CHEMICAL COMPLEMENTATION: A GENETIC SELECTION SYSTEM FOR DRUG DISCOVERY, PROTEIN ENGINEERING AND DECIPHERING BIOSYNTHETIC PATHWAYS

A Thesis
Presented to
The Academic Faculty

By

Bahareh Azizi

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

Georgia Institute of Technology
August 2005

CHEMICAL COMPLEMENTATION: A GENETIC SELECTION SYSTEM FOR DRUG DISCOVERY, PROTEIN ENGINEERING AND DECIPHERING BIOSYNTHETIC PATHWAYS

Approved By:

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

Dr. Mostafa El-Sayed
School of Chemistry and Biochemistry
Georgia Institute of Technology

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

Dr. Jung Choi
School of Biology
Georgia Institute of Technology

Dr. Allen. Orville
School of Chemistry and Biochemistry
Georgia Institute of Technology

Date Approved: June 27, 2005

In the land of repute, our passage they will dispute
If this will not suit, don't stay mute, and transform dictates of fate.

When destitute and in need, let your love and passion breed
Life's alchemy, essence and seed, unimagined wealth shall create.

- Persian Poet, Hafez, Ghazal 5

I dedicate this work to my parents, Mohammed Reza Azizi and Zahra Barez, for their hard work and sacrifices for my education. Thank you for your endless love and support, and all that you have taught me throughout my life. I could not have completed this work if it were not for your hard work and all of your love. Thank you for teaching me the beauty and value of an education, and for believing in me always. *Mochakaram, Mochakaram, Mochakaram.*

ACKNOWLEDGEMENTS

Throughout my graduate school years, I have been blessed to have the opportunity to meet people that have touched my life and from whom I have gained scientific and general knowledge which have allowed me to grow as an individual. First of all, I would like to thank my advisor, Dr. Donald Doyle, for his guidance and assistance throughout these years. I have learned so much under his mentorship and thank him for the support of a mentor, as well as a friend. I will cherish the friendship that we have developed throughout my life. I would like to thank my committee members: Dr. Sheldon May, Dr. Allen Orville, Dr. Mostafa El-Sayed, and Dr. Jung Choi for their support during my graduate school years. I would like to thank my group members for all the help and fun that we have in our lab. I especially have to thank Eileen Chang, for being my right hand, especially for her help with testing of the RXR variants. I would like to thank Priyanka Rohatgi for the training in the cell culture room and for being such a good friend. I would like to thank Lauren Schwimmer for her help with modeling and for synthesizing LG335. I would like to thank Kenyetta Johnson for being more than a good friend, but a sister to me, and for her continual help in editing and listening to my presentations. Also, thanks go to Terry Watt, Jesse Kumar, Francis Swain, and Micheal Yeager for all of their useful discussions about science and life. Many thanks go to members of the Bommarius group, starting with Dr. Andreas Bommarius, for his support of my research throughout these years. I would like to thank Karen Polizzi and Tracey Thaler for all of their help, and also for the good times and talks that we have had. I have to thank Dr. Philip Gibbs for the challenging scientific discussions that we have had and

our great friendship. I would like to thank Dr. Farrokh Ayazi for showing me the beauty of research and inspiring me to go to graduate school. I gratefully thank the School of Chemistry and Biochemistry for their continual support of my work throughout these years, especially Dr. Cam Tyson and Dr. Robert Braga. My family has been my strength throughout these years, especially my mother and father who have always inspired me to dream big and work hard. Thank you for all the support. My sister, Banafsheh, is the sunshine in my life and has always been by my side and I can't thank her enough. Many thanks go to my aunt Fariba Barez, for her caring and long-distance calls. I would like to thank Mike and Kathy Broering, and my extended family in Mercer County, Ohio for their support and love. Finally, last but certainly not least, to my husband, James Broering, whom is an angel sent to me. Thank you for your endless love, patience, and understanding. Thank you for believing in me and pushing me when I could no longer push myself. This work would not have been possible without your love and support. As for the future, *Ay Khodayeh Man, beh omideh Shoma.*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

ACTR	Activator for Thyroid Hormone and Retinoid Receptors
AD	activation domain
AR	Human Androgen Receptor
CAR	Human Constitutive Androstane Receptor
DBD	DNA Binding Domain
ERβ	rat Estrogen Receptor β
FOA	5'-fluoroorotic acid
FU	5'-fluorouracil
FXR	Human Farnesoid X Receptor
GAD	Gal4 Activation Domain
LBD	Ligand Binding Domain
LiAc	Lithium Acetate
LXR	Human Liver X Receptor
PPAR γ	Mouse Peroxisome Proliferator Activated Receptor γ
PXR	Human Pregnane X Receptor
SF-1	Bovine Steriodgenic Factor-1
SRC-1	Steroid Receptor Coactivator-1
3-AT	3-amino 1,2,4-triazole
9cRA	9-*cis* retinoic acid

SUMMARY

Chemical complementation is a general system for detecting protein-small molecule interactions, and linking that interaction to genetic selection in yeast. Chemical complementation provides a general system for detecting a small molecule for a protein, such as a receptor or an enzyme. Here, chemical complementation is used to identify the interaction of a small molecule ligand with a nuclear receptor. In first generation chemical complementation, a two-component assay was developed where the Gal4 DNA-binding domain is fused to the ligand binding domains of nuclear receptors, and expressed in the *S. cerevisiae* strain PJ69-4A. The Gal4 DNA binding domain binds to a Gal4 response element controlling transcription of a selective marker, and the nuclear receptor ligand-binding domain binds its ligand. This system was developed using the retinoid X receptor (RXR), the pregnane X receptor (PXR), and the liver X receptor (LXR) and their ligands, 9-*cis* retinoic acid, paclitaxel, and oxysterols respectively. Yeast survive on selective plates only in the presence of two components: a nuclear receptor and the corresponding ligand. Growth was observed at the highest concentration of ligand (10^{-5} M) and the growth density was less than Gal4-activated growth. The growth time was also longer than Gal4-activated growth.

The 2nd generation chemical complementation system is a three-component system comprising a nuclear receptor ligand-binding domain fused to the Gal4 DNA binding domain, and a nuclear receptor coactivator fused to the yeast Gal4 activation domain. After the ligand binds to the nuclear receptor ligand-binding domain, a conformational change recruits the nuclear receptor coactivator fusion protein, inducing

transcription via RNA polymerase. This system has been developed using the retinoid X receptor and 9-*cis* retinoic acid and has been extended to and tested with a several other nuclear receptors. The sensitivity of chemical complementation is increased 1000-fold, and growth time and density are equivalent to Gal4 activated growth. An assay of chemical complementation was developed to provide a quantitative high-throughput assay for evaluating nuclear receptor-ligand interactions, and measuring EC_{50} values for the ligand-receptor pairs. A correlation between the functions of nuclear receptors in yeast and mammalian cell assays is observed, therefore, chemical complementation an attractive system for developing and evaluating nuclear receptor-ligand interactions.

Chemical complementation can be extended to a variety of applications. The first application is for drug discovery, because nuclear receptors are implicated in a number of diseases, ranging from metabolic disorders to a variety of cancers, and are currently targets for ~10% of commonly prescribed drugs and potentially many more. This system provides a high-throughput assay for the discovery of potential nuclear receptor ligands, such as agonists, and with the use of negative chemical complementation, the discovery of nuclear receptor antagonists. Drug discovery assays may be extended to enzyme targets by engineering receptors that activate transcription in response to the small molecule product of the enzyme-catalyzed reaction. In protein engineering applications, chemical complementation offers a general method of engineering receptors that activate transcription in response to arbitrary small molecules. This system uses the power of genetic selection to analyze libraries of protein variants in a more time and effort efficient method, allowing the survival of the variants with the desired functions. The discovery of these engineered protein-ligand pairs can serve as candidates for small molecule based

gene regulation, with applications in gene therapy. Other applications of engineered receptors extend to biotechnological applications, such as biosensors, and for engineering enzymes with enhanced or novel functions. Finally, with the discovery of more and more natural products, chemical complementation can serve as a tool for deciphering and assembling the biosynthetic pathways of these natural products, and as a tool for creating heterologous hosts for the production of natural products. This application will advance the use of rare natural products for therapeutic applications.

Chapter 1

The Power of Genetic Selection

1.1. Genetic Selection

Genetic selection is one of the most powerful tools and techniques used in multiple aspects of molecular biology for evaluating macromolecular interactions [1-7]. The general idea behind genetic selection is that the host cell will only survive if the desired entity, such as a desired function of the macromolecule, is present inside the cell [1-7]. Genetic selection can be applied toward evaluating macromolecular interactions or other cellular interactions, such that the desired interactions or desired entities are linked to cell survival.

With the more recent advancements of protein and enzyme engineering, genetic selection provides advantages for analyzing and discovering proteins with novel functions [1,7-9]. Two methods can be used to analyze large combinatorial protein libraries: genetic screens and genetic selection [1,7-9]. In genetic screening techniques, the interactions of macromolecules (or of the entity of interest) are not linked to cell survival; every member of the library needs to be evaluated to determine whether the desired function is present. In terms of protein libraries, if the best screening protocol is used to evaluate a protein library of interest using a chromogenic substrate, the maximum number of library members assayed will be about 10^5 members. The presence of undesired members of the library lowers the signal to noise ratio, increasing the amount of effort and time required to evaluate a large protein library [1,7,8].

On the other hand, in genetic selection techniques, conditions are such that only the desired members of the library will survive, eliminating all the unwanted members of

the library and in some ways compared with true Darwinian evolution [8]. The signal to noise ratio increases, and much less time and effort is put forth towards evaluating the members of a library. Using a selection method in *Escherichia coli* (*E. coli*), will allow the evaluation of about 10^{10} clones, based on an optimal transformation efficiency and technique [10,11]. One problem with selection-based strategies is in developing a generalizable selection method for various applications. With the advances in genetic selection techniques and with the numerous microbial strains present, more genetic selection in vivo systems can be developed, providing effective tools for the creation of novel proteins with various functions, greatly advancing protein and enzyme engineering [1,7,8,12,13].

Traditionally, genetic selection has been widely used in techniques such as cloning. For introducing exogenous DNA into bacteria, these pieces of DNA are cloned into circular pieces of DNA, known as vectors, which can be transcribed in vivo using the host cell's transcription machinery [1,8]. The vector containing the desired piece of DNA contains an antibiotic resistance gene, such as the ampicillin resistance gene (Amp^R). This antibiotic resistance gene when transcribed and translated will allow the microbe to survive on medium containing the antibiotic. When the vector is introduced into the host organism, by a process called transformation, not every cell will successfully take-up the vector DNA, thus the transformation is plated onto medium containing the antibiotic. Only the cells that have successfully taken-up the vector, these cells will survive on the medium containing the antibiotic, because they contain the antibiotic resistance gene [1].

Other genetic selection systems have been developed for evaluating a number of macromolecular interactions, such as protein-protein, protein-DNA, and protein-RNA interactions [14-18]. The remainder of this chapter will discuss classical genetic selection techniques; such as the yeast two-hybrid assay and more recent derivatives of this assay that have been developed to analyze protein-small molecule interactions.

1.2. Yeast Two-Hybrid System

Protein networks and interactions are key to the function of cellular processes, providing the physiological basis for organisms. Proteins form large complexes, which interact with other large complex networks of macromolecules, such as deoxyribonucleic acids (DNAs), ribonucleic acids (RNAs), and other proteins. With advances in genome projects of different species, particularly with the advances in the human genome project, a demand is made for a system to evaluate protein-protein interactions, as well as determining protein function. Toward this end, one development is the yeast two-hybrid system.

The yeast two-hybrid assay was developed by Fields and Song, and since the birth of the system more than a decade ago, there have been over 3000 publications that have used this system for characterizing protein/protein interactions [10,11,19-21]. This assay provides a general system for analyzing protein function, finding novel partners for a particular protein, and for discovering and deciphering protein-protein networks. The original system developed was tailored towards protein-protein interactions, however, a number of laboratories have developed derivatives of the system, adapting the system to

detect the interactions between proteins and other macromolecules, such as RNA, and other small molecule ligands [12,13,20-23].

The basic concept of the two-hybrid assay is the creation of a transcriptional activator based on the fusion of two proteins. The creation of a transcriptional activator consists of two protein hybrids that need to be brought in close proximity to one another, as shown in Figure 1.1 [19,20]. The first hybrid is a fusion of the protein of interest (Protein X) to a DNA binding domain (DBD); this entire hybrid protein is called the "bait". In the second hybrid, the other protein of interest (Protein Y) is fused to a transcriptional activation domain (AD) creating a fusion protein referred to as "prey". Through the association of protein X and Y, the DBD and AD are brought into close proximity to one another, activating transcription of the genes, which are under the control of a promoter containing a binding site for the DBD (Figure 1.1) [19]. For this system to be efficient, the bait and prey cannot activate transcription independently, and can only activate transcription in the presence of each other [19]. The genes, which are under the control of these promoters, can range from colorimetric markers, such as the *lacZ* gene that encodes for an enzyme that can cleave its substrate and form a color change, or a selective marker, such as the *HIS3* or *ADE2* gene[10,11,24]. These markers allow for yeast growth on medium lacking histidine or adenine respectively, and provide a cell death/growth diagnostic assay [11,19,20].

As mentioned above, a wide range of applications and crucial protein/protein interactions have been identified by using the two-hybrid assay and the impact of this

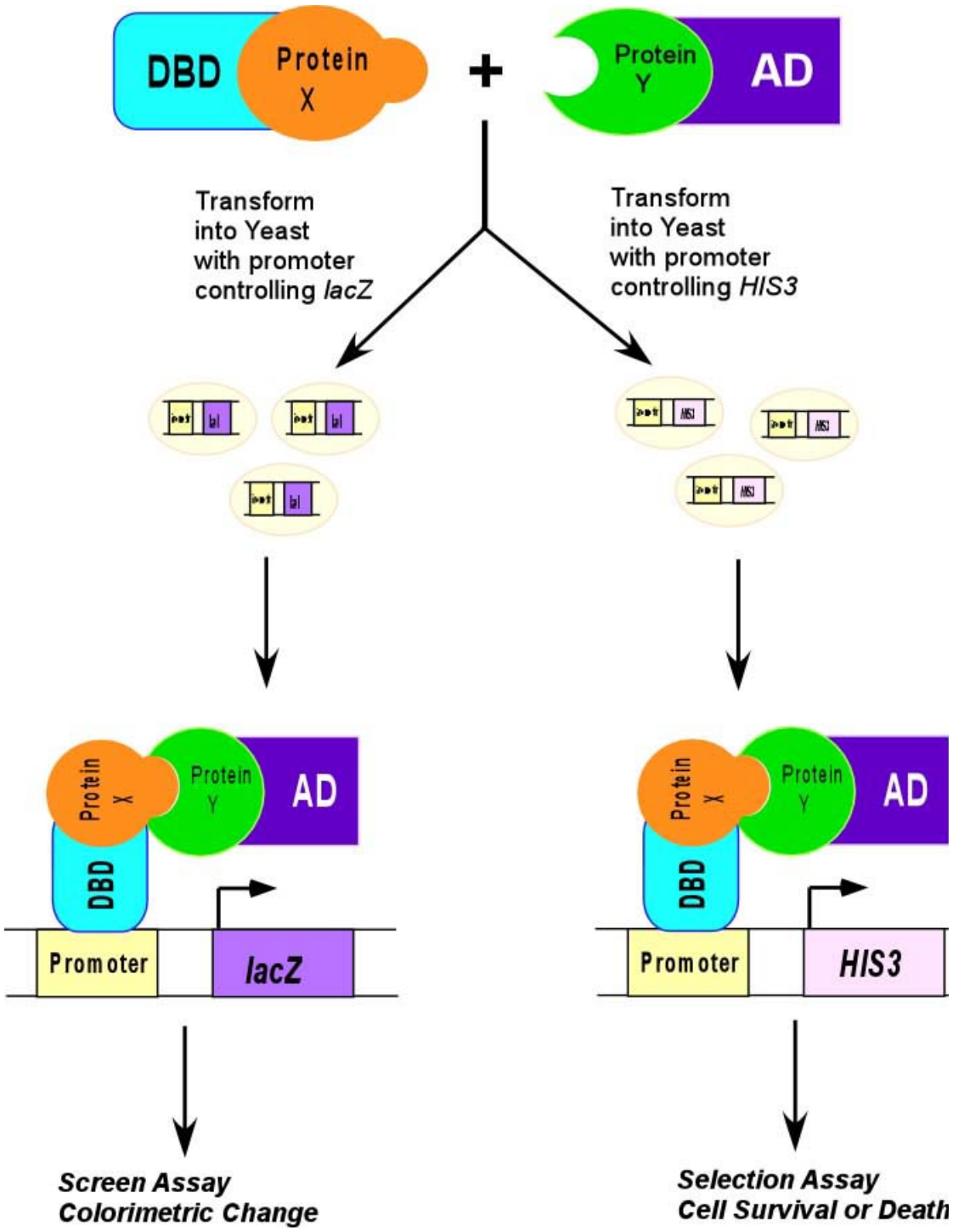


Figure 1.1. Yeast two-hybrid assays using selection or screen markers.

system in identifying the mechanism of physiological protein interactions is tremendous. One recent application of the yeast two-hybrid assay involves creating an automated version of this assay for the identification of all possible protein/protein interactions in *S. cerevisiae* (yeast) [20]. In this work, Uetz and coworkers screened every open reading frame encoding a protein (there are about 6000 in yeast) against itself. Each open reading frame was fused to both a Gal4 DBD and a Gal4 AD, but to transform all of 6000 combinations of DBD and AD fusions would be extremely challenging. Uetz and coworkers took advantage of the yeast mating strains and mated cells that had a combination of DBD with those cells containing a library of AD, and vice versa. Cells that survived the selection pressure were then selected and further tested [20].

Despite the versatility and applicability of the yeast two-hybrid system, several limitations do exist. In some cases, growth of the yeast cell may be observed but the two proteins or entities do not interact, creating the presence of "false positives". On the other hand, "false negatives" may be present, where the lack of growth may not always indicate that the proteins or entities are not interacting, and could be caused by other factors, such as protein expression. Although these limitations are present, the yeast two-hybrid assay still provides a powerful system for evaluating the interaction between two entities.

1.3. Yeast Three- Hybrid Assays

The advantage of yeast three hybrid assays

Recently, derivatives of the yeast two-hybrid assay have been developed, extending the traditional role of this assay from detecting protein-protein interactions to detecting protein-RNA or protein-small molecule interactions, forming "yeast three-hybrid assays" [15,16,22,23]. In these examples, the small molecule or RNA molecule serves as a bridge linking two proteins, each fused to the respective DBD or AD (Figure 1.2) [25]. Using three-hybrid assays for evaluating protein-ligand interactions provides a feasible and relatively easy method for evaluating macromolecular interactions. Using genetic selection provides an alternative solution to *in vitro* or *in vivo* mammalian cell culture assays, which are more expensive and time-consuming.

Yeast three-hybrid assays for protein-small molecule interactions

Small molecules can be used to regulate gene expression, providing a powerful tool for use in clinical applications such as gene therapy. For gene therapy to become a useful tool, pharmacological control of the transgene to be introduced into the human body needs would be beneficial and the ability to finely adjust the levels of transgene expression will be highly advantageous [26]. One idea behind the regulation of the transgene expression is through the use of a small molecule, where only in the presence of the small molecule will expression of the gene occur. Using small molecules as regulating gene therapy will also provide a dose-regulation, such that as the concentration of the small molecule varies, the amount of expression would vary accordingly. Small molecules also provide the ability to create and synthesize a range of these molecules with minimized immune responses and with drug-like pharmacokinetics. Several research groups have developed assays for evaluating the interactions between small

molecule ligands and receptors, creating a tool for discovering and developing small molecule based gene regulators. These assays have also been extended to applications in enzyme and protein engineering [27-30].

Chemical Inducers of Dimerization (CIDs)

To create these small molecule-based gene regulators, attempts have been made by several research groups to develop assays for developing and evaluating these drug-inducible transcription factors [15,22,23,26,30-32]. Basically, the small molecules (or chemicals) in these systems serve as a bridge between two proteins, and by binding the small molecule these proteins dimerize, in turn causing transcription to occur; thus creating a chemical inducer of dimerization (CID). As shown in Figure 1.2, two fusion proteins are involved in the function of this system. The first fusion consists of a DNA binding domain (DBD) fused to a small molecule binding protein. The other fusion consists of an activation domain (AD) fused to a small molecule binding protein. In some instances, the same small molecule binding protein is fused to either the DBD or AD and homodimerization occurs [22,23,30]. In more recent developments, different small molecule binding proteins have been used. The DBD will bind a DNA sequence in the promoter region of the target gene, and in the presence of the small molecule, the two small molecule-binding proteins dimerize, recruiting the AD protein to the promoter region [22,23,33,34]. Thus, transcription of the target gene will occur.

Several modifications using CIDs have been described in various host organisms, ranging from mammalian cell lines to yeast and *E. coli*. The most widely used protein/small molecule combination has been the immunophilin FKBP, a cytoplasmic

intracellular receptor for FK506 and rapamycin, which can homodimerize with another protein, called FK1012 [26,35-37]. Rapamycin was used to further extend the system and develop a method for clinical uses in controlling gene therapy. The rapamycin-inducible system uses a zinc finger pair and homeodomain (ZFHD1) to serve as a DBD, which is fused to the human FKBP (FK506 binding protein). The other fusion protein consists of the p65 subunit of human NF-κB gene fused to a domain of the FRAP protein (FKBP rapamycin-associated protein). These constructs provided a system with a dose-responsive robust expression system, which is currently in clinical trials for its activity as a drug-induced transcriptional regulator of gene expression [34,36-39].

More recently, another CID-based system has been developed also called chemical complementation, where a linker molecule is used to recruit an activation domain and DNA binding domain in close proximity to one another [40]. The system is based on two fusions. The first fusion protein consists of dihydrofolate reductase (DHFR) fused to the LexA DNA binding domain, in which the LexA DBD is able to bind to a DNA sequence that controls expression of a *LEU2* gene. The *LEU2* gene is essential for leucine biosynthesis, and a strain is utilized where the *LEU2* gene is under the control of the LexA inducible promoter. Thus, leucine is produced when the LexA protein binds to the promoter, activating transcription of the *LEU2* gene, and if the yeast cells are not supplemented with leucine, this process needs to occur, such that the yeast will be able to produce their own leucine for survival. DHFR is able to bind to methotrexate. The second fusion protein consists of the glucocorticoid receptor (GR) fused to the B42 functional activation domain (AD). GR binds to dexamethasone (Dex). In this chemical

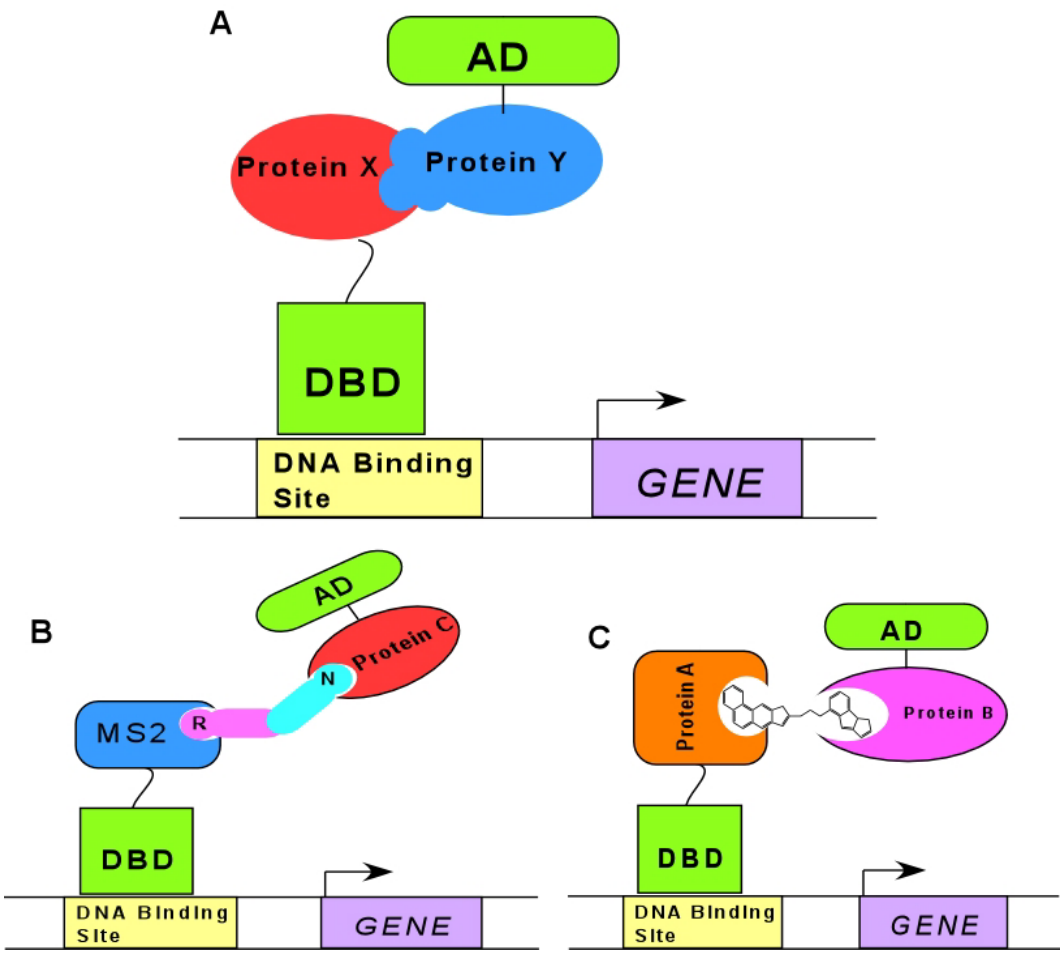


Figure 1.2. Yeast two and three- hybrid assays.

complementation developed by Cornish and co-workers, synthetic ligands were created with Mtx and Dex on the ends of the small molecules and a linker region in the middle. This small molecule serves as a bridge between the DHFR: DBD and the GR: B42 AD, where in the presence of this linker small molecule, the two fusions are brought together and transcription of the *LEU2* gene occurs. Thus, yeast survival is linked to the presence of a small molecule linker [30,33].

The chemical complementation technique developed by Cornish and co-workers was extended to applications in enzyme engineering. Using their system, a glycosidase variant, an enzyme involved in carbohydrate biosynthesis, is known to act as a glycosynthase, which forms glycosidic bonds. Cornish and coworkers used their chemical complementation and used a methotrexate- disaccharide fluoride donor and a dexamethasone-disaccharide acceptor to create their linker ligand. The glycosynthase can form a glycosidic bond between these two ligands, creating the linker molecule, and bridging the two fusion proteins together. Thus, yeast survival is linked to the functionality of an enzyme; if the enzyme is able to form a glycosidic bond and create the linker molecule, transcription occurs allowing the cells to survive on selective media [7,30,33,41].

Chemical Complementation using nuclear receptors

The focus of this work has been the development of a genetic selection system for detecting the interaction of a small molecule with a protein; in this case a nuclear receptor. This system, described thoroughly in the chapters ahead, takes advantage of nuclear receptor's function and links this function to genetic selection. In short, nuclear

receptors are ligand-activated transcription factors, where upon the presence of ligand, these receptors undergo a conformational change, indirectly recruiting RNA polymerase, and initiating transcription. In the chemical complementation described throughout this dissertation, the nuclear receptor ligand-binding domain is fused to a Gal4 DNA binding domain. A yeast strain that contains a Gal4 inducible promoter controlling expression of a selective gene is used, such that binding of the Gal4 protein to the promoter site leads to transcriptional activation of a genetic selection marker. Therefore, in the presence of the nuclear receptor ligand binding and activating a nuclear receptor will allow survival of yeast in selective media [12].

Chemical complementation is a powerful tool for analyzing nuclear receptor-ligand interactions. This system can be extended to applications in drug discovery and protein engineering, especially in creating small molecule regulators of gene expression. Furthermore, chemical complementation can be extended towards deciphering and assembling biosynthetic pathways and for applications in enzyme engineering. The following chapters will detail chemical complementation and the many applications that have been and can be used with this system.

Chapter 2

Nuclear Receptors

2.1 The Nuclear Receptor Superfamily

Nuclear receptors are a superfamily of ligand-activated transcription factors involved in diverse cellular and biological processes, such as cell proliferation, differentiation, and intracellular signaling [42-45]. These proteins are abundant in various organisms, from humans to worms, and are involved in the regulation of crucial physiological genes. The role of nuclear receptors in multiple molecular processes leads to the complex network of molecular events in the process of transcription [42-45].

To date, there are 65 nuclear receptors that have been identified, and these receptors range from steroid receptors, such as the estrogen receptors, to non-steroidal receptors, such as the thyroid hormone receptor and the retinoic acid receptor [43,44]. Most of these nuclear receptors contain various isotypes (such as α, β, γ) and even further various isoforms (for example $\alpha 1$ or $\gamma 2$), which are all involved in specific regulatory biological pathways. The distinguishing feature of nuclear receptors in comparison to other transcription factors is that these proteins are ligand-dependent transcription factors. These receptors bind small hydrophobic ligands, and these small lipophilic molecules are involved in signaling pathways of various physiological and biological regulatory mechanisms [42-44]. A great deal of research has been forth to discovering the function, determining the structure and discovering ligands for the various nuclear receptors, leading to the further understanding of the superfamily of these receptors.

Besides their obvious role in transcription of regulatory biological pathways, nuclear receptors are also implicated in a number of diseases [43,44]. Dysfunction of

nuclear receptors leads to multiple complications in various biological pathways, such as reproduction and metabolism, spanning a range of medical conditions, such as cancer, diabetes, and obesity [43,44,46]. Pharmaceutical industries have a vast interest in the development and the discovery of agonists and antagonists for a variety of nuclear receptors as possible therapeutics for nuclear receptors-based diseases.

Currently about 10% of the commonly prescribed drugs are targeted towards nuclear receptors [43,44,47]. Some examples of nuclear receptor based drugs currently used in therapeutics are: (1) tamoxifen: an antagonist for the estrogen receptor used as a drug for breast cancer treatment, (2) thiazolidinediones: agonists for the peroxisome proliferators activated receptor γ (PPARγ) used as drugs for type II diabetes, and (3) dexamethasone: agonist for the glucocorticoid receptor used for targeting inflammatory diseases [43,44]. The challenge lies in the discovery of novel chemical compounds that contain the desired effect and pharmokinetics; in other words contain desirable absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties of small molecules. Furthermore, more recently, the extent of drug discovery targeted towards nuclear receptors has extended toward the development of small molecules that are inverse agonists or antagonists; these molecules are able to stabilize an inactive conformation of a receptor, causing tighter association of the nuclear receptor with proteins involved in repressing transcription [43,44]. An interest also exists in discovering selective nuclear receptor modulators (SNuRMS), which are small molecules that can modulate different receptor subtypes and that can act in a tissue-specific manner [43,44,48]. Overall, nuclear receptors drug targets hold promise for understanding disease mechanism and for developing therapeutic remedies, allowing further

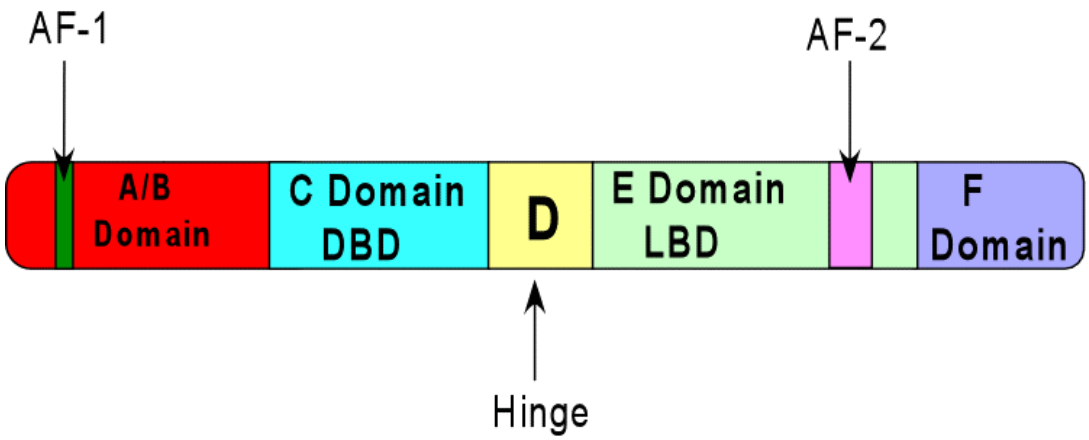


Figure 2.1. Nuclear receptor domains.

understanding of the disease and more importantly of the instrumental physiological role of nuclear receptors.

2.2 Structure of Nuclear Receptors

Typically nuclear receptors consist of several functional domains, shown in Figure 2.1 [44,49]. These domains can be classified as domains A-F, which include a variable N-terminal A/B domain, a DNA binding domain (DBD) referred to as the C domain, a flexible hinge D domain, and a C-terminal ligand-binding domain (LBD) referred to as the E domain, and the F domain. Each of these domains contributes to the structure and function of nuclear receptors.

The N-terminal A/B domain of nuclear receptors consists of the AF-1 region, the ligand-independent trans-activation function, which is the least conserved region among nuclear receptors, both in the size of the region as well as the sequence homology [43,44,50,51]. Recent studies have shown that the AF-1 domain is involved in binding coactivators, however the structural basis for these interactions are not fully understood [43,50]. The C domain of nuclear receptors is the DBD, which is a highly conserved region involved in binding to specific DNA sequences, known as response elements (REs) [43]. As shown in Figure 2.2, the DBD mainly consists of two zinc finger modules; in each module four cysteine residues chelate to one zinc ion. Various X-ray and NMR solution structures of this domain have been solved and analyzed, allowing deeper understanding of the structure and function of the DBD. Nuclear receptors contain a flexible hinge region that lies between the DBD and the ligand-binding domain (LBD). The hinge region is less conserved than the DBD or LBD and is thought to

provide some flexibility for the DBD and LBD, preventing structural issues, such as steric hindrance [42-44,52].

The hallmark of nuclear receptors is their LBD, which is responsible for interacting with the transcription machinery such as the RNA polymerase holoenzyme complex. The LBD of nuclear receptors consists of a canonical fold containing 10-13 alpha helices, 2-5 β-strands, and variable loops; these structural features are arranged into an anti-parallel three-layered sandwich (Figure 2.3) [43,52,53]. Within the LBD lies the ligand-binding pocket (LBP), involved mainly in the ligand-binding function of these nuclear receptors. The LBP is defined by its primarily hydrophobic residues and several different alpha helices involved in determining the specificity and efficiency of ligand-binding [43,53-55]. Another important structural feature of the LBD is a ligand-dependent activation function domain, referred to as the AF-2 domain. This domain, unlike the AF-1 domain is highly structured, and is in involved directly in the interaction with transcriptional activators. The AF-2 domain in the LBD is part of helix 12, and the proper positioning of this helix is crucial for transcriptional activation and for nuclear receptor function [43,50,53-56].

The first three crystal structures of nuclear receptor LBDs solved were that of the *apo* (without ligand) retinoid X receptor, the retinoic acid receptor bound to its ligand *all trans* retinoic acid, and of the *holo* thyroid receptor [57,58]. These crystal structures revealed the various features of the LBD of these nuclear receptors, as well as showing that in the *apo* (without ligand bound) form of the nuclear receptor and in the *holo* (with ligand bound) form of the nuclear receptor, the main structural fold of these nuclear receptors are the same [59-61]. As more three-dimensional structures of the different

nuclear receptors are solved, the common LBD fold is observed in these nuclear receptors, with variability in the sizes and shapes of the LBD (Figure 2.3). This variation in the LBD of nuclear receptors is expected, since most of the nuclear receptors are involved in diverse physiological pathways and are responsible for binding small molecule ligands that are members of those pathways [43,44,55].

Some of the nuclear receptors also contain a C-terminus domain called the F domain; a domain that contains very little evolutionary conservation [43]. Like some of the more undefined structural features of nuclear receptors, the A/B domain in the N-terminus and the F-domain in the C-terminus are thought to be involved in aiding the other main domains of nuclear receptors in carrying out their function. For example, recent studies have shown that the F domain of the retinoic acid receptor alpha (RARα) is involved in maintaining the open conformation of the LBD without the presence of ligand, by stabilizing corepressor binding and destabilizing coactivator binding [62]. Some recent studies also have implicated these domains in post-translational modification of nuclear receptors [43].

2.3 The Nuclear Receptor Transcriptional Model

Several biochemical experiments and approaches have been used to decipher the role of nuclear receptors in the process of transcription. With a vast oversimplification of the actual mechanism of transcriptional control by nuclear receptors, the model of transcriptional activation by nuclear receptors is based on the ability of these receptors to interact with a series of transcriptional coregulators. More specifically, in the absence of ligand, or in an antagonized state, nuclear receptors recruit corepressor complexes

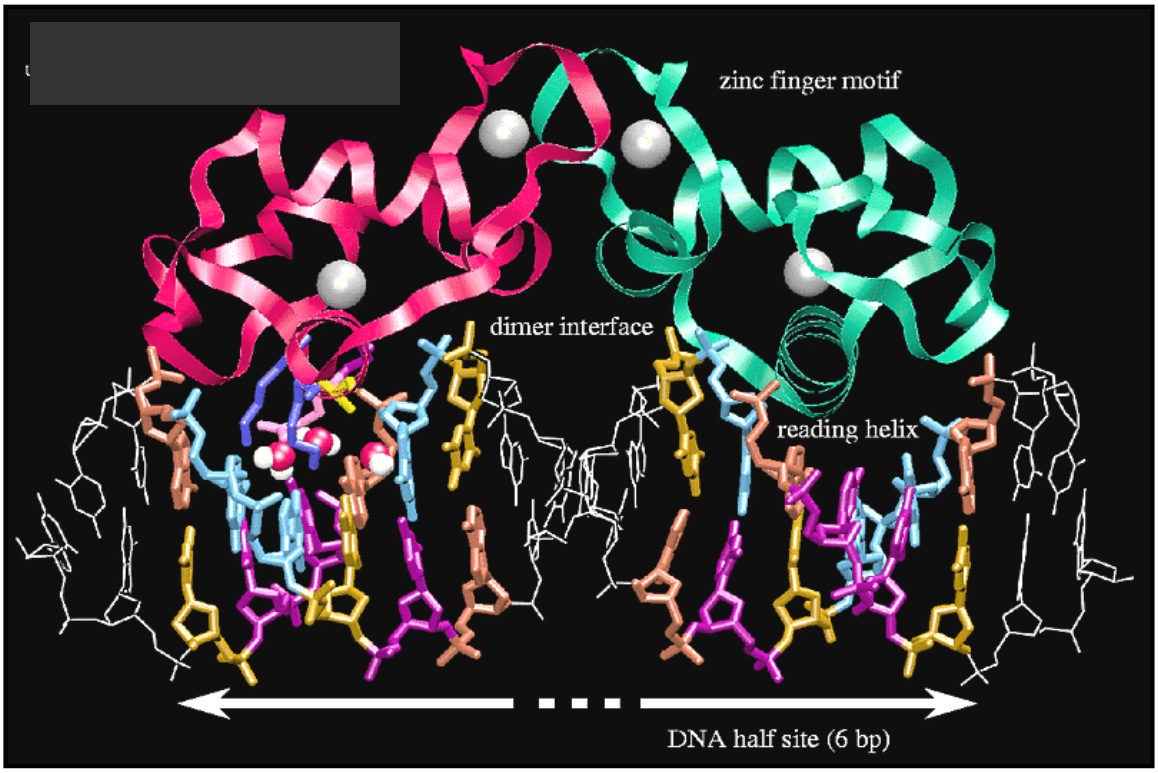


Figure 2.2. The DNA binding domain (DBD) of nuclear receptors.
Adapted from ***http://www.ks.uiuc.edu/Research/pro_DNA/ster_horm_rec***

(Figure 2.3). Two major corepresssors involved in interacting with nuclear receptors have been identified: the nuclear receptor corepressor (NCoR) and the silencing mediator of retinoid and thyroid hormone receptors (SMRT) [63-65]. These corepresssors interact with a region in the nuclear receptor structure known as the CoR box, which is located near the hinge region of the receptor. Thus, in the *apo* form of the nuclear receptor, the structure of the nuclear receptor allows the CoR region to be exposed for binding by the various corepressors. These corepressors, such as NCoR1 or SMRT, are then involved in recruiting a complex of additional proteins, known as histone deacetlytransferases (HDACs) [66,67]. HDACs actively repress transcription through chemical modification of the histones, resulting in the tighter association of the chromatin wound around the histones, preventing RNA polymerase exposure to the transcribed DNA [66,68-70]. These events lead to the repression of transcription.

Upon ligand binding, the nuclear receptor LBD undergoes a conformational change, leading to a major structural change in the binding domain shown in Figure 2.3. The structural conformational change involves several of the helices present in the LBD, especially helix 12. In the *apo* form of the receptor, helix 12 of the LBD is freely floating around in solution. However, once the ligand binds to the receptor, helix 12 swings towards the protein, forming a seal to the binding cavity, shown in Figure 2.3 [71-75]. The structural change by helix 12 is accompanied by other minor structural changes in helices 3, 6, 10 and 11 [72-74]. The conformational change leads to the formation of a seal caused by helix 12 provides a stabilizing effect for the ligand in the pocket by contributing to the hydrophobic pocket. More importantly, the AF-2 domain of helix 12 is completely repositioned with the presence of ligand, and forms a

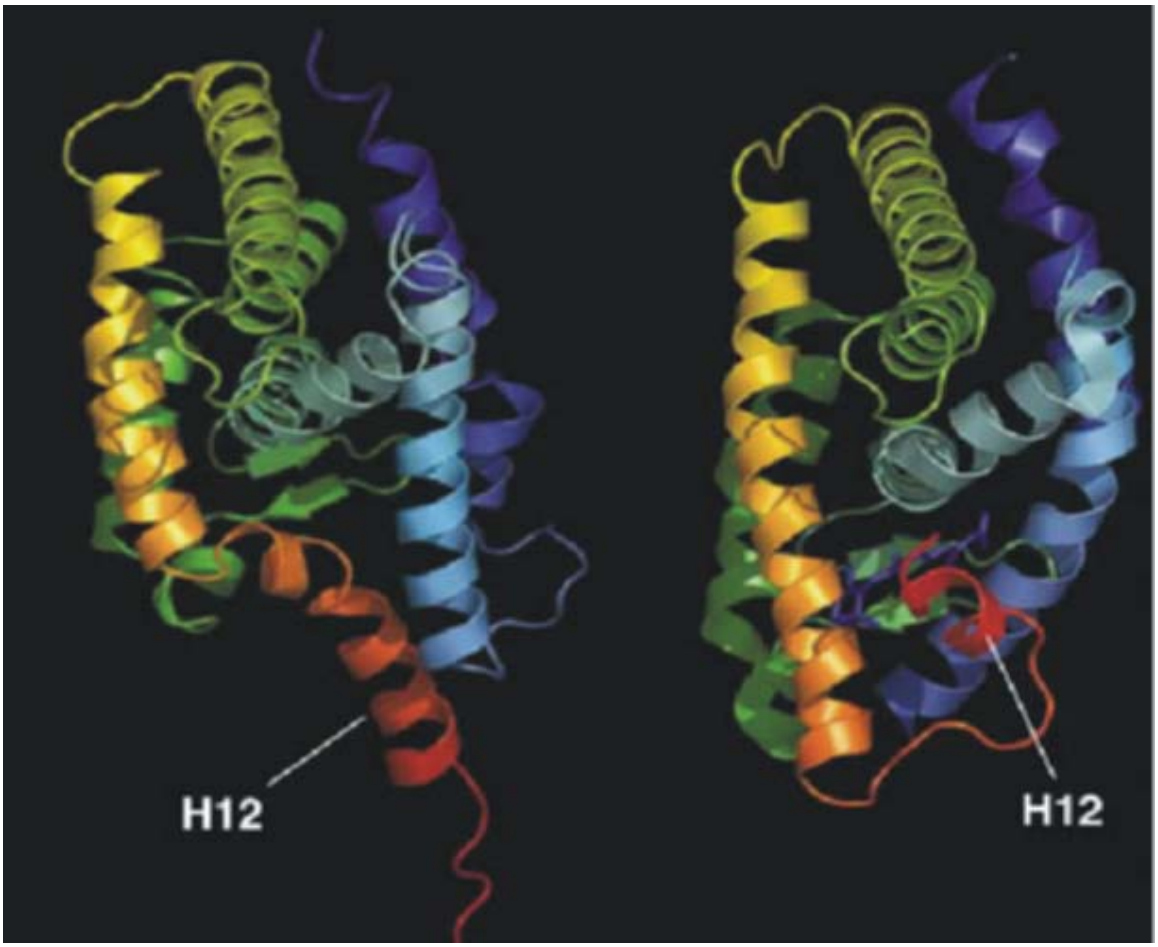


Figure 2.3. The Ligand Binding Domain (LBD) of Nuclear Receptors in the *apo* and *holo* Form

hydrophobic cleft that together with other surface exposed amino acids forms a coactivator binding region, known as a "LXXLL NR box" (Figure 2.4) [43,55,72-74].

The coactivators of nuclear receptors can be divided into two general classes: the switch/sucrose non-fermentable family and the histone acetyltransferase (HATs) family of proteins [43,55,72-74]. Both of these families of proteins have the ability to modify the chromatin, however, histone acetyltransferases were identified based on their interaction with the ligand- binding domain of nuclear receptors [43,55,72-74]. The mechanism of coactivator recruitment involves the following events: (1) ligand-binding induces a conformational change, which displaces the bound corepressors, (2) the conformational changes leads to the exposing of the LXXLL NR box, and (3) a complex of coactivator proteins is recruited to the nuclear receptor LBD and transcription is eventually initiated. Nuclear receptor coactivators contain several LXXLL motifs, which play an important role in mediating the interaction between the coactivators and receptors through the LXXLL NR box (Figure 2.4). The coactivator proteins are capable of acetylating specific residues in the histones, allowing RNA polymerase to initiate transcription. If the coactivators cannot perform the function of acetylation, they are involved in recruiting other proteins which contain intrinsic acetylation activity [71-74,76-78]

2.4 Summary

Nuclear receptors are transcription factors that play a pivotal role in molecular processes, such as intracellular signaling, and are also involved in the various metabolic and physiological pathways, such as differentiation and development. Small

predominantly hydrophobic molecules that bind to these receptors are crucial for regulating these proteins by responding to direct physical association with other signaling pathways. These proteins have also been implicated in a number of diseases, ranging from obesity to cancer, and pharmaceutical industries are focused on developing new drug targets towards the superfamily of nuclear receptors. As the structure of more nuclear receptors are being solved through methods, such as X-ray crystallography, more information is determined about the basic structure of this family of proteins, as well as the various unique structural features of the different members of this family. Furthermore, knowledge on the function and structure of these receptors will eventually lead to the understanding of the evolution of these receptors and to the exact molecular mechanisms involved in their natural function as well as in their role in disease.

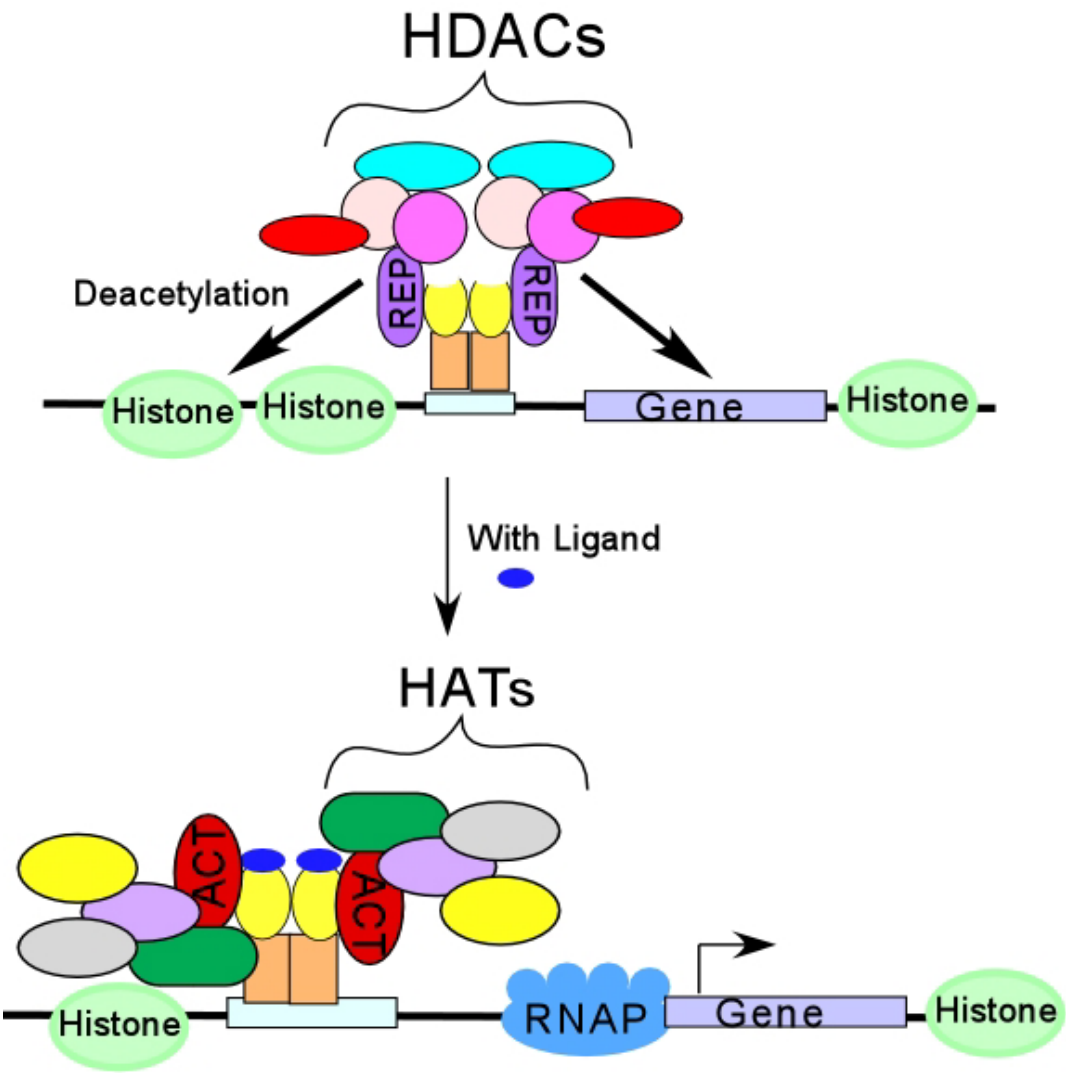


Figure 2.4. Molecular mechanism of transcriptional regulation by nuclear receptors.

Chapter 3

First Generation Chemical Complementation

3.1 Defining Chemical Complementation

Chemical complementation is broadly defined as genetic selection using small molecules. More specifically, the chemical complementation system developed and described here, involves a one-hybrid yeast assay where genetic selection is used as a tool for evaluating the interaction between a protein and a small molecule [40,79]. The survival of yeast is linked to the ability of the ligand to bind and activate the nuclear receptor. The activation of the nuclear receptor leads to the expression of a selective marker. Chemical complementation combines the power of genetic selection with nuclear receptors, which serve as a rich source of proteins for ligand-activated genetic selection, permitting the ability to link molecular recognition of a small molecule by a protein to survival of yeast. In classical genetic complementation a functional gene complements a non-functional gene in the yeast (e.g. an auxotroph). Chemical complementation can be compared to classical genetic complementation in yeast, where both of these systems are similar in that the complementation involved allows the yeast to grow. The difference is that in chemical complementation a small molecule chemical allows the yeast to grow by activating transcription of a functional gene through the action of a nuclear receptor, whereas in classical genetic complementation a functional gene is delivered to complement a non-functional gene. The interaction between the small molecule ligand and the receptor provides an effective, general system for genetic selection.

Detection of ligand binding to nuclear receptors and transcriptional activation has been performed in mammalian cell culture assays and in yeast. These methods use reporter plasmids containing a nuclear receptor response element (RE) upstream of a minimal promoter and a reporter gene, such as luciferase in mammalian cell assays. Binding of the ligand to the nuclear receptor initiates transcription of the reporter gene. Yeast are ideal organisms for studying nuclear receptor-ligand interactions because there are no nuclear receptors in yeast to interfere with nuclear receptor assays, and because the yeast transcription machinery is sufficiently similar to the mammalian transcription machinery to permit the expression and function of mammalian nuclear receptors in yeast.

To establish chemical complementation, the PJ69-4A yeast strain was used. This is a third generation yeast two-hybrid strain consisting of Gal4 response elements (Gal4 REs) controlling the expression of genetic selection and screening genes, such as the *HIS3* and *lacZ* genes respectively [80]. The Gal4 protein is an endogenous yeast ligand-independent transcriptional activator, which contains a DNA binding domain (DBD) and an activation domain (AD). The DBD, like that of nuclear receptors, binds to the Gal4 REs and the activation domain, is involved in interacting with RNA polymerase complex.

To develop chemical complementation, we utilized the *HIS3* gene of PJ69-4A, which encodes a key enzyme in the histidine biosynthetic pathway, imidazoleglycerol-phosphate dehydratase [12]. Expression of the *HIS3* gene allows yeast to produce histidine and to survive in media lacking histidine. We utilize this strain by fusing the nuclear receptor's ligand-binding domain (LBD) to the Gal4 DNA binding domain (Gal4 DBD), as shown in Figure 3.1. This fusion protein binds to the Gal4 response elements

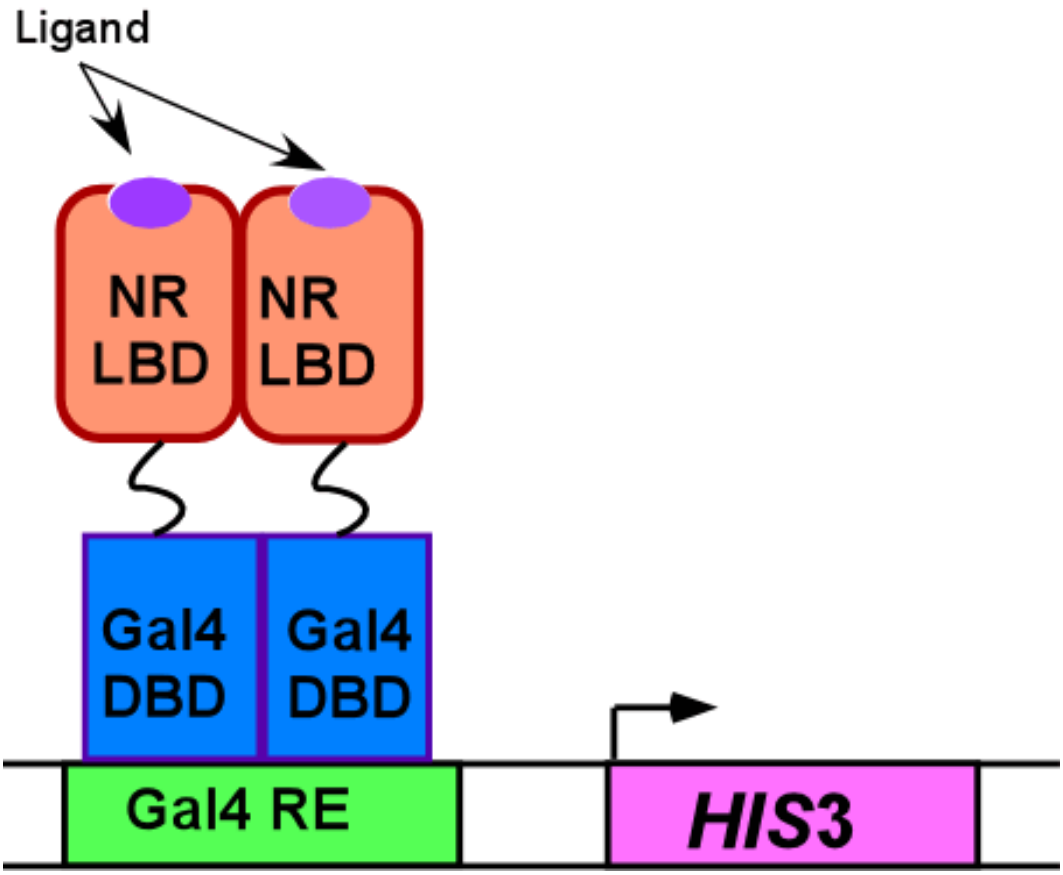
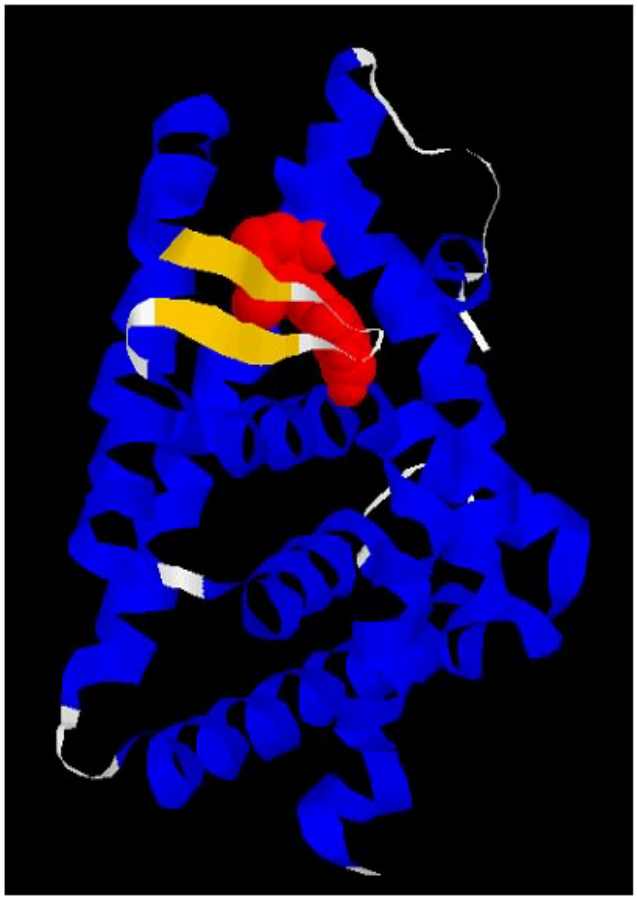


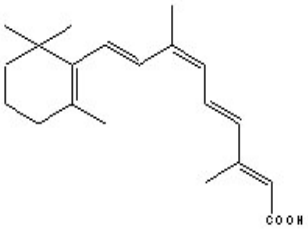
Figure 3.1. 1st generation chemical complementation.

controlling the expression of the *HIS3* gene. Upon binding ligand the nuclear receptor activates transcription of the regulated genes [12]. Thus, only ligand binding and activation of the nuclear receptor will confer yeast survival. We investigate this system of genetic selection in yeast using the retinoid X receptor (RXR), the pregnane X receptor (PXR), and the liver X receptor (LXR).

The retinoid X receptor (RXR) is a member of the nuclear receptor superfamily, under the subclass of retinoid receptors [43,46,81]. RXR is involved in various pathways throughout the body, playing a crucial role in differentiation, development, and in maintenance of homeostasis [82-85]. RXR's unique characteristic is that this receptor has been shown to form heterodimers with many of the other members of the nuclear receptors family, and even more interesting, RXR has been known to self-associate into homo-tetramers [86-89]. The formation of the different oligomeric states of RXR was found to be tightly regulated by the presence ligand, due to the structural changes that occur in the presence of ligand leading to the loss of tetramerization [86-89]. The natural known ligands for RXR include phytanic acid, docosahexaenoic acid and 9-*cis* retinoic acid (9cRA) [82-85,90]. A series of synthetic ligands have also been discovered for RXR, such as LG100268, shown in Figure 3.2 [81].

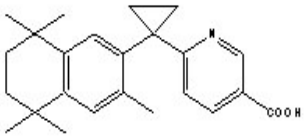
The structure of RXR is similar to other members of the nuclear receptors superfamily containing a variable N-terminal region, DBD, a hinge, and a relatively conserved LBD. The crystal structure of RXR bound to 9cRA is one of the first nuclear receptor *holo* crystal structures to be solved, where key interactions of the binding pocket with the ligand were determined (Figure 3.2) [81,83,91]. Several other nuclear receptors categorized under the same family as RXR tend to have larger binding pockets



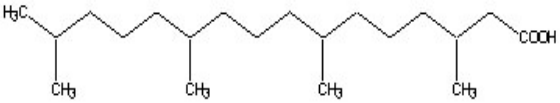


9-*cis* retinoic acid

Docahexaenoic acid



LG100268



Phytanic acid

Figure 3.2. The crystal structure of RXR and known ligands.

in which their respective ligands do not fully occupy the volume of the binding pocket. However, in the case of RXR and 9cRA, there is a much tighter association between this ligand-receptor pair; the RXR binding pocket is relatively smaller, generally about 400-500 A^3. Overall, the structure of the RXR LBD resembles the traditional nuclear receptor LBD (Figure 3.2) [83,91].

Characterized as an orphan nuclear receptor, the pregnane X receptor (PXR) is a member of the nuclear receptor family involved in the regulation of various cytochrome P450s (CYPs) [92]. The CYP enzymes are large family of heme monooxygenases that are involved in the body's defense mechanism, particularly in the role of eliminating xenobiotics from the body. PXR response elements have been found in the regulatory regions of CYP3A genes [92,93]. This receptor acts as a heterodimer with RXR. PXR has a promiscuous LBD, i.e. it binds a variety of small molecules (Figure 3.3). The promiscuity of PXR is due to the fact that unlike the smaller three-stranded β-sheet found in most nuclear receptors, structural data shows that the PXR LBD contains a five-stranded anti-parallel β sheet, allowing more room for accommodating both small and large ligands in the binding pocket (Figure 3.3). Rifampicin (EC_{50}=3 μM) and paclitaxel (EC_{50}=5 μM) are known ligands for human PXR (Figure 3.3). More details about this nuclear receptor are discussed in Chapter 7 [94,95].

The liver X receptor (LXR) is a nuclear receptor that regulates genes involved in cholesterol absorption, transport, excretion, and fatty acid metabolism [96,97]. Two isotypes of this nuclear receptor have been identified, LXRα and β, which are expressed differently in tissues. Both isotypes heterodimerize with RXR [43]. LXR also functions

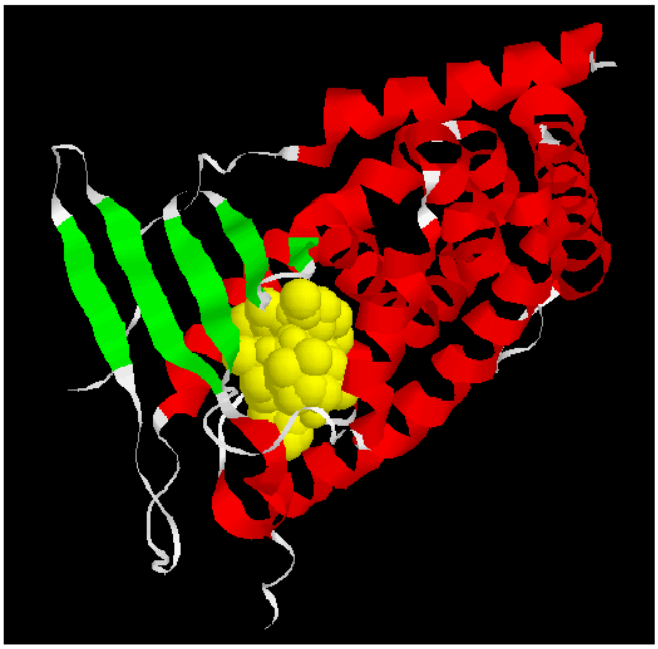
as a sensor of cellular oxysterols, such as lanosterol and 4β-hydroxycholesterol shown in Figure 3.5 [98-102]. More details about this receptor are discussed in Chapter 7.

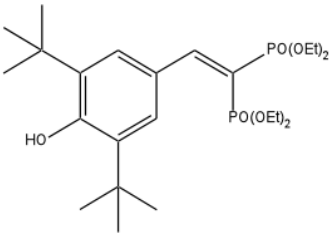
3.2 Results of Chemical Complementation

Chemical Complementation using wild-type RXR and PXR LBDs.

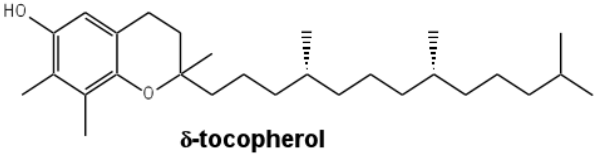
Yeast expression plasmids containing a tryptophan (Trp) marker were constructed. The Gal4DBD (residues 1-151) was fused to wild-type RXR LBD (residues 44-462), denoted as pGBDRXR, or fused to the PXR LBD, denoted as pGBDPXR. These plasmids were transformed and were plated on SC minus tryptophan (SC -Trp) to select for the plasmid. After three days of growth, yeast cells containing pGBDRXR were streaked on SC minus histidine and tryptophan (SC –His -Trp) with 10 μM 9-*cis* retinoic acid (9cRA), selecting for ligand-activated growth. The cells were also streaked on a SC plate as a positive control, and a SC –His -Trp plate without ligand, as a negative control. Plates were incubated at 30 ºC for five days. Each plate consisted of four sectors: yeast cells containing pGBDRXR, pGBDMT, pGBT9Gal4, and the last sector with just yeast cells with no plasmids. The pGBDMT plasmid is an empty plasmid and the sector containing this plasmid, as well as the sector with no plasmids, is used as a negative control. The pGBT9Gal4 plasmid contains the entire Gal4 protein and is used as a positive control.

Ligand activated growth is observed with pGBDRXR and 9cRA. Growth on the SC plate indicates efficient transfer of the cells across selection plates and the viability of the yeast (Figure 3.4). On the plate with no ligand, growth is only seen on the sector of

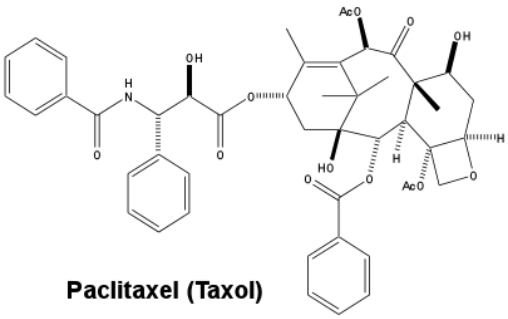




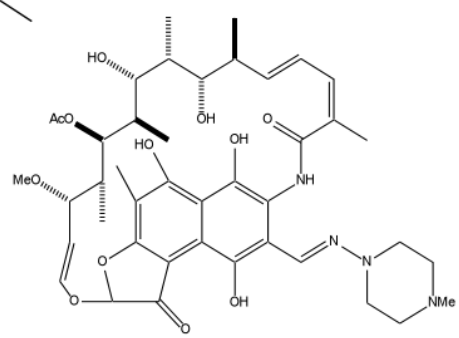
SR12813



δ-tocopherol



Paclitaxel (Taxol)



Rifampicin

Figure 3.3. The crystal structure of PXR and known ligands.

the plate with Gal4, as expected. Growth with the Gal4 protein was detected after two days of incubation. On the plate with 9cRA, growth is observed on the sector containing RXR. However, the growth observed is not as intense as the dense growth seen in the sector with Gal4 and was observed after five days of incubation.

Genetic selection using human PXR LBD was performed using 10 μM paclitaxel and δ-tocopherol, two known ligands for the receptor. Yeast cells containing pGBDPXR were streaked on selective plates, SC-His-Trp plates. The same controls were used as with RXR. Yeast cells grow on all sectors of the SC plate, as expected (Figure 3.4). No growth is seen on the PXR sector of the plate without the presence of any ligand, but good growth is seen on the Gal4 sector of the plate as expected (Figure 3.4). In the presence of 10 μM paclitaxel (Figure 3.4) growth occurs in the pGBDPXR sector of the plate and the Gal4 sector of the plate as expected. On the plate with 10 μM δ-tocopherol, ligand activated growth was also detected on the pGBDPXR sector of the plate and the Gal4 sector of the plate. As with RXR and 9cRA, growth with PXR and both ligands were seen after five days of incubation, and were not as intense as the Gal4 activated growth seen after only two days. Chemical complementation was also tested with the *ADE2* gene, which is the more stringent selection media in PJ69 [80]. No growth was observed with wild-type RXR and 9cRA or with PXR and the two ligands.

Chemical Complementation with wild-type RXR and LXR.

