TRANSCRIPTION FACTOR ENGINEERING: TOOLS AND APPLICATIONS

BY

MICHAEL JAMES MCLACHLAN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biophysics & Computational Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Professor Huimin Zhao, Chair
Professor John A. Katzenellenbogen
Assistant Professor Sheng Zhong
Professor John A. Gerlt

Abstract

Transcription factors play a vital role in the biology of every organism. By controlling gene expression they regulate growth, development, metabolism, reproduction, signaling, and response to the environment. They have also provided the basis for many useful tools in molecular biology. The estrogen receptor alpha is one of the most studied human transcription factors and acts as a ligand controlled regulator of transcription. The modular design of this and other transcription factors allows for the rational design of artificial gene switches to control expression of desired genes.

In this thesis I explore some tools for, and applications of, engineering the estrogen receptor. Beginning with two ligand binding domain mutants previously engineered to recognize the small molecules 4,4'-dihydroxybenzil (DHB) or 2,4-di(4-hydroxyphenyl)-5-ethylthiazole (L9), I showed that they could be used as a gene switch to independently control reporter genes in yeast and mammalian cells. By using different DNA binding domains, activation and repression domains, promoter elements, and a luciferase reporter I implemented the logic functions AND, OR, NAND, and NOR in HeLa cell culture. My research revealed some of the limitations of both the modular engineering approach, and the yeast two-hybrid screening assay used to engineer the ligand binding domain. I explored the feasibility of performing directed evolution of gene switches in mammalian cells through a protoplast fusion method, which combines the benefits of simple library creation with screening in a functionally relevant system. Although individual steps of the process were successful, the method proved unsuitable for large scale screening of libraries. The endogenous gene vascular endothelial growth factor-A (*VEGF-A*) was targeted for control by a gene switch. The effect of construct design was evaluated using a *VEGF-A* promoter controlled luciferase gene and performance was impacted by the choice of DNA binding domain, activation domain, and the order of domain use. Endogenous VEGF-A protein secretion in HeLa cells was successfully upregulated twofold by a DHB ligand controlled gene switch. Finally, I developed a useful biosensor for estrogenic compound detection that has the advantage of requiring no added substrates for signal generation. Through fusing the N and C terminal halves of the fluorescent protein Venus to the receptor ligand binding domain, fluorescence complementation generated a robust signal upon addition of an estrogenic ligand.

The biosensor was capable of responding to a range of endogenous, pharmaceutical, environmental, and industrial compounds with sensitivities that correlated with their relative binding affinity. The signal characteristics were seen to depend on the length of the LBD region used, with some constructs distinguishing between agonists and antagonistic ligands.

Acknowledgements

Foremost, I would like to thank my wife Traci for her support and encouragement. There have been many ups and downs during my graduate studies and you have helped me through them all with your love, friendship, and advice. You have kept me on track when things were hard and I will be forever grateful. Thank you to my parents Bruce and Colleen, who raised me with a love for learning and helped me reach my goals, and my brother Stuart and sister Helen for all your support. To my old advisor and friend Helen Davey, you always wanted the best for me and I hope this would make you proud.

To my advisor, Professor Huimin Zhao, thank you for the opportunity to work in your laboratory and be a part of so many interesting projects, I have learned a lot from my time here. Thank you to my committee members Professor John Katzenellenbogen, Professor Sheng Zhong, and Professor John Gerlt for your comments and guidance. I would also like to thank Professor Ido Golding for serving on my Prelim committee.

I am grateful to all the past and present members of the Zhao laboratory for their years of support, suggestions, and companionship. In particular Professor Tyler Johannes who I worked with on the phosphite dehydrogenase engineering project, Dr. Karuppiah Chockalingam for getting me started on the gene switch project, and Jing Liang for his help with the protoplast fusion experiments. I am grateful to Victor Gonzalez for his work on the gene switch project and Dr. Norio Yasui, John Comninos, and Kathy Carlson for help with ligands. Xiaoyi Cao and Zhanar Abil assisted me with cloning some of the constructs. Dr. Barbara Pilas and Dr. Bernard Montez were of great assistance for the flow cytometry experiments. I appreciate the gift of plasmids from Professor Wolfgang Hillen (p65), Professor Carlos Barbas (KRAB and VP64), Professor Lee Ellis (VEGF-A reporter for comparison), Professor Yingxiao Wang (fluorescent proteins), and Dr. Changkyu Oh (F435p). Thanks to Sandy McMasters for all the media.

