novel functions. Several methods of random mutagenesis, such as gene shuffling, provides the ability to produce variants by combining related DNA sequences via PCR and allowing crossovers from complementary ends of the similar sequences to produce novel DNA sequences [62]. In error prone PCR, another common random mutagenesis method, mutations are introduced into the DNA template through amplification performed by *Taq* polymerase, which lacks proof-reading capabilities, generating mutations in the product DNA through multiple rounds of amplification [63]. Error prone PCR has previously been used to engineer hERα variants activated by the novel synthetic molecule, 4,4'-dihydroxybenzil (DHB) [45].

Error-prone PCR using manganese chloride ($MnCl_2$) concentrations of 20 μM, 100 μM, and 200 μM was performed on wild-type hERα, with the hypothesis that higher concentrations of $MnCl_2$ cause higher mutational rates. Manganese can be added to the error-prone PCR reaction mixture to increase the error rate and thereby increasing the overall number of mutations [63]. A greater diversity of variants can be discovered through having libraries that introduce different numbers of mutations. The error-prone PCR reaction generated cassettes with errors introduced into the hERα gene as well as a cassette with ends that are complementary to the background plasmid. The cassettes produced from the PCRs were transformed alongside the linearized background vector,

Table 4.2 Summary of sequencing of variants from error prone PCR on hERα.

Variant	Receptor Template	Concentration of $MnCl_2$ (μM)	Mutations	Fold Activations
13	Wild-type hERα	20	L454F	CA
14	Wild-type hERα	20	M342I	5.8
15	Wild-type hERα	20	L462Q	CA
16	Wild-type hERα	20	L462P	CA
17	Wild-type hERα	100	P406P, G442R	NA
18	Wild-type hERα	100	V376A	CA
19	Wild-type hERα	40*	E336K, H373Y, M438T, A505S, I510F, R515R, H547Y	6.4
20	Wild-type hERα	40*	L346Q,V355F	7.0
21	Wild-type hERα	40*	L384P	CA
22	Wild-type hERα	40*	V478L, I487V	CA
23	Wild-type hERα	40*	T431I, M438V, S527S	7.7
24	Wild-type hERα	40*	L466Q, L536P	CA
25	Wild-type hERα	40*	N359S, N407D	4.2
26	L346E	40*	L308S, A546A	CA
27	L346E	40*	N348K, G400G	7.0
28	L384E; M388L	40*	L454F, S464G, C530R	1.6

* The dNTPs for this transformation contained a biased mixture of 45% T, 25 % A, 15 % G, and 15 % C.
CA: constitutively active
NA: not available

and cells were plated onto adenine selective plates containing the same antibiotic sets as used for the rational hERα library. Colonies were rescued and sequenced as before.

Table 4.2 shows sequencing for variants, numbered 13-18, which were produced from the transformation of the error-prone PCR libraries with 20 μM, 100 μM, and 200 μM $MnCl_2$ (Table 4.2). Table 4.2 also shows which concentration of $MnCl_2$ was used to obtain that variant. Overall, these variants have one or two mutations located outside the binding pocket. Interestingly, variants 15 and 16, containing the mutations L462Q and L462P, respectively, have a mutation at the same position (Table 4.2). Overall, none of the variants rescued from the error prone library using 200 μM $MnCl_2$ provided clean and reliable sequencing, and consequently this concentration of $MnCl_2$ was determined to be too high to produce functional hERα variants. Variants 13-18 all contain one mutation, except variant 17, which has the silent mutation P406P in addition to the G442R mutation, indicating that 20 μM $MnCl_2$ or 100 μM $MnCl_2$ does not produce the expected lower or higher number of mutations.

Although the sequencing from the first set of error-prone libraries produced mutational diversity, additional methods to vary the sequence space when performing error prone PCR can be performed to provide different mutational options. One method is to vary the concentrations of the dNTPs in the dNTP mixture for the PCR reaction. Analysis of the nucleotide changes in previous hERα variants produced from a PCR reaction with equal concentration of dNTPs showed higher substitutions to cytosine and guanine and many to adenine, with few changes to thymine. Therefore, an additional error-prone PCR was performed on wild-type hERα using 40 μM $MnCl_2$ and a dNTP mixture containing different concentrations of the dNTPs, specifically, 45 % thymine, 25

% adenine, 15 % guanine, and 15 % cytosine. This library was transformed and colonies were rescued and sequenced. Variants 19-25 in Table 4.2 are from the error-prone PCR with the biased dNTP mixture (Table 4.2). Many of the variants contain two to three mutations, located throughout the protein as seen in the previous error prone libraries (Table 4.2). Conversely, variant 19 has seven mutations, one of which is a silent mutation. Interestingly, variants 20 and 21 have a mutation at L346 and L384, respectively, both positions that were chosen to mutate in the rationally designed library.

Analysis of the nucleotide substitutions for variants 19-28 shows more mutations to thymine, which is not surprising since the dNTP mixture included 45 % thymine. Mutation to adenine, guanine, and cytosine were equal among the variants, indicating a preference for substitutions to guanine and cytosine, especially since the dNTP mixture contained the lowest concentrations of guanine and cytosine. In summary, the error-prone PCR reaction with varying concentrations of dNTPs produced mutational diversity and varying numbers of mutations within each variant.

Both the rational and random mutagenesis libraries did not produce a variant able to activate transcription in response to the β-lactam antibiotics. Therefore, the next approach was to determine if the variants that displayed different activation profiles with E2 than wild-type hERα could be used as a starting template to generate a variant activated by the antibiotics. As seen in Figure 4.10, variants 3 (L346E) and 9 (L384E;M388L) displayed decreased sensitivity toward E2, having EC_{50} values of 3 μM and 1 μM, respectively, whereas wild type has an EC_{50} with E2 at 300 pM (Figure 4.10). While these variants have some activation toward E2, they are significantly less sensitive to E2 than wild-type hERα. Therefore, error-prone PCRs using 40 μM $MnCl_2$ and a

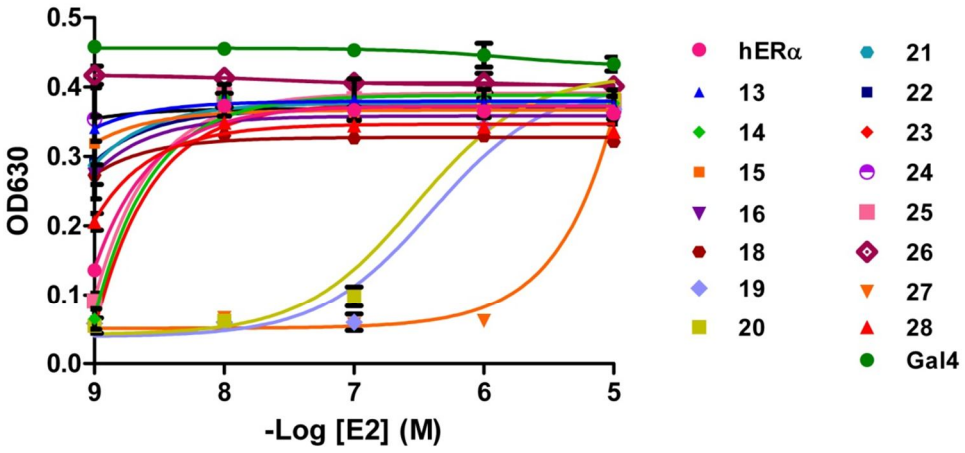


Figure 4.11 Growth profiles in yeast cells with E2 for wild type hERα variants obtained from random mutagenesis in histidine selective media containing 10 mM 3-AT. Sequencing for each variant is found in Table 4.2.

dNTP mixture containing 45 % thymine, 25 % adenine, 15 % guanine, and 15 % cytosine, as done with wild-type hERα, were performed on variants 3 and 9. The libraries were transformed into yeast and plated onto selective media containing the antibiotics. Variants 26-28 have two or three additional mutations to the template variant. Interestingly, variant 28, produced from the error prone PCR with 40 μM $MnCl_2$ using variant 9 (L384E;M388L) as the template, displayed decreased E2 sensitivity. Variant 13 has the mutation L454F, however this variant was produced from the error prone PCR with 20 μM $MnCl_2$ and wild-type hERα as the template.

All the variants obtained from the random mutagenesis on hERα and variants 3 and 9 were tested in the liquid quantitation growth assay in yeast. Figure 4.11 shows a dose response with E2 concentrations ranging from 10 nM to 10 μM (Figure 4.11). As previously mentioned, this analysis is to determine if the variants are ligand activated or constitutively active. This assay does not include E2 concentrations low enough to compare the variants to the 300 pM EC_{50} value for wild-type hERα. More significant variants can then be tested at lower concentrations E2 as well as with the β-lactam antibiotics. Half of the variants tested were constitutively active and the remainder had fold activations ranging from approximately 1.6 to 7.7-fold (Table 4.2). Some variants displayed the same growth levels as wild-type hERα, showing growth at 10 nM (Figure 4.11). Many variants have very high basal levels in the media without ligand and were therefore considered constitutively active (Table 4.2).

Three variants, 19, 20, and 27, show decreased sensitivity toward E2 (Figure 4.11). The decreased sensitivity for variant 19 may be due to the significantly higher number of mutations in this variant (Table 4.2). Variant 20 has one mutation in the

binding pocket, which could be one reason for the influence on activity toward E2. Interestingly, variant 26 was made from the template containing the mutation L346E. As previously mentioned, L346E was chosen for random mutagenesis because this variant had significantly reduced sensitivity toward E2. The mutations introduced from the error-prone PCR, specifically L308S and A546A, caused variant 26 to become constitutively active (Table 4.2 and Figure 4.11). The random mutagenesis libraries provide multiple hERα variants containing a wide variety of mutations, however when tested in a liquid assay with the antibiotics, none of the variants display ligand activated growth (data not shown).

4.4 Summary and Future Work

To engineer a nuclear receptor to be activated by the β-lactam antibiotics, a library of hERα variants using rational design based on molecular modeling analysis with the β-lactam antibiotics was created and tested in chemical complementation. No variants were discovered that were activated by the antibiotics. Random mutagenesis, therefore, was performed on hERα, producing functional variants with diverse mutations, however, as with the rational library, none of the variants showed ligand activated growth with the antibiotics.

The rational library produced three variants that contain a point mutation at residue L346: L346M, L346V, and L346E. When tested in the liquid quantitation assay in yeast, the variant L346E displayed decreased sensitivity to E2. L346M, however, maintained wild type activity with E2. Structurally, L346 is located on helix 3, one of the helices involved in forming the LBP and therefore influences ligand binding (Figure 4.12) [5, 8]. Multiple crystal structures of the hERα LBD also show that L346 makes

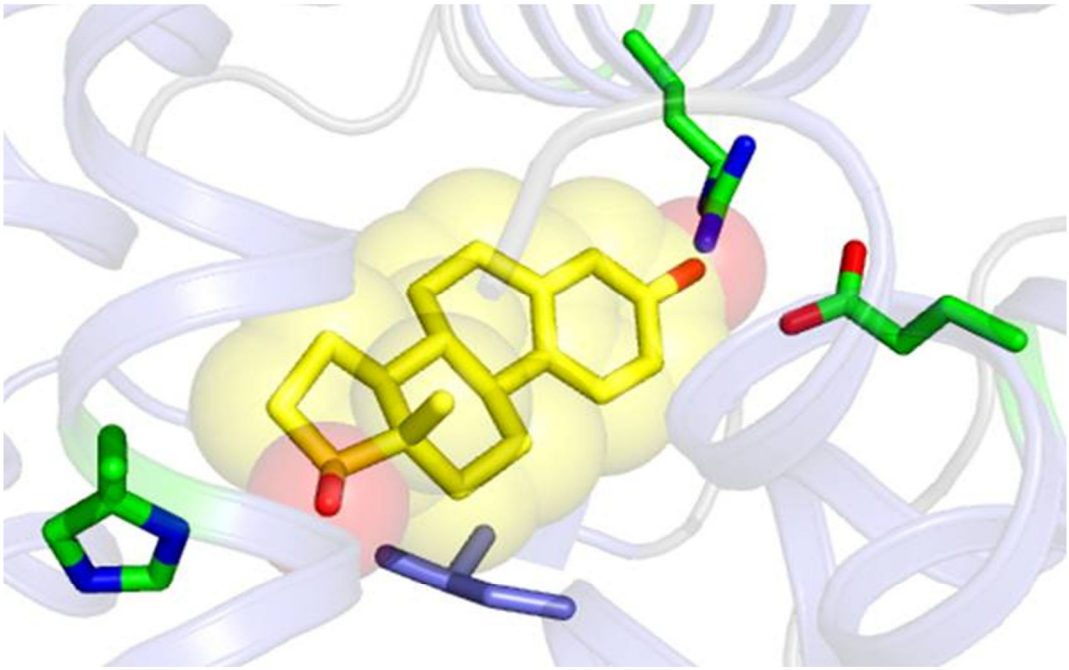


Figure 4.12. Ligand binding pocket of the crystal structure of the hERα LBD bound to 17 β-estradiol (E2) showing residue L346 (PDB 1ERE). L346 is shown in blue and the hydrogen bonding residues are shown in green. E2 is shown in yellow sticks with spheres.

hydrophobic contacts with the ligand in the pocket [26, 29, 46, 48, 64]. Due to the marked differences in activity toward E2, this position was saturated to investigate the role this residue has in ligand binding and activity. The characterization of position 346 will be discussed in the next chapter.

4.5 Materials and methods

4.5.1 Plasmids and Primers

The plasmid pGBDhERαLBD was previously cloned [65]. Primers containing a *NheI* site at the 5' end and *SpeI* site at the 3' end were designed to amplify the ERαLBD fusion gene (containing residues 301-595 of ERα) from the pSG5-HEGO vector. The gene was ligated into the pGBDRXR vector using the *NheI* and *SpeI* restriction enzymes.

The yeast expression vector containing GAD fused to a peptide of the SRC-1 coactivator, designated pGAD10BALXXLLSRC-1, was previously made [65]. Primers containing a *BglII* or a *NotI* site were used to amplify the SRC-1 fragment (residues 594-821 of SRC-1) containing three LXXLL motif repeats from the pGAD10BASRC-1 vector constructed previously [66]. The yeast expression vector pGAD10BAACTR was digested with restriction enzymes *BglII* and *NotI* and ligated with the SRC-1 fragment with the LXXLL motifs to generate the pGAD10BALXXLLSRC-1 plasmid [65, 66].

To make the pGBD hERαLBD background plasmid, QuikChange™ (Stratagene) site-directed mutagenesis was used to insert *SacII* and *KpnI* sites into the hERα gene in pGBD hERαLBD, producing the pGBD hERαLBDSacIIKpnI plasmid. The plasmid was digested with *SacII* and *KpnI*. A 145 bp random sequence of DNA was amplified from the plasmid pMSCVeGFP. The random DNA sequence was digested with *SacII* and *KpnI* and ligated into the pGBD hERαSacIIKpnI plasmid between the two restriction sites,

generating three STOP codons. The resulting plasmid was confirmed by sequencing and named pGBDhERαbackground (Operon, Huntsville, AL).

To make pCMXGBDhERαLBD, the GBDhERαLBD fusion was amplified from pGBDhERαLBD using PCR with the following primers: GBDSacIIFor 5' tccccgcggatgaagctactgtcttctatcgaacaag 3' and NotIrev 5' tcagactggcagggaaacccgccggcgaaaaggaaaa3'. The plasmid pCMXrERβ as well as the GBDhERαLBD cassette were digested with *SacII* and *NotI*, and ligated to create the pCMXGBDhERαLBD plasmid. The plasmid was confirmed by sequencing (Operon, Huntsville, AL).

4.5.2 Ligands

17-β estradiol (MW= 272.4 g/mol) was purchased from MP Biomedicals (Aurora, OH). Amoxicillin (MW=365.41 g/mol) and Oxacillin (MW = 423.40 g/mol) were purchased from Fluka (Milwaukee, MI). Ampicillin (MW=371.4), Penicillin G (MW=372.48 g/mol), Cloxacillin (MW=475.9 g/mol), and Nafcillin (MW=454.5 g/mol) were purchased from Sigma (St. Louis, MO). 10 mM stocks of each ligand were dissolved in 80% ethanol:20% DMSO and stored at 4°C.

4.5.3 Creating Designed hERα library

Oligonucleotides were ordered from Operon (Huntsville, AL) with overlapping ends. To create the cassette containing the randomized codons, overlap extension PCR was utilized. Through PCR, the fragments were amplified in a PCR reaction containing 100 ng of each oligo, 125 ng of each primer, 100 mM dNTPs, and 1 μL *Pfu* polymerase. The PCR program was 95 °C for 2 minutes, 95 °C for 1 minute, 55 °C for 1 minute, 72 °C for 1 minute, repeated 30 times followed by 72 °C for 2 minutes. The resulting

cassette was 896 bp and was gel purified using the Zymo Gel DNA Recovery Kit (Zymo Research, Orange, CA).

The yeast transformation was performed using the lithium acetate transformation method and as previously described [67, 68].

4.5.4 Error-Prone PCR

The plasmid pGBDhERαLBD was used as the template DNA for error-prone PCR random mutagenesis using ERPCR1For 5' ggtcaaagacagttgactgtatcgc 3' and ERPCR8Rev 5' gcctcccccgtgatgtaatact 3' primers. The PCR reaction contained 0.5 μM of each primer, 500 μM dNPTs, 7 nM $MgCl_2$, 250 ng of template DNA, 1x *Taq* buffer, and *Taq* polymerase (Promega), and varied concentrations of $MnCl_2$. The PCR program was 95 °C for 1 minute, 95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 1 minute, repeated 20 times followed by 72 °C for 2 minutes. The resulting cassette was gel purified using the Zymo Gel DNA Recovery Kit (Zymo Research, Orange, CA).

4.5.5 Yeast Transformation

The yeast transformation was done using the lithium acetate method as performed in Chapter 2 and as previously described [67].

4.5.6 Liquid Yeast Growth Assay

The yeast growth assay was performed as previously described [67]. Briefly, transformants are grown overnight in nonselective media containing synthetic complete media lacking leucine and trypophan (SC-LW) at 30 °C with shaking at 300 rpm. Yeast are added to 96-well plates containing selective media without ligand and with increasing concentrations of ligand in a 4:1 ratio of media to yeast. Plates are incubated for 48 hours

at 30 °C with shaking at 170 rpm. Optical density readings at 630 nm (OD630) are taken at 0, 24, and 48 hours as a measure of growth.

4.5.7 Modeling hERα using Autodock Vina

The protein was prepared using AutoDockTools to add polar hydrogens and Gasteiger charges to set the partial charge property of each atom. The ligands were prepared using ChemBioDraw Ultra 11.0 and ChemBio3D Ultra 11.0 (Cambridge Soft, USA) and minimized using MMFF94 force fields [69]. Ligands were modified in AutoDockTools by adding Gasteiger charges to set the partial charge property of each atom. AutoDock Vina was used to perform the docking simulations using the default parameters [58].

4.6 References

1. Bai, Z.L. and R. Gust, *Breast Cancer, Estrogen Receptor and Ligands.* Archiv Der Pharmazie, 2009. **342**(3): p. 133-149.

2. Hall, J.M. and D.P. McDonnell, *Coregulators in nuclear estrogen receptor action - From concept to therapeutic targeting.* Molecular Interventions, 2005. **5**(6): p. 343-357.

3. Huang, P.X., V. Chandra, and F. Rastinejad, *Structural Overview of the Nuclear Receptor Superfamily: Insights into Physiology and Therapeutics.* Annual Review of Physiology, 2010. **72**: p. 247-272.

4. Stanisic, V., D. Lonard, and B. O'Malley, *Modulation of Steroid Hormone Receptor Activity.* Progress in Brain Research, 2010. **181**: p. 153-176.

5. Ekena, K., J.A. Katzenellenbogen, and B.S. Katzenellenbogen, *Determinants of ligand specificity of estrogen receptor-alpha: Estrogen versus androgen discrimination.* Journal of Biological Chemistry, 1998. **273**(2): p. 693-699.

6. Katzenellenbogen, B.S. and J.A. Katzenellenbogen, *Estrogen receptor transcription and transactivation Estrogen receptor alpha and estrogen receptor*

beta: regulation by selective estrogen receptor modulators and importance in breast cancer. Breast Cancer Research, 2000. **2**(5): p. 335-344.

7. Pike, A.C.W., A.M. Brzozowski, and R.E. Hubbard, *A structural biologist's view of the oestrogen receptor.* Journal of Steroid Biochemistry & Molecular Biology, 2000. **74**: p. 261-268.

8. Pike, A.C.W., *Lessons learnt from structural studies of the oestrogen receptor.* Best Practice & Research Clinical Endocrinology & Metabolism, 2006. **20**(1): p. 1-14.

9. Edwards, D.P., *Regulation of signal transduction pathways by estrogen and progesterone.* Annual Review of Physiology, 2005. **67**: p. 335-376.

10. Mueller-Fahrnow, A. and U. Egner, *Ligand-binding domain of estrogen receptors.* Current Opinion in Biotechnology, 1999. **10**(6): p. 550-556.

11. Chen, Z.L., et al., *Directed evolution of human estrogen receptor variants with significantly enhanced androgen specificity and affinity.* Journal of Biological Chemistry, 2004. **279**(32): p. 33855-33864.

12. Dai, S.Y., et al., *Prediction of the tissue-specificity of selective estrogen receptor modulators by using a single biochemical method.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(20): p. 7171-7176.

13. Do Minh, S., et al., *Novel mammalian cell lines expressing reporter genes for the detection of environmental chemicals activating endogenous aryl hydrocarbon receptors (ArhR) or estrogen receptors (ER).* Toxicology in Vitro, 2008. **22**(8): p. 1935-1947.