Certain oxysterols, which are part of the steroid hormone biosynthesis and bile acid synthesis, are known ligands of LXR [101]. Lanosterol, ergosterol, 25-

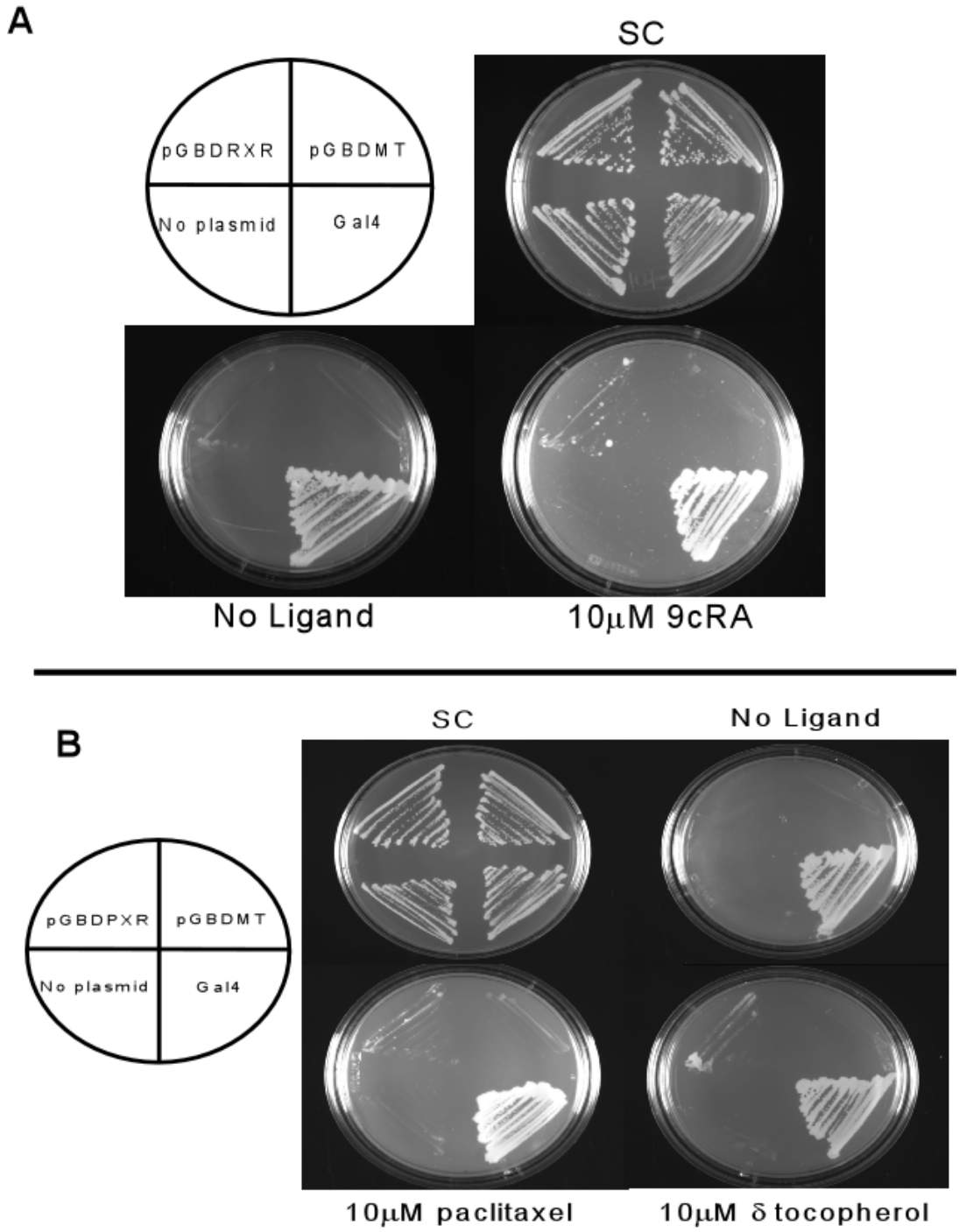


Figure 3.4. Chemical complementation with RXR and PXR and their ligands.

hydroxycholesterol, and 4β-hydroxycholesterol were used for performing genetic selection using the LXR LBD. These ligands have previously been evaluated for their transcriptional activation with LXR in mammalian (CV-1) co-transfection assays [101]. To determine whether genetic selection produced the same result, yeast cells were transformed with pGBDLXRα, a yeast plasmid that contains the Gal4 DBD fused to the LXR LBD. Cells were also transformed with both pGBDRXR and pGBDLXR because LXR is known to function as a heterodimer with RXR. These cells were streaked onto SC -His -Trp plates with 10 μM of each of the ligands. Unlike RXR and PXR, little background growth was observed on the LXRα sector of the SC -His -Trp without any ligand (Figure 3.5); therefore, 3-AT was not used in this case. Using the same controls as previously stated, yeast cells were able to grow as expected on the SC plates after two days of incubation. After five days of incubation, growth was observed on the LXR and LXR+RXR sectors of the plate containing each of the oxysterols (Figure 3.5). The results using cell culture assays had shown that ergosterol and lanosterol were weak activators of LXR, whereas, 4β- hydroxycholesterol was a good activator of the receptor. Judging from the growth that is observed on the selective plates, more growth is seen on the plate with 4β- hydroxycholesterol relative to lanosterol or ergosterol, showing the same trend as in cell culture assays. Growth is also observed on the RXR sector of the plate.

Chemical Complementation with RXR mutants.

Several RXR LBD mutants were cloned into the yeast pGBD expression vector (Table 3.1). These mutants have substituted amino acid residues that have been shown

through examination of the RXR-9cRA crystal structure to directly interact with the ligand or are believed to be crucial in binding the ligand. They include single and multiple amino acid substitutions, and have been previously tested in mammalian cell culture assays for activity in the presence of 9cRA. Table 3.1 summarizes the transcriptional activation that was measured with the different mutants and 9cRA in mammalian cell assays. To investigate the correlation between genetic selection and mammalian cell assays, the mutants were streaked onto SC-His-Trp plates with and without 10 μM 9cRA.

Dense growth is seen on the SC plate after two days (Figure 3.5). Without any ligand, no growth was detected in any of the sectors containing the mutants (Figure 3.6). With 10 μM 9cRA, after three days of incubation, ligand-activated growth was observed on the sector of the plate containing the F439L mutant with 9cRA. After five days of incubation, ligand- activated growth was also observed on the wild-type RXR sector of the plate (Figure 3.6). Some growth was also detected on the sector of the plate containing the pGBDRXR F313I; F439L mutant. Higher density growth was observed on the sector with the F439L mutant and 9cRA than wild-type RXR and 9cRA. These results correlate with transcriptional activation measured through mammalian cell assays (Table 3.1). However, there is an incomplete correlation with the other mutants tested.

Several observations are notable about the ligand-activated growth seen with RXR F439L mutant. The growth time for 9cRA activated growth is less for the wild-type; growth occurs after three days of incubation versus five days of incubation. Growth is more dense with F439L than for wild-type. To compare the activation threshold for the F439L mutant, with that of wild-type RXR, a chemical complementation dose response

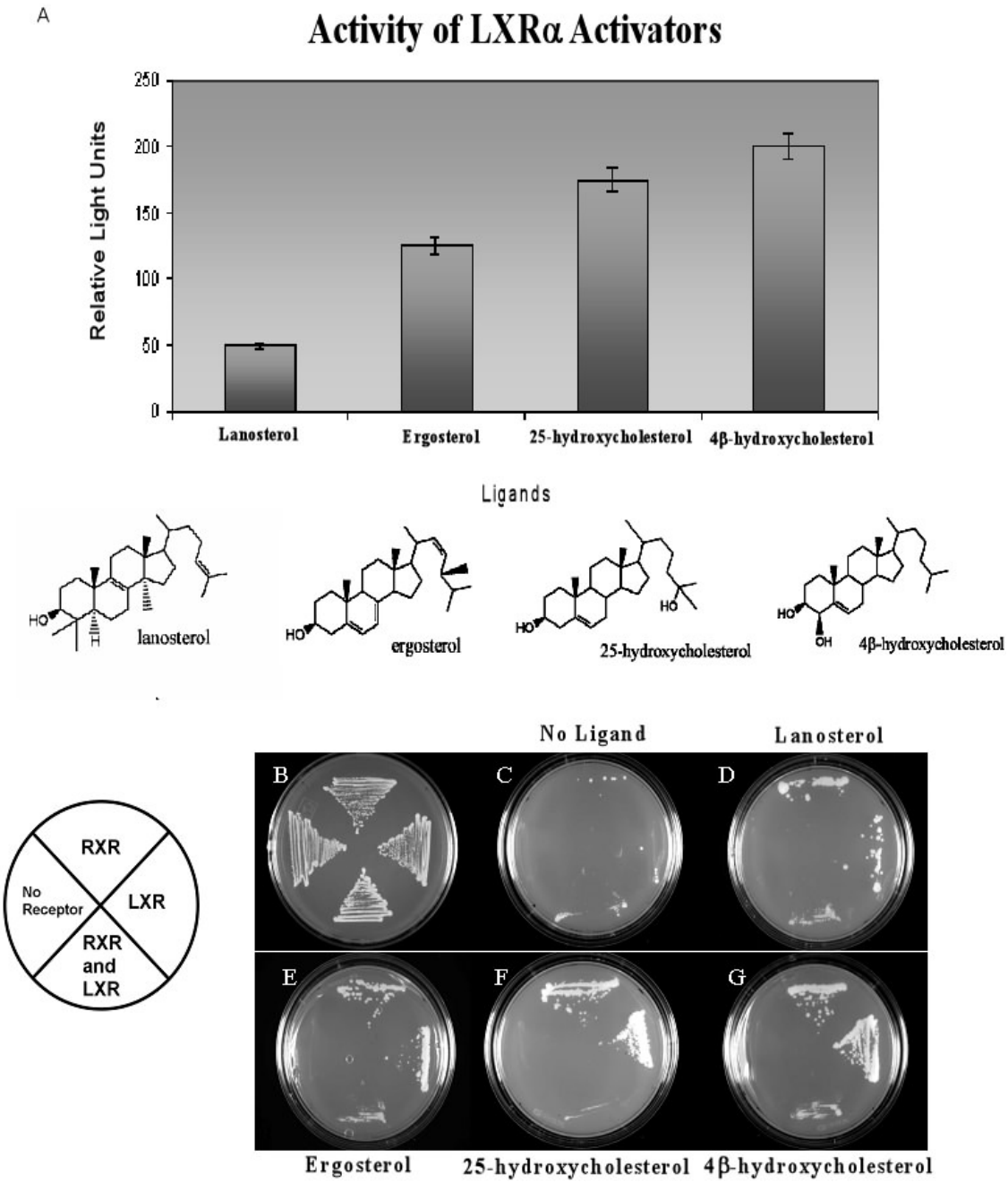


Figure 3.5. Activity of LXR in Mammalian Cell Culture (A) and with chemical complementation (B-G).

was performed with 0.1 μM, 1 μM, and 10 μM 9cRA (Figure 3.7). At 1 μM 9cRA (Figure 3.7), growth occurs in the F439L sector of the plate, but no growth occurs on the wild-type RXR sector. At 10 μM 9cRA, growth is detected on both sectors as expected (Figure 3.7). The F439L mutant was also streaked onto minus adenine media and some growth was detected (data not shown).

3.3 Summary of Chemical Complementation

Genetic selection in yeast is a proven, powerful tool for evaluating interactions between proteins in the yeast two-hybrid system. We have developed a system in which the recognition of a small molecule ligand by a protein is necessary for the survival of the yeast. In this system, nuclear receptors serve as a rich source of proteins for ligand-activated genetic selection, permitting the ability to link molecular recognition of a small molecule by a protein to survival of yeast in a process called "chemical complementation." In genetic complementation a functional gene complements a non-functional gene in the yeast (e.g. an auxotroph). Chemical complementation is similar to genetic complementation in that complementation allows the yeast to grow. Chemical complementation differs from genetic complementation because instead of delivering a functional gene to complement a non-functional gene, a small molecule chemical allows the yeast to grow by activating transcription of a functional gene through the action of a nuclear receptor. The interaction between the small molecule ligand and the receptor provides an effective, general system for genetic selection.

Correlation between mammalian cell assays and chemical complementation.

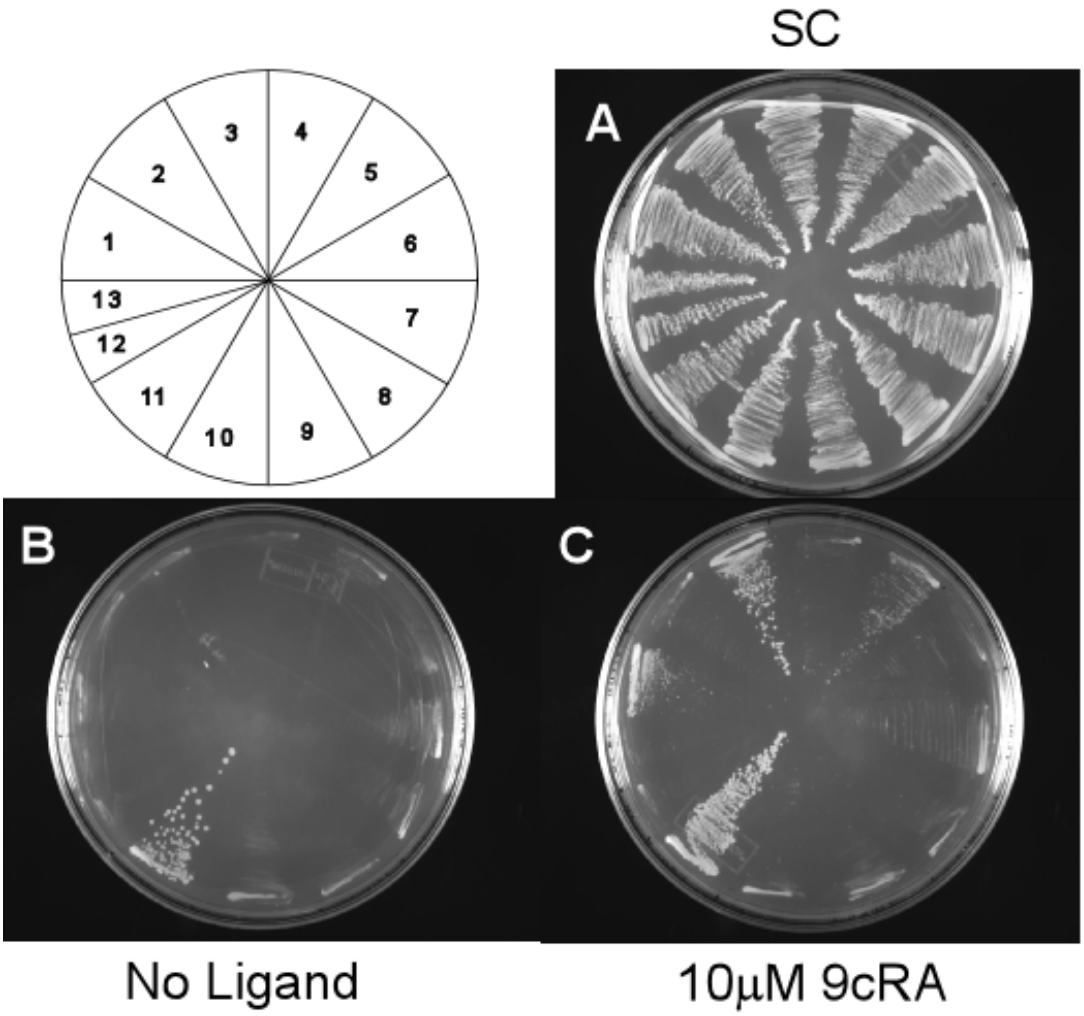


Figure 3.6. Chemical complementation with RXR variants.

We demonstrate chemical complementation with two nuclear receptors and their various ligands. The first case of chemical complementation is with RXR. The ligand 9cRA activates wild-type RXR in mammalian cell assays, and yeast are able to grow on selective plates (SC-His-Trp) in the presence of 9cRA (Figure 3.4). The second case of chemical complementation is with PXR. Paclitaxel and δ-tocopherol activate PXR in mammalian cell assays and yeast cells are able to grow on selective plates made with both of these ligands (Figure 3.4). The same correlation is seen with LXR and its oxysterol ligands, as shown in Figure 3.5, where the LXR ligand, 4β-hydroxycholesterol that has the highest transcriptional activity in cell culture, also showed the most dense growth using chemical complementation. In summary, chemical complementation has been achieved with three distinct nuclear receptors and their ligands, and transcriptional activation in mammalian cell assays correlates with ligand-activated growth in yeast.

Mutagenesis increases the sensitivity of chemical complementation.

Mutagenesis can be used to increase the sensitivity of chemical complementation. Through the modification of one amino acid of wild-type RXR, replacing a phenylalanine with a leucine at position 439 (F439L), yeast cells grow in response to 9cRA better than wild-type, as shown in Figure 3.6. Ligand-activated growth on the selective plates with the RXR F439L mutant and the F313I;F439L mutant correlates with transcriptional activation measurements seen with mammalian cell assays. However, this correlation does not extend to all mutants tested (Table 3.1) [28,29]. For example, the RXR mutant F313I does not show 9cRA-activated growth, which is in contrast to activation patterns in

mammalian cell assays. The incomplete correlation of mutant nuclear receptor function in mammalian cells and yeast shows that although yeast and mammals are closely related enough for mammalian receptors to function in yeast, they have diverged enough so that their function is not identical to their function in mammalian cells. We plan to investigate this further.

Nuclear receptors function as homodimers in yeast.

The fusion of the Gal4 DBD to the nuclear receptor LBD enables nuclear receptors that naturally serve as obligate heterodimers to function as homodimers. Specifically, both PXR and LXR heterodimerize with RXR. However, as can be seen in Figures 3.4, ligand-activated growth seen with PXR and both paclitaxel and δ-tocopherol is with PXR serving as a homodimer, and the same trend is observed for LXR and the oxysterols (Figure 3.5). The Gal4 DBD evidently provides enough dimerization interface to allow PXR to act as a homodimer in yeast.

In conclusion, we demonstrate chemical complementation in *Saccharomyces cerevisiae*, linking small molecules to genetic selection. In this genetic selection system, the survival of the yeast depends on a small molecule to bind and activate a nuclear receptor, thus allowing transcription of the regulated genes. We have shown that a correlation between transcriptional activation in mammalian cell assays and ligand-activated growth through genetic selection exists, though it is not perfect. Mutagenesis of the receptor can increase sensitivity of chemical complementation in yeast. Fusing nuclear receptor LBDs to the Gal4 DBD permits an obligate heterodimer to act as a homodimer in yeast. Using genetic selection is a relatively simple method of evaluating

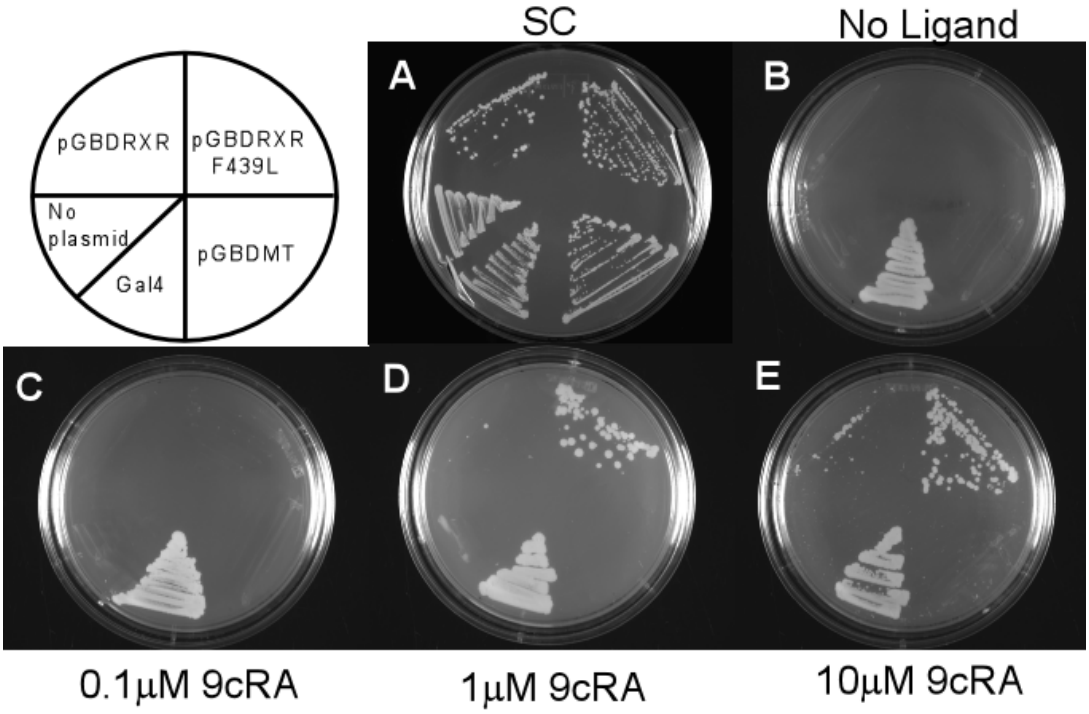


Figure 3.7. Wild-type RXR and the F439L Variant on histidine selective plates with 9cRA.

ligand-activated transcription between a small molecule and a protein. The system should be useful for engineering nuclear receptors for practically any small molecule through genetic selection, basing the survival of the yeast on chemical complementation. Mutant libraries of proteins can be transformed into yeast and tested against a series of small molecules, selecting for functional receptors. Similar to antibodies and anticalins, the nuclear receptor fold provides a versatile scaffold for engineering molecular recognition of small molecules. Protein and metabolic engineering should greatly benefit from this general system linking the presence of a small molecule to the power of genetic selection. This system shows the feasibility of engineering nuclear receptors for practically any small molecule through directed evolution coupled to genetic selection, and for performing metabolic engineering in yeast.

3.4 Future Work

Chemical complementation has displayed the ability to link the presence of a small molecule to genetic selection in yeast providing a powerful tool for various applications involving protein ligand interactions. The system developed has shown that the fundamental theory of controlling cell survival to a small molecule is possible. However, a more complete correlation with mammalian cell culture assays, as well as an increased sensitivity would allow the system to be more versatile. In the next chapter, these issues will be investigated and discussed.

3.5 Experimental Techniques and Methods

Yeast Strain

Yeast strain, 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). This strain contains the Gal4 responsive promoter Gal1 controlling the *HIS3* gene[80].

Ligands

9-*cis* retinoic acid (MW=300.44 g/mol) was purchased from ICN Biomedicals (Aurora, OH). Paclitaxel (Taxol®) (MW=853.9 g/mol) was purchased from LKT Laboratories (St.Paul, MN). 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 (4:1v/v) and stored at 4°C. The 9-*cis* retinoic acid and rifampicin stocks were protected from light. Lanosterol (426.7 MW=402.6 g/mol), Ergosterol (MW=369.7 g/mol), and 25-hydroxycholesterol (MW=402.7 g/mol), and 4β-hydroxycholesterol (MW=402.6 g/mol) was purchased from Steraloids Inc. (Newport, RI). A 10 mM stock of the ligand was made with 100% ethanol and stored at 4°C.

Expression plasmids

pCMX-hRXRα was a kind gift from Dr. Ronald Evans (Salk Institute for Biological Studies, La Jolla, CA) [103]. pGDBRXRα (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]. pCMXRXRα mutants were made as described previously. These mutants were cloned into the pGBDRXRα vector using the pCMXRXRα vectors with the *Sal*I and *Pst*I restriction enzymes. The pGBDRXRα vector and pCMXRXRα insert cassettes were purified and ligated. Ligations were transformed

Table 3.1. Transcriptional Activation in Mammalian Cells and Genetic Selection in Yeast.

Mutant	Activation with 9cRA in mammalian cell assays	Growth with 9cRA using genetic selection in yeast
Wild-Type	+++++	Yes
L436V	+	No
F313I	++++++	No
F439L	++++++	Yes
Q275C;I310M	+++	No
Q275C;F313I	+++	No
F313I;L436V	+	No
F313I;F439L	++++++	Yes
Q275C;I310M;F313I	+	No
Q275C;I310M;F439L	+++	No

Activation of transcription key based on % of wild-type activation by 9cRA:

++++++ >110% ;+++++ 90-110%; ++++ 70-90%; +++ 50-70%, ++ 30-50%, + 10-30%.
Genetic selection in yeast based on growth of cells in presence of 9cRA and 1mM 3-AT.

into XL1-Blue competent cells. DNA from the transformants was purified using QIAprep® Spin Miniprep Kit (Qiagen Inc., Valencia, CA) and sequenced from the *Sal*I site to the stop codon. To construct pGBDPXR, site-directed mutagenesis using the QuikChange™ Site-Directed Mutagenesis method (Stratagene, La Jolla, CA) was used to introduce an *Afl*II site in the 3' untranslated region of pGBDRXRα. The *Afl*II site at the 3'end and the *EcoRI* site at the 5'end were used for removing the RXR gene, and for inserting the PXR gene. pRESTA-PXR (containing residues 130-434) was a kind gift from Dr. Matthew Redinbo (University of North Carolina, Chapel Hill, NC). Primers were designed to amplify the PXR gene with *EcoRI* and *Afl*II sites incorporated into the primers for cloning. After digestion and standard ligation, the DNA was transformed into competent XL1-Blue cells. Purified plasmids from the transformants were sequenced.

pCMXmLXRα (containing murine LXRα LBD) was a kind gift from Dr. David Mangelsdorf (University of Texas, Southwestern Medical Center; Dallas, TX) [101]. The LXRα gene was excised using the two *EcoRI* sites. pGBDRXRα contains a unique *EcoRI* site after the Gal4 DBD. A new *EcoRI* site was inserted into pGBDRXRα after the stop codon through site directed mutagenesis. Digestion of pGBDRXRα with *EcoRI* removed the RXR LBD, which was replaced with the LXRα gene, forming pGBDLXRα. pGBDLXRα transformants were sequenced. All pGBD expression vectors contain a tyrptophan (Trp) marker.

Two different plasmids were constructed to serve as controls. Gal4 activated growth was used as a ligand independent positive control in these experiments. The Gal4 gene was inserted into the same vector as RXR and PXR. Site-directed mutagenesis was performed on a derivative of the pGBDRXRα plasmid, pGBDRXRα-LSH, to introduce

an *Xma*I site in the 3' UTR of the plasmid. A unique *Nhe*I site and *Xma*I site were used to clone in the Gal4 gene. The Gal4 gene was amplified from the plasmid pCL1 using PCR with primers containing *Nhe*I and *Xma*I sites. As a negative control, an empty plasmid was constructed. pGBDRXRα-LSH was digested with the restriction enzymes *Nhe*I and *Pst*I, removing the entire Gal4 DBD and RXR LBD. The overhangs were filled with T4 DNA polymerase, ligated and transformed into bacteria. Both of these plasmids contain tryptophan markers.

Yeast selective plates and transformation

Synthetic complete (SC) media and plates were made as described previously [105]. Selective plates were made of SC media minus any one or two of the five nutrients (histidine, leucine, tryptophan, adenine, and uracil). Ligands were added to the plates after the media was cooled to about 50 °C. 3-amino-1, 2,4-triazole (3-AT) was purchased from Sigma (St. Louis, MO), dissolved in water to make 100 mM stock solutions, filter sterilized and stored at room temperature. Final concentration used on the plates was 1 mM. pGBD expression vectors were introduced into PJ69-4A using the LiAc transformation method [106] using 0.5 μg of DNA. Cells were plated onto SC -Trp plates. Plates were incubated at 30 °C for three days.

Streaking of cells onto selection plates

Yeast transformants containing the plasmid were streaked onto SC -His -Trp plates with and without 3-AT and the various ligands using sterile toothpicks. Plates were divided into four quadrants, each containing yeast cells with plasmids, plus a quadrant without any plasmids serving as a negative control. The same colony was used for streaking on all the plates, ending with a SC plate to confirm efficient transfer of the cells

to each plate. Selective plates were incubated at 30°C for up to five days; SC plates were only incubated for two days. Each set of selective plates was replicated at least once.

Chapter 4
2nd Generation Chemical Complementation

4.1 Motive for a More Sensitive Chemical Complementation

First generation chemical complementation proved that the survival of the yeast cell can be linked to the presence of a small molecule [12]. This system was shown to work with the retinoid X receptor (RXR), the pregnane X receptor (PXR), and the liver X receptor (LXR) which were compared to the Gal4 *holo* protein as the control. Although chemical complementation did work with these three receptors and their respective ligands, the characteristics of the nuclear receptor -based genetic selection varied greatly compared to Gal4-based genetic selection. With nuclear receptor chemical complementation, growth occurred after *five* days of incubation on histidine selective plates and not at all on the more stringent adenine selective plates. Gal4 stimulated growth occurred on both adenine and histidine selective media and after *two* days of incubation. The growth density on the plates were not equivalent; with the nuclear receptors and their ligands sparse growth was observed with most receptor-ligand combinations, whereas, with the Gal4 *holo* protein dense growth was observed on any selective media. Finally, in regard to ligand concentrations, growth was observed with the wild-type receptors and their corresponding ligands only at the highest concentration of ligand (10^{-5} M). This concentration is well above the EC_{50} value observed in mammalian cell culture assays; for example, the EC_{50} value for wild-type RXR and its ligand 9-*cis* retinoic acid (9cRA) is observed to be 60 nM in mammalian cell assays [84,85,103,107,108]. When a series of RXR variants were tested to determine whether a

correlation existed between chemical complementation and mammalian cell assays, a correlation was observed in some cases, but not in all cases [12].

We hypothesized that the difference in growth might be due to the evolutionary divergence between yeast and mammals, specifically in relation to the protein complexes that are involved and formed to initiate transcription. Recalling from Chapter 2, after nuclear receptors have bound ligand, these proteins undergo a conformational change. This conformational change leads to the formation of a binding site on the surface of the nuclear receptor crucial for coactivators. The coactivator proteins recruit the transcriptional machinery, including RNA polymerase. In the first generation chemical complementation system (Chapter 3), when the ligand bound to the nuclear receptor, yeast coactivators were recruited to the nuclear receptor and transcription was initiated. The coactivator proteins in yeast that bind the human nuclear receptors are similar enough to human coactivator proteins to allow chemical complementation to occur. However, divergence has occurred between the human and yeast coactivators so the interactions between the human and yeast coactivators are likely not optimal, producing the growth disparity compared to Gal4.

In an attempt to overcome these differences in the transcription machinery between the two organisms, an adapter protein consisting of a fusion between a human coactivator and a yeast transcriptional activator is introduced to the system. With the addition of the fusion protein, after the nuclear receptor binds to its ligand and induces a conformational change, the human receptor will be in direct contact with a human coactivator, recreating the interactive environment between the two human proteins.

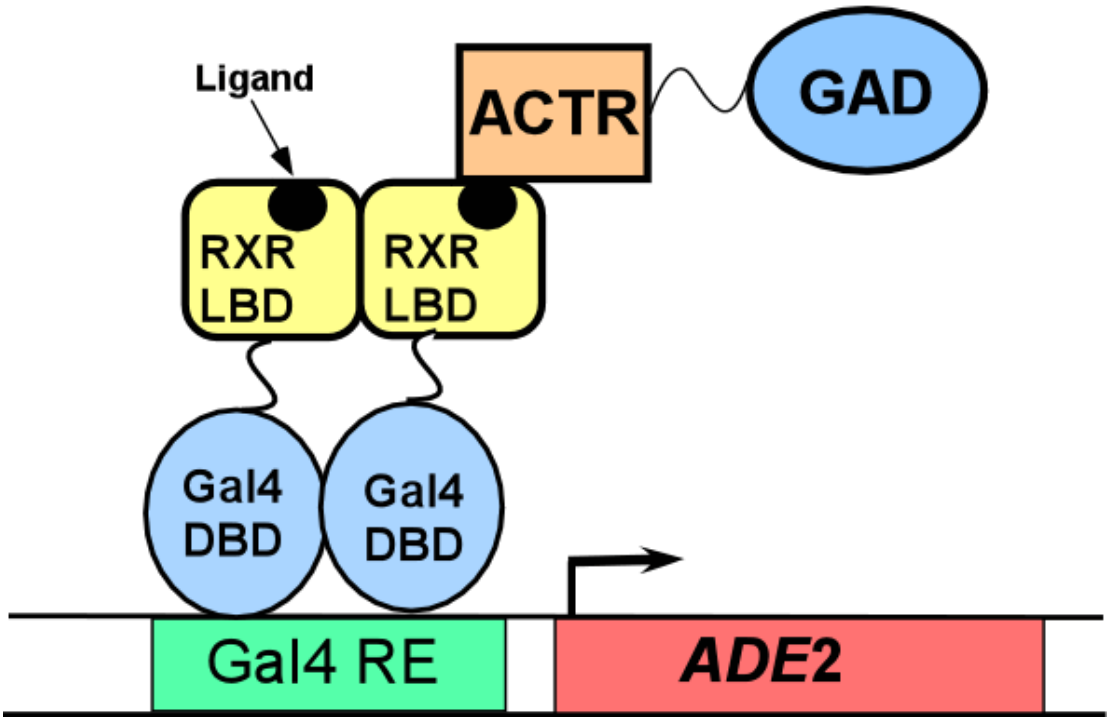


Figure 4.1. 2nd Generation Chemical Complementation.

The human coactivator fused to the yeast transcriptional activator will also allow optimal interactions between this protein and yeast coactivators involved in transcription.

This three-component chemical complementation system was developed to determine whether the sensitivity of chemical complementation can be increased and more Gal4-like growth to occur. This 2nd generation chemical complementation system comprises the nuclear receptor LBD fused to the Gal4 DBD, an adapter fusion protein between the human coactivator (SRC-1 or ACTR) and the Gal4 activation domain GAD), and a small molecule ligand (Figure 4.1) [72,76]. This three-component system serves as a bridge between the human nuclear receptor and the yeast transcriptional machinery, creating a human model of transcriptional activation in yeast, thus humanizing the yeast response.

4.2 Results with 2nd Generation Chemical Complementation

Creation of a Highly Sensitive Chemical Complementation with ACTR.

The human nuclear receptor coactivator ACTR was fused to the yeast Gal4 activation domain. This plasmid, pGAD10BAACTR, expresses the ACTR:GAD fusion protein and contains a leucine marker. This plasmid was co-transformed into yeast with the plasmid pGBDRXR, which expresses the Gal4 DNA binding domain (DBD) fused to the RXR ligand binding domain (GBD:RXR) and contains a tryptophan marker. Transformants were selected on SC -Leu -Trp plates, and were streaked onto adenine selective plates (SC -Ade) containing 10^{-5} M 9cRA, a known ligand for RXR (Figure 4.2). Yeast containing just the pGBDRXR plasmid, the pGAD10BAACTR plasmid, a plasmid with just the Gal4DBD (pGBDMT), and a plasmid containing the Gal4 *holo* protein (pGBT9Gal4) were also streaked onto these plates as controls.

After two days of incubation, growth occurs on the sector of the plate containing ACTR:GAD with GBD:RXR and on the sector of the plate with Gal4, whereas no growth occurs on the sector of the plate with GBD:RXR alone (Figure 4.2). The growth density produced by GBD:RXR and ACTR:GAD is the same as the growth produced by the *holo* Gal4. Importantly, GBD:RXR and ACTR:GAD produced no growth on plates without 9cRA (Figure 4.2).

To determine the sensitivity of this system, a dose response experiment was performed on adenine selective plates (SC -Ade) containing ligand concentrations ranging from 10^{-5} M to 10^{-9} M. After two days of incubation, a clear dose response occurs on the plates (Figure 4.2). Without ligand, growth occurs only on the Gal4 sector of the plate, as expected. At concentrations as low as 10^{-8} M 9cRA, ligand-activated growth occurs only on the sector of the plate containing both GBD:RXR with ACTR:GAD (Figure 4.2). At concentrations of ligand above 10^{-8} M, higher density growth is observed on the sector of the plate containing GBD:RXR with ACTR:GAD. No growth occurs with GBD:RXR alone as expected.

Creation of a Highly Sensitive Chemical Complementation with SRC-1.

Another RXR coactivator was tested to increase the sensitivity of chemical complementation. Residues 54 to 1442 of the human nuclear receptor coactivator, SRC-1, were fused to the Gal4 activation domain to construct the plasmid pGAD10BASRC1. This plasmid, which expresses SRC1:GAD in yeast and contains a leucine marker was transformed with GBD:RXR; transformants selected from SC -Leu -Trp were streaked onto adenine selective plates (SC -Ade) with various concentrations of 9cRA (Figure

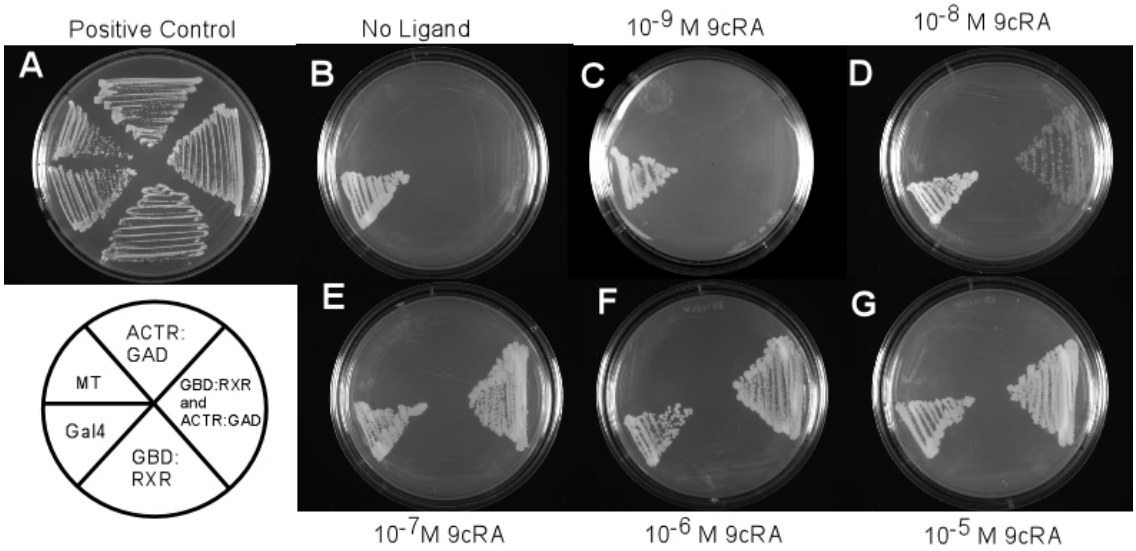


Figure 4.2. Chemical Complementation with Wild-Type RXR and ACTR.

4.3). Ligand-activated growth is observed only in the sector of the plate containing both GBD:RXR with SRC1:GAD, and the same trend is observed with SRC-1 as the ACTR coactivator (Figure 4.3).

To verify that the increased sensitivity is from specific interactions between the coactivator and the active conformation of the receptor, a series of further controls was devised and implemented (Figure 4.4). pGAD10, a plasmid containing the Gal4 activation domain (GAD) without a coactivator domain was cotransformed with pGBDRXR. The plasmid was also transformed alone. pGAD10BAACTR, pGAD10BASRC1, pGBT9Gal4, and pGBDMT were all transformed individually. These controls were streaked onto adenine selective plates (SC -Ade) with and without 9cRA. In the absence of ligand, only the entire Gal4 gene (pGBT9Gal4) grows as expected. In the presence of 10^{-5} M 9cRA, growth occurs with the GBD:RXR with ACTR:GAD and GBD:RXR with SRC1:GAD. The Gal4 AD only with GBD:RXR displays no growth (Figure 4.4). These results verify that the increase in chemical complementation is specifically due to the interaction of the coactivator fusion protein with the ligand-bound nuclear receptor.

Determination of specific interactions.

To verify that the increased sensitivity is from specific interactions between the coactivator and the active conformation of the receptor, a series of controls was devised. pGAD10, a plasmid containing the Gal4 activation domain (GAD) without a coactivator domain was cotransformed with pGBDRXR. The plasmid was also transformed alone. pGAD10BAACTR, pGAD10BASRC1, pGBT9Gal4, and pGBDMT were all transformed

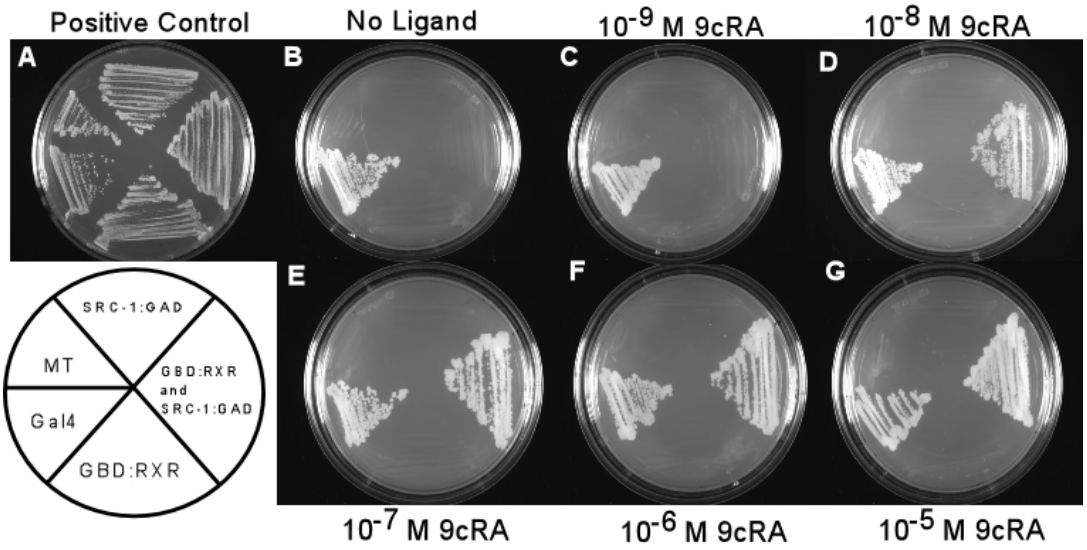


Figure 4.3. Chemical Complementation with Wild-Type RXR and SRC-1.

individually. These controls were streaked onto adenine selective plates (SC -Ade) with and without 9cRA. In the absence of ligand, only the sector of the plate containing the entire Gal4 gene (pGBT9Gal4) grows as expected. In the presence of 10^{-5} M 9cRA, growth occurs on the sector of the plate containing GBD:RXR with ACTR:GAD and GBD:RXR with SRC1:GAD. The sector of the plate containing the Gal4 AD only (without the coactivator domain) with GBD:RXR displays no growth. These results verify that the increase in chemical complementation is specifically due to the interaction of the coactivator fusion protein with the ligand-bound nuclear receptor.

Correlating Variant RXR Function in Yeast and Mammalian Cells.

Several RXR variants previously tested in both mammalian cell assays and with the two-component chemical complementation in yeast (without the coactivator fusion protein) showed different activities in the two cell types [12,29]. Without the coactivator fusion protein, ligand-activated growth was observed only with wild-type RXR and the F439L mutant after five days of incubation; none of the other variants showed ligand-activated growth. We postulated that the differences in the transcription machinery lead to the different patterns in activation. To test whether the adapter fusion protein could overcome the differences and show a direct correlation of activities in the two cell types, all the variants in Table 4.1 were cloned into pGBD vectors and cotransformed into yeast with pGAD10BAACTR. Transformants were selected from -Leu -Trp plates and then streaked onto adenine selective plates (SC -Ade -Trp). These mutants were tested with 9cRA and LG335 (a near-drug, a synthetic compound

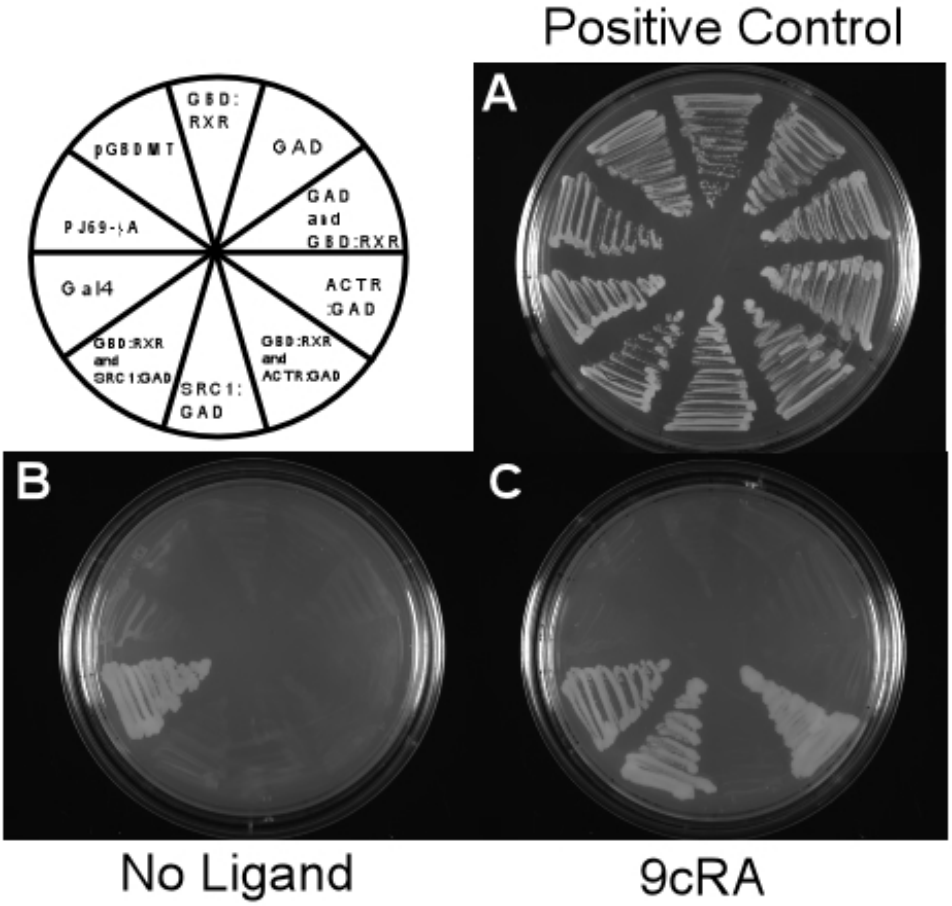


Figure 4.4. Verifying Specific Interactions of Chemical Complementation.

structurally similar to the RXR agonist but that does not activate wild-type RXR) (Table 4.1).

The transcriptional activation patterns of these variants in chemical complementation with the addition of ACTR:GAD was tested on dose response solid media plates containing increasing concentrations of 9cRA or the synthetic ligand, LG335 (Figure 4.5). On the plate without ligand, growth occurs on the sector of the plate containing Gal4, but growth also occurs on the sector of the plate with the two mutants F313I and F313I;F439L. We speculate that this is a result of the mutations causing a structural modification to the binding pocket that is favorable for the binding of an endogenous small molecule in yeast. At 10^{-5} M 9cRA, growth occurs on the sectors of the plate with the single mutants, C432G, Q275C, I268F, I310M, V342F, and F439L, as well as some of the triple mutants I310M;F313I;F439L and Q275C;F313I;V342F. As the concentration of ligand decreases, some mutants no longer show ligand-activated growth. At 10^{-7} M 9cRA, growth is observed with the F439L variant as well as wild-type RXR (Figure 4.5). At the lowest concentration of ligand, 10^{-8} M 9cRA, growth is observed in the Gal4, F313I, and F313I;F439L sectors of the plates.

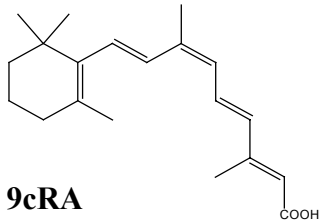
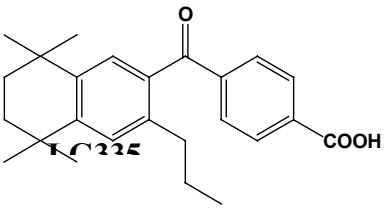
For the synthetic ligand LG335, growth is observed with several of the single, double and triple variants at 10^{-5} M (Figure 4.5). At lower concentrations of ligand, the single substitution variants do not show much growth. However, several of the double and triple substituted variants I310M;F313I;F439L, Q275C;F313I, and I310M;F313I display ligand- activated growth at 10^{-7} M LG335. At 10^{-8} M LG335, some growth is still observed in the I310M;F313I;F439L sector of the plate.

4.3 Integration of the ACTR gene into PJ69-4A

The fusion protein consisting of the nuclear receptor coactivator ACTR and the Gal4 activation domain (GAD) increased the sensitivity of chemical complementation and provided a more complete correlation with mammalian cell assays. When this three-component system was initially developed, the GBD:RXR LBD fusion was placed on a yeast expression plasmid containing a tryptophan marker, and the coactivator fusion protein, ACTR:GAD, was on a separate yeast expression plasmid containing a leucine marker, and for chemical complementation both plasmids had to be transformed into yeast. Yeast transformations are quite efficient, but as the amount of DNA added to the competent yeast cells increases, the transformation efficiency decreases. For creating large combinatorial libraries of proteins using chemical complementation, higher transformation efficiencies are desired for creating more complete libraries. Furthermore, as the complexity of the chemical complementation system increases and more proteins are analyzed in this system, yeast expression plasmids with various selective markers need to be available. Therefore, to make chemical complementation more generally applicable for protein and enzyme engineering, a yeast strain containing the integrated nuclear receptor coactivator: Gal4 activation domain would provide more versatility. This strain will allow higher transformation efficiencies for creating large protein libraries for various applications, such as protein engineering, because only one plasmid will be introduced rather than two and the leucine marker can be used for other plasmids.

To create a new strain containing the ACTR gene, a method known as homologous recombination was used to integrate the ACTR gene into PJ69-4A.

Table 4.1. Activation of Transcription by Wild-Type RXR and Variant Receptors in CV-1 Cells and in Yeast Strain PJ69-4α with pGAD-ACTR Plasmid and pGBDRXR-Variant Plasmids.

Ligand	 **9cRA**			
Variant	Activation in Mammalian Cell Assays	Chemical Complementation in Yeast	Activation in Mammalian Cell Assays	Chemical Complementation in Yeast
Wild-type	+++++	+++++	---	---
I268F	+++++	+++++	---	---
Q275C	++++++	++++++	---	---
L326F	++	+	---	---
C432G	++++++	++++++	---	---
L309V	+++	---	---	---
L436V	+	---	---	---
I310M	++++	+++++	+	+
F313I	++++++	Constitutive	++++	Constitutive
V342F	++	++	---	---
F439L	++++++	+++++	+++	+++
Q275C;I310M	+++	---	+	---
Q275C;F313I	+++	---	+++++	+++++
I310M;F313I	++++	++	+++++	+++++
F313I;L426V	+	---	+++	+++++
F313I;F439L	++++++	Constitutive	++++++	Constitutive
Q275C;I310M;F313I	+	---	++++	+++++
Q275C;F313I;V342F	+++	+++	+++	+++++
Q275C;L436V;F439L	+++	+	+	---
I310M;F313I;F439L	+++++	++++	+++++	+++++

†Key: ++++++ > 110% of wild-type activation by ligand; +++++ 90-110% of wild-type activation by ligand; ++++ 70-90% of wild-type activation by ligand; +++ 50-70% of wild-type activation by ligand; ++ 30-50% of wild-type activation by ligand; + 10-30% of wild-type activation by ligand (estimated uncertainty one +).

‡Key: ++++++ > 110% of pGBT9-Gal 4 activation by ligand; +++++ 90-110% of pGBT9-Gal 4 activation by ligand; ++++ 70-90% of pGBT9-Gal4 activation by ligand; +++ 50-70% of pGBT9-Gal 4 activation by ligand; ++ 30-50% of pGBT9-Gal 4 activation by ligand; + 10-30% of pGBT9-Gal 4 activation by ligand (estimated uncertainty one +).--- = no growth

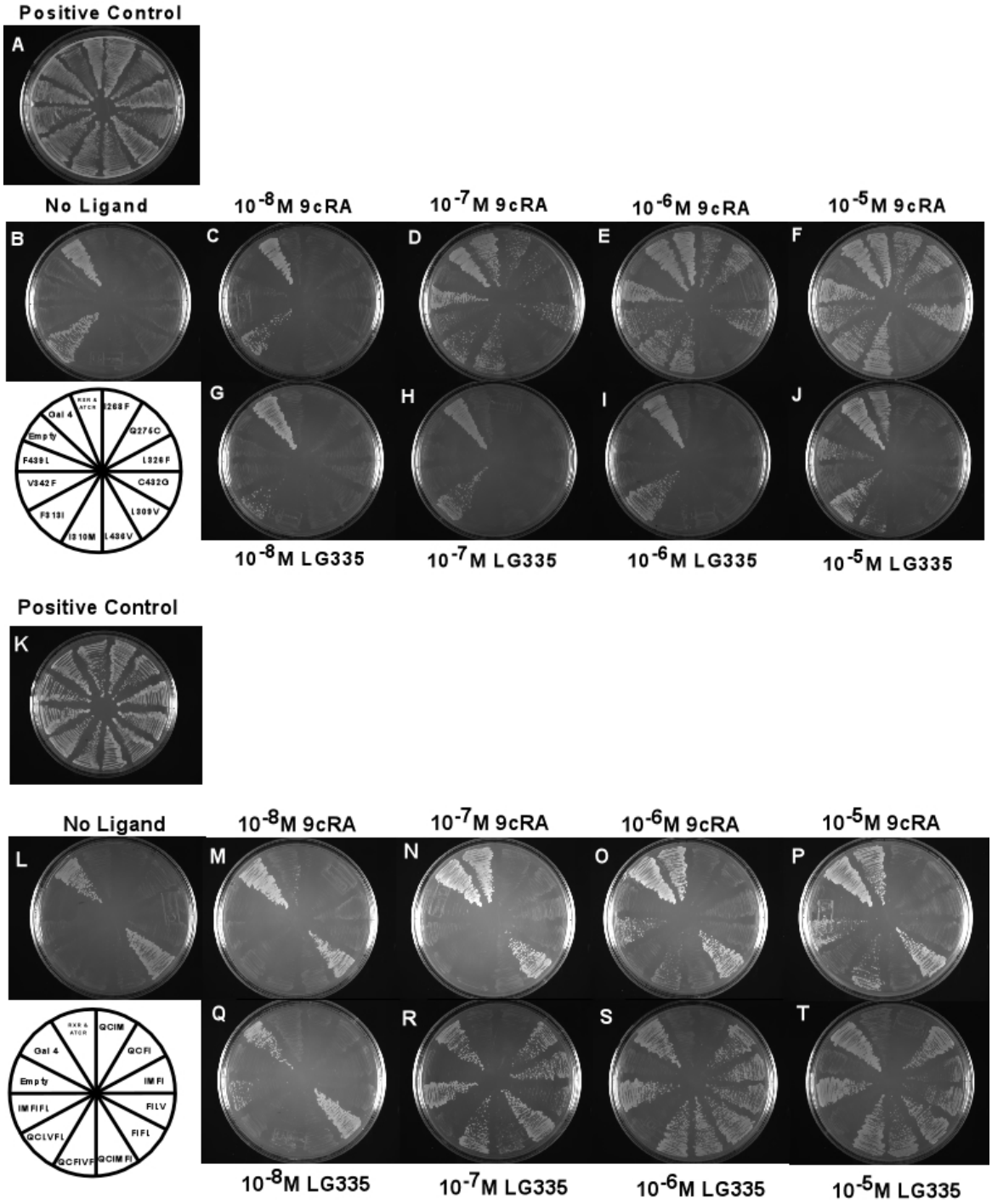


Figure 4.5. RXR Variants in Chemical Complementation with ACTR.

Homologous recombination is a tool endogenous in yeast that can be used to introduce exogenous DNA into an organism's genome, creating new mutant strains that can be used for various applications, such as cloning and genetic selection and screening techniques. During homologous recombination, the DNA of interest is integrated at a specific location in the genome by flanking the exogenous DNA sequence of interest with sequences of the genome target site. We decided to integrate the GAD:ACTR fragment into the mutated *leu2-3* gene of PJ69-4A. The *leu2-3* gene is a derivative of the wild-type *LEU2* gene, a gene involved in leucine biosynthesis. A mutation introduced into the *LEU2* gene causes a frame shift, leading to the inactive leucine gene, *leu2-3*. The pGAD10BAACTR plasmid contains a *LEU2* gene as a selection marker, and linearizing this plasmid with a restriction enzyme in the *LEU2* gene would create a fragment containing the GAD:ACTR gene flanked by portions of the *LEU2* gene (Figure 4.6). Studies have shown that integration of a DNA sequence into a genome is enhanced by introducing a double-stranded break into the plasmid, increasing the frequency of transformation by 10 to 1000-fold compared to uncut circular plasmid [109].

Two steps were taken to ensure and increase the efficiency of integration of the linear fragment containing the Gal4AD: ACTR into PJ69-4A. In the first step, the pGAD10BAACTR plasmid contains a 2μ origin of replication, a sequence of DNA that allows plasmid to replicate autonomously in yeast [109]. In the absence of these autonomously replicating sequences (ARS), DNA transformed into the yeast cell integrates into the yeast genome by homologous recombination, therefore greatly reducing the number of false positives; these are cells which contain the plasmid- based Gal4 AD:ACTR fragment. The pGAD10BAACTR was digested with two restriction

enzymes, *AatII* and *SnaBI*, which encompass the 2μ origin of replication. Once the 2μ origin of replication was removed from the plasmid, the plasmid was religated to form the pGAD10BAACTRwout2μ plasmid (the plasmid with the Gal4AD: ACTR but without the 2μ origin of replication). The second step to ensure the integration of the Gal4AD: ACTR fragment into PJ69-4A involves treating the ends of the linearized fragment with calf intestinal phosphatase (CIP). CIP catalyzes the removal of 5'-phosphate groups from DNA, and since CIP-treated fragments lack the 5'-phosphoryl termini required by ligases these fragments can not self-ligate.

A *SalI* site was added to the *LEU2* gene of pGAD10BAACTRwout2μ plasmid using site-directed mutagenesis (Stratagene, USA). The *SalI* site was added to allow the formation of a linear fragment containing portions of the *LEU2* gene encompassing the Gal4 AD: ACTR fragment. This plasmid was digested with the *SalI* enzyme, and subjected to CIP (Figure 4.6). The linear fragment was then transformed into PJ69-4A, along with pGBDPXR using the LiAc yeast transformation protocol [106].

Transformants were initially selected for growth on growth on histidine and adenine selective plates (-His -Trp and -Ade -Trp respectively). PXR was found to be constitutively active in 2nd generation chemical complementation with ACTR; constitutively active in chemical complementation implies that without the presence of ligand, the receptor is activated. It is possible these receptors are activated by some endogenous ligand (discussed more in Chapter 7). Therefore, transforming PXR into this strain would provide an assay, which would not require a ligand on the selective media; cells containing the pGBDPXR plasmid and the ACTR: GAD fusion are able to grow on selective media without ligand. Once the transformants are plated onto selective plates,

adenine or histidine plates without ligand, growth will only be observed if the ACTR: GAD fragment has integrated into the genome. This is because the only source of the ACTR: GAD fusion will be the integrated gene and cells containing PXR can only survive on selective media if the ACTR: GAD fusion is present. Depending on the efficiency of the CIP, a population of the linear fragment may not have treated ends, therefore, contamination by the pGAD10BAACTRwout2μ plasmid (full circular plasmid containing ACTR: GAD) might be present. To address this issue, a diagnosis on the transformants was performed on leucine selective media. If integration of ACTR was successful at the *leu2-3* in the genome, this gene would remain non-functional due to the integration of the ACTR: GAD fragment in between the two portions of the leucine gene. If however, the linear fragment had religated to form a functional plasmid, the *LEU2* gene on the plasmid would be functional, allowing the cells to survive on media lacking leucine. Several hundred colonies were obtained as a result of the transformation and the colonies were screened to ensure their -Leu phenotype.

Solid media selective plates containing the transformants were replica plated onto selective media lacking leucine and tryptophan (-Leu -Trp) and on non-selective media (synthetic complete, SC media). Colonies that survived on the SC plates but were unable to grow on the leucine and tryptophan (-Leu -Trp) selective media were chosen, because these transformants have a non-functional leucine gene. From this analysis, about 50 possible transformants showed a non-functional leucine gene, and were streaked onto non-selective media (YPD media) to allow these cells to lose the pGBDPXR plasmid. This is because only the transformants that are able to maintain the integrated fragment stably were of interest. Thus to ensure that the linear fragment had been integrated and

maintained stably into the yeast genome, once these candidates had lost the pGBDPXR plasmid, these candidates would then be retransformed with pGBDPXR and tested again. These cells were grown in non-selective liquid media for 180 doubling times and plated onto non-selective plates (YPD) and on selective plates lacking tryptophan (SC -Trp).

Prior to re-transformation of these candidates with the nuclear receptor, yeast genomic DNA was isolated from each of the candidates. These genomic preparations were then subjected to multiple rounds of PCR with different primers to determine whether the linear fragment had actually integrated into the genome, and to determine whether integration had occurred at the right location in the genome. As shown in Figure 4.6, two sets of PCR primers were used to determine the following: (1) primers C and D showed whether the ACTR gene had actually integrated into the genome and (2) primers A and B: showed that the integration had actually occurred in the *leu2-3* gene. The PCR results with primers A and B show that the integration did not occur in the *leu2-3* gene, because if integration had occurred, a band of 12 kilobases (kb) would have been observed on the gel. However, in looking at the gel, a band size of about 1.5 kb is observed, indicating the size of the *leu2-3* gene (Figure 4.6). Furthermore, PCR results from the amplification of the genome using primer C and D and C and E, which amplify a portion of the ACTR and the full ACTR gene respectively, both show that the gene is present in the genome. A PCR product of about 1.5 kb and 4 kb are expected to be seen when amplified with primers C and D and C and E respectively, and both of these PCR fragments are observed on a gel, as shown in Figure 4.6. Therefore, although the exact integration site of the ACTR gene is unknown, the gene has been integrated into the PJ69-4A genome and the strain was named BAH14RPJ69.

After BAH14RPJ69 has lost the initial PXR plasmid that had been transformed, the strain was re-transformed with PXR. pGBDPXR was transformed into BAH14RPJ69 using the standard LiAc transformation protocol, and the transformation mix was plated onto adenine and histidine selective plates (-Ade -Trp and -His -Trp) [106]. Growth was observed on both of the selective plates, no growth was seen with the negative control, the strain without pGBDPXR, on the selective media plates.

4.4 Discussion of 2nd Generation Chemical Complementation

Development of a Highly Sensitive Three-component Chemical Complementation.

We present a novel, highly sensitive three-component chemical complementation system in *Saccharomyces cerevisiae*. Previously, we reported a two-component chemical complementation in yeast using wild type RXR and 9cRA. While the survival of the yeast could indeed be linked to a small molecule ligand, ligand-activated growth was only observed at 10^{-5} M, the highest possible concentration of ligand (higher concentrations precipitate). Furthermore, the growth time was much longer and the density much less than the growth due to the ligand-independent yeast transcription factor Gal4. Growth using the *HIS3* marker required five days of incubation and no ligand-activated growth occurred with the more stringent *ADE2* marker, whereas Gal4 growth required only two days of incubation and produced growth with the adenine marker. We hypothesized that the evolutionary gap between yeast and humans is small enough to allow human proteins to function in yeast, but have diverged enough that transcriptional interactions are not optimal.

The introduction of the adapter fusion protein ACTR: GAD in chemical complementation increases the sensitivity 1000-fold over the two-component system, because growth occurs at concentrations as low as 10^{-8} M 9cRA instead of 10^{-5} M 9cRA. Additionally, growth occurs on adenine selective plates with 9cRA after two days of incubation (Figure 4.2) versus no growth within five days for the two-component system[12]. With chemical complementation, an apparent EC_{50} value occurs between 10^{-8} M and 10^{-7} M for wild-type RXR and 9cRA (Figure 4.2), which is comparable to the EC_{50} value measured for wild-type RXR in mammalian cell assays (~10^{-7} M) [110]. The growth density and rate with the ACTR:GAD fusion protein is comparable to Gal4 activated growth. The same results were obtained on adenine selective plates (SC -Ade -Trp and SC -Ade -Leu -Trp) and on histidine selective plates.

The ACTR:GAD fusion protein also humanizes the yeast response, allowing a complete correlation of nuclear receptor function in yeast and in mammalian cell culture. A variety of RXR variants that had been previously tested in mammalian cells with two-component chemical complementation showed different activities in the two cell types (only two of the variants showed the same activity in yeast as mammalian cells). With the exception of constitutively active variants, a complete correlation is apparent between yeast growth and transcriptional activation in mammalian cells (Table 4.1). We speculate that the constitutively active variants are activated by some endogenous small molecule in yeast that is not present (or present at a lower concentration) in mammalian cells. The otherwise complete correlation between chemical complementation and mammalian cell assays shows that the coactivator fusion protein (ACTR:GAD) serves to bridge millions

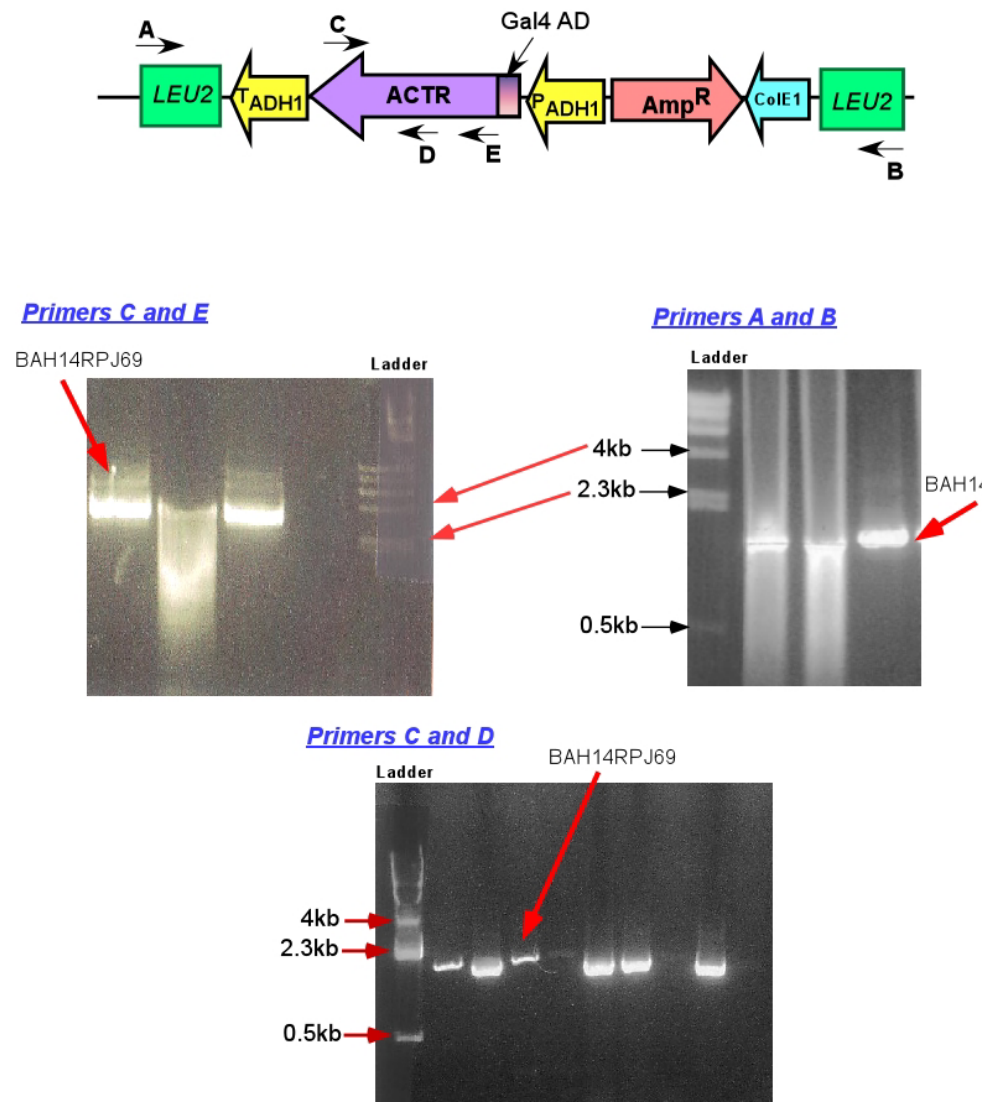


Figure 4.6. Creation of BAH14RPJ69, the PJ69-4A strain with the integrated ACTR

of years of evolution by adapting mammalian nuclear receptor function to the yeast transcription machinery.

In summary, introducing an adapter fusion protein of the human coactivator with the Gal4 activation domain increases the sensitivity of chemical complementation 1000-fold, making this system highly efficient for detection and analysis of protein-ligand interactions. Additionally, introduction of the adapter fusion protein humanizes the yeast response. This chemical complementation system is highly sensitive (producing growth at 10 nM 9cRA), with zero background (no growth without ligand), and very high dynamic range (growth density and time equivalent to Gal4 induced growth).

4.5 Discussion on BAH14RPJ69

To enhance the versatility of chemical complementation, we developed a derivative of the PJ69-4A strain that contains an ACTR gene integrated into the genome. The integration of the ACTR gene would increase the versatility of chemical complementation in two main ways: (1) allow one less plasmid to be transformed each time chemical complementation is utilized and (2) allow the use of plasmids with a leucine selective marker for transforming other forms of exogenous DNA. These two reasons are important for the various applications of chemical complementation. For protein and enzyme engineering, large combinatorial protein libraries are analyzed. Using chemical complementation greater transformation efficiencies would allow larger protein libraries to be analyzed. Transforming one plasmid rather than two plasmids would increase the transformation efficiency because as more plasmids are introduced into a cell, the transformation efficiency decreases. Yeast tend to have lower

transformation efficiencies than other microorganisms used for protein engineering, such as *E. coli* (10^6 transformants compared to 10^9 transformants), therefore, any measure that can increase the transformation efficiency should be used. We have been successful in creating the BAH14RPJ69 strain that contains the ACTR: GAD linear fragment in its genome, and the construction of this strain would greatly enhance the versatility of chemical complementation in various applications, such as protein and enzyme engineering.

4.6 Future Work

We have shown that chemical complementation can link the presence of a small molecule to genetic selection, and have improved the sensitivity of chemical complementation to correlate with mammalian cell culture assays. Thus far, chemical complementation has analytically been based on qualitative observation; the transcriptional activation measurements for chemical complementation is based on growth observed on plates by the human eye. In the next chapter, the focus is on developing a quantitation assay of chemical complementation, where we quantitate the amount of transcriptional activation by measuring optical density of growth as well as using the *lacZ* reporter gene.

4.7 Materials and Methods

Yeast Strain.

Yeast strain, 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 and Dr. Elizabeth Craig (University of Wisconsin, Madison)[80]. This strain contains three reporter genes, *HIS3, ADE2,* and *lacZ*, which are under the control of different Gal4 inducible promoters[80].

Ligands.

9-*cis* retinoic acid (9cRA) was purchased from ICN Biomedicals. LG335 was synthesized[13]. 10 μM stocks of each ligand were protected from light and dissolved in 80% ethanol:20% DMSO (4:1 v/v) and stored at 4° C.

Expression plasmids.

pCMX-hRXRα, pGDBRXRα, and pCMXRXRα mutants were described previously. The pCMXRXRα mutants were cloned into the pGBDRXRα vector using standard cloning methods. Two different plasmids were constructed to serve as controls: pGBT9Gal4, expressing *holo* Gal4 serves as a positive control and pGBDMT, which is the pGBD plasmid without the Gal4 DBD: RXR LBD construct serving as a negative control. Both of these plasmids have been described previously [5]. Both control plasmids and all the pGBD vectors contain tryptophan markers.

Plasmids expressing the fusion protein of the Gal4 activation domain with the coactivators are based on the commercial plasmid pGAD10 (Clontech, USA). The pGAD10 vector contains the Gal4 activation domain (residues 491-829) fused to a multiple cloning site (MCS) and uses a leucine marker. Additional restriction enzyme sites were added to the MCS of the plasmid via site directed mutagenesis. Primers were designed to add the following restriction enzymes: *NdeI, EagI, EclXI, NotI, XmaIII,*

XmaI, and *SmaI*, forming a new plasmid known as pGAD10BA. This plasmid was sequenced and used for specific interaction studies mentioned in the results.

pCMX-ACTR, the expression plasmid for the human nuclear receptor coactivator ACTR, was a kind gift from Dr. Ron Evans (Salk Institute for Biological Studies, La Jolla, CA)[72]. pCR3.1hSRC-1, the expression plasmid for the human nuclear receptor coactivator SRC-1, was a kind gift from Dr. Bert O'Malley (Baylor College of Medicine, Houston, TX) [76]. Both ACTR (residues 1-1413) and SRC-1 (residues 54-1442) genes were amplified via PCR with primers that contained *BglII* and *NotI* sites. The PCR products were digested with the two restriction enzymes and cleaned using the Zymo "DNA Clean and Concentrator Kit" (Zymo Research, Orange, CA) spin columns. pGAD10BA was digested with *BglII* and *NotI* and ligated with both the ACTR and SRC-1 products. Ligations were transformed into Z-competent (Zymo Research, Orange, CA) XL1-Blue cells (Stratagene, La Jolla, CA). Transformants were rescued and sequenced. The final plasmids are called pGAD10BAACTR and pGAD10BASRC1.

Yeast selection plates and transformation.

Synthetic complete (SC) media and plates were made as described [105]. Selective plates were made of SC media minus any of five nutrients (histidine, leucine, tryptophan, adenine, and uracil). Ligands were added to the plates after the media cooled to 50 °C.

Yeast transformations introduced the pGBD expression vectors into PJ69-4A and MaV103 using the LiAc transformation method with 0.5 μg of DNA [106]. Cells with pGBD vectors were plated onto SC -Trp plates. Cells transformed with both pGBD and

pGAD vectors were plated onto SC -Leu -Trp plates. Transformation plates were incubated at 30 °C for three days.

Streaking cells onto adenine selective plates using PJ69-4A.

Yeast transformants containing the plasmids were streaked onto the selective plates (SC -Ade) with different ligand concentrations using sterile toothpicks. Plates were divided into sectors for the samples and controls; the control sectors contain pGBDMT and pGBT9Gal4. The same colony was used for streaking on all the plates, ending with a SC plate to confirm efficient transfer of the cells to each plate. Both selective and non-selective plates were incubated at 30 °C for two days. Each set of genetic selection plates was replicated at least once.

Creating the BAH14RPJ69 strain.

To create the BAH14RPJ69 strain, the pGAD10BAACTR plasmid (2 μg) described above was digested with *SnaBI* and *AatII* to remove the ARS sequence from the plasmid. After an overnight digestion with the restriction enzymes, the vector cassette was purified using the Gel Recovery Kit (Zymo Research, Orange, CA). The vector cassette was treated with 1 μL T4 DNA polymerase to remove the overhangs and purified using the DNA Clean and Concentrator Kit (Zymo Research, Orange, CA). The purified vector was subjected to a ligation using the Quick Ligase Kit (New England Biolabs, MA) with incubation at 22 °C overnight. 10 μL of the ligation mix was transformed into Z-competent XL1-Blue cells and transformants were sequenced for confirmation. This plasmid was name pGAD10BAACTRwout2μ.

A *SalI* site was inserted into the *LEU2* gene on the pGAD10BAACTRwout2μ plasmid via site-directed mutagenesis. Primers were ordered from MWG Biotech (High Point, NC) with the *SalI* site for insertion in the *LEU2* gene and these primer stocks were made (100ng/ μL). The primers, dNTP, 10 X *Pfu* Buffer, and 100 ng of pGAD10BAACTRwout2μ were added to a PCR and 1 μL of *Pfu polymerase* (Stratagene, USA) was added. The PCR reaction was digested with *DpnI* to digest away any remaining parent plasmid, and 5 μL of the PCR mixture was transformed into Z-competent XL1-Blue cells (ZymoResearch, CA). Transformants were subjected to digest with *SalI* to make sure that the mutagenesis had occurred. This plasmid was called pGAD10BAACTRwout2μ*SalI*.

After the pGAD10BAACTRwout2μ*SalI* had been made, 4 μg of the vector cassette was digested with *SalI* overnight. The digest was treated with 1 μL of *calf intestinal phosphatase* (*CIP*), and purified with the DNA Clean and Concentrator Kit (ZymoResearch, CA). The DNA was then transformed into PJ69-4A and transformants were selected on histidine and adenine selective plates (see above for details on making the plates). The plates were incubated at 30 °C for three to five days. Transformants were streaked onto adenine selective plates and leucine selective plates were selected for PCR and for removal of the pGBDPXR plasmid.

To determine whether integration in the genome had occurred, yeast chromosomal DNA was isolated by growing the yeast in yeast peptone dextrose (YPD) media overnight. The cells were harvested after an overnight growth period and resuspended in 0.5 mL of water. The resuspended cells were broken open with 200 μL of "breaking buffer", which consists of 2% Triton X 100, 1% SDS, 100 mM NaCl, 10 mM

Tris-HCl pH 8, and 1 mM EDTA pH 8. Glass beads (0.3 g) and 200 μL of a phenol/c*holo*form/isoamyl alcohol were also added. This mixture was vortexed at the highest speed for 3 minutes. 200 μL of TE (Tris, EDTA) buffer was added and again the mixture was vortexed. These cells were then centrifuged and the aqueous layer was transferred to a clean microcentrifuge tube. 1 mL of 100% ethanol was added to the aqueous layer. This mixture was centrifuged again and the supernatant was removed and the pellet was resuspended in 0.4 mL TE buffer. 30 μL RNase (1 mg mL^{-1}) was added to the mixture and incubated at 37 °C for five minutes. 4 M ammonium acetate (10 μL) and 1 mL 100% ethanol were added and mixed. The mixture was centrifuged at high speed (room temperature) and the supernatant was discarded. The pellet was dried and resuspended in 100 μL of TE buffer. The isolated chromosomal DNA was run on an agarose gel to determine its purity, and it was determined that the yield of the chromosomal DNA is about 5 μg. The chromosomal DNA along with primers A, B, C, D, and E were all placed in tubes with FailSafe Kit (Epicentre, USA) and PCR fragments were amplified. The candidates with the correct band sizes were chosen for retransformation with pGBDPXR.

To segregate the pGBDPXR plasmid from the potential candidates, the candidates were grown in 5 mL YPD media overnight (30 °C, shaking 300 rpm) and plated onto YPD plates after the overnight growth period. If the cells contain the pGBDPXR plasmid, the cells will have a white phenotype, and if they have lost the plasmid, they will have a red phenotype. This is because wild-type PJ69-4A is deficient in making adenine and the build-up of one of the adenine precursors gives the cells its red phenotype. If the adenine biosynthetic pathway is restored and functional, the

intermediate will no longer be accumulating and the will confer the cells to a white phenotype. Therefore, if pGBDPXR is present along with ACTR, these cells will contain a functional adenine biosynthetic pathway and confer a white phenotype. If however, the pGBDPXR plasmid is lost, the cells will no longer have a functional adenine biosynthetic pathway, and the cells will have a red phenotype. After the initial plate onto YPD plates, the cells displayed either a white or light pink phenotype. The cells with a light pink phenotype were grown again in YPD and plated onto YPD plates after the overnight growth. Cells that displayed a red phenotype and showed a positive result in the chromosomal PCR were re-transformed with pGBDPXR.

1 μg of pGBDPXR was transformed into the candidate cells using the standard LiAc transformation protocol. As a control, the cells were not transformed with pGBDPXR. Both sets of samples were plated onto selective media, and growth was observed on the selective plates only in the presence of pGBDPXR and no growth was observed without the plasmid.

Chapter 5
Quantitation of Chemical Complementation

5.1 Motive for a Quantitation Assay

In the previous chapters, we defined chemical complementation as a generalizable genetic selection system in *S. cerevisiae* (yeast) for evaluating protein-ligand interactions, where only in the presence of the ligand do the yeast survive. In the initial development of chemical complementation, observations were made using selective plates where the amount of growth was strictly made on a qualitative, visual basis. In most cases, the amount of growth was based on the density of the cells on the sector of the plate in comparison to the control Gal4. Although this qualitative analysis was adequate for the initial stages of developing chemical complementation, to enhance this system for applications in drug discovery and protein engineering, a more quantitative analysis of chemical complementation is required. This quantitative assay for chemical complementation aims to: (1) differentiate between strong and weak agonists/antagonists, (2) develop a high-throughput assay for drug discovery using nuclear receptors, and (3) provide an assay for analyzing individual colonies of combinatorial protein libraries.

5.2 Quantitation Assays of Chemical Complementation

Growth Assay

The initial focus on developing a quantitation assay of chemical complementation was to establish an efficient growth assay. In this assay, the PJ69 yeast cells transformed with the nuclear receptor of interest are grown in selective medium (e.g. without histidine or adenine) in the presence of a ligand(s) analogous to the chemical complementation on

selective plates. If the ligand of interest is an agonist, these cells will be able to survive and grow, as shown in Figure 2.1 of Chapter 2. If the ligand(s) is a small molecule that binds but does not activate (an antagonist), a small molecule that cannot bind the receptor, or a toxic small molecule, then the yeast cells will not be able to grow. Therefore, the assay is growth in liquid medium, where only in the presence of the small molecule ligand will the cells be able to grow.

The assay provides an initial analysis of the compound(s) of interest through growth. The small molecule(s) that is able to cause growth of the yeast cells can then be subjected to a dose-response growth assay. In the dose-response growth assay, various ligand concentrations can be used and as the ligand concentration increases, typically more growth should occur. Quantitative measurements of growth assays can be made by measuring cell density at an optical density (OD) of 630 nm, where the more growth that occurs the higher the OD_{630} measurements.

LacZ Assay

One of the advantages of using the PJ69-4A strain for chemical complementation is that the strain contains a Gal4 response element (Gal4 RE) controlling expression of a *lacZ* gene [80]. The *lacZ* gene, one of the first large genes to be completely sequenced (1024 amino acids), encodes the enzyme β-galactosidase. This enzyme cleaves β-D-galactoside between the glycosyl oxygen and the anomeric carbon as shown in Figure 5.1. Its analytical substrates are generally colorimetric, such that upon the conversion of the substrate to the product, a color change is produced and can be measured [24]. The *lacZ* gene is used in a range of applications and has several advantages for use in

diagnostic assays. One of these advantages is that the *lacZ* protein is tolerated by many microorganisms at extremely high levels of expression, and this protein is relatively easy to purify for in vitro applications. The *lacZ* protein is also extremely malleable; previous studies have shown that the *lacZ* protein consists of domains, which due to their modularity can be separated and fused to other proteins or peptides [24]. Once these domains are brought into close proximity, a full and functional enzyme is restored. Furthermore, another advantage of the *lacZ* gene is its use as a transcription reporter gene due to the number of substrates that are available for β-galactosidase. The substrates range from naturally occurring substrates to chemically synthesized ones (Figure 5.1).

The yeast two-hybrid assay often utilizes the *lacZ* gene for reporting on protein-protein interactions, and a number of protocols and derivatives of these protocols have been developed to optimize the β-galactosidase activity in yeast [11,24]. These β-galactosidase assays have been developed on solid media (agar plates) assays and liquid media (tubes or 96-well microtiter plates) assays. Protocols have been developed for an efficient β-galactosidase assay in yeast and multiple cell lysis protocols for yeast have been developed, such that the β-galactosidase activity measured is efficient. Substrates used include colorimetric, chemiluminscence, or fluorescence detection techniques. One of the most commonly used substrates in liquid assays and for measuring β-galactosidase activity is ONPG, *ortho*-nitrophenyl-β-D-galactoside, which when cleaved by β-galactosidase is converted to the yellow compound ONP, *ortho*-nitrophenol. For agar plates, the use of X-gal, 5-bromo-4-chloro-3-indolyl-β-D-galactoside, is used quite frequently, where the conversion of X-gal by β-gal to 5-bromo-4-chloro-indigo, a blue

colored pigment, provides a visual qualitative analysis, where the colonies turn blue if the enzyme is present [24,111].

Since PJ69-4A contains a Gal4 RE controlling expression of the lacZ gene, the combination of a growth assay along with a β-galactosidase assay would provide an efficient quantitative analysis of the amount of transcriptional activation (Figure 5.2). The quantitation protocol would have to be optimized for the initial growth assay and for the reporter assay using β-galactosidase. The combination of these two assays provides the development of a high-throughput chemical complementation assay for various applications, such as drug discovery and protein engineering.

To develop the high-throughput chemical complementation assay using an initial growth assay and a β-galactosidase assay, a protocol would have to be developed with near optimal conditions for both the growth assay and the β-galactosidase assays. Various parameters in both assays were finely tuned to provide efficient conditions for each of these assays.

Media for Growth Assays

There are two types of media that can be employed in the growth quantitation assays of chemical complementation. The first is a selection growth assay, where the cells are grown under selective pressure, utilizing the Gal4 REs controlling expression of the *HIS3* or *ADE2* genes. The selective media will contain a concentration(s) of the ligand of interest, and only cells containing a receptor able to bind and activate in response to the ligand will be able to survive. Selection assays are generally a more efficient assay because only cells containing receptors with the ability to bind and

activate in response to the ligand will be able to survive. These characteristics of selective growth assays make selective media more efficient because only receptors with the desired function are able to survive. However, one disadvantage to using a selection based assay, is that the β-gal activity that is measured is not a true testament to the amount of transcriptional activation that is induced by the receptor and its corresponding ligand. This is due to the fact that the number of cells in each well is not constant; at lower concentrations of ligand, a very small amount of cells are present. Therefore, in cases of low expression, the β-galactosidase activity is undetectable.

The 2nd type of media used in growth assays is when yeast cells are grown in media selecting for the plasmids only, a non-selective growth assay. For example, if two yeast expression plasmids are used that contain a *URA3* (involved in uracil biosynthesis) and *LEU2* (leucine biosynthesis) marker, then the yeast cells would be grown in media lacking uracil and leucine. In the non-selective quantitation growth assay, cells would be grown in medium selecting for the plasmids containing the receptor and the coactivator, and in the presence of ligand. These cells will then be lysed to assay for β-gal activity. A disadvantage to using the non-selective growth assay is that every variant will be growing independent of the receptor's ability to bind ligand. Therefore, every variant will have to be assayed for its ability to bind the ligand through its activation of the *lacZ* gene. However, the one advantage of utilizing this system for chemical complementation is the constant number of cells in each well, and in cases of low expression, β-galactosidase activity is detected.

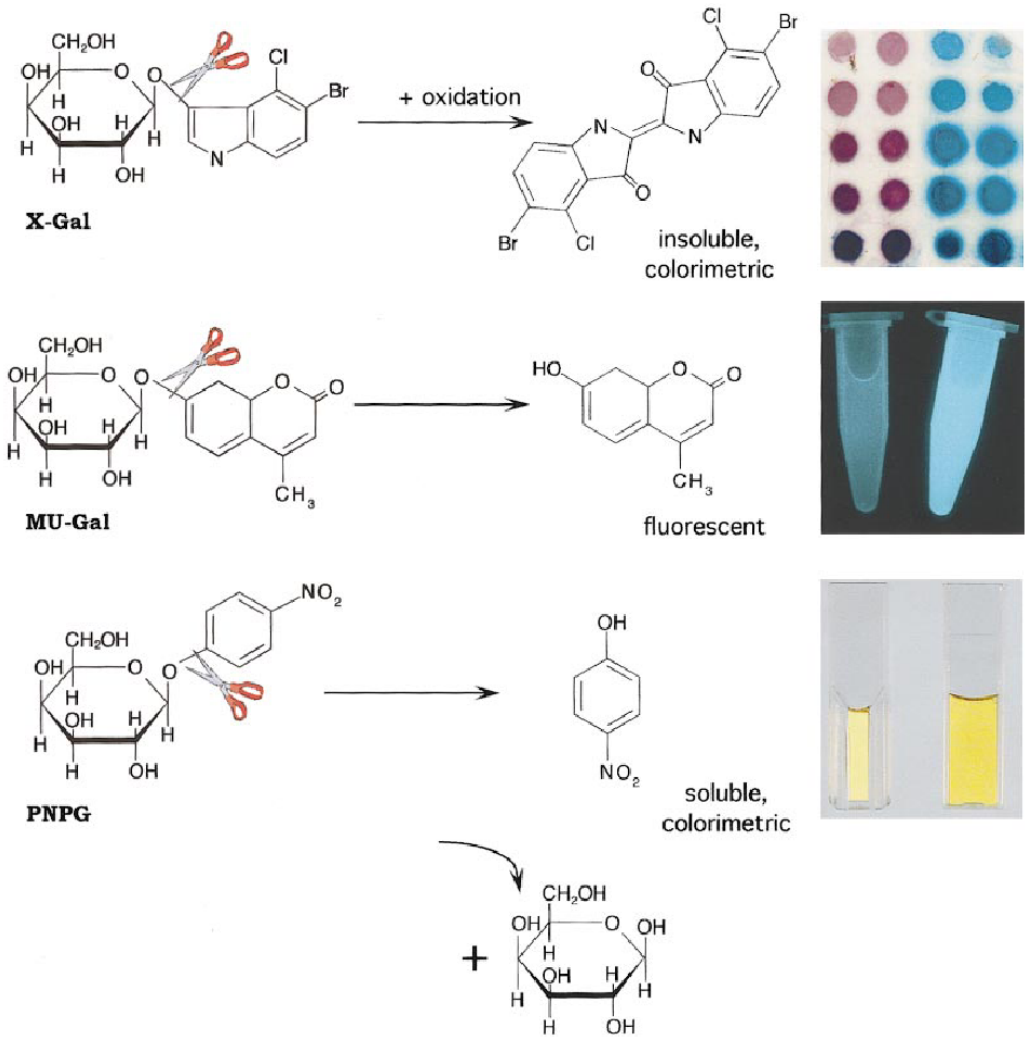


Figure 5.1. Substrates of β-galactosidase.
Adapted from Serebriiskii et al_AnalBioch_2000

and the β-gal activity measured. In the quantitation assay developed for chemical complementation both growth assays were used and compared for β-gal activity (see below).

Using 3-AT for added selective pressure

PJ69-4A is a useful yeast strain for chemical complementation because this strain contains both a *HIS3* gene and an *ADE2* gene under the control of a Gal4 response element (RE). Each of these selective markers has its disadvantages and advantages. The *ADE2* gene has a relatively low basal expression associated with it, and provides a more stringent selection system. With the *ADE2* gene, the receptor either turns the gene on or off, displaying either growth or no growth. The *HIS3* gene in PJ69 contains basal growth associated with this gene, thus the need to reduce the basal expression. 3-amino 1,2 triazole (3-AT) is an inhibitor of the *HIS3* enzyme, imidazoleglycerol-phosphate dehydratase, a key enzyme in the histidine biosynthetic pathway. 3-AT can be used to reduce the background growth associated with the *HIS3* gene. The activity of the *HIS3* enzyme can be tuned, such that if a ligand-receptor pair with higher efficacy is desired, a higher concentration of 3-AT can be added to the media to increase the threshold level of activation (Figure 5.3). In cases where receptor libraries are created for their ability to bind and activate in response to a ligand with variable activation profiles, these receptors will be assayed in media containing different concentrations of 3-AT. Receptor variants with variable efficacies can be selected by adding higher or lower concentrations of 3-AT. For example, in the hypothetical example shown in Figure 5.3, if the desired ligand-receptor pair is a variant that is able to grow to a

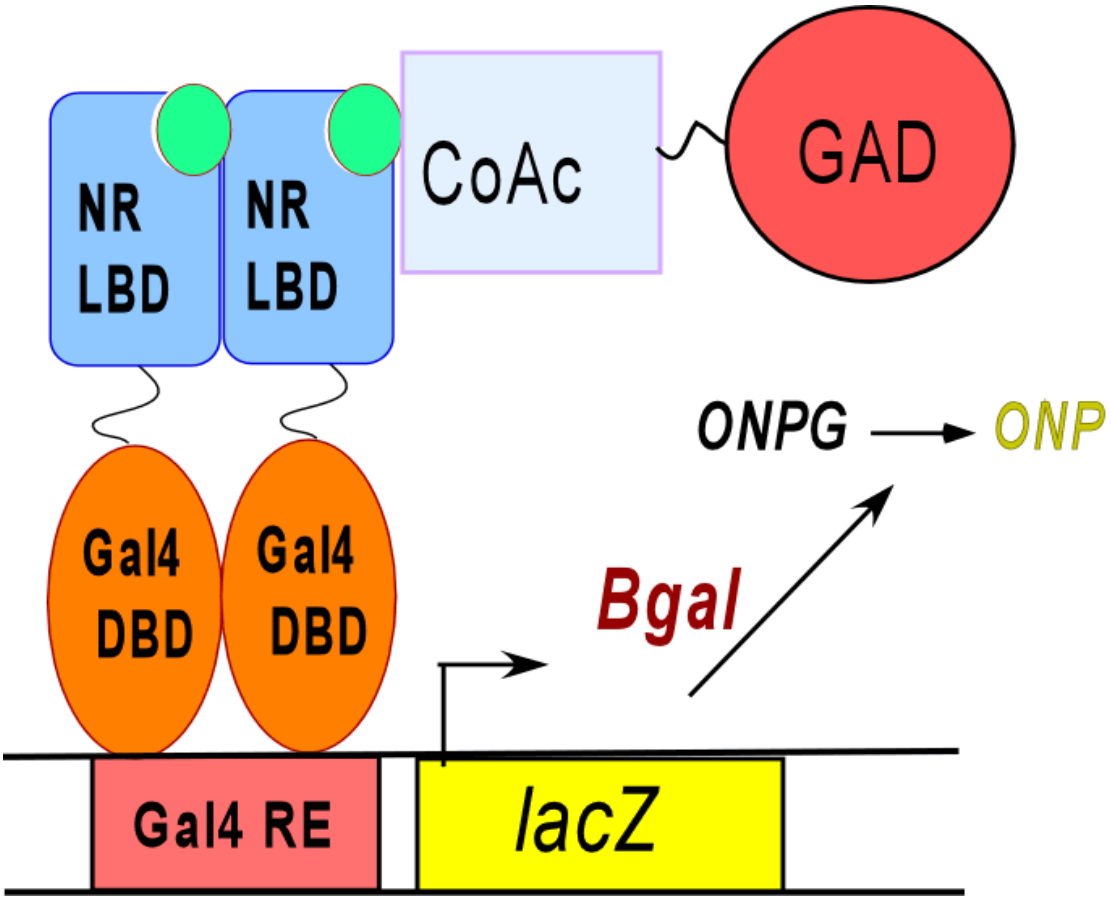


Figure 5.2. Chemical complementation with the *lacZ* gene.

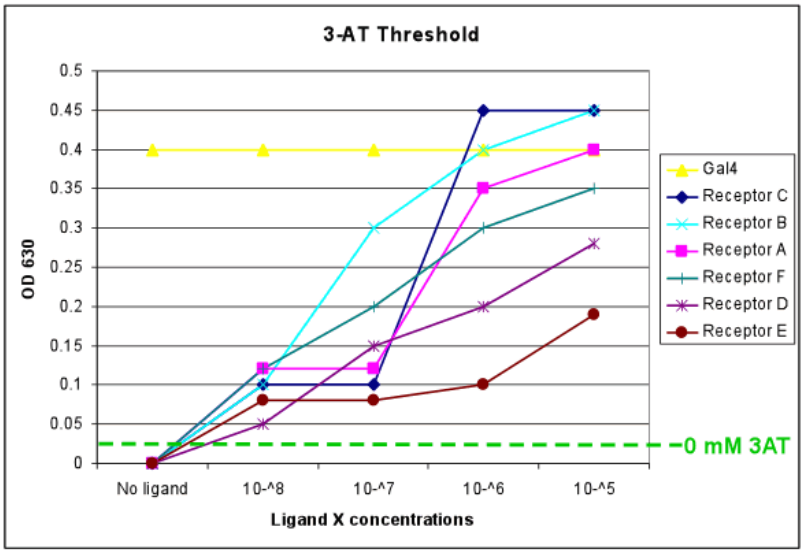
cell density (OD_{630} of 0.3) at a ligand concentration of 10^{-6} M, then 50 mM 3-AT can be used in the medium to eliminate all other variants that are unable to activate at this level (Figure 5.3). The ability to tune the efficacy of the ligand-receptor pairs using the *HIS3* gene and 3-AT expands the versatility of chemical complementation for applications in protein and enzyme engineering, as well for drug discovery and for creating biosensors (discussed in Chapters 8 and 9).

5.3 Developing the quantitation assay using wild-type RXR

The pGBDRXR and pGAD10BACTR plasmids described in the previous chapters were used to develop the quantitation protocol for chemical complementation. The protocol listed below was developed in 96-well microtiter plates, consisting of a growth assay and a β-galactosidase assay. A quick summary of the protocol is that during the growth assay, cells are grown overnight at 30 °C in either selective or non-selective media. After the initial growth assay, the cells are lysed using yeast lytic enzyme and the detergent Triton X 100. The cell lysates are then incubated with ONPG and β-galactosidase activity is measured by absorbance spectroscopy at 405 nm. For the growth assay, the following conditions were determined to be the suitable conditions for chemical complementation.

Initial cell concentration.

The number of cells initially added to the 96-well microtiter plate's wells is important for maintaining a viable yeast cell population. If the concentration of the yeast cells are too high in the well initially, the growth of the yeast can be hindered by multiple



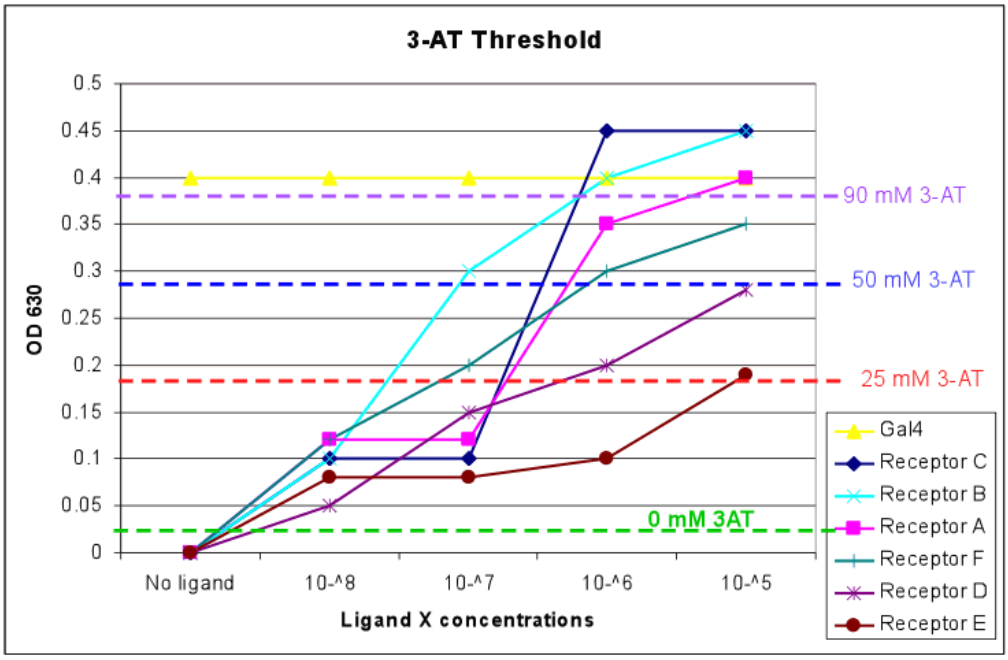
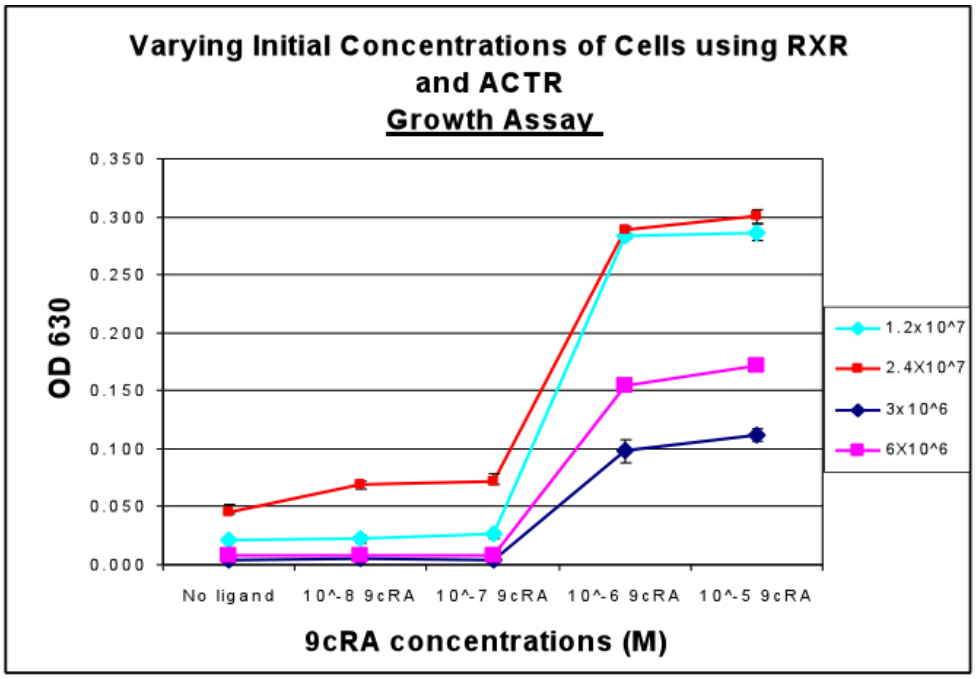


Figure 5.3. A hypothetical example of using 3-AT to increase the threshold in the quantitation assay.

side effects, such as the elimination of resources in the media. Another effect of too high a concentration of yeast in the wells, is that these cells will obtain their nutrients from cells that have already died, and yeast survival will not be ligand-dependent. Therefore, the right population of cells in a small volume space is crucial for a sensitive growth assay.

To determine the appropriate concentration of cells to add to each of the wells, overnight yeast cultures containing pGBDRXR and pGAD10BAACTR and pGBT9Gal4 were grown in 2 mL media in three tubes selecting for the plasmids (-Leu -Trp and -Trp media respectively). These cells were grown for about 24 hours at 30 °C with shaking (300 rpm) and were centrifuged (4000 rpm, 7 minutes) to remove the media. The cells were resuspended in 1 mL of water and the concentration of the cells was determined at an OD_{600}. Various concentrations of cells were added to each of the wells and the final concentrations were targeted to 3×10^6, 6×10^6, 1.2×10^7, and 2.4×10^7 cells mL^{-1}. Working stocks were made such that adding 20 μL from the working stocks to the wells gives the desired final concentrations mentioned above. Once these cells were added to the 96-well microtiter plates, the cells were incubated at 30 °C with shaking at 150 rotations per minute (rpm). The cells were grown for about 20 hours lysed, and β-galactosidase activity measured.

The growth assay showed a correlation with cell concentration; more growth was observed as the initial concentration of the cells increased (Figure 5.4). The growth observed with 1.2×10^7 and 2.4×10^7 cells mL^{-1} was greater (≅ OD_{630} 0.3) than in the wells with 3×10^6 and 6×10^6 cells mL^{-1} (≅ OD_{630} 0.15). However, once these cells were lysed, the lower concentration of cells showed greater β-galactosidase activity in



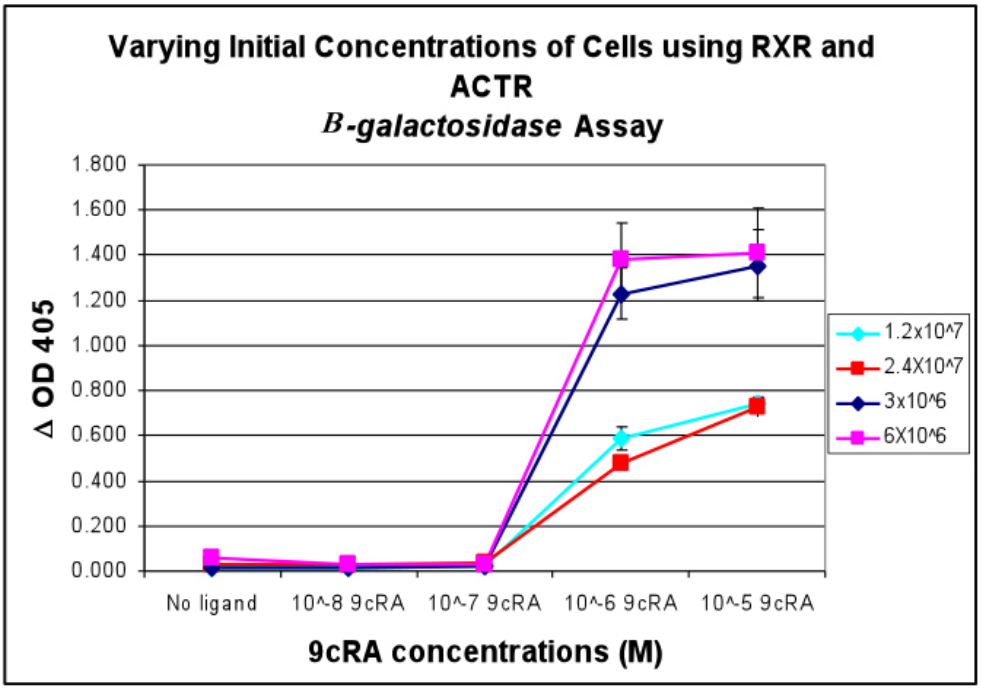


Figure 5.4. Varying initial concentration of cells in the quantitation growth assay.

comparison to the higher concentration of cells (OD_{405} of ≅ 1.4 compared to ≅ 0.8 respectively). When this assay was repeated, the lower concentrations of cells showed comparable β-galactosidase activities in comparison to the wells with higher concentrations of cells. Therefore, to eliminate issues resulting from an overcrowded population of yeast, we decided that the efficient initial concentration of cells would be $3x10^6$cells mL^{-1}. In terms of growth and β-galactosidase activity, both were observed with RXR and ACTR only at 10^{-5} M and 10^{-6} M concentrations of 9cRA. On agar plates, growth was observed at 10-fold lower concentrations of ligand.

Growth Time

To increase the sensitivity of the quantitation protocol such that ligand-activated growth is observed at lower concentrations of ligand, an attempt was made to increase the growth time. The hypothesis is that if the growth time is extended to 36 or 48 hours, the wells with the lower ligand concentrations will be able to overcome the stringent selective pressure and will start to grow.

Cells containing the pGBDRXR and pGAD10BAACTR plasmids and the pGBT9Gal4 plasmid were grown overnight in their selective media (-Leu-Trp and –Trp respectively). The cells were harvested in the morning, resuspended in water, and a final concentration of $3x10^6$ cells mL^{-1} was added to the wells. In addition to the cells, adenine selective media containing the appropriate ligand concentration was added, and the 96-well microtiter plates were incubated at 30 °C with shaking. The cells were lysed after 24, 36, and 48 hours of growth.

The longer incubation times increased the amount of growth (Figure 5.5). After 24 hours of incubation, the growth at the highest concentration of ligand is about OD_{630} of 0.08 whereas after 36 and 48 hours, the OD_{630} of 0.3. The β-galactosidase activity for the three time points also follows the same trend. The longer the incubation time, the more β-galactosidase activity, varying from an OD_{630} of 0.4 for 24 hours to an OD_{630} of 1.2 for 48 hours. However, the activation occurs was only at the two highest concentrations of ligand, therefore increasing the growth assay time did not affect growth or activation at lower concentrations.

Growth media for growth assay

In a final attempt to increase the sensitivity of the quantitation assay with lower concentrations of ligand, two types of growth media were used in the quantitation growth assay. Specifically, we wanted to determine whether a non-selective growth assay versus a selective growth assay increases the sensitivity of the quantitation assay. Maintaining selective pressure in the media would permit only the "producers" to survive, whereas eliminating the selective pressure would allow even the non-producers to remain in the medium. Without selective pressure in non-selective media, the β-galactosidase expression will drift, causing individual cells to vary in the population, whereas maintenance of the selective pressure, removes the non-producers, and β-galactosidase expression may be more uniform. In the non-selective assay, β-galactosidase expression will cause higher error bars due to the variance in the population. On the other hand, the stringency of a selective assay, despite the genetic uniformity and because the non-producers are removed, causes lower β-galactosidase activity because of the stringency in

growth. However, in this case, the error bars are lower because of the uniform population.

A non-selective growth assay versus a selective growth assay was set-up to determine whether an increase in sensitivity is seen with one assay versus another. Cells that contain the pGBDRXR and pGAD10BAACTR plasmid were placed in 96-well microtiter plates at 3 x 10^6 cells mL^{-1}. The non-selective assay contained media lacking leucine and tryptophan (-Leu -Trp), selecting only for the plasmids. The selective growth assay contains adenine selective media (-Ade –Leu -Trp), and each media contained various concentrations of 9cRA. Both 96-well microtiter plates were grown for about 24 hours (Figure 5.6). The results of the non-selective growth assay with RXR and ACTR showed the same amount of growth for every sample as expected, and a ligand-activated β-galactosidase activity; a measurable increase in β-galactosidase activity was observed at 10^{-7} M concentration of 9cRA. The β-galactosidase activity at the highest concentration of ligand is observed to be at about 0.7 at an OD_{405}. In the selection assay a more ligand-activated growth pattern is observed, where as the ligand concentration increases, the growth increases as well. Growth and β-galactosidase activity are barely observed at 10^{-7} M. The amount of β-galactosidase activity at the highest concentration of ligand is equal in both growth assays (Figure 5.6).

Other parameters of the β-galactosidase assay of the quantitation protocol were also experimented with to determine the most efficient assay for measuring β-galactosidase activity using chemical complementation. The parameters tested included the amount of yeast lytic enzyme and the duration of incubation with yeast lytic enzyme,

to ensure the most efficient cell lysis and least amount of loss of β-galactosidase activity. The amount of Triton X 100 and the amount of ONPG were also varied.

Most of the lysis protocol was based on previous β-galactosidase assays for yeast [24,112-114]. The most efficient concentration of yeast lytic enzyme was determined to be 10 mg mL^{-1} with an incubation time at 37 °C for 30 minutes. A 1:1000 fold dilution of Triton X 100 was determined to enhance the cell lysis process initiated by the yeast lytic enzyme. The incubation time was varied initially from 15-45 minutes, and the 15-minute incubation at room temperature was determined to be sufficient for completing the yeast lysis process. The amount of ONPG added is at a concentration of 2 mg mL^{-1}; this concentration is sufficient to display good β-galactosidase activity.

5.4 Quantitation Assay with RXR variants

Comparing EC_{50} values with mammalian cell assays

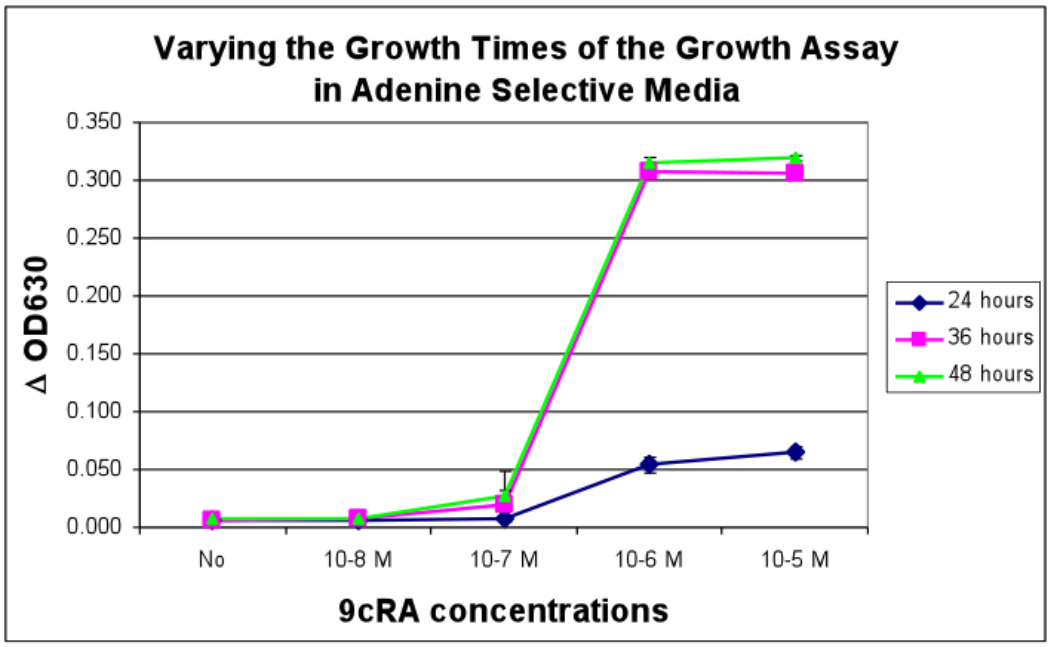
The quantitation protocol was developed and optimized using RXR wild type and ACTR, as well as the Gal4 protein as a control. One of the advantages to developing a quantitative assay is to be able to report on the levels of transcriptional activation. Furthermore, a useful tool in evaluating the characteristic of a ligand-receptor pair is EC_{50} values, and using a quantitative assay will allow the determination of this value. These values can be compared to mammalian cell assays, and can determine how well chemical complementation correlates with mammalian cell assays.

The RXR variants were cloned into yeast expression plasmids (pGBD) and transformed into yeast with the pGAD10BAACTR vector. Transformants were selected from selective media lacking leucine and tryptophan, and inoculated in the same media

for overnight growth. The quantitation protocol was set-up for each of the variants, along with wild-type RXR and pGAD10BAACTR and pGBT9Gal4 as controls. These cells at a concentration of $3x10^6$ cells mL^{-1}, were placed in 96-well microtiter plates with adenine selective media containing varying concentrations of 9cRA and the synthetic ligand LG335. The concentrations ranged from 10^{-5} M to 10^{-8} M. The plates were incubated at 30°C with shaking for about 24 hours and then lysed using the protocol below. β-galactosidase activity was determined for each of these variants and the EC_{50} values obtained for each of these variants was compared to results in mammalian cell culture assays.

For mammalian cell culture assays, mammalian vectors containing each of the variants had previously been made [29]. HEK293 (human embryonic kidney) cells were transfected with a vector (20 ng) expressing each variant, along with a plasmid containing the β-galactosidase gene (40 ng), and the luciferase reporter (40 ng). The cells were also placed in media containing varying concentrations of both 9cRA and LG335 (see materials and methods below). The cells were harvested and luciferase activity and β-galactosidase activity were both measured. Transcriptional activation was reported in relative light units (RLUs).

Most of the EC_{50} values obtained using mammalian cell culture assays are comparable to the values obtained from chemical complementation, using both the integrated ACTR strain and plasmid borne ACTR (Table 5.1) [29]. With some of the RXR variants, such as QCIM and F439L, there are distinct differences in the EC_{50} values among the two systems; mammalian cell culture assays report EC_{50}s in the nanomolar



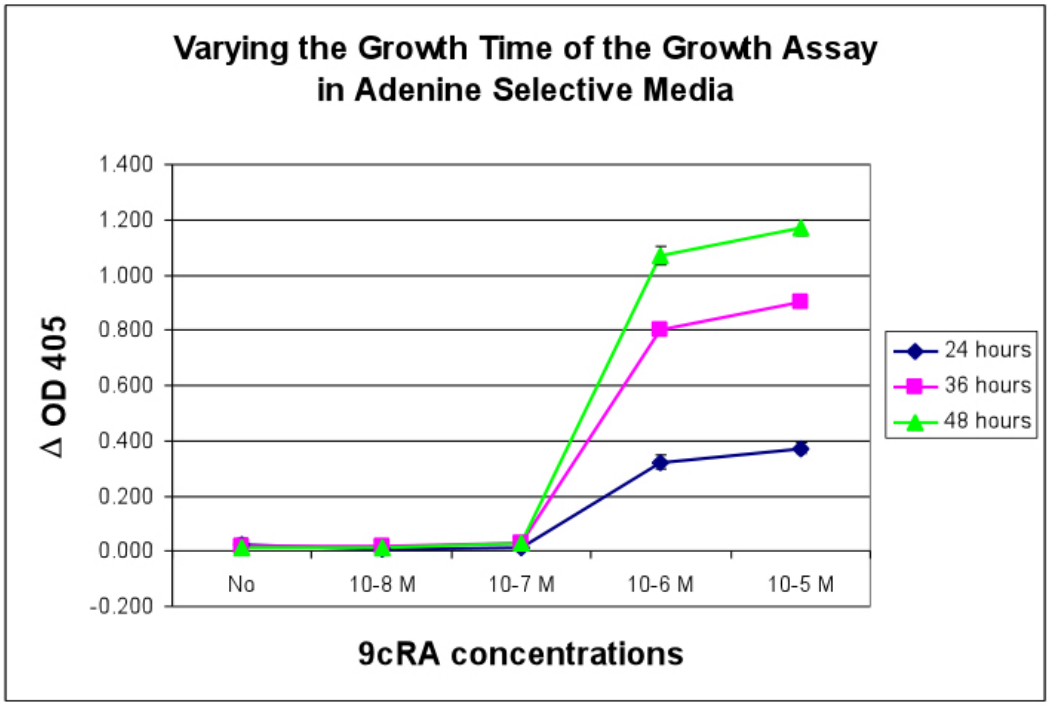


Figure 5.5. Varying incubation times in the quantitation growth assay.

range with LG335, whereas no activity is seen with that variant and LG335 in chemical complementation. In most cases, the values are comparable.

Comparing BAH14RPJ69 strain with Plasmid ACTR.

The RXR variants that were transformed with pGAD10BAACTR into the PJ69-4A strain were transformed alone into the BAH14RPJ69 strain, the strain with the integrated ACTR gene (See Chapter 4 for more detail). The transformants from both strains were inoculated in the selective media overnight and plated onto 96-well microtiter plates with media also containing varying concentrations of 9cRA and LG335. The concentrations ranged from 10^{-5} M to 10^{-8} M. The plates were incubated at 30 °C with shaking for about 24 hours and then lysed using the protocol shown below. β-galactosidase activity was determined for each of these variants and the EC_{50} values obtained for each of these variants in both of the strains. The idea was to determine whether the plasmid borne ACTR and the integrated ACTR gene function equivalently and whether the expression level of the ACTR gene varies from expression from the plasmid to being integrated.

As shown in Figure 5.7, the BAH14RPJ69 strain seems to have relatively the same correlation as when the ACTR is introduced into PJ69-4A in the pGAD10BAACTR. In Figure 5.7, most of the RXR variants have the same transcriptional profiles using the ACTR integrated strain and on the plasmid. Yet, a significant difference is the expression levels of ACTR, which seem to be much lower in the integrated strain. This is clearly seen in the graphs shown in Figure 5.7; the

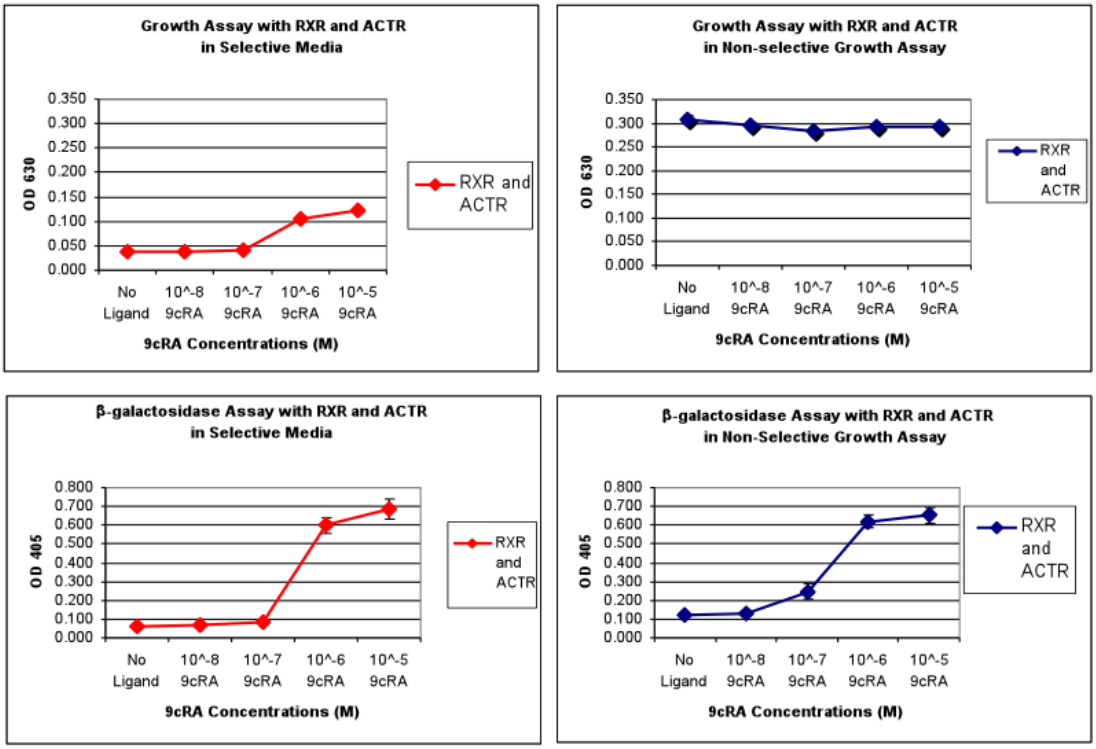


Figure 5.6. Varying the media type in the quantitation growth assay.

transcriptional activation of the β-galactosidase genes in the integrations strains with both 9cRA and LG335 were around an OD_{405} of 0.12. However, in the graphs that contain the ACTR gene as part of a plasmid, the β-galactosidase activity that is measured is around an OD of 1 for 9cRA and wild-type RXR and about an OD_{405} 0.2 for the other variants. This is not surprising since plasmids are known to contain sequences for higher copy numbers, and integration of the gene into the strain reduces the expression level of the gene dramatically. Despite the differences in expression, the EC_{50} values for RXR variants measured in both strains are relatively the same, further support that the expression of ACTR is the cause of reduced transcriptional activity.

5.5 Discussion of Quantitation

Developing a quantitative analysis for chemical complementation increased the versatility of the system greatly. Previous analysis of ligand-activated growth was based on a qualitative human "eye" analysis. When comparing various receptors and their ligands with the qualitative approach, only the larger variations between the different receptors were detected because subjective evaluations were made on the density of growth. Slighter variations in growth are not detectable through the qualitative approach. By developing this quantitative based approach for ligand-receptor pairs, large and small variations can be observed using both the growth and β-galactosidase assays.

The growth assay provides the initial test to determining whether the ligand-receptor pair are interacting and functioning properly. This assay would be extremely useful in applications such as drug discovery, where a combinatorial library of synthetic compounds is placed onto a 96-well microtiter plate, such that each well contains a

different ligand. The wells will all have selective media, such as adenine or histidine selective media. Yeast cells containing the receptor of interest and its coactivator would then be added to each well. These cells would be grown for about 24 hours and after the incubation period the plates will be analyzed for growth. The wells on the plate that grew show initial results that the ligand can bind and activate the receptor. These ligand-receptor pair can then be further analyzed for activation profiles through the β-galactosidase assay.

Measuring the β-galactosidase activity in chemical complementation is analogous to using the luciferase reporter in cell culture. Both of these reporter genes provide information on how well the ligand can activate the receptor, indirectly giving information on the ability of the receptor to bind the ligand at lower concentrations. Importantly, these reporters allow EC_{50} values to be observed and measured, which serves as a direct comparison between chemical complementation and mammalian cell assays. The EC_{50} values are important to obtain when discovering a new ligand for a receptor, or for creating novel ligand-receptor pairs with chemical complementation, because these values determine whether the ligand-receptor pair is a good pair for applications in drug discovery or medical applications, such as gene therapy. Chemical complementation serves as a quick, relatively simple, and cheaper system for evaluating libraries of proteins and compounds in comparison to mammalian cell assays. Therefore, these systems should be comparable, such that a good "hit" in chemical complementation also is a good 'hit" in mammalian cell culture assays. With the β-galactosidase assay, the correlation between chemical complementation and mammalian cell assays can be

Table 5.1. Comparing EC_{50} values from the PJ69-4A transformed with variants and pGAD10BAACTR, BAH14RPJ69 strain with integrated with ACTR, and HEK293 mammalian cell assays.

	PJ69-4A transformed with ACTR plasmid		**BAH14R PJ69 ACTR integrated in the strain**		**HEK 293 cells Mammalian Cell Culture**	
	9cRA	***LG335***	***9cRA***	***LG335***	***9cRA***	***LG335***
Wild-Type RXR	330 nM	-----	330 nM	-----	630 nM	310 nM
I268F	3 μM	-----	3 μM	-----	2.5 μM	1 μM
L326F	-----	-----	-----	-----	-----	-----
C432G	1 μM	-----	1 μM	-----	1.5 μM	-----
Q275C	3 μM	-----	3 μM	-----	1 μM	-----
L309V	-----	-----	-----	-----	-----	-----
L436V	-----	-----	-----	-----	-----	-----
F439L	1 μM	-----	300 nM	-----	150 nM	120 nM
F313I	1μM	1 μM	1 μM	1 μM	1.5μM	63 nM
V342F	3.5 μM	-----	3.5 μM	-----	800 nM	63 nM
I310M	3.5 μM	-----	3.5 μM	-----	1.5 μM	3 μM
F313I;F439L	3 μM	1 μM	3 μM	1 μM	1 μM	6.3 nM
I310M;F313I	6 μM	3.3 μM	5 μM	3.3 μM	3.9 μM	------
Q275C;F313I	-----	1 μM	-----	3 μM	4 μM	33 nM
Q275C;I310M	-----	-----	-----	-----	3.9 μM	150 nM
F313I;L436V	3.3 μM	-----	3.3 μM	-----	3.3 μM	33 nM
Q275C;I310M;F313I	-----	1 μM	-----	1 μM	------	33 nM
Q275C;F313I;V342F	-----	100 nM	-----	100 nM	790 nM	63 nM
Q275C;L436V;F439L	-----	-----	-----	-----	3.3 μM	33 nM
I310M;F313I;F439L	-----	1 μM	-----	1 μM	3.3 μM	330 nM

- Each of the measurements in both yeast strains and in mammalian cell culture assays were repeated twice.
- Based on ΔOD_{405} measurements in the β-galactosidase assay in selective media for chemical complementation.
- (-----):no transcriptional activation measured

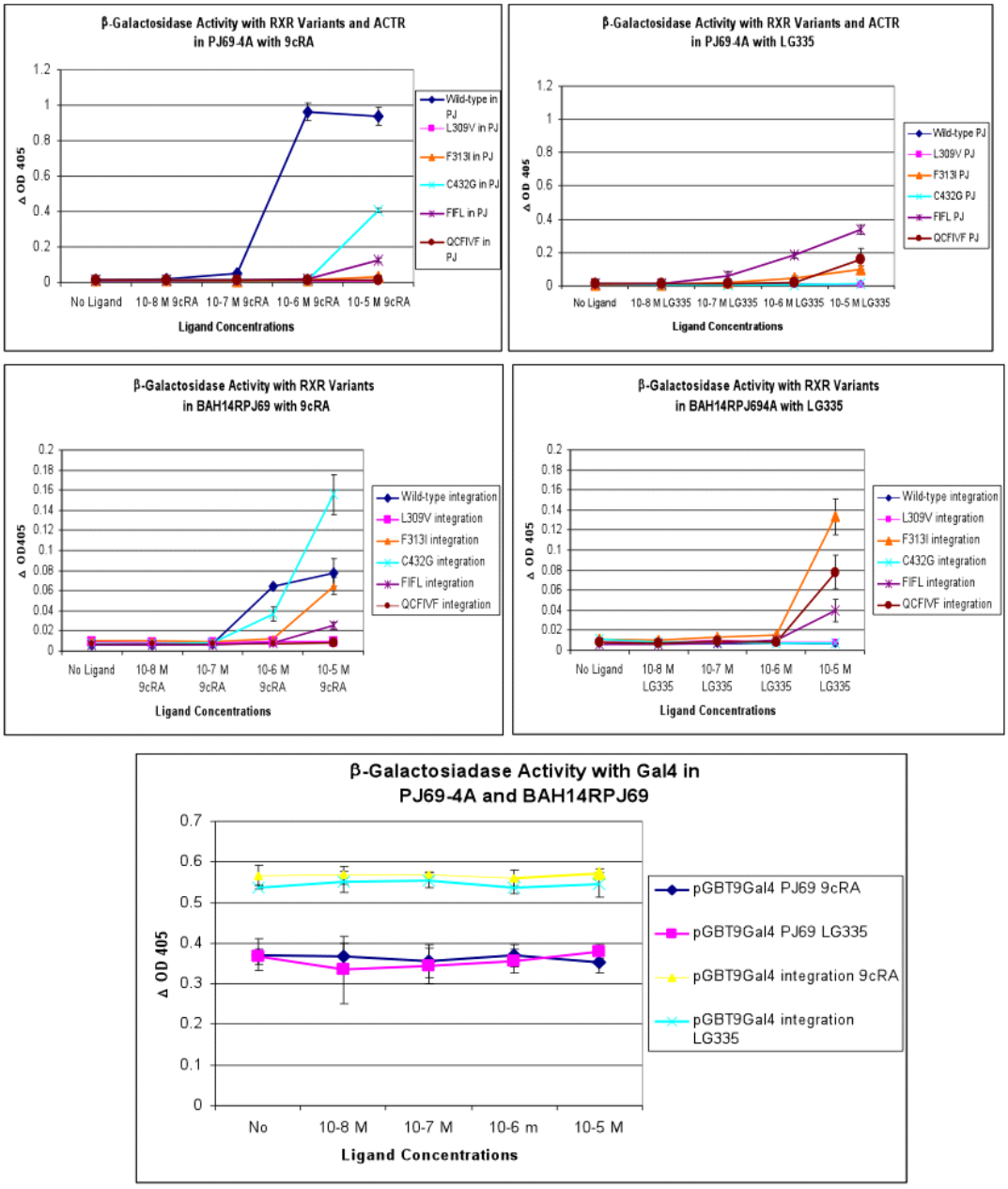


Figure 5.7. Comparing the β-galactosidase activity of RXR variants and ACTR in PJ69-4A and in the BAH14RPJ69 strain in adenine selective media.

made on a quantitative basis, providing a more efficient and complete way for determining ligand-receptor interactions. We have been able to show that an efficient quantitation assay for chemical complementation has been developed. This assay consists of a growth assay varying from 24 hours to 48 hours where ligand-activated growth can be detected in the presence of a selective media with ligand. A screen growth assay can also be used, where cells are grown in media selecting for the plasmid's selective markers. Screen media may enhance the transcriptional activity at lower concentrations of ligand as shown in Figure 5.6, but in a screen assay the presence of the non-producers can often result in a mixed population of cells, leading to higher noise. Selection provides the stringency that allows only the producers to survive, leading to reduced transcriptional profiles at lower concentrations of ligand, but also less noise is associated with selection assays. The growth assay of the quantitation protocol can be tuned to meet the desired conditions whether these conditions are a stronger signal or higher sensitivity.

The β-galactosidase assay is the second part of the quantitation assay. Once these cells are grown to the desired density, the cells can be lysed using a combination of yeast lytic enzyme and the detergent, Triton X 100. The cell lysates can then be evaluated for β-galactosidase activity, because of the presence of a Gal4 response element controlling expression of the *lacZ* gene. In most instances, as shown in Figures 4-8, there is ligand-dependence activation; more ligand increases the amount of β-galactosidase activity. This assay provides the advantage of quantitatively measuring transcriptional activation, which can be, used as a comparison to other transcriptional activation *in vivo* assays, such

as mammalian cell assays. The EC_{50} values obtained in comparison to mammalian cell assays show that these results are comparable, but some differences do exist.

To determine a high-throughput assay quality, a Z-factor can be calculated for the assay [115]. A Z-factor is dimensionless, simple, statistical characteristic for each high-throughput assay, and measures the quality of a particular high-throughput assay. This factor can also be used to compare the assay developed with other high-throughput assays present and can be used in optimization and validation of the assay [115]. The Z-factor of the assay can be calculated using the following equation:

$$\textbf{Z factor} = \frac{\textbf{3SD of sample} - \textbf{3SD of control}}{\textbf{Mean of sample} - \textbf{Mean of control}} \qquad \textbf{(1)}$$

where the standard deviation and mean of the sample are both used to determine the Z factor value. When this value was calculated for the chemical complementation quantitation assay, this value was determined to be 0.5. According to Table 5.2, a value of 0.5 indicates that this is an excellent assay [115].

Developing a quantitation protocol for chemical complementation increases the versatility of this system, providing a high throughput method for evaluating protein-ligand interactions in yeast. This can be useful towards applications in drug discovery, such that agonists and antagonists of nuclear receptors can be discovered in combinatorial libraries of synthetic compounds. For applications in protein engineering, protein libraries can be evaluated in a high-throughput fashion using the quantitation protocol, evaluating the transcriptional activation of the proteins with a ligand(s). Furthermore, by using 3-AT, protein libraries can be placed under a genetic selection, such that only certain receptor variants that have the desired transcriptional profile survive (Figure 5.3).

All of the tools of the quantitation assay mentioned above expand the applications of chemical complementation and the efficiency of chemical complementation as a generalizable system for evaluating ligand-protein interactions.

5.6 Materials and Methods

Expression Plasmid

Yeast expression plasmids, pGBDRXR, pGAD10BAACTR, and RXR variants in yeast expression vectors are described previously (Chapters 3 and 4). The mammalian expression plasmid, pCMXRXR, and mammalian expression vectors were previously made by Doyle and co-workers .[29]

Mammalian Cell Culture Assays

HEK293 (Human embryonic kidney cells) were transfected with the RXR variant (20 ng), along with the 20 ng of the pCMXβgal and pLucCRBPIIMCS plasmids, coding for β-galactosidase and luciferase respectively. The transactivation of these variants in response to 9cRA and LG335, were measured as described by Peet et al. [110].

5.7 Quantitation Protocol

Day 1:

1. Resuspend cells in 1 mL sterile water, pellet again at the same conditions as Step 1, and decant water.
2. Resuspend pellet in 1 mL sterile water.
3. Take 20 μL of the sample for reading at OD_{630} using the spectrophotometer.

Table 5.2. Characterizing Assays Based on Z-factor Values
Reproduced from Zhang et al. Journal of Biomoelcular Screening, 1999

Z-Factor value	Related to Screening
1	An ideal assay
$1 > Z \geq 0.5$	An excellent assay
$0.5 > Z > 0$	A double assay
0	A "yes/no" type assay
< 0	Screening is essentially impossible

4. Determine the concentration of cells using the conversion factor 3.0 x 10^6 cells/ mL/ 0.1 OD_{630}
5. The final concentration of yeast cells in the wells should be 3.0 x 10^6cells mL^{-1}. Since the volume to add to each well is 20 μL, you have to make a 1.5 x 10^7 cells mL^{-1} stock solution of yeast. From that stock, add 20μL to the appropriate wells.
6. Prepare the ligand/media mixes. Determine the amount of volume needed. Ligand/media stocks are made via serial dilutions. For example, if the ligand stock is a 10^{-2} M, and you are looking to do a dose response from 10^{-5} M to 10^{-8} M, this is procedure.
    i. First determine the total volume that you will need for the plates. In this case let's assume that it is 4 mL. You will need to remember that 4 mL is the total volume for these ligand dilutions.
    ii. To determine how much of the 10^{-2} M ligand stock you will need to add to the 10^{-5} M solution, use the following equation: *M1V1=M2V2*; in this case it would be set up as

$$(1 \times 10^{-2}\ M)\ (x) = (1 \times 10^{-5})(4mL)$$

$$x = 4\mu L$$

Therefore you add 4μL of the 10^{-2} M stock to 4 mL of the media that you are going to use. For the next set of dilutions, a serial dilution can be performed. Use the same equation as above. For example, for the 10^{-6} M ligand solution, you will need to add 400μL of the 10^{-5} M to 3.6 mL media. The same will be done for the other dilutions.

7. Add 80 μL of the media/ligand mixes to the wells.
8. Add 20μL of the cell stocks.
9. Add water to the wells that don't have cells or media.
10. Measure OD_{630} of wells with plate reader. This is OD_{630} at t = 0.
11. Incubate plate(s) overnight (30 °C, 150 RPM) in 96-well microtiter plate holder. Place beakers of water into incubator for additional humidity. Place aluminum foil on window of incubator to reduce light exposure, if necessary.