14. Harris, H.A., J.A. Katzenellenbogen, and B.S. Katzenellenbogen, *Characterization of the biological roles of the estrogen receptors, ER alpha and ER beta, in estrogen target tissues in vivo through the use of an ER alpha-selective ligand.* Endocrinology, 2002. **143**(11): p. 4172-4177.

15. Levenson, A.S., et al., *Resveratrol acts as an estrogen receptor (ER) agonist in breast cancer cells stably transfected with ER alpha.* International Journal of Cancer, 2003. **104**(5): p. 587-596.

16. Manas, E.S., et al., *Understanding the selectivity of genistein for human estrogen receptor-beta using X-ray crystallography and computational methods.* Structure, 2004. **12**(12): p. 2197-2207.

17. Celik, L., J.D.D. Lund, and B. Schiott, *Exploring Interactions of Endocrine-Disrupting Compounds with Different Conformations of the Human Estrogen Receptor alpha Ligand Binding Domain: A Molecular Docking Study.* Chemical Research in Toxicology, 2008. **21**(11): p. 2195-2206.

18. Steinmetz, A.C.U., J.P. Renaud, and D. Moras, *Binding of ligands and activation of transcription by nuclear receptors.* Annual Review of Biophysics and Biomolecular Structure, 2001. **30**: p. 329-359.

19. Gross, J.M. and D. Yee, *How does the estrogen receptor work?* Breast Cancer Research, 2002. **4**(2): p. 62-64.

20. Martin-Santamaria, S., et al., *New scaffolds for the design of selective estrogen receptor modulators.* Organic & Biomolecular Chemistry, 2008. **6**(19): p. 3486-3496.

21. Shen, J., et al., *Discovery of potent ligands for estrogen receptor beta by structure-based virtual screening.* J Med Chem, 2010. **53**(14): p. 5361-5.

22. Pike, A.C.W., et al., *Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist.* Embo Journal, 1999. **18**(17): p. 4608-4618.

23. Kostelac, D., G. Rechkemmer, and K. Briviba, *Phytoestrogens modulate binding response of estrogen receptors alpha and beta to the estrogen response element.* Journal of Agricultural and Food Chemistry, 2003. **51**(26): p. 7632-7635.

24. Zeng, J., et al., *Insights into ligand selectivity in estrogen receptor isoforms: Molecular dynamics simulations and binding free energy calculations.* Journal of Physical Chemistry B, 2008. **112**(9): p. 2719-2726.

25. Barone, I., L. Brusco, and S.A. Fuqua, *Estrogen receptor mutations and changes in downstream gene expression and signaling.* Clin Cancer Res, 2010. **16**(10): p. 2702-8.

26. Shiau, A.K., et al., *The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen.* Cell, 1998. **95**(7): p. 927-937.

27. Salum, L.D., I. Polikarpov, and A.D. Andricopulo, *Structural and chemical basis for enhanced affinity and potency for a large series of estrogen receptor ligands: 2D and 3D QSAR studies.* Journal of Molecular Graphics & Modelling, 2007. **26**(2): p. 434-442.

28. Mukherjee S, M.D., *Computational molecular docking assessment of hormone receptor adjuvant drugs: Breast cancer as an example.* Pathophysiology, 2009. **16**(1): p. 19-29.

29. Brzozowski, A.M., et al., *Molecular basis of agonism and antagonism in the oestrogen receptor.* Nature, 1997. **389**(6652): p. 753-758.

30. Chambon, P., *The nuclear receptor superfamily: A personal retrospect on the first two decades.* Molecular Endocrinology, 2005. **19**(6): p. 1418-1428.

31. Gronemeyer, H., J.A. Gustafsson, and V. Laudet, *Principles for modulation of the nuclear receptor superfamily.* Nature Reviews Drug Discovery, 2004. **3**(11): p. 950-964.

32. Moore, J.T., J.L. Collins, and K.H. Pearce, *The nuclear receptor superfamily and drug discovery.* ChemMedChem, 2006. **1**(5): p. 504.

33. Folkertsma, S., et al., *A family-based approach reveals the function of residues in the nuclear receptor ligand-binding domain.* Journal of Molecular Biology, 2004. **341**(2): p. 321-335.

34. Azizi, B., E.I. Chang, and D.F. Doyle, *Chemical complementation: small-molecule-based genetic selection in yeast.* Biochemical and Biophysical Research Communications, 2003. **306**(3): p. 774-780.

35. Leo, C. and J.D. Chen, *The SRC family of nuclear receptor coactivators.* Gene, 2000. **245**(1): p. 1-11.

36. McInerney, E.M., et al., *Determinants of coactivator LXXLL motif specificity in nuclear receptor transcriptional activation.* Genes & Development, 1998. **12**(21): p. 3357-3368.

37. McKenna, N.J. and B.W. O'Malley, *Combinatorial control of gene expression by nuclear receptors and coregulators.* Cell, 2002. **108**(4): p. 465-474.

38. Rosenfeld, M.G., V.V. Lunyak, and C.K. Glass, *Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptionaf response.* Genes & Development, 2006. **20**(11): p. 1405-1428.

39. Germain, P., et al., *Overview of nomenclature of nuclear receptors.* Pharmacological Reviews, 2006. **58**(4): p. 685-704.

40. Chen, Z.L. and H.M. Zhao, *Rapid creation of a novel protein function by in vitro coevolution.* Journal of Molecular Biology, 2005. **348**(5): p. 1273-1282.

41. Chen, Z.L. and H.M. Zhao, *A highly efficient and sensitive screening method for trans-activation activity of estrogen receptors.* Gene, 2003. **306**: p. 127-134.

42. Gallinari, P., et al., *A functionally orthogonal estrogen receptor-based transcription switch specifically induced by a nonsteroid synthetic ligand.* Chemistry & Biology, 2005. **12**(8): p. 883-893.

43. Herynk, M.H. and S.A.W. Fuqua, *Estrogen receptor mutations in human disease.* Endocrine Reviews, 2004. **25**(6): p. 869-898.

44. Islam, K.M.D., et al., *Directed evolution of estrogen receptor proteins with altered ligand-binding specificities.* Protein Engineering Design & Selection, 2009. **22**(1): p. 45-52.

45. Chockalingam, K., et al., *Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(16): p. 5691-5696.

46. Tanenbaum, D.M., et al., *Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains.* Proceedings of the National

Academy of Sciences of the United States of America, 1998. **95**(11): p. 5998-6003.

47. Zhao, C.Q., J. Abrams, and D.F. Skafar, *Targeted mutation of key residues at the start of helix 12 in the hER alpha ligand-binding domain identifies the role of hydrogen-bonding and hydrophobic interactions in the activity of the protein.* Journal of Steroid Biochemistry and Molecular Biology, 2006. **98**(1): p. 1-11.

48. D'Ursi, P., et al., *Modelling the interaction of steroid receptors with endocrine disrupting chemicals.* Bmc Bioinformatics, 2005. **6**: p. 8.

49. Aliau, S., et al., *Differential interactions of estrogens and antiestrogens at the 17 beta-hydroxyl or counterpart hydroxyl with histidine 524 of the human estrogen receptor alpha.* Biochemistry, 2002. **41**(25): p. 7979-7988.

50. McLachlan, M.J., et al., *Directed Evolution of Orthogonal Ligand Specificity in a Single Scaffold.* Angewandte Chemie-International Edition, 2009. **48**(42): p. 7783-7786.

51. Shadnia, H., J. Wright, and J. Anderson, *Interaction force diagrams: new insight into ligand-receptor binding.* Journal of Computer-Aided Molecular Design, 2009. **23**(3): p. 185-194.

52. Wrenn, C.K. and B.S. Katzenellenbogen, *Structure-function analysis of the hormone binding domain of the human estrogen receptor by region-specific mutagenesis and phenotypic screening in yeast.* Journal of Biological Chemistry, 1993. **268**(32): p. 24089-24098.

53. Tremblay, G.B., et al., *Ligand-independent activation of the estrogen receptors alpha and beta by mutations of a conserved tyrosine can be abolished by antiestrogens.* Cancer Research, 1998. **58**(5): p. 877-881.

54. Paulmurugan, R., et al., *A human estrogen receptor (ER)alpha mutation with differential responsiveness to nonsteroidal ligands: Novel approaches for studying mechanism of ER action.* Molecular Endocrinology, 2008. **22**(7): p. 1552-1564.

55. Zhong, L. and D.F. Skafar, *Mutations of tyrosine 537 in the human estrogen receptor-alpha selectively alter the receptor's affinity for estradiol and the kinetics of the interaction.* Biochemistry, 2002. **41**(13): p. 4209-4217.

56. Zhao, C.Q., et al., *Mutation of Leu-536 in human estrogen receptor-alpha alters the coupling between ligand binding, transcription activation, and receptor conformation.* Journal of Biological Chemistry, 2003. **278**(29): p. 27278-27286.

57. Gangloff, M., et al., *Crystal structure of a mutant hER alpha ligand-binding domain reveals key structural features for the mechanism of partial agonism.* Journal of Biological Chemistry, 2001. **276**(18): p. 15059-15065.

58. Trott, O. and A.J. Olson, *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.* J Comput Chem, 2010. **31**(2): p. 455-61.

59. Reichert, J., et al., *The IMB Jena Image Library of biological macromolecules.* Nucleic Acids Res, 2000. **28**(1): p. 246-9.

60. Higuchi, R., B. Krummel, and R.K. Saiki, *A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions.* Nucleic Acids Res, 1988. **16**(15): p. 7351-67.

61. Struhl, K. and R.W. Davis, *Prodution of a functional eukaryotic enzyme in escherichia coli cloning and expression of a yeast structural gene for imidazoleglycerolphosphate dehydratase (HIS3).* Proceedings of the National Academy of Sciences of the United States of America, 1977. **74**(12): p. 5255-5259.

62. Wang, T.W., et al., *Mutant library construction in directed molecular evolution: casting a wider net.* Mol Biotechnol, 2006. **34**(1): p. 55-68.

63. Pritchard, L., et al., *A general model of error-prone PCR.* J Theor Biol, 2005. **234**(4): p. 497-509.

64. Voth, W.P., et al., *Yeast vectors for integration at the HO locus.* Nucleic Acids Research, 2001. **29**(12).

65. Johnson, K., *Extending chemical complementation to bacteria and furthering nuclear receptor based protein engineering and drug discovery (PhD)*. 2009, Georgia Institute of Technology: Atlanta.

66. Azizi, B., *Chemical complementation: A genetic selection system for drug discovery, protein engineering, and deciphering biosynthetic pathways (PhD)*. 2005, Georgia Institute of Technology: Atlanta, GA.

67. Schwimmer, L.J., et al., *Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(41): p. 14707-14712.

68. Gietz, R.D. and R.A. Woods, *Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method*, in *Guide to Yeast Genetics and Molecular and Cell Biology, Pt B.* 2002, Academic Press Inc: San Diego. p. 87-96.

69. Sanner, M.F., *Python: a programming language for software integration and development.* J Mol Graph Model, 1999. **17**(1): p. 57-61.

CHAPTER 5

CHARACTERIZATION OF L346 IN HUMAN ESTROGEN RECEPTOR ALPHA: DETERMINING TRANSCRIPTIONAL ACTIVATION WITH STEROID LIGANDS

5.1 Previous Mutational Analysis on Human Estrogen Receptor Alpha

As mentioned in Chapter 4, a significant amount of research has been reported on the human estrogen receptor alpha (hERα), specifically in terms of understanding the structural mechanism of activation as well as important ligand-receptor interactions [1-10]. For example, Aliau *et al.* mutated H524 to alanine, one of the hydrogen bonding residues in hERα, which caused a reduced level of activation toward 17-β estradiol (E2), confirming the significant interaction between H524 and the 17-hydroxy group on E2 [11]. In another study, Kim *et al.* mutated D351, a residue involved in coactivator binding and antagonist interaction, to a lysine as well as a leucine, causing significantly reduced E2 activation and eliminating coactivator interaction [6]. Studies similar to these determine the importance of specific residues or regions within various α-helices and have aided in understanding receptor activation.

Directed protein evolution on hERα has resulted in changed ligand specificity, providing insight into the effect of different residues on ligand binding and activation [1, 2, 10, 12, 13]. For example, E353, one of the hydrogen bonding residues in the pocket, has altered specificity toward androgens when mutated to a glutamine, demonstrating the significant role one position has on ligand affinity [1]. Some residues, such as L536 and Y537 were shown to interact differently with the ligand in agonist versus antagonist bound crystal structures [10, 14]. Additionally, different point mutations at Y537 have

resulted in receptors that have a range of constitutive activity [14]. Further investigation of the residues involved in ligand activation, however, is necessary to fully understand the role of individual amino acids in the pocket.

A rational library for hERα was previously designed and assessed in Chapter 4. The goal was to engineer a hERα variant to activate transcription in response to the β-lactam antibiotics ampicillin, amoxicillin, penicillin G, oxacillin, cloxacillin, and nafcillin. Amino acid residues from different α-helices surrounding the ligand in the LBP were mutated in the hERα library (Figure 5.1). Transformation of the library of mutations produced three variants, L346M, L346V, and L346E, containing a single point mutation at the same position. More importantly, variant L346E displayed significantly decreased sensitivity toward E2, showing that a point mutation at position 346 can significantly affect ligand interaction and activation of the receptor.

Structurally, L346 is located on helix 3, one of the helices involved in forming the ligand binding pocket (LBP) [15-17]. Most of the structural analysis previously performed has been focused on helix 12, as well as 10/11 and 7 [5, 10, 14, 18, 19]. Some work has been done on the role of helix 3 and helix 5 [20-22]. Previously, mutations in helix 3 of the androgen receptor (AR) have proven useful in understanding androgen insensitivity in the male reproductive system [23]. Several crystal structures of the mineralocorticoid receptor (MR) have shown the importance of residues in helix 3 in a hydrogen bond network with the ligand [24]. In addition, among the steroid receptors, an interaction occurs between an alanine on helix 3 and a leucine on helix 5. For example, A350 in helix 3 contacts L387 in helix 5 in the estrogen receptor, indicating an important role interactions with helix 3 have on receptor activation [16].

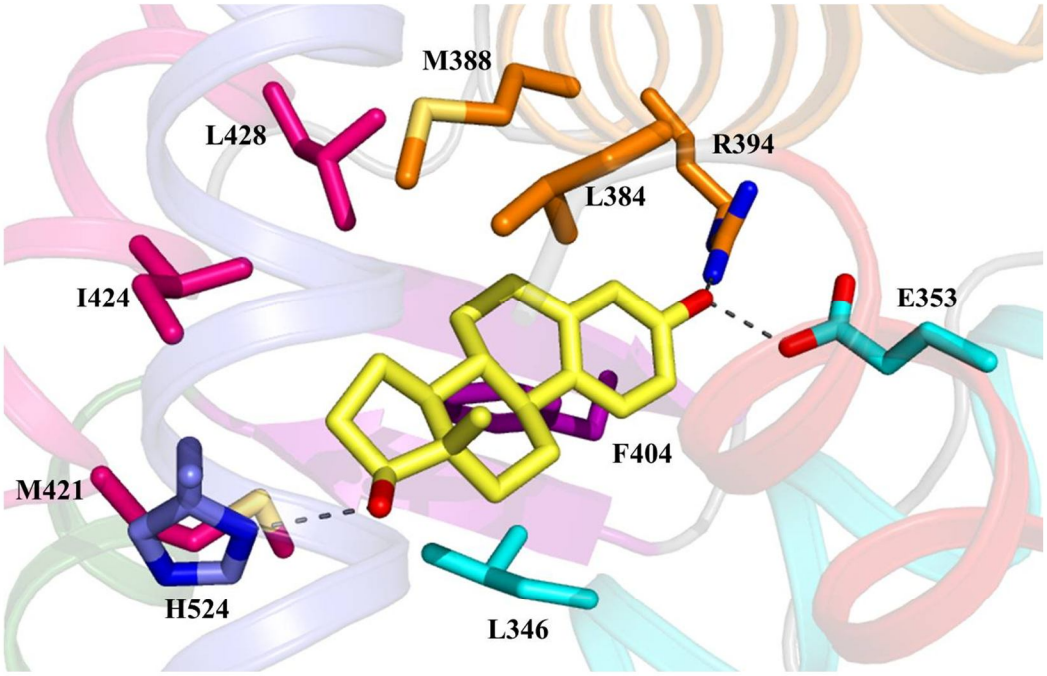


Figure 5.1. The LBP of the crystal structure of hERα with 17 β-estradiol (E2) (PDB 1ERE). The different α-helices are colored cyan, orange, green, pink, blue, and red, for 3, 5, 6, 7, 10/11, and 12, respectively. The β-sheet is shown in purple. Hydrogen bonds are shown as black dotted lines. E2 is shown in yellow sticks. Surrounding residues are shown in sticks.

As a residue located on helix 3, studies on the effect of L346 on ligand binding would provide additional information on the involvement of helix 3 in ligand interaction [5, 7, 17]. Multiple crystal structures of the hERα LBD show that L346 makes hydrophobic contacts with the ligand in the pocket (Figure 5.2) [4, 9, 16, 25, 26]. In addition, the leucine at position 346 is fully conserved among the androgen receptor (AR), progesterone receptor (PR), glucocorticoid receptor (GR), and mineralocorticoid receptor (MR), all steroid receptors within the nuclear receptor superfamily [26]. Moreover, D'Ursi *et al.* showed that L346 made contacts with a variety of docked polychlorinated compounds, maintaining that L346 can interact with ligands having different structures [4].

A recent study by Shadnia and coworkers developed an algorithm to determine energy based interactions between proteins or enzymes with their respective ligands or substrates [8]. The work performed by Shadnia and coworkers discovered a series of amino acid residues in the pocket of hERα with significant molecular interactions with E2, designated as force interactions [8]. The known hydrogen bonding residues E353, R394, and H524 had some of the highest force interactions. Furthermore, two other residues with high force interactions were L346 and L387, providing additional support for the importance of L346 in ligand interaction [8]. As a result of the previous analyses on L346 and the marked differences in activity toward E2 amongst the three variants containing a mutation at position 346 and to investigate the role this residue has in ligand binding and activity, L346 was saturated via site-directed mutagenesis using primers containing codons for each different amino acid and tested in yeast as well as mammalian cells.

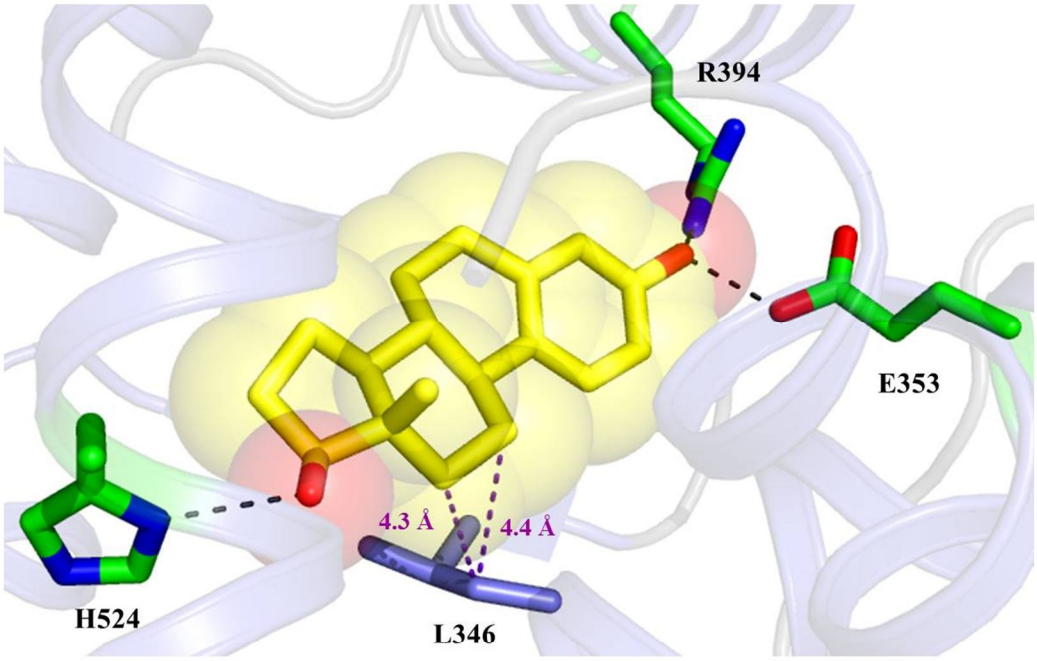


Figure 5.2. **Image of the crystal structure of the hERα LBD bound to E2, showing the ligand binding pocket (PDB 1ERE).** E2 is shown in yellow sticks with spheres. The three hydrogen bonding residues are shown in green sticks. Hydrogen bonds are shown as black dotted lines. The distances between the beta carbon of L346 and carbon 11 and 12 of the steroid ring of E2 are shown as purple dotted lines with the distances indicated.

5.2 Analysis of L346 Variants

5.2.1 L346 Variants in Chemical Complementation

All twenty variants containing the different mutations at position 346 were transformed into yeast to test in the chemical complementation yeast assay with E2. As mentioned in Chapter 4, chemical complementation uses two fusion proteins, a plasmid containing the yeast Gal4 DNA binding domain (GBD) fused to the hERα LBD, containing a tryptophan marker, and a plasmid containing a human coactivator fused to the Gal4 activation domain (GAD), containing a leucine marker. The coactivator in this system contains peptide motif repeats of LXXLL of the steroid receptor coactivator-1 (SRC-1). As previously mentioned, in the liquid assay, yeast containing the variants are grown in a 96-well plate in selective media without ligand as well as with increasing concentrations of ligand.