Day 3:

A. *Preparation of solutions:*

i. Prepare Z buffer: 100 mM Na_2PO_4, 10mM KCl, 1mM $MgCl_2$, 50mM βME, pH 7.0

ii. Prepare 25mM Pipes stock solution.

iii. **Prepare lyticase stock solution**. Prepare a 10mg mL^{-1} lyticase solution in Z buffer.

iv. β-galactosidase solution: Dissolve 0.5 μL β-galactosidase (Sigma G-4155) in 200 μL 25 mM Pipes buffer (pH 7.5). Dilute this stock 10-fold in additional Pipes buffer (make ~100 μL stock).

v. ONPG solution: Prepare a 2mg mL^{-1} ONPG (Sigma N-1127) solution in 25 mM Pipes (pH 7.5). Make sure to make enough ONPG for all your samples. If you are doing one 96-well plate, you generally want to make about 3 mL total volume.

vi. Triton X-100: Dissolve Triton X-100 to 0.1% (v/v) in H_2O. Vortex solution until Triton is completely dissolved.

vii. 1M Na_2CO_3 solution: Dissolve 12.4 g $Na_2CO_3 \cdot H_2O$ into 100 mL H_2O.

Day 3:

1. Measure OD_{630} of 96-well microtiter plate.
2. Pellet cells in 96-well microtiter plate (15 min @ 1500RPM (~430xg's)). Carefully remove media from cells with multi-channel pipettor. (**TIP**: To remove the media, go straight down to the bottom of the well and pipet up the media. Try to avoid the sides of the well, because the cells are more concentrated in that region.)
3. Add 20 μL of diluted lyticase to each well. Shake the plate so that the pellet will become resuspsended. Let plate(s) sit at 37°C for 30 minutes. Include extra column for β-galactosidase positive control.
4. Add 100 μL diluted Triton X-100 solution to each well. Let sit at room temperature for 10 min (not optimized). (No need to agitate.)
5. Centrifuge 96-well microtiter plate (10 min @ 1500 RPM (~500xg's)).
6. Aliquot 100 μL of lysed cell solution into a clean 96-well microtiter plate.
7. Add 20 μL ONPG solution to each well. (**TIP**: Add the ONPG in rows rather than columns.
8. Quickly measure OD_{405} of each well for t = 0 min data point.

9. Add 10 μL diluted β-galactosidase to positive control column of wells.

10. Incubate 96-well microtiter plate until definite yellow formation in non control wells. Typically, 15-30 min, but no longer than 60 min. Ideally, OD_{405} should be between 0.6 and 0.9. (The β-galactosidase can be stopped before the other samples)

11. Add 50 μL 1M Na_2CO_3 solutions to each well to "stop" reaction. Note time. Get rid of any bubbles. Measure OD_{405} of each well.

Chapter 6

Negative Chemical Complementation

6.1 Defining Negative Chemical Complementation

Chemical complementation has thus far been shown to link the presence of a small molecule to the survival of yeast cells, implying positive genetic selection. When a small molecule binds and activates the nuclear receptor, known as an agonist, leads to the recruitment of the coactivator complex, and the initiation of transcription. Positive genetic selection could lead to the discovery of new agonists for nuclear receptors as well as for discovering new agonist-receptor pairs. This tool is also extremely useful in engineering new receptors to bind novel small molecules. Both of these concepts are further described in Chapters 7 and 8 respectively.

Chemical complementation has been developed for discovering agonists and new activators of nuclear receptors. However, nuclear receptor ligands can either act as agonists or antagonists, and more recently some small molecules have been defined to be inverse agonists. Antagonists are small molecules that bind to the nuclear receptor but do not induce the proper conformational change that activates transcription; thus the structure of the nuclear receptor may remain intact for the binding of corepresssors [43,68,70]. Potent antagonists are able to displace the ligand, returning the nuclear receptor from an active conformation to an inactive conformation, and again corepressor proteins are able to bind to the receptor. An active search exists for nuclear receptor agonists, and even more to also discover nuclear receptor antagonists. If chemical complementation were to be used for discovering antagonists, the antagonists would

displace the agonists, and turn off transcription, causing lack of growth. Therefore, to extend chemical complementation to discover small molecules that binds but do not activate nuclear receptors, a derivative of chemical complementation needs to be developed where the cells survive on selective media when transcription is turned off.

Besides applications for discovering new antagonists, a derivative of chemical complementation where a small molecule is not able to bind to the nuclear receptor, or one that causes an inactive conformational change, would also be useful for applications in protein and enzyme engineering. Specifically for protein engineering, if a receptor is being engineered to acquire a new function such that it binds and activates in response to a novel synthetic ligand and not respond to its natural ligand, the library of mutated receptors can be placed initially on selective media containing its natural ligand. In this case, chemical complementation will be designed such that those receptors that are unable to bind and activate in response to the natural ligand will survive; those that can bind and be activate by the natural ligand will die. This in essence, is the exact opposite of the chemical complementation as it has been defined thus far, and is called negative chemical complementation.

Negative chemical complementation is similar to chemical complementation in that the presence of a small molecule allows the yeast to survive. The difference is that in chemical complementation small molecule activators (e.g. agonists) allow the yeast to survive whereas in negative chemical complementation, small molecules that do not activate the receptor allow cell survival (e.g. antagonists). Negative chemical complementation is developed using the *URA3* selective marker [116,117]. The *URA3* gene codes for orotidine-5'-phosphate decarboxylase, an enzyme in the uracil

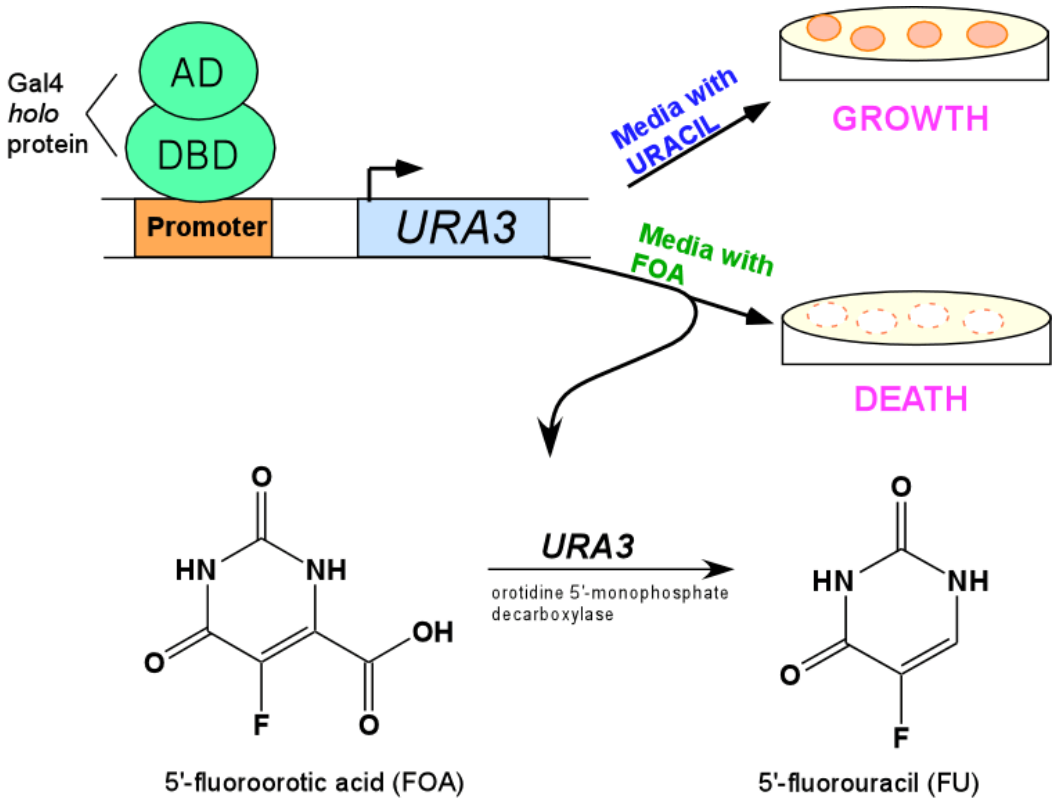


Figure 6.1. Mechanism of the *URA3* gene in Genetic Selection in Yeast.

biosynthetic pathway. This gene can be used for both positive and negative selection. For positive selection, yeast expressing this gene will survive in the absence of uracil in the media. For negative selection, uracil and 5-fluoroorotic acid (FOA) is added to the media, and expression of orotidine-5'-phosphate decarboxylase coverts FOA to the toxin 5-fluorouracil, which kills the yeast (Figure 6.1). Therefore, depending on the type of media that is used in selection assay, the *URA3* provides a package where both chemical complementation and negative chemical complementation can be performed using this selective marker, depending in the desired output [116,117]. If the desired outcome of chemical complementation is to discover a small molecule binder and activator, then the traditional chemical complementation can be used with positive genetic selection, using the *URA3* gene (Figure 6.2). In this case the cells would be grown in selective media lacking uracil. If however, the desired output is to discover a small molecule that can displace an agonist, bind but not activate the receptor, then negative chemical complementation is the proper assay. In this case, the cells would be grown in FOA medium with uracil, where only small molecules that do not activate the receptor will allow cell survival (Figure 6.2).

6.2 Negative Chemical Complementation with yeast strain, MaV103

Negative chemical complementation requires a Gal4 response element (RE) controlling expression of the *URA3* gene. For negative genetic selection in yeast, Vidal and coworkers created yeast strains that contain Gal4 response elements (REs) fused to the *URA3* gene. When fused to *URA3* genes, most yeast promoter tend to have high

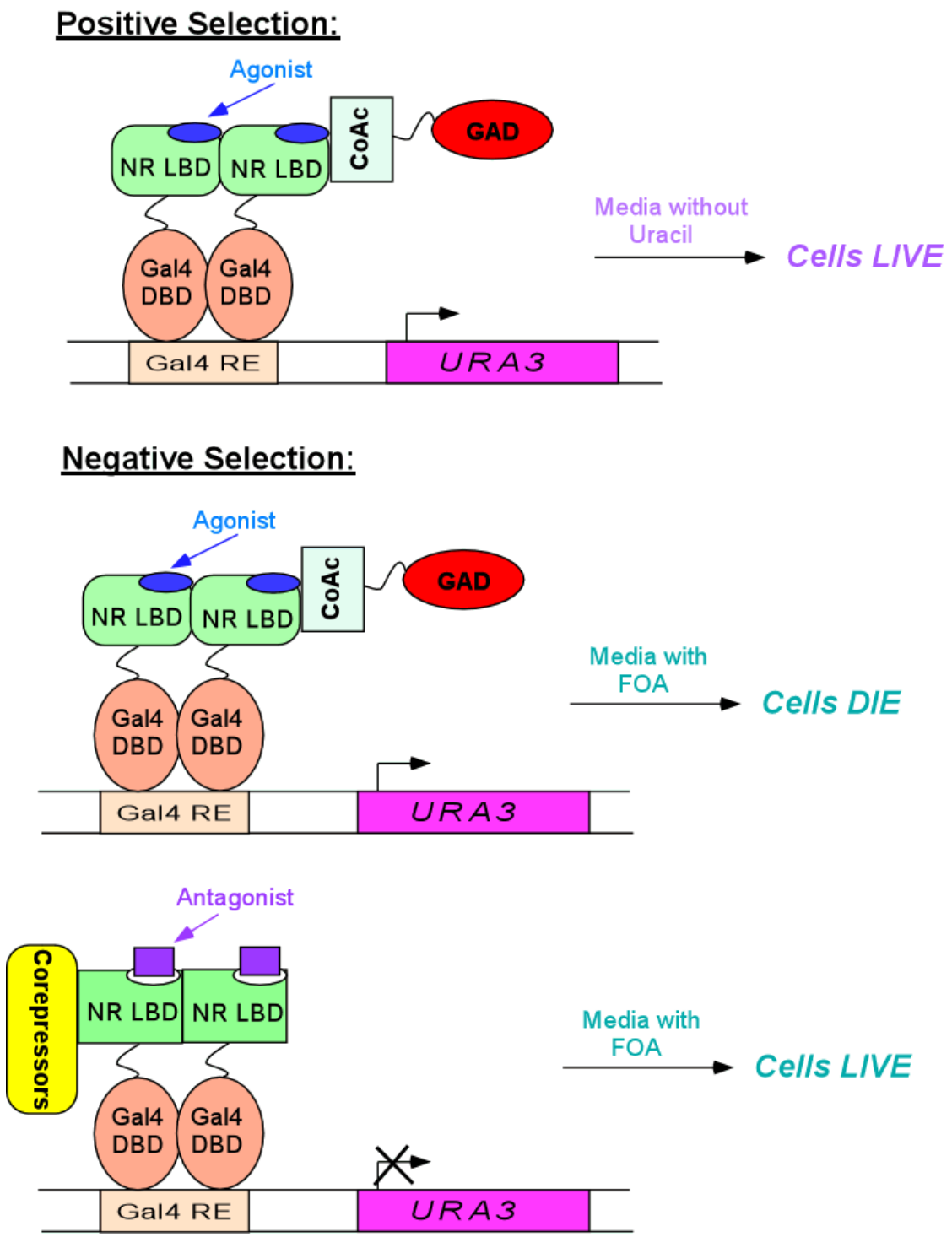


Figure 6.2. The *URA3* gene and Negative Chemical Complementation.

basal levels of expression of *URA3*, causing growth of the yeast cells on media lacking uracil in the absence of a transcriptional activator [116,117]. Therefore, to eliminate background basal growth, Vidal and co-workers fused the *SPO13* promoter to the *URA3* gene; this promoter is only activated under sporulation conditions and tightly repressed under normal conditions, creating a strain ideal for controlling transcriptional activation of the *URA3* gene for both positive and negative selection. Derivatives of the *SPO13* promoter were made to create a Gal4 response element, such that a strain could be developed for use in yeast two-hybrid assays with the Gal4 DNA binding domain (DBD) as part of the fusion protein [116,117]. The strain that was created, MaV103, was tested using the Gal4 *holo* protein as a control, along with fusion proteins of the Gal4 DBD and the Gal4 activation domain (AD) with strong and weak interacting protein pairs, and Vidal and co-workers were able to show that upon association of these proteins cell death occurred in media with FOA. However, if mutations or molecules were present that caused the dissociation of the two proteins, the cells were able to survive on media with FOA, thus creating a reverse two-hybrid assay such that the dissociation of the two fusion proteins causes cell survival. The Mav103 strain also contains a Gal4 response element controlling expression of the *HIS3* gene and the *lacZ* gene [116,117].

We decided to utilize the MaV103 strain in our chemical complementation assay. Plasmids pGBDRXR and pGAD10BAACTR were individually transformed and co-transformed into MaV103. Transformants were streaked onto uracil selective plates (SC -Ura -Trp) with 9cRA for positive selection. The same trend was seen with the ACTR: GAD with GBD: RXR in the yeast strain MaV103 strain as seen previously with the

PJ69-4A strain. The same transformants were streaked onto selective plates (SC -Leu -Trp) with FOA for negative chemical complementation. Varying concentrations of 9cRA were also added to the plates, ranging from 10^{-85} M to 10^{-5} M. In the absence of ligand (Figure 6.3), yeast grow on the sector of the plate containing ACTR: GAD with GBD: RXR as expected. This is expected because uracil is provided and in the absence of ligand RXR maintains its inactive conformation, preventing ACTR: GAD from binding and transcription does not occur. Without expression of the *URA3* gene, 5-fluorouracil is not produced and the yeast survive. However, as the concentration of ligand increases (Figure 6.3), less growth occurs and at the highest concentration of ligand, 10^{-5} M, very little growth occurs. The small amount of growth that is observed is due to background growth associated with negative selection in this strain.

With the MaV103 strain, we were able to show that in the presence of a small molecule ligand, we can perform negative chemical complementation, where the cells die in the presence of a small molecule agonist. However, the MaV103 strain has only the expression of the *HIS3* and *lacZ* genes under the control of Gal4 response elements, and the *HIS3* gene seems to have a high basal (data shown in Figure 6.10; discussed later in the chapter). The *ADE2* gene in this strain was not inducible by the Gal4 protein, and for these reasons we decided to create a derivative of PJ69-4A, which has a Gal4 RE, controlling expression of the *URA3* gene integrated into its genome.

6.3 Construction of Negative Strain, BAPJ69

We decided to develop a derivative of PJ69-4A to create a new yeast strain for negative chemical complementation. This strain was going to be designed such that three selective markers: *HIS3*, *ADE2*, and *URA3* were under the control of Gal4 inducible promoters. Furthermore, this strain will also have a Gal4 response element controlling the expression of the *lacZ* gene, used for the quantitation of negative chemical complementation.

To create this strain the plasmid pGH1, which was used by Phil James and co-workers to create PJ69-4A, was modified. This plasmid contains a *URA3* selective marker and contains the Gal4 response element controlling the *HIS3* gene. The *URA3* gene was amplified with restriction enzyme sites at the ends of the amplification, and cloned into the plasmid where the *HIS3* gene had been; creating a vector that contained a Gal4 response element controlling transcription of the *URA3* gene, as well as a constitutively expressed *URA3* selective marker on the plasmid. The *HIS3* gene in the pGH1 was digested out of the plasmid with the *BssHII* and *SacI* sites creating the vector cassette, and the *URA3* gene was amplified with primers containing the same restriction enzyme sites creating the insert cassette. The resulting plasmid, pGH1Gal1:*URA3* was confirmed using restriction enzymes and with sequencing.

Using the *URA3* for integration sites was difficult, because a number of clones just integrated a functional *URA3* gene into the mutated *ura3-52* gene on the chromosome, creating a strain in which the *URA3* gene was constitutively active and not controllable. Furthermore, we noticed that through this recombination strategy, the integrations seemed to be lost after multiple rounds of growth and budding. This is not uncommon, because it has been shown that recombination can occur between tandemly

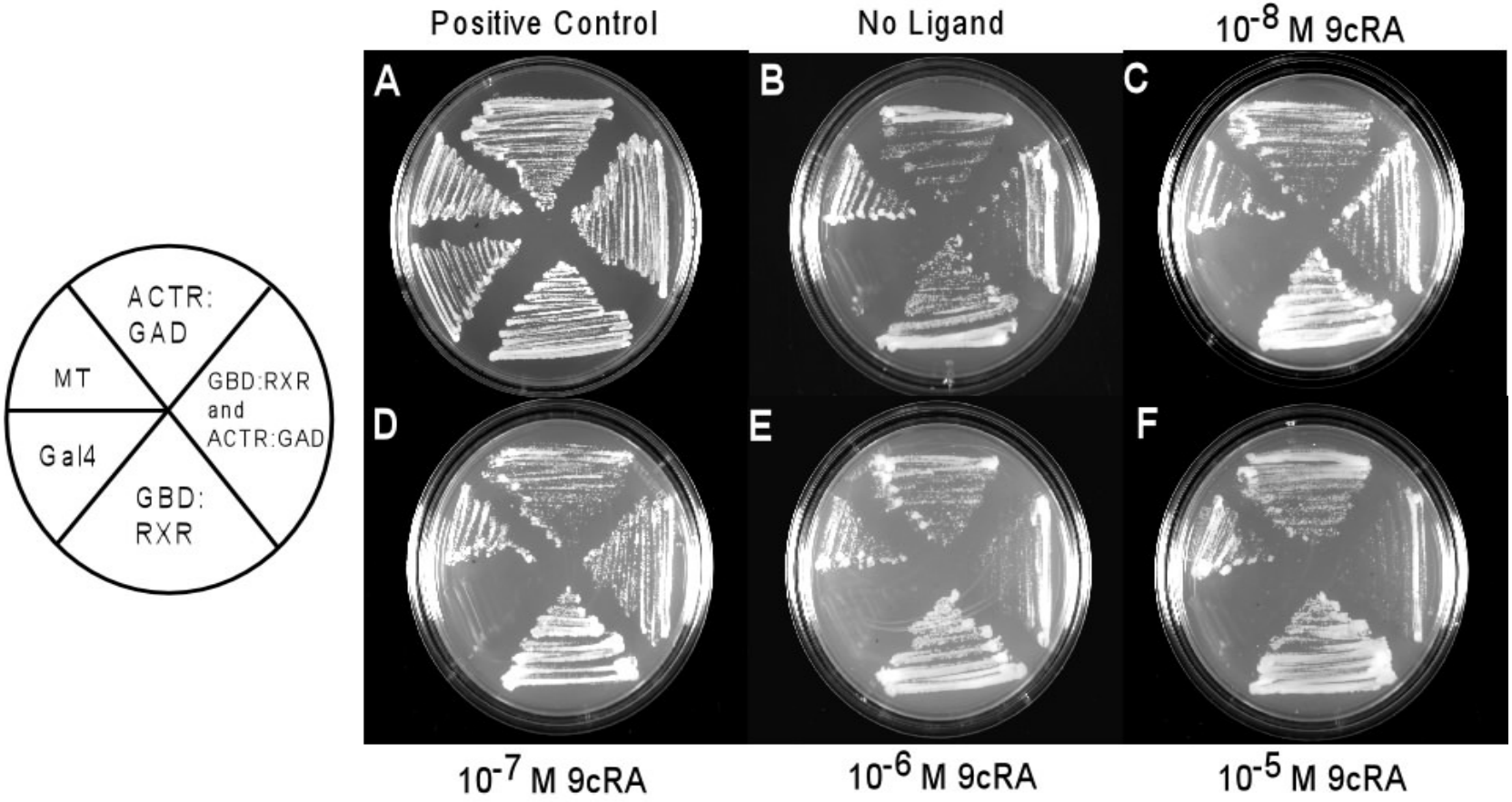


Figure 6.3. Negative Chemical Complementation with MaV103 with Gal4 and both RXR and ACTR.

integrated sequences, between the integrated *URA3* gene and the mutated *ura3*-552 gene in the yeast genome, resulting in the loss of the integrated gene. In some instances, the integrant was shown to be lost with colony PCR, but the strain's phenotype showed to contain a functional *URA3* gene; the strain was able to survive on selective media lacking uracil. This was not unlikely, because depending upon where the recombination excision had occurred, the part of the gene required for functionality may have still remained intact, leading to the uracil prototrophy.

To create this strain, plasmids were constructed at David's Stillman's laboratory (Utah) with target integration at the *HO* locus of the yeast genome [118]. This particular site was used for several reasons. The *HO* gene encodes an endonuclease involved in the introversion of the mating-type locus and is involved in promoting haploid cells to diploid [118]. Disruption of this part of this gene has shown to have no effect on growth rate, or the functionality of the organisms. As a matter of fact, the majority of yeast strains used in the laboratory for research have a mutation at the *HO* locus [118].

The HO-poly-HO plasmid contains portions of the HO gene with a multiple cloning site (MCS) inserted in between the two HO gene portions, as shown in Figure 6.4. The plasmid does not contain a yeast selective marker for plasmid-linked survival in yeast or autonomously replicated sequences (ARS) required for plasmids to independently survive and be passed onto daughter cells, therefore, leaving integration as the only option for the plasmid to be propagated [118]. The Gal1:*URA3* fragment from the plasmid pGH1:Gal1*URA3* plasmid was amplified using PCR primers containing *BamHI* and *BglII* restriction enzyme sites. The HO-poly-HO plasmid was digested with the same restriction enzymes, and these two fragments were ligated. The resulting

plasmid, HO-Gal1*URA3*-HO was confirmed by restriction enzyme diagnostics and sequencing.

Once the HO-Gal1*URA3*-HO plasmid had been created, the plasmid was digested with *PvuI* as shown in Figure 6.4 and a linear fragment containing the HO-Gal1:*URA3*-HO was transformed into PJ69-4A with the plasmid containing the Gal4 *holo* protein (pGBT9Gal4). These yeast transformants were plated onto selective media lacking uracil and tryptophan. Several transformants were obtained and were analyzed through colony PCR. These transformants were also grown in non-selective media and plates to loose the Gal4 plasmid; this experiment is performed to ensure that the integration is stable that the integration is able to pass onto daughter cells and remains intact after several rounds of growth. The colony PCR was performed with primers in the HO locus of the genome and would help confirm that the linear fragment had indeed integrated into the HO locus. If the HO locus on the genome is intact without the integrant present, a PCR band of about three kilobases would be present. If however, the integration had occurred, then the PCR fragment with the integrant present would be about two kilobases. The colony PCR for the transformants looked favorable (data not shown) implying that the integration had actually occurred. However, once the cells were grown under multiple rounds of non-selective conditions to remove the Gal4 plasmid, and reintroduced with the Gal4 plasmid to confirm the stability of the integration, no transformants were observed. This indicated that cells had lost the integration.

We hypothesized that the integration was possibly lost due to recombination between tandemly integrated sequences of the integrated fragment and the genome. This could be due in part to the additional pieces of the plasmid that are integrated along with

the HO-Gal1:*URA3*-HO fragment. Removing the additional fragments of the plasmid and just transforming the HO-Gal1:*URA3*-HO would eliminate any possible excision of the integrant with other sequences in the genome. Furthermore, if the HO-Gal1:*URA3*-HO plasmid could be passed onto daughter cells, eliminating the majority of the plasmid from the integration fragment, could prevent any possible contamination from the linear fragment re-ligating and possibly being passed onto other daughter cells.

The HO-Gal1:*URA3*-HO plasmid was digested with *NotI* and PvuI and the digest gel purified to obtain a clean HO-Gal1:*URA3*-HO fragment as shown in Figure 6.5. The fragment was then transformed into PJ69-4A with the pGBT9Gal4 plasmid containing the *holo* Gal4 gene. Transformants were selected on media lacking uracil and tryptophan (-Leu-URA), and then tested for the integration, using colony PCR using the primers shown in Figure 6.6. At the same time, the transformants picked for colony PCR are also grown in non-selective media to allow the yeast cells to loose the Gal4 plasmid. These cells are grown in Yeast Peptone Dextrose media (YPD) (rich; non-selective) and plated onto YPD plates. Once the cells have grown, they are grown again in YPD media and plated for another round of growth and replication. If integration has occurred and the integration is stable, the newly formed daughter cells will contain the integrant in their genome.

As shown in Figure 6.6, the colony PCR shows that several of the candidates have the two-kilobase fragment indicating the integration has occurred. However, in most of the samples, an additional band is present at about three kilobases, similar to the size of the band seen with just PJ69-4A without integration. The two samples, which do not contain the additional band, are samples #26 and #5, shown in Figure 6.6. These two

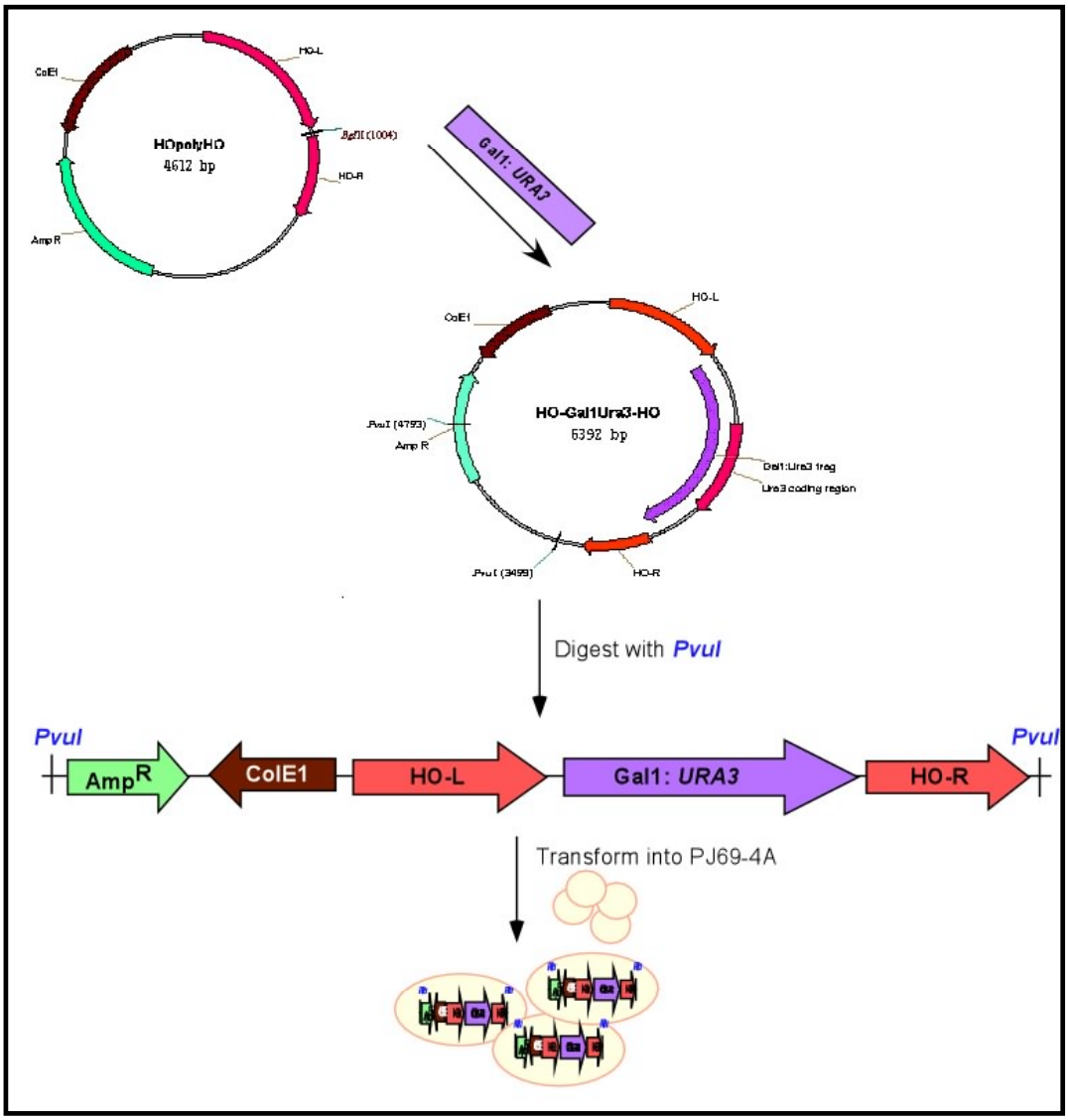


Figure 6.4. Creating *HO*-Gal1:*URA3*-*HO* fragment by digesting with *PvuI.*

possible candidates were then retransformed with and without the Gal4 plasmid (pGBT9Gal4) once the initial growth in non-selective media had indicated that the samples had lost the plasmid. Transformants were again selected for growth on media lacking uracil and tryptophan. Growth was seen on the selective media with both candidates (#26 and #5) only when the Gal4 plasmid was present, indicating that the integration of the HO-Gal1:*URA3*-HO fragment had indeed occurred and that the fragment has been stably integrated into the genome. More importantly, the integration has produced a Gal4 promoter inducible *URA3* gene and not one that is constitutively active. The #26 candidate seemed to have higher transformation efficiency and was pursued in further studies. The strain was named BAPJ69.

6.4 Comparing Strains: PJ69-4A and BAPJ69

The newly formed negative strain BAPJ69 was then subjected to a series of tests to compare the activity of its other selective markers with the parent strain, PJ69-4A. Both strains were transformed with pGBT9Gal4 and the combination of pGBDRXR and pGAD10BAACTR (see above). The pGBT9Gal4 samples were also transformed with pGAD10BA, a plasmid that contains a leucine marker. This transformation was performed so that when the cells were streaked onto selective media, selecting for both the pGBDRXR plasmid and pGAD10BAACTR plasmids (i.e. media lacking histidine, leucine and tryptophan) the Gal4 samples will also be able to survive. Transformants were then selected for growth on adenine, histidine, uracil and FOA selective media with and without 9cRA.

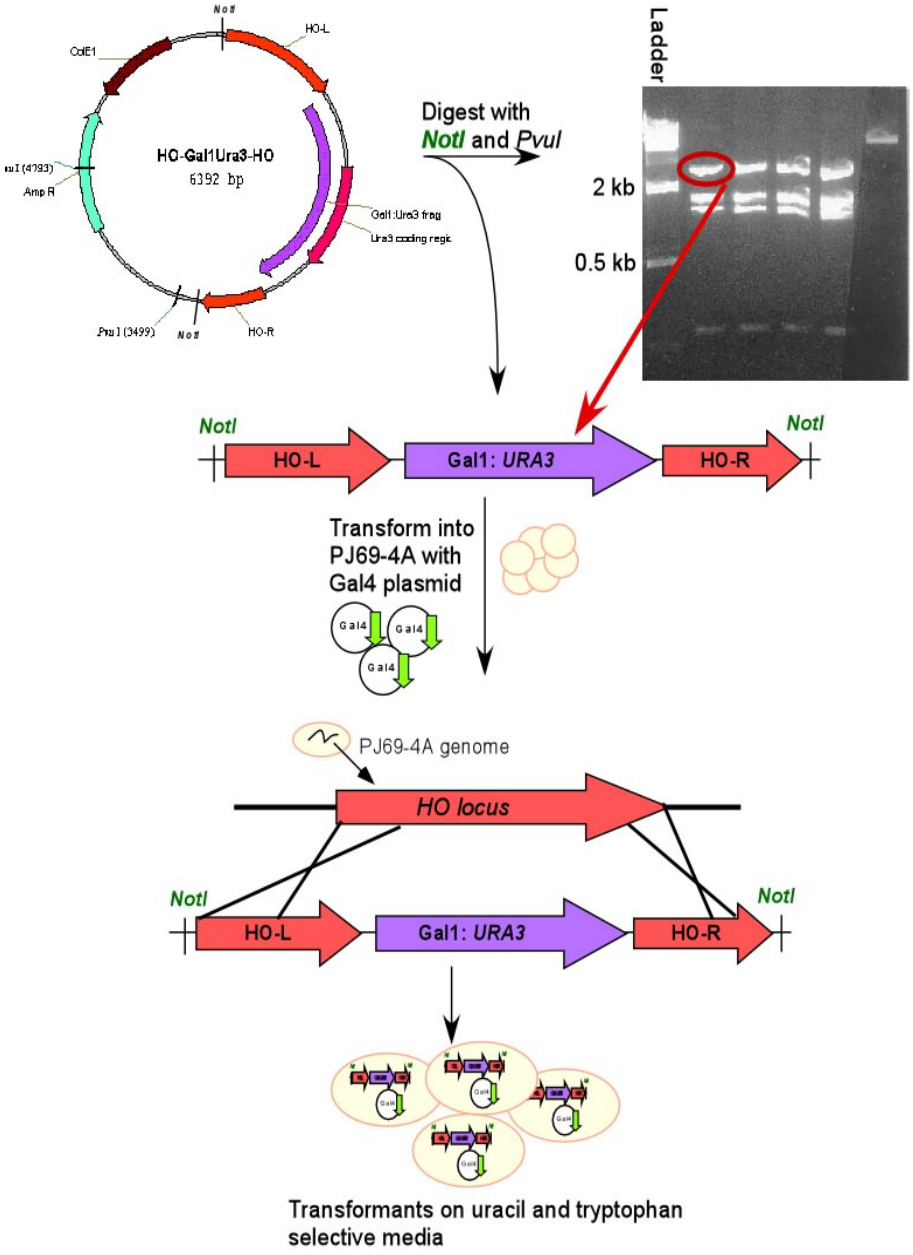


Figure 6.5. Creating *HO*-Gal1:*URA3-HO* fragment by digesting with *PvuI* and *NotI* for recombination.

The histidine selective media contains 1 mM 3-aminotriazole (3-AT) that is used to reduce the basal histidine expression previously observed. The adenine and histidine selective media both contain varying concentrations of 9cRA. As expected, in the absence of ligand, only the sectors of the plate that contain the Gal4 *holo* protein are growing (6.7). However once ligand is added to the plates, the quadrants of the plate containing the RXR and ACTR fusion proteins. Growth occurs at 10^{-8} M 9cRA with both strains, PJ69-4A and BAPJ69, and the growth density and rate are the same for both of these strains. The same results are seen on adenine selective media (Figure 6.7).

Once these cells were streaked onto uracil and FOA selective media, a similar trend was observed for negative chemical complementation. As shown in Figure 8, on uracil selective plates, growth is only observed on the sectors of the plates with the BAPJ69 strain. On the uracil selective plate with no ligand present, growth again is only observed on the sector of the plate containing the Gal4 plasmid in the BAPJ69 strain. Once ligand is added, growth again is seen at concentrations as low as 10^{-8} M 9cRA and growth is seen on all ligand plates (Figure 6.8).

With negative chemical complementation and the FOA plates, if the *URA3* gene is expressed and FOA is present, the toxin, fluorouracil is produced, causing cell death. The FOA selective plates in this experiment do not contain uracil, leucine or tryptophan, but the FOA solution itself contains a concentration of uracil, therefore, uracil is present in the media. Due to the presence of uracil in the FOA media, the sectors of the plates that contain PJ69-4A cells are growing. These cells contain the Gal4 *holo* protein and the pGAD10BA (plasmid with leucine marker) and the RXR and ACTR plasmids. Thus, they are able to overcome the selective pressure because the pGBT9Gal4 and pGBDRXR

plasmids contain tryptophan markers and pGAD10BA and pGAD10BAACTR contain a leucine marker.

On the FOA selective plates, the sectors of the plate containing the PJ69-4A cells are able to survive because of the following reason: (1) they are not expressing the *URA3* gene, (2) the addition of uracil in the FOA media allows then to overcome that selective pressure, and (3) they contain plasmids to overcome both the leucine and tryptophan deficiencies. On the sectors of the plate containing the BAPJ69 strains, the exact opposite effect is seen. These cells contain a Gal4 response element controlling Transcription of the *URA3* gene. Without ligand present, the only sector of the plate in which growth is not seen is the sector of the plate with the Gal4 plasmid. This is expected because Gal4 is ligand-independent and is expressing the *URA3* gene, causing cell death. Gal4 growth is not observed on any of the FOA plates; in some instance there is some leaky expression of the *URA3* gene observed. Once ligand is added, growth is observed at the lower concentrations of ligand (10^{-8} M and 10^{-7} M) on the sector of the plate containing RXR and ACTR, and less growth is observed on the plate containing a 9cRA concentration of 10^{-6} M, and the least amount of growth is observed at the highest concentration of ligand as expected (Figure 6.8). These results reaffirm that the BAPJ69 strain functions as the PJ69-4A strain does, and also has the additional *URA3* selective marker, which functions as a positive selective marker for chemical complementation and a negative selective marker for negative chemical complementation.

6.5 Comparing Strains: PJ69-4A, BAPJ69, and MaV103

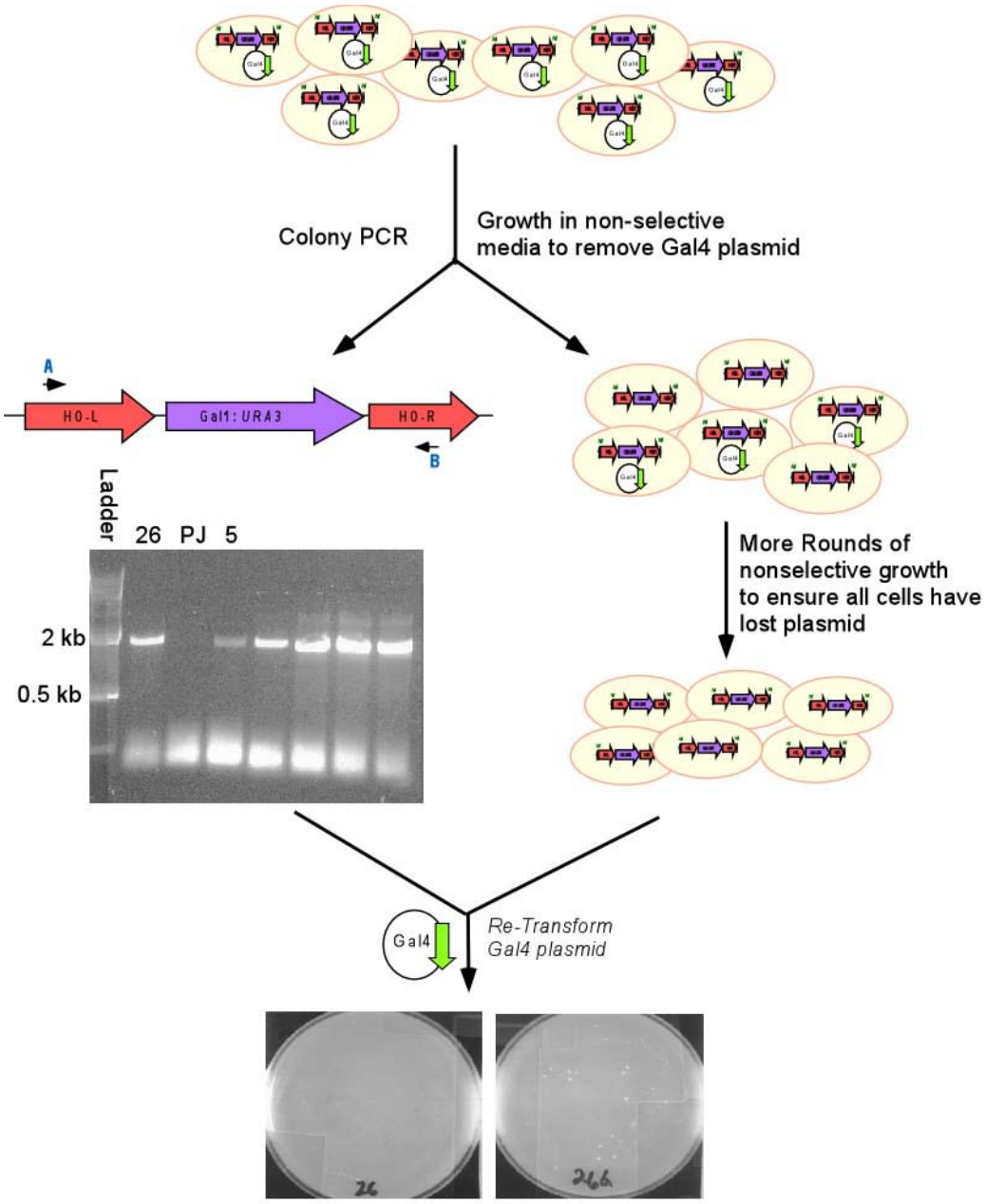


Figure 6.6. Confirmation of BAPJ69 using PCR and Yeast Transformations.

We had originally tested negative chemical complementation with the MaV103 strain and wanted to compare BAPJ69 with the MaV103 strain. In looking at each of the genotypes, the MaV103 contains a *HIS3* and *URA3* selective genes under the control of Gal4 inducible promoter, whereas BAPJ69 contains both of those selective genes as well as the *ADE2* gene. Both strains contain Gal4 response elements controlling expression of the *lacZ* gene used in β-Galactosidase assays.

The Gal4 plasmid was transformed into the MaV103, BAPJ69, and PJ69-4A strains and the transformants were streaked onto various selective media: histidine (with 1mM 3-AT), adenine, uracil, and FOA media (Figure 6.9). Growth was observed on all of the sectors of the plate on the histidine and adenine selective media with all three strains, as expected. On the uracil selective plate, only the sectors of the plate containing MaV103 and BAPJ69 showed growth, as excepted since PJ69-4A does not contain a Gal4 response element controlling expression of the *URA3* gene. On the FOA selective media, both sectors that contain the negative strains are dead; this is also expected because *URA3* is being expressed and with FOA present, the yeast cells are dead. The sector of the plate containing PJ69-4A shows growth because uracil is present in the FOA media and the lack of tryptophan in the media is compromised by the Gal4 plasmid Trp marker. A complete correlation is seen on these plates with the MaV103 and BAPJ69 strains.

A second set of experiments were performed using the quantitation assay of chemical complementation. The cells containing Gal4 and both RXR and ACTR were transformed into each of the strains, and transformants were grown in media selecting for the plasmids (media lacking leucine and tryptophan). After an overnight growth, these

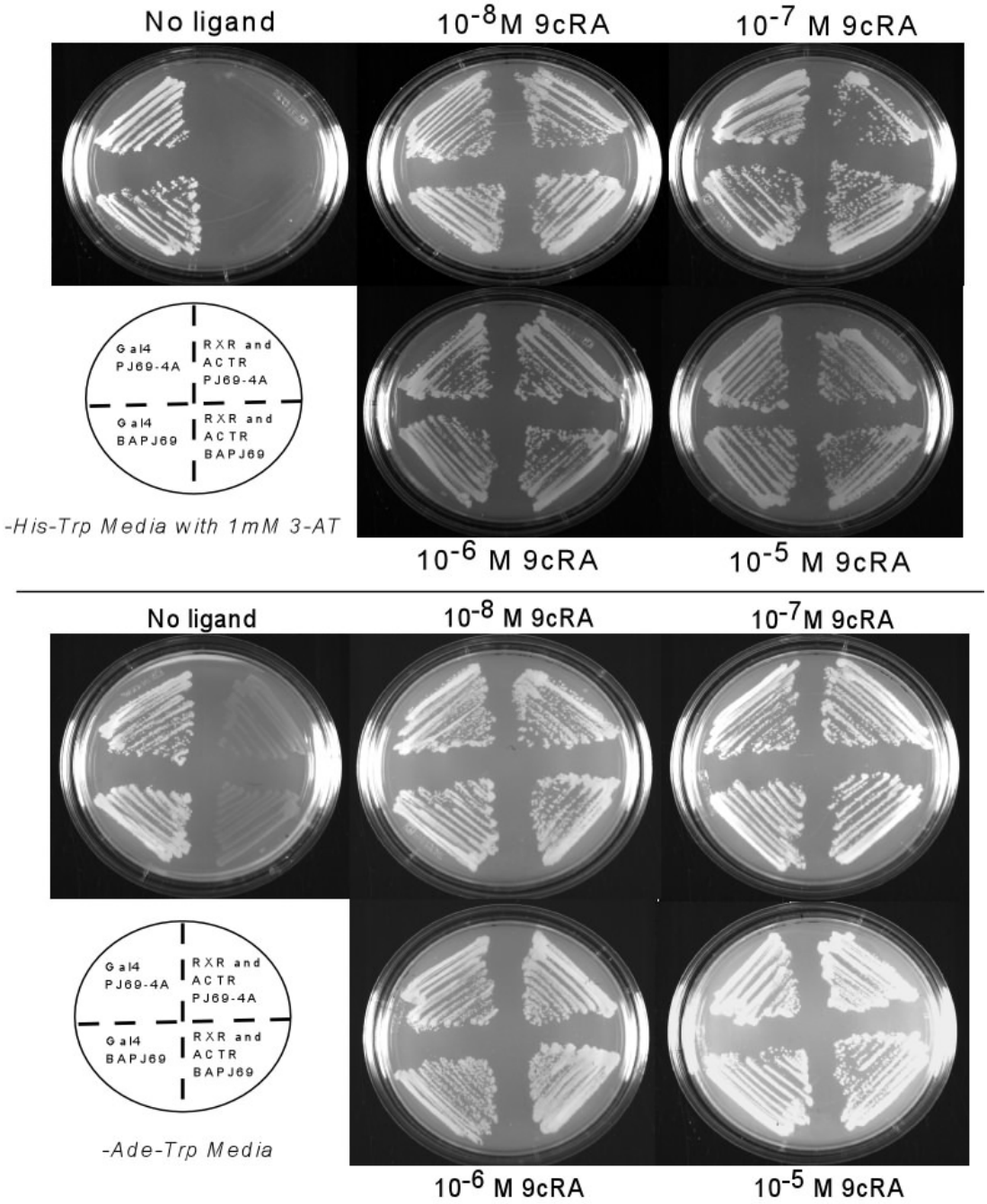


Figure 6.7. Comparison of PJ69-4A and BAPJ69 in Histidine and Adenine Selective Media

cells were placed in 96-well microtiter plates containing histidine, uracil, and FOA selective media with and without varying concentrations of 9cRA. This experiment was set-up to determine whether the same results were seen in liquid media as were seen on solid media (plates).

In the first set of experiments in histidine selective media containing 1mM 3-AT, the cells were grown in the 96- well s for about 40 hours and the difference in growth is observed in Figure 6. 10. As expected, a dose response curve is seen with all the strains containing RXR and ACTR. The does response in the BAPJ69 and PJ69-4A strains are almost identical, whereas the RXR and ACTR in the MaV103 strain shows an extremely high basal. Without the presence of ligand, growth is seen at an OD of about 0.3 with MaV103, whereas in the two strains, growth is not observed without ligand present. A nice dose-response curve is observed with both BAPJ69 and PJ69-4A where transcriptional growth is observed at about 10^{-7} M concentrations of 9cRA, and as the concentration of ligand increases, the growth increases as well. The same relative trend is observed with the MaV103 strain, except the presence of the high basal. The strains that contain Gal4 show a steady growth as expected, and each strain seems to have relatively the same amount of growth (Figure 6.10).

After about 40 hours of growth, these cells were subjected to the β-Galactosidase assay. The cells were lysed with yeast lytic enzyme and Triton X 100 and the cell lysates were placed in the presence of ONPG (See Chapter 5). The colorless ONPG is converted to ONP, a yellow product in the presence of β-galactosidase, and this color change is measured at an OD of 405nm. As shown in Figure 6.10, in the β-Galactosidase assay, a dose-response curve is observed for all three strains containing RXR and ACTR, but the

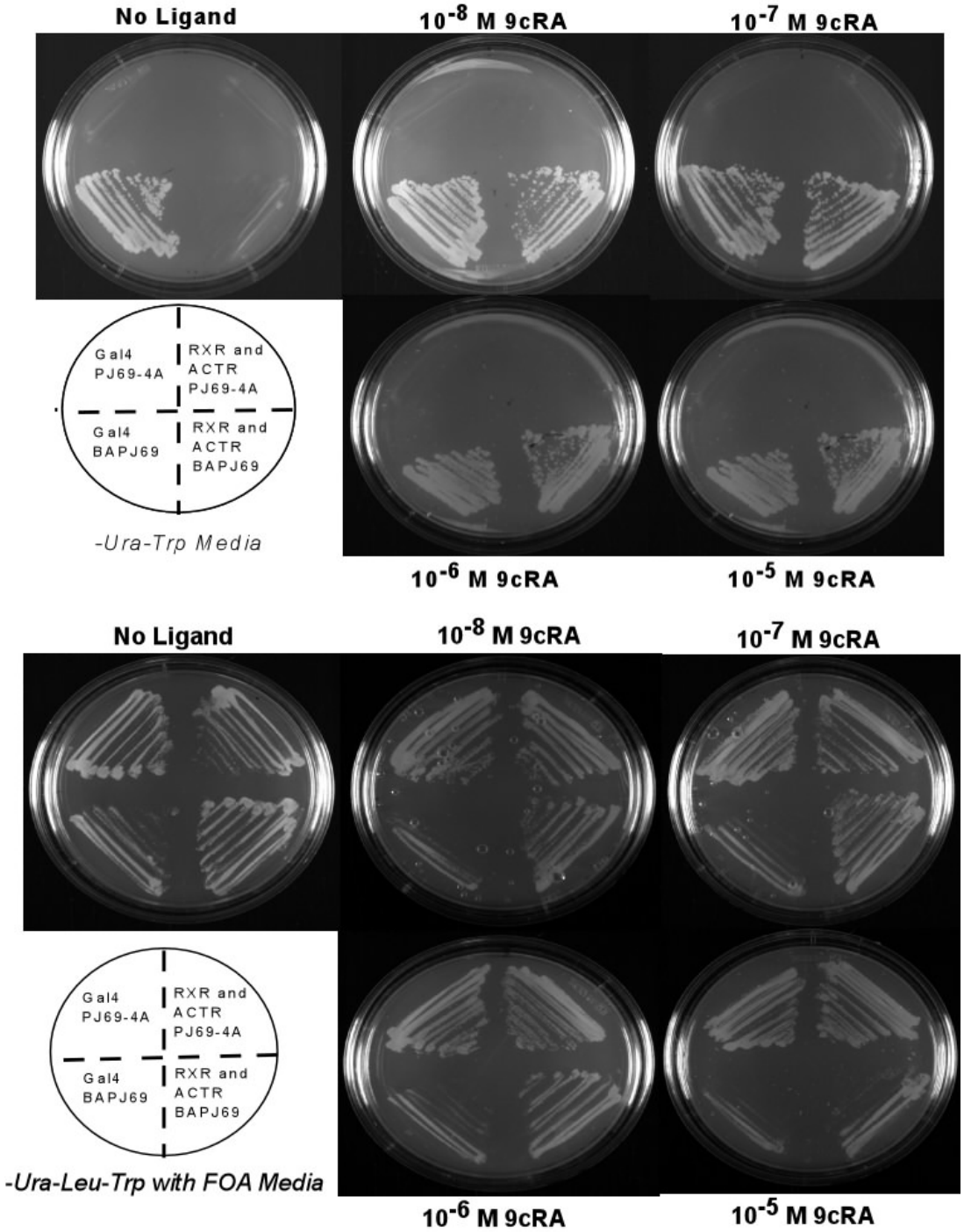


Figure 6.8. Comparison of PJ69-4A and BAPJ69 in Uracil and FOA Selective Media.

dose-response curve observed with the MaV103 strain has much lower transcriptional activation in comparison to the other two strains. The BAPJ69 and PJ69-4A strain both show good levels of transcriptional activation, comparable to Gal4 activated growth in the BAPJ69 strain. The β-Galactosidase activity observed with Gal4 in the MaV103 strain showed higher levels of activation than the other two strains, which is relatively surprising since all the strains seemed to show the same amount of growth. However, this may be due to the presence of more Gal4 response elements controlling the *lacZ* gene in the MaV103 strain. Overall, the correlation between solid plates and liquid media exists, but with the quantitation assay, we can observe the slight differences in growth and β-Galactosidase activity which is not detectable in a qualitative assay.

In the uracil selective media quantitation growth assay, most of the expected results were observed. The PJ69-4A cells showed no growth as expected (Figure 6.11). The RXR and ACTR sample in BAPJ69 showed a dose-response with 9cRA; as the ligand concentration increases the growth increases as well. However, one aspect of surprise is that the MaV103 strain containing RXR and ACTR showed no ligand-activated growth, which was surprising since negative chemical complementation has been performed using this strain, and the same result was obtained when these results were repeated. The MaV103 strain with Gal4 showed growth as expected, and growth was observed with the Gal4 plasmid in the BAPJ69 strain, as expected (Figure 6.11).

Once these samples were tested in the β-galactosidase assay, the same relative trend as the growth assay was observed. No activity was observed with the MaV103 strain with RXR and ACTR. A nice dose response curve was observed with 9cRA in the BAPJ69 strain with RXR and ACTR. Gal4 in both the BAPJ69 and MaV103 strains

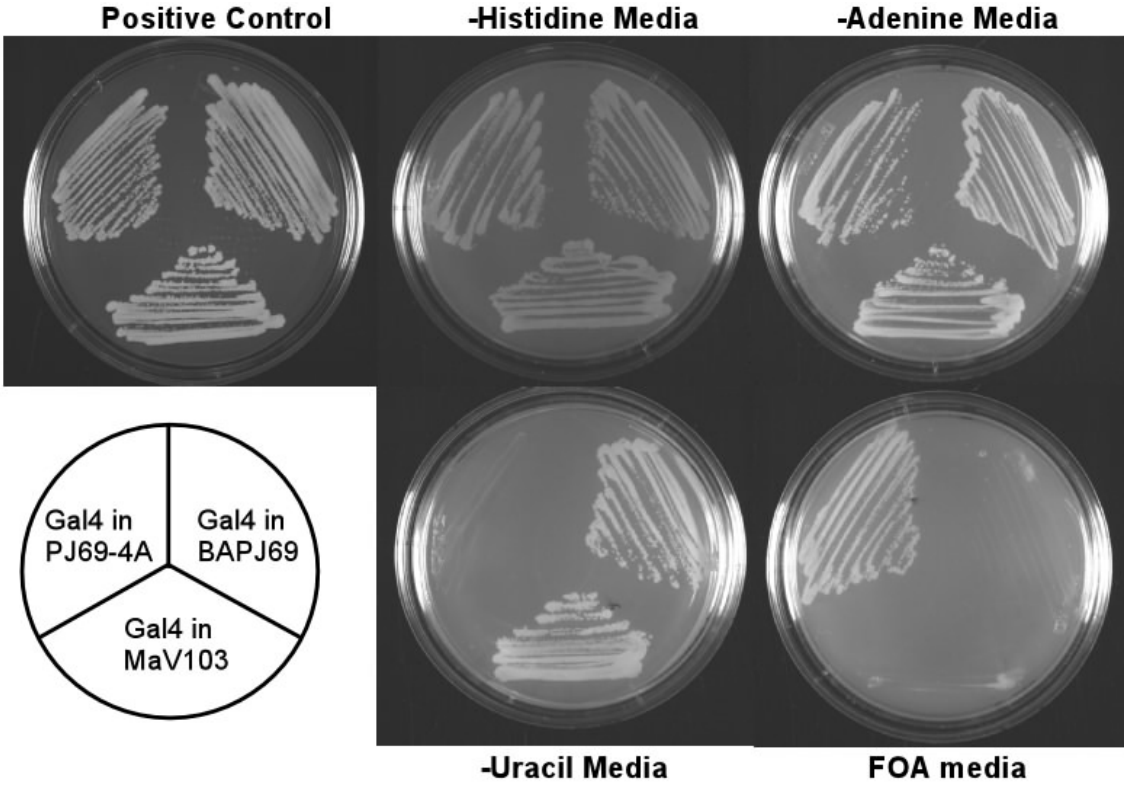
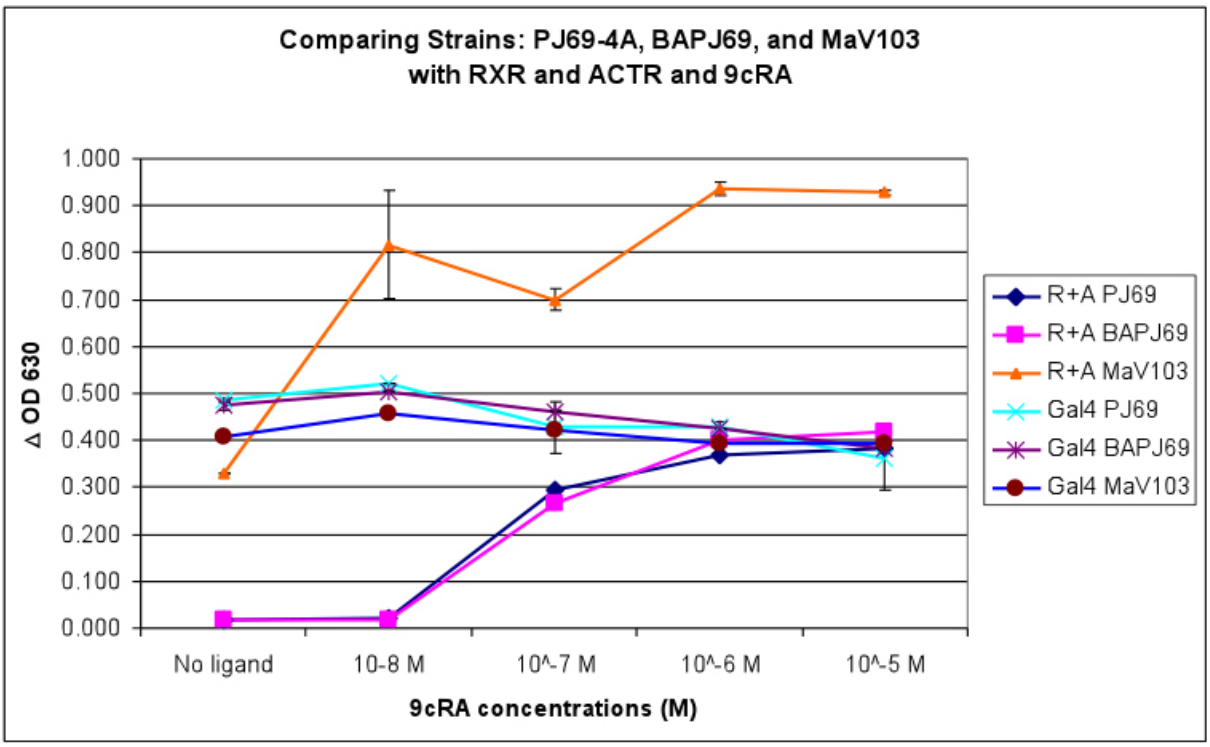


Figure 6.9. Comparing Gal4 activity in these three strains: PJ69-4A, BAPJ69 and MaV103.

showed a consistent β-galactosidase activity, but again MaV103 is observed to have a higher β-galactosidase activity than the BAPJ69 strain, but nevertheless, the activation seen with both strains and Gal4 is consistent (Figure 6.11).

A growth assay was performed in FOA selective media, with the media lacking uracil, tryptophan, and leucine, but the FOA solution contains some uracil. Thus, the PJ69-4A cells are able to grow, since they are able to overcome the leucine and tryptophan deficiencies, since some uracil is present in the media (Figure 6.12). This is observed with the PJ69-4A cells containing Gal4 and both RXR and ACTR. The growth with PJ69-4A and both RXR and ACTR is not ligand-dependent, since there is no additional selective pressure present (e.g. adenine or histidine). The BAPJ69 strain with Gal4 shows no growth as expected, and the same trend is observed with the Gal4 plasmid in the MaV103 strain. On the other hand, with both RXR and ACTR in the Mav103 strain, there is a slight growth that is observed but is not ligand-dependent; a dose-response is not observed with increasing ligand concentrations (Figure 6.12). With the BAPJ69 strain though, a ligand-dependent growth response is observed. With the FOA present in the media, as the concentration of ligand increases, the yeast cells begin to die. The cells are growing when no ligand or very little ligand (10^{-8} M) is present, but as the concentration increases, less and less growth is observed. At the highest concentration of ligand, no growth is observed (Figure 6.12). In comparison to the results observed on solid media shown in Figure 6.8, the quantitation assay shows greater sensitivity than the results on solid media, because on the plates, a difference in the density of yeast cells was noticed prominently at the highest concentration of ligand,



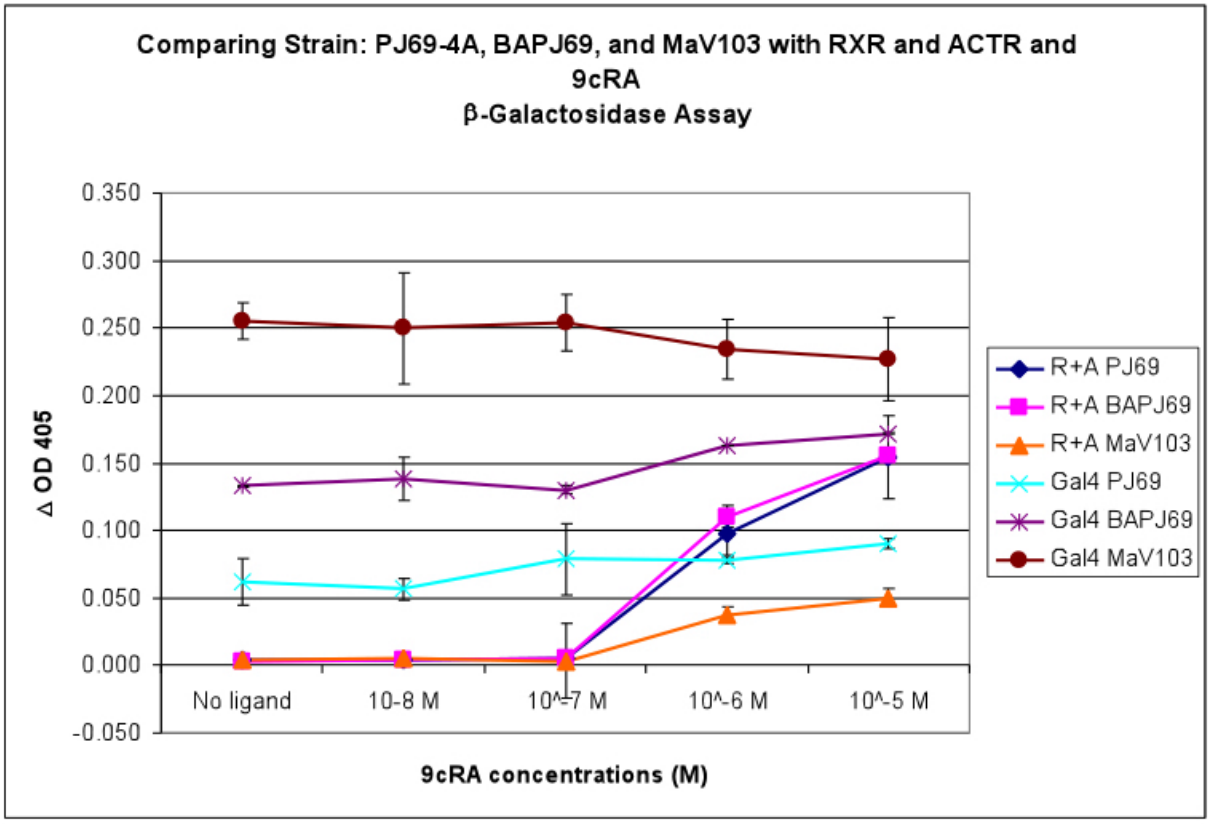


Figure 6.10. Comparing PJ69-4A, BAPJ69, and MaV103 in quantitation using histidine selective media.

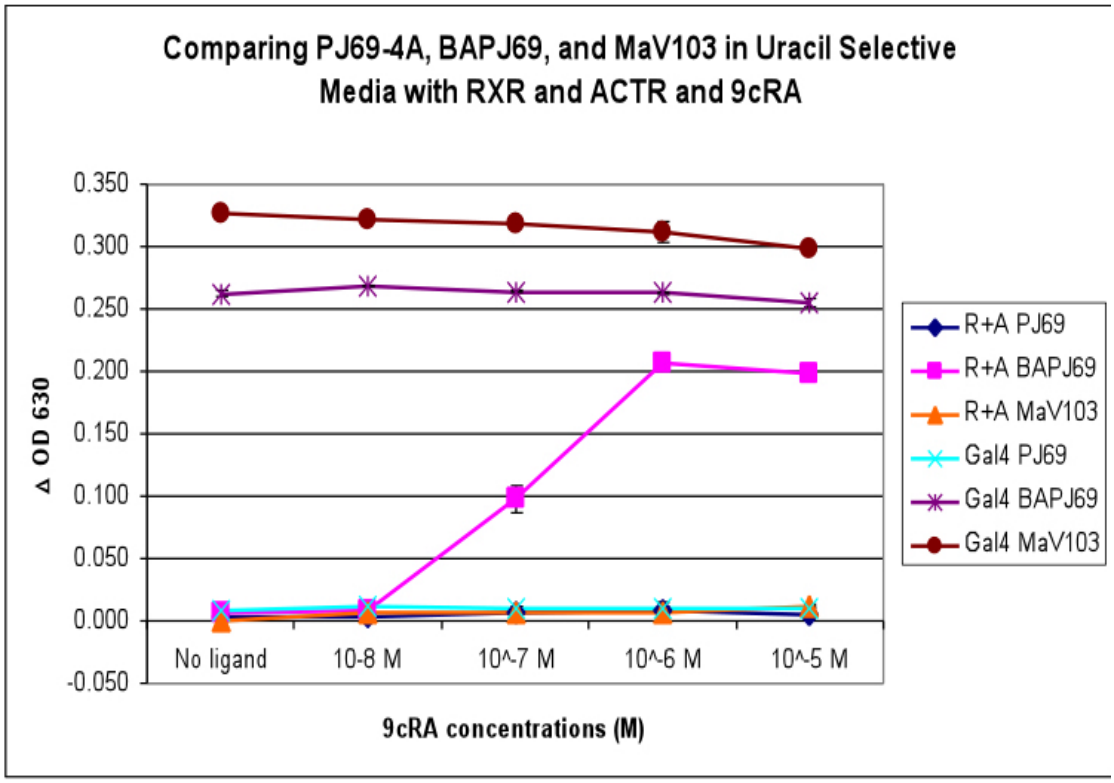


Figure 6.11. Comparing PJ69-4A, BAPJ69, and MaV103 in quantitation using uracil selective media.

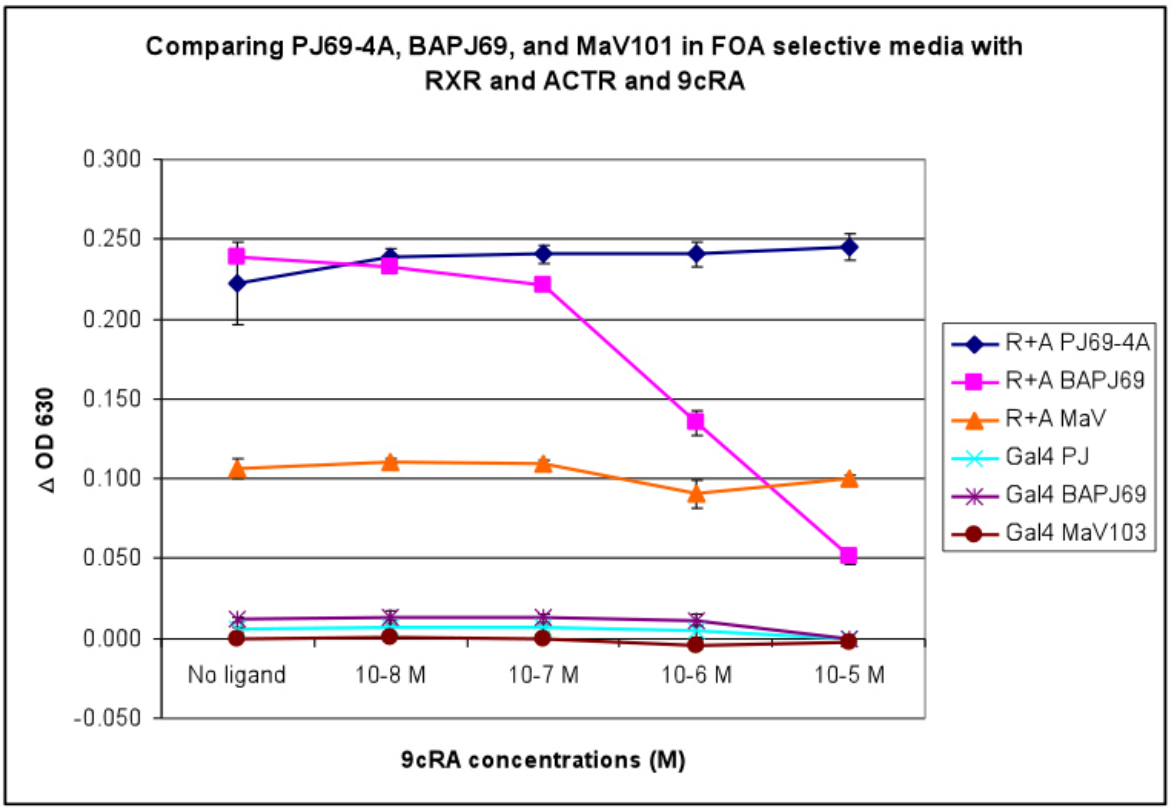


Figure 6.12. Comparing PJ69-4A, BAPJ69, and MaV103 in quantitation using FOA selective media.

whereas with the quantitation assay, a marked decrease in growth is observed at about 10^{-7} M (6. 12). Overall, these experiments show that the new negative strain, BAPJ69, functions as expected with a Gal4 inducible promoters controlling expression of the *HIS3*, *ADE2*, *lacZ*, and *URA3* genes (Figure 6.13).

We have been able to improve and expand the versatility of chemical complementation in this chapter and in the previous chapters. We have developed chemical complementation, where agonists of nuclear receptors allow the survival of the organism. We have also developed negative chemical complementation, where antagonists of nuclear receptors will allow the yeast to survive. Both chemical complementation and negative chemical complementation provide a powerful system for evaluating protein-ligand interactions. In the following chapters, applications of both chemical complementation and negative chemical complementation are described and demonstrated.

6.6 Materials and Methods

MaV103 Strain with RXR and ACTR. and Gal4

Yeast transformants containing the plasmids were streaked onto selective plates, -Leu -Trp, containing 5-fluororotic acid, FOA, and different ligand concentrations. Plates were also divided into sectors, with pGBT9Gal4 and pGBDMT as controls. The same procedure was used for streaking as for the adenine selection plates. Plates were incubated for two days. Each set of the genetic selection plates was replicated at least once.

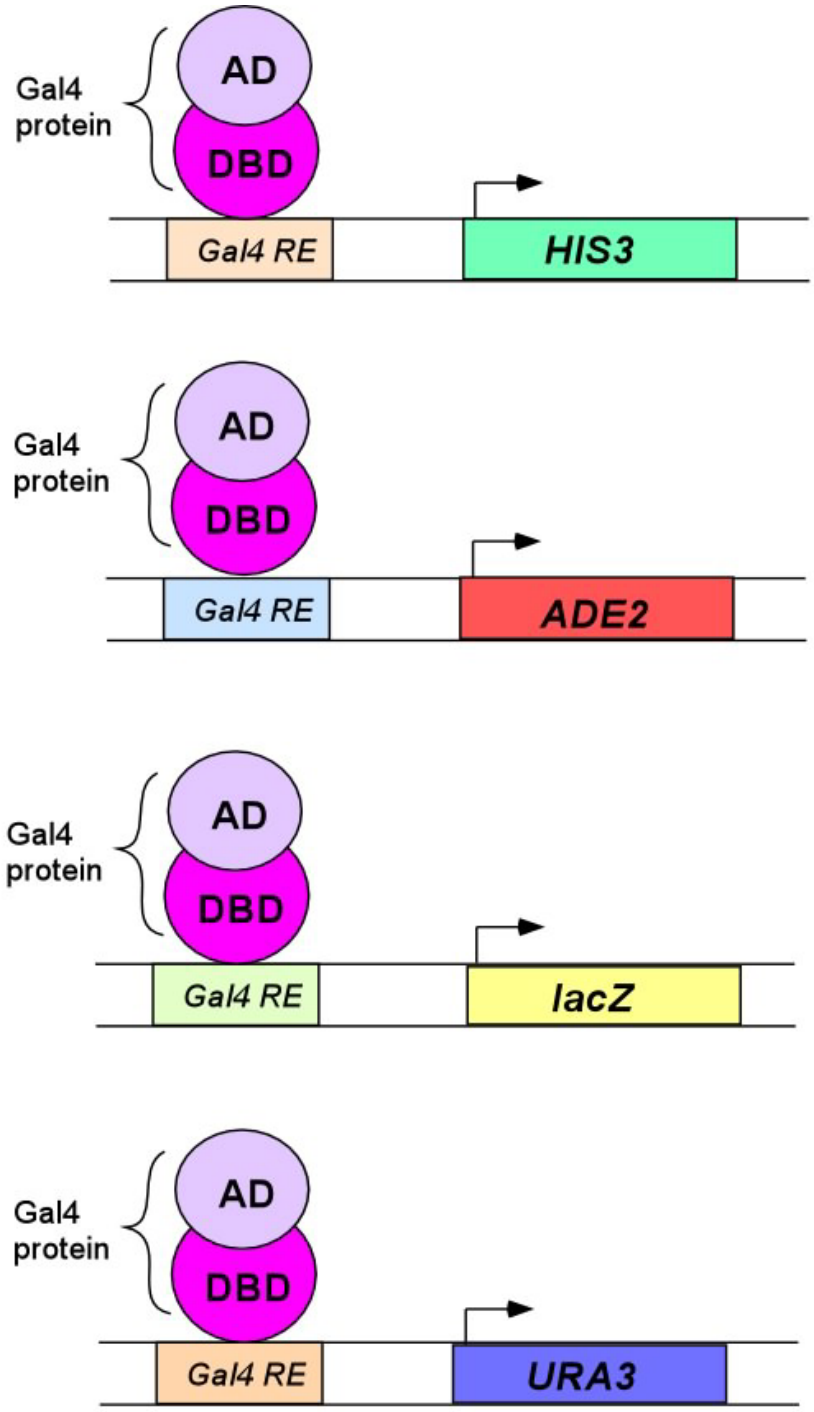


Figure 6.13. Genotype of BAPJ69.

Creating HO-Gal1URA3-HO fragment

The pGH1:Gal1:*URA3* plasmid was constructed using standard cloning techniques. The pGH1 plasmid (2 μg), which contains a Gal1 promoter controlling expression of the *HIS3* gene, was digested using *BssHII* and *SacI*, creating the vector cassette. After incubation at 37° overnight, the vector cassette was treated with 1 μL of calf-intestinal phosphatase (CIP), and gel cleaned using the Zymogen Gel DNA Recovery Kit (Zymo Research, Orange, CA). For the insert cassette, the pGH1 plasmid also contains a *URA3* selective marker, and the *URA3* gene was amplified from the pGH1 plasmid (100ng) with 125 ng of PCR primers containing *BssHII* and *SacI* sites on both ends, and a PCR reaction was performed with *Pfu* polymerase using the following PCR reaction: 95 °C for 3 min, 95 °C for 1 min, 55 °C for 1 min, and 72 °C for 6 minutes, repeated 25 times. The PCR fragment was cleaned using the DNA Clean and concentrator Kit (Zymo Research, Orange, CA). The *URA3* PCR fragment was digested with *BssHII* and *SacI* and incubated overnight at 37 °C, and again purified with the DNA Clean and concentrator Kit (Zymo Research, Orange, CA). After the concentrations of each fragment were determined, about 150 ng of the *URA3* digested fragment was legated to about 45 ng of the vector cassette using the Quick Legation Kit (New England Bolas, MA USA). The ligations were incubated at 22 °C overnight and transformed into Z-competent cells (Zymo Research, Orange, CA) XL1-Blue cells. Transformants were selected for growth onto LB plates with ampicillin (Amp) and the transformants that grew were selected and grown in LB Amp media. The plasmids were isolated from these samples using the QIAprep Spin Miniprep Kit (Qiagen Sciences, MD USA) and were diagnosed with restriction enzymes and eventually sequences for confirmation.

To construct the HO-Gal1*URA3*-HO plasmid, PCR primers were designed to amplify the Gal1:*URA3* fragment from pGH1:Gal1*URA3* with *BamHI* and *BglII* sites at the ends of the primers. The PCR product was cleaned using the Zymo Gel DNA Recovery Kit (Zymo Research, Orange, CA) and the digested overnight with *BamHI* and *BglII*. The Gal1:*URA3* PCR fragment was then cleaned with Zymo Research DNA Clean and Concentrator Kit (Zymo Research, Orange, CA). The HO-poly-HO plasmid was digested with *BamHI* and *BglII* and CIP was added to the overnight digest. The vector was cleaned using the Zymo Research DNA Clean and Concentrator Kit (Zymo Research, Orange, CA). The Gal1:*URA3* and HO-poly-HO plasmid cut with *BamHI* and *BglII* were ligated together using the Quick Ligation Kit (New England Biolabs, M USA) and ligations were incubated at 22 °C overnight and transformed into Z-competent cells (Zymo Research, Orange, CA) XL1-Blue cells. Transformants were selected for growth onto LB plates with ampicillin (Amp) and the transformants that grew were selected and grown in LB Amp media. The plasmids were isolated from these samples using the QIAprep Spin Miniprep Kit (Qiagen Sciences, MD USA) and were diagnosed with restriction enzymes and eventually sequences for confirmation. This plasmid is named HO-Gal1:*URA3*-HO.

Creating BAPJ69

The HO-Gal1:*URA3*-HO (6 μg) was digested with PvuI and *NotI* and subjected to CIP. The fragment of interest was gel purified using the Zymo Gel DNA Recovery Kit (Zymo Research, Orange, CA) and about 2 μg of the linear fragment was transformed into PJ69-4A using the LiAc Transformation protocol with pGBT9Gal4 (Gal4 *holo*

plasmid) [106]. Yeast transformants were selected on selective media lacking uracil and tryptophan (SC -Ura -Trp).

Transformants that grew on selective media (SC-Leu-Trp) were then analyzed to determine whether integration had occurred using two strategies. In the first method, colony PCR was used with primers A and B (Figure 6.6). Transformants were picked with sterile pipette tips and placed in PCR tubes with 10X Taq Buffer, 1mM $MgCl_2$, 100 mM dNTPs, 125 ng of each primer and 1 μL of Taq Polymerase. The PCR program used was 95 °C for 3 min, 95 °C for 1 min, 55 °C for 1 min, and 72 °C for 6 minutes, repeated 25 times. The PCR was run on a 1.2% agarose gel and if a band of 2.1 kb was observed then these samples had the integration present.

A couple of samples, #26 and #5 were shown to have the 2 kb band. These cells were then grown in non-selective media, yeast peptone dextrose (YPD), overnight and plated onto YPD plates. Samples that had lost the Gal4 plasmid would turn red on YPD media, and cells that looked red were picked again and inoculated into YPD media and grown overnight. The cells were plated onto YPD plates to ensure that the pGBT9Gal4 plasmid was lost.

The red-colored cells were selected and pGBT9Gal4 (1 μg) and pGBDRXR wild-type (1 μg) and pGAD10BAACTR (1 μg) were transformed into these cells as well as with a control where no DNA was added. Transformants were selected for growth on selective media lacking uracil and tryptophan (SC -Ura -Trp) for the Gal4 samples and SC- Leu-Trp plates for the cells with both RXR and ACTR. The no DNA sample was plated onto uracil selective plates (SC -Ura -Trp). The #26 sample showed growth with

Gal4 onto the uracil selective plates whereas no growth was observed without DNA on uracil selective plates. #26 was used for further evaluation and was name BAPJ69.

Testing BAPJ69 on selective plates

BAPJ69 was streaked onto adenine, uracil, histidine, and FOA selective plates with and without varying concentrations of 9cRA. As a control, cells containing pGBT9Gal4 are also streaked onto the plates. Plates are incubated at 30 °C for two to three days.

Quantitation with BAPJ69

MaV103, PJ69-4A, and BAPJ69 cells containing both RXR and ACTR and Gal4 were inoculated in selective media (SC-Leu-Trp) overnight. The quantitation protocol in Chapter 5 was followed where these cells were incubated in selective media containing varying concentrations of 9cRA. The 96-well microtiter plates were incubated at 30 °C with shaking at 150 rpm. After 40 hours of growth, the cells were lysed and subjected to the β-galactosidase assay as shown in Chapter 5.

Chapter 7

Applications of Chemical Complementation:

Drug Discovery

7.1 Nuclear Receptors and Disease

Nuclear receptors play a vital role in controlling many key physiological pathways for multiple organisms [42-45,56]. These receptors are involved in a number of signaling pathways, as well as cell proliferation, differentiation, and metabolism. Several nuclear receptors are involved in maintaining lipid and glucose homeostasis and inflammation. Due to their pivotal role in a number of biological pathways, these receptors are also involved in many diseases, such as diabetes, obesity, and artherosclerosis and have more recently been implicated in a number of cancers [43,44,119,120]. The role of nuclear receptors in these various diseases varies from mutations within the actual nuclear receptor that cause the defect, as in the case with the peroxisome proliferators activated receptor γ (PPARγ) and its role in Type II Diabetes [121-123]. The interaction of the nuclear receptor with other proteins can also cause a defect, such as in acute promyelotyic leukemia; the cause of this disease is a result of the fusion of retinoic acid receptor α (RARα) and other transcriptional proteins[46].

Pharmaceutical industries have a vast interest in discovering and developing small molecules that can be used for treating conditions such as diabetes and cancer. Nuclear receptors form the second largest class of drug targets, where the current market for nuclear receptor drug targets are estimated to be at 10%-15% of the global pharmaceutical market [43,44]. The most commonly used drugs in the market as this time that function towards nuclear receptors are tamoxifen for breast cancer treatment,

thiazolidinediones for diabetes, and dexamethasone for inflammation [48,121-125]. Yet, as more research is being performed about these nuclear receptors, more dysfunctions of nuclear receptors are being discovered leading to other medical conditions, therefore, the ongoing need for drugs, which target nuclear receptors remains prominent. The following paragraphs outline some of the nuclear receptors and their role in disease.

The estrogen receptor was one of the earlier nuclear receptors discovered and has been implicated primarily in breast cancer [43,44,120,126,127]. Breast cancer is one of the leading causes of death for women in the world, and novel effective therapies remain elusive despite the fact that the mortality rate for women with breast cancer has decreased due to the methods in earlier detection and better treatment options. A large percentage of breast cancers are dependent upon estrogen for growth, and the inhibition of the estrogen signaling pathway would decrease the size of the breast tumors. Inhibition or reduction of the estrogen signaling pathways may help decrease the size of the tumors, however, estrogens are involved in multiple signaling pathways and many have beneficial effects towards the development and maintenance of physiology. For example, estrogens are known to reduce artherosclerosis and known to reduce the severity of Alzheimer diseases and maintaining normal cognitive function in post-menopausal women. Due to the positive and negative effects of estrogen, the need for selective estrogen receptors modulators (SERMs), small molecules that would produce tissue-selective effects on various diseases, would be beneficial. These small molecules would target the certain pathologies and not affect normal cellular function. Several SERMs are currently in clinical trials for a possible treatment towards osteoporosis and cancer, such as tamoxifen

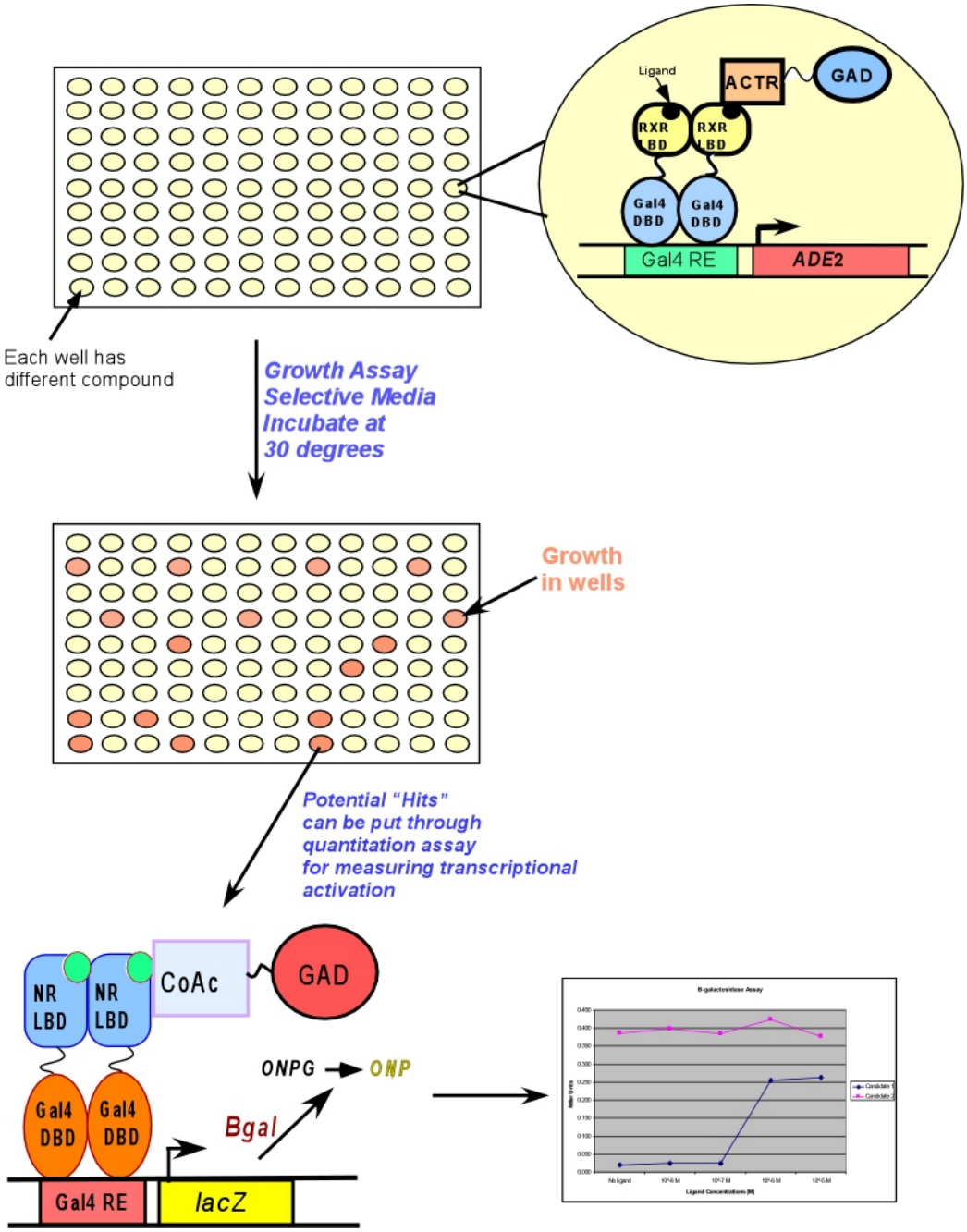


Figure 7.1. High-throughput chemical complementation assay for discovering agonists of nuclear receptors.

(Nolvadex ®, AstraZeneca), which displays antagonistic effects in breast, and agonist effects in other tissues, such as bone uterine, and cardiovascular tissue. Another SERM currently in clinical trials is Raloxifene (Evista ®, Eli Lilly), which also exhibits agonistic activity in the bone and antagonistic activity in the treatment against breast cancer. The success of these two SERMs, especially with tamoxifen in the cure of breast cancer, more research has been forth towards developing these small molecule modulators with more versatile applications, such as for the long-term prevention of breast cancer reoccurrence [48,120,124,125,128,129].

The estrogen receptor is definitely not the only nuclear receptor with a role in the pathogenesis of cancer. The vitamin D receptor (VDR) which binds 1α, 25-dihydroxycholecalciferol D_3 (the most active metabolite of vitamin D) is involved in calcium endocrinology and bone formation [43]. Recently, it has been shown that vitamin D analogs can reduce the occurrence of mammary tumors, and also enhance the effectiveness of tamoxifen to prevent tumors. Research has been forth towards understanding the effects of vitamin D on breast cancer, and for using synthetic analogs of vitamin D to create novel agonists for the receptor. However, a number of these small molecules that have been developed also induce hypocalcaemia and may have other possible long-term effects. The VDR is also shown to have an effect in colorectal cancer; vitamin D has been shown to reduce incidents of colorectal cancer.

The androgen receptor (AR) is a nuclear receptor able to bind androgens that mediate a wide range of the male physiology, such as male sexual development [43,119]. Androgens are essential for the normal function and development of the prostate, over-activation of the AR signaling pathway is also involved in prostate cancer. Initially,

Huggins and Hodges reported that the progression of prostate cancer involved the presence of androgens, androgen ablation in patients with prostate cancer resulted in tumor regression [43,119]. However, in most cases, although the tumor regresses initially, there is a development of androgen-independent growth of the cells, despite the androgen ablation. This is believed to be due to the presence of androgens in other organs, such as adrenal androgens. Currently, anti-androgens as well as castration are possible therapeutics towards the cure of androgen-independent prostate cancer [43,119].

Besides cancers, nuclear receptors are implicated in a number of metabolic diseases, such as diabetes and inflammation [43]. In particular, PPARs, especially α and γ are involved in these particular medical conditions. PPARγ is a nuclear receptor, specifically binds TZDs, and studies have show that this receptor plays a vital role in the development of type II diabetes. Several variants of PPARγ have been shown to destabilize the receptor, leading to hypertension, insulin resistance and diabetes [121-123]. More recently, a PPARγ variant (Pro 12 changed to Ala) has been shown to show a decreased risk of type II diabetes, probably due to inhibition of PPARγ regulated genes [121-123]. PPARγ has been widely studied and a number of synthetic analogs of TZDs have been developed as possible cures for diabetes and insulin resistance, and are currently in clinical trials. More recently, these receptors have been implicated in cancers as well [121-123].

7.2 Assays for Discovering Nuclear Receptor Ligands

Nuclear receptors play a major role in the mammalian physiology, and developing synthetic agonists and antagonists for clinical use against nuclear receptor based

pat*holo*gies are critical. Developing novel nuclear receptor ligands has been facilitated by advancements in combinatorial chemistry, *in silico* modeling and ligand docking into the nuclear receptors. Furthermore, with recent developments of high-throughput screening methods, the efficiency of discovering nuclear receptor ligands has greatly increased. Currently, most nuclear receptor agonists and antagonists are developed and discovered through high-throughput screening methods in mammalian cell assays. These assays are more time-consuming and relatively expensive to perform. With chemical complementation, we have developed a high-throughput assay in yeast, which provides a less time-consuming the less expensive method for discovering and developing nuclear receptor ligands.

7.3 Chemical Complementation for Discovering Agonists

Chemical complementation provides a high-throughput assay for evaluating nuclear receptor ligand interactions (Figure 7.1). Combinatorial libraries of synthetic or natural compounds can be obtained in 96-well microtiter plates, to which the yeast cells containing the nuclear receptor of interest will be added. If the particular compound in that well is indeed an agonist for the nuclear receptor, the ligand will be able to bind and activate transcription of the selective gene (*ADE2*, *HIS3*, or *URA3*). Cells in the well will be able to survive in selective media. If the ligand is not an agonist, the cells in those wells will not be able to grow (Figure 7.1).

Using chemical complementation provides a quick and relatively easy method for determining agonists in a high-throughput format. With the power of genetic selection, only the compounds that are actual agonists for the receptor will be able to survive.

These potential candidates will be then analyzed through the quantitation assay using the Gal4 response element controlling expression of the *lacZ* gene for determining the transcriptional activation profile of that particular compound (Figure 7.1). Using chemical complementation is more efficient than some of the other screen assays, where every compound is assayed for its transcriptional profile with the receptor. With genetic selection, time and effort will only be put towards those wells that are growing, indicating the presence of a potential agonist.

7.4 Chemical Complementation for Discovering Antagonists

Negative chemical complementation can also be utilized as a high-throughout assay for discovering antagonists. As in the case with agonists, compounds of interest will allow cell survival. Libraries of compounds can be obtained in 96-well microtiter plates, where to each well yeast cells containing the receptor of interest will be added. In addition to the yeast cells, FOA selective media with a concentration of the nuclear receptor's agonist will be added to each well. In each well, with the presence of an agonist, the cells will die; this is because the agonist will be able to activate the *URA3* gene. The Ura3 protein will then convert the FOA present in the media to 5'-fluorouracil, killing the yeast. The wells with cells growing are able to do so because the compound present in that well is able to replace the agonist, as an antagonist (Figure 7.2). The antagonist binding will induce a conformational change in the nuclear receptor, allowing the repression of transcription (Figure 7.2). This will prevent the expression of the Ura3 protein, and yeast cells will be able to survive because 5'-fluorouracil is not being produced (Figure 7.2). Again, as with the case of discovering agonists, genetic

selection provides a powerful tool where only the desired variants are able to survive. This provides a time-efficient system for discovering and developing nuclear receptor antagonists.

7.5 Extending Chemical Complementation to other Nuclear Receptors

Human Androgen Receptor (hAR)

The hAR is categorized as a steroid nuclear receptor involved in regulating prostate development and function [43]. The hAR binds to androgens, which are involved in regulating a series of genes essential for male reproductive development, such as the genes for prostatic specific antigen (PSA) and probasin [43,119]. hAR binds to the response elements of these regulated genes as homodimers, and upon binding of the small molecule, the nuclear receptor LBD undergoes a conformational change as with most other nuclear receptors, eventually, initiating transcription [130]. The structure of the hAR is relatively the same as the rest of the nuclear receptor superfamily, consisting of a DBD and LBD, both, which play vital roles in its function as a nuclear receptor. There are two transactivation domains, AF-1 and AF-2, where as shown in some other nuclear receptors, the AF-1 domain is involved in the ligand-independent transactivation of AR, whereas, the AF-2 domain is involved in a more general ligand dependent transactivation of the AR [43,119,130].

The growth and differentiation of prostate tissue are dependent on the AR and certain androgens, such as testosterone and dihydrotestosterone (DHT), the more active form of testosterone [43,119]. As mentioned above, the hAR is involved in the progression of prostate cancer, where a dysfunctional regulation and expression of

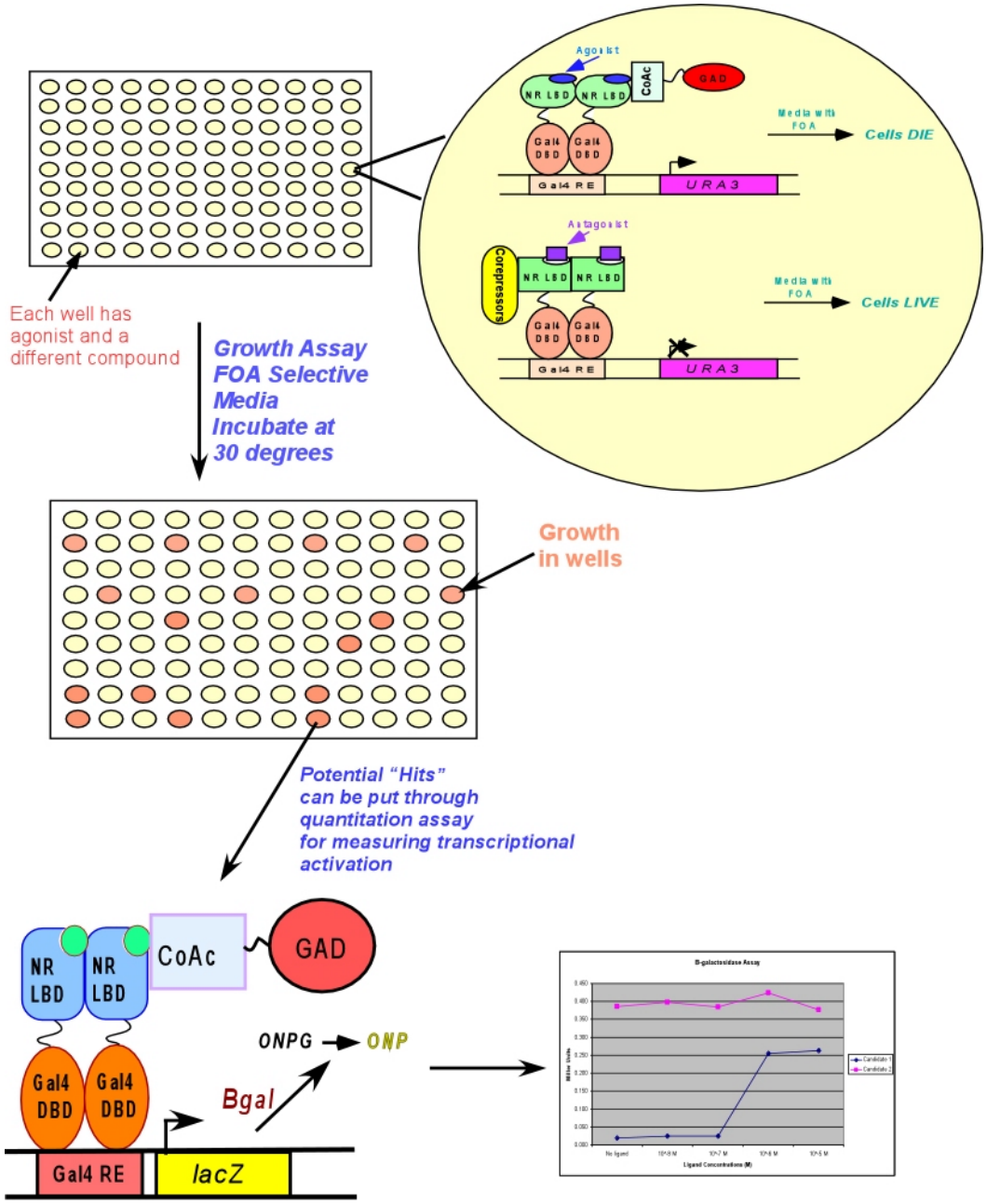


Figure 7.2. High-throughput chemical complementation assay for discovering antagonists of nuclear receptors.

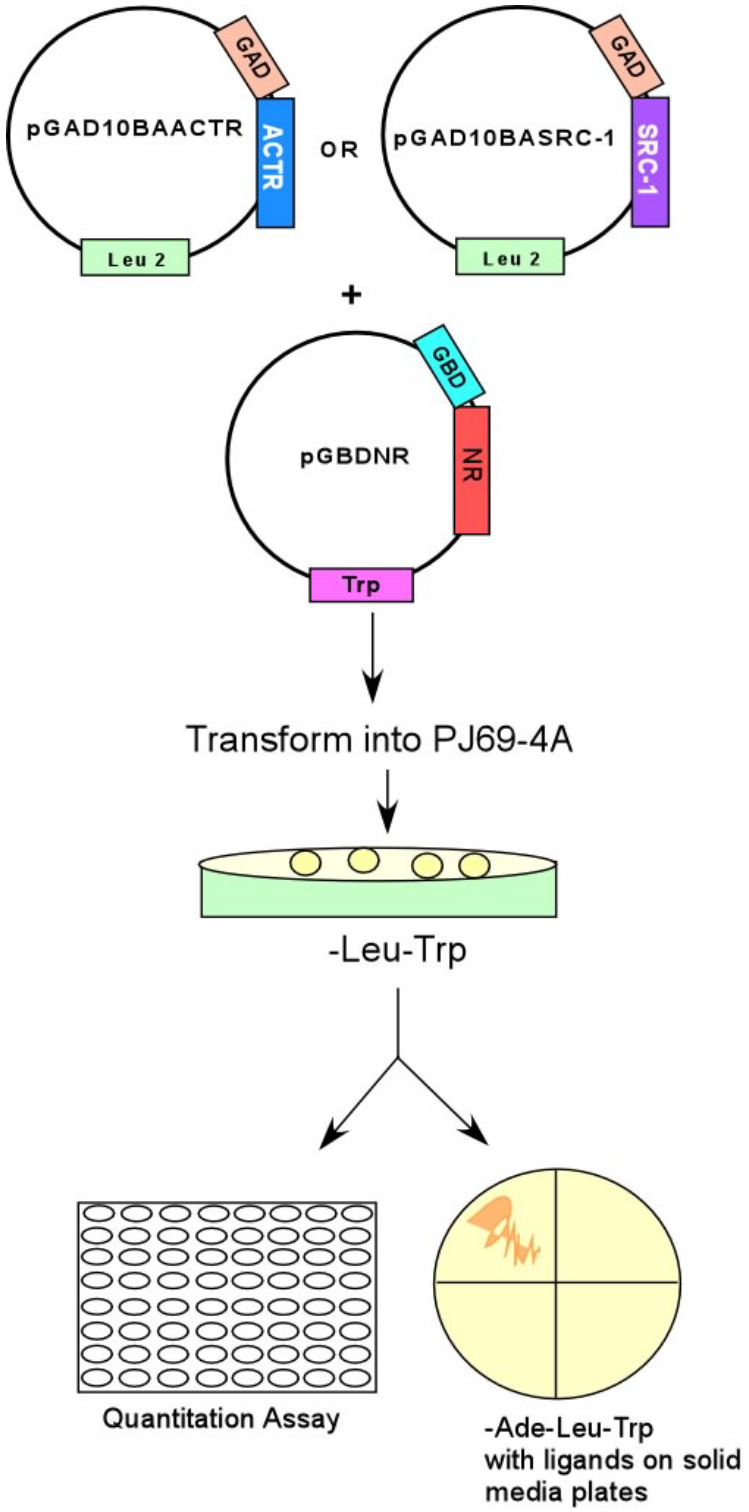


Figure 7.3. General experimental technique for evaluating nuclear receptors with chemical complementation.

AR co-regulators may lead to prostate cancer, and a more common cure for the earlier stages of prostate cancer is the ablation of androgen. Eventually, during treatment a development of androgen-resistant cells becomes present, which can make the treatment and cure of prostate cancer more challenging [119].

As a target for drug discovery based on its role in prostate cancer, we decided to test the androgen receptor in chemical complementation. A mammalian vector containing the androgen receptor (pCMVhAR) was obtained from Dr. Elizabeth Wilson (University of North Carolina, Chapel Hill) [131]. A plasmid was constructed containing the Gal4 DBD fused to the AR (full protein), name pGBDhAR (Figure 7.3). This plasmid, along with pGAD10BA ACTR, was transformed into PJ69-4A, where transformants were selected from media lacking leucine (pGAD10BASRC-1 has a leucine selective marker) and tryptophan (the pGBDhAR plasmid has a tryptophan selective marker). Transformants containing both plasmids were then inoculated in media selecting for both plasmids and subjected to the quantitation assay with histidine selective media. Histidine selective media was used with 3-AT in order to see determine whether the cells are constitutively active. Constitutively active receptors are receptors that are transcriptionally active without the presence of ligand; we hypothesize that this is possible due to the presence of an endogenous ligand that binds and activates the receptor.

Without the presence of ligand, the AR is transcriptionally active, and varying concentrations of 3-AT reduces this activation (Figure 7.4). As the concentration of 3-AT increases, the growth decreases as expected. The relative amount of growth observed with AR is comparable to Gal4 growth at the different concentrations of 3-AT (Figure

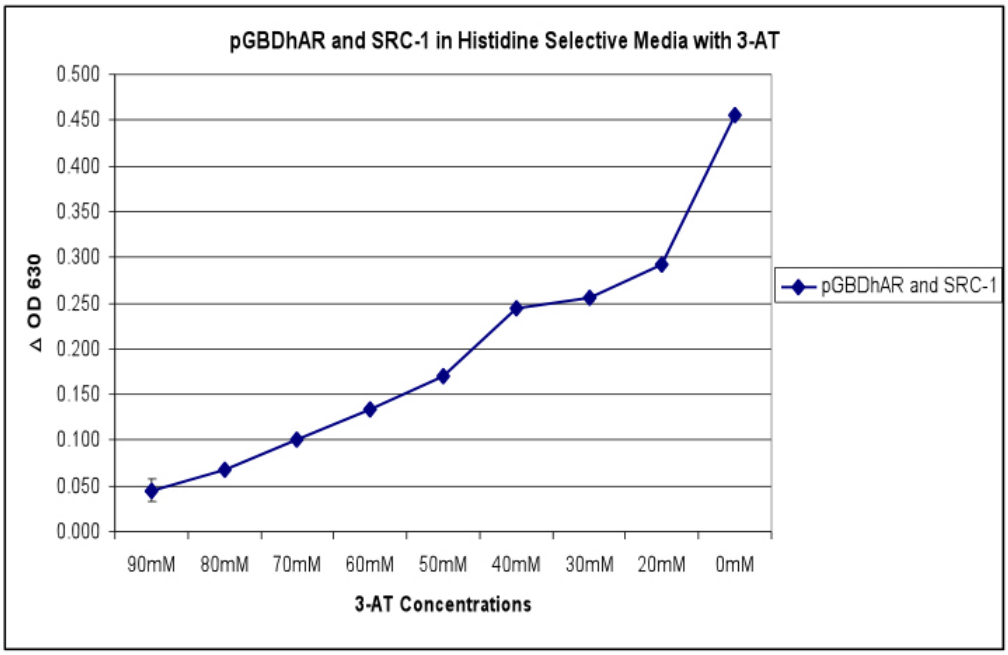


Figure 7.4. AR and SRC-1 in chemical complementation.

7.4). This indicates that the possible endogenous ligand binding to AR, binds and activates the receptor with high efficacy.

Nuclear receptor antagonists are involved in repressing transcriptional activation by displacing the agonist and recruiting proteins, known as corepresssors. These proteins are then able to recruit proteins known as histone deacetlytransferases (HDACs), which chemically modify the histones causing repression of transcription. Since, the AR is constitutively active, we decided to use a known antagonist of the receptor, resveratrol, and determine whether the antagonist can replace the endogenous agonist, and turn off the transcriptional repression. For antagonist assays of nuclear receptors, the agonist must be present at the optimal EC_{50} values to be able to detect the antagonistic effect of the ligand, but in the case with constitutively active receptors in chemical complementation, the endogenous agonist concentration is not controllable. Since the identity of the endogenous agonist is unknown, the EC_{50} value of the unknown ligand cannot be determined. Therefore, the antagonistic assay in this case would probably not be effective.

Yeast cells containing pGBDhAR and pGAD10BASRC-1 were grown in histidine selective media (-His-Leu-Trp) with varying concentrations of 3-AT and with and without resveratrol. With and without resveratrol, the growth curve was almost identical (Figure 7.5). If resveratrol was acting as an antagonists as was expected, the graph with resveratrol would shown the growth curve shifted lower, indicating less growth, or a curve that is flat at an OD of about 0.05, indicating a strong antagonist, causing complete repression of growth. However, as can be seen in Figure 7.5, we were unable to see a decrease in growth with the presence of resveratrol. This indicates that

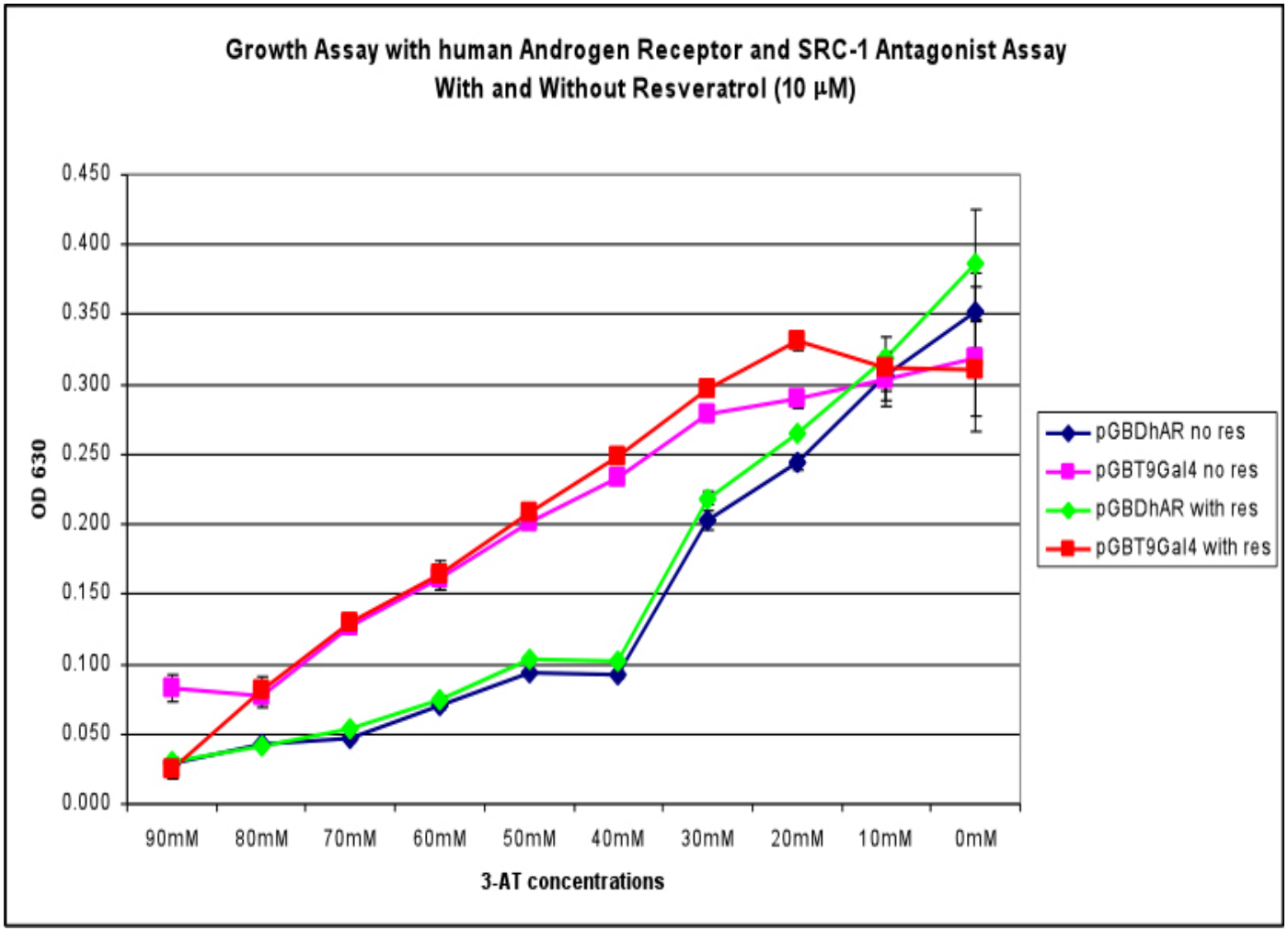


Figure 7. 5. AR with known antagonist, resveratrol

the endogenous ligand that is able to bind and activate the AR is indeed a potent agonist of this receptor because a known antagonist, resveratrol, was unable to displace the agonist present in the binding pocket, and recruit corepressors to repress transcription. Therefore, although we were able to show that the AR functions in yeast and in our chemical complementation system, the use of this receptor in chemical complementation for drug discovery is challenging because of the presence of the possible potent agonist present in yeast. This hypothesis is unlikely because yeast contain metabolites, which resemble the androgens, which may have the ability to activate this receptor.

Human Constitutive Androstane Receptor

The constitutive androstane receptor (CAR) is an orphan nuclear receptor involved in a number of biological pathways, especially in the metabolism of xenobiotics [43]. Activation of this receptor is involved in the regulation and expression of cytochrome P450 enzymes, along with other enzymes involved in dealing with xenobiotic stress [132,133]. The CAR is also involved in regulating thyroid hormone and bilirubin metabolism, making this receptor an attractive target for the treatment of various diseases, such as obesity and neonatal jaundice [43].

CAR is oftentimes compared to the pregnane X receptor (PXR), a nuclear receptor whose primary function is xenobiotic clearance from the body. CAR and PXR both share a highly conserved DBD, and moderately conserved LBD [134]. Both of these receptors form heterodimers with the retinoid X receptor (RXR. PXR and CAR both bind a number of the same small molecules, as well as number of similar molecules, but different activation profile for these small molecules [135,136]. In looking at the

PXR crystal structure, the PXR has a large, elliptically shaped binding pocket (volume ~1200 A^3), and suited to bind a range of larger compounds, such as paclitaxel and rifampicin, and smaller small molecules, such as steroid macrolides [134]. CAR, on the other hand, has a much smaller binding pocket and is able to bind diverse chemicals, including phenobarbital, the anti-inflammatory drug, acetaminophen, and the insecticide contaminant TCPOBOP [134]. However, similar to PXR, the activation profile of CAR is also species dependents; ligands that are able to bind and activate mouse CAR, such as TCPOBOP, can not activate the human isoforms of the receptor [137].

A unique feature of the CAR LBD is that this receptor is constitutively active in mammalian cell assays [43,132-134]. In most nuclear receptors, the transcriptional activation is induced by the conformational change upon ligand binding; this change properly positions the AF-2 helix for coactivators to dock into. In the case with CAR, the AF-2 helix is properly positioned independent of ligand binding, and mutations in the AF-2 helix have eliminated the constitutive activity of the receptor [43,132-134]. The crystal structure of CAR has recently been solved and through this structure, three distinct features have been shown to contribute to the constitutive activation of CAR. The first feature is the presence of a short linker helix before the CAR AF-2, which in most other nuclear receptors is a loop, but this short helix turns into a stable two-turn helix in CAR. This short helix packs tightly across helices 3 and 10, placing helix 12 in the active conformation [43,132-134]. The second feature of CAR that contribute to its constitutive activity is the short AF-2 helix, which forms direct hydrogen bonds with lysine 205 form helix 4 and with serine 337 from helix 10. These two interactions are crucial for maintaining the stability of the AF-2 helix and towards its constitutive

activation. Finally, the last feature involved in contributing to the constitutive activity of CAR is that in CAR helix 2 packs tightly against helix 3, causing stabilizing the N-terminal portion of helix 3, placing AF-2 in the active conformation [43,132-134].

The high basal activity of CAR is further enhanced by the ligand, TCPOBOP, stabilizing the AF-2 helix and enhancing CAR transcriptional activity. The constitutive activation of CAR can also be repressed by androstanol. This compound has been shown to be an antagonist for CAR, fitting deeply into the CAR pocket [43,132-134]. In comparison to the agonist TCPOBOP, androstanol can fit into the pocket shape of TCPOBOP but leaves a void between androstanol and helix 10. This void causes the repositioning of the linker helix, placing the AF-2 helix in an inactive conformation, not providing the proper interface for coactivator binding [43,132-134].

With its unique structural and functional properties, we decided to test CAR in our chemical complementation system. The human CAR gene was amplified from a liver cDNA library with primers and ligated into a pGBD vector, creating a pGBDhCAR plasmid. Possible pGBDhCAR candidates were confirmed through sequencing. The pGBDhCAR plasmid was transformed into yeast along with a plasmid containing the Gal4 activation domain fused to the SRC-1 coactivator (pGAD10BASRC-1). Yeast cells were selected for on media lacking leucine and tryptophan, and cells that grew on that media were then subjected to the quantitation assay. Cells containing both pGBDhCAR and pGAD10BASRC-1 were inoculated in selective media lacking leucine and tryptophan overnight and then placed into 96-well microtiter plates containing histidine selective media. Since CAR is known to be constitutively active in mammalian cells, we hypothesized that this receptor would also be constitutively active in yeast, and wanted to

test that. We decided to use varying concentrations of 3-AT in the histidine selective media to determine whether we can reduce the constitutively activation of CAR.

Yeast cell with both CAR and SRC-1 and a plasmid containing the *holo* Gal4 as a control (pGBT9Gal4) were grown in histidine selective media with varying concentrations of 3-AT (Figure 7.6). As expected, CAR is constitutively active, where without the presence of ligand, the yeast cells are growing. The amount of growth is comparable to Gal4 growth; for example at with 10mM 3-AT, the cells containing Gal4 are growing to an OD of 0.36 and the cells containing both CAR and SRC-1, are growing to an OD of 0.3. As the 3-AT concentration increases, less growth is observed with both sets of cells, as expected (Figure 7.6). This shows that there is a correlation between mammalian cell assays and chemical complementation in yeast. There also may be an endogenous ligand binding to CAR, which enhances its transcriptional activation.

Androstanol was shown to be an antagonist for CAR and we decided to determine whether the addition of androstanol to the chemical complementation assay would repress the constitutive activation of CAR [134]. We also decided to evaluate a series of androstanol derivatives: androsterone, epiandrosterone, and 5-α, 3-androstane, 17-diol (5 alpha). Since the endogenous ligand binding to CAR was unknown, a true antagonist assay could not be performed with androstanol. Therefore, each of the androstanol derivatives was added at a concentration of 10uM. Cells containing pGBDhCAR and pGAD10BASRC-1 along with Gal4 (pGBT9Gal4) were grown in histidine selective media (-His -Trp) with varying concentrations of 3-AT. Each of the candidate antagonists were added at a concentration of 10 μM and grown for 48 hours. If any of these ligands were actual antagonists for CAR, the growth of the cells containing the

small molecule would be shifted lower in comparison to the growth of the cells without ligand. This is because the antagonist would be displacing the ligand that is activating transcription, and actually turning transcription off. We would expect to see this effect with androstanol, since this small molecule has been shown to be an antagonist for CAR.

Each of the four derivatives that were tested did not kill the yeast, as the cells with Gal4 seem to be growing regularly; less growth is observed as the 3-AT concentration decreases (Figure 7.7). In the case with androstanol, there was not a decrease in growth in comparison to the cells with no ligand, which is not what we expected to see. Androsterone and 5-alpha also showed no decrease in growth in comparison to the curve without ligand. However, one interesting feature in this assay with CAR and SRC-1 is that in the presence of epiandrosterone, a slight shift of the growth is seen in comparison to the growth curve without ligand. This slight shift was observed again when the assay was repeated (Figure 7.7).

Although, we expected an antagonistic effect with androstanol, we were not able to detect any growth difference in comparison to the no ligand curves, but were able to see a difference with epiandrosterone. We hypothesize that this could be due to the presence of a different agonist in the binding pocket endogenous in yeast, which may not be displaced by androstanol. This agonist may be more potent than known CAR agonists, thus androstanol may not be efficient in displacing this ligand. On the other hand, epiandrosterone may be a more potent antagonist of CAR, being able to displace the endogenous ligand binding in the pocket, causing the repression of transcription.

Rat Estrogen Receptor β (rERβ)

The human estrogen receptor was one of the first nuclear receptors to be cloned by Pierre Chambon and co-workers [138-140]. For about 10 years, the idea was hat only one isoform of the estrogen receptor existed (ERα), but Jan-Ake Gustafsson and co-workers reported the β isoform of this receptor [141]. The DNA binding domain (DBD) of both isoforms are highly homologous and the region, which is critical for target-DNA recognition and specificity, is identical between the two isoforms. In regards to the ligand-binding domain (LBD), there is a high degree of similarity between the two isoforms, and in most cases, these receptors are able to bind to the same ligands. A number of agonists exist for both isoforms, such as 17-β estradiol and hexestrol [120,142,143].

Both ERα and β interact with the p160/SRC coactivator family, which consists primarily of SRC-1, -2, and –3 [43]. Although both receptors can interact with most of the members of this family, there have been notable differences between the isoforms and coactivator binding. For example, ERβ has a lower affinity for SRC-3, and studies have shown that ERβ can act in a ligand-independent manner with SRC-1 and the short heterodimer partner (SHP) [43].

The first crystal structure of the ERα was solved with the receptor bound to estradiol [143-145]. The structure of ERα revealed that this receptor had a similar architecture and binding modes to the other nuclear receptors. Helix 12 is properly positioned to allow coactivator binding for the initiation of transcription. The crystal structure of ERα and its antagonist, raloxifene, was also solved,

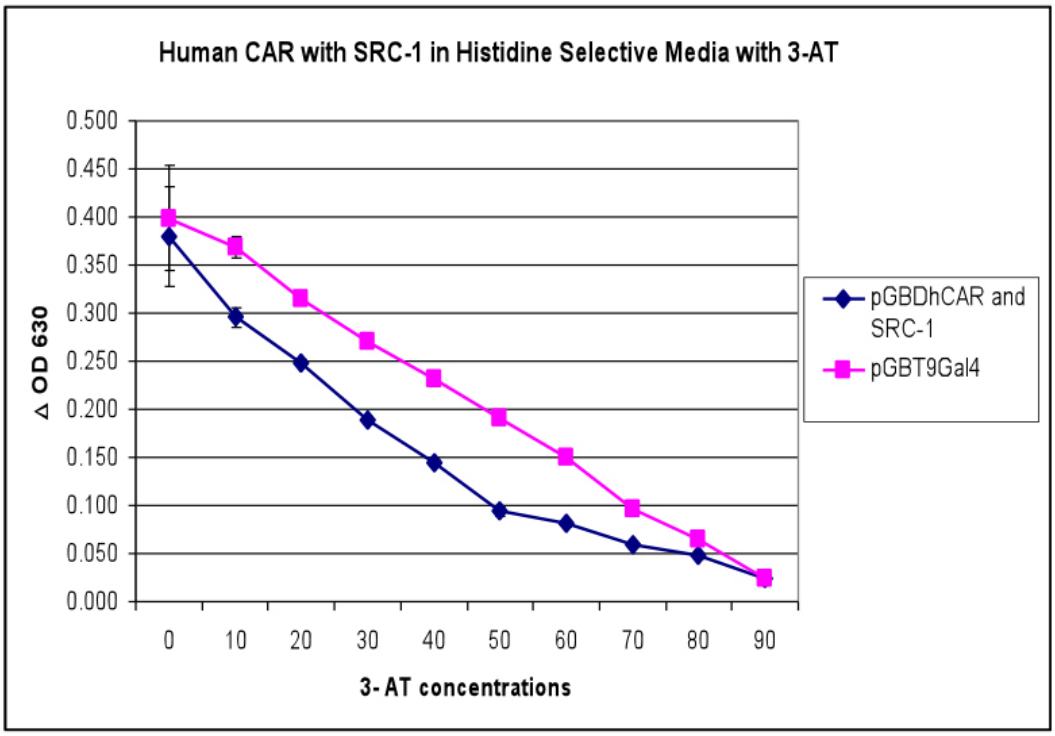


Figure 7.6. Chemical complementation with CAR and SRC-1 in histidine selective media with 3-AT.

showing a completely different conformation of the ligand-binding domain. In particular, helix 12 is inhibited from properly closing the hydrophobic pocket containing the ligand. This new conformation of helix 12 impairs coactivator recruitment [143-145]. The crystal structure of ERβ with its agonist and antagonist genistein has been solved [143,144]. An interesting feature of the ERβ bound to genistein is that this ligand is buried deep into the pocket, but helix 12 is unable to adopt the proper conformation, and adopts and orientation similar to antagonists defining genistein as a partial agonist for ERβ [143,144].

Estrogen receptors are associated with a number of disease, but these receptors are implicated mainly in breast cancer and osteoporosis as mentioned above. Besides their roles in several pathologies, estrogen receptors are involved in maintaining and exert positive effects in a number of physiological pathways, including the central nervous system cognitive function. The development of SERMs would greatly benefit the advancement of estrogen receptor related disease, targeting specific effects on the pathologies of interest. Discovering agonists or antagonists of these receptors would greatly benefit pathologies, such as breast cancer.

The rat ERβ was obtained from the lab of Jan-Ake Gustafson, and PCR was used to amplify the rat ERβ with primers containing restriction enzymes sites on either ends of the gene. This PCR product was cloned into a pGBD plasmid and with ligation and transformation into competent XL1-Blue cells, the pGBDratERβ plasmid was created. This plasmid was confirmed through

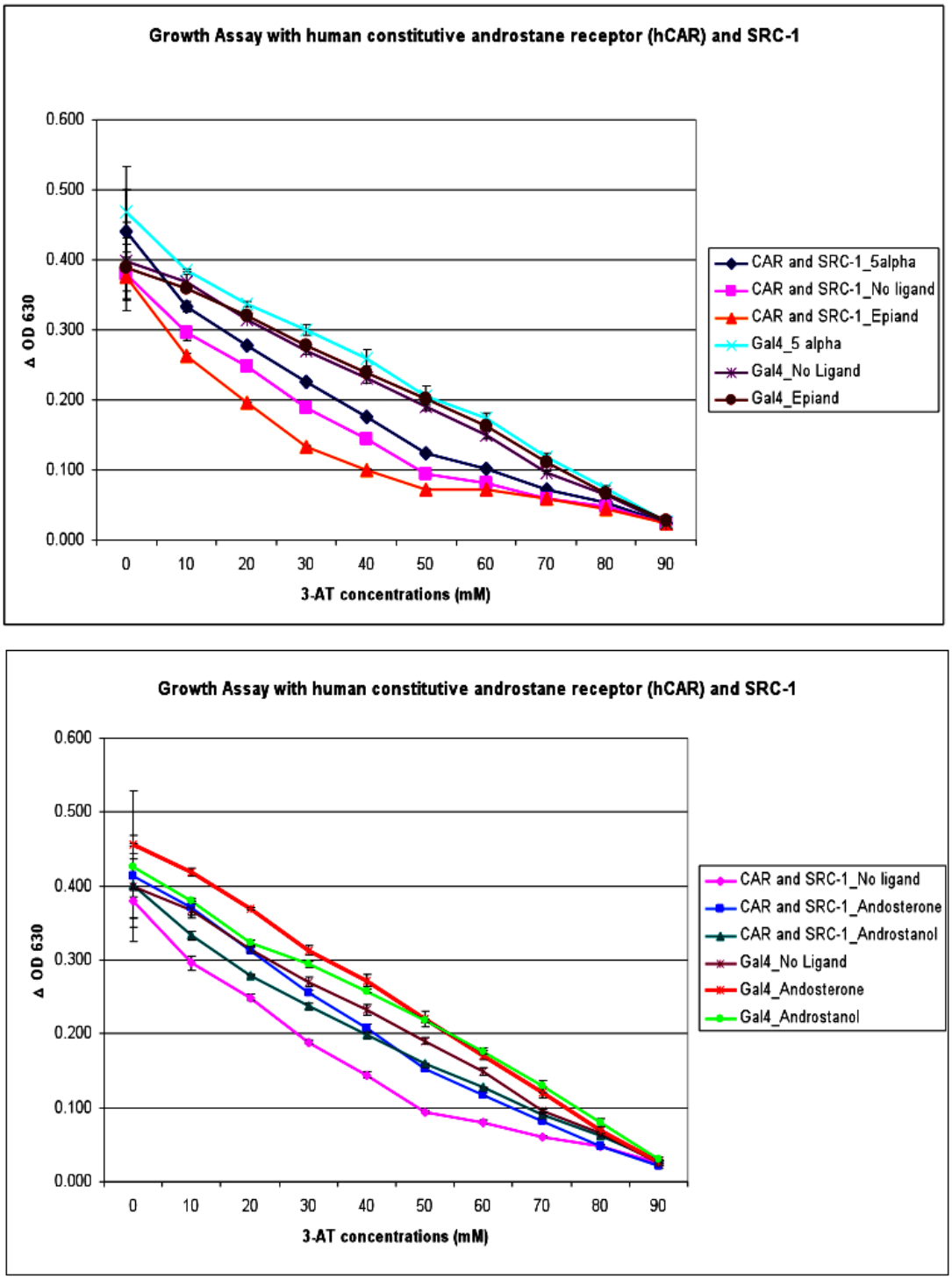


Figure 7.7. Constitutive androstane receptor and SRC-1 in chemical complementation with androstanol derivatives.

sequencing.

Once the pGBDratERβ plasmid was created, this plasmid was transformed into yeast with SRC-1 (pGAD10BASRC-1). Yeast transformants were plated onto leucine and tryptophan selective media. Transformants containing both plasmids were then tested for growth on adenine selective medium containing known agonists for ERβ: estradiol and hexestrol. As shown in Figure 7.8, growth was observed with ERβ and SRC-1 on the plates containing both ligands and no growth was observed on the plates without ligand.

Yeast cells containing both ERβ and SRC-1 were then inoculated and grown in leucine and tyrptophan selective media (-Leu-Trp) for setting up a quantitation assay. A plasmid containing the Gal4 *holo* protein was again used as a control. After an overnight growth period in the leucine and tyrptophan selective media, these cells were then added to a 96-well microtiter plate containing histidine selective media, with varying concentrations of 3-AT. 3-AT was added to eliminate the any growth seen with the ERβ due to the estrogenic effects of polystyrene. 3-AT concentrations from 10mM 3-AT were used to eliminate any of these effects. As the concentration of ligand increased, more growth was observed with ERβ and estradiol and hexestrol showed growth at 10^{-8} M concentrations of ligand as expected (Figure 7.9). We have been able to show that ERβ is functional in with chemical complementation and can be used for discovering possible new agonists for this receptor.

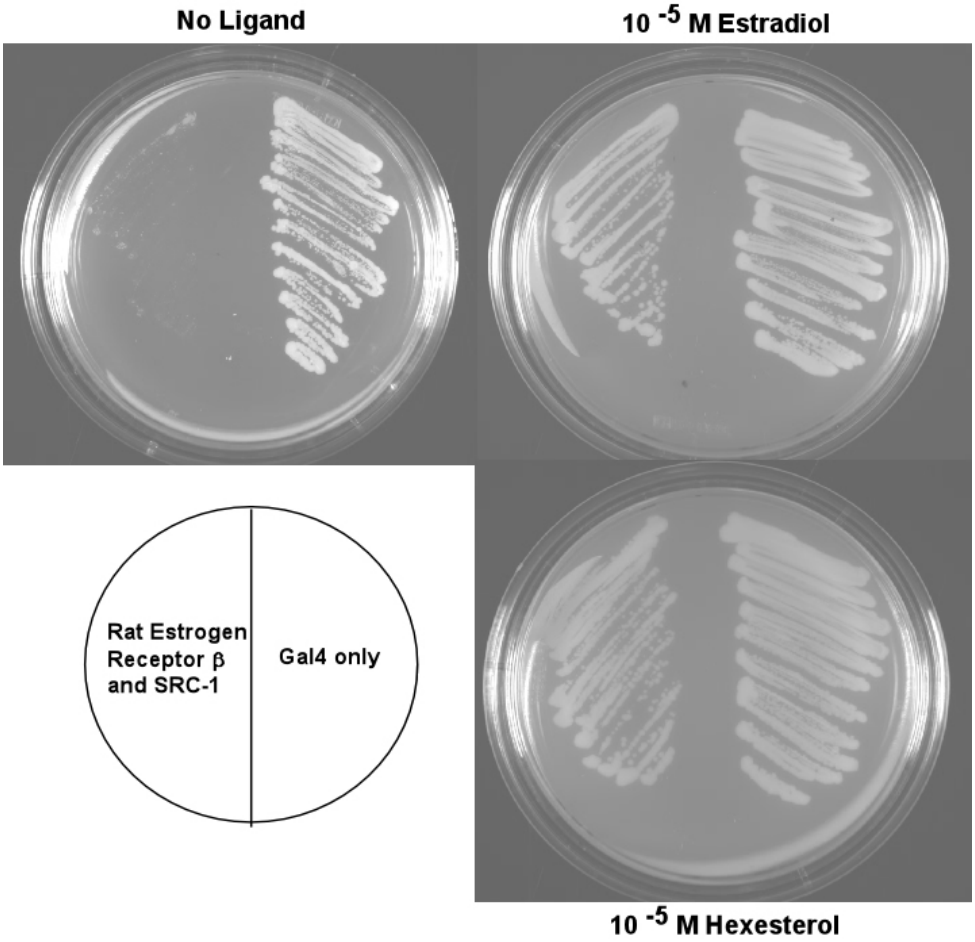


Figure 7.8. Rat ERβ and SRC-1 on adenine selective plates with and without ligand.

Human Farnesoid Receptor (FXR)

FXR is a member of the nuclear receptor superfamily involved in sensing bile acids [146,147]. This receptor is activated by various bile acids, such as chenodeoxycholic acid (CDCA) and lithocholic acid (LCA), and the genes regulated by FXR are involved in bile acid biosynthesis and recycling. For example, cholesterol 7α-hydroxlase (CYP7A1), a rate-limiting step of bile acid biosynthesis is controlled by a negative feedback loop [146,147]. The mechanism in CYP7A1 regulation involves collaboration between three different nuclear receptors. FXR is involved in bile acid biosynthesis and recycling involves two other nuclear receptors. When the bile acids bind and activate FXR, FXR induces the transcription of the orphan nuclear receptor, short heterodimer partner (SHP). This receptor forms a complex with another nuclear receptor, liver receptor homolog-1 (LRH-1), which is required for CYP7A1 expression, and the formation of the complex prohibits the function of LRH-1 [43]. Thus, as the concentration of SHP increases, the expression of the CYP7A1 gene decreases.

The crystal structure of FXR has not yet been solved. The FXR gene displays a 76% amino-acid identity with LXR and about a 44-55% identity to PXR and CAR [43,146,147]. In the ligand-binding domain, LXR has a 37% identity with FXR and both PXR and CAR have about a 27-30% amino acid identity. Like most of the other nuclear receptors, FXR forms a heterodimer with RXR, and SRC-1 was shown to interact with FXR in the presence of its ligand, CDCA[43,146,147].