The twenty hERα variants were tested in a dose response with E2 ranging from 10 nM to 10 μM in histidine selective media (SC-HLW+3-AT). As a positive control, Gal4, a ligand-independent transcription factor, was used, and displays growth with and without the presence of ligand (Figure 5.3A, olive green square). Ligand activated growth was observed for all variants except L346D, L346R, and L346P, displaying a specific trend in sensitivity toward E2 based on the type of mutation (Figure 5.3A-C). The observed pattern showing different levels of sensitivity toward E2 can be separated into three different categories. In general, variants with the large and hydrophobic amino acids isoleucine, phenylalanine, tryptophan, methionine, and valine display similar sensitivity toward E2 as the wild type leucine (Figure 5.3A, purple triangle, green diamond, orange

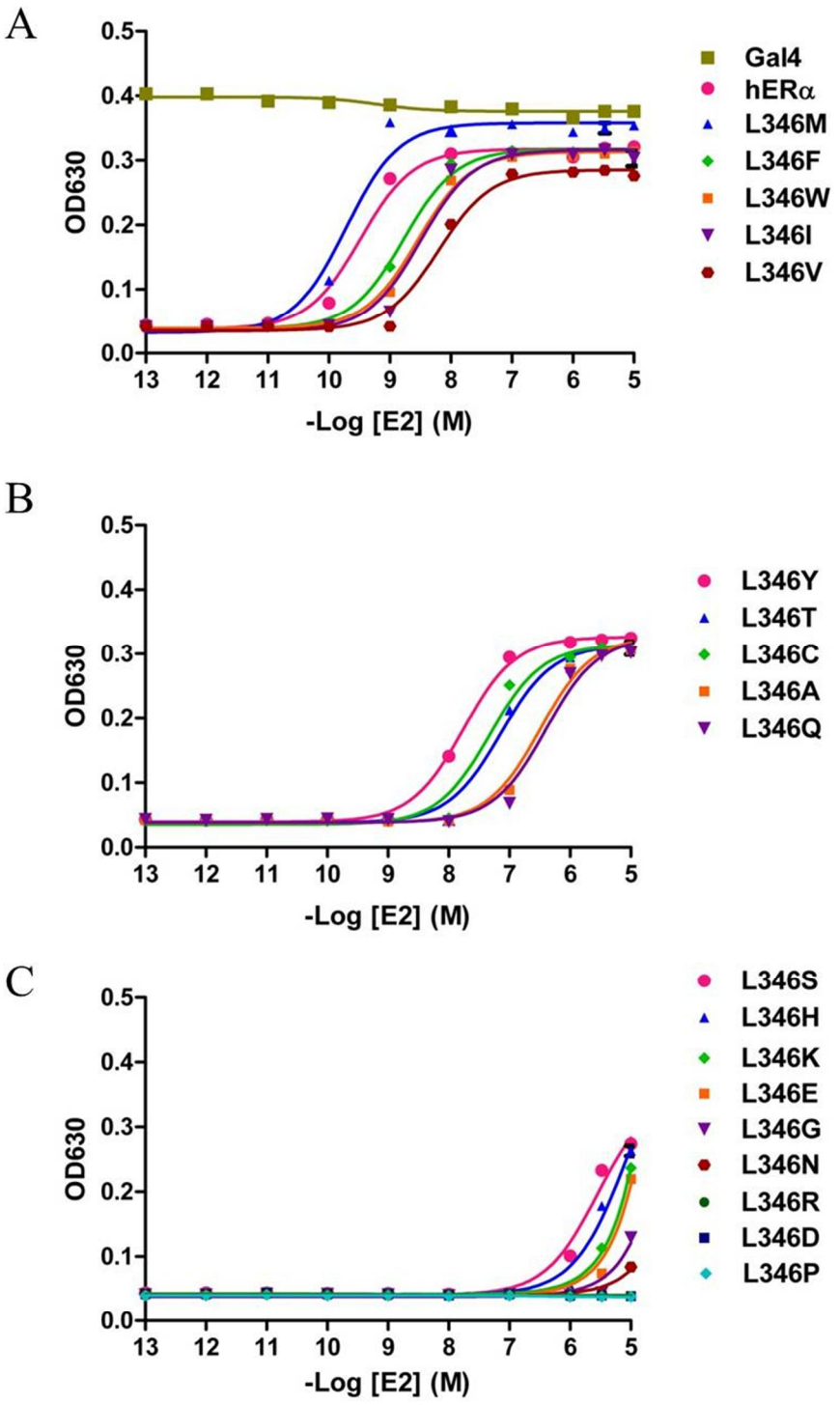


Figure 5.3. Growth profiles in PJ69-4A yeast cells with E2 for wild type hERα and L346 variants in histidine selective media containing 10 mM 3-AT. (A) Gal4, wild type hERα, L346M, L346F, L346W, L346I, and L346V. (B) L346 Y, L346C, L346T, L346A, and L346Q. (C) L346S, L346H, L346K, L346E, L346G, L346N, L346R, L346G, and L346P.

Table 5.1. EC_{50} values, based on growth, and fold activations in PJ69-4A yeast cells and HEK293T mammalian cells for wild type hERα and L346 variants with E2. Values are in nanomolar range except where indicated.

	E2 Yeast		E2 HEK293T	
Variant	EC_{50}	Fold Activations	EC_{50}	Fold Activations
hERα	0.3	7.3	0.2	8±2
L346I	3.2	7.4	1.4	16±3
L346F	1.7	7.2	1.2	56±9
L346W	2.9	7.5	2.2	68±18
L346M	0.2	8.4	0.6	50±13
L346V	6.3	6.5	4.6	77±6
L346Y	17	7.6	42	83±6
L346T	72	7.0	139	73±39
L346C	48	7.3	55	78±11
L346A	321	7.3	392	79±3
L346Q	399	7.0	678	82±5
L346S	2.8 μM	6.2	> 10 μM	67±3
L346H	8.5 μM	6.0	> 10 μM	48±5
L346K	3 μM	5.4	> 10 μM	41±2
L346E	3 μM	5.0	> 10 μM	25±2
L346N	> 10 μM	2.0	> 10 μM	4±0
L346G	> 10 μM	3.1	> 10 μM	9±0
L346D	Dead	Dead	Dead	Dead
L346P	Dead	Dead	Dead	Dead
L346R	Dead	Dead	Dead	Dead

square, blue diamond, and maroon hexagon, respectively). The EC_{50} values for the hydrophobic substitutions range from approximately 300 pM to 6 nM, with wild-type having an EC_{50} value of 0.33 nM (Table 5.1). For example, variants with the larger hydrophobic amino acids, phenylalanine and methionine, show sensitivity similar to wild-type leucine, having EC_{50} values of 1.68 nM and 0.19 nM, respectively (Figure 5.3A, green diamond and blue triangle, respectively).

Polar or small amino acids display decreased sensitivity, having EC_{50} values of approximately 17 nM, 72 nM, 48 nM, 321 nM, and 399 nM, for tyrosine, threonine, cysteine, alanine, and glutamine, respectively (Figure 5.3B and Table 5.1). In this category, for example, tyrosine has the lowest EC_{50}, most likely due to the larger volume; however, the polar hydroxyl group causes this amino acid to have a higher EC_{50} than the first category (Figure 5.3B, pink circle). The fold activations for both the larger hydrophobic as well as the polar and small amino acids are approximately 7-fold (Table 5.1).

Lastly, charged and smaller amino acids show the lowest sensitivity toward E2, having EC_{50} values of 2.8 μM, 8.5 μM, 3 μM, and 3 μM for serine, histidine, lysine, and glutamic acid, respectively (Figure 5.3C, pink circle, blue triangle, green diamond, and orange square, respectively, and Table 5.1). L346 mutated to glycine and asparagine have EC_{50} values greater than 10 μM (Figure 5.3C, purple triangle, maroon hexagon, respectively, and Table 5.1). The fold activations for the charged and smaller amino acids are lower than the larger hydrophobic and polar amino acids, ranging from 2 to 6-fold (Table 5.1).To determine if the significant variation of sensitivity toward E2 seen in yeast

was consistent as well as to observe whether the results display the three distinct levels of activation, these variants were next tested in mammalian cell culture.

5.2.2 L346 Variants in HEK293T Mammalian Cells

To assess whether the variants containing the mutations at position 346 displayed the same activation levels as the growth levels observed in the yeast assay, all twenty variants were tested in mammalian cell culture. The variants comprising the Gal4 DBD fused to the hERα LBD were cloned into the mammalian expression plasmid pCMX, containing a CMV promoter, using *SacII* and *NotI* restriction sites. This plasmid, pCMXGBDhERαLBD, was transfected alongside a reporter plasmid containing Gal4 response elements controlling a luciferase gene (p17*4TATALuc), as previously described [27].

The same activation profiles occur in mammalian cells with the different variants as were seen in chemical complementation (Figure 5.4A-C). The three categories of sensitivity toward E2 are consistent, displaying the trend where position 346 favored hydrophobic bulkier amino acids and have a lower tolerance for smaller or charged amino acids. The EC_{50} values for the hydrophobic substitutions range from approximately 150 pM to 5 nM, more specifically 0.15 nM, 1.4 nM, 1.2 nM, 2.2 nM, 0.6 nM, and 4.6 nM, for wild-type hERα, isoleucine, phenylalanine, tryptophan, methionine, and valine, respectively (Table 5.1). The polar or small amino acids have EC_{50} values slightly higher than those observed in yeast, however the trend is the same (Table 5.1). L346 mutated to tyrosine, threonine, cysteine, alanine, and glutamine, have EC_{50} values of 42 nM, 139 nM, 55 nM, 392 nM, and 678 nM, respectively (Table 5.1). In the last

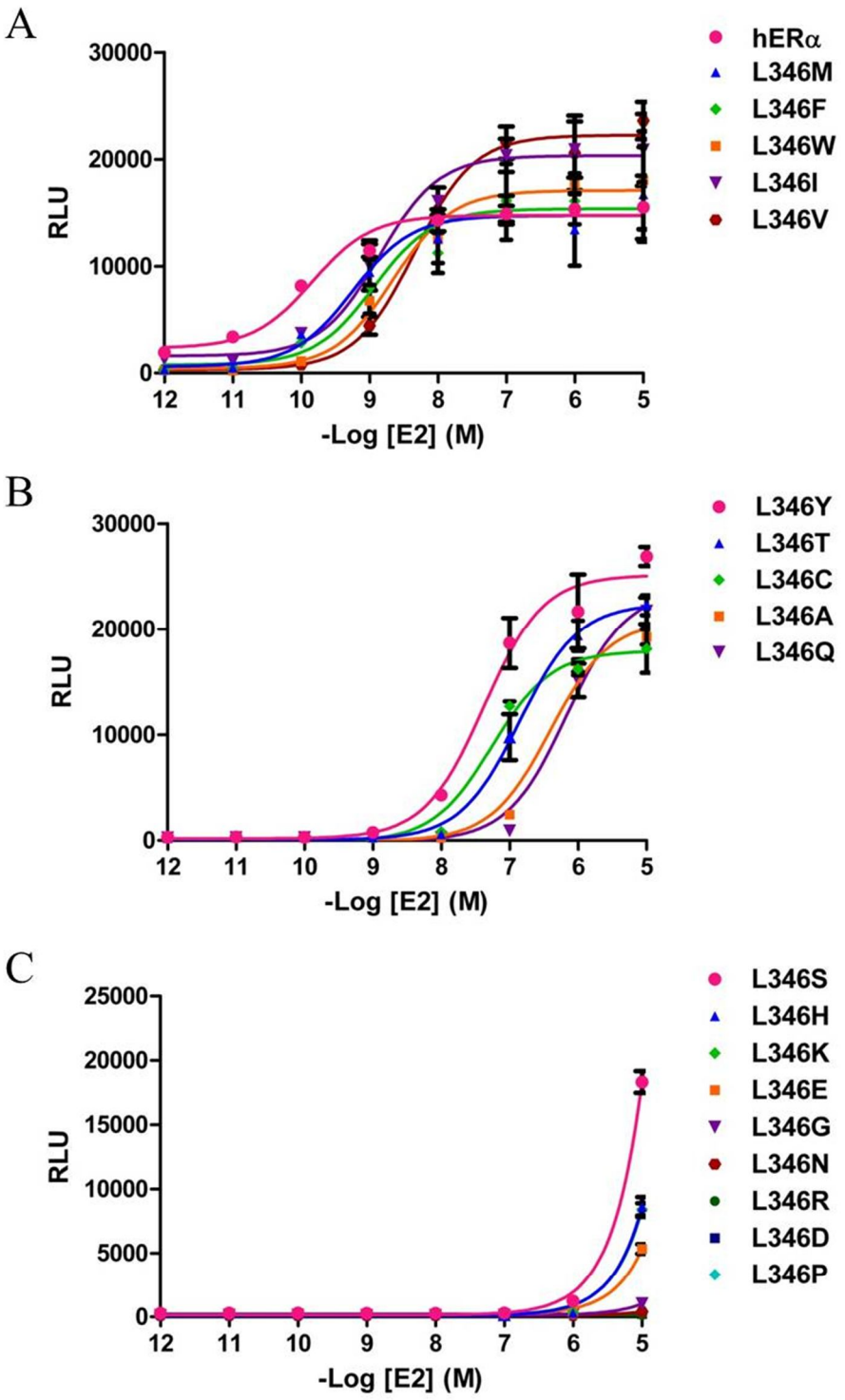


Figure 5.4. Activation profiles in HEK293T mammalian cells with E2 for wild type hERα and L346 variants. (A) Wild type hERα, L346M, L346F, L346W, L346I, and L346V. (B) L346 Y, L346C, L346T, L346A, and L346Q. (C) L346S, L346H, L346K, L346E, L346G, L346N, L346R, L346G, and L346P.

category, the charged amino acids, serine, histidine, lysine, glutamic acid, asparagine, and glycine have EC_{50} values greater than 10 μM (Table 5.1).

Interestingly, one difference observed between wild-type L346 and all the variants is the basal level of activity. Wild-type leucine has higher basal, thus leading to a significantly lower fold activation than the variants. For example, wild-type leucine has a fold activation of 8±2 with E2, whereas L346F has a fold activation of 56±9 (Table 5.1). The higher fold activations are consistent among all the variants except L346N and L346G, which have 4±0 and 9±0 fold activation, respectively (Tablel 5.1). The higher basal level for wild-type hERα, however, is not surprising due to the conserved amino acid at this position among the steroid receptors, perhaps indicating an evolutionarily significant role of the biophysical chemistry of that area of the pocket.

Thus far, the yeast data and the mammalian assays show the tolerance for different types of mutations at position 346, with a preference for larger hydrophobic residues. Activation occurs with the smaller and polar amino acid substitutions, with little to none seen for the charged amino acids. In addition, these results also demonstrate the preference for the wild-type amino acid leucine at position 346, thus indicating the significant influence an evolutionarily conserved residue has on receptor-ligand interactions. More importantly, both the yeast data and the mammalian data confirm that a single point mutation causes drastic changes in the sensitivity of hERα toward the wild-type ligand E2. Therefore to determine if the same difference in activity was seen with additional ligands, these variants were tested with a variety of steroids, which have a similar core structure as E2, but different functional groups. The hypothesis was that changing the chemical nature of the amino acid at position 346 would cause the variant to

bind another steroid. Due to the response seen in yeast and mammalian cells toward E2 for L346D, L346R, and L346P, these three variants were not included in further analyses.

5.3 Activation with Additional Steroid Ligands

As previously mentioned, ER, AR, PR, GR, and MR belong to the same steroid receptor subgroup in the nuclear receptor superfamily [28, 29]. Classically, the estrogen receptor is believed to be the ancestral steroid receptor, however, AR, PR, GR, and MR bind 3-keto steroids, whereas ER binds the 3-hydroxy steroid (Figure 5.5) [30]. Genome mapping and phylogenetic analyses have determined that the ancestral vertebrate ER diverged into the current ERα and ERβ isotypes [30]. These studies also show that evolutionarily, AR and PR diverged from an ancestor that bound androgens, while MR and GR came from a receptor that bound corticoids [30]. The wild-type ligands for AR, PR, and GR include testosterone, progesterone, cortisol, and corticosterone [30]. The steroid receptors are also known to have high ligand specificity toward each wild-type ligand, thereby allowing each receptor to regulate distinct metabolic pathways [1, 2, 31]. As mentioned in Chapter 4, however, mutations within these various steroid receptors have resulted in altered steroid receptor specificity [1, 5, 31]. Several L346 variants displayed different activation profiles with E2 than wild-type hERα. For example, since L346E had drastically reduced sensitivity toward E2, perhaps this variant would have enhanced activity toward other steroid molecules. Therefore, due to these differences, the L346 variants were tested in HEK293T cells with 10 μM testosterone, progesterone, cortisol, and corticosterone, the wild-type ligands for the steroid receptors AR, PR, and

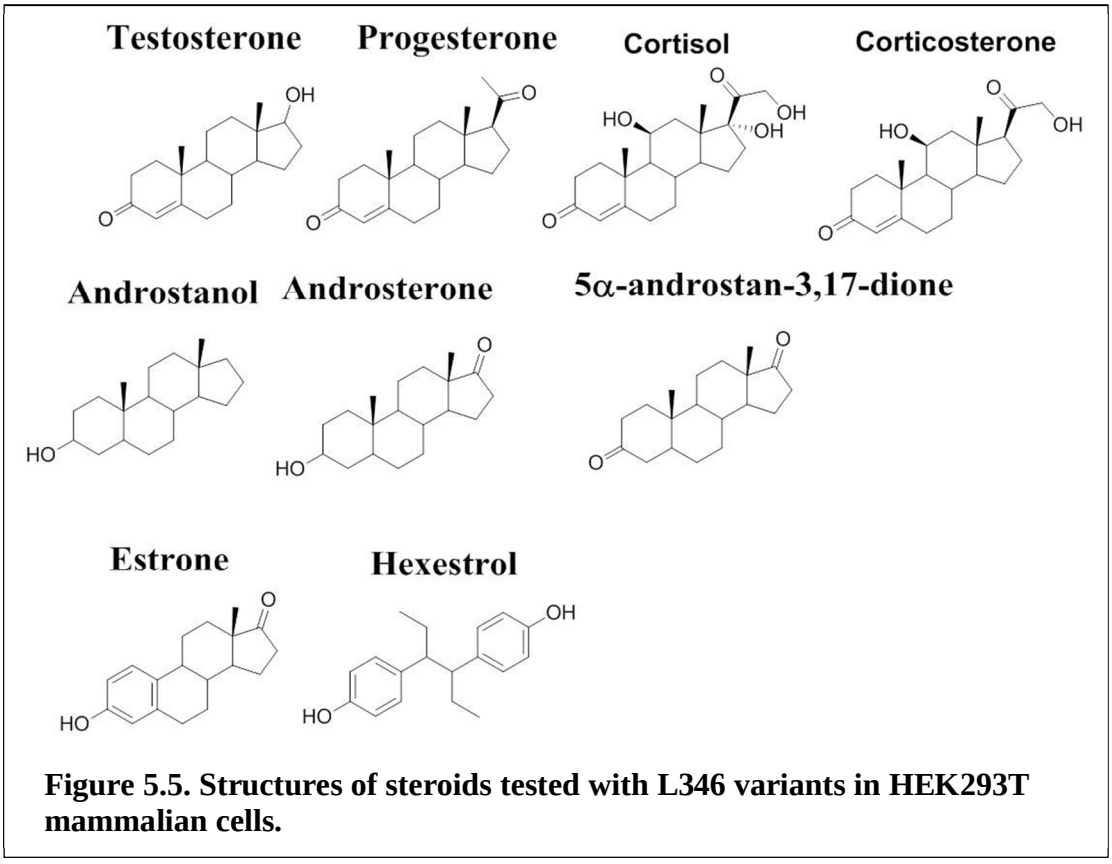


Figure 5.5. Structures of steroids tested with L346 variants in HEK293T mammalian cells.

GR, respectively (Figure 5.5). The hypothesis was that the variants containing amino acids with different substitutions at position 346 could be activated by these similar steroid ligands.

5.3.1 Activity in Mammalian Cells with Steroid Receptor Ligands and Androstanols

Wild-type hERα was activated by testosterone, having 9-fold activation, consistent with literature values (Figure 5.6A,B and Table 5.2) [2]. Variants L346I, L346F, L346W, and L346M have 1, 7, 3, and 4-fold activations with testosterone, respectively (Table 5.2) No activation was observed for wild-type hERα with progesterone, cortisol, and corticosterone (Figure 5.6BC). All other L346 variants were not activated by any of the four steroids, confirming the known high ligand specificity of the steroid receptors (Figure 5.6B,C).