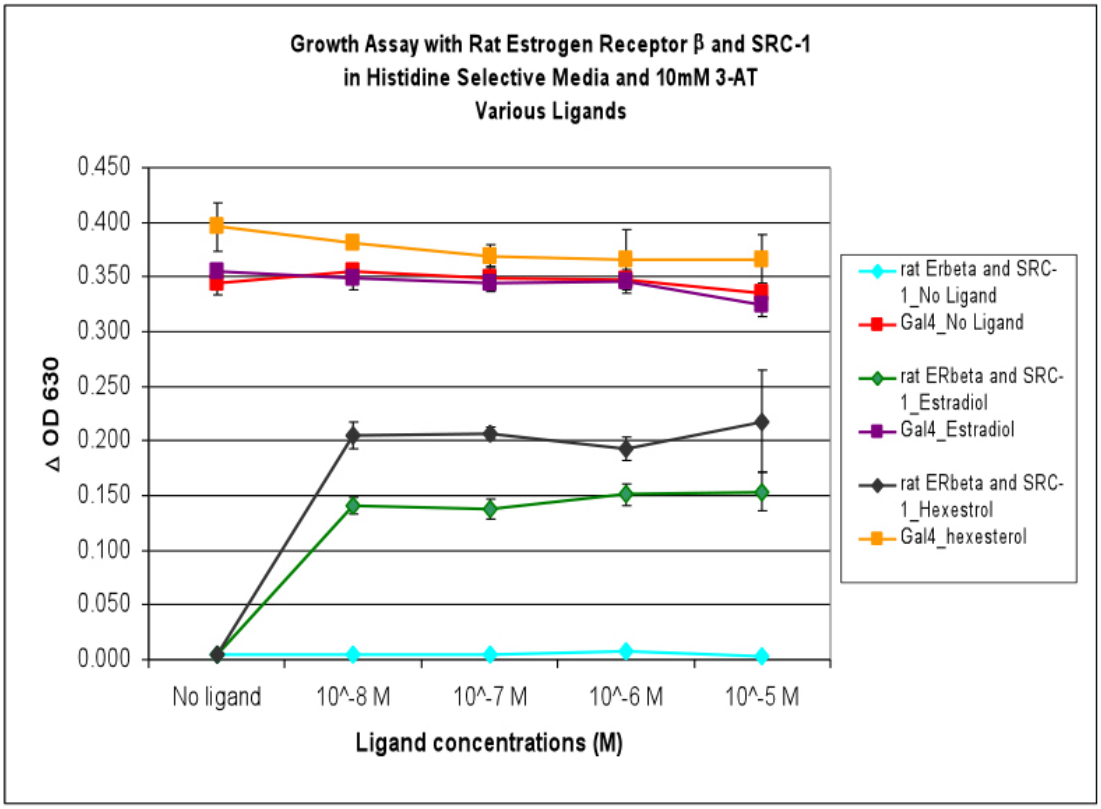
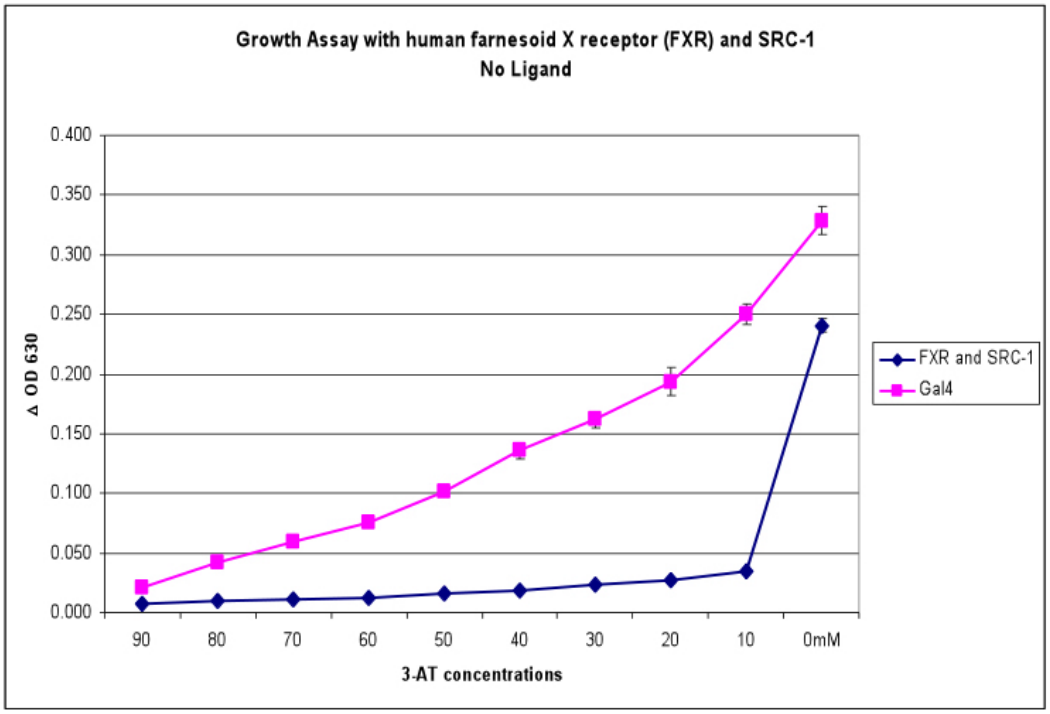


Figure 7.9. Rat ERβ and SRC-1 with estradiol and hexesterol.

To increase the versatility of chemical complementation, FXR was cloned from a liver cDNA library containing *NheI* and *SpeI* sites at either end of the FXR gene. The amplified PCR product was then digested with both of the enzymes creating the insert cassette. The pGBDRXRNS was also digested with *NheI* and *SpeI*, cutting out the entire RXR gene, creating the vector cassette. The vector and insert cassette were ligated together using T4 DNA ligase and transformed into competent XL1-Blue cells. The cells containing pGBDhFXR were confirmed through sequencing. This plasmid was then transformed into yeast with a vector containing the Gal4 activation domain fused to the SRC-1 gene (pGAD10BASRC-1). Yeast transformants were selected onto media lacking tryptophan and leucine, the selective markers of both expression vectors respectively.

Yeast transformants containing both of the vectors were grown in histidine selective media containing varying concentrations of 3-AT. To our surprise this receptor was constitutively active at 0 mM 3-AT, but the constitutive activation was dramatically repressed at 10-30 mM 3-AT. This was very different from what we have observed using AR and CAR. We wanted to determine whether we would be able to observe ligand-activated growth after the basal constitutive activation was decreased using 3-AT. A recent study in mammalian cell assays had shown that the portent LXR agonist, T0901317, was also an agonist for FXR. Thus, T0901317 was added at a concentration of 10uM to the histidine selective media with varying concentrations of 3-AT.



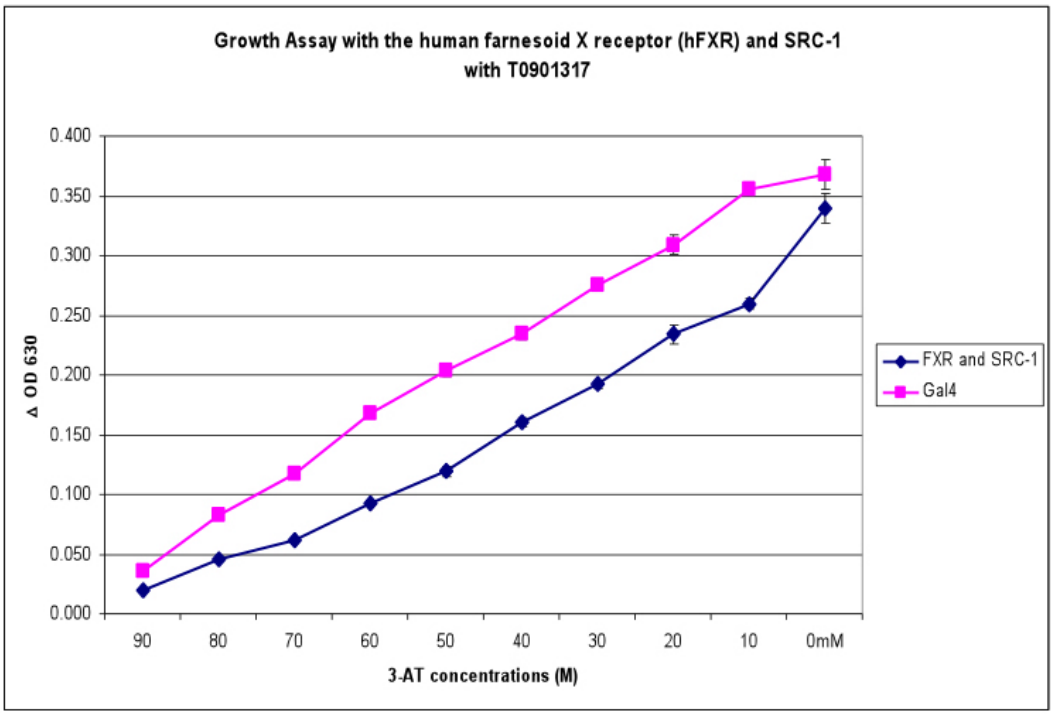


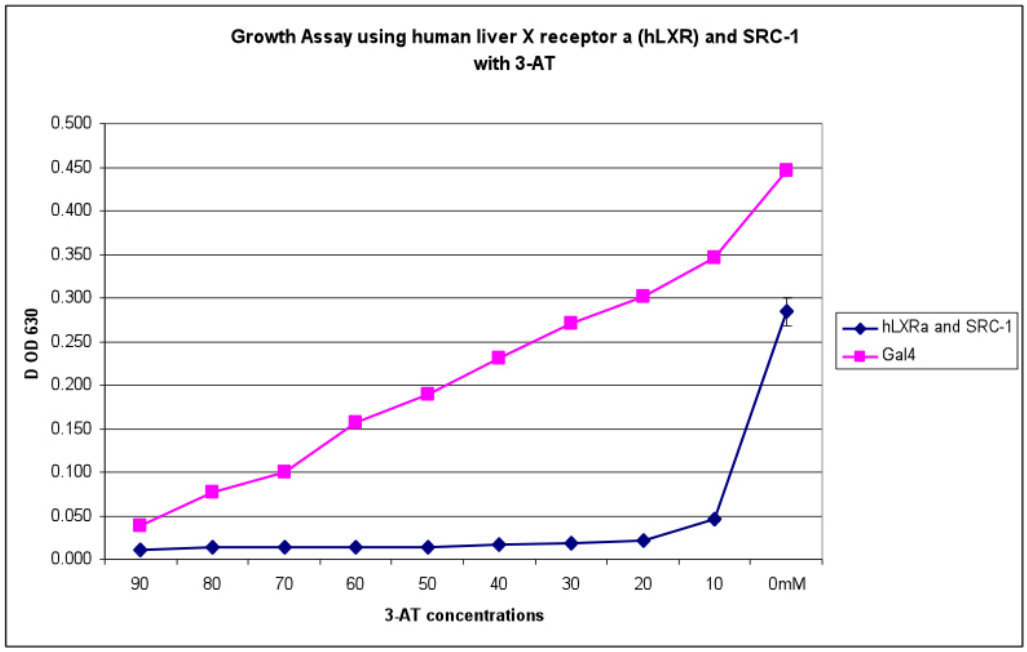
Figure 7.10. FXR and T0901317 with chemical complementation.

In the presence of 10 μM T0901317 and 10 mM 3-AT we were able to observe an increase the growth of the cells. As shown in Figure 7.10, without the presence of a ligand, at 10 mM 3-AT, the growth measured is about 0.04 OD_{630}, whereas in the presence of 10 μM T0901317 and 10 mM 3-AT, the amount of growth measured is at an OD_{630} of 0.250. Therefore, we are able to use this receptor in chemical complementation once the background constitutive activation is reduced with 10 mM 3-AT.

Human Liver X Receptor α (LXRα)

The LXRα is a nuclear receptor, which plays a crucial part in the maintenance of cholesterol homeostasis, where genes that are activated by this receptor are involved in cholesterol efflux from peripheral tissues [96,97,100,101]. There are two main isotypes of this receptor that have been identified: LXRα and LXRβ [103]. LXRβ is ubiquitously expressed in mammalian cells, where as LXRα is more confined to certain tissues, such as the liver, small intestine and adrenal gland. Both of these receptors heterodimerize with RXR and upon ligand binding are able to recruit SRC-1 as their primary coactivator target. LXRs are able to bind a diverse natural ligands thought to be oxidized derivatives of cholesterol, such as 24(*S*), 25-epoxycholesterol and 27-hydroxycholesterol. A potent synthetic ligand of LXR, known as T0901317, was also identified, and is more frequently used in studies using LXR [43,97].

LXRα was cloned from a liver cDNA library containing a *NheI* and *SpeI* site at either end of the gene for evaluation using chemical complementation. The amplified PCR product was then digested with both of the enzymes creating the insert cassette. The



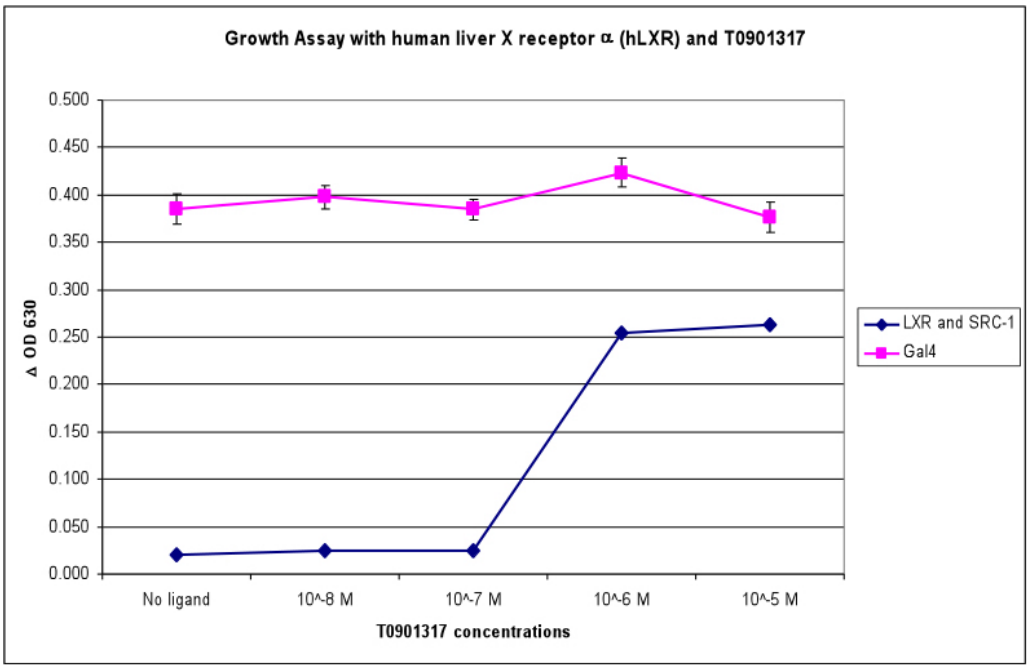


Figure 7.11. LXR and T0901317 with chemical complementation.

pGBDRXRNS was also digested with *NheI* and *SpeI*, cutting out the entire RXR gene, creating the vector cassette. pGBDhLXRα transformants were confirmed sequencing. This plasmid was then transformed into yeast with a vector containing the Gal4 activation domain fused to the SRC-1 gene (pGAD10BASRC-1). Yeast transformants were selected onto media lacking tryptophan and leucine, the selective markers of both expression vectors respectively.

Yeast transformants containing both of the vectors were grown in histidine selective media containing varying concentrations of 3-AT. The activation profiles seen with LXR are similar to FXR, where LXRα was constitutively active at 0 mM 3-AT, but the constitutive activation was dramatically repressed at 10 mM 3-AT. Using 10mM 3-AT, a dose response with T0901317 was performed, where varying concentrations of T0901317 were used. The concentrations of ligand ranged from 10-8 M to 10-5 M. Without the presence of ligand at 10mM 3-AT, the yeast cells containing LXRα and SRC-1 were not growing. As the concentration of T0901317 increased, growth was observed at about 500nM (5×10^{-7} M). As the concentration of ligand increased more growth was observed as expected (Figure 7.11). The same trend was also observed on solid media plates.

While testing CAR with the androstanol derivatives was ongoing, we decided to determine whether any of these small molecule ligands were also antagonists of LXRα. To create an effective antagonist assay, the EC_{50} value of the agonist should be used, such that the antagonist can displace the agonist, and repressing transcription. Based on the dose-response curve using T0901317, we were able to see that the EC_{50} values for T0901317 using chemical complementation is between 1 and 5 μM; we settled for an

EC_{50} value of 3uM. The androstanol derivatives were all analyzed a concentration of 10 μM.

Cells containing LXRα and SRC-1, as well as Gal4 as a control, were grown in leucine and tryptophan selective media selecting for the two expression plasmids. These cells were then placed in 96-well microtiter plates containing histidine selective media (-His -Leu -Trp) with 10mM 3-AT and 3 μM T0901317. Varying concentrations of each of the androstanol derivatives were added to the wells, ranging from 10^{-5} M to 10^{-8} M. The cells were grown for about 48 hours and the cell density was compared to the initial cell density in each well.

Some interesting results were obtained with certain androstanol derivatives (Figure 7.12). Without the presence of any androstanol derivative, the cells are able tot grow in the histidine selective media as expected. If androstanol or one of its derivatives actually has the ability to antagonize LXRα, a drop in the growth will be observed. In looking at Figure 7.12, the cells with androstanol are growing identical in comparison to the cells without any of the derivatives present. However, in the presence of the other compounds: androsterone, epiandrosterone, and 5-alpha, there is a decrease in the growth and that decrease is observed at concentration of about 5 μM with all three ligands (7.12). This decrease in growth is only seen with the yeast cells containing LXRα and not with Gal4, therefore implying that the compounds are not having a overall general effect on the yeast organism. From the data, there seems to be an antagonistic effect present with each of the three compounds and LXRα, where these compounds are able to displace T0901317, and repress transcription. Thus, we have been able to show that T0901317 with chemical complementation displays ligand-activated growth with LXRα

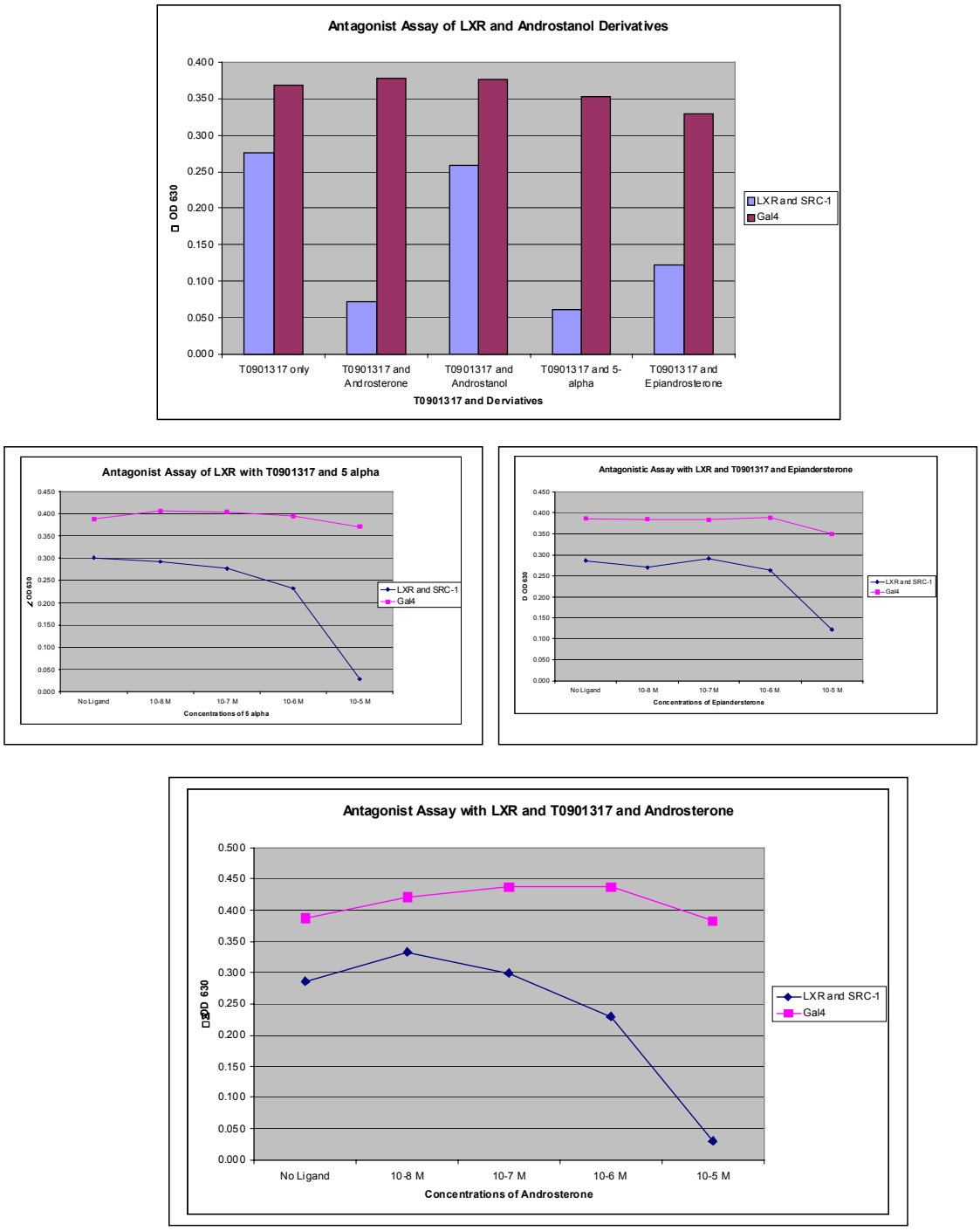


Figure 7.12. LXR and Androstanol derivatives.

as expected. Furthermore, we have also identified possible antagonists of LXRα using chemical complementation.

Mouse Peroxisome Proliferator Activated Receor γ (PPARγ)

The peroxisome proliferator activated receptors are members of the nuclear receptor family, which play a crucial role in adipogenesis and glucose homeostasis [43,121-123]. These nuclear receptors are involved in binding a variety of fatty acids and other metabolites. There are three isotypes of these receptors: PPARα, PPARδ, and PPARγ. Each of these nuclear receptors are involved in the regulation of cell differentiation, cell growth and homeostasis [43,121-123].

The best characterized PPAR isotype is PPARγ, a key regulator of transcriptional pathways involved in adipogenesis. PPARγ has been implicated in playing a major role in type II diabetes [43]. Synthetic PPARγ ligands, thiazolidinediones (TZDs) have been shown to reduce the insulin resistance and lower plasma glucose levels in type II diabetics. Due to its prominent role in this disease and several other diseases, discovering agonists and antagonists for these receptors are in demand [43,121-123] .

The structure of the PPARγ LBD has been solved displaying some of the features of the LBD of this nuclear receptor [43,148]. For the most part, the PPARγ LBD contains the common fold present in most of the other nuclear receptors. However, the crystal structure of these receptors also displayed some distinct features of this receptor. The PPARγ structure revealed that its ligand-binding pocket is much larger than some of the other traditional nuclear receptors, with a volume of $1300A^3$ [43,148]. The ligands that bind to the receptor only occupy about 40% of the actual binding pocket, whereas in

other nuclear receptors, about 90% of the pocket is occupied by the ligand. The binding pocket has either a T or Y shape, and also contains an additional helix, helix 2'. Helix 2 is also placed in an different orientation in comparison to other nuclear receptors. These structural features of PPARγ add to the uniqueness of this nuclear receptor and allow for the binding pocket to accommodate longer fatty acid molecules, as well as synthetic antidiabetic drugs [43,148].

The mouse PPARγ gene was amplified with PCR with primers containing an *EcoR1* and *AflII* site on either end of the gene. The amplified PCR product was digested with both *EcoRI* and *AflII* and ligated into a pGBD vector containing both of those restriction sites after the Gal4 DBD. The digested mouse PPARγ gene was then ligated into the vector and the ligations were transformed into competent XL1-blue cells. Transformants containing pGBDmPPARγ2 were confirmed through sequencing.

The pGBDmPPAR was transformed into yeast with its coactivator PGC-1a. A plasmid containing the Gal4 activation domain fused to the PGC1a gene was constructed (pGAD10BAmPGC1a) and transformed along with pGBDmPPARγ2 into yeast. Yeast transformants were plated onto leucine and tryptophan selective media. Cells containing both PPARV and PGC1a were then picked for the quantitation assay, where the cells were grown in histidine selective media with varying concentrations of 3-AT. Mouse PPARγ showed constitutive activity without the presence of ligand (Figure 7.13). The constitutive activation was also observed when the ACTR coactivator was used. The fact that PPARγ is constitutively active is not surprising, since fatty acids are present in yeast and this activity is probably due to an endogenous ligand present in the pocket.

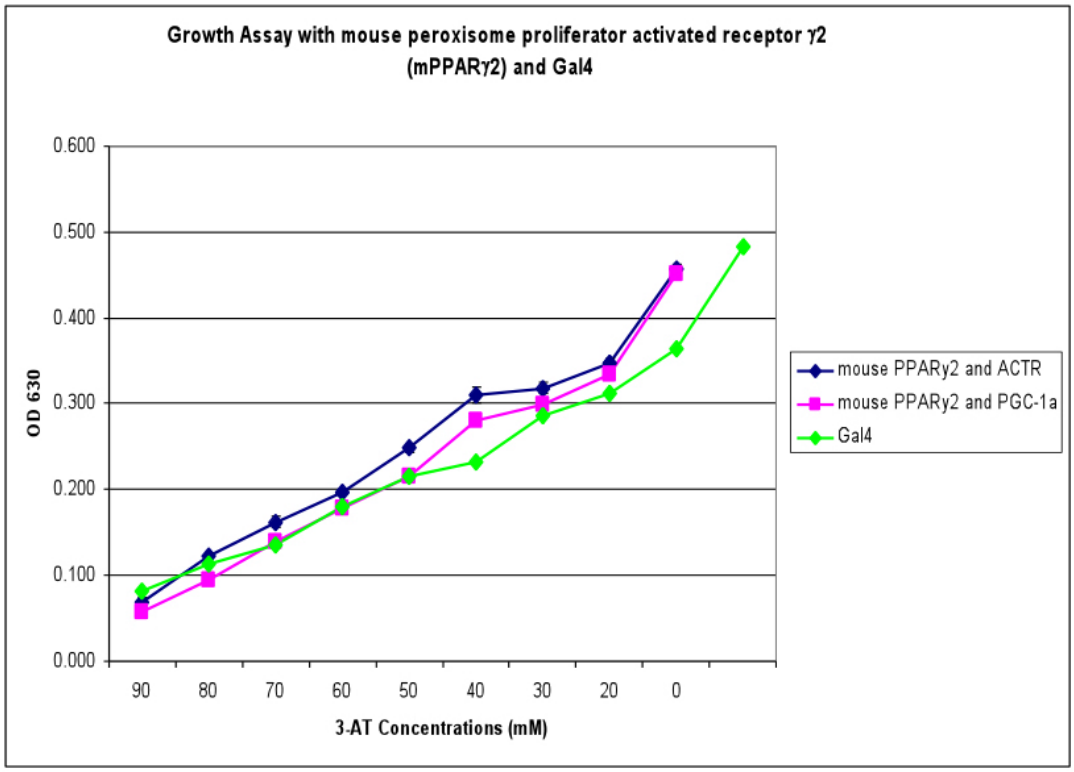


Figure 7.13. Mouse PPARγ2 and PGC-1a in chemical complementation with histidine selective media and 3-AT.

Human Pregnane X Receptor (PXR or SXR)

PXR is involved in a number of pathways, particularly towards the clearance of xenobiotics from the body [149,150]. A number of PXR's target genes involve the detoxifying enzymes of the cytochrome P450 family, such as CYP2A and CYP3A [149,150]. PXR seems to collaborate with CAR in the complete regulation of the cytochrome P450 enzymes, particularly CYP3A and CYP2B genes [149,150].

When the crystal structure of PXR was solved, a large promiscuous binding pocket was discovered for this nuclear receptor [151]. This is not surprising since PXR is able to bind some of the larger, bulkier compounds, such as rifampicin and paclitaxel (See Chapter 3 for more details). Initially, when PXR was tested in chemical complementation using tocopherol and paclitaxel, two of its known ligands, ligand-activated growth was observed in 1st generation chemical complementation. However, in 2nd generation chemical complementation PXR becomes constitutively active, and even varying 3-AT concentrations cannot eliminate the basal activation (Figure 7.14). The fact that this receptor exhibits such a profile is not uprising, since this receptor is known for its promiscuity and is able to bind a diverse set of ligand, and plenty of yeast metabolites would be able to bind and activate the receptor [151]. Yet, we were hoping that the basal constitutively activation would be reduced with the higher concentrations of 3-AT.

The constitutive activation observed with PXR, CAR and AR were all compared with the removal of the coactivator from chemical complementation (i.e. 1st generation chemical complementation). Without the presence of the coactivator, AR was still quite active without the presence of a ligand or the coactivator fusion protein (Figure 7.15).

On the other hand, the constitutive activity observed with both CAR and PXR was eliminated; basal activity was only observed at 0 mM 3-AT (Figure 7.15). These results imply that the endogenous agonist that binds to AR induces a conformational change that is optimal for both the yeast and mammalian coactivator proteins. In the case with both CAR and PXR, the conformational change induced by the binding of an endogenous ligand, is only suitable for binding human coactivators and not yeast coactivators. Thus, this further supports the idea that although yeast and mammalian transcriptional machinery is quite similar, these two systems have evolved to contain a number of differences and with unique specificities and characteristics, allowing for the differences between these two systems.

Bovine Steroidgenic Factor-1 (bSF-1)

The steroidgenic factor is a nuclear receptor from the 5A subfamily and these receptors are also involved in multiple physiological pathways, such as in male sexual development and endocrine metabolism [152,153]. These nuclear receptors have for the most part been identified as being constitutively active and act as monomers, rather than the traditional dimers. SF-1 and other members of this subfamily are characterized by the unusually large hinge domain, which is subjected to posttranslational modifications [43,154,155].

Ligands for members of this subfamily, including SF-1, were sought after. Due to the role of SF-1 in steroid synthesis, speculations were made on whether oxysterols which activate LXR, could also be ligands for SF-1 [156,157]. However, research was performed in different cells lines showing that oxysterols are not ligands for SF-1. More

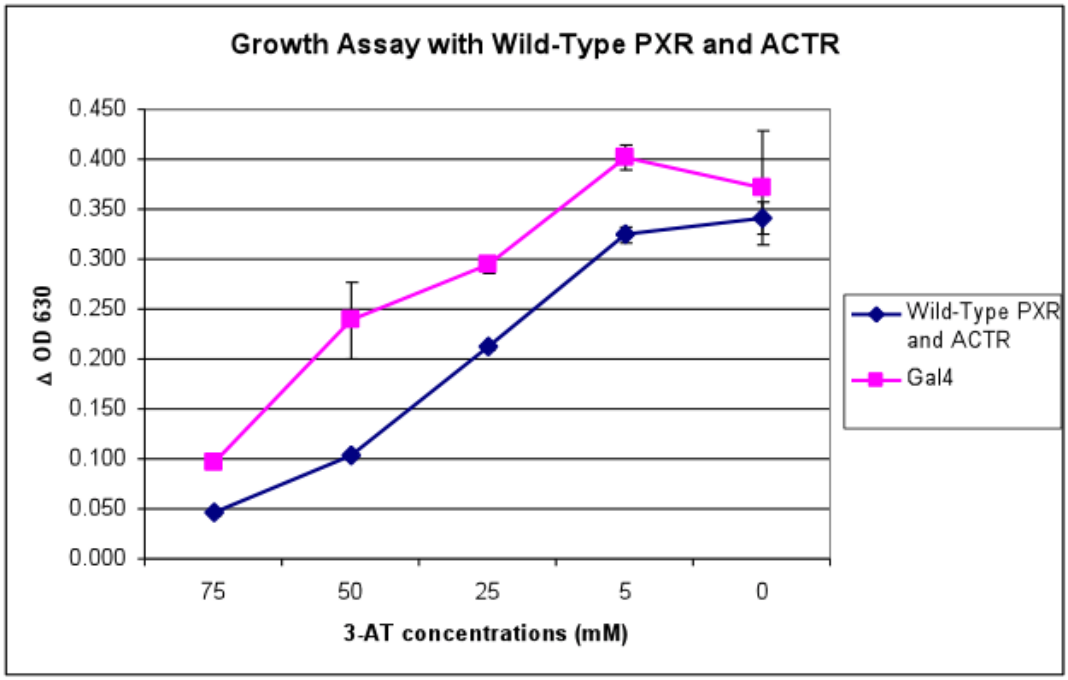


Figure 7.14. PXR and ACTR in Chemical Complementation.

recently, the crystal structure of SF-1 was determined and the structure revealed the presence of phosphitadyl insotiol as the ligand for SF-1. Further testing revealed that SF-1 did indeed bind these ligands, and maximal activity of this receptor was achieved with ligand binding [153,158]. The presence of the phosphotidyl inositol in the binding pocket of SF-1 put forth into perspective the role of some of the other members of this subfamily as phospholipid sensors [153,158].

With the discovery of a phospholipid as a ligand for the SF-1, we decided to clone SF-1 into a yeast expression plasmid and evaluate its activity with chemical complementation. The bovine SF-1 gene was amplified with PCR products from a plasmid containing the gene. The PCR product contained restriction enzyme sites, *EcoRI* and *AflII*, which were also present in the vector cassette. The amplified PCR product was digested with both *EcoRI* and *AflII* and ligated into a pGBD vector containing both of those restriction sites after the Gal4 DBD. The digested bovine SF-1 gene was then ligated into the vector and the ligations were transformed into competent XL1-blue cells. Transformants containing pGBDbSF-1 were confirmed through sequencing.

This plasmid was then transformed into yeast with a vector containing the Gal4 activation domain fused to the SRC-1 gene (pGAD10BASRC-1). SRC-1 had previously been shown to be a coactivator for SF-1. Yeast transformants were selected onto media lacking tryptophan and leucine, the selective markers of both expression vectors respectively. Yeast transformants containing both of the vectors were grown in histidine selective media containing varying concentrations of 3-AT. To our surprise this receptor was constitutively active at 0 mM 3-AT, but the constitutive activation was dramatically repressed at 50 mM 3-AT. This was not surprising because of the presence of

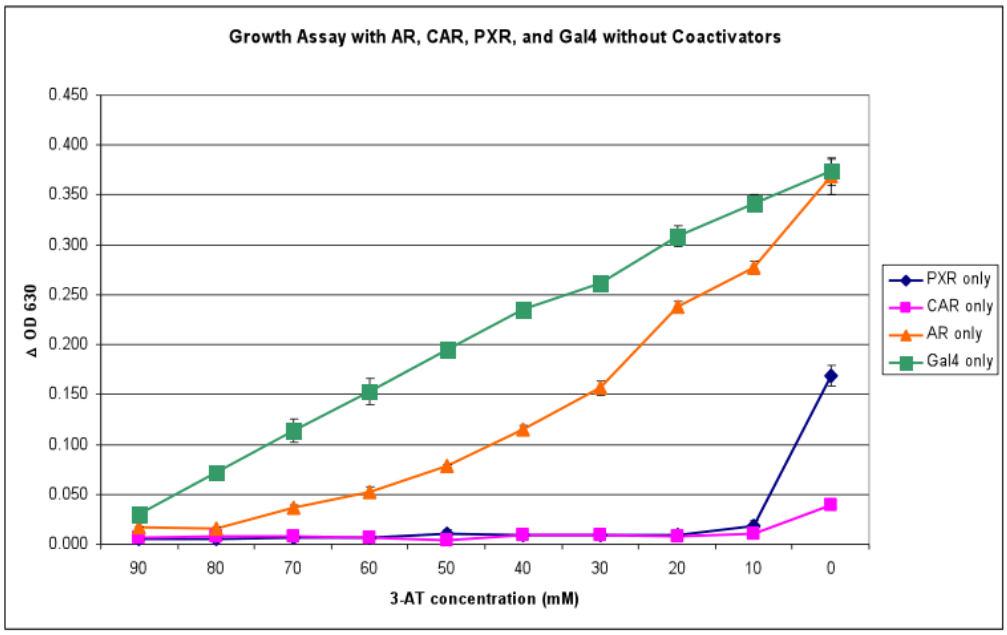
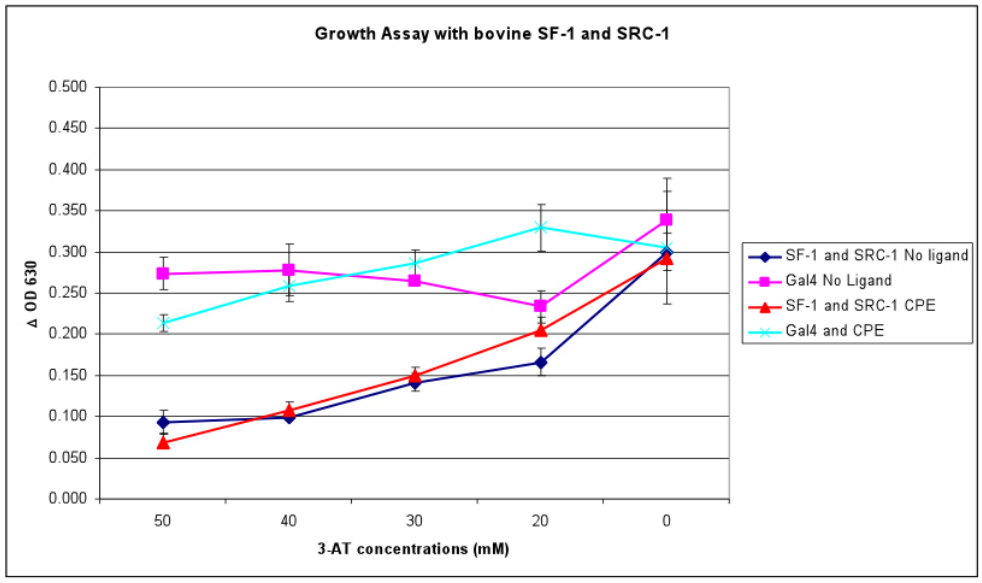


Figure 7.15. 1st Generation Chemical Complementation with PXR, CAR and AR



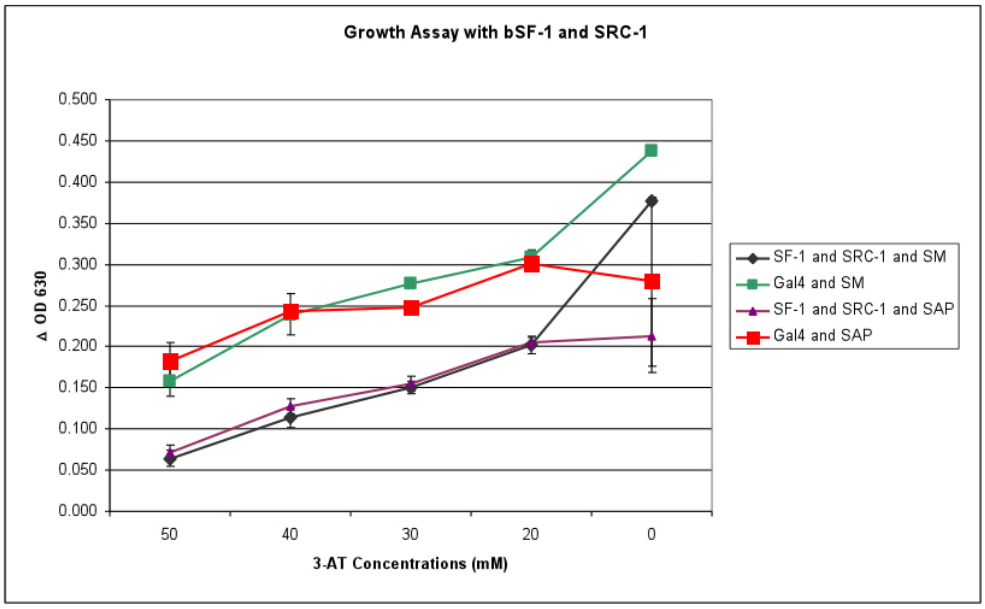


Figure 7.16. SF-1 and SRC-1 in chemical complementation with various sphingolipids.

endogenous phospholipids that could be activating the receptor. Once the constitutive activity was reduced with 50 mM 3-AT, as series of sphingolipids were tested to see whether ligand-activated growth could be observed above background. No ligand-dependent activity was detected above background (Figure 7.16). This indicated that either the ligands were no agonists for SF-1 or that the phospholipids were not able to penetrate the cell membrane.

7.6 Summary of Drug Discovery

Chemical complementation has been extended to a variety of nuclear receptors, ranging from steroid receptors, such as the androgen and estrogen receptor, to non-steroidal receptors, such as CAR and SF-1. Most of the nuclear receptors are constitutively active, in which we hypothesize the presence of an endogenous small molecule ligand binds to these receptors. In some instances, the constitutive activation can be reduced with 3-AT, as in the cases of LXR and FXR, and ligand-activated growth was detected once the constitutive activation was reduced. Most of the nuclear receptors have similar activity between yeast and mammalian cell assays, indicating that we have indeed created a system, which can be utilized for discovering agonists and antagonists of nuclear receptors in mammalian cells. Furthermore, these ligands could also serve as possible therapeutics.

7.7 Future Work

Chemical complementation can be used as a high-throughput assay for discovering agonists of nuclear receptors, as well as using negative chemical

complementation for discovering antagonists. Both of these assays utilize the power of genetic selection to eliminate wasted effort and time towards ligands that are not agonists or antagonists of interest. Combinatorial libraries of different compounds can be analyzed for agonist or antagonist activity simply by the addition of yeast cells containing the nuclear receptor and coactivator expression plasmids will be added to each of the wells. This will help enhance the discovery and further development of nuclear receptor ligands, which could in turn become a therapeutic drug.

The majority of the nuclear receptors assayed through chemical complementation in yeast were constitutively active. Mass spectrometry analysis could be performed to determine the presence of a ligand. These receptors could also be crystallized and structures for these receptors could be determined. The crystal structure would allow the determination of the endogenous ligand present in the binding pocket, or if structural modifications in yeast change the fold of these receptors, causing these receptors to be constitutively active. Overall, these structures could provide further insight into the structure and function of these nuclear receptors.

The androstanol derivatives, androsterone, epiandrosterone, and 5-α,3-androstane, 17-diol, show some antagonistic activity with LXR and CAR. The next step would be to evaluate these receptors and ligands in mammalian cell assays to determine whether these ligands are indeed antagonists.

7.8 Materials and Methods

Ligands

9-*cis* retinoic acid (MW=300.44 g/mol), β-estradiol (MW= 272.4 g/mol) and

hexestrol (MW= 270.4 g/mol) were purchased from MP Biomedicals (Aurora, OH). Paclitaxel (Taxol®) (MW=853.9 g/mol) was purchased from LKT Laboratories (St.Paul, MN). Resveratrol (MW= 228 g/mol), epiandrosterone (MW= 290.4 g/mol), androsterone (MW= 290.4 g/mol), and 5α-androstane-3, 17 dione (MW= 288.4 g/mol) were purchased from Sigma (St. Louis, MO). Androstanol (MW= 276.5 g/mol) was purchased from BIOMOL (Plymouth Meeting, PA). Rosglitazone (MW= 357.4 g/mol) and T0901317 (481.3 g/mol) were purchased from Cayman Chemicals (Ann Arbor, MI). 10 mM stocks of each ligand were dissolved in 80% ethanol: 20% DMSO (4:1v/v) and stored at 4°C.

Human Androgen Receptor.

A mammalian vector containing the androgen receptor gene (pCMVhAR) was obtained from Elizabeth Wilson, UNC Chapel Hill. The androgen receptor was amplified from this plasmid using primers with *NheI* and *SpeI* sites at the ends of the primers. A derivative of the pGBDRXR plasmid was made with a *NheI* and *SpeI* site inserted into the plasmid. The *NheI* site was inserted after the Gal4 DNA binding domain (Gal4 DBD) and the *SpeI* was inserted into the 3' UTR (untranslated region). This plasmid was named pGBDRXRNS. The pGBDRXRNS plasmid (2 μg) was digested with *NheI* and *SpeI*, along with the PCR amplified AR gene (150 ng). The digested pGBDRXRNS plasmid was treated with 1 μL of calf intestinal phosphatase (CIP). These digests were cleaned using the Gel DNA Recovery Kit (Zymo Research, Orange, CA), and ligated using the Quick Ligation Kit (New England Biolabs, MA). The ligations were incubated overnight at 22 °C and transformed into Z-competent XL1-Blue cells (Zymo Research, Orange, CA). Transformants were confirmed through restriction enzyme analysis and sequencing. The plasmid was name pGBDhAR.

Once pGBDhAR had been made, this plasmid was transformed into PJ69-4A with pGAD10BASRC-1 using the standard LiAc transformation protocol. Transformants were plated on media lacking leucine and tryptophan. The cells that were able to grow on this media were then inoculated in selective media lacking leucine and tryptophan and subjected to the quantitation protocol in Chapter 5.

CAR

The human CAR gene was amplified from a liver cDNA library (Clontech, CA) primers containing *EcoRI* and *SpeI* sites on either ends of the gene. The pGBDRXRSpeI plasmid (contains a *SpeI* site at the 3' UTR) was digested with *EcoRI* and *SpeI*, treated with CIP, and cleaned with the Gel DNA Recovery Kit (Zymo Research, Orange, CA). The CAR PCR fragment was also digested with the same restriction enzymes and cleaned with the same kit. These pieces of DNA are then ligated together and transformed into Z-competent XL1-Blue cells (Zymo Research, Orange, CA). Transformants were sequenced for confirmation. The plasmid is called pGBDhCAR.

Once pGBDhCAR had been made, this plasmid was transformed into PJ69-4A with pGAD10BASRC-1 using the standard LiAc transformation protocol. Transformants were plated on media lacking leucine and tryptophan. The cells that were able to grow on this media were then inoculated in selective media lacking leucine and tryptophan and subjected to the quantitation protocol in Chapter 5.

Rat Estrogen Receptor β.

The rat ERβ was obtained in a mammalian vector from Jan-Ake Gustaffson's lab and the rat ERβ was amplified with primers containing *NheI* and *AflII* sites on either ends of the gene. The PCR fragment was digested with both restriction enzymes and cleaned using the DNA Clean and Concentrator Kit (Zymo Research, Orange, CA). pGBDRXRNA is a vector that contains a NheI site after the Gal 4 DNA binding domain and an *AflII* site in the 3- UTR. This vector was digested with *NheI* and *AflIII* and treated with CIP (described above). This vector cassette was ligated with the rat ERβ and transformed into Z-competent XL1-Blue (Zymo Research, Orange, CA). pGBDratERβ were confirmed through restriction enzyme analysis and sequenced.

pGBDratERβ were transformed with pGAD10BASRC-1 into yeast using the traditional LiAc yeast transformation method. The transformants were plated on selective media lacking leucine and tryptophan. Cells that grew on this media were picked for streaking onto ligand plates. Adenine selective plates (-Ade -Trp) were made with 10^{-5} M estradiol and hexestrol, and without ligand. Cells containing both pGBDratERβ and pGAD10BASRC-1 and cells containing pGBT9Gal4 were streaked on these plates. These plates were incubated at 30 °C for two to three days. Cells containing both pGBDratERβ and pGAD10BASRC-1 were inoculated for the quantitation protocol.

FXR

FXR was cloned from a liver cDNA library containing a *NheI* and *SpeI* site at either end of the FXR gene. The amplified PCR product was then digested with both of the enzymes creating the insert cassette. The pGBDRXRNS was also digested with *NheI*

and *SpeI*, cutting out the entire RXR gene, creating the vector cassette. The vector and insert cassette were ligated together using T4 DNA ligase and transformed into competent XL1-Blue cells. The cells containing pGBDhFXR were confirmed through sequencing. This plasmid was then transformed into yeast with a vector containing the Gal4 activation domain fused to the SRC-1 gene (pGAD10BASRC-1). Yeast transformants were selected onto media lacking tryptophan and leucine, the selective markers of both expression vectors respectively.

Human LXRα

LXR was cloned from a liver cDNA library containing a *NheI* and *SpeI* site at either end of the FXR gene. The amplified PCR product was then digested with both of the enzymes creating the insert cassette. The pGBDRXRNS was also digested with *NheI* and *SpeI*, cutting out the entire RXR gene, creating the vector cassette. The vector and insert cassette were ligated together using T4 DNA ligase and transformed into competent XL1-Blue cells. The cells containing pGBDhFXR were confirmed through sequencing. This plasmid was then transformed into yeast with a vector containing the Gal4 activation domain fused to the SRC-1 gene (pGAD10BASRC-1). Yeast transformants were selected onto media lacking tryptophan and leucine, the selective markers of both expression vectors respectively.

Mouse PPARγ

The mouse PPARγ was obtained from Bruce Speigelman's lab (Harvard University) in the pSVSPORT-PPARγ2. The mouse PPARγ was amplified with primers containing an *EcoRI* site and *AflII* sites at the end of the primers. The PCR product was

digest with the restriction enzymes and ligated into the pGBDRXRAflII; this plasmid contains an *AflII* site in the 3' UTR of the plasmid. The ligation was transformed into competent cells as above. The plasmid was called pGBDmPPARγ2.

The PGC1a gene was amplified from the pSVSPORTPGC1a plasmid obtained from Bruce Speigelman's lab. This gene was cloned into the pGAD10 plasmid (Clontech, CA) with the restriction enzymes, *XhoI* and *EagI*, ligated and transformed into competent XL1-Blue cells. Transformants were sequenced for confirmation, and this plasmid was named pGAD10BAPGC1a.

pGAD10BAPGC1a and pGBDmPPARγ2 were transformed into yeast using the standard LiAc transformation protocol. These cells were plated onto media lacking leucine and tryptophan and transformants containing both plasmids were selected for the quantitation protocol in Chapter 5.

PXR.

PXR was cloned as described in Chapter 3. This plasmid, pGBDPXR, was transformed into yeast with pGAD10BA ACTR using the standard LiAc transformation protocol mentioned above, and followed the quantitation protocol in Chapter 5.

Bovine SF-1

The SF-1 gene was obtained from Marion Sewer's lab (Georgia Institute of Technology) in the pET24 vector. The SF-1 gene was amplified through PCR using primers with *EcoRI* and *AflII* and cloned into the pGBD vector with *EcoRI* and *AflII* sites. The vector and insert cassettes were ligated and transformed into competent XL1-Blue

cells. Transformants were analyzed through restriction enzyme digests and sequenced. This plasmid was named pGBDbSF-1. pGBDbSF-1 and pGAD10BASRC-1 were transformed into yeast in the same procedure as above and subjected to a quantitation assay.

Chapter 8

Applications of Chemical Complementation : Protein Engineering

8.1 Chemical Complementation for Protein Engineering

Thus far, chemical complementation can serve as a high-throughput assay for discovering agonists and antagonists of nuclear receptors. These applications can be a powerful tool for drug discovering targeting nuclear receptors, helping advance the development and discovery of these drugs. These applications lead to further understanding of the function and structure of nuclear receptors. Another beneficial application of chemical complementation is for protein engineering. This system can be used to evaluate combinatorial libraries of engineered proteins for their function.

Over the past few years, advancements have been made in creating large combinatorial libraries of proteins with novel structures and functions. The key to a successful analysis of a large combinatorial protein library is the assay used to evaluate the members of the library. Most of the common assays developed and used for evaluating members of the library are screens, where each of the members is individually analyzed for their function. These assays are mostly based on a colorimetric or fluorometric change, or involve a reporter gene. Yet, this method despite its commonality is extremely time-consuming and not all of the members of the library assayed are of interest.

A more efficient assay for deciphering protein libraries would be the use of a genetic selection assay [1,7,159]. Genetic selection only allows the members of the library, which are of interest to survive, because in genetic selection assays protein

function is linked to the viability of the cell. Thus, only the members of the combinatorial protein library with the function of interest will allow the cells to survive. Genetic selection provides a powerful tool where only the desired variants are evaluated and assayed for their function. This assay enhances the ability to evaluate a protein library by eliminating variants that are not of interest, providing a more time and materials efficient method for fully evaluating a protein library.

Chemical complementation provides an ideal assay for analyzing and creating large protein libraries with small molecule interactions. This system takes the power of genetic selection and links the yeast survival to the presence of a small molecule ligand. Chemical complementation can be used for engineering nuclear receptors to bind a novel small molecule ligand, or this system can be tailored toward engineering enzymes for creating novel small molecules (See Chapter 9). The method in which chemical complementation is used for engineering nuclear receptors towards a target small molecule is shown in Figure 8.1. A desired small molecule is selected as the target compound. A nuclear receptor chosen as the candidate for protein engineering is subjected to a range of mutagenic protocols, creating a library of engineered receptors. These engineered receptors are fused to the Gal4 DNA binding domain (DBD) and transformed into yeast with the appropriate coactivator: Gal4 activation domain (AD) fusion. The yeast transformants are plated on selective media (i.e. adenine selective media) with a certain concentration of the target compound. Those variants able to bind and activate the receptor will express the *ADE2* gene (or some other selective gene), allowing the yeast cell to survive on the selective media. If the variant is unable to bind the target compound, the cells will not survive, thus with chemical

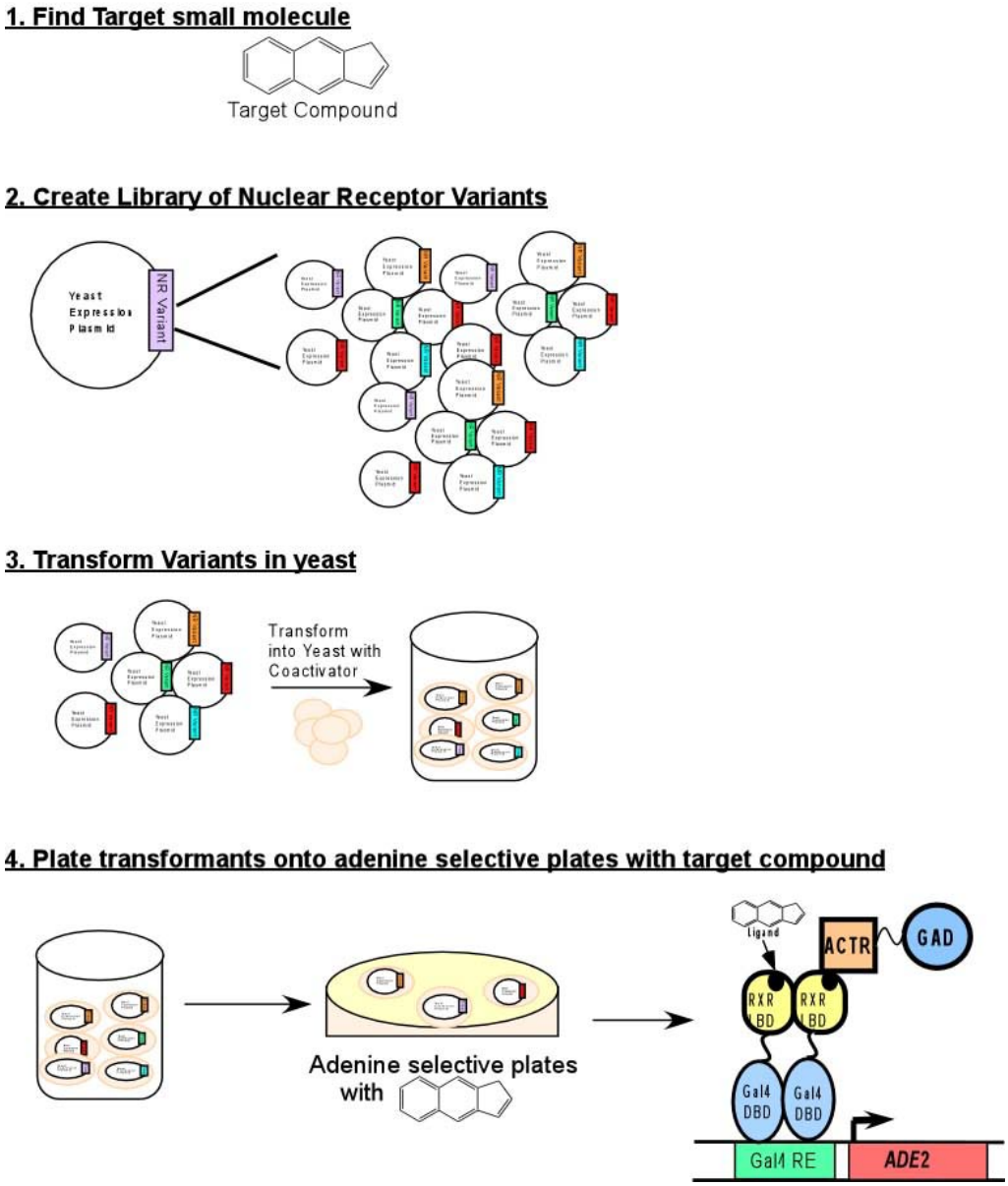


Figure 8.1. Chemical complementation for engineering nuclear receptors.

complementation, only the desired variants will survive. Although chemical complementation is advantageous in that only the variants with the desired function will survive, the survival of the yeast can also be due to the presence of an endogenous ligand binding and activating the receptor. Constitutively active receptors will be able to survive as well. To eliminate the possibility of constitutively active receptors or "false-positives", negative chemical complementation can be used in two ways. First, the target small molecule can be placed on uracil selective media with the desired small molecule, such that only if the receptor binds to the small molecule and activate transcription of the *URA3* gene will the yeast be able to survive (analogous to the *ADE2* or *HIS3* genes). Secondly, negative chemical complementation can also be applied towards these protein libraries by using the 5'-fluoroorotic acid (FOA) selective media and selecting for receptors that are not able to bind a certain small molecule. For example, if a library of nuclear receptor A variants is created to bind to the target compound X and not to the wild-type ligand, compound Y, the protein library will be put through two rounds of selection. In the first round, receptor A variants will be plated on FOA selective media with Compound Y (Figure 8.2). Variants that bind to the wild-type ligand, Compound Y, and those variants that are constitutively active will cause cell death. This is because these variants will be activated either by the wild-type ligand, compound Y, or by an endogenous ligand within the cells, activating the *URA3* gene. In media with FOA, the activation of the *URA3* gene will cause to conversion of FOA into 5'-fluorouracil, a toxin killing the yeast (Figure 8.2). Only those variants that are unable to bind to the wild-type ligand or an endogenous ligand will be able to survive, because they will not activate the *URA3* gene and 5-fluorouracil will not be produced. Those variants that survive the FOA

selection will then be put through chemical complementation with the target compound X in adenine, histidine, or uracil selective media (Figure 8.2). In this case, only those samples that are able to bind and activate in response to the target compound X were able to survive. In both instances, only the desired variants are able to survive, proving even more the power of genetic selection and chemical complementation in protein engineering.

In this chapter, chemical complementation is utilized for protein engineering. Specifically, two nuclear receptors, the pregnane X receptor (PXR) and the retinoid X receptor (RXR) are engineered to target two small molecules, paclitaxel and resveratrol respectively.

8.2 PXR library for Paclitaxel

Motivation to create PXR Library

Paclitaxel is a natural product used as a chemotherapeutic agent for several types of cancers, especially in breast cancer [160-163]. This compound is predominantly found in the bark of Pacific Yew trees, and recently more biosynthetic and synthetic measures have been taken to produce ample amounts of this compound, since relying on the Pacific Yew trees, as the only source of this compound would be environmentally unfavorable. Rodney Croteau and his co-workers at the University of Washington, Seattle, have been trying to decipher the biosynthetic pathway of paclitaxel from cDNA libraries of Pacific Yew trees [160-163]. In particular, these researchers have been trying to identify the substrates and enzymes, which can convert the initial precursors of this pathway, geranyl

geranyl diphosphate, to paclitaxel [160-163]. A number of the enzymes involved in this biosynthetic pathway have been identified, but the entire pathway is yet to be completed.

Chemical complementation can also be used to decipher biosynthetic pathways, which will be discussed in detail in Chapter 9. Briefly, a nuclear receptor can be engineered to bind the target natural compound, in this case paclitaxel. The cDNA library from the organism that produces the natural compound, such as the Yew tress, will be inserted into yeast expression plasmids. The cells will then be subjected to chemical complementation; cells will be expressing the genes for producing paclitaxel and once the paclitaxel is produced, paclitaxel will bind to and activate the receptor controlling expression of a selective marker (e.g. adenine). Thus, these cells will be able to survive on media lacking adenine and maintain the complete biosynthetic pathway.

Our aim was to use chemical complementation to recreate the biosynthetic pathway of paclitaxel in yeast. The human pregnane X receptor (PXR) is activated by paclitaxel in mammalian cell culture assays [95]. However in 2nd generation chemical complementation, PXR is constitutively active (as shown in Chapter 7) due, we hypothesize, to the presence of an endogenous ligand activates the receptor. This is not surprising since PXR has been shown to be a promiscuous receptor, where a variety of ligands activate this receptor, and more than likely, a yeast metabolite is able to bind and activate this receptor. To use PXR as the nuclear receptor that binds the target compound in the pathway, we would have to create a library of PXR variants that

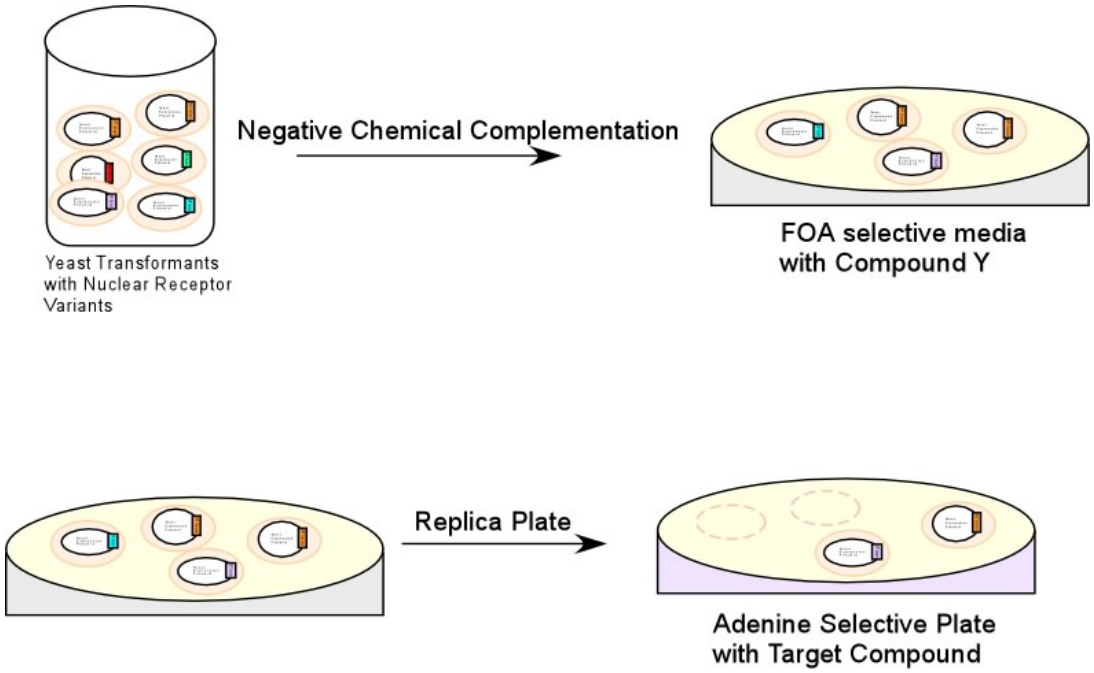


Figure 8.2. Negative chemical complementation for engineering nuclear receptors.

are no longer constitutively active, but that utilize paclitaxel as a ligand.

In silico Design of PXR Library

To create this library of variants, we choose to randomize codons in the binding pocket from the crystal structures of PXR with the synthetic agonist, SR12813, and hyperforin (PDB files 1ILH and 1M13 respectively) [151,164]. Both of these are known agonists for PXR, and their structures were evaluated using Rasmol (RASWIN).

The ligands in these structures seemed to be buried deep in the binding pocket, and the interactions between the side chains of the residues with the ligand were evaluated. We determined which of the residues were within 4 Å and 6 Å of both SR12813 and hyperforin. These amino acids residues were identified and compared, and residues in contact with the both ligands were chosen for randomization. Six amino acids were chosen on their ability to interact with SR12813 and hyperforin: M243, S247,F281, C284, F288, and W299 (Figure 8.3).

Each of the six amino acids were substituted with an alanine, phenylalanine, isoleucine, leucine, methionine, serine, threonine, valine, cysteine, tryptophan, glycine, and alanine (LIVMFAST) (Table 8.1). The majority of these substitutions are hydrophobic residues, and these residues were chosen because of the hypothesis that hydrophobic interactions enhance stability of the ligand-protein complex. Furthermore, substituting with smaller amino acids such as alanine and larger amino acids such as tryptophan, allows the modification of the binding

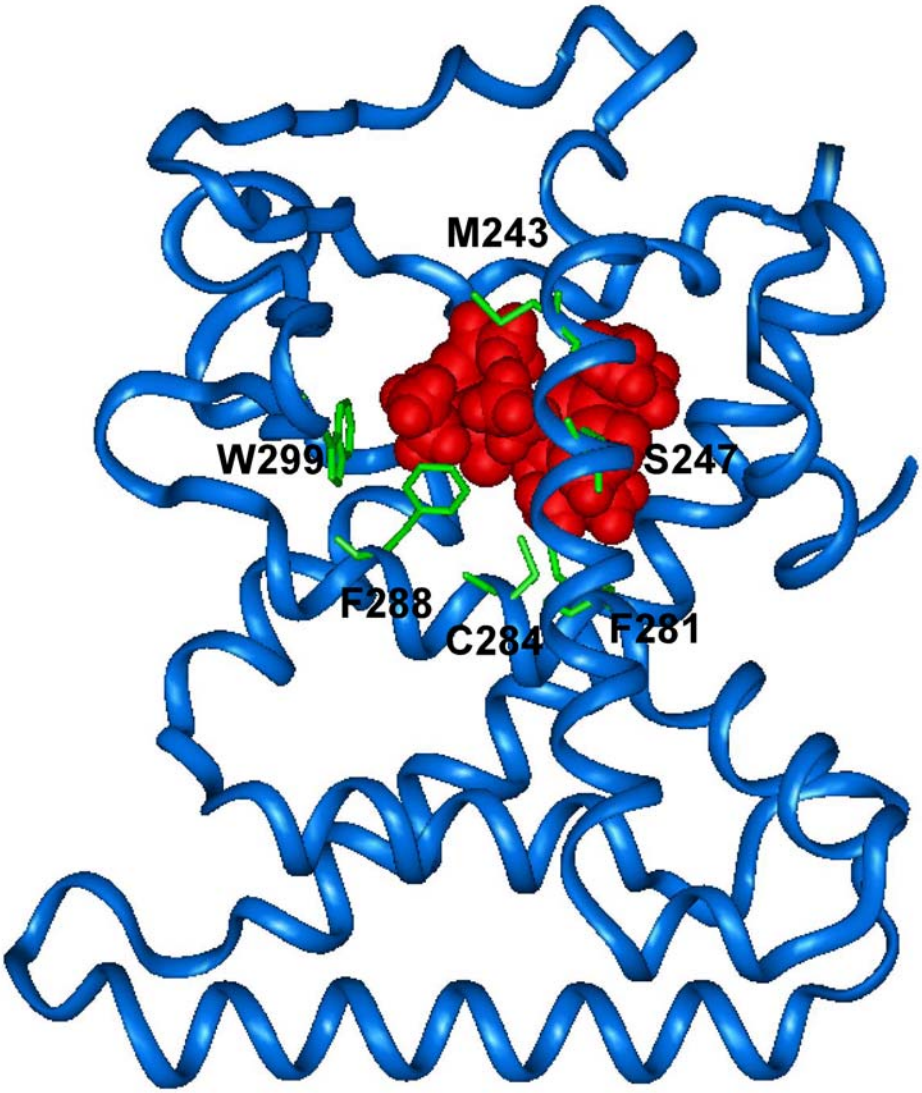


Figure 8.3. Structure of PXR bound to SR12813 with the randomized amino acids.

pocket shape and size, which may result in more selectivity towards paclitaxel and not the endogenous molecules present inside the cell.

Experimental Techniques for design of library

PXR variants were constructed through synthetic oligonucleotides that created insert cassettes with randomized codons using splice overlap extension (SOEd) [165-167]. These insert cassettes contain ends that are complementary to sequences on the background vector cassette, which contains the Gal4 DBD fused to a portions of the PXR LBD sequence. The background plasmid has a random DNA sequence inserted in between portions of the PXR LBD creating multiple STOP codons. The purpose of creating such a background plasmid is to eliminate contamination by vectors containing wild-type PXR. The ends of the insert cassettes are complementary to the portions of the PXR LBD that are present in the background plasmid, such that once the insert cassettes are transformed into yeast with along with the background vector cassettes, the insert cassettes would be inserted into the background vector via homologous recombination. This creates yeast expression plasmids containing the Gal4 DBD fused to the PXR variants.

Oligonucleotides containing the randomized codons at the amino acids positions listed in Table 8.1 were ordered and designed such that oligonucleotides #1 and #8 contained regions of homology to the vector cassette. Each of the oligonucleotides were rehydrated with distilled water and then paired with complementary ends; thus oligonucleotide #1 and #2 are paired, #3 and #4 and so forth (Figure 8.4). These oligonucleotide pairs were subjected to a hybridization protocol, with heating at 95 °C for

3 minutes and then cooled to 10 °C. As the temperature cools to 10 °C, the complementary ends of the pairs hybridize. Once the hybridization is complete, the pairs are put together with *Klenow* polymerase and dNTPs to create full double-stranded cassettes (Figure 8.4). *Klenow* polymerase is an enzyme that is able to fill in overhangs from the 5' to 3' direction. The addition of *Klenow* creates small cassettes, shown in Figure 8.4, which were then placed in a single PCR tube, and with the addition of *Pfu* polymerase and oligonucleotides 1 and 8, these small cassettes were made into the larger insert cassette, shown in Figure 8.4. 30 rounds of PCR were performed to generate more copies of the insert cassette for transformation into yeast.

Table 8.1. PXR Library Design. The amino acid residues modified are listed and the possible the codon changes at each position are listed.

Wild-type PXR Residue	Possible Codon Changes
M 243	LIVMFASTCRGW
S 247	LIVMFASTCRGW
F 281	LIVMFASTCRGW
C 284	LIVMFASTCRGW
F 288	LIVMFASTCRGW
W 299	LIVMFASTCRGW

The vector cassette was digested with the restriction enzyme, *KpnI*. This digests out of the vector cassette the stop codons leaving the Gal4 DBD fused to portions of the PXR LBD. These portions of the PXR LBD are homologous to the ends of the insert

cassettes. The linearized vector cassette was treated with calf intestinal phosphatase (CIP) so that re-ligation does not occur.

The linearized vector cassette and randomized insert cassettes were transformed into PJ69-4A cells, which contain the pGAD10BAACTR plasmid (this plasmid expresses the Gal4 activation domain fused to the ACTR gene). The amount of linearized vector cassette was 1 μg and about 27 μg of the randomized insert cassettes were added. The vector and insert cassettes were transformed into yeast using the large-scale transformation protocol [106].

Transformants were plated onto selective media lacking leucine and tryptophan (-Leu -Trp). This is used to confirm that recombination has occurred; if recombination occurs, the linearized vector will become circular, allowing the expression of the tryptophan marker. The pGAD10BAACTR plasmid contains the leucine marker. Dilutions of 1:20 and 1:200 of the transformation mixes were made such that the colonies were spread out and could be used for replica plating. The plates (standard size of were incubated at 30 °C for three to five days.

Results and Discussion of PXR Library

A transformation efficiency of 1 x 10^5 colonies μg^{-1} vector DNA was obtained, creating a library size of about a million transformants. The theoretical library size is 3.5 x 10^7 possible codon combinations. The

Table 8.2. Sequences from the PXR Library Oligonucleotides with overlapping regions

Oligo #	Oligonucleotide Sequence (Overlap regions Underlined and T_M noted) Target Codons in Bold and Highlighted
1	5'- CCGGAAAGATCTGTGCTCTTTGAAGGTCTCTCTGCAGCTGCGGGG<u>GGGAGGATGGCAGTGTCTGG</u> –3' T_M = 56.5 °C
2	3'-<u>CCCTCCTACCGTCACAAGACC</u>TTGATGTTTGGGGGTCGGCTGTCACCGCCCTTT<u>CTCTAGAAGAGGGACGACGGG</u>-5' T_M = 56.5 °C T_M = 56.6 °C
3	5'- <u>GAGATCTTCTCCCTGCTGCCC</u>CAC**ATG**GCTGACATG**TCA**ACCTACATGTTCAAA<u>GGCATCATCAGCTTTGCCAAAG</u>–3' T_M = 56.6 °C **M243** **S247** T_M = 58.4 °C
4	3'-<u>CCGTACTACTCGAAACGGTTTC</u>AGTAGAGGATGAAGTCCCTGAACGGGTAGCTC<u>CTGGTCTAGAGGGACGACTTCC</u>-5' T_M = 58.4 °C T_M = 56.6 °C
5	5'-<u>GACCAGATCTCCCTGCTGAAGG</u>GGGCCGCT**TTC**GAGCTG**TGT**CAACTGAGA**TTC**AA<u>CACAGTGTTCAACGCGGAGGACTG</u> 3' T_M = 56.6 °C **F281** **C284** **F288** T_M = 59.3 °C
6	3'-<u>GTGTCACAAGTTGCGCCTCTGAC</u>CTTGG**ACC**CTCACACACCGGCCGACAGGAT<u>GACGATCATTCTGTGACGTCCACCG</u>-5' T_M = 59.3 °C **F299** T_M = 58.3 °C
7	5'- <u>CTGCTAGTAAGACACTGCAGGTGGC</u>TTCCAGCAACTTCTACTGGAGCC<u>CATGCTGAAATTCCACTACATGCTG</u>–3' T_M = 58.3 °C T_M = 57.0 °C
8	3'-<u>GTACGACTTTAAGGTGATGTACGAC</u>TTCTTCGACGTCGACGTACTCCTCATACACGACTACGTCCGGTAGAGGG-5' T_M = 57.0 °C

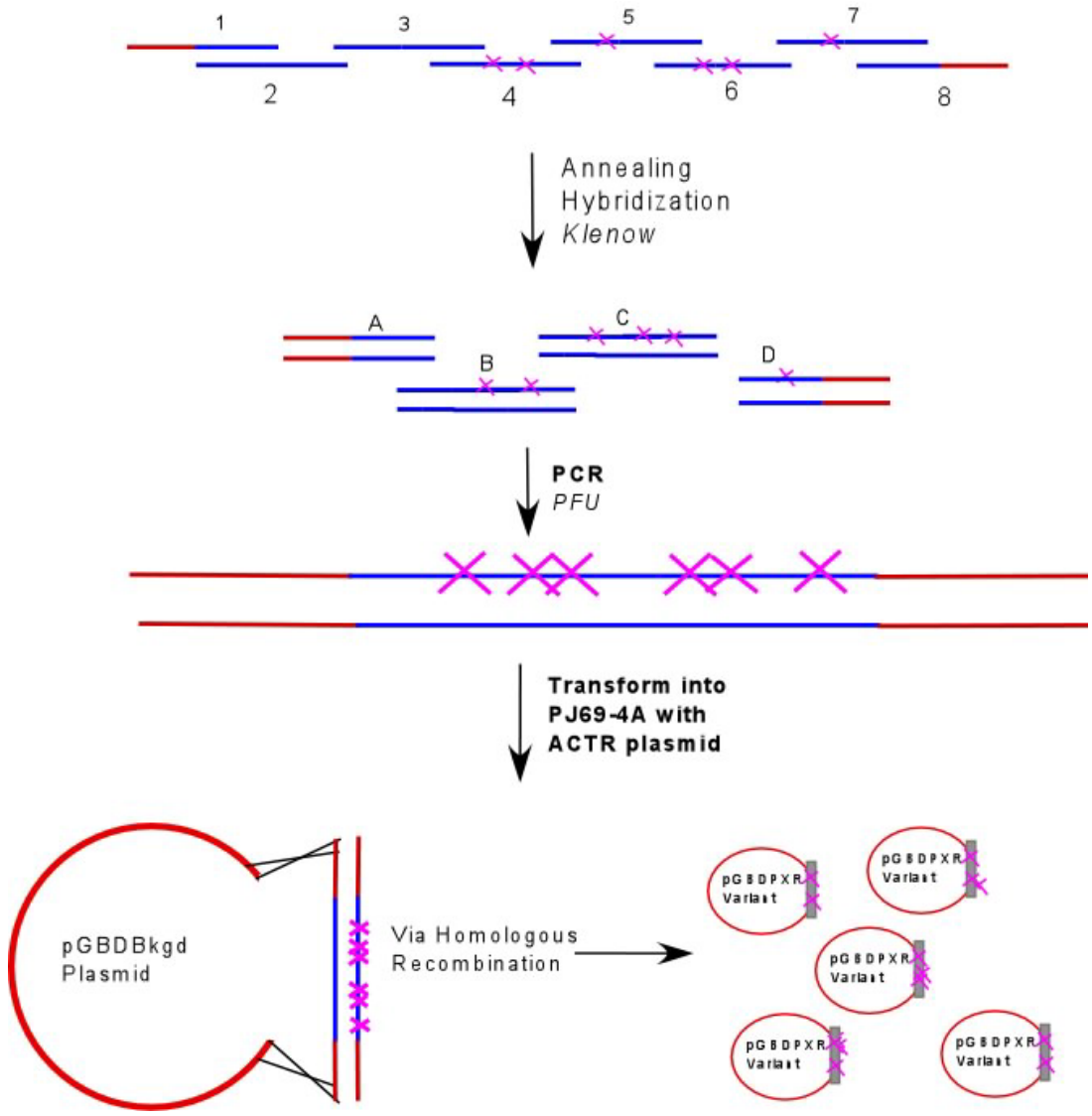


Figure 8.4. Creating PXR library of randomized codons.

transformants were replica plated onto adenine selective media containing 10^{-5} M concentrations of a paclitaxel, tocopherol and resveratrol. Tocopherol is another known agonist for PXR and resveratrol is a natural product found in grapes another compound of interest (discussed in more detail later in the chapter). From the replica plating results, it was determined that the all of the variants were constitutively active. This was not surprising since wild-type PXR is constitutively active as well. A few of the transformants were sequenced to characterize the diversity in the library. From the sequencing results of four transformants, we were able to determine these transformants contained wild-type sequences with additional insertions and deletions.

The PXR library results were disappointing in that the diversity was not present, but this library was also not fully characterized; only 4 out of the million transformants were sequenced. The rest of the variants were not characterized because the majority of the transformants were constitutively active, and not of interest. From these results, we were able to see that starting with a receptor that is constitutively active, and creating a library of variants that are not constitutively active, but ligand dependent is not easy. A better design would be to incorporate a receptor that is not constitutively active and try to engineer the receptor to bind the desired ligand. The next section focuses on engineering the retinoid X receptor, RXR, which is not constitutively active, towards resveratrol.

8.3 **RXR Library for Resveratrol**

Benefits of Resveratrol

Engineering PXR variants without constitutive activity and also with ligand activated growth with paclitaxel proved to be unsuccessful, since most the variants

retained the constitutive activity. We hypothesized that starting with a compound that is not constitutively active in chemical complementation would be a better starting point, such as the retinoid X receptor (RXR) or the liver X receptor (LXR). The ligand-binding pocket of both of these receptors is not naturally accommodating to bulkier small molecules, like paclitaxel. Therefore, we decided to targeted another natural product, resveratrol.

Resveratrol is a stilbene phytoalexin found in many dietary compounds and red wine (Figure 8.5) [168]. This compound was implicated in epidemiological data demonstrating the inverse correlation between the consumption of red wine and cardiovascular disease, known as the "French paradox" [168]. The "French paradox" was established because the French have a diet rich in fatty foods, and have a 30% lower incidence of heart disease than North Americans. Obesity and other cardiovascular diseases within the French population are also considerably lower in comparison to North Americans. Researchers believe that compounds similar to resveratrol are responsible for the "French paradox" [168].

Resveratrol's mechanism in the prevention of many human pat*holo*gical processes, including inflammation, atherosclerosis, and cardiovascular disease stem from this compound's ability to act as an antioxidant and prevent the oxidation of low density lipoproteins (LDL) [168]. Oxidized LDL result in the abnormal accumulation of foam cells and provide an environment for artherosclerosis.

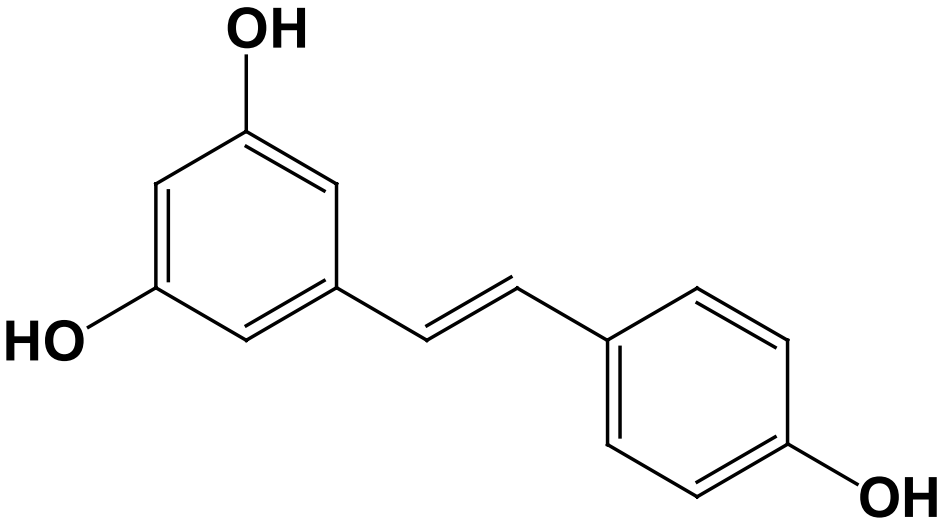


Figure 8.5. Structure of Resveratrol.

Resveratrol has also been shown to decrease the toxic effects of reactive oxygen species (ROI) in living cells; abnormal accumulation of ROI molecules results in oxidative stress eventually causing deleterious effects within the cell [168].

More recently, resveratrol has been shown to have anticancer activities and has been shown to possibly have a remarkable effect as a chemopreventive and chemotherapeutic agent. Researchers are trying to decipher the anti-cancer mechanism of this compound. Several key proteins are possible targets for the anti-cancer mechanism of resveratrol, such as the aryl hydrocarbon receptor, the transcription factor NF-κB, and through the androgen receptor in prostate cancer cells [168]. In general, biochemical pathways involved in differentiation, transformation and cell cycle growth seem to be potential targets of the resveratrol, and in vivo analysis in mice have shown that rat treated with resveratrol develop fewer tumors [168].

We decided to use resveratrol as a target for engineering nuclear receptors to bind a natural product. The resveratrol molecule resembles estradiol structures, agonists of the estrogen receptors. We decided to engineer a RXR variant to bind resveratrol, since we had already worked with RXR in chemical complementation and this receptor is not constitutively active.

In silico modeling of RXR and resveratrol

Protein engineering has greatly benefited from the advancement of in silico modeling programs, which have allowed researchers to evaluate possible protein-protein or protein-ligand interactions. Protein-protein and protein-small molecule interactions can

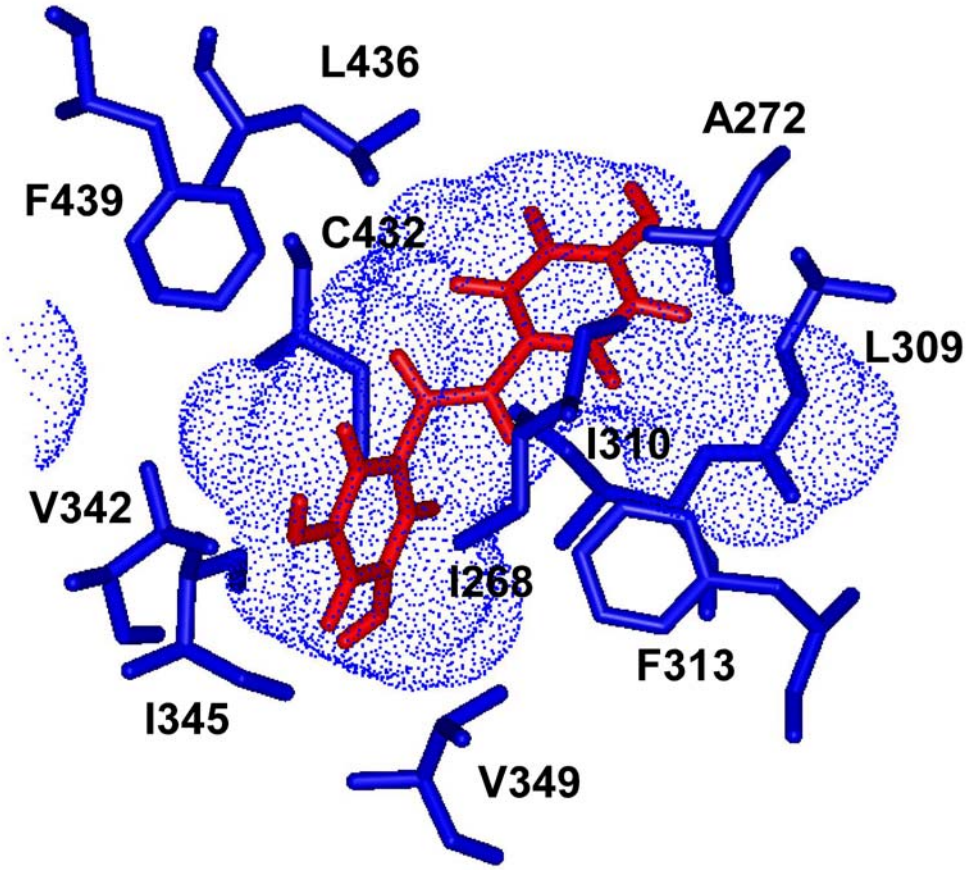


Figure 8.6. Molecular surface diagram of the RXR LBD docked with resveratrol where the residues used for randomization are shown. Stippled outline is the space occupied by 9cRA. Resveratrol is shown in red.

be evaluated using several algorithms and modeling parameters, allowing the evaluation of key molecular interactions. For analyzing the interface between small molecules and receptors, docking has emerged as a valuable tool allowing the modeling of the interactions between the ligand of interest and the receptor. Furthermore, for protein engineering, these programs allow local variations in protein structure, such as point mutations.

To engineer RXR variants that are able to bind and activate in response to resveratrol, we decided to use AFFINITY (Insight), an in silico program that docks a conformational flexible ligand into the binding site of the desired protein. The energies of the protein-ligand structures are calculated, where the lower the energy of the protein-ligand complex, the more likely that those molecular interactions occur.

Resveratrol was docked into the binding pocket of the wild-type RXR structure (PDB file: 1FBY). From the results of the molecular docking, we were able view the molecular interactions between the ligand and the amino acid residues in the binding pocket of RXR. The results of the docking procedure indicated a resveratrol- wild-type RXR complex with energy of –19 kcal mol^{-1}, and this structure was used to postulate which amino acid residues would be in contact with resveratrol in the binding pocket. A molecular surface diagram was generated for resveratrol into the binding pocket of RXR, with the RXR wild-type ligand, 9-*cis* retinoic acid (9cRA), shown as a space fill model (Figure 8.6). Resveratrol docks into the binding pocket in an orientation similar to that of 9cRA, however, the 9cRA fills part of the pocket that is not occupied with resveratrol, particularly at position A272.

Table 8.3. RXR Library Design. The amino acid residues modified are listed and the possible the codon changes at each position are listed.

Wild-type RXR Residues	Possible Codon Changes
I 268	LIVMFAST
A 272	LIVMFASTCRGW
L 309	LIVMFAST
I 310	LIVMFAST
F 313	LIVMFAST
V 342	LIVMFAST
I 345	LIVMFAST
F 346	LIVMFAST
V 349	LIVMFAST
C 432	LIVMFAST
L 436	LIVMFAST
F 439	LIVMFAST

To determine which specific amino acids residues to target for randomization, we used the docked structure of RXR and resveratrol, and determined that the following residues in the binding pocket were within 2.5 Å and 4 Å of the ligand: I268, A272, L309, I310, F313, V342, I345, F346, V349, C432, L436, and F439. From that analysis, these 12 amino acids in the binding pocket were chosen for mutagenesis based on their location and possible interactions with the ligand. These 12 amino acids are listed in Table 8.2.

Each of the amino acid residues that were chosen was mutated to a series of hydrophobic residues. This is based on the fact that the RXR ligand binding pocket consists primarily of hydrophobic residues with a few positively charged side chains that are able to stabilize the negative charge of its wild-type ligand, 9cRA. Changing the natural amino acid side chains to other hydrophobic residues is based on the hypothesis that hydrophobic contacts increase the binding affinity. Selectivity for the desired ligand is achieved by the size and shape of the pocket, as well as through electrostatic interactions and hydrogen bonding.

Of the residues listed in Table 8.2, 11 out of the 12 amino acid residues were changed from the wild-type residue to an alanine, phenylalanine, isoleucine, leucine, methionine, serine, threonine, and valine. At alanine 272 (A272), in addition to the eight amino acid residues substituted at each of the other sites, a glycine, arginine, tryptophan and cysteine were added to the possible amino acid residues that were possible changes that could be made at that position. These additional four amino acids were added to position A272, there is a part of the binding pocket that is not filled with the resveratrol

molecule, which is occupied by 9cRA (Figure 8.6). Therefore, substituting larger amino acids in that region, such as tryptophan, may occupy the empty space present with the resveratrol in the binding pocket, and create a more stable binding pocket for resveratrol.

Experimental technique for creating and evaluating library

The library of RXR variants was created through the techniques known as homologous recombination. Yeast, along with other organisms, have the ability to recombine sequences, which are complementary. Therefore, synthetic oligonucleotides were ordered to re-create the portion of the RXR LBD that contained the residues that were going to be subjected to the mutagenesis. A vector cassette containing the Gal4 DBD fused to a part of the RXR LBD was created. This vector contained a small region of the RXR LBD, and then contained random stop codons, followed by another part of the RXR LBD. This type of a vector cassette was created to prevent contamination by wild-type RXR, and designed such that homologous recombination would occur. The ends of the synthetic RXR insert cassettes are complimentary to the parts of the RXR LBD on the vector cassette, thus when these two pieces are introduced into yeast, homologous recombination will occur.

Oligonucleotide pairs were mixed together to create small cassettes. All of the oligonucleotides were synthetically made except oligonucleotide 7, which was created through amplification of a 200 bp fragment of the RXR gene. In this region, no modifications were made to any of the amino acid residues. All of the other oligonucleotides were paired with another oligonucleotide that had a complementary end (e.g. #1 and #2 were paired, #3 and #4, etc.).

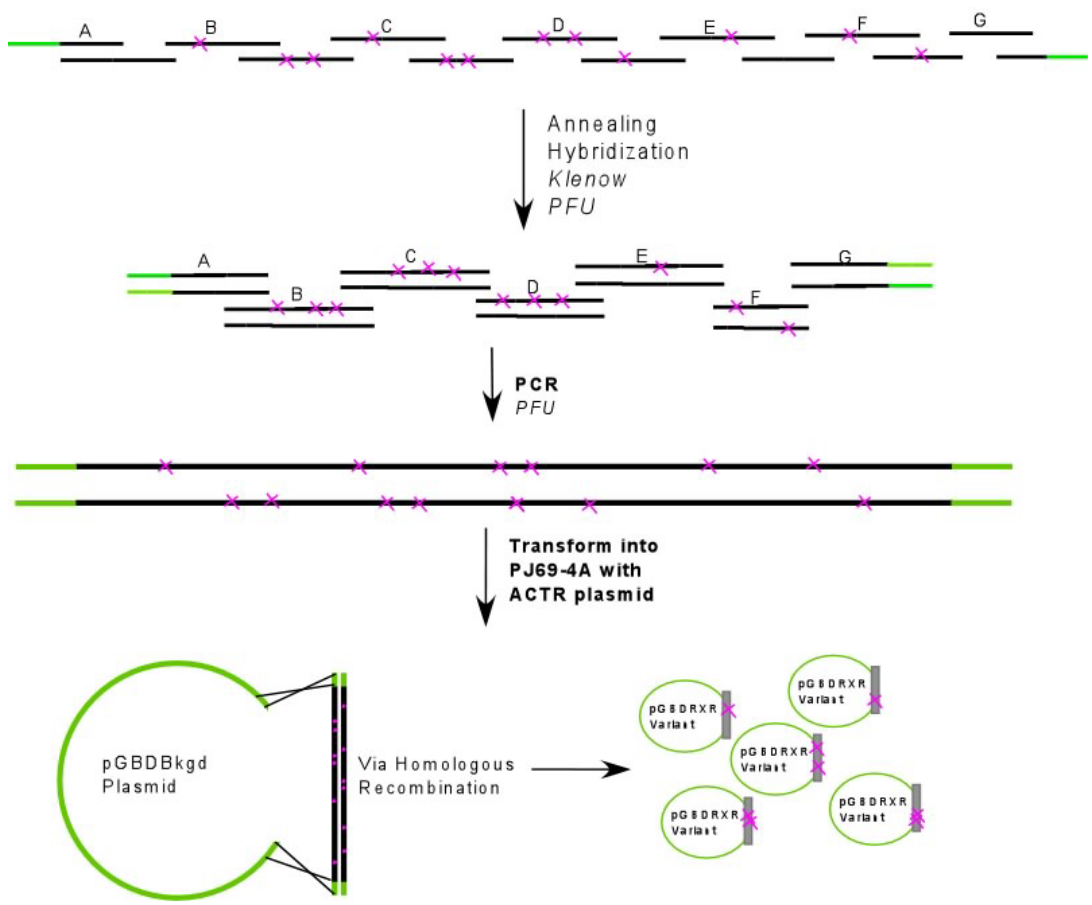


Figure 8.7. Experimental technique for RXR library.

Table 8.4. Sequences from the RXR library oligonucleotides with overlapping regions

Oligo Number	Oligonucleotide Sequence (Overlap regions Underlined and T_M noted) Target Codons in Highlighted and Underlined
1	5'-CGAGGACATGCCGGTGGAGAGGATCCTGGAGGCTGAGCTGGCCGTGGAGCC -3' T_M = 52.6 °C
2	5'-GCCCCATGTTTGCCTCCACGTAGGTCTCGGTCTTGGGCTCCACGGCCAGC-3' T_M = 53.2 °C T_M = 52.6 °C
3	5'-GGAGGCAAACATGGGGCTGAACCCCAGCTCGCCGAACGACCCGGTCACCA-3' T_M = 53. 2 °C T_M = 51.7 °C
4	5'-TGGCCCACTCCACTAGTGTGAAAAGCTGTTTGTCGGCTGCTTGGCAAATGTTGGTGACCGGGTCGTTC-3' T_M = 53. 1°C A272 I268 T_M = 51.7 °C
5	5'-CACACTAGTGGAGTGGGCCAAGCGGATCCCACACTTCTCAGAGCTGCCCC-3' T_M = 53. 1°C T_M = 49 °C
6	5'-GCCTGCCCGCAGCAGGATGACCTGGTCGTCCAGGGGCAGCTCTGAGAAGT-3' T_M = 52.5 °C T_M = 49 °C
7	5'-TGCTGCGGGCAGGCTGGAATGAGCTCCTCATCGCCTCCTTCTCCCACCGCTCCATCGCCG-3' T_M = 52.5 °C L309 I310 F313 T_M = 51.8 °C
8	5'-CGTGCAGCCCGGTGGCCAGGAGAATTCCGTCCTTCACGGCGATGGAGCGG-3' T_M = 54.6 °C T_M = 51.8 °C
9	5'-CCACCGGGCTGCACGTCCACCGGAACAGCGCCCACAGCGCAGGG-3' T_M = 54.6 °C T_M = 53.5 °C
10	5'-CCGCATCTTGGACACAAGCTCCGTCAGCACCCTGTCAAAGATGGCGCCCACCCCTGCGCTGTGGGC -3' T_M = 53.7 °C V349 F436 I345 V342 T_M = 53.5 °C
11	11 For: 5'-CTTGTGTCCAAGATGCGGG -3' 11 Rev: 5'-GCCCGATGGAGCGC-3' T_M = 53.7 °C T_M = 50.6 °C
12	5'-GGTGTGTCCCCGATGAGCTTGAAGAAGAAGAGATGTTCCAGGCACTTAAGCCCGATGGAGCGC-3' T_M = 53.1 °C F439 L436 C432 T_M = 50.6 °C
13	5'-GCTCATCGGGGACACACCAATTGACACCTTCCTTATGGAGATGCTGGAGGCGCCGCACC-3' T_M = 53.1 °C T_M = 53.5 °C
14	5'-GGCAGGGTGGCCAGAACGGGTGGGCACAAAGGATGGGCCCGCAGGCCTAAGTCATTTGGTGCGGCGCCTCC-3' T_M = 53.5 °C

About a microgram of each of these oligonucleotides were initially subjected to a hybridization protocol, where they were placed in the same tube and incubated at 95 °C for three minutes and then cooled to 10 °C. The hybridized oligonucleotides were placed in the presence of *Klenow* polymerase and dNTPs, which then fill in the overhangs of the hybridized oligonucleotides, extending the oligonucleotides and created cassettes, named A-G.

These cassettes (concentration of about 100 ng) were added together for PCR amplification using *Pfu* polymerase. Each of the cassettes had complementary overhang sequences, which were hybridized during the PCR procedure. *Pfu polymerase* extends the hybridized cassettes, creating the full synthetic cassette. The full cassette was further amplified using oligonucleotides 1 and 14 (Figure 8.7). Again, 30 rounds of PCR were performed to generate more copies of the insert cassette for transformation into yeast.

The background plasmid was linearized with the *SacI* and *SpeI* restriction enzyme sites and then treated with calf intestinal phosphatase to prevent the religation of the plasmid. These restriction enzyme sites digest out the stop codon sequence that was placed in the vector, and leave the RXR portions necessary for recombination intact. The purpose of a linearized vector cassette is that linear fragments are shown to have a higher transformation and recombination frequency [109]. Also, only if recombination occurs will the cells be able to survive on the selective media because the plasmid will become circular upon recombination, permitting replication and transcription of the selective tryptophan marker.

Results of Library #1

PJ69-4A cells already transformed with a yeast expression vector containing the Gal4 activation domain fused to the ACTR gene (pGAD10BAACTR) were grown in leucine selective media, selecting for the ACTR plasmid. These cells were made competent for a 30 X LiAc transformation procedure in which about 50 μg DNA can be transformed into yeast cells. 25 μg of the insert cassettes amplified through PCR and 2 μg of the linearized vector cassette were transformed into the yeast cells containing the ACTR plasmid. Transformants were spread onto two types of selective plates: leucine and tryptophan selective media and adenine selective media with 10 μM resveratrol. The first type of selective plates were to determine the transformation and recombination efficiency of the experiment; the tryptophan deficiency is overcome if recombination occurs and the plasmid with the Gal4 DBD fused to the RXR LBD is able to become circular. Yeast cells with RXR variants able to bind and activate in response to resveratrol will be selected for on the adenine selective plates with resveratrol. Plates were incubated at 30 °C for three to five days.

A transformation efficiency of 4×10^4 colonies μg^{-1} of vector DNA was obtained in that procedure, creating a library size of 1.6×10^5 transformants. On the adenine selective plates with resveratrol, 125 colonies grew and all of these colonies were constitutively active. Transformants from the leucine and tryptophan plates, as well as from the adenine selective plates with resveratrol were picked and plasmids from these cells were rescued, retransformed into *E.coli* and sequenced. Table 8.3 shows the results of the sequences obtained from this library.

The constitutively active variants do not contain amino acid changes at all 12 positions (Table 8.3). Only a couple of the transformants from the leucine and

tryptophan selective plates contained a mutation at every position. This could be due to the possibility that mutations at certain positions, particularly V342, I345, F346, and V349 are deleterious to the function of these receptors. However, a few of the transformants that were sequenced from both sets of plates contained wild-type silent mutations, previously made for adding restriction enzyme sites. This contamination of wild-type RXR could have hindered the variety in the library. Thus, to overcome the contamination by wild-type RXR, additional restriction enzyme analyses were performed to ensure that only the background plasmid remained, and another transformation was performed.

Results of Library #2

To eliminate the presence of the plasmid containing wild-type RXR, stock of background plasmid was digested with two restriction enzymes only present in the wild-type RXR plasmid. This digested sample was retransformed into *E.coli* and a single colony was selected for a miniprep procedure. The background vector was digested with *SacI* and *SpeI* restriction enzyme sites and then treated with calf intestinal phosphatase to prevent the religation of the plasmid. Another transformation procedure was done with 2.75 μg of the background-linearized fragment and 25 μg of the randomized insert cassettes. The same procedures were performed as in the last transformation (see above). A transformation efficiency of 7.4×10^4 colonies μg^{-1} of vector DNA was obtained in this procedure, creating a total library size of 2×10^5 transformants.

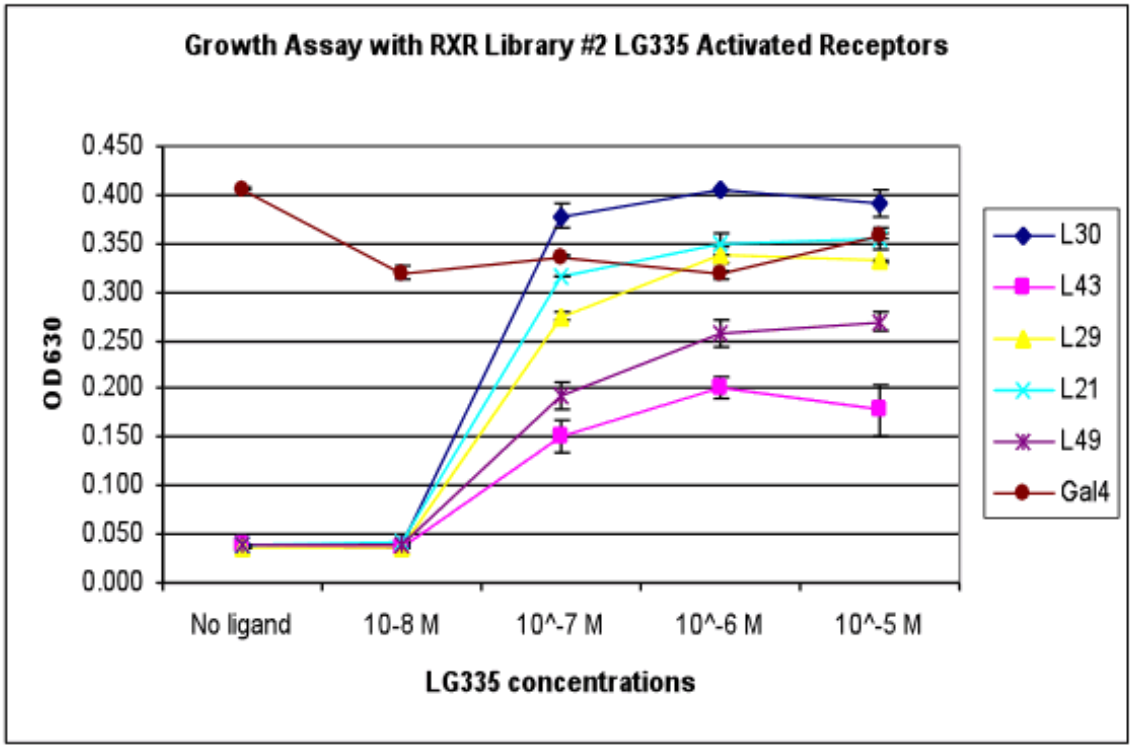
Table 8.5. Sequence of RXR Variants in RXR Library #1

Wild-type Residues	I 268	A 272	L 309	I 310	F 313	V 342	I 345	F 346	V 349	C 432	L 436	F 439
Variants from Naïve Library												
1										C432T	L436A	F439M
2	I268A	A272G	L309A	I310T	F313A			F346L	V349I			
3	Wild-Type RXR											
4		A 272G	L309V	I310V	F313V	V342S	I345M	F346S	V349T	C432V	L436A	F439A
5			L309V	I310S	F313V	V342T	I345F	F346T	V349S	C432L	L436I	F439V
Variants from Selected Library												
1	I 268I (1 bp change)	A272T		I310S	F313V						L436F	
2				I310M	F313M							
3	I268A			I310M	F313A					C432F		F439T
4				I310L	F313M							
6			L309L (3 bp change)	I310L	F313V					C432F	L436I	F439I
7	Wild-Type RXR											
8			L309L (2bp change)	I310L	F313A					C432T	L436L (3 bp change)	F439V
9	I268A	A272A (1 bp change)		I310M	F313A	V342M		F346T	V349V (1 bp change)	C432F	L436L (2 bp change)	F439L
10				I310L	F313M					C432F		

On the adenine selective plates with resveratrol, 140 colonies grew, and these colonies were streaked onto adenine plates with and without resveratrol. All of the colonies were able to grow in the absence of ligand, making them all constitutively active.

To determine whether any of the transformants were activated in response to any ligands, adenine selective plates were made containing resveratrol, LG335, androstanol, γ-1-oxo-pyrenebutyric acid (γ-OPBA), 6-amino penicillanic acid (6-APA), Penicillin G (PEN G), and naficillin, all at 10 μM concentrations. Transformants that were able to grow on these plates were then tested on adenine plates without ligand. From this analysis, we observed that 30 LG335 activated receptors and 5 γ-OPBA activated receptors existed in the library. These candidates were then analyzed in the quantitation growth assay to determine the potency and efficacy.

The LG335 activated variants and γ-OPBA variants have different activation profiles (Figure 8.8 and 8.9). The LG335-activated variant, L24, is growing at a concentration lower than 10^{-8} M making these receptors more potent than others (Figure 8.8). Variants D3 and D4 seem to have the same transcriptional growth profile, but variant E10 is activated at γ-OPBA concentrations as low as 10^{-8} M (Figure 8.9). Six variants from both the LG335 activated and γ-OPBA activated receptors were sequenced (Table 8.4). The mutations found in these variants did not contain any mutations in the V342, I345, F346, and V349 residues, and none of these variants had a mutation present at every single randomized codon. This could be due to a couple of reasons. First, a combination of all 12 mutations may have a deleterious effect on the structure and



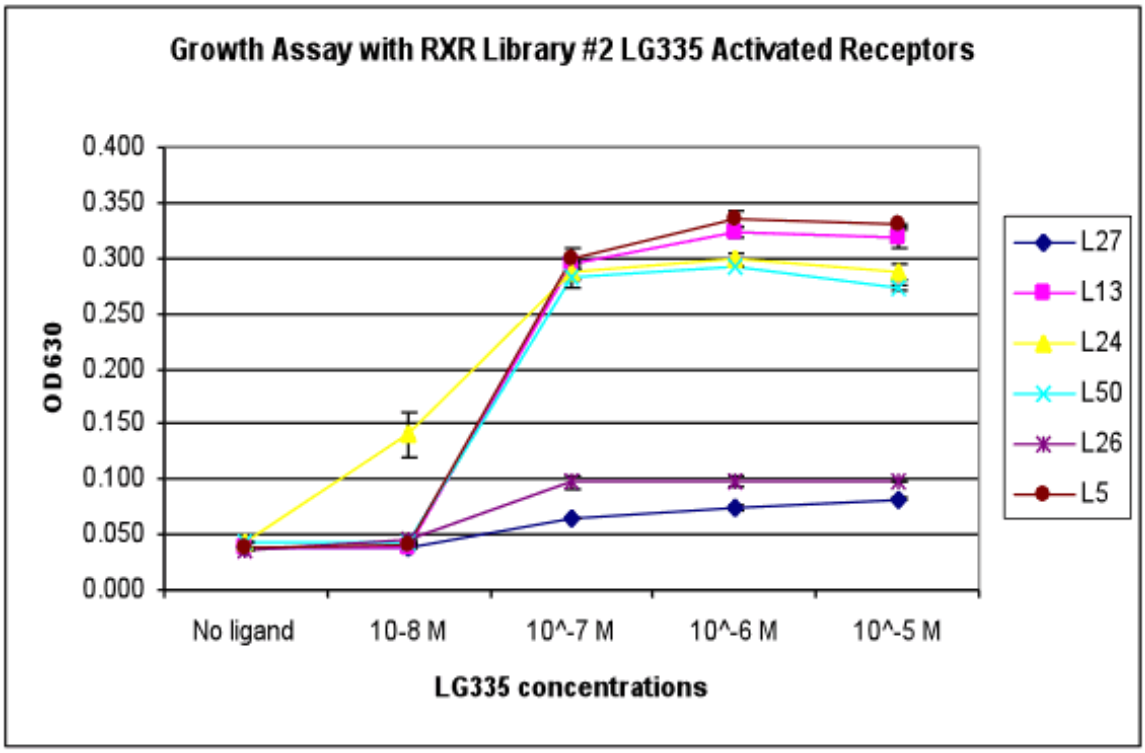


Figure 8.8. Subset of LG335 activated variants from RXR library #2.

function of the receptor; therefore these variants are not present on the adenine selective plates. The V342, I345, F346, and V349 residues are all located on helix 7 in the RXR LBD. Mutating these residues may hinder the proper alignment of helices for coactivator recruitment, or affect the stability of the receptor. The second reason could be due to an incomplete library, which is clearly seen through the sequences of the LG335 and the γ-OPBA variants. Multiple rounds of PCR were performed to amplify these insert cassettes and multiple rounds of PCR are known to introduce a bias in the library.

Discussion of RXR library toward resveratrol

The goal of discovering a novel RXR variant that can bind to resveratrol was not achieved through this library design. However, the library does contain functional variants. Most of these variants were constitutively active, probably due to the ability to be activated by an endogenous ligand. In the library, LG335-activated receptors were present with a variety of transcriptional profiles. γ-OPBA activated variants were also found in this library with a couple of interesting variants. One of these variants was activated by γ-OPBA at a concentration of 10^{-8} M or lower.

Chemical complementation is a powerful tool for engineering proteins, especially towards engineering nuclear receptors. A library was created with 200,000 possible candidates, and chemical complementation uses the power of genetic selection to allow only yeast with the active variants to survive. From the 200,000 transformants, eventually only about 35 were active variants. When the sequences from both RXR library #1 and #2 are analyzed, it is obvious that

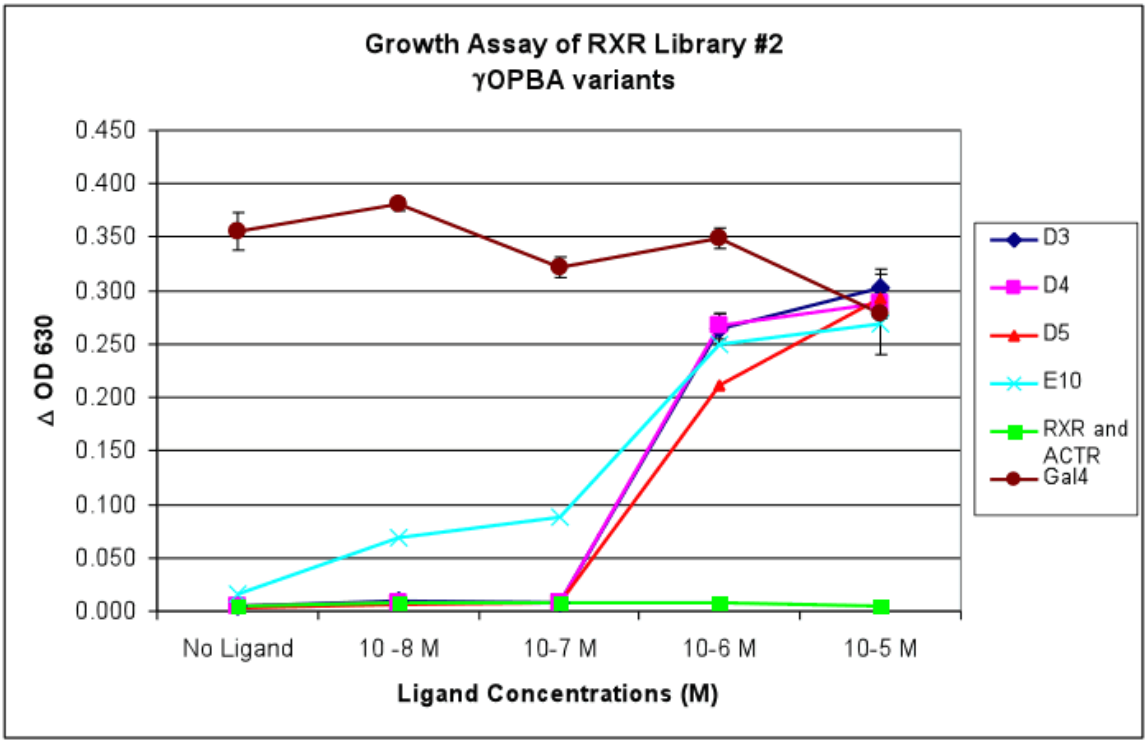


Figure 8.9. γ-OPBA activated variants from RXR Library #2.

the residues that are more tolerant of modifications are present in the active receptors, and those mutations that are functionally or structurally deleterious are present in the non-functional plates. By analyzing sequences from selective adenine plates (-Ade-Leu-Trp) and non-selective plates (-Leu-Trp plates), we are able to determine which residues may be causing a problem and may not be good candidates for engineering RXR.

In most assays where "dead variants tell no tales", with chemical complementation, "dead variants" do tell a story. From looking at the data set obtained from these libraries, mutations were not found in any of the residues mutated in helix 7. Looking back at the docking results from the RXR-resveratrol complex, the residues in helix 7 played a significant role in contacting the ligand. Therefore, modifications in this helix may be required for resveratrol to bind, yet mutations in these positions may lead to non-functional variants. Therefore, another library design may overcome these effects, providing variants that are able to bind and activate in the presence of resveratrol, or RXR may not be a suitable receptor for creating resveratrol variants. RXR seems to be a good candidate for engineering LG335 variants and γ-OPBA as demonstrated in this library and previously published ones, but variations in the binding pocket to accommodate these ligands may not be major changes, which may have to be made to bind compounds like resveratrol.

Table 8.6. Sequence of RXR Variants in RXR Library #2

Wild-type Residue	I 268	A 272	L 309	I 310	F 313	V 342	I 345	F 346	V 349	C 432	L 436	F 439
γ-OPBA Activated Variants												
D3	I268V	A272V		I310M	F313S					C432S		
D4				I310M	F313S						L436M	
D5	I268V	A272V		I310M	F313S							
LG335 Activated Variants												
L27	I268A			I310M	F313A							
L15	I268A			I310M	F313A							
L29	I268V	A272V		I310L	F313M							

8.4 Negative Chemical Complementation and Protein Engineering

As shown in the two libraries above, one of the problems when analyzing these libraries is the presence of constitutively active receptors. These constitutively active receptors can be removed from the library, leaving only ligand-activated receptors using negative chemical complementation. Negative chemical complementation for protein engineering adds a powerful new dimension to using genetic selection for protein engineering.

Cells are normally re-suspended in water after a heat-shock period and before they are plated onto selective plates. For the addition of negative chemical complementation to the protocol, after the heat-shock the cells are re-suspended in selective FOA media containing either no ligand, or with a compound that is desirable not to activate transcription (i.e. wild-type ligand or a stereoisomer of the target ligand). The cells are then grown for a period of time before these cells are plated onto positive selective plates (e.g. adenine, histidine, or uracil) containing the target ligand. Negative chemical complementation will eliminate all the undesirable variants. Cells that contain constitutively active receptors or receptors that are activated by the undesired ligand will die because these cells will express the Ura3 protein, which produces the fluorouracil toxin (Chapter 6).

To incorporate negative chemical complementation into the transformation protocol, we wanted to determine how long the cells have to be grown in the FOA selective media such that all the undesired variants are killed and only ones that are not able to respond to the unwanted ligand are able to survive. To optimize the negative

chemical complementation protocol for protein libraries, a plasmid expressing the Gal4 *holo* protein, pCL1, was used which contains a leucine selective marker. The Gal4 protein was used for experimentation because Gal4 is a ligand-independent transcription factor, eliminating the ample amounts of ligand that would be required for optimization of this protocol. This plasmid was transformed into BAPJ69, and after the heat-shock stage, the sample was resuspended in selective media lacking uracil and leucine with FOA (FOA media contains some concentration of uracil so technically the media does contain uracil). These cells were then grown for various time periods and plated onto adenine and uracil selective media.

The number of colonies at each time point was determined. The expected result would be that as the incubation time in the FOA media increases, the number of living cells would decrease, due to the fact that these cells have the Gal4 plasmid and are able to turn on the *URA3* gene expression and produce the toxin, killing the yeast. As shown in Figure 8.10, at the initial time point, there were about 200 colonies, and after 15 minutes the number of colonies has decreased by about 60% to 80 colonies. From then onward, the number of cells decrease but at smaller increments, and at the last incubation time point, there number of colonies left are about 22% of the original colonies.

Although 78% of the undesired cells had been eliminated, for large protein libraries, the remaining 22% would be "false positives", and interfere with library evaluation. Thus, we wanted to increase the efficiency of negative chemical complementation. One possibility is to decrease the concentration of uracil in the FOA media to have an effect in eliminating the remaining 22%. BAPJ69 yeast cells were transformed with pCL1, and after the heat-shock incubation period, these cells were then

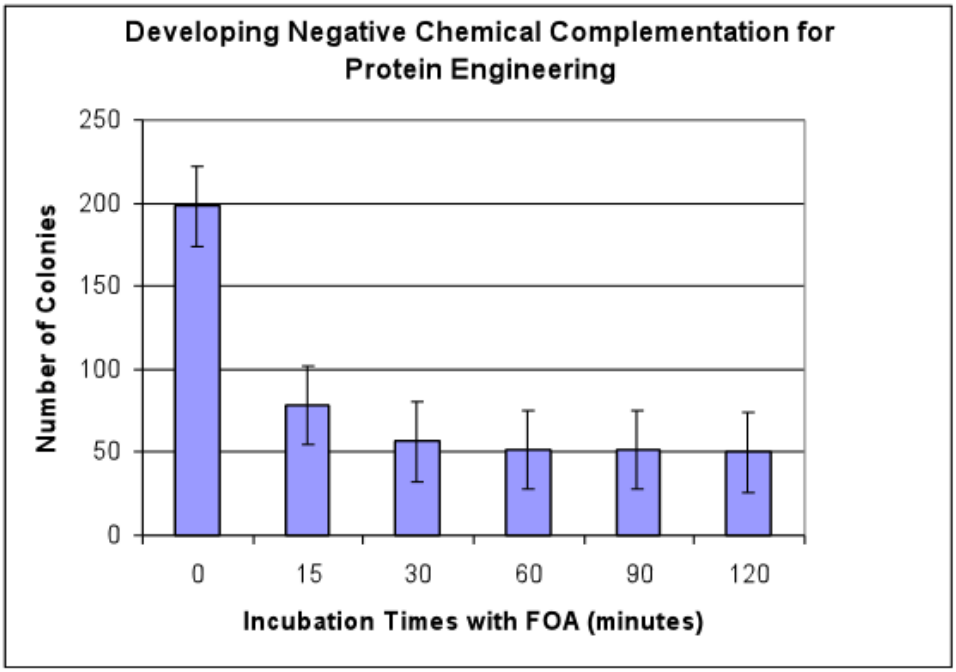


Figure 8.10. The number of colonies at each time point with FOA.

grown in FOA media with varying concentrations of uracil for different incubation times. These cells were grown in media containing the regular amount of uracil in the FOA (20 μL) and also 15 μL, 10 μL, 5 μL, and with no uracil in the FOA. The results are shown in Table 8.7. The kill rate increase to about 91%, but there is no real correlation between the amount of uracil present and the number of cells dying. At certain time points, such as after 30 minutes of incubation, clearly there are much less colonies present on the plates without any uracil compared to the other concentrations. Yet, this was not consistent at all time points. There is more of a correlation between the incubation times, such that as the incubation time increases, more cells tend to die. These results will help develop negative chemical complementation as a tool for protein engineering. This protocol eliminates around 90% of the undesired variants and may be further optimized. The addition of negative chemical complementation to the protein engineering toolbox allows the creation of a novel method for eliminating false positives present in most assays.

8.5 Summary of Chemical Complementation for Protein Engineering

The field of protein engineering is rapidly emerging as a method for engineering proteins with novel functions with applications in medicine and biotechnology. Chemical complementation provides a general system for evaluating libraries of nuclear receptors that have been engineered to bind a small molecule. This system utilizes the power of genetic selection to allow only the functional variants to survive, eliminating time, materials and effort in evaluating non-functional library members. Negative chemical complementation adds a new dimension to this system for protein engineering, providing

a method of eliminating receptors that is able to bind and activate in response to endogenous ligands or undesired ligands. Together, these two systems provide protein engineers a new system for creating large protein libraries with the added benefits of less effort and time in reaching the desired goal.

8.6 Materials and Methods

Ligands

9-*cis* retinoic acid (MW=300.44 g/mol) was purchased from ICN Biomedicals (Aurora, OH). Paclitaxel (Taxol®) (MW=853.9 g/mol) was purchased from LKT Laboratories (St.Paul, MN). Resveratrol (MW= 228 g/mol), 6-amino penicillanic acid (6-APA) (MW= 216.3 g/mol), Penicillin G (PEN G) (MW= 372.5 g/mol), and Naficillin (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 (4:1v/v) and stored at 4°C.

PXR Background Cassette

A pGBDPXR vector was created with a KpnI site inserted in the middle of the PXR gene using site-directed mutagenesis (Stratagene, USA). The PCR samples were treated with *DpnI* and transformed into Z-competent XL1-Blue cells (Zymo Research, Orange, CA) and transformants were selected based on restriction enzyme analysis. The plasmid was name pGBDPXRKpnI. PCR primers were ordered with KpnI sites at the ends of the primer to amplify a section of the Gal4 activation domain gene. The pGAD10BAACTR gene was placed in a PCR tube with these primers, 10X *Pfu* Buffer, dNTPs, and *Pfu* polymerase and using the standard protocol (95 °C 5 minutes, 95 °C 1

minute, 55 °C 1 minute, 72 °C 3 minutes with 25 cycles) the fragment from the Gal4 activation domain was amplified. This band was purified using the Gel Recovery DNA Kit (Zymo Research, Orange, CA) and this fragment, along with the pGBDPXRKpnI plasmid were digested with KpnI. The vector cassette was treated with 1 μL of CIP, and incubated at 37 °C for 30 minutes. Both the insert and vector cassette were purified using the DNA Clean and Concentrator Kit (Zymo Research, Orange, CA), and 1 μg of the vector cassette and 3 μg of the insert cassette were ligated together. The ligations were transformed into Z-competent XL1-Blue cells and transformants were sequenced. The plasmid was names pGBDPXRBkgd.

PXR Insert Cassette

Oligonucleotides were ordered from MWG Biotech (High Point, NC) with overlapping ends as shown in Table 8.2. 5 μg of each oligonucleotide pair with overlapping ends (i.e. oligonucleotide 1 and 2, 3 and 4, and so forth) were mixed with 100mM NaCL and TE to a total volume of 50 μL into a tube and heated at 95 °C for 5 minutes using the Thermomixer (Eppendorf, CA). After 5 minutes, the temperature was lowered to 10 °C. to allow annealing to occur Once the temperature reached 10 °C, the annealed mixture was combined with 20 μL of 10X EcoPol Buffer, 0.5 μL 100mM dNTPs, 2 μL of *Klenow* polymerase (New England Biolabs, MA) and with distilled water to a total volume of 200 μL. This mixture was incubated at 22 °C for 45 minutes, and then heat inactivated at 75 °C for 20 minutes. The fragments were purified using BioSpin-6 columns (BioRad, CA), and the purified fragments were combined in a PCR tube with 10X *Pfu* Buffer, dNTPs, and *Pfu* polymerase and excess of the end primers, and

with using the following protocol (95 °C 5 minutes, 95 °C 1 minute, 60 °C 1 minute, 72 °C 3 minutes repeat 15 cycles, 95 °C 1 minute, 50 °C 1 minute, 72 °C 3 minutes repeat 15 cycles) the complete insert cassette was created. Once the insert cassette was made, 30 PCR tubes were set-up with 10X *Pfu* Buffer, dNTPs, and *Pfu* polymerase and excess of the end primers, and subjected to 30 rounds of PCR to make more insert cassette.

PXR Transformation

The vector cassette, 1 μg, and randomized insert cassettes, 27 μg, were transformed into PJ69-4A cells, with the pGAD10BAACTR plasmid (this plasmid expresses the Gal4 activation domain fused to the ACTR gene). The vector and insert cassettes were transformed into yeast using the large-scale transformation protocol [169]. Transformants were plated onto selective media described in Chapters 3 and 4.

RXR Background Cassette

This plasmid is described in Schwimmer et al. [13].

RXR Insert Cassette

Oligonucleotides were ordered from MWG Biotech (High Point, NC) with overlapping ends as shown in Table 8.4 and these oligonucleotides were HPLC purified. 2 μg of each oligonucleotide pair with overlapping ends (i.e. oligonucleotide 1 and 2, 3 and 4, and so forth) were mixed with 100mM NaCl and TE to a total volume of 50 μL into a tube and heated at 95 °C for 5 minutes using the Thermomixer (Eppendorf, CA). After 5 minutes, the temperature was lowered to 10 °C. to allow annealing to occur.

Once the temperature reached 10 °C, the annealed mixture was combined with 20 μL of 10X EcoPol Buffer, 0.5 μL 100mM dNTPs, 2 μL of *Klenow* polymerase (New England Biolabs, MA) and with distilled water to a total volume of 200 μL. This mixture was incubated at 22 °C for 45 minutes, and then heat inactivated at 75 °C for 20 minutes. The fragments were purified using BioSpin-6 columns (BioRad, CA), and the purified fragments were combined in a PCR tube with 10X *Pfu* Buffer, dNTPs, and *Pfu* polymerase and excess of the end primers, and with using the following protocol (95 °C 5 minutes, 95 °C 1 minute, 72 °C 1 minute, 72 °C 3 minutes repeat 15 cycles, 95 °C 1 minute, 62 °C 1 minute, 72 °C 3 minutes repeat 15 cycles) the complete insert cassette was created. Once the insert cassette was made, 30 PCR tubes were set-up with 10X *Pfu* Buffer, dNTPs, and *Pfu* polymerase and excess of the end primers, and subjected to 30 rounds of PCR (95 °C 1 minute, 72 °C 1 minute, 72 °C 3 minutes) to make more insert cassette. These cassettes were purified with the Gel DNA Recovery Kit (Zymo Research, Orange, CA). This procedure was repeated for all the oligonucleotides, except oligonucleotide 7, which was created by PCR. PCR primers with overlapping ends to Oligonucleotides 6 and 8 were ordered, and these primers along with 10X *Pfu* Buffer, dNTPs, and *Pfu* polymerase were placed in a PCR tube for amplification with the following PCR program (95 °C 5 minutes, 95 °C 1 minute, 52 °C 1 minute, 72 °C 3 minutes, repeat 25 times). This oligonucleotide was purified with the Gel DNA recovery Kit (Zymo Research, Orange, CA), and added to the PCR mixture to create the entire insert cassette.

RXR Transformation

The background plasmid was digested with *SacI* and *SpeI* to eliminate the stop codons. This linear vector cassette was then treated with 1 μL of CIP, and incubated at 37 °C for 30 minutes. 25 μg of the insert cassettes amplified through PCR and 2 μg of the linearized vector cassette were transformed into the yeast cells containing the ACTR plasmid using the 30X LiAc transformation protocol [106]. Transformants were spread onto two types of selective plates: leucine and tryptophan selective media and adenine selective media as described in Chapters 3 and 4.

Negative Chemical Complementation and FOA incubation times

BAPJ69 was transformed with 1.5 μg of the pCL1 plasmid using the standard 30 X LiAc transformation protocol [106]. The cells were subjected to heat shock and after the heat shock, the cells were pelleted and resuspended in 2 mL FOA solution. The cells were grown for various time points and at each time point were plated onto selective plates.

Negative Chemical Complementation and FOA and Uracil Concentrations

BAPJ69 was transformed with 1.5 μg of the pCL1 plasmid using the standard 30 X LiAc transformation protocol [106]. The cells were subjected to heat shock and after the heat shock, the cells were pelleted and resuspended in 2 mL FOA solution with varying concentrations of uracil. Each concentration was grown to a certain time point, and at each time point were plated onto selective plates.

Table 8.7. Negative chemical complementation in protein engineering: varying incubation times and the amount of uracil in the FOA stock.

Incubation Times (minutes)	Amount of Uracil in FOA	Number of Colonies *
0	20 μL	160
	15 μL	280
	10 μL	240
	5 μL	200
	0 μL	240
30	20 μL	80
	15 μL	120
	10 μL	80
	5 μL	80
	0 μL	28
60	20 μL	50
	15 μL	40
	10 μL	40
	5 μL	55
	0 μL	51
90	20 μL	37
	15 μL	51
	10 μL	45
	5 μL	34
	0 μL	53
120	20 μL	36
	15 μL	28
	10 μL	20
	5 μL	5
	0 μL	9

*These numbers are averages based on two sets of experiments where the number of colonies at each time point was averaged.

Chapter 9

Applications of Chemical Complementation: Deciphering and Assembling Biosynthetic Pathways

9.1 Deciphering metabolic pathways with chemical complementation

Natural products are extremely useful compounds in a variety of applications, particularly in the field of medicine. Various natural products have been utilized as chemotherapy agents and as drugs for a number of other pathologies, such as heart disease and towards viral pathogens [170-173].

The use of natural products for therapy is extremely beneficial, yet ample amounts of these natural products need to be available to become a potential therapy for the disease of interest. As the demand for these compounds increases, so do environmental concerns for the elimination of the natural source. The prime example of this case is the natural product paclitaxel, the chief source of which is the Pacific Yew tree [174-177]. Based on data from the U.S. Forest Service from 1992, 36,000 yew trees are required to provide 327,200 kg of bark (about 9 kg/tree) from which 24 kg of paclitaxel can be extracted. Approximately, 2.08 g of paclitaxel is needed per person for treatment or 3.15 trees per person. Others have claimed this number can be as high as 10 trees per person depending on the size of the trees [178]. Therefore, paclitaxel chemotherapy is extremely expensive and involves the possible elimination of the yew trees, causing a big environmental concern and elimination of the natural product source for future patients.

In the case of paclitaxel, a number of alternative chemical synthetic methods of paclitaxel production have been developed and are currently being used. However, for more newly identified natural products with medicinal applications, synthetic routes are yet to be identified. An alternative approach for the mass production of the natural products is the identification of the biosynthetic pathways used to produce these natural compounds. Once the genes encoding the enzymes that produce the compounds have been identified, the genes can be introduced into a heterologous organism, such as bacterium or yeast. These organisms can serve as a "factory" for an alternative means of obtaining large amounts of these compounds for medicinal applications. For example, epithilones are potential anticancer drugs that stabilize microtubules in a manner similar to paclitaxel and are produced from the myxobacterium, *Sorangium cellulosum*. Researchers at Kosan Biosciences (CA) have been able to express the biosynthetic pathway from *Sorangium cellulosum*, its natural host organism, to make epithilone A and B in a heterologous host, *Myxococcus xanthus*, providing a more efficient organism for producing these two compounds in comparison to their natural host [179-181]

Chemical complementation can play a vital role in the identification of biosynthetic genes of a desired natural product (Figures 9.1 and 9.2). First, a nuclear receptor is engineered towards the target natural product, such that this receptor can only bind and activate in response to the natural product of interest. The engineered receptor must also not activate in response to any of the substrates or precursors that are placed in the yeast medium for production of the natural product. Once a receptor is engineered to bind the natural product, yeast expression plasmids are made with the enzymes involved in making the natural product. These expression plasmids will be introduced into yeast

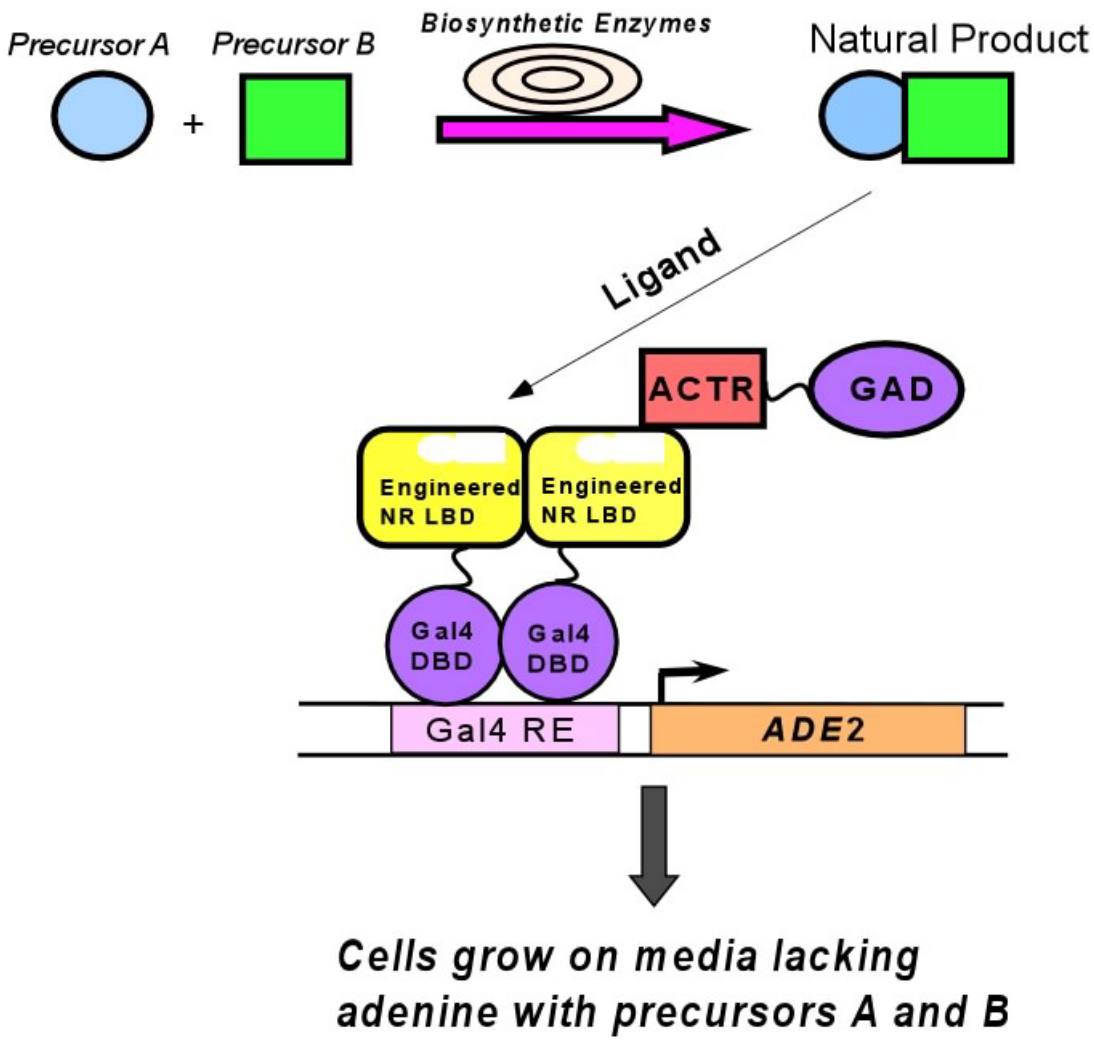


Figure 9.1. Chemical Complementation in producing natural products.

along with the engineered receptor and its coactivator (Figure 9.1). The yeast cells will be plated onto selective medium (adenine, uracil, or histidine) containing any necessary substrates or precursors. The yeast will express the enzymes required to produce the natural product from expression plasmids containing them, and taking precursors (if present) or endogenous yeast metabolites that can be used as substrates for these enzymes and converting them into the natural product. The natural product will then bind to the engineered receptor and activate transcription of the selective marker (Figure 9.1). Therefore, the survival of the yeast cells depends on their ability to produce the natural product.

The method mentioned above should be extremely useful for creating heterologous hosts as an alternative approach to obtaining vast amounts of these natural products. To make the heterologous host ideal for industrial scale production of these natural products, the biosynthetic genes can be integrated into a heterologous host's genome. Though the practice of developing heterologous hosts for producing natural products is not new, chemical complementation brings an added advantage to maintain the production of these compounds in the heterologous host. Genes integrated into chromosomes can be lost or modified, such that they no longer are functional (Figure 9.1). With chemical complementation only the cells that have not lost the integrated genes, and that produce the natural product will survive. Cells that have undergone modifications or have otherwise lost the genes will die. Therefore, chemical complementation can serve as a method of controlling the fermentation batch for the mass production of these natural products.

In most cases, the complete biosynthetic pathways of natural products have not yet been discovered. Much time and effort is being put forth for discovering the biosynthetic pathways of multiple natural products, such that these enzymes could be utilized to develop large amounts of these products. Chemical complementation can play an extremely useful role in discovering biosynthetic genes and deciphering the biosynthetic pathways involved in making these natural products. A nuclear receptor is engineered towards the natural product of interest. In cases, where the biosynthetic pathway is unknown, a cDNA library will be made from the organism that produces the natural product (Figure 9.2). This cDNA library will then be fragmented and cloned into yeast expression plasmids. These yeast expression plasmids containing fragments of the cDNA library will be introduced into yeast cells along with the receptor and coactivator (Figure 9.2).