All L346 variants except L346D, L346R, L346P, and L346N were tested in HEK293T cells at 10 μM for all the other steroids, specifically androstanol and the androstanol derivatives androsterone and 5α androstan-3,17-dione (Figure 5.5) [32, 33]. Androstanol is a testosterone metabolite as well as a ligand for both the pregnane X receptor and the constitutive androstane receptor [34, 35]. These ligands maintain the steroid ring system, however, contain different substituents than testosterone, progesterone, cortisol, and corticosterone (Figure 5.5). Structurally, androstanol and androsterone maintain the hydroxy group at the 3-carbon position, but do not have the π ring system in the A ring of the steroid as with E2 (Figure 5.5). In addition, the hydroxy group at carbon 17 is a 17-deoxy or 17-keto group in androstanol or androsterone, respectively (Figure 5.5). In addition, these functional groups are smaller than those on

Table 5.2. Fold activations in HEK293T mammalian cells for wild type hERα and L346 variants with Testosterone

Variant	Testosterone
wtER	9 ± 1
L346I	1.2 ± 0.1
L346F	7 ± 1
L346W	3.1 ± 0.5
L346M	3.6 ± 0.5
L346V	0.9 ± 0
L346Y	0.8 ± 0
L346T	0.9 ± 0.1
L346C	1.0 ± 0.2
L346A	0.9 ± 0.2
L346Q	1.0 ± 0
L346S	1.0 ± 0.1
L346H	1.0 ± 0.1
L346K	1.0 ± 0.1
L346E	0.9 ± 0
L346G	0.9 ± 0.1

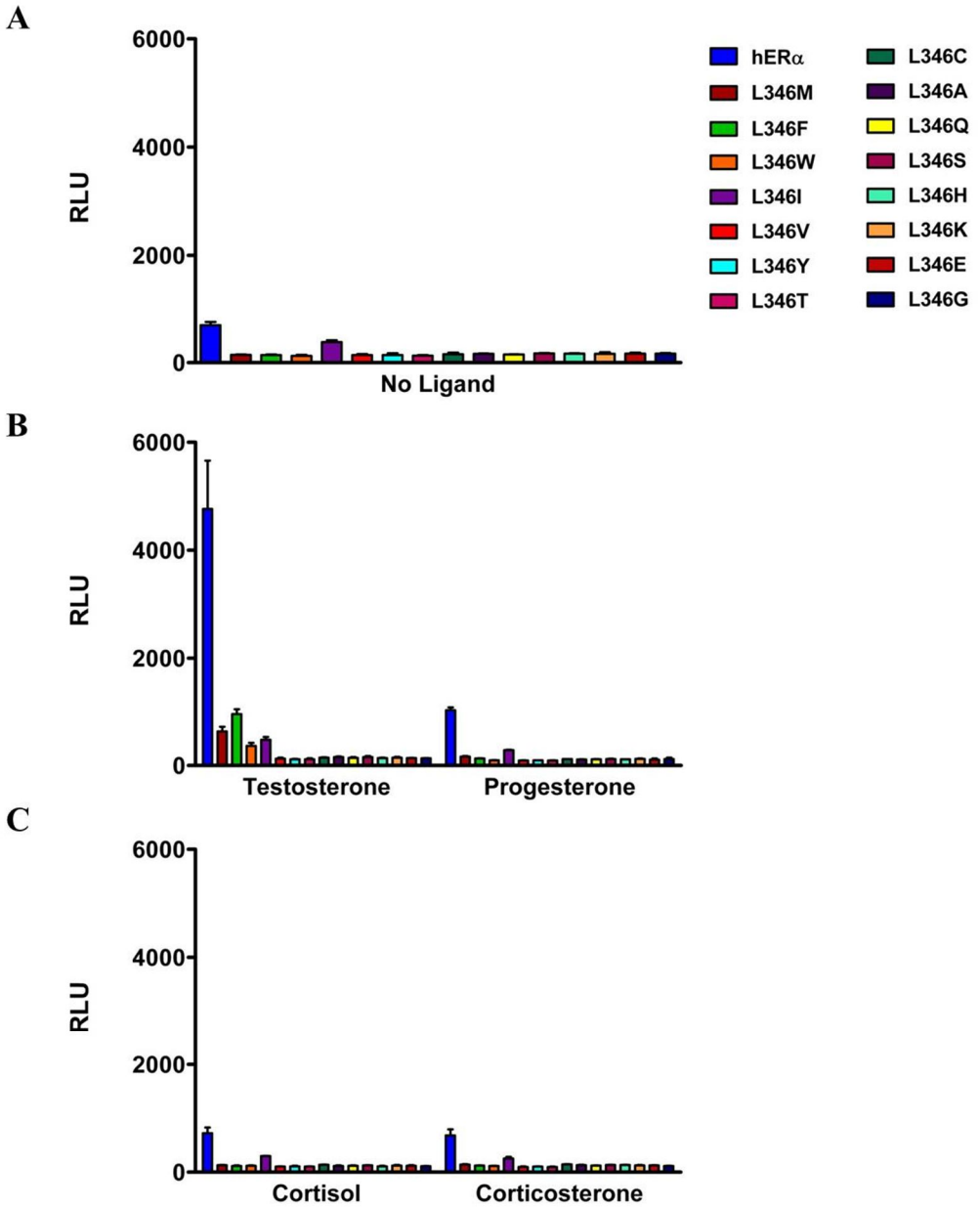


Figure 5.6. Activation profiles for L346 variants in HEK293T mammalian cells with (A) no ligand and with 10 μM of (B) Testosterone and Progesterone and (C) Cortisol and Corticosterone.

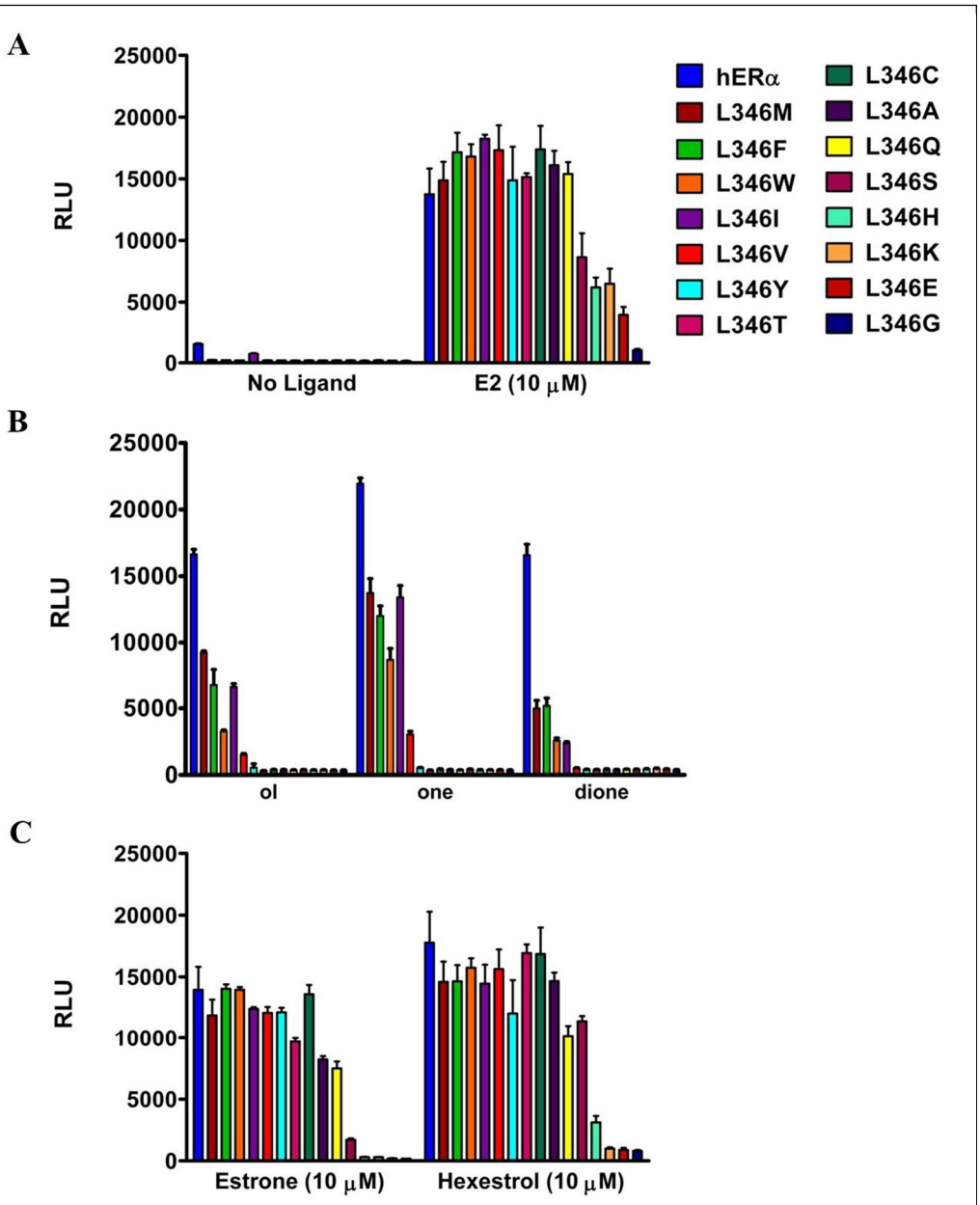


Figure 5.7. Activation profiles for L346 variants in HEK293T mammalian cells with 10 μM ligand for (A) E2 (B) Androstanol, Androsterone, and 5α-androstan-3,17-dione and (C) Estrone and Hexestrol.

progesterone, cortisol, and corticosterone making them more similar in structure to E2 and therefore providing ligands that could activate the L346 variants.

As a control, the variants were also tested with 10 μM E2, displaying activation levels consistent with those in Table 5.1 (Figure 5.7A and Table 5.4). Only wild-type hERα and variants L346I, L346F, L346W, and L346M showed activation with androstanol, androsterone, and 5α androstan-3,17-dione, confirming the preference of a hydrophobic residue at position 346 (Figure 5.7B). Wild-type hERα displayed fold activations of 18-fold, 24-fold, and 18-fold for androstanol, androsterone, and 5α androstan-3,17-dione, respectively (Table 5.3). Fold activations for variants L346I, L346F, L346W, and L346M for are 27, 24, 11, and 17-fold activations for androstanol, respectively (Table 5.3). With slightly less fold-activations, L346I, L346F, L346W, and L346M displayed 15, 18, 9, and 6-fold activations for 5α androstan-3,17-dione, respectively (Table 5.3). Of these three ligands, androsterone displayed the highest level of activation, which could be as a result of maintaining the 3-hydroxy group as well as having a polar substituent at carbon 17 (Figure 5.5). L346I, L346F, L346W, and L346M displayed fold activations of 40, 42, 30, and 35-fold for androsterone (Table 5.3). In summary, these data show the ability for hydrophobic amino acids, which have properties similar to the wild-type leucine, at position 346 to display some level of activation toward different steroid ligands. However, the varying levels of activation toward these ligands in comparison to the natural ligand E2 demonstrates that even slight differences in the structures of the steroids affects activity, specificity, and sensitivity of the receptor [36].

5.3.2 Activity in Mammalian Cells with Estrone and Hexestrol

Table 5.3. Fold activations in HEK293T mammalian cells for wild type hERα and L346 variants with various ligands. E2 is shown as a control.

Variant	E2	Androstanol	Androsterone	5α-androstan-3,17-dione
hERα	32.17	18.5	24.4	18.4
L346I	74.36	27.1	40.3	14.7
L346F	88.34	23.7	41.8	18.1
L346W	92.34	11.3	30.1	8.9
L346M	72.84	17.4	35.0	6.3
L346V	82.88	4.4	8.8	1.5
L346Y	86.92	1.5	1.5	1.1
L346T	77.93	0.9	1.0	1.2
L346C	87.68	1.2	1.2	1.2
L346A	73.99	1.0	1.1	1.1
L346Q	82.44	1.1	1.2	1.2
L346S	71.66	1.0	1.0	1.2
L346H	51.06	1.0	1.0	1.2
L346K	60.82	1.1	1.1	1.3
L346E	30.48	1.0	1.1	1.2
L346G	11.29	0.9	0.9	1.0

Since most of the variants were not activated by testosterone, progesterone, cortisol, and corticosterone, or the androstanol compounds, two steroid ligands known to interact with ER were chosen to test in mammalian cells. The L346 variants were tested with the estrogen hormone, estrone, and the non-steroidal ER ligand hexestrol in HEK293T cells with 10 μM of each ligand [37, 38]. Estrone maintains the π ring system in the A ring as E2 as well as the 3-hydroxy group, but has a ketone at carbon 17 (Figure 5.5). Hexestrol also has the same A ring hydroxyl group as E2, but does not have the core steroid structure (Figure 5.5).

These data show that the activation levels were more similar to those seen with E2. With estrone, L346 mutated to isoleucine, phenylalanine, tryptophan, methionine, valine, tyrosine, threonine, cysteine, alanine, and glutamine all show activation levels comparable to wild-type hERα (Figure 5.7C). Fold activations, however, are significantly higher for the variants than wild-type hERa, ranging from 15-fold to 74-fold for L346 variants mutated to isoleucine, phenylalanine, tryptophan, methionine, valine, tyrosine, threonine, cysteine, alanine, and glutamine, whereas wild-type hERα only has a 8.7-fold activation, due to the higher basal levels previously mentioned (Table 5.4). A significant decrease in fold activation occurs at L346S, having a fold-activation of 9-fold (Table 5.4). L346H, L346K, and L346G were not activated by estrone (Figure 5.7C and Table 5.4).

Hexestrol activation levels display a similar trend as estrone, with slightly higher fold activations (Table 5.4). L346 mutated to isoleucine, phenylalanine, tryptophan, methionine, valine, tyrosine, threonine, cysteine, alanine, and glutamine, serine, and histidine had fold activations ranging from 18.7-90.5-fold, whereas wild-type had a 11-

Table 5.4. Fold activations in HEK293T mammalian cells for wild type hERα and L346 variants with Estrone and Hexestrol. E2 is shown as a control.

Variant	E2	Estrone	Hexestrol
hERα	8.59	8.7	11.1
L346I	22.71	15.0	19.4
L346F	82.46	67.5	70.4
L346W	88.69	73.6	83.1
L346M	63.57	50.7	62.3
L346V	89.45	64.1	74.7
L346Y	79.22	64.5	64.1
L346T	80.94	51.8	90.5
L346C	89.84	70.2	87.0
L346A	83.87	42.9	76.4
L346Q	78.29	38.1	51.5
L346S	44.79	8.9	59.1
L346H	36.99	1.8	18.7
L346K	33.52	1.5	5.1
L346E	22.74	1.1	5.0
L346G	7.36	1.2	5.3

fold activation (Table 5.4). L346K, L346E, and L346G displayed very low activation, having fold activations of approximately 5-fold (Figure 5.7C and Table 5.4). These data demonstrate the importance of maintaining the phenolic A-ring, although more ligands would need to be tested for additional validation. Both estrone and hexestrol also confirm the trend that charged amino acids are not preferred at position 346.

5.3.3 Activity in Mammalian Cells Adrosterone and 17-α Estradiol

Since androsterone showed the highest level of activation among the three androstanol steroids tested, the variants containing the hydrophobic mutations, wild-type hERα, L346I, L346F, L346W, L346M, and L346V, were analyzed in a dose response with this ligand alongside 17-α estradiol (17αE2) to determine sensitivity levels toward these two ligands. Moreover, we wanted to confirm the preference for a hydrophobic amino acid at position 346. Wild-type hERα displayed the lowest EC_{50} value toward 17αE2, having an EC_{50} of 3.25 nM, although due to the basal level of expression, the fold activation of 10±0 was lower than the other variants containing a hydrophobic amino acid Figure 5.8 and Table 5.5). L346I, L346F, L346W, L346M, and L346 V displayed EC_{50} values of 12.33 nM, 32.81 nM, 53.58 nM, 10.62 nM, and 26.73 nM, respectively (Figure 5.8 and Table 5.5). L346I, L346F, L346W, L346M, and L346V have fold activations ranging from 18-fold to 95-fold (Table 5.5).

Wild-type hERα, L346I, L346M, L346F, and L346W displayed some activation toward androsterone (Figure 5.9). Wild-type hERa has an EC_{50} value at 7.2 μM, whereas L346I, L346M, L346F, and L346W had EC_{50} values greater than 10 μM (Table 5.5). The fold activations for wild-type hERα, L346I, L346M, L346F, and L346W are 4, 4, 11, 6,

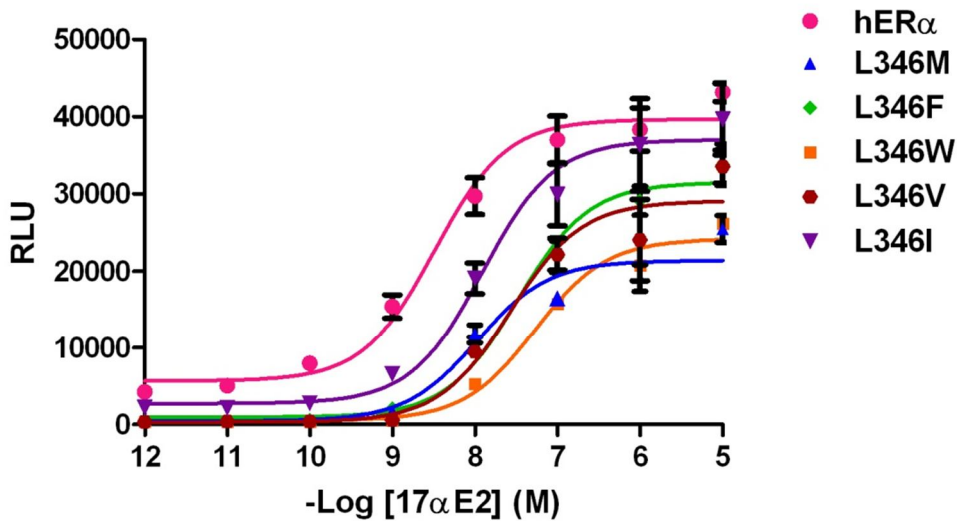


Figure 5.8. Activation profiles in HEK293T mammalian cells with 17 α-estradiol (17αE2) for wild type hERα and L346 variants containing hydrophobic substitutions.

Table 5.5. EC_{50} values and fold activations in HEK293T mammalian cells for wild type hERα and L346 variants with 17αE2 and Androsterone. Values are in nanomolar range except where indicated.

	17αE2	HEK293T	Androsterone	HEK293T
Variant	EC_{50}	Fold Activations	EC_{50}	Fold Activations
hERα	3.25	10±0	7.2 μM	4±0
L346I	12.33	18±2	> 10 μM	4±0
L346F	32.81	79±12	> 10 μM	11±1
L346W	53.58	63±1	> 10 μM	6±1
L346M	10.62	51±4	> 10 μM	14±3
L346V	26.73	95±13	> 10 μM	2±0
L346Y	--	--	--	--
L346T	--	--	--	--
L346C	--	--	--	--
L346A	--	--	--	--
L346Q	--	--	--	--
L346S	--	--	--	--
L346H	--	--	--	--
L346K	--	--	--	--
L346E	--	--	--	--
L346N	--	--	--	--
L346G	--	--	--	--
L346D	--	--	--	--
L346P	--	--	--	--
L346R	--	--	--	--

-- Not tested

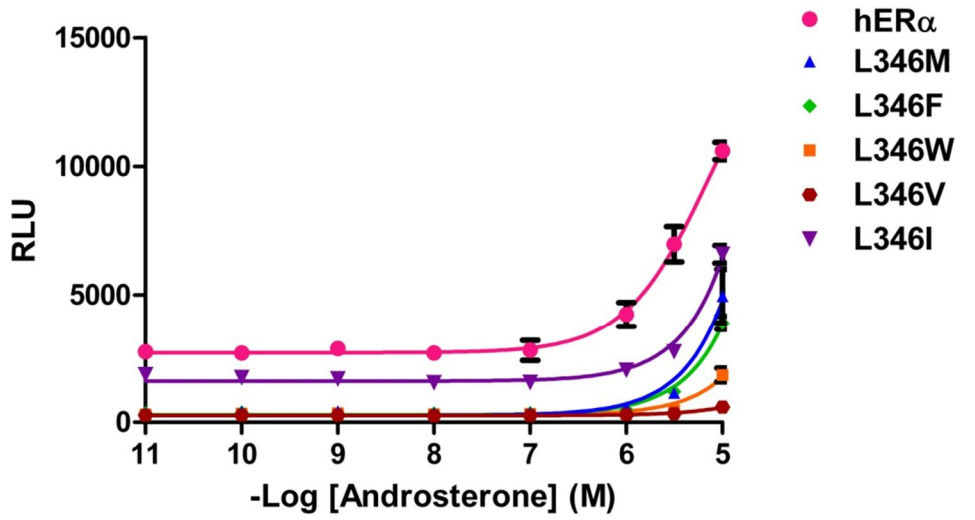


Figure 5.9. Activation profiles in HEK293T mammalian cells with androsterone for wild type hERα and L346 variants containing hydrophobic substitutions.

and 14-fold, respectively (Table 5.5). The variant L346V was not activated by androsterone (Figure 5.9 and Table 5.5). As previously mentioned, these data confirm the ability for hydrophobic amino acids at position 346 to display some level of activation toward different steroid ligands.

5.4 Summary and Future Work

Sequencing from the rational library in Chapter 4 revealed three variants containing a single point mutation at residue L346: L346M, L346V, and L346E. These variants were tested in a dose response with E2 in liquid growth assays in yeast. Variant L346E displayed significantly decreased sensitivity toward E2, having an EC_{50} value of 3 μM, whereas wild type has an EC_{50} value of 0.3 nM, based on growth (Figure 5.3A). L346V also had slightly decreased sensitivity toward E2 having an EC_{50} value around 6 nM, based on growth. L346M, however, maintained wild type activity with E2.

Therefore, to expand on the role this residue plays on the interaction between hERα and the ligand, this position was saturated and all variants were tested in yeast and mammalian cells with E2. The activation profiles in yeast and mammalian cells displayed a specific trend. Hydrophobic bulkier amino acids had activation levels similar to the wild-type leucine, whereas variants containing polar or smaller amino acids, specifically L346Y, L346T, L346C, L346A, and L346Q had decreased sensitivity toward E2. The variants with charged or smaller amino acids, L346S, L346H, L346K, L346E, L346N, and L346G had significantly decreased sensitivity toward E2.