The yeast cells will be plated onto selective medium (adenine, uracil, or histidine) containing any necessary substrates or precursors. In this case, the yeast cell that expresses the plasmid containing the cDNA fragment, which codes for the biosynthetic enzyme(s) will grow. This is because that yeast cell will be able to convert the precursors, endogenous or present in the media, into the natural product, which binds to the engineered receptor and the selective marker expresses (Figure 9.2). The yeast expression plasmid containing the cDNA fragment can then be rescued, sequenced and used for further analysis. Thus, chemical complementation provides a system in which biosynthetic pathways can be discovered and maintained in a heterologous host. This approach provides the pressure of genetic selection that will force these organisms to produce the natural products or express proteins, which are foreign to the yeast. The

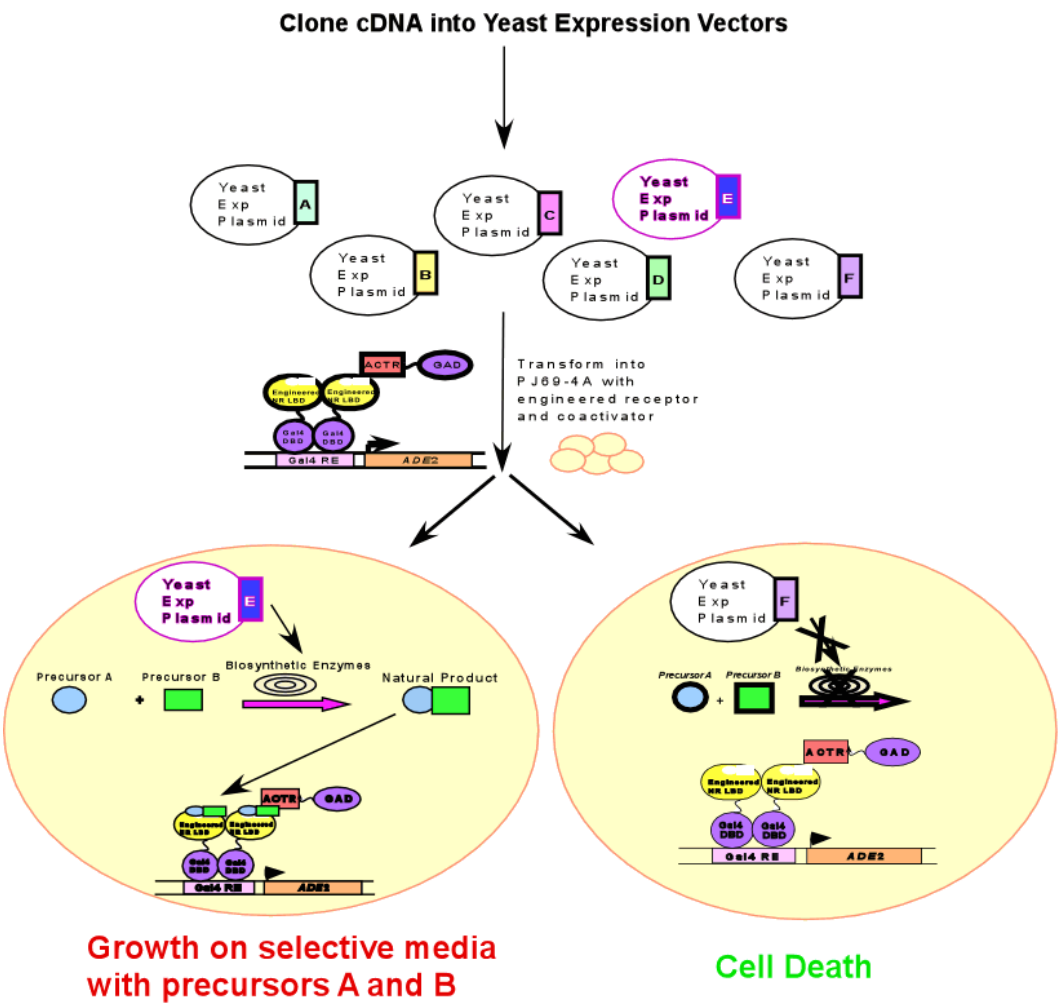


Figure 9.2. Chemical complementation for deciphering and assembling biosynthetic Pathways with cDNA Libraries.

added pressure genetic selection provides allows researchers to manipulate these organisms to express proteins of interest and create small molecules.

9.2 Chemical Complementation and Resveratrol Biosynthesis

Our initial attempt to develop chemical complementation for deciphering biosynthetic pathways was to complete the biosynthetic pathway for paclitaxel. Rodney Croteau and co-workers at the University of Washington have been able to piece together the majority of the biosynthetic pathway for this compound, yet the entire pathway is not complete. It is arguable that paclitaxel was an easy target because the human pregnane X receptor (PXR or SXR) was already shown to bind and activate in response to paclitaxel. However, in 2nd generation chemical complementation, PXR was determined to be constitutively active. An attempt was made to create a library of PXR variants, which were no longer constitutively active and were activate in response to paclitaxel. However, the library produced only constitutively active receptors, and proved to be unsuccessful. Since paclitaxel is a fairly larger compound (MW= 854 g/mol), the ability to engineer a receptor that would also not be constitutively active seemed challenging. Thus, we decided to change our target to a small molecule whose biosynthetic pathway has already been determined, as a proof of principle that chemical complementation can indeed be used to decipher biosynthetic pathways. We decided to use resveratrol as our target compound.

Resveratrol is a small molecule found in many dietary compounds and red wine, which has multiple mechanisms in the prevention of many human diseases, including inflammation, atheroscelorosis, and cardiovascular disease (Figure 9.3). This compound

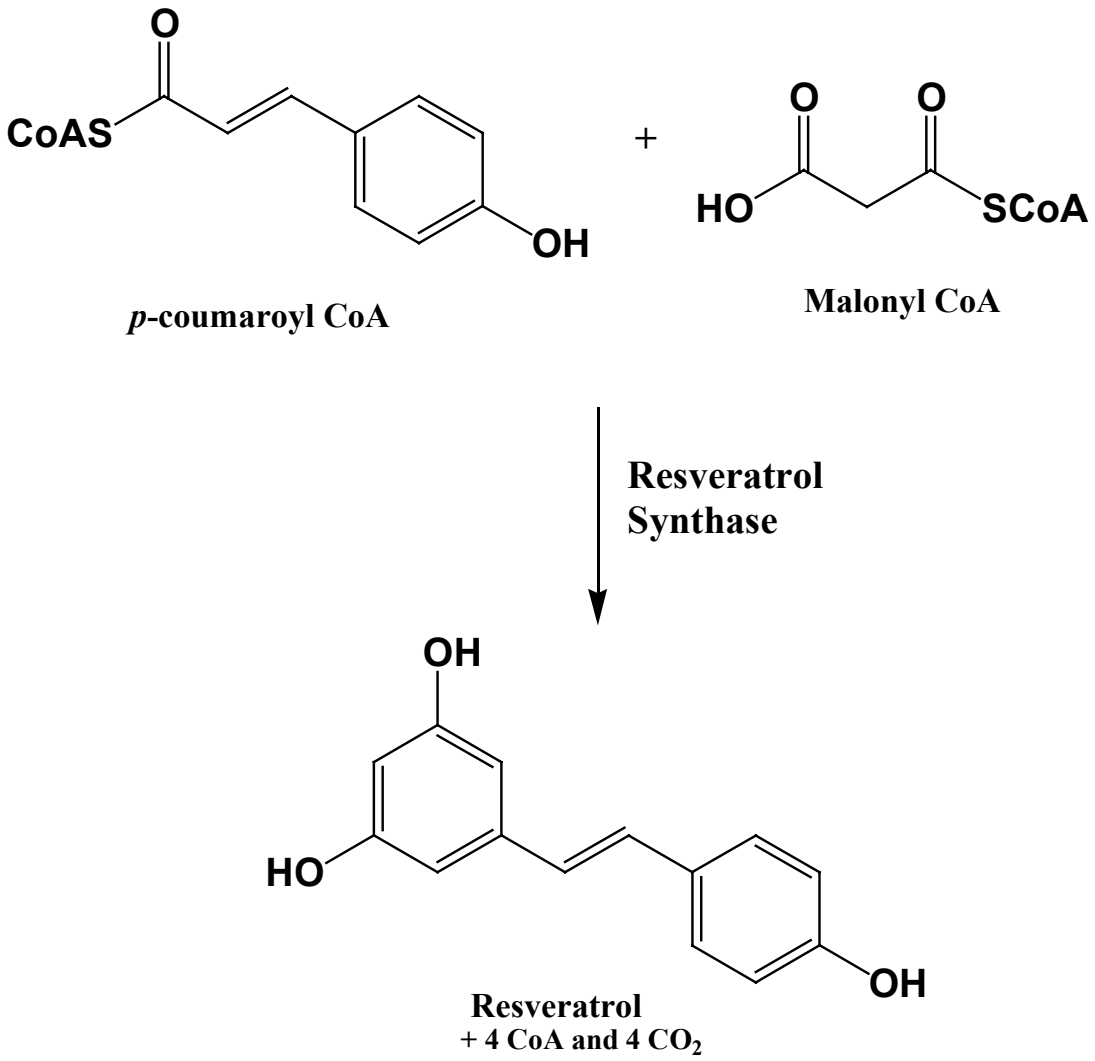


Figure 9.3. Resveratrol biosynthetic pathway.

has also been shown to have anticancer activities and has been shown to possibly have a remarkable effect as a chemopreventive and chemotherapeutic agent. Resveratrol is thoroughly discussed in Chapter 8. The biosynthetic pathway of resveratrol has been determined. Resveratrol is produced by resveratrol synthase, which synthesizes the backbone of stilbene phytoalexins (Figure 9.3) [168,182]. Thus enzyme is closely related to chalcone synthase, a plant-specific enzyme that synthesizes naringenin chalcone, a precursor of flavanoids [183-185]. Resveratrol is produced from the two substrates: coumaroyl CoA and malonyl CoA, which is eventually converted to resveratrol via the resveratrol synthase (Figure 9.3).

With the biosynthetic pathway of resveratrol already determined, this small molecule is a good target molecule as a proof of principle for chemical complementation as a tool for deciphering biosynthetic pathways. Furthermore, if the yeast cells are able to produce resveratrol with the help of chemical complementation, this system also could serve as an alternative approach to obtaining resveratrol.

9.3 Results

Step One: Engineer a receptor to bind resveratrol.

A library of RXR variants was engineered to bind and activate in response to resveratrol (Chapter 8). However, none of the variants activated in response to resveratrol, but were able to bind other synthetic ligands, such as LG335 and γOPBA. During the creation of the RXR library, the rat estrogen receptor β (rat ERβ) was cloned into a yeast expression vector containing the Gal4 DNA binding domain fused to the rat

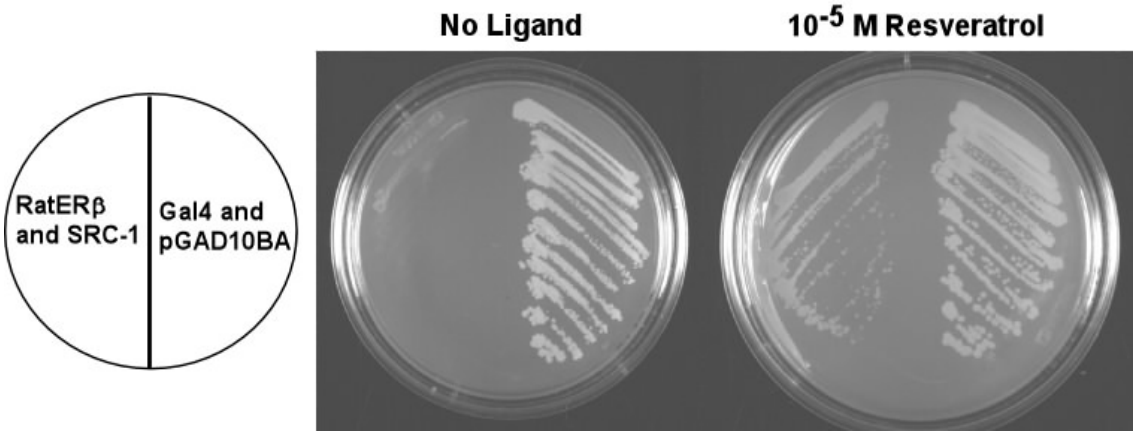


Figure 9.4. Rat estrogen receptor β and resveratrol on adenine selective plates.

ERβ LBD (pGBDratERβ) (discussed in Chapter 7). This receptor was tested with chemical complementation with its coactivator SRC-1, and was found induce ligand-activated growth in the presence of its agonists, estradiol and hexestrol, and not constitutively active (Chapter 7).

Resveratrol is a known weak agonist for estrogen receptors (3-10 μM), and was tested in chemical complementation on selective plates and in liquid quantitation assays. Yeast cells containing pGBDratERβ and pGAD10BASRC-1 (the SRC-1 gene fused to the Gal4 activation domain) were streaked onto adenine selective plates (-Ade –Leu -Trp) with and without 10^{-5} M resveratrol. The leucine and tryptophan selections complement the two expression plasmids, pGAD10BASRC-1 and pGBDratERβ respectively. Yeast cells containing the *holo* Gal4 protein in the plasmid pGBT9Gal4, and the pGAD10BA plasmid (for the leucine deficiency) were also streaked as a positive control. These plates were incubated at 30 °C for two days.

Resveratrol-activated growth was observed with rat ERβ and SRC-1 and no growth was observed on the plate without ligand (Figure 9.4). This was promising, so a dose response growth assay was performed to determine the lowest ligand concentration that allows the cells to survive. Yeast cells containing rat ERβ and SRC-1 were inoculated in media (-Leu –Trp) selecting for both expression plasmids. Gal4 was used as a control. After an overnight growth period, these cells were placed into 96-well microtiter plates containing histidine selective media with 10 mM concentration of 3-AT, 3-amino 1,2 4-triazole, an inhibitor of the *HIS3* enzyme. 3-AT was used because the 96-well microtiter plates are made from polystyrene, and polystyrene has been shown to

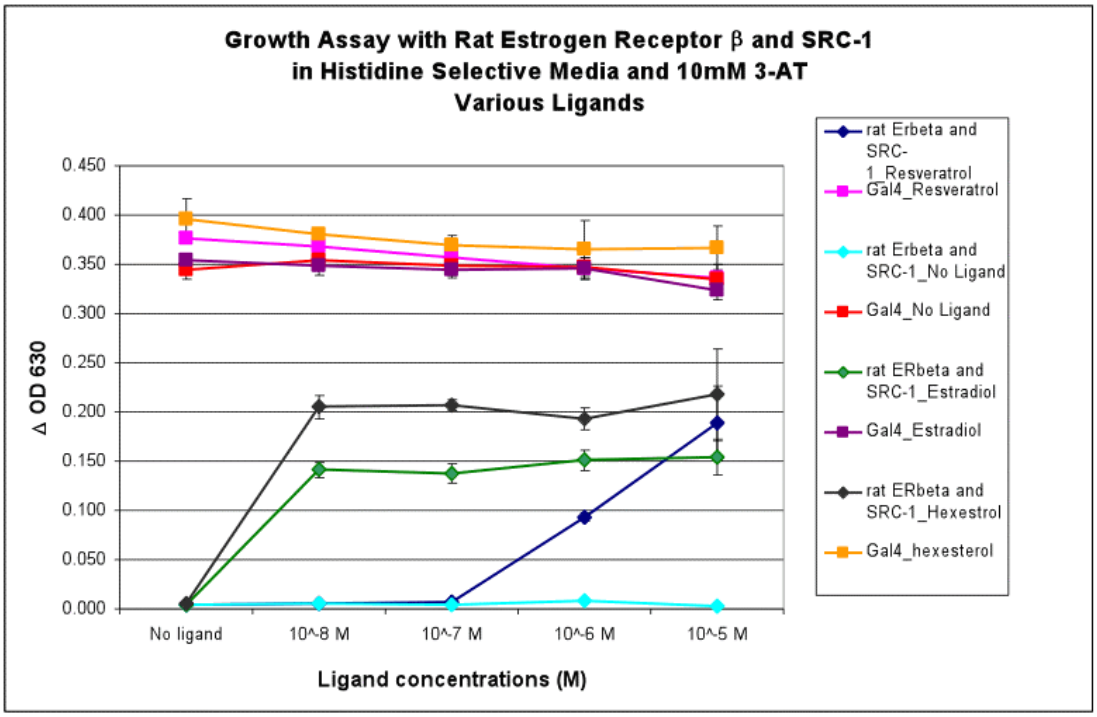


Figure 9.5. Rat estrogen receptor β in a quantitation assay with estradiol, hexesterol and resveratrol.

have estrogenic effects, and 10 mM 3-AT eliminates any of the background growth due to the polystyrene. Concentrations of estradiol, hexestrol, and resveratrol ranging from 10^{-5} M to 10^{-8} M were added to the wells. These plates were grown for 48 hours with shaking at 30 °C.

Ligand-activated growth was observed in the wells containing ligand as expected and no growth was observed without the presence of ligand (Figure 9.5). The EC_{50} values for both estradiol and hexestrol are much lower than the EC_{50} value for resveratrol (~50 nM compared to ~1 μM). This is expected because in mammalian cell culture assays, resveratrol is known to be a weaker agonist of the estrogen receptor compared to the potent agonists estradiol and hexestrol. Nevertheless, ligand-activated growth with rat ERβ and resveratrol is observed on both solid media (plates) and in the liquid quantitation assay. This is a step forward in the use of resveratrol for using chemical complementation in deciphering biosynthetic pathways, because we have obtained the receptor that binds and activates in response to the natural product resveratrol.

One potential problem that may need to be addressed is that the sensitivity of rat ERβ towards resveratrol. If the concentration of resveratrol produced from the biosynthetic pathway is in the nanomolar range, this system will not work because a high enough concentration of ligand will not be present for the receptor to be activated. In a recent publication by Becker and co-workers, a yeast strain was metabolically engineered to produce resveratrol [182]. These researchers were able to introduce the resveratrol biosynthetic enzymes required to produce the compound, along with enzymes to produce the precursor p-coumaric acid [182]. These genes were transformed into yeast and through high purification liquid chromatography (HPLC) and mass spectrometry analysis

Becker and co-workers were able to detect resveratrol production from the organic extracts of the yeast cells [182]. Table 9.1 shows the amount of resveratrol detected in the different cultures of the yeast strains containing the resveratrol biosynthetic enzymes [182]. The amount of resveratrol detected was not extremely high, where the highest concentration of resveratrol detected is in the nanomolar range [182]. Therefore, the sensitivity of rat ERβ will have to be increased through mutagenesis protocols, where libraries of these receptors can be generated with a higher potency for resveratrol. This work is currently in progress.

Step Two: Testing Precursors with Rat ERβ

With the identification of the rat ERβ as a potential ligand for binding resveratrol, the next step is to investigate whether any precursors or substrates not required for biosynthesis to occur also activate ERβ. In the creation of the yeast strain metabolically engineered to produce resveratrol by Becker and co-workers, they had found that p-coumaric acid needed to be added to the medium to produce the substrate p-coumaroyl-CoA. In the resveratrol biosynthetic pathway, two substrates are required: malonyl CoA and p-coumaroyl CoA. Malonyl CoA is present in the yeast and is actively involved in fatty acid biosynthesis pathways. The other substrate, p-coumaroyl CoA has been shown to accumulate in yeast, but Becker et al. had felt that this substrate might be the limiting factor. Therefore, p-coumaric acid was added to the cultures they used and introduced an enzyme able to convert the p-coumaric acid to the substrate p-coumaroyl CoA [182]. Since p-coumaric acid would be added to the media to produce resveratrol, we had to

Table 9.1. Amount of resveratrol detected in yeast cultures from the engineered metabolic strain created by Becker et al. *Adapted from Becker et al.; FEMS Yeast Research, 2003*

Yeast Culture	Resveratrol produced ($\mu g\ L^{-1}$ culture volume)	Resveratrol produced (nM)
S. cerevisiae FY23 (control strain)	0	0
Recombinant Strain 1	0.61	2.7
Recombinant Strain 2	0.84	3.7
Recombinant Strain 3	1.30	5.7
Recombinant Strain 4	0.32	1.4
Recombinant Strain 5	1.45	6.3
Recombinant Strain 6	0.83	3.6
Recombinant Strain 7	1.38	6.0

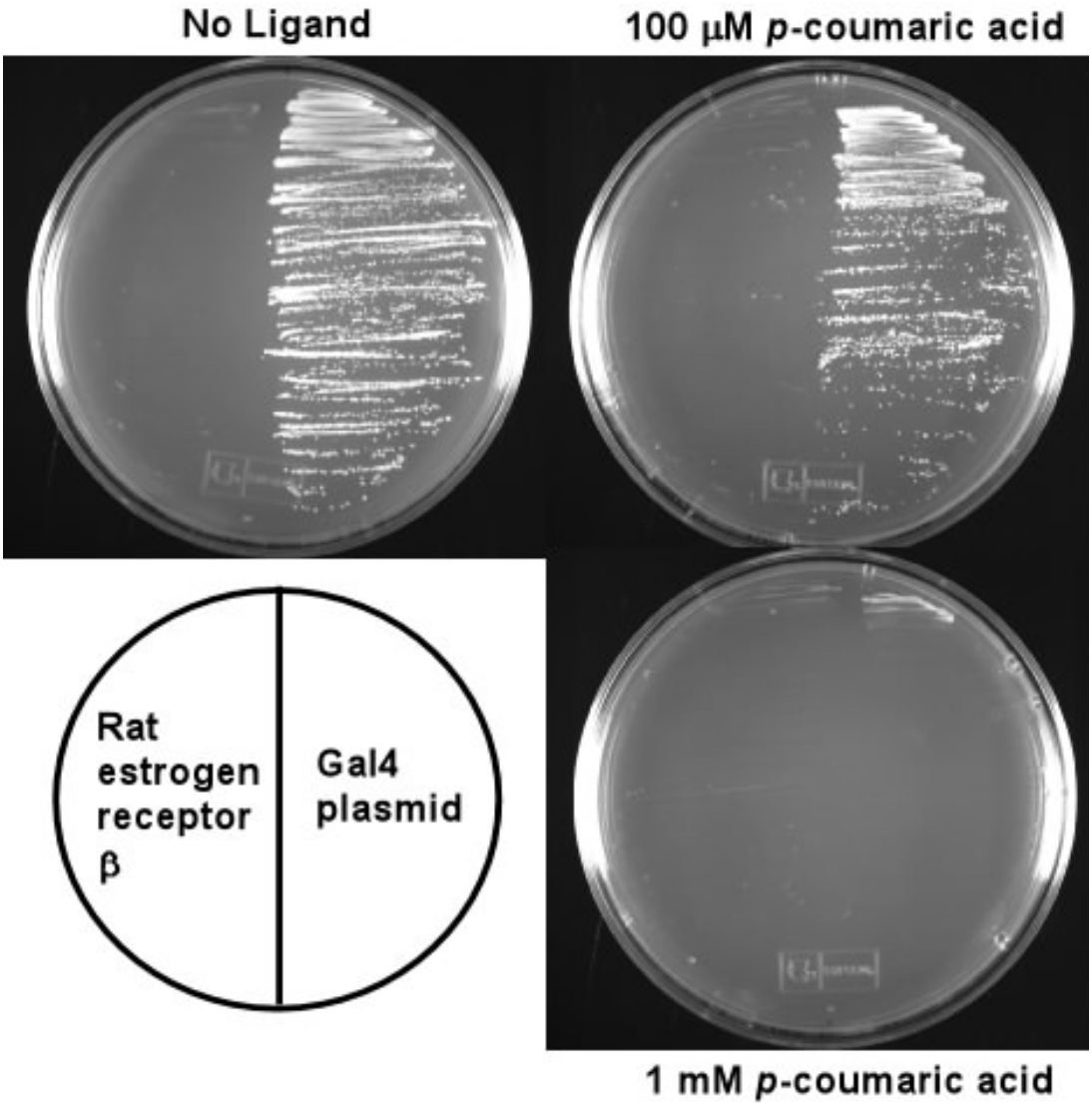


Figure 9.6. Rat estrogen receptor β and *p*-coumaric acid.

confirm that p-coumaric acid was not able to activate rat ERβ, so that the growth seen with chemical complementation in deciphering these pathways are only due to the production of resveratrol. Adenine selective plates were made with 10 μM, 100 μM and 1mM concentrations of p-coumaric acid. Growth was not observed on any of the sectors containing the rat ERβ and at a concentration of 1mM, the sector of the plate containing Gal4 is not able to grow either, indicating some lethal effect at that concentration of p-coumaric acid (Figure 9.6) Thus, the precursors are not able to bind and activate rat ERβ in response to the precursor, p-coumaric acid, and any ligand-activated growth seen with chemical complementation and the resveratrol biosynthetic pathway should be due to the production of resveratrol.

9.4 Future Work

Thus far, we have been able to bring together pieces of the chemical complementation system for deciphering biosynthetic pathways, in this case for resveratrol. We are able to use the rat ERβ for its ability to bind resveratrol, although attempts are currently being made to increase the sensitivity of the receptor towards resveratrol. We have shown that the precursor, p-coumaric acid, is not able to activate rat ERβ, therefore when ligand activated growth is observed with the system, this growth is due only to the production of resveratrol.

The next step is to clone the resveratrol synthase gene into a yeast expression plasmid, along with the enzyme that is able to convert p-coumaric acid into p-coumaroyl CoA. Once these clones are made, they can be introduced into yeast along with the ERβ and SRC-1 for detection of resveratrol production (Figure 9.7). If this works, chemical

complementation will provide a powerful system for discovering and assembling biosynthetic genes, as well as providing a system for fermenting yeast cells to produce ample amounts of scarce natural products.

9.5 Materials and Methods

Ligands

Resveratrol (MW= 228 g/mol) and *p*-coumaric acid (MW= 164.2 g/mol) were purchased from Sigma (St. Louis, MO). 10 mM stocks of each ligand were dissolved in 80% ethanol: 20% DMSO (4:1v/v) and stored at 4°C.

Yeast selective plates and quantitation data

Adenine selective plates with and without ligand were made as described in Chapter 4, and the quantitation was set-up as described in Chapter 5. pGBDratERβ and pGAD10BASRC-1 are described in Chapters 7 and 4 respectively.

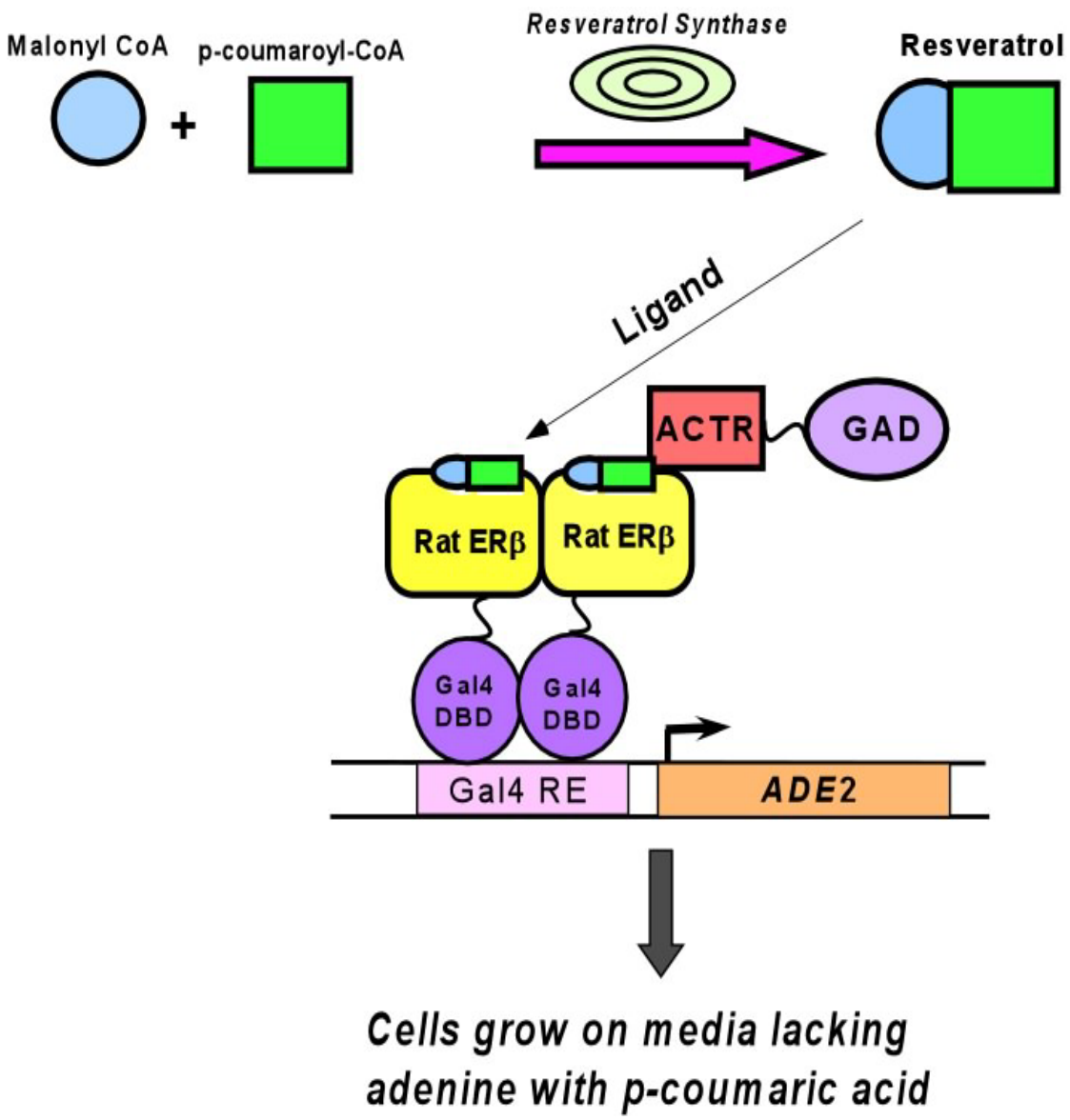


Figure 9.7. Chemical Complementation for the production of resveratrol.

CONCLUSIONS

Chemical complementation is a general system that links the presence of a small molecule to genetic selection. In our chemical complementation system, the interaction of a nuclear receptor and ligand are linked to yeast survival. This three-component system was initially developed using the retinoid X receptor and extended to other members of the nuclear receptor family.

Nuclear receptors are implicated in a number of diseases, and are currently targets for ~10% of commonly prescribed drugs. Using chemical complementation and negative chemical complementation, a high-throughput assay for drug discovery is possible, where both agonists and antagonists can be discovered as potential therapeutics targeting nuclear receptor pathologies. Drug discovery assays may be extended to enzyme targets by engineering receptors that activate transcription in response to the small molecule product of the enzyme-catalyzed reaction.

Chemical complementation offers a general method of engineering receptors that activate transcription in response to arbitrary small molecules for protein engineering applications. The discovery of these engineered proteins-ligand pairs can serve as candidates for small molecule based gene regulation, with applications gene therapy, and can be extended towards other applications, such as engineering enzymes with enhanced or novel functions, and for use in biosensors. Finally, with the discovery of more and more natural products as potential drug candidates towards a number of pathologies, chemical complementation can serve as a tool for deciphering the biosynthetic pathways

of these natural products, and as tool for creating heterologous hosts for the production of natural products.

Throughout this dissertation, chemical complementation has been developed and demonstrated, as well as a high-throughput quantitation assay of chemical complementation, and negative chemical complementation. Chemical complementation has been shown and used for protein engineering, specifically engineering receptors with the ability to bind and activate in response to non-natural small molecule ligands. The preliminary steps for the use of chemical complementation for deciphering and assembling biosynthetic pathways have also been demonstrated. In the future, other applications of chemical complementation can be further developed, allowing the creation of a versatile, generalizable genetic selection system for evaluating protein-small molecule interactions.

REFERENCES

1. Woycechowsky KJ, Hilvert D: **Deciphering enzymes - Genetic selection as a probe of structure and mechanism**. *European Journal of Biochemistry* 2004, **271**:1630-1637.

2. Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, Botstein D: **Characteristic genome rearrangements in experimental evolution of Saccharomyces cerevisiae**. *Proceedings of the National Academy of Sciences of the United States of America* 2002, **99**:16144-16149.

3. Omalley K, Douglas MG: **Selection and Characterization of Nuclear Genes-Coding Mitochondrial Proteins - Genetic Complementation of Yeast Pet Mutants**. *Methods in Enzymology* 1983, **97**:344-355.

4. Omalley K, Pratt P, Robertson J, Lilly M, Douglas MG: **Selection of the Nuclear Gene for the Mitochondrial Adenine-Nucleotide Translocator by Genetic Complementation of the Op1 Mutation in Yeast**. *Journal of Biological Chemistry* 1982, **257**:2097-2103.

5. Miller AM, Savinelli EA, Couture SM, Hannigan GM, Han ZY, Selden RF, Treco DA: **Recombination Walking - Genetic Selection of Clones from Pooled Libraries of Yeast Artificial Chromosomes by Homologous Recombination**. *Proceedings of the National Academy of Sciences of the United States of America* 1993, **90**:8118-8122.

6. Gellerfors PL, Saltzgabermuller J, Douglas MG: **Selection by Genetic Complementation and Characterization of the Gene Coding for the Yeast Porphobilinogen Deaminase**. *Biochemical Journal* 1986, **240**:673-677.

7. Lin HN, Cornish VW: **Screening and selection methods for large-scale analysis of protein function**. *Angewandte Chemie-International Edition* 2002, **41**:4403-4425.

8. Taylor SV, Kast P, Hilvert D: **Investigating and engineering enzymes by genetic selection**. *Angewandte Chemie-International Edition* 2001, **40**:3311-3335.

9. Tao HY, Cornish VW: **Milestones in directed enzyme evolution**. *Current Opinion in Chemical Biology* 2002, **6**:858-864.

10. Serebriiskii IG, Fang R, Latypova E, Hopkins R, Vinson C, Joung JK, Golemis EA: **A Combined Yeast/Bacteria Two-hybrid System: Development and Evaluation**. *Mol Cell Proteomics* 2005, **4**:819-826.

11. Toby GG, Golemis EA: **Using the yeast interaction trap and other two-hybrid-based approaches to study protein-protein interactions**. *Methods* 2001, **24**:201-217.

12. Azizi B, Chang EI, Doyle DF: **Chemical complementation: small-molecule-based genetic selection in yeast**. *Biochemical and Biophysical Research Communications* 2003, **306**:774-780.

13. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL, Doyle DF: **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**:14707-14712.

14. Norman TC, Smith DL, Sorger PK, Drees BL, O'Rourke SM, Hughes TR, Roberts CJ, Friend SH, Fields S, Murray AW: **Genetic selection of peptide inhibitors of biological pathways**. *Science* 1999, **285**:591-595.

15. Drees BL: **Progress and variations in two-hybrid and three-hybrid technologies**. *Current Opinion in Chemical Biology* 1999, **3**:64-70.

16. McNabb DS, Guarente L: **Genetic and biochemical probes for protein-protein interactions**. *Current Opinion in Biotechnology* 1996, **7**:554-559.

17. Warbrick E: **Two's company, three's a crowd: The yeast two hybrid system for mapping molecular interactions**. *Structure* 1997, **5**:13-17.

18. Youderian P, Vershon A, Bouvier S, Sauer RT, Susskind MM: **Changing the DNA-Binding Specificity of a Repressor**. *Cell* 1983, **35**:777-783.

19. Fields S, Song O: **A novel genetic system to detect protein-protein interactions**. *Nature* 1989, **340**:245-246.

20. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, et al.: **A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae**. *Nature* 2000, **403**:623-627.

21. Uetz P: **Two-hybrid arrays**. *Curr Opin Chem Biol* 2002, **6**:57-62.

22. Licitra EJ, Liu JO: **A three-hybrid system for detecting small ligand-protein receptor interactions**. *Proc Natl Acad Sci U S A* 1996, **93**:12817-12821.

23. Griffith EC, Licitra EJ, Liu JO: **Yeast three-hybrid system for detecting ligand-receptor interactions**. In *Applications of Chimeric Genes and Hybrid Proteins, Pt C*. Edited by; 2000:89-103. Methods in Enzymology, vol 328.]

24. Serebriiskii IG, Golemis EA: **Uses of lacZ to study gene function: evaluation of beta-galactosidase assays employed in the yeast two-hybrid system**. *Anal Biochem* 2000, **285**:1-15.

25. Spencer DM, Wandless TJ, Schreiber SL, Crabtree GR: **Controlling signal transduction with synthetic ligands**. *Science* 1993, **262**:1019-1024.

26. Pollock R, Clackson T: **Dimerizer-regulated gene expression**. *Current Opinion in Biotechnology* 2002, **13**:459-467.

27. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL, Doyle DF: **Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules**. *Proc Natl Acad Sci U S A* 2004, **101**:14707-14712.

28. Peet DJ, Doyle DF, Corey DR, Mangelsdorf DJ: **Engineering novel specificities for ligand-activated transcription in the nuclear hormone receptor RXR**. *Chem.Biol.* 1998, **5**:13-21.

29. Doyle DF, Braasch DA, Jackson LK, Weiss HE, Boehm MF, Mangelsdorf DJ, Corey DR: **Engineering orthogonal ligand-receptor pairs from "near drugs"**. *Journal of the American Chemical Society* 2001, **123**:11367-11371.

30. Abida WM, Carter BT, Althoff EA, Lin H, Cornish VW: **Receptor-dependence of the transcription read-out in a small-molecule three-hybrid system**. *Chembiochem* 2002, **3**:887-895.

31. Muddana SS, Peterson BR: **Facile synthesis of cids: biotinylated estrone oximes efficiently heterodimerize estrogen receptor and streptavidin proteins in yeast three hybrid systems**. *Org Lett* 2004, **6**:1409-1412.

32. Athavankar S, Peterson BR: **Control of gene expression with small molecules: Biotin-mediated acylation of targeted lysine residues in recombinant yeast**. *Chemistry & Biology* 2003, **10**:1245-1253.

33. Althoff EA, Cornish VW: **A bacterial small-molecule three-hybrid system**. *Angew Chem Int Ed Engl* 2002, **41**:2327-2330.

34. Horswill AR, Savinov SN, Benkovic SJ: **A systematic method for identifying small-molecule modulators of protein-protein interactions**. *Proc Natl Acad Sci U S A* 2004, **101**:15591-15596.

35. Ho SN, Biggar SR, Spencer DM, Schreiber SL, Crabtree GR: **Dimeric ligands define a role for transcriptional activation domains in reinitiation**. *Nature* 1996, **382**:822-826.

36. Diver ST, Schreiber SL: **Single-step synthesis of cell-permeable protein dimerizers that activate signal transduction and gene expression**. *Journal of the American Chemical Society* 1997, **119**:5106-5109.

37. Freiberg RA, Ho SN, Khavari PA: **Transcriptional control in keratinocytes and fibroblasts using synthetic ligands**. *J Clin Invest* 1997, **99**:2610-2615.

38. Firestine SM, Salinas F, Nixon AE, Baker SJ, Benkovic SJ: **Using an AraC-based three-hybrid system to detect biocatalysts in vivo**. *Nature Biotechnology* 2000, **18**:544-547.

39. Firestine SM, Salinas F, Benkovic SJ: **Using transcriptional-based systems for in vivo enzyme screening**. *Methods Mol Biol* 2003, **205**:315-327.

40. Baker K, Bleczinski C, Lin H, Salazar-Jimenez G, Sengupta D, Krane S, Cornish VW: **Chemical complementation: a reaction-independent genetic assay for enzyme catalysis**. *Proc Natl Acad Sci U S A* 2002, **99**:16537-16542.

41. de Felipe KS, Carter BT, Althoff EA, Cornish VW: **Correlation between ligand-receptor affinity and the transcription readout in a yeast three-hybrid system**. *Biochemistry* 2004, **43**:10353-10363.

42. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, et al.: **The nuclear receptor superfamily: the second decade**. *Cell* 1995, **83**:835-839.

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

44. Gronemeyer H, Gustafsson JA, Laudet V: **Principles for modulation of the nuclear receptor superfamily**. *Nature Reviews Drug Discovery* 2004, **3**:950-964.

45. Chambon P: **(T)he nuclear receptor superfamily: A personal retrospect on the first two decades**. *Molecular Endocrinology* 2005, **19**:1418-1428.

46. Soprano DR, Qin P, Soprano KJ: **Retinoic acid receptors and cancers**. *Annual Review of Nutrition* 2004, **24**:201-221.

47. McDonnell DP, Vegeto E, Gleeson MAG: **Nuclear Hormone Receptors as Targets for New Drug Discovery**. *Bio-Technology* 1993, **11**:1256-1261.

48. Jordan VC: **Antiestrogens and selective estrogen receptor modulators as multifunctional medicines. 2. Clinical considerations and new agents**. *J Med Chem* 2003, **46**:1081-1111.

49. Krust A, Green S, Argos P, Kumar V, Walter P, Bornert JM, Chambon P: **The chicken oestrogen receptor sequence: homology with v-erbA and the human oestrogen and glucocorticoid receptors**. *Embo J* 1986, **5**:891-897.

50. Warnmark A, Treuter E, Wright APH, Gustafsson JA: **Activation functions 1 and 2 of nuclear receptors: Molecular strategies for transcriptional activation**. *Molecular Endocrinology* 2003, **17**:1901-1909.

51. Folkertsma S, van Noort P, Van Durme J, Joosten HJ, Bettler E, Fleuren W, Oliveira L, Horn F, de Vlieg J, Vriend G: **A family-based approach reveals the function of residues in the nuclear receptor ligand-binding domain**. *Journal of Molecular Biology* 2004, **341**:321-335.

52. Moras D, Gronemeyer H: **The nuclear receptor ligand-binding domain: structure and function**. *Current Opinion in Cell Biology* 1998, **10**:384-391.

53. Benoit G, Malewicz M, Perlmann T: **Digging deep into the pockets of orphan nuclear receptors: insights from structural studies**. *Trends in Cell Biology* 2004, **14**:369-376.

54. Steinmetz ACU, Renaud JP, Moras D: **Binding of ligands and activation of transcription by nuclear receptors**. *Annual Review of Biophysics and Biomolecular Structure* 2001, **30**:329-359.

55. Moras D, Gronemeyer H: **The nuclear receptor ligand-binding domain: structure and function**. *Curr Opin Cell Biol* 1998, **10**:384-391.

56. Chambon P: **A decade of molecular biology of retinoic acid receptors**. *Faseb J* 1996, **10**:940-954.

57. Glass CK, Holloway JM, Devary OV, Rosenfeld MG: **The Thyroid-Hormone Receptor Binds with Opposite Transcriptional Effects to a Common Sequence Motif in Thyroid-Hormone and Estrogen Response Elements**. *Cell* 1988, **54**:313-323.

58. Sap J, Munoz A, Schmitt J, Stunnenberg H, Vennstrom B: **Repression of Transcription Mediated at a Thyroid-Hormone Response Element by the V-Erb-a Oncogene Product**. *Nature* 1989, **340**:242-244.

59. Bourguet W, Ruff M, Chambon P, Gronemeyer H, Moras D: **Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-alpha**. *Nature* 1995, **375**:377-382.

60. Renaud JP, Rochel N, Ruff M, Vivat V, Chambon P, Gronemeyer H, Moras D: **Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid**. *Nature* 1995, **378**:681-689.

61. Wagner RL, Apriletti JW, McGrath ME, West BL, Baxter JD, Fletterick RJ: **A structural role for hormone in the thyroid hormone receptor**. *Nature* 1995, **378**:690-697.

62. Farboud B, Privalsky ML: **Retinoic acid receptor-alpha is stabilized in a repressive state by its C-terminal, isotype-specific F domain**. *Molecular Endocrinology* 2004, **18**:2839-2853.

63. Chen JD, Evans RM: **A Transcriptional Co-Repressor That Interacts with Nuclear Hormone Receptors**. *Nature* 1995, **377**:454-457.

64. Horlein AJ, Naar AM, Heinzel T, Torchia J, Gloss B, Kurokawa R, Ryan A, Kamel Y, Soderstrom M, Glass CK, et al.: **Ligand-Independent Repression by the Thyroid-Hormone Receptor-Mediated by a Nuclear Receptor Co-Repressor**. *Nature* 1995, **377**:397-404.

65. Zamir I, Harding HP, Atkins GB, Horlein A, Glass CK, Rosenfeld MG, Lazar MA: **A nuclear hormone receptor corepressor mediates transcriptional silencing by receptors with distinct repression domains**. *Molecular and Cellular Biology* 1996, **16**:5458-5465.

66. Guenther MG, Barak O, Lazar MA: **The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3**. *Molecular and Cellular Biology* 2001, **21**:6091-6101.

67. Hinojos CAD, Sharp ZD, Mancini MA: **Molecular dynamics and nuclear receptor function**. *Trends in Endocrinology and Metabolism* 2005, **16**:12-18.

68. Nagy L, Kao HY, Love JD, Li C, Banayo E, Gooch JT, Krishna V, Chatterjee K, Evans RM, Schwabe JW: **Mechanism of corepressor binding and release from nuclear hormone receptors**. *Genes Dev* 1999, **13**:3209-3216.

69. Nagy L, Kao HY, Chakravarti D, Lin RJ, Hassig CA, Ayer DE, Schreiber SL, Evans RM: **Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase**. *Cell* 1997, **89**:373-380.

70. Love JD, Gooch JT, Nagy L, Chatterjee VKK, Schwabe JWR: **Transcriptional repression by nuclear receptors: mechanisms and role in disease**. *Biochemical Society Transactions* 2000, **28**:390-396.

71. Onate SA, Tsai SY, Tsai MJ, Omalley BW: **Sequence and Characterization of a Coactivator for the Steroid-Hormone Receptor Superfamily**. *Science* 1995, **270**:1354-1357.

72. Chen HW, Lin RJ, Schiltz RL, Chakravarti D, Nash A, Nagy L, Privalsky ML, Nakatani Y, Evans RM: **Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300**. *Cell* 1997, **90**:569-580.

73. Stallcup MR, Kim JH, Teyssier C, Lee YH, Ma H, Chen DG: **The roles of protein-protein interactions and protein methylation in transcriptional activation by nuclear receptors and their coactivators**. *Journal of Steroid Biochemistry and Molecular Biology* 2003, **85**:139-145.

74. Leo C, Chen JD: **The SRC family of nuclear receptor coactivators**. *Gene* 2000, **245**:1-11.

75. Belandia B, Parker MG: **Nuclear receptors: A rendezvous for chromatin remodeling factors**. *Cell* 2003, **114**:277-280.

76. Liu Z, Wong JM, Tsai SY, Tsai MJ, O'Malley BW: **Steroid receptor coactivator-1 (SRC-1) enhances ligand-dependent and receptor-dependent cell-free transcription of chromatin**. *Proceedings of the National Academy of Sciences of the United States of America* 1999, **96**:9485-9490.

77. Collingwood TN, Urnov FD, Chatterjee VKK, Wolffe AP: **Chromatin remodeling by the thyroid hormone receptor in regulation of the thyroid-stimulating hormone alpha-subunit promoter**. *Journal of Biological Chemistry* 2001, **276**:34227-34234.

78. Collingwood TN, Urnov FD, Wolffe AP: **Nuclear receptors: coactivators, corepressors and chromatin remodeling in the control of transcription**. *Journal of Molecular Endocrinology* 1999, **23**:255-275.

79. Kim S, Liang J, Barry BA: **Chemical complementation identifies a proton acceptor for redox-active tyrosine D in photosystem II**. *Proc Natl Acad Sci U S A* 1997, **94**:14406-14411.

80. James P, Halladay J, Craig EA: **Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast**. *Genetics* 1996, **144**:1425-1436.

81. Szanto A, Narkar V, Shen Q, Uray IP, Davies PJA, Nagy L: **Retinoid X receptors: X-ploring their (patho)physiological functions**. *Cell Death and Differentiation* 2004, **11**:S126-S143.

82. Hallenbeck PL, Minucci S, Lippoldt P, Phyillaier M, Horn V, Ozato K, Nikodem VM: **Differential 9-cis-retinoic acid-dependent transcriptional activation by murine retinoid X receptor alpha (RXR alpha) and RXR beta - Role of cell type and RXR domains**. *Journal of Biological Chemistry* 1996, **271**:10503-10507.

83. Egea PF, Mitschler A, Moras D: **Molecular recognition of agonist Ligands by RXRs**. *Molecular Endocrinology* 2002, **16**:987-997.

84. Levin AA, Sturzenbecker LJ, Kazmer S, Bosakowski T, Huselton C, Allenby G, Speck J, Kratzeisen C, Rosenberger M, Lovey A, et al.: **9-Cis Retinoic Acid Stereoisomer Binds and Activates the Nuclear Receptor Rxr-Alpha**. *Nature* 1992, **355**:359-361.

85. Allenby G, Bocquel MT, Saunders M, Kazmer S, Speck J, Rosenberger M, Lovey A, Kastner P, Grippo JF, Chambon P, et al.: **Retinoic Acid Receptors and Retinoid X-Receptors - Interactions with Endogenous Retinoic Acids**. *Proceedings of the National Academy of Sciences of the United States of America* 1993, **90**:30-34.

86. Kersten S, Dong D, Lee W, Reczek PR, Noy N: **Auto-silencing by the retinoid X receptor**. *J Mol Biol* 1998, **284**:21-32.

87. Kersten S, Kelleher D, Chambon P, Gronemeyer H, Noy N: **Retinoid X receptor alpha forms tetramers in solution**. *Proc Natl Acad Sci U S A* 1995, **92**:8645-8649.

88. Kersten S, Pan L, Chambon P, Gronemeyer H, Noy N: **Role of ligand in retinoid signaling. 9-cis-retinoic acid modulates the oligomeric state of the retinoid X receptor**. *Biochemistry* 1995, **34**:13717-13721.

89. Kersten S, Reczek PR, Noy N: **The tetramerization region of the retinoid X receptor is important for transcriptional activation by the receptor**. *J Biol Chem* 1997, **272**:29759-29768.

90. de Urquiza AM, Liu SY, Sjoberg M, Zetterstrom RH, Griffiths W, Sjovall J, Perlmann T: **Docosahexaenoic acid, a ligand for the retinoid X receptor in mouse brain**. *Science* 2000, **290**:2140-2144.

91. Egea PF, Mitschler A, Rochel N, Ruff M, Chambon P, Moras D: **Crystal structure of the human RXR alpha ligand-binding domain bound to its natural ligand: 9-cis retinoic acid**. *Embo Journal* 2000, **19**:2592-2601.

92. Lehmann JM, McKee DD, Watson MA, Willson TM, Moore JT, Kliewer SA: **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**:1016-1023.

93. LeCluyse EL: **Pregnane X receptor: molecular basis for species differences in CYP3A induction by xenobiotics**. *Chemico-Biological Interactions* 2001, **134**:283-289.

94. Landes N, Pfluger P, Kluth D, Birringer M, Ruhl R, Bol GF, Glatt H, Brigelius-Flohe R: **Vitamin E activates gene expression via the pregnane X receptor**. *Biochemical Pharmacology* 2003, **65**:269-273.

95. Synold TW, Dussault I, Forman BM: **The orphan nuclear receptor SXR coordinately regulates drug metabolism and efflux**. *Nature Medicine* 2001, **7**:584-590.

96. Chawla A, Repa JJ, Evans RM, Mangelsdorf DJ: **Nuclear receptors and lipid physiology: opening the X-files**. *Science* 2001, **294**:1866-1870.

97. Schultz JR, Tu H, Luk A, Repa JJ, Medina JC, Li LP, Schwendner S, Wang S, Thoolen M, Mangelsdorf DJ, et al.: **Role of LXRs in control of lipogenesis**. *Genes & Development* 2000, **14**:2831-2838.

98. Peet DJ, Turley SD, Ma W, Janowski BA, Lobaccaro JM, Hammer RE, Mangelsdorf DJ: **Cholesterol and bile acid metabolism are impaired in mice lacking the nuclear oxysterol receptor LXR alpha**. *Cell* 1998, **93**:693-704.

99. Janowski BA: **The hypocholesterolemic agent LY295427 up-regulates INSIG-1, identifying the INSIG-1 protein as a mediator of cholesterol homeostasis through SREBP**. *Proc Natl Acad Sci U S A* 2002, **99**:12675-12680.

100. Peet DJ, Janowski BA, Mangelsdorf DJ: **The LXRs: a new class of oxysterol receptors**. *Curr Opin Genet Dev* 1998, **8**:571-575.

101. Janowski BA, Willy PJ, Devi TR, Falck JR, Mangelsdorf DJ: **An oxysterol signalling pathway mediated by the nuclear receptor LXR alpha**. *Nature* 1996, **383**:728-731.

102. Janowski BA, Grogan MJ, Jones SA, Wisely GB, Kliewer SA, Corey EJ, Mangelsdorf DJ: **Structural requirements of ligands for the oxysterol liver X receptors LXRalpha and LXRbeta**. *Proc Natl Acad Sci U S A* 1999, **96**:266-271.

103. Heyman RA, Mangelsdorf DJ, Dyck JA, Stein RB, Eichele G, Evans RM, Thaller C: **9-cis retinoic acid is a high affinity ligand for the retinoid X receptor**. *Cell* 1992, **68**:397-406.

104. Miyata KS, McCaw SE, Marcus SL, Rachubinski RA, Capone JP: **The Peroxisome Proliferator-Activated Receptor Interacts with the Retinoid-X Receptor in-Vivo**. *Gene* 1994, **148**:327-330.

105. Lundblad V: In *Short Protocols in Molecular Biology*. Edited by: Wiley; 1999:13-11 - 13-57. [Ausubel F, Brent R, Kingston R, Moore D, Seidman J, Smith J, Struhl K (Series Editor):

106. Gietz RD, Woods RA: **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*. Edited by; 2002:87-96. Methods in Enzymology, vol 350.]

107. Mangelsdorf DJ, Borgmeyer U, Heyman RA, Zhou JY, Ong ES, Oro AE, Kakizuka A, Evans RM: **Characterization of three RXR genes that mediate the action of 9-cis retinoic acid**. *Genes Dev* 1992, **6**:329-344.

108. Allegretto EA, McClurg MR, Lazarchik SB, Clemm DL, Kerner SA, Elgort MG, Boehm MF, White SK, Pike JW, Heyman RA: **Transactivation Properties of Retinoic Acid and Retinoid-X Receptors in Mammalian-Cells and Yeast - Correlation with Hormone-Binding and Effects of Metabolism**. *Journal of Biological Chemistry* 1993, **268**:26625-26633.

109. Guthrie C, Fink GR: *Guide to yeast genetics and molecular biology*, vol 194. Edited by Abelson JN, Simon MI. San Diego: Academic Press; 1991.

110. Peet DJ, Doyle DF, Corey DR, Mangelsdorf DJ: **Engineering novel specificities for ligand-activated transcription in the nuclear hormone receptor RXR**. *Chemistry & Biology* 1998, **5**:13-21.

111. Toby GG, Golemis EA: **Using the yeast interaction trap and other two-hybrid-based approaches to study protein-protein interactions**. *Methods* 2001, **24**:201-217.

112. Diaz-Camino C, Risseeuw EP, Liu E, Crosby WL: **A high-throughput system for two-hybrid screening based on growth curve analysis in microtiter plates**. *Analytical Biochemistry* 2003, **316**:171-174.

113. Rupp S: **LacZ assays in yeast**. In *Guide to Yeast Genetics and Molecular and Cell Biology, Pt B*. Edited by; 2002:112-131. Methods in Enzymology, vol 350.]

114. Serebriiskii IG, Toby GG, Golemis EA: **Streamlined yeast colorimetric reporter activity assays using scanners and plate readers**. *Biotechniques* 2000, **29**:278-+.

115. Zhang JH, Chung TDY, Oldenburg KR: **A simple statistical parameter for use in evaluation and validation of high throughput screening assays**. *Journal of Biomolecular Screening* 1999, **4**:67-73.

116. Vidal M, Braun P, Chen E, Boeke JD, Harlow E: **Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system**. *Proceedings of the National Academy of Sciences of the United States of America* 1996, **93**:10321-10326.

117. Vidal M, Brachmann RK, Fattaey A, Harlow E, Boeke JD: **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**:10315-10320.

118. Voth WP, Richards JD, Shaw JM, Stillman DJ: **Yeast vectors for integration at the HO locus**. *Nucleic Acids Research* 2001, **29**:e59.

119. Heinlein CA, Chang CS: **Androgen receptor in prostate cancer**. *Endocrine Reviews* 2004, **25**:276-308.

120. Katzenellenbogen BS, Katzenellenbogen JA: **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**:335-344.

121. Rosen ED, Hsu CH, Wang XZ, Sakai S, Freeman MW, Gonzalez FJ, Spiegelman BM: **C/EBP alpha induces adipogenesis through PPAR gamma: a unified pathway**. *Genes & Development* 2002, **16**:22-26.

122. Rosen ED, Sarraf P, Troy AE, Bradwin G, Moore K, Milstone DS, Spiegelman BM, Mortensen RM: **PPAR gamma is required for the differentiation of adipose tissue in vivo and in vitro**. *Molecular Cell* 1999, **4**:611-617.

123. Ge K, Guermah M, Yuan CX, Ito M, Wallberg AE, Spiegelman BM, Roeder RG: **Transcription coactivator TRAP220 is required for PPAR gamma 2-stimulated adipogenesis**. *Nature* 2002, **417**:563-567.

124. Jordan VC: **Is tamoxifen the Rosetta stone for breast cancer?** *J Natl Cancer Inst* 2003, **95**:338-340.

125. Schafer JM, Bentrem DJ, Takei H, Gajdos C, Badve S, Jordan VC: **A mechanism of drug resistance to tamoxifen in breast cancer**. *J Steroid Biochem Mol Biol* 2002, **83**:75-83.

126. Ekena K, Weis KE, Katzenellenbogen JA, Katzenellenbogen BS: **Different residues of the human estrogen receptor are involved in the recognition of structurally diverse estrogens and antiestrogens**. *Journal of Biological Chemistry* 1997, **272**:5069-5075.

127. Tamrazi A, Carlson KE, Katzenellenbogen JA: **Molecular sensors of estrogen receptor conformations and dynamics**. *Molecular Endocrinology* 2003, **17**:2593-2602.

128. Bentrem DJ, Jordan VC: **Antiestrogens: the past and the future**. *Zentralbl Gynakol* 2002, **124**:551-558.

129. Jensen EV, Jordan VC: **The estrogen receptor: a model for molecular medicine**. *Clin Cancer Res* 2003, **9**:1980-1989.

130. Rahman M, Miyamoto H, Chang CS: **Androgen receptor coregulators in prostate cancer: Mechanisms and clinical implications**. *Clinical Cancer Research* 2004, **10**:2208-2219.

131. Lubahn DB, Joseph DR, Sullivan PM, Willard HF, French FS, Wilson EM: **Cloning of human androgen receptor complementary DNA and localization to the X chromosome**. *Science* 1988, **240**:327-330.

132. Frank C, Molnar F, Matilainen M, Lempiainen H, Carlberg C: **Agonist-dependent and agonist-independent transactivations of the human constitutive androstane receptor are modulated by specific amino acid pairs**. *Journal of Biological Chemistry* 2004, **279**:33558-33566.

133. Frank C, Gonzalez MM, Oinonen C, Dunlop TW, Carlberg C: **Characterization of DNA complexes formed by the nuclear receptor constitutive androstane receptor**. *Journal of Biological Chemistry* 2003, **278**:43299-43310.

134. Suino K, Peng L, Reynolds R, Li Y, Cha JY, Repa JJ, Kliewer SA, Xu HE: **The nuclear xenobiotic receptor CAR: Structural determinants of constitutive activation and heterodimerization**. *Molecular Cell* 2004, **16**:893-905.

135. Handschin C, Meyer UA: **Induction of drug metabolism: The role of nuclear receptors**. *Pharmacological Reviews* 2003, **55**:649-673.

136. Handschin C, Meyer UA: **Regulatory network of lipid-sensing nuclear receptors: roles for CAR, PXR, LXR, and FXR**. *Archives of Biochemistry and Biophysics* 2005, **433**:387-396.

137. Huang WD, Zhang J, Wei P, Schrader WT, Moore DD: **Meclizine is an agonist ligand for mouse constitutive androstane receptor (CAR) and an inverse agonist for human CAR**. *Molecular Endocrinology* 2004, **18**:2402-2408.

138. Green S, Walter P, Greene G, Krust A, Goffin C, Jensen E, Scrace G, Waterfield M, Chambon P: **Cloning of the human oestrogen receptor cDNA**. *J Steroid Biochem* 1986, **24**:77-83.

139. Greene GL, Gilna P, Waterfield M, Baker A, Hort Y, Shine J: **Sequence and expression of human estrogen receptor complementary DNA**. *Science* 1986, **231**:1150-1154.

140. Green S, Chambon P: **A superfamily of potentially oncogenic hormone receptors**. *Nature* 1986, **324**:615-617.

141. Kuiper G, Enmark E, PeltoHuikko M, Nilsson S, Gustafsson JA: **Cloning of a novel estrogen receptor expressed in rat prostate and ovary**. *Proceedings of the National Academy of Sciences of the United States of America* 1996, **93**:5925-5930.

142. Bapat AR, Frail DE: **Full-length estrogen receptor alpha and its ligand-binding domain adopt different conformations upon binding ligand**. *Journal of Steroid Biochemistry and Molecular Biology* 2003, **86**:143-149.

143. Pike ACW, Brzozowski AM, Hubbard RE, Bonn T, Thorsell AG, Engstrom O, Ljunggren J, Gustafsson JK, Carlquist M: **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**:4608-4618.

144. Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engstrom O, Ohman L, Greene GL, Gustafsson JA, Carlquist M: **Molecular basis of agonism and antagonism in the oestrogen receptor**. *Nature* 1997, **389**:753-758.

145. Tanenbaum DM, Wang Y, Williams SP, Sigler PB: **Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains**. *Proc Natl Acad Sci U S A* 1998, **95**:5998-6003.

146. Makishima M, Okamoto AY, Repa JJ, Tu H, Learned RM, Luk A, Hull MV, Lustig KD, Mangelsdorf DJ, Shan B: **Identification of a nuclear receptor for bile acids**. *Science* 1999, **284**:1362-1365.

147. Fujino T, Une M, Imanaka T, Inoue K, Nishimaki-Mogami T: **Structure-activity relationship of bile acids and bile acid analogs in regard to FXR activation**. *Journal of Lipid Research* 2004, **45**:132-138.

148. Weatherman RV, Fletterick RJ, Scanlan TS: **Nuclear-receptor ligands and ligand-binding domains**. *Annu Rev Biochem* 1999, **68**:559-581.

149. Kliewer SA, Goodwin B, Willson TM: **The nuclear pregnane X receptor: A key regulator of xenobiotic metabolism**. *Endocrine Reviews* 2002, **23**:687-702.

150. Takeshita A, Taguchi M, Koibuchi N, Ozawa Y: **Putative role of the orphan nuclear receptor SXR (steroid and xenobiotic receptor) in the mechanism of CYP3A4 inhibition by xenobiotics**. *Journal of Biological Chemistry* 2002, **277**:32453-32458.

151. Watkins RE, Wisely GB, Moore LB, Collins JL, Lambert MH, Williams SP, Willson TM, Kliewer SA, Redinbo MR: **The human nuclear xenobiotic**

receptor PXR: Structural determinants of directed promiscuity. *Science* 2001, **292**:2329-2333.

152. Hammer GD, Ingraham HA: **Steroidogenic factor-1: its role in endocrine organ development and differentiation**. *Front Neuroendocrinol* 1999, **20**:199-223.

153. Sablin EP, Krylova IN, Fletterick RJ, Ingraham HA: **Structural basis for ligand-independent activation of the orphan nuclear receptor LRH-1**. *Mol Cell* 2003, **11**:1575-1585.

154. Ingraham HA, Lala DS, Ikeda Y, Luo X, Shen WH, Nachtigal MW, Abbud R, Nilson JH, Parker KL: **The nuclear receptor steroidogenic factor 1 acts at multiple levels of the reproductive axis**. *Genes Dev* 1994, **8**:2302-2312.

155. Desclozeaux M, Krylova IN, Horn F, Fletterick RJ, Ingraham HA: **Phosphorylation and intramolecular stabilization of the ligand binding domain in the nuclear receptor steroidogenic factor 1**. *Mol Cell Biol* 2002, **22**:7193-7203.

156. Lala DS, Syka PM, Lazarchik SB, Mangelsdorf DJ, Parker KL, Heyman RA: **Activation of the orphan nuclear receptor steroidogenic factor 1 by oxysterols**. *Proc Natl Acad Sci U S A* 1997, **94**:4895-4900.

157. Mellon SH, Bair SR: **25-Hydroxycholesterol is not a ligand for the orphan nuclear receptor steroidogenic factor-1 (SF-1)**. *Endocrinology* 1998, **139**:3026-3029.

158. Krylova IN, Sablin EP, Moore J, Xu RX, Waitt GM, MacKay JA, Juzumiene D, Bynum JM, Madauss K, Montana V, et al.: **Structural analyses reveal phosphatidyl inositols as ligands for the NR5 orphan receptors SF-1 and LRH-1**. *Cell* 2005, **120**:343-355.

159. Desai SK, Gallivan JP: **Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation**. *Journal of the American Chemical Society* 2004, **126**:13247-13254.

160. Walker KD, Klettke K, Akiyama T, Croteau R: **Cloning, heterologous expression, and characterization of a phenylalanine aminomutase involved in Taxol biosynthesis**. *J Biol Chem* 2004, **279**:53947-53954.

161. Jennewein S, Park H, DeJong JM, Long RM, Bollon AP, Croteau RB: **Coexpression in yeast of Taxus cytochrome P450 reductase with cytochrome P450 oxygenases involved in Taxol biosynthesis**. *Biotechnol Bioeng* 2005, **89**:588-598.

162. Hezari M, Croteau R: **Taxol biosynthesis: an update**. *Planta Med* 1997, **63**:291-295.

163. Walker K, Croteau R: **Taxol biosynthetic genes**. *Phytochemistry* 2001, **58**:1-7.

164. Watkins RE, Maglich JM, Moore LB, Wisely GB, Noble SM, Davis-Searles PR, Lambert MH, Kliewer SA, Redinbo MR: **2.1 angstrom crystal structure of human PXR in complex with the St. John's wort compound hyperforin**. *Biochemistry* 2003, **42**:1430-1438.

165. Horton RM, Ho SN, Pullen JK, Hunt HD, Cai ZL, Pease LR: **Gene-Splicing by Overlap Extension**. *Methods in Enzymology* 1993, **217**:270-279.

166. Horton RM, Cai ZL, Ho SN, Pease LR: **Gene-Splicing by Overlap Extension - Tailor-Made Genes Using the Polymerase Chain-Reaction**. *Biotechniques* 1990, **8**:528-&.

167. Horton RM: **In vitro recombination and mutagenesis of DNA. SOEing together tailor-made genes**. *Methods Mol Biol* 1997, **67**:141-149.

168. Pervaiz S: **Resveratrol: from grapevines to mammalian biology**. *Faseb Journal* 2003, **17**:1975-1985.

169. Yamada K, Wang JC, Osawa H, Scott DK, Granner DK: **Efficient large-scale transformation of yeast**. *Biotechniques* 1998, **24**:596-598, 600.

170. Butler MS: **Natural products to drugs: natural product derived compounds in clinical trials**. *Natural Product Reports* 2005, **22**:162-195.

171. Butler MS: **The role of natural product chemistry in drug discovery**. *Journal of Natural Products* 2004, **67**:2141-2153.

172. Clardy J, Walsh C: **Lessons from natural molecules**. *Nature* 2004, **432**:829-837.

173. Pelzer S, Vente A, Bechthold A: **Novel natural compounds obtained by genome-based screening and genetic engineering**. *Current Opinion in Drug Discovery & Development* 2005, **8**:228-238.

174. Kohler DR, Goldspiel BR: **Paclitaxel (Taxol)**. *Pharmacotherapy* 1994, **14**:3-34.

175. Cragg GM: **Paclitaxel (Taxol): a success story with valuable lessons for natural product drug discovery and development**. *Med Res Rev* 1998, **18**:315-331.

176. Rowinsky EK, Donehower RC: **The clinical pharmacology of paclitaxel (Taxol)**. *Semin Oncol* 1993, **20**:16-25.

177. Martin V: **Overview of paclitaxel (TAXOL)**. *Semin Oncol Nurs* 1993, **9**:2-5.

178. Paclitaxel statistics on World Wide Web URL: www.pfc.forestry.ca/ecology/yew/qanda_e.html

179. Tang L, Shah S, Chung L, Carney J, Katz L, Khosla C, Julien B: **Cloning and heterologous expression of the epothilone gene cluster**. *Science* 2000, **287**:640-642.

180. Julien B, Shah S: **Heterologous expression of epothilone biosynthetic genes in Myxococcus xanthus**. *Antimicrob Agents Chemother* 2002, **46**:2772-2778.

181. Julien B, Shah S, Ziermann R, Goldman R, Katz L, Khosla C: **Isolation and characterization of the epothilone biosynthetic gene cluster from Sorangium cellulosum**. *Gene* 2000, **249**:153-160.

182. Becker JVW, Armstrong GO, Van der Merwe MJ, Lambrechts MG, Vivier MA, Pretorius IS: **Metabolic engineering of Saccharomyces cerevisiae for the synthesis of the wine-related antioxidant resveratrol**. *Fems Yeast Research* 2003, **4**:79-85.

183. Jez JM, Ferrer JL, Bowman ME, Austin MB, Schroder J, Dixon RA, Noel JP: **Structure and mechanism of chalcone synthase-like polyketide synthases**. *Journal of Industrial Microbiology & Biotechnology* 2001, **27**:393-398.

184. Jez JM, Bowman ME, Dixon RA, Noel JP: **Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase**. *Nature Structural Biology* 2000, **7**:786-791.

185. Ferrer JL, Jez JM, Bowman ME, Dixon RA, Noel JP: **Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis**. *Nat Struct Biol* 1999, **6**:775-784.