When tested in HEK293T cells with additional steroid and non-steroidal ligands, similar activation profiles were observed. Wild-type hERα and variants L346I, L346F, L346W, and L346M showed activation with the steroids androstanol, androsterone, and

5α androstan-3,17-dione, also confirming the importance of a hydrophobic amino acid at position 346. All other variants were not activated by this set of steroid ligands. When tested with estrone or the non-steroidal ligand hexestrol, the charged amino acids had little to no activity toward these two ligands.

Lastly, the variants containing hydrophobic amino acid mutations were tested in a dose response with androsterone and 17αE2 in HEK293T cells. With androsterone, wild-type hERα had an EC_{50} around 7 μM toward androsterone whereas the variants containing the hydrophobic amino acids were activated by androsterone at approximately 10 μM, displaying the preference for wild-type leucine. In summary, the analysis with the additional steroid ligands confirms the preference toward the wild-type leucine as well as hydrophobic amino acids at position 346. More importantly, this work shows how a point mutation at position 346 significantly affects ligand activation.

5.5 Materials and methods

5.5.1 Plasmids and Primers

The plasmids pGBDhERαLBD, pGAD10BALXXLLSRC-1, pGBDhERαLBD background plasmid, and pCMXGBDhERαLBD were made as described in Chapter 4. All single mutants were generated through QuikChange™ (Stratagene) site-directed PCR on pGBDhERαLBD or pCMXGBDhERαLBD using the forward primer 5’ agcttcgatgat gggcttannnaccaacctggcagaca 3’ and reverse primer 5’ tgtctgccaggttggtnnntaagcccatcatcgaagct 3’, each containing the different codons for each amino acid at the indicated nnn location.

5.5.2 Ligands

17-β estradiol (MW=272.4 g/mol) was purchased from MP Biomedicals (Aurora, OH). Androstanol was purchased from BIOMOL (Plymouth Meeting, PA). Androsterone (MW=290.44 g/mol) and 5α androstan-3,17-dione (MW=288.42 g/mol) were a kind gift from Dr. Andreas Bommarius (Georgia Institute of Technology, Atlanta, GA). Estrone (MW=270.4 g/mol) and 17-α estradiol (MW=272.4 g/mol) were purchased from Sigma Aldrich (St. Louis, MO). 10 mM stocks of each ligand were dissolved in 80% ethanol:20% DMSO and stored at 4°C.

5.5.3 Liquid Yeast Growth Assay

The yeast growth assay was performed as described in Chapter 4 and as previously described [39]. Briefly, transformants are grown overnight in nonselective media containing synthetic complete media lacking leucine and trypophan (SC-LW) at 30 °C with shaking at 300 rpm. Yeast are added to 96-well plates containing selective media without ligand and with increasing concentrations of ligand in a 4:1 ratio of media to yeast. Plates are incubated for 48 hours at 30 °C with shaking at 170 rpm. Optical density readings at 630 nm (OD630) are taken at 0, 24, and 48 hours as a measure of growth.

5.5.4 Transfection Assay in HEK293T Cells

Transfections with HEK293T cells (ATCC, USA) were performed as previously described [27]. Briefly, the plasmid containing wild-type hERα, pCMXhERαLBD, as well as the variants containing the L346 mutants were transfected with the reporter plasmid p17*4TATALuc, containing four Gal4 response elements controlling a *Renilla* luciferase gene as well as the pCMXβgal plasmid, containing a β-galactosidase gene under the control of a CMV promoter. The cationic lipid used was lipofectamine 2000 (Invitrogen, USA).

5.6 References

1. Chen, Z.L., et al., *Directed evolution of human estrogen receptor variants with significantly enhanced androgen specificity and affinity.* Journal of Biological Chemistry, 2004. **279**(32): p. 33855-33864.

2. Chen, Z.L. and H.M. Zhao, *Rapid creation of a novel protein function by in vitro coevolution.* Journal of Molecular Biology, 2005. **348**(5): p. 1273-1282.

3. Chockalingam, K., et al., *Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(16): p. 5691-5696.

4. D'Ursi, P., et al., *Modelling the interaction of steroid receptors with endocrine disrupting chemicals.* Bmc Bioinformatics, 2005. **6**: p. 8.

5. Ekena, K., J.A. Katzenellenbogen, and B.S. Katzenellenbogen, *Determinants of ligand specificity of estrogen receptor-alpha: Estrogen versus androgen discrimination.* Journal of Biological Chemistry, 1998. **273**(2): p. 693-699.

6. Kim, J.H., et al., *Role of aspartate 351 in transactivation and active conformation of estrogen receptor alpha.* J Mol Endocrinol, 2005. **35**(3): p. 449-64.

7. Pike, A.C.W., *Lessons learnt from structural studies of the oestrogen receptor.* Best Practice & Research Clinical Endocrinology & Metabolism, 2006. **20**(1): p. 1-14.

8. Shadnia, H., J. Wright, and J. Anderson, *Interaction force diagrams: new insight into ligand-receptor binding.* Journal of Computer-Aided Molecular Design, 2009. **23**(3): p. 185-194.

9. Shiau, A.K., et al., *The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen.* Cell, 1998. **95**(7): p. 927-937.

10. Zhao, C.Q., et al., *Mutation of Leu-536 in human estrogen receptor-alpha alters the coupling between ligand binding, transcription activation, and receptor conformation.* Journal of Biological Chemistry, 2003. **278**(29): p. 27278-27286.

11. Aliau, S., et al., *Differential interactions of estrogens and antiestrogens at the 17 beta-hydroxyl or counterpart hydroxyl with histidine 524 of the human estrogen receptor alpha.* Biochemistry, 2002. **41**(25): p. 7979-7988.

12. Islam, K.M.D., et al., *Directed evolution of estrogen receptor proteins with altered ligand-binding specificities.* Protein Engineering Design & Selection, 2009. **22**(1): p. 45-52.

13. Oehler, M.K., et al., *Functional characterization of somatic point mutations of the human estrogen receptor a (hER alpha) in adenomyosis uteri.* Molecular Human Reproduction, 2004. **10**(12): p. 853-860.

14. Zhong, L. and D.F. Skafar, *Mutations of tyrosine 537 in the human estrogen receptor-alpha selectively alter the receptor's affinity for estradiol and the kinetics of the interaction.* Biochemistry, 2002. **41**(13): p. 4209-4217.

15. Bai, Z.L. and R. Gust, *Breast Cancer, Estrogen Receptor and Ligands.* Archiv Der Pharmazie, 2009. **342**(3): p. 133-149.

16. Brzozowski, A.M., et al., *Molecular basis of agonism and antagonism in the oestrogen receptor.* Nature, 1997. **389**(6652): p. 753-758.

17. Huang, P.X., V. Chandra, and F. Rastinejad, *Structural Overview of the Nuclear Receptor Superfamily: Insights into Physiology and Therapeutics.* Annual Review of Physiology, 2010. **72**: p. 247-272.

18. Paulmurugan, R., et al., *A human estrogen receptor (ER)alpha mutation with differential responsiveness to nonsteroidal ligands: Novel approaches for studying mechanism of ER action.* Molecular Endocrinology, 2008. **22**(7): p. 1552-1564.

19. Nettles, K.W., et al., *NFkappaB selectivity of estrogen receptor ligands revealed by comparative crystallographic analyses.* Nat Chem Biol, 2008. **4**(4): p. 241-7.

20. Steinmetz, A.C.U., J.P. Renaud, and D. Moras, *Binding of ligands and activation of transcription by nuclear receptors.* Annual Review of Biophysics and Biomolecular Structure, 2001. **30**: p. 329-359.

21. Zhang, J.H. and D.S. Geller, *Helix 3-helix 5 interactions in steroid hormone receptor function.* Journal of Steroid Biochemistry and Molecular Biology, 2008. **109**(3-5): p. 279-285.

22. Zhang, J.H., et al., *A critical role of helix 3-helix 5 interaction in steroid hormone receptor function.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(8): p. 2707-2712.

23. Thin, T.H., et al., *Mutations in the helix 3 region of the androgen receptor abrogate ARA70 promotion of 17 beta-estradiol-induced androgen receptor transactivation.* Journal of Biological Chemistry, 2002. **277**(39): p. 36499-36508.

24. Bledsoe, R.K., et al., *A ligand-mediated hydrogen bond network required for the activation of the mineralocorticoid receptor.* Journal of Biological Chemistry, 2005. **280**(35): p. 31283-31293.

25. Voth, W.P., et al., *Yeast vectors for integration at the HO locus.* Nucleic Acids Research, 2001. **29**(12).

26. Tanenbaum, D.M., et al., *Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(11): p. 5998-6003.

27. Taylor, J.L., et al., *Characterization of a molecular switch system that regulates gene expression in mammalian cells through a small molecule.* BMC Biotechnology, 2010. **10**: p. 12.

28. Baker, M.E., C. Chandsawangbhuwana, and N. Ollikainen, *Structural analysis of the evolution of steroid specificity in the mineralocorticoid and glucocorticoid receptors.* Bmc Evolutionary Biology, 2007. **7**.

29. Kumar, R. and G. Litwack, *Structural and functional relationships of the steroid hormone receptors' N-terminal transactivation domain.* Steroids, 2009. **74**(12): p. 877-83.

30. Thornton, J.W., *Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(10): p. 5671-5676.

31. Tedesco, R., et al., *The estrogen receptor: a structure-based approach to the design of new specific hormone-receptor combinations.* Chemistry & Biology, 2001. **8**(3): p. 277-287.

32. Moore, D.D., et al., *International Union of Pharmacology. LXII. The NR1H and NR1I receptors: Constitutive androstane receptor, pregnene X receptor, farnesoid X receptor alpha, farnesoid X receptor beta, liver X receptor alpha, liver X receptor beta, and vitamin D receptor.* Pharmacological Reviews, 2006. **58**(4): p. 742-759.

33. Ekins, S., et al., *Evolution of pharmacologic specificity in the pregnane X receptor.* Bmc Evolutionary Biology, 2008. **8**: p. 21.

34. Moore, L.B., et al., *Orphan nuclear receptors constitutive androstane receptor and pregnane X receptor share xenobiotic and steroid ligands.* J Biol Chem, 2000. **275**(20): p. 15122-7.

35. Ekins, S., et al., *Evolution of pharmacologic specificity in the pregnane X receptor.* BMC Evolutionary Biology, 2008. **8**.

36. Pike, A.C.W., A.M. Brzozowski, and R.E. Hubbard, *A structural biologist's view of the oestrogen receptor.* Journal of Steroid Biochemistry & Molecular Biology, 2000. **74**: p. 261-268.

37. Dubey, R.K. and E.K. Jackson, *Genome and hormones: Gender differences in physiology - Invited review: Cardiovascular protective effects of 17 beta-estradiol metabolites.* Journal of Applied Physiology, 2001. **91**(4): p. 1868-1883.

38. Jan, S.T., et al., *Metabolic activation and formation of DNA adducts of hexestrol, a synthetic nonsteroidal carcinogenic estrogen.* Chem Res Toxicol, 1998. **11**(5): p. 412-9.

39. Schwimmer, L.J., et al., *Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(41): p. 14707-14712.

CHAPTER 6

NEGATIVE CHEMICAL COMPLEMENTATION WITH ANTAGONISTS

6.1 Genetic Selection and Negative Chemical Complementation

Genetic selection systems have been demonstrated to be powerful tools for discovering and deciphering protein-protein interactions and small molecule-protein interactions [1-8]. One such system, chemical complementation, was developed as a selection assay for deciphering small molecule receptor interactions and can be used for applications such as drug discovery or protein engineering [3-5, 9, 10]. Due to their significant involvement in diseases, approximately 10-15% of the pharmaceutical drugs currently on the market are targeted toward nuclear receptors either as agonists or antagonists [11]. As previously mentioned, agonists are small molecule ligands that bind and activate the nuclear receptor LBD through inducing the proper conformational change, and antagonists are small molecule ligands that do not cause the proper conformational change to occur, recruiting corepressors (CoR) to associate with the LBD and repressing transcription [12, 13]. Both agonists and antagonists play significant roles in nuclear receptor regulation, hence the importance for drug discovery in nuclear receptor based diseases.

To assess nuclear receptor antagonists for drug discovery purposes, a yeast strain, called BAPJ69-4A, previously developed in our lab, as an extension of the chemical complementation selection system, called negative chemical complementation was utilized [14]. BAPJ69-4A contains Gal4 REs controlling expression of the *URA3* gene [14, 15]. The *URA3* gene codes for the enzyme orotidine-5'-phosphate decarboxylase, responsible for the production of uracil in the uracil biosynthetic pathway [16].

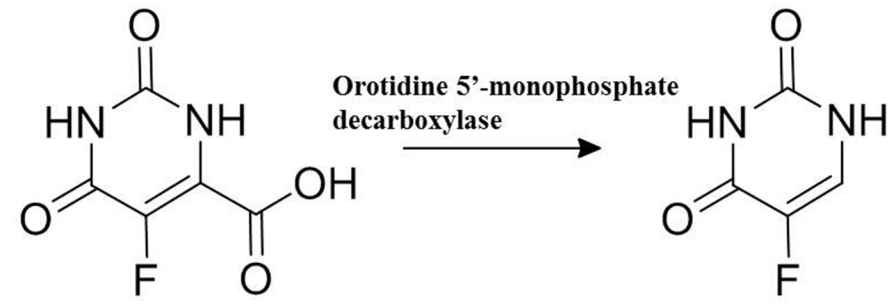


Figure 6.1. 5'Floroorotic acid (5'FOA) is converted to 5'Fluorouracil (5'FU) when the *URA3* gene is expressed.

Previously, Vidal *et al*. developed a reverse-two hybrid selection assay using the *URA3* gene to analyze small molecule inhibitor-protein interactions [17]. If the *URA3* gene is expressed in media containing 5'fluoorotic acid (5'FOA), the enzyme orotidine-5'-phosphate decarboxylase converts 5'FOA into the toxin 5-fluorouracil (5'FU), leading to cell death (Figure 6.1) [17]. With positive selection, in media lacking uracil, two associating partners, such as the nuclear receptor fusion protein and the coactivator (CoAc) fusion protein, would lead to the positive effect of cell growth. Conversely, in positive selection two dissociating partners, where the nuclear receptor and the coactivator fusion protein do not interact, would lead to the negative effect of cell death.

In negative chemical complementation, like chemical complementation, the Gal4-DBD (GBD) is fused to the nuclear receptor LBD and the nuclear receptor coactivator (CoAc) is fused to the Gal4 activation domain (GAD). When an agonist activates the nuclear receptor in selective media lacking uracil and the *URA3* gene is activated, the yeast survive (Figure 6.2A). As a result, in positive selection novel drugs or functional protein variants can be discovered through the survival of the yeast. In negative selection on uracil selective media containing 5'FOA, activation of transcription of the *URA3* gene through an agonist binding the LBD would lead to cell death (Figure 6.2B). Therefore, for protein-protein interactions or protein-ligand interactions, cell survival or cell death can be regulated based on the media used. Negative chemical complementation, therefore, can be used as a high throughput selection method for nuclear receptor drug discovery and protein and enzyme engineering.

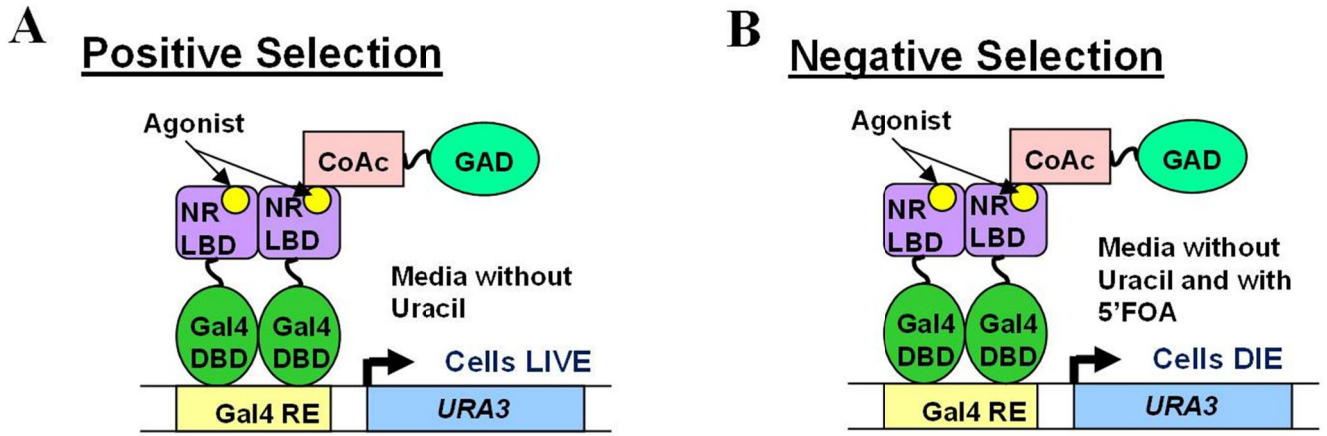


Figure 6.2. Negative chemical complementation high-throughput selection system in yeast with an agonist ligand (A) Positive selection. In positive chemical complementation, a small molecule agonist activates transcription of the *URA3* gene and the cells survive on uracil selective media. **(B) Negative selection.** In negative selection, a small molecule agonist activates the receptor causing transcription of the *URA3* gene, and in media containing 5'FOA, the toxic compound 5'FU is produced, causing cell death.

A new *Saccharomyces cerevisiae* strain, containing three selective markers *ADE2*, *HIS3*, and *URA3* as well as the *lacZ* gene controlled by Gal4 promoters was previously developed [14]. In this work, the retinoid X receptor (RXR) was characterized in negative chemical complementation with the wild-type agonist 9-*cis* retinoic acid (9cRA). The human estrogen receptor alpha (hERα) was also tested with the wild-type agonist 17-β estradiol (E2) in negative chemical complementation. Negative chemical complementation was also evaluated through analysis of hERα with the selective estrogen receptor modulator (SERM) 4-hydroxytamoxifen (OHT) (Figure 6.3) [12, 18, 19]. OHT is currently used as an antagonist for the estrogen receptor in breast cancer therapy [19]. A potent antagonist, ICI 182,780, also known as Fulvestrant, was also chosen to test with hERα in negative chemical complementation (Figure 6.3) [20]. Fulvestrant (Faslodex®) is recommended as treatment of breast cancer after OHT becomes ineffective [19].

Negative chemical complementation, like chemical complementation, can also be used as a tool for protein engineering. Chemical complementation has previously been used to engineer novel ligand-receptor pairs, selecting for nuclear receptor variants able to bind and activate transcription in response to small molecule ligands [10, 21]. One common limitation to yeast two-hybrid systems like chemical complementation, however, is the production of false-positives [17, 22, 23]. As observed with the pregnane X receptor (PXR) and the human estrogen receptor alpha (hERα) libraries in Chapters 2 and 4, many of the variants obtained were constitutively active. Therefore, one approach toward overcoming the constitutive activity in chemical complementation was to modify

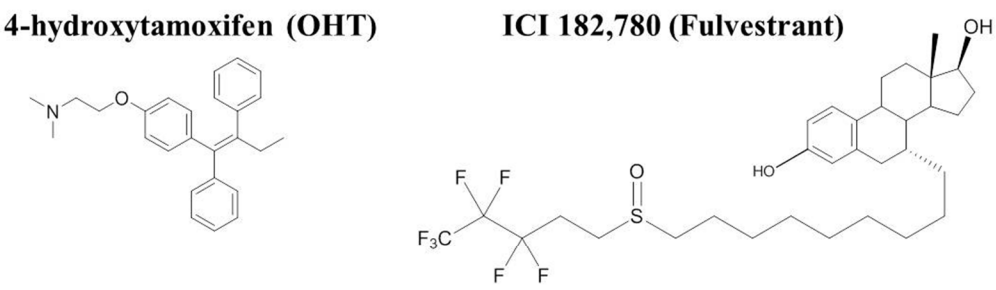


Figure 6.3. Structures of the SERM 4-hydroxytamoxifen (OHT) and antagonist ICI 182,780 (Fulvestrant). OHT is currently used as an antagonist for the estrogen receptor in breast cancer therapy. Fulvestrant is a more potent antagonist used for OHT resistant breast cancer.

the current chemical complementation transformation, such that constitutively active receptors would be eliminated from the library. To do this, 5'FOA selective media was used to remove the constitutively active variants through expression of the *URA3* gene.

Two applications for negative chemical complementation, therefore, were tested. This system can be used in a high throughput method to discover novel small molecule ligands as potential drugs for nuclear receptors. RXR and hERα were tested in negative chemical complementation with agonist ligands. More importantly, hERα was tested with the current drugs OHT and Fulvestrant to assess the use of this system for drug discovery of potential antagonists in the future. Additionally, negative chemical complementation was used toward protein engineering, through using 5'FOA selective media to remove constitutively active variants as well as engineer functional hERα variants.

6.2 Negative Chemical Complementation with Agonists

As previously mentioned, in positive selection with uracil selective media, when an agonist activates the nuclear receptor leading to activation of the *URA3* gene, the yeast survive (Figure 6.2A). In negative selection with uracil selective media containing 5'FOA, activation of the *URA3* gene leads to the formation of 5'FU and cell death (Figure 6.2B).

6.2.1 Negative Chemical Complementation with RXR

As mentioned in Chapters 2, 4, and 5, prior to testing the nuclear receptors in yeast, the plasmids were transformed into BAPJ69-4A. The yeast expression plasmid containing the Gal4 DNA binding domain (GBD) fused to the RXR LBD (pGBDRXR), containing a tryptophan marker, was co-transformed with a plasmid containing a coactivator fusion protein. The activator for thyroid hormone and retinoid receptors

(ACTR), a nuclear receptor coactivator, is fused to the Gal4 activation domain (GAD) on a yeast expression plasmid with a leucine marker. As a positive control, a plasmid containing the *holo* Gal4 gene (pGBT9Gal4), containing a tryptophan marker, was utilized. The transformants were then tested in negative chemical complementation in synthetic complete media lacking uracil (SC-ULW) and SC-ULW media containing 5'FOA.

As expected, Gal4 displays growth in the uracil selective media, but no growth is observed in the uracil selective media containing 5'FOA, due to the production of 5'FU (Figure 6.4, closed orange diamond and open orange diamond, respectively). RXR displays a dose response in positive selection with uracil selective media containing 9cRA, with an EC_{50} value of approximately 300 nM, based on growth (Figure 6.4, closed cyan square). In uracil selective media containing 5'FOA, cell death is observed at higher concentrations of 9cRA as a result of expression of the *URA3* gene, showing a negative effect from the association of the two fusion proteins caused by the presence of the ligand (Figure 6.4, open cyan square). At the highest concentration of 9cRA (10 μM), no growth is observed (Figure 6.4, open cyan square). These data are consistent with previous work [14].

6.2.2 Negative Chemical Complementation with hERα

Next, hERα was tested in the BAPJ69-4A strain in uracil selective media and uracil selective media containing 5'FOA with E2 concentrations ranging from 10 μM to 1 pM of E2. The plasmid containing the fusion of the GBD with the hERαLBD alongside the fusion of the GAD with coactivator were transformed into BAPJ69-4A. The

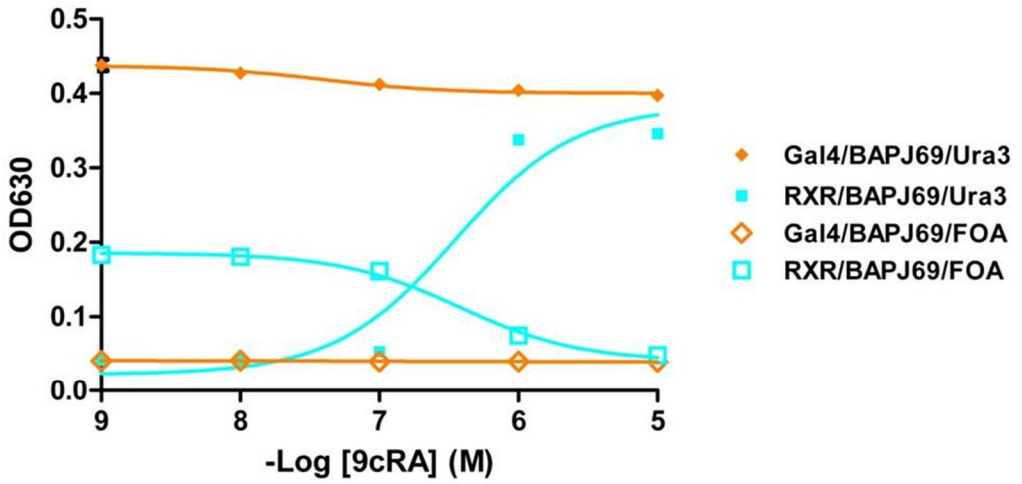


Figure 6.4 Negative chemical complementation assay with RXR and Gal4 in uracil selective media and uracil selective media containing 5'FOA. Activation profiles for RXR in uracil selective media with 9cRA showing an EC_{50} value of 300 nM. In uracil selective media containing 5'FOA with 9cRA, cell death occurs at the same concentration as growth is seen in the uracil selective media without 5'FOA.

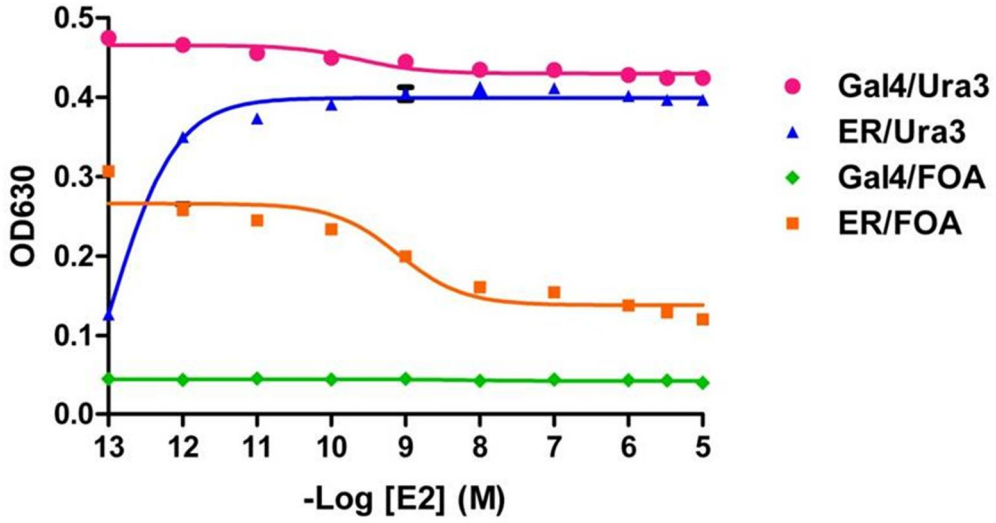


Figure 6.5. Negative chemical complementation assay with hERα. In uracil selective media with E2 hERα shows agonist activation with E2 as low as 10^{-12} M. In uracil media containing 5'FOA, hERα shows cell death with increasing concentrations of E2 displaying an IC_{50} value of 5 nM.

coactivator in this case consists of repeats of the peptide LXXLL of the steroid receptor coactivator-1 (SRC-1). As mentioned in Chapter 4, SRC-1 peptide motifs were shown to be effective in activating transcription, therefore the smaller coactivator peptide was chosen for analysis with hERα [24].

Gal4 displays growth in the uracil selective media without ligand and no growth is observed in the uracil selective media containing 5'FOA, consistent with the data seen in Figure 6.4 (Figure 6.5, pink circle and green diamond, respectively). As expected, in the uracil selective media, cells containing hERα show growth with the wild-type agonist E2 at concentrations as low as 1 pM, having a 3-fold activation (Figure 6.5, blue diamond). Conversely, in uracil selective media containing 5'FOA, cell death is observed as the concentration of E2 increases (Figure 6.5, orange square). Basal growth levels are observed in the uracil selective media without E2, due to the low expression levels seen with the *URA3* gene in the absence of a transcription factor [17, 25, 26].

Both RXR and hERα confirm the functionality of the *URA3* gene in uracil selective media and 5'FOA selective media with an agonist. BAPJ69-4A had not, however, been tested in negative chemical complementation with an antagonist.

6.3 Negative Chemical Complementation with Antagonists: Applications in Drug Discovery

6.3.1 hERα Analysis with the SERM 4-Hydroxytamoxifen

An antagonist for RXR was not readily available, therefore, the wild type hERα was tested in negative chemical complementation with the SERM OHT, an ER antagonist currently used to treat breast cancer [18]. When testing an antagonist in the liquid assay, the yeast are grown in selective media with a constant concentration of the E2 agonist

with increasing concentrations of the antagonist. Structurally, the binding of the antagonist displaces the agonist, resulting in an inactive conformation of the nuclear receptor, leading to the recruitment of corepressors [27-29]. In negative chemical complementation using the BAPJ69-4A yeast strain, therefore, cell death is expected in uracil selective media in the presence of an antagonist due to lack of expression of the *URA3* gene (Figure 6.6A). Whereas, in uracil selective media containing 5'FOA, cell growth is expected because the toxic compound 5'FU is no longer being produced (Figure 6.6B).

When tested with OHT in the uracil selective media with hERα displays a dose response with an EC_{50} value of approximately 10 μM, indicating that, in chemical complementation, OHT is acting as an agonist (Figure 6.7, closed blue triangle). In addition, when tested in uracil selective media containing 5'FOA with OHT, cell death occurs at the higher concentrations of OHT, confirming the agonist activity of OHT in the yeast system (Figure 6.7, open blue triangle). As previously mentioned, OHT is a SERM, meaning this compound can function as either an agonist or antagonist depending on the tissue-type [19]. Therefore, to determine if OHT functioned as an agonist in the presence of E2, hERα was tested with OHT in negative chemical complementation with a constant concentration of 300 pM E2, the known *in vitro* EC_{50} value for hERα [30]. When 300 pM E2 is added to the uracil selective media with various concentrations of OHT, growth is observed for hERα at all concentrations of OHT, also indicating that OHT is not functioning as an antagonist (Figure 6.8, closed cyan square). In addition, Figure 6.8 also shows an additive effect of agonist activity is seen in the uracil selective

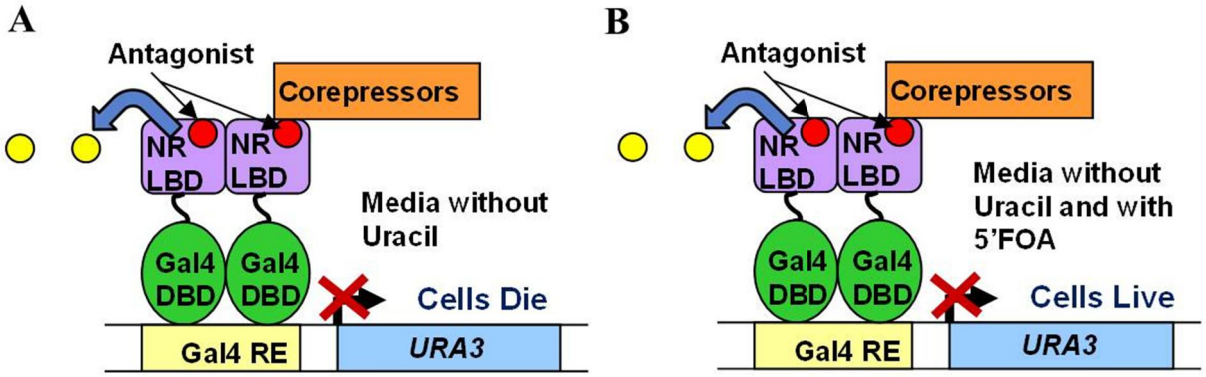


Figure 6.6 Negative chemical complementation with an antagonist ligand. (A) In uracil selective media, an antagonist (red circle) displaces the agonist (yellow circle), inactivating the *URA3* gene causing the yeast to die. (B) In the presence of an antagonist in uracil selective media containing 5'FOA, transcription of the *URA3* gene is repressed, inhibiting the conversion of 5'FOA to 5'FU, allowing the cells to survive in media containing 5'FOA.

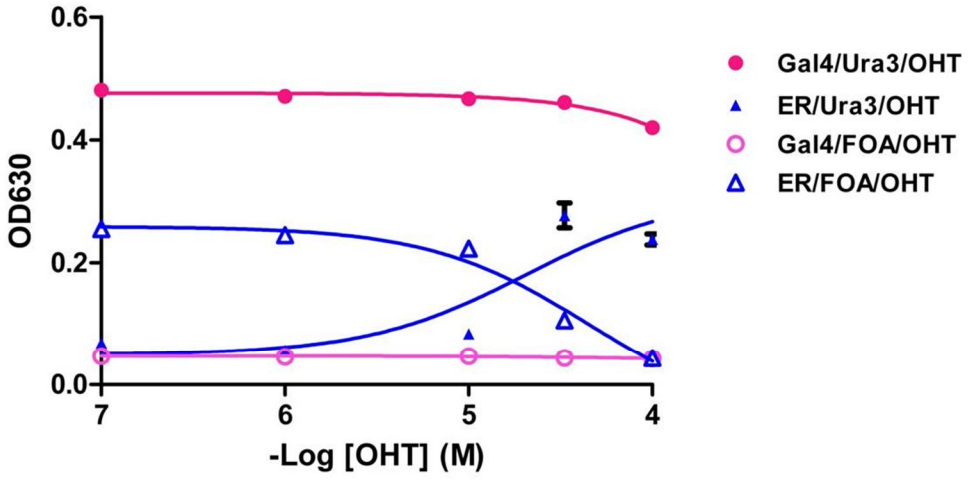


Figure 6.7. Negative chemical complementation activation profiles for hERα and Gal4 in uracil selective media with OHT. In uracil selective media with OHT, hERα shows agonist activity with an EC_{50} of 10 µM. In selective media containing 5'FOA with OHT, cell death occurs at higher concentrations of OHT for hERα.

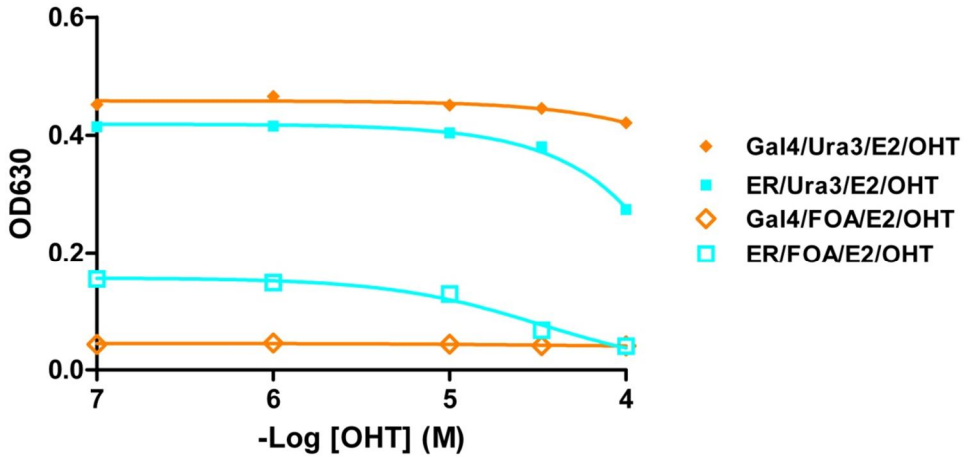


Figure 6.8. Negative chemical complementation activation profiles for hERα and Gal4 in selective media with 300 pM E2 with increasing concentrations of OHT. In uracil selective media containing 300 pM E2 and increasing concentrations of OHT, hERα shows growth at all conditions (closed cyan square). Activation profile for hERα in uracil selective media containing 5'FOA and 300 pM E2 with increasing concentrations of OHT showing cell death at higher concentrations of OHT (open cyan square). An additive affect is seen with the two agonists present in the media shown by the lower growth level.

media with 5'FOA containing a constant concentration of 300 pM E2 showing less growth than the uracil selective media only with OHT, as shown in Figure 6.7 (Figure 6.8, open cyan square). OHT could be functioning as an agonist because only one coactivator is used in the chemical complementation system, as opposed to a combination of coactivators found in mammalian cells. In summary, OHT functioned as an agonist in yeast and therefore did not show ligand activated growth in uracil selective media containing 5'FOA, a necessary component for antagonist drug discovery using negative chemical complementation.

6.3.2 hERα with the Antagonist Fulvestrant

Fulvestrant has a much stronger affinity for ER than OHT, and is used to treat OHT resistance breast cancers [20, 31]. Moreover, Fulvestrant has a similar affinity for ER as E2, and therefore, when treating breast cancer, competes with E2 [31]. The goal was that, as a more potent antagonist, Fulvestrant would function as an antagonist in the yeast system. As shown in Figure 6.6B, the desired effect is cell growth in uracil selective media containing 5'FOA selective media in the presence of an antagonist.

6.3.2.1 hERα in HEK293T Mammalian Cell Culture with Fulvestrant

To assess the antagonistic ability of Fulvestrant with the GBD-hERαLBD construct, hERα was tested with Fulvestrant in HEK293T mammalian cells (Figure 6.9). hERα shows activity when tested with E2, having an EC_{50} of 17 pM (Figure 6.9, pink circle). When tested with Fulvestrant in the presence of 300 pM E2, hERα activity is inhibited with an IC_{50} of 50 pM, confirming the strength of Fulvestrant as an antagonist (Figure 6.9, blue triangle).

6.3.2.2 hERα in Yeast with Fulvestrant

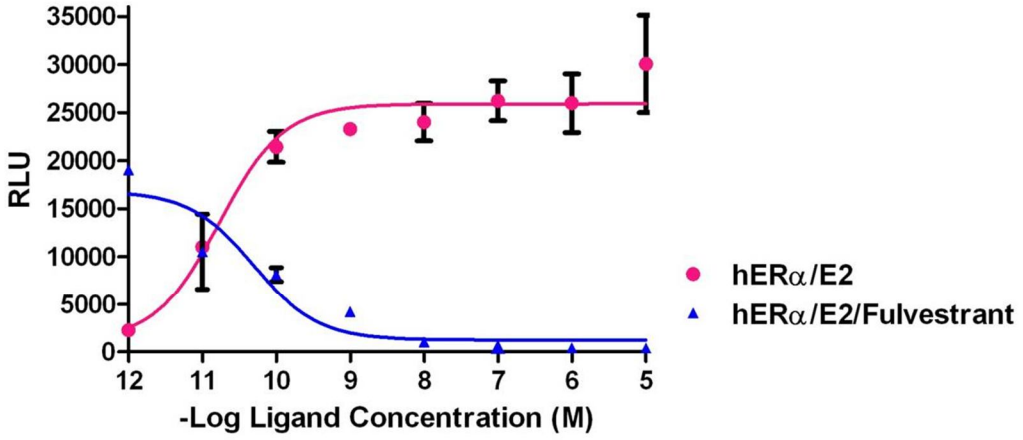


Figure 6.9. Gal4DBD-hERαLBD fusion protein in HEK293T cells with E2 and Fulvestrant. hERα is activated by E2 having an EC_{50} value of 17 pM and inhibited by Fulvestrant with an IC_{50} of 50 pM.

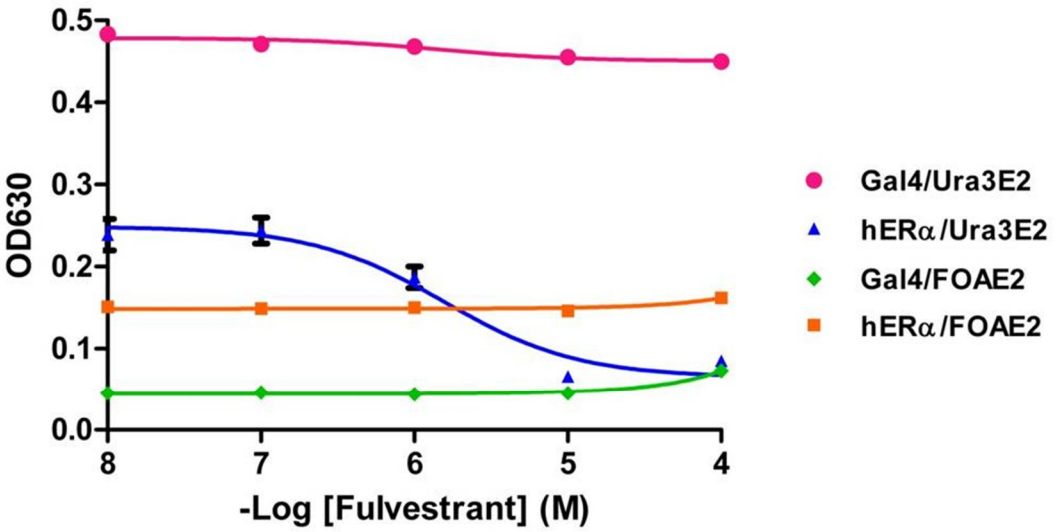


Figure 6.10. Negative chemical complementation with hERα and Gal4 in BAPJ69-4A. Uracil selective media with 300 pM E2 and increasing concentrations of Fulvestrant show antagonist activity (blue diamond). Uracil selective media containing 5'FOA with 300 pM E2 and increasing concentrations of Fulvestrant show no growth above basal (orange square).

Due to the agonist response with OHT in negative chemical complementation, Fulvestrant was tested with hERα. In uracil selective media containing 300 pM E2 and increasing concentrations of Fulvestrant, antagonistic activity is seen, showing cell death with higher concentrations of Fulvestrant (Figure 6.10, blue triangle). These data confirm that Fulvestrant functions as an antagonist in uracil selective media. In 5'FOA selective media, however, cell growth in the presence of an antagonist, as hypothesized in Figure 6.6B, should be observed. On the other hand, in 5'FOA selective media, no growth is observed above basal, indicating cell survival as would be expected with an antagonist (Figure 6.10, orange square). One hypothesis for the inability of the yeast to show growth above basal with of the addition of antagonist, is that the yeast are surviving on the small amount of uracil present in the 5'FOA media, thus the yeast could potentially efflux the ligands out of the cells, since the small molecule is not needed for survival. In summary, hERα in the negative chemical complementation selection system showed cell death in the presence of an antagonist in uracil selective media. However, in 5'FOA selective media, cell growth was not observed in the presence of an antagonist above basal growth.

6.4 Using Negative Chemical Complementation for Protein Engineering

Another application of negative chemical complementation is to subject protein libraries to selection for protein engineering. Unlike chemical complementation, negative chemical complementation has an advantage of removing eliminating constitutively active variants using the 5'FOA selective media because the constitutively active variants would lead to activation of the *URA3* gene, producing the toxic 5'FU compound. Therefore hERα was tested in negative chemical complementation with the goal of

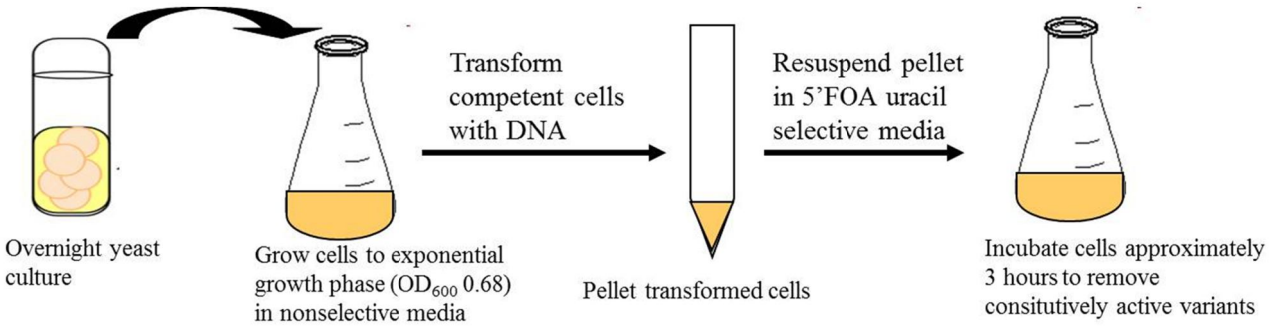


Figure 6.11. Scheme of Using Negative Chemical Complementation for Protein Engineering. An overnight yeast culture containing the coactivator fusion plasmid was grown and used to inoculate a larger culture, which was growth to competency. The competent yeast cells were then transformed with the randomized nuclear receptor DNA. The transformed cells are then pelleted and re-grown in uracil selective media containing 5'FOA to remove the constitutively active variants.

Table 6.1 Sequencing from negative chemical complementation transformation

Selective Variants	Sequencing
17	M342I
18	H474H
26	T347A, R352K, F367Y, V392I, I510V
27	Wild-type hERα
28	N359I, L378V, G390D
33	F425del, S433P, H476R, K492R, 532del, 533ins, L539L
Nonselective Variants	
3	H476H, I482L, V533V
4	L379L, V422M, L489P

del: deletion, ins: insertion

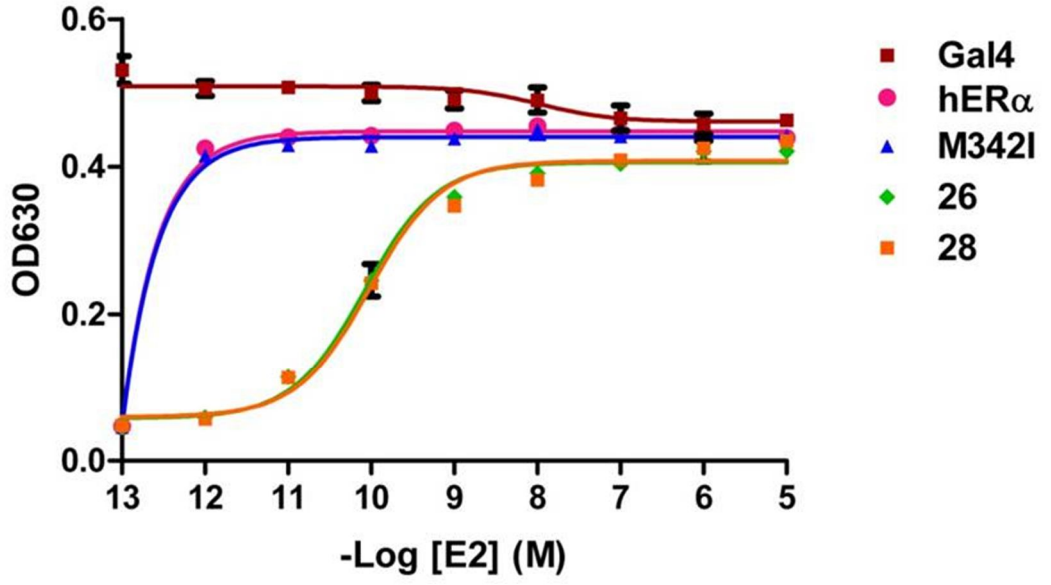


Figure 6.12. Activation profiles in yeast for hERα variants M342I, 26 (T347A, R352K, F367Y, V392I, I510V) and 28 (N359I, L378V, G390D) in uracil selective media with E2.

removing constitutively active variants as well as produce functional ligand activation receptors.

As mentioned in Chapter 4, error-prone PCR, one method of random mutagenesis that introduces mutations into a gene, can be used for protein engineering. Error-prone PCR was performed on hERα and the cassette was transformed into competent yeast cells with the linearized background plasmid. The transformed cells were then incubated in uracil selective media containing 5'FOA to remove constitutively active variants (Figure 6.11). After the three hour incubation period, the yeast were plated onto uracil selective plates containing 5'FOA with 300 pM E2 and potential antagonists for hERα (e.g. Fulvestrant) as well as uracil selective plates with potential agonists and incubated at 30 °C for 4 days. A plate with synthetic complete media lacking leucine and tryptophan (SC-LW) was also used to assess the library size.

The transformation produced a library size of 1.68×10^3. Thirty-six mutants total grew on the uracil selective plates and the uracil selective plates containing 5'FOA. All were streaked for constitutive activity onto uracil selective plates (SC-ULW) and SC-LW nonselective plates. Only 5 of the colonies analyzed were constitutively active. Several colonies were rescued and sequenced (Table 6.1). Three variants were discovered that display ligand activated growth with E2 (Figure 6.12). One variant, M342I, had the same growth levels as wild-type, whereas the other two variants, 26 and 28, displayed decreased sensitivity toward E2 (Figure 6.12, blue triangle, green diamond, and orange square, respectively). This method removed approximately 90 % of the constitutively active variants, but many of the variants tested were non-functional.

6.5 Nuclear Receptors and Corepressors

Previous yeast two-hybrid systems have been developed and successfully used to engineer proteins with novel functions [6, 32, 33]. This work and the data shown in previous chapters demonstrate how chemical complementation and negative chemical complementation, using the PJ69-4A and BAPJ69-4A strains, respectively, can be used for protein engineering. In addition, both chemical complementation and negative chemical complementation are useful for positive selection with agonists, providing an effective tool for the advancement of drug discovery. To use negative chemical complementation toward the advancement of antagonist drug discovery, therefore, the system was modified by using corepressor proteins instead of the current CoAc:GAD fusion protein. Two examples of corepressors that nuclear receptors interact with are the silencing mediator of retinoid and thyroid hormone receptors (SMRT) and the nuclear receptor corepressor (NCoRI) [28, 34, 35]. As mentioned in Chapter 1, in the absence of ligand, most nuclear receptors are bound to corepressor proteins. Corepressors are also able to associate with nuclear receptors when antagonists are bound, due to the altered conformation of the receptor in the presence of antagonist [12, 28, 36].

Both PJ69-4A and BAPJ69-4A show ligand activated growth in positive selection as seen in this chapter as well as Chapters 2, 4, and 5. Growth above basal was not observed in the negative selection system using uracil selective media with 5'FOA with an antagonist (Figure 6.10, orange square). Therefore, the goal was to change the system back to a positive selection method, where gene activation results in growth.

In the new system, corepressors are fused to the activation domains to allow cell growth in the presence of an antagonist. The yeast corepressor TUPI will also be tested in the modified negative chemical complementation system to determine the effects a yeast

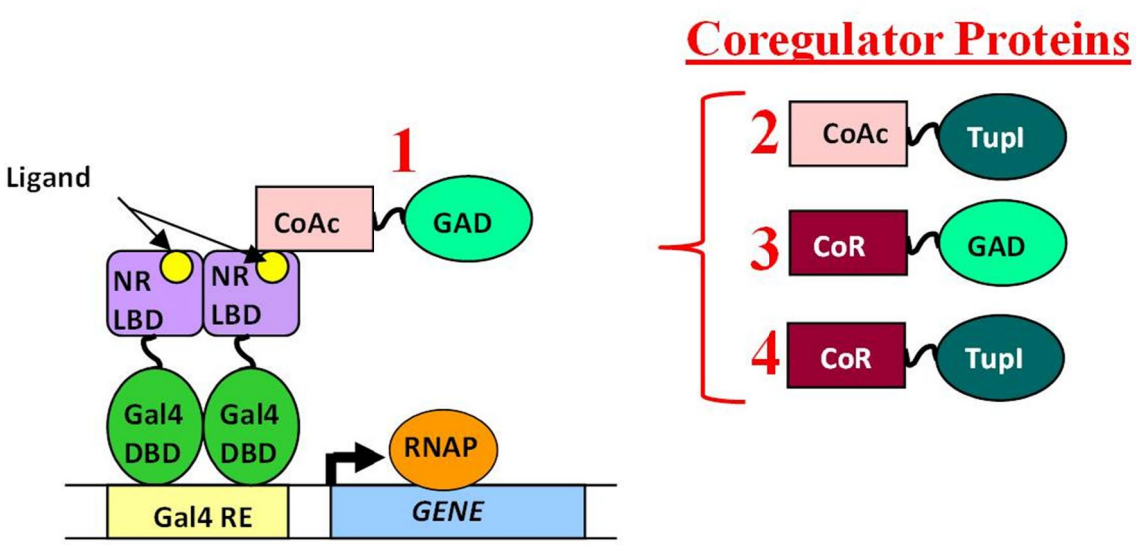


Figure 6.13. Fusion proteins used in the modified negative chemical complementation selection system. 1 is the current system uses both human an yeast coactivators (1). 2 is a human coactivator fused to a yeast repressor . 3 has a human corepressor fused to a yeast activator and 4 has both corepressors fused together.

repressor has on cell growth and cell death. TUPI has previously been used to inhibit transcription in yeast in a two-hybrid system [37-39]. Here, the design of the negative chemical complementation corepressor selection system and the work done to date on this project is discussed.

6.5.1 Developing Negative Chemical Complementation with Corepressors

Several fusion proteins were used in the modified negative chemical complementation selection system (Figure 6.13). Figure 6.13 shows the current chemical complementation system using the human CoAc:GAD fusion construct as one option to regulate transcription of the selective gene (Figure 6.13). Three additional coregulator fusion proteins for use in chemical complementation are shown with different combinations of the yeast and human coactivators and corepressors (Figure 6.13). The second construct combines the human coactivator ACTR and yeast repressor TUPI, which can be used as a control to assess transcriptional repression using a yeast repressor. When an agonist activates the nuclear receptor, the ACTR:TUPI fusion is recruited and transcription of the *URA3* gene is repressed by the TUPI leading to cell death. The third construct consists of the human corepressor NCoRI fused to the yeast GAD (Figure 6.13). In this system, when an antagonist binds the nuclear receptor, the nuclear receptor changes to the inactive conformation, thus allowing the human corepressor, NCoRI, to associate with the LBD. The GAD can turn on transcription of the selective gene, resulting in ligand activated growth where the ligand is an antagonist. As a control, the final construct contains both corepressors, having NCoRI fused to TUPI (Figure 6.13). As previously mentioned, when an antagonist binds, the NCoRI will associate with the LBD. This brings TUPI into proximity with the DNA, causing repression of transcription.

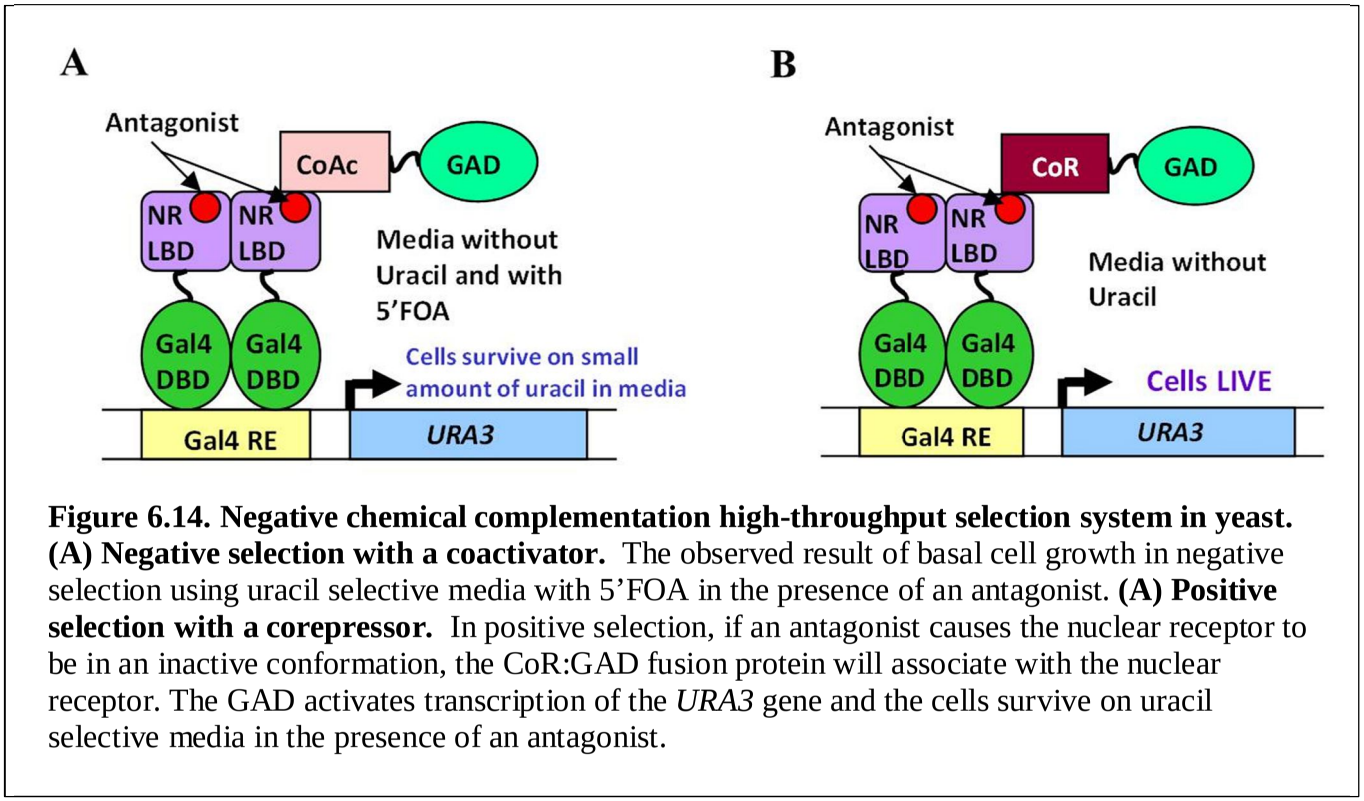


Figure 6.14. Negative chemical complementation high-throughput selection system in yeast. (A) Negative selection with a coactivator. The observed result of basal cell growth in negative selection using uracil selective media with 5'FOA in the presence of an antagonist. **(A) Positive selection with a corepressor.** In positive selection, if an antagonist causes the nuclear receptor to be in an inactive conformation, the CoR:GAD fusion protein will associate with the nuclear receptor. The GAD activates transcription of the *URA3* gene and the cells survive on uracil selective media in the presence of an antagonist.

As mentioned above, the current negative chemical complementation system only shows basal growth levels, not the expected increase in dose response above basal, in uracil selective media with 5'FOA in the presence of an antagonist (Figure 6.14A). The objective of the modified system is to return to a system with ligand-dependent growth, where the ligand is an antagonist. As seen in Figure 6.14B, in the presence of an antagonist, the NcoRI:GAD fusion protein will associate with the nuclear receptor. Since the corepressor is fused to the GAD, in uracil selective media, the GAD will recruit yeast transcriptional machinery to turn on the selective *URA3* gene and the yeast will survive (Figure 6.14B). The aim, therefore, is to use the NCoRI:GAD fusion protein in the system to cause ligand activated growth with an antagonist.

Conversely, the CoAc:TUPI fusion protein previously made in our lab has ACTR as the coactivator [40]. With the ACTR:TUPI construct as the coregulator protein, in the presence of an agonist in uracil selective media, the hypothesis is that cell death would occur as a result of ACTR being recruited and the TUPI yeast respressor turning off transcription of the *URA3* gene. In 5'FOA selective media with an agonist, however, repression of the *URA3* gene would inhibit formation of 5'FU, and allow the yeast to survive on the small amount of uracil present in the 5'FOA media.

6.5.1.1 hERα with ACTR:TUPI in Negative Chemical Complementation with E2

Gal4 and hERα were tested in negative chemical complementation with the ACTR:TUPI fusion protein in uracil selective media as well as uracil selective media with 5'FOA with increasing concentrations of E2. As a control, the Gal4 and hERα with the SRC-1:GAD fusion protein was tested in the same conditions. As previously seen with hERα and the SRC-1:GAD fusion protein, ligand activated growth or cell death is

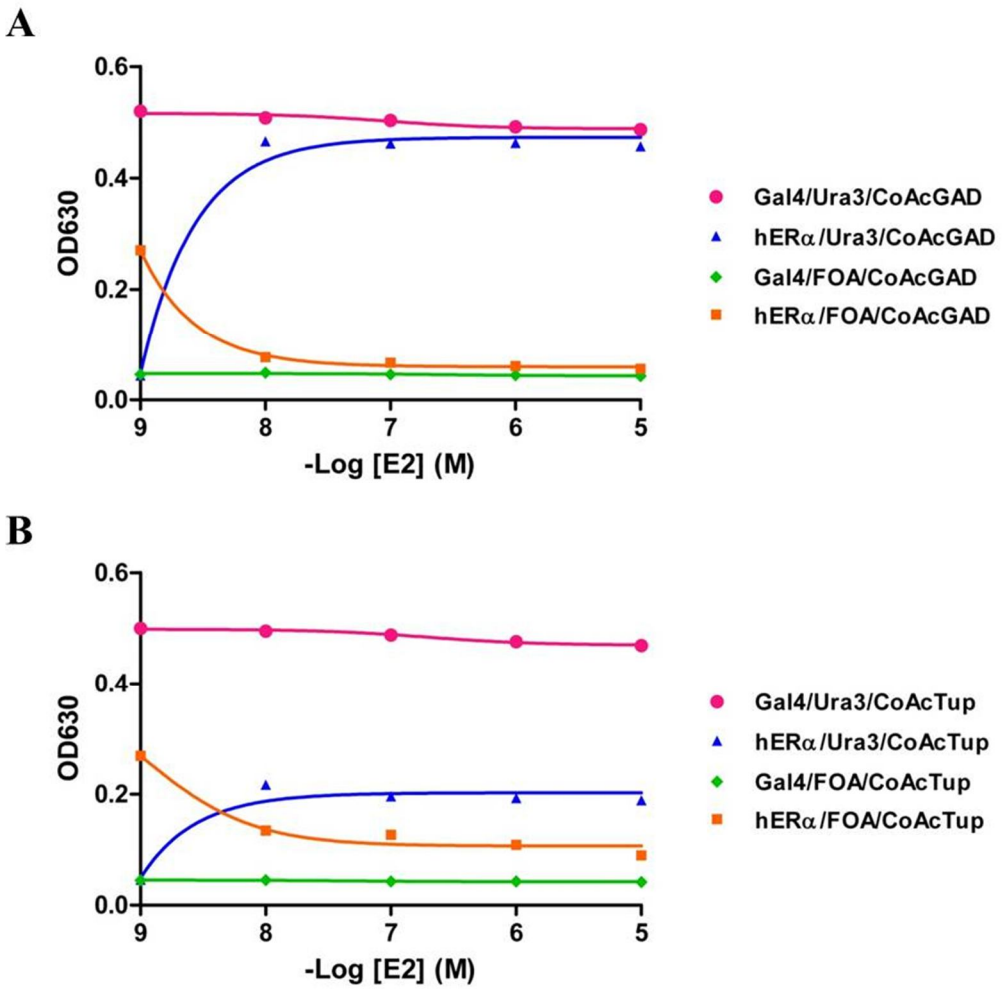


Figure 6.15. Negative chemical complementation with hERα and Gal4 with the CoAc:TUPI fusion protein in BAPJ69-4A. (A) hERα with CoAc:GAD displays cell growth in uracil selective media with E2 and cell death in uracil selective media containing 5'FOA (orange line). These data are consistent with Figure 6.5. (B) hERα with CoAC:TUPI in uracil selective media with displays decreased cell growth (blue line), indicating some repression through the TUPI fusion protein. In uracil selective media containing 5'FOA, cell death is observed.

observed in uracil selective media and uracil selective media with 5'FOA, respectively (Figure 6.15A, blue diamond and orange square, respectively). Gal4 displays growth with and without ligand in uracil selective media and does not grow in uracil selective media containing 5'FOA (Figure 6.15A, pink circle and green diamond, respectively). Gal4 growth levels are consistent among all coregulator constructs. When tested with the ACTR:TUPI construct in uracil selective media with E2, hERα shows ligand activated growth, however, the growth levels are reduced by approximately 50 % in comparison to the growth observed with the SRC-1:GAD construct (Figure 6.15B, blue triangle). This indicates that TUPI is leading to some repression in the negative chemical complementation system, but the expected complete cell death does not occur. In uracil selective media with 5'FOA, the expected result for hERa with the ACTR:TUPI construct is for the yeast to survive on the small amount of uracil present in the 5'FOA media, showing basal growth levels with and without ligand. As seen in Figure 6.15B, basal cell growth is observed without E2 ligand, and cell death is observed with the addition of E2 to the media (Figure 6.15B, orange square). The cell death in the presence of E2 could be due to the agonist turning on expression of the *URA3* gene independent of the interaction with the ACTR:TUPI protein, by perhaps indirectly recruiting yeast coactivators. This indicates that perhaps the human coactivator is capable of indirectly recruiting yeast coactivators to initiate transcription. In conclusion, the ACTR:TUPI coregulator fusion protein decreases cell growth levels by approximately half the level seen with the SRC-1:GAD in the presence of an agonist in uracil selective media, but does not fully turn off transcription.

6.5.1.2 hERα with ACTR:TUPI in Negative Chemical Complementation with Fulvestrant

To compare repression with the ACTR:TUPI construct to the previous results using the SRC-1:GAD construct with the antagonist Fulvestrant in uracil selective media containing 300 pM E2, hERα with ACTR:TUPI was assessed using the same media conditions. In the presence of an antagonist, the ACTR:TUPI coregulator protein would not interact with the nuclear receptor. Cell death is not observed with increasing concentrations of Fulvestrant, as was observed with the SRC-1:GAD coregulator in Figure 6.10 (Figure 6.16, blue triangle). The observed cell growth could occur through yeast transcriptional coactivators causing transcription of the selective gene independent of the ACTR:TUPI fusion protein.

In uracil selective media with 5'FOA, hERα with the ACTR:TUPI is consistent with the SRC-1:GAD system in Figure 6.10, showing basal growth levels with and without the presence of antagonist (Figure 6.16, orange square). Interestingly, one additional difference between the coactivator with TUPI and the coactivator with GAD is the type of human coactivator. The TUPI construct contains ACTR, whereas the GAD construct contains the SRC-1 peptide motifs. Thus far, a coregulator protein containing the SRC-1 peptide motifs fused to the TUPI has not been made. The constant growth seen in Figure 6.16 with the ACTR:TUPI construct, therefore, may be due to the fact that the coactivator is ACTR instead of the SRC-1 peptide motifs. Therefore, hERα with the ACTR:GAD construct was tested in uracil selective media with 300 pM E2 and increasing concentrations of Fulvestrant. In contrast to the antagonistic effect seen with hERα with the SRC-1:GAD construct, hERα with ACTR:GAD does not display cell death as a result of the antagonist, indicating the influence of the different coactivator in the yeast system (Figure 6.17, blue triangle). In uracil selective media with 5'FOA, hERα

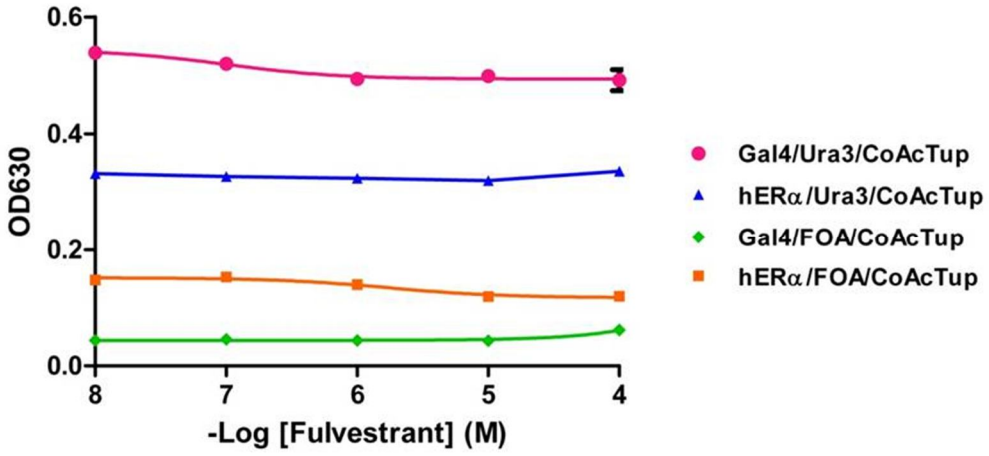


Figure 6.16. Negative chemical complementation with hERα and Gal4 in BAPJ69-4A with ACTR:TUPI construct. hERα in uracil selective media with 300 pM E2 and increasing concentrations of Fulvestrant does not display cell death (blue triangle) as was observed with the CoAc:GAD fusion protein in Figure 6.10. In uracil selective media containing 5'FOA with 300 pM E2 and increasing concentrations of Fulvestrant hERα does not display above basal (orange square).

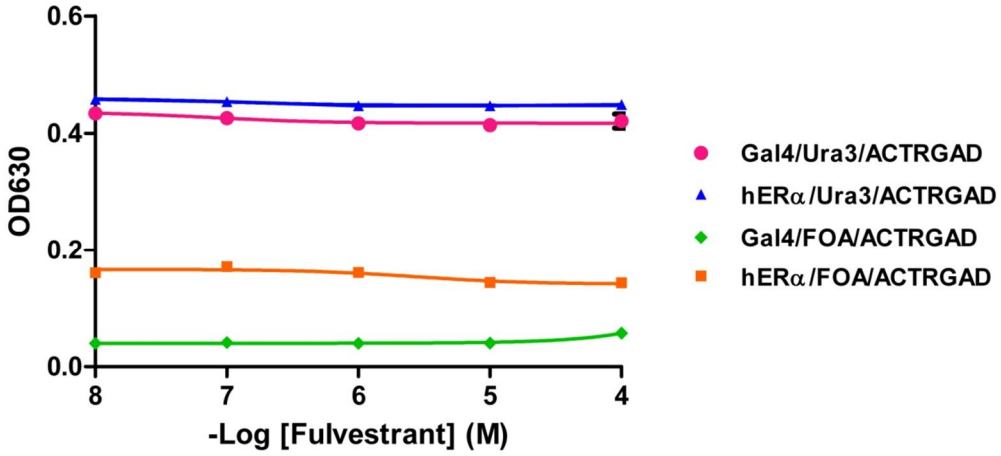


Figure 6.17. Negative chemical complementation with hERα and Gal4 in BAPJ69-4A with GAD:ACTR construct. hERα in uracil selective media with 300 pM E2 and increasing concentrations of Fulvestrant does not show the expected antagonistic activity (blue triangle) as was observed with the CoAc:GAD fusion protein in Figure 6.10. In uracil selective media containing 5'FOA with 300 pM E2 and increasing concentrations of Fulvestrant hERα shows basal growth levels (orange square).

with the ACTR:GAD, basal growth levels are seen, consistent with ACTR:TUPI and SRC-1:GAD in 5'FOA selective media (Figure 6.17, orange square). As previously mentioned, this is most likely due to the yeast surviving on the uracil present in the 5'FOA media.

6.5.1.3 hERα with NCoRI:GAD in Negative Chemical Complementation with Fulvestrant

The focus of assessing hERα with corepresors toward antagonist drug discovery was to test hERα in negative chemical complementation with the NCoRI:GAD fusion protein. As previously mentioned, cell growth in the presence of an antagonist had not been observed in the negative chemical complementation system. As seen in Figure 6.14B, the expected result with the NCoRI:GAD construct in uracil selective media with an antagonist is cell growth as a result of the antagonist bound nuclear receptor recruiting the corepressor and the GAD turning on expression of the *URA3* gene (Figure 6.14B).

hERα with NCoRI:GAD was next tested in uracil selective media as well as 5'FOA selective media containing 300 pM E2 with increasing concentrations of Fulvestrant. Cell death, not the expected cell growth, is observed for hERα in uracil selective media containing E2 with and without antagonist (Figure 6.18, blue triangle). The cell death could be occurring because the NCoRI:GAD fusion protein is not associating with the nuclear receptor and therefore the GAD is not in proximity to turn on expression of the *URA3* gene. Furthermore, the NCoRI protein has not been tested for proper expression or folding in yeast. As previously seen, basal growth levels are observed for hERα in uracil selective media containing 5'FOA (Figure 6.18, cyan

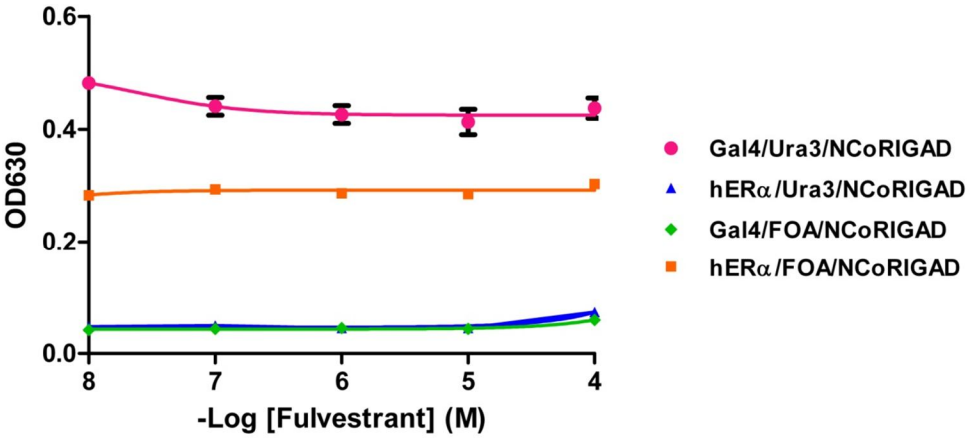


Figure 6.18. Negative chemical complementation with hERα and Gal4 with the NCoRI:GAD fusion protein in BAPJ69-4A. hERα with NCoRI:GAD in uracil selective media with 300 pM E2 and increasing concentrations of Fulvestrant do not display ligand dependent-growth (blue triange). Uracil selective media containing 5'FOA with 300 pM E2 and increasing concentrations of Fulvestrant show basal growth with and without the antagonist ligand ligand (cyan square).

square). In this case, the yeast do not need the NCoRI:GAD construct to turn on the *URA3* gene and therefore, the cells survive on the uracil present in the 5'FOA media (Figure 6.18, cyan square). In summary, ligand-dependent growth with an antagonist was not observed with hERα using the NCoRI:GAD construct.

6.6 Summary and Future Work

The goal was to assess negative chemical complementation for drug discovery and protein engineering applications. RXR and hERα were tested in uracil selective media and 5'FOA selective media with the agonist ligands 9cRA and E2, respectively. Both displayed cell growth in uracil selective media and cell death in FOA selective media. hERα was next tested with the SERM OHT and the antagonist Fulvestrant to assess negative chemical complementation with current drugs. OHT functioned as an agonist, also displaying cell growth in uracil selective media and cell death in 5'FOA selective media. When hERα was tested in uracil selective media containing 300 pM E2, the yeast displayed cell death with increasing concentrations of the antagonist Fulvestrant. However, in uracil selective media containing 5'FOA with 300 pM E2, growth above basal was not observed in uracil selective media with increasing concentrations of the antagonist. Both methods allow selection of small molecule-protein and protein-protein interactions, which can be used for drug discovery.

Negative chemical complementation was then tested to be used for protein engineering applications. Although the method using incubation in 5'FOA media removed most constitutively active variants as well as produced ligand activated variants, the system is inefficient in generating large libraries of functional protein variants.

To use negative chemical complementation toward antagonist drug discovery, the system was modified to use corepressor proteins as opposed to the current coactivator system. hERα with the CoAc:TUPI construct displayed some growth repression in uracil selective media with E2, but the cell death was not complete. When hERα was tested with the NCoRI:GAD construct in uracil selective media containing 300 pM E2, the yeast did not display cell growth in the presence of the antagonist Fulvestrant. Therefore, the current system with the corepressors cannot be used for antagonist drug discovery. The absence of cell growth could be due to improper folding of the NCoRI protein within the yeast. An additional method to continue toward obtaining growth in the presence of an antagonist, however could be to use additional corepressors, such as SMRT, instead of NCoRI.

6.7 Materials and methods

6.7.1 Strains

The genotype of the BAPJ69-4A strain is *MATa trp1-901 leu2-3, 112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ ho::GAL1-URA3* [14, 15].

6.7.2 Plasmids and PCR Reactions

pGBDhERaLBD and pGAD10SRC-1LXXLL plasmids were made as described in Chapter 4.

The plasmid pGBDhERαLBD was used as the template DNA for error-prone PCR random mutagenesis as described in Chapter 4. The PCR reaction contained 0.5 μM of each primer, 500 μM dNPTs, 7 nM $MgCl_2$, 250 ng of template DNA, 1x *Taq* buffer,

and *Taq* polymerase (Promega), and 20 μM $MnCl_2$. The PCR program was 95 °C for 1 minute, 95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 1 minute, repeated 20 times followed by 72 °C for 2 minutes.

A plasmid containing the human coactivator ACTR and yeast repressor TUPI, designated pTupACTR, was previously engineered in our lab [40]. To make the pGADNCoRI plasmid, QuikChange™ (Stratagene) site-directed mutagenesis was used to insert a *PacI* site into pGAD10BASRC-1LXXLL plasmid previously made in our lab to generate the pGADSRC1LXXLLPacI plasmid [14]. The mouse NCoRI gene was amplified from YEmNCoRI using PCR with the following primers: NCoRIPacIFor 5' ccttaattaaatgtcaagttcaggttatcctccca 3' and NCoRINotIrev 5' aaggaaaaaagcggccgc tcagtcgtcactatcagacagtgtctc 3' yielding a cassette approximately 1800 bp. The plasmid pGADSRC1LXXLLPacI as well as the NCoRI cassette were digested with *PacI* and *NotI*, and ligated to create the pGADNCoRI plasmid. The plasmid was confirmed by sequencing (Operon, Huntsville, AL).

6.7.3 Ligands

17-β estradiol (MW= 272.4 g/mol) was purchased from MP Biomedicals (Aurora, OH). Tamoxifen (371.5 g/mol) and Fulvestrant (MW=606.77 g/mol) were purchased from Sigma-Aldrich (St. Louis, MO). 9-*cis* retinoic acid (MW=300.44 g/mol) was purchased from ICN Biomedicals (Aurora, OH). 10 mM stocks of each ligand were dissolved in 80% ethanol:20% DMSO and stored at 4°C.

6.7.4 Liquid Yeast Growth Assay

The yeast growth assay was performed as done in chapter 4 and as previously described [10]. For uracil selective media containing, 5'FOA, a solution of 0.2 % w/v

5’FOA was added to selective media in a 1:1 ratio to produce media containing 0.1 % w/v 5’FOA as the final concentration, as previously described [41].

6.7.5 Yeast Transformation using 5’FOA

Competent cells were transformed with the cassette containing from the error-prone PCR product and the linearized background vector using the lithium acetate method [42]. The transformed cells were then grown in uracil selective media containing 5’FOA at 30 °C with shaking at 300 rpm until the cell density (OD600) was approximately 0.68, indicating that the cells were in exponential growth. Cells were then plated onto uracil selective plates containing 5’FOA with 300 pM E2 and either Fulvestrant or other potential hERα antagonists and uracil selective plates with potential agonist ligands and incubated at 30 °C for 4 days. A plate with synthetic complete media lacking leucine and tryptophan (SC-LW) was also used to assess the library size.

6.7.6 Transfection in HEK293T Cells

The transfection assay was performed as in Chapter 5 and as previously described [43].

6.8 References

1. Fields, S. and O.K. Song, *A novel genetic system to detect protein-protein interactions.* Nature, 1989. **340**(6230): p. 245-246.

2. McDonnell, D.P., E. Vegeto, and M.A.G. Gleeson, *Nuclear hormone receptors as targets for new drug discovery.* Bio-Technology, 1993. **11**(11): p. 1256-1261.

3. Woycechowsky, K.J. and D. Hilvert, *Deciphering enzymes - Genetic selection as a probe of structure and mechanism.* European Journal of Biochemistry, 2004. **271**(9): p. 1630-1637.

4. Lin, H.N., H.Y. Tao, and V.W. Cornish, *Directed evolution of a glycosynthase via chemical complementation.* Journal of the American Chemical Society, 2004. **126**(46): p. 15051-15059.

5. Taylor, S.V., P. Kast, and D. Hilvert, *Investigating and engineering enzymes by genetic selection.* Angewandte Chemie-International Edition, 2001. **40**(18): p. 3310-3335.

6. Drees, B.L., *Progress and variations in two-hybrid and three-hybrid technologies.* Current Opinion in Chemical Biology, 1999. **3**(1): p. 64-70.

7. Uetz, P., et al., *A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.* Nature, 2000. **403**(6770): p. 623-627.

8. Licitra, E.J. and J.O. Liu, *A three-hybrid system for detecting small ligand-protein receptor interactions.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(23): p. 12817-12821.

9. Azizi, B., E.I. Chang, and D.F. Doyle, *Chemical complementation: small-molecule-based genetic selection in yeast.* Biochemical and Biophysical Research Communications, 2003. **306**(3): p. 774-780.

10. Schwimmer, L.J., et al., *Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(41): p. 14707-14712.

11. Chen, T.S., et al., *Coactivators in assay design for nuclear hormone receptor drug discovery.* Assay and Drug Development Technologies, 2003. **1**(6): p. 835-842.

12. Shiau, A.K., et al., *The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen.* Cell, 1998. **95**(7): p. 927-937.

13. Smith, C.L. and B.W. O'Malley, *Coregulator function: A key to understanding tissue specificity of selective receptor modulators.* Endocrine Reviews, 2004. **25**(1): p. 45-71.

14. Azizi, B., *Chemical complementation: A genetic selection system for drug discovery, protein engineering, and deciphering biosynthetic pathways (PhD)*. 2005, Georgia Institute of Technology: Atlanta, GA.

15. Shaffer, H.A., B. Azizi, and D.F. Doyle, *BAPJ69-4A: A Yeast Two-Hybrid Strain for Both Positive and Negative Genetic Selection.* To be submitted.

16. Boeke, J.D., F. Lacroute, and G.R. Fink, *A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance.* Molecular & General Genetics, 1984. **197**(2): p. 345-346.

17. Vidal, M., et al., *Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(19): p. 10315-10320.

18. Dai, S.Y., et al., *Prediction of the tissue-specificity of selective estrogen receptor modulators by using a single biochemical method.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(20): p. 7171-7176.

19. Bai, Z.L. and R. Gust, *Breast Cancer, Estrogen Receptor and Ligands.* Archiv Der Pharmazie, 2009. **342**(3): p. 133-149.

20. Wakeling, A.E., M. Dukes, and J. Bowler, *A potent specific pure antiestrogen with clinical potential.* Cancer Research, 1991. **51**(15): p. 3867-3873.

21. Chockalingam, K., et al., *Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(16): p. 5691-5696.

22. Bruckner, A., et al., *Yeast two-hybrid, a powerful tool for systems biology.* International Journal of Molecular Sciences, 2009. **10**(6): p. 2763-2788.

23. Zeng, J., et al., *Genome wide screens in yeast to identify potential binding sites and target genes of DNA-binding proteins.* Nucleic Acids Research, 2008. **36**(1): p. 8.

24. McInerney, E.M., et al., *Determinants of coactivator LXXLL motif specificity in nuclear receptor transcriptional activation.* Genes & Development, 1998. **12**(21): p. 3357-3368.

25. Aparicio, O.M. and D.E. Gottschling, *Overcoming Telomeric silencing a transactivator competes to establish gene expresion in a cell cycle dependent way.* Genes & Development, 1994. **8**(10): p. 1133-1146.

26. Voth, W.P., et al., *Yeast vectors for integration at the HO locus.* Nucleic Acids Research, 2001. **29**(12).

27. Love, J.D., et al. *Transcriptional repression by nuclear receptors: mechanisms and role in disease*. 2000: Portland Press.

28. Nagy, L., et al., *Mechanism of corepressor binding and release from nuclear hormone receptors.* Genes & Development, 1999. **13**(24): p. 3209-3216.

29. Pike, A.C., et al., *Structural aspects of agonism and antagonism in the oestrogen receptor.* Biochem Soc Trans, 2000. **28**(4): p. 396-400.

30. Kostelac, D., G. Rechkemmer, and K. Briviba, *Phytoestrogens modulate binding response of estrogen receptors alpha and beta to the estrogen response element.* Journal of Agricultural and Food Chemistry, 2003. **51**(26): p. 7632-7635.

31. McKeage, K., M.P. Curran, and G.L. Plosker, *Fulvestrant: a review of its use in hormone receptor-positive metastatic breast cancer in postmenopausal women with disease progression following antiestrogen therapy.* Drugs, 2004. **64**(6): p. 633-48.

32. Peralta-Yahya, P., et al., *High-throughput selection for cellulase catalysts using chemical complementation.* Journal of the American Chemical Society, 2008. **130**(51): p. 17446-17452.

33. Gillies, A.R., G. Skretas, and D.W. Wood, *Engineered systems for detection and discovery of nuclear hormone-like compounds.* Biotechnology Progress, 2008. **24**: p. 8-16.

34. McKenna, N.J. and B.W. O'Malley, *Combinatorial control of gene expression by nuclear receptors and coregulators.* Cell, 2002. **108**(4): p. 465-474.

35. Privalsky, M.L., *The role of corepressors in transcriptional regulation by nuclear hormone receptors.* Annual Review of Physiology, 2004. **66**: p. 315-360.

36. Guenther, M.G., O. Barak, and M.A. Lazar, *The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3.* Molecular and Cellular Biology, 2001. **21**(18): p. 6091-6101.

37. Hirst, M., et al., *A two-hybrid system for transactivator bait proteins.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(15): p. 8726-8731.

38. Tzamarias, D. and K. Struhl, *Functional dissection of the yeast Cyc8-Tup1 transcriptional co-repressor complex.* Nature, 1994. **369**(6483): p. 758-61.

39. Joshi, P.B., et al., *Identification of protein interaction antagonists using the repressed transactivator two-hybrid system.* Biotechniques, 2007. **42**(5): p. 635-644.

40. Johnson, K., *Extending chemical complementation to bacteria and furthering nuclear receptor based protein engineering and drug discovery (PhD)*. 2009, Georgia Institute of Technology: Atlanta.

41. Ausubel, F.M., et al., eds. *Short Protocols in Molecular Biology*. Fourth ed. 1999, John Wiley & Sons, Inc.

42. Gietz, R.D. and R.A. Woods, *Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method*, in *Guide to Yeast Genetics and Molecular and Cell Biology, Pt B*. 2002, Academic Press Inc: San Diego. p. 87-96.

43. Taylor, J.L., et al., *Characterization of a molecular switch system that regulates gene expression in mammalian cells through a small molecule.* BMC Biotechnology, 2010. **10**: p. 12.