Table of Contents

Chapter 1: Introduction

1.1 Transcription Factors

Transcription factors are proteins that control the timing and abundance of gene expression and thus lie at the heart of metabolism, development, and differentiation (1, 2). Their mutation or misregulation can lead to abnormal states such as cancer, inflammatory diseases, metabolic disorders, and congenital birth defects (3, 4, 5, 6). They are typically characterized by the presence of a DNA binding domain (DBD), activation or repression domains capable of influencing transcription, and protein interaction or regulatory domains. There are estimated to be around 1900 transcription factors in humans, most of which remain poorly characterized, that act in conjunction with the general transcriptional machinery to determine which genes are expressed at any given time (7). Their regulation and differential expression allow cells to respond to their environment and generate multiple cell phenotypes from a single genome. For these reasons, studies of transcription factors, and the gene regulatory networks they form, are of great interest.

Research on transcription factors has resulted in a number of useful biological tools. Model transcription factors and promoter elements are widely used as inducible gene expression systems, for example protein production in *E. coli* controlled by LacI (8) or AraC (9), or gene expression in mammalian cells controlled via TetR (10, 11). The yeast two hybrid system was developed with the yeast GAL4 protein by recognizing that transcription factors consist of separable domains controlling DNA binding and transcriptional activation, and provides a useful assay to detect protein-protein interaction. The fusion of the GAL4 DBD to one protein and the GAL4 activation domain (AD) to another leads to induction of a reporter gene if the two proteins interact (12, 13). Many transcription factors are regulated by the binding of a small molecule, such as a metabolite or signaling compound, and thus can be converted into biosensors for compound detection or library screening purposes (14, 15). Site-specific genome modifications are challenging to perform in higher eukaryotes, but combining a DNA binding domain with the FokI nuclease domain has provided a solution. By cleaving the chromosome at a desired location using a zinc finger nuclease, homologous recombination with a DNA construct containing the desired alteration is greatly stimulated (16, 17, 18, 19, 20, 21).

Controlling gene expression through the use of transcription factors has many applications. Metabolic engineering involves the alteration of metabolism through the introduction of novel enzymes/pathways or the alteration of expression of existing genes within an organism, for the purpose of producing commercial or nutritional compounds of interest. Depending on the approach taken, transcription factors can be used to modulate expression of either individual genes or whole pathways in plants (22, 23, 24), *E. coli* (25, 26), or yeast (27, 28). The recent and rapidly growing field of induced pluripotent stem cells (iPSC) was spurred by the discovery that the introduction of only four transcription factors (Oct3/4, Sox2, c-Myc, and Klf4) has the ability to reset the gene regulatory network and convert fibroblast cells into stem cells (29), and expands on observations that cellular identity can be modified by introducing alternate regulatory proteins (30). The iPSC field holds much promise for basic science, drug discovery, and therapeutic applications. Gene therapy, while slow to produce many clinical treatments, is being actively pursued and often involves the introduction of transcription factors. The first approved gene therapy treatment was in China and uses the p53 tumor suppressor transcription factor in the treatment of head and neck squamous cell cancer (31).

1.2 Estrogen Receptor Alpha

The estrogen receptor alpha (ERα) is a member of the steroid binding nuclear receptor family of transcriptional regulators and is one of the most highly studied transcription factors (7, 32, 33, 34). It plays an important role in development and reproduction, and is implicated in a number of cancers. ERα was first cloned in 1985 (35, 36) and was subsequently discovered to act in conjunction with a related gene that was named estrogen receptor β to mediate estrogen signaling (37). The receptors are expressed in most tissues of the body, with ERα expressed mainly in reproductive tissues (uterus, ovary, testes), bone, breast, liver, kidney, fat, and the brain.(38). While gene deletion of the estrogen receptor in mice is not lethal, females were infertile and males had decreased fertility (39).

Nuclear receptors have a common structure consisting of an N-terminal region, a conserved Cys_4 zinc finger DNA binding domain, a hinge region, a ligand-binding domain (LBD), and a C terminal region (Figure 1.1). Due to the highly dynamic nature of the receptors, the only full length structure is for the PPARγ-RXRα heterodimer (40). However, structures for various

domains have been published from a number of nuclear receptors (33). The N terminal A/B region contains a ligand independent activation function and regions involved in domain and coactivator interactions. The DNA binding C domain is highly conserved across all nuclear receptors. Crystal structures determined of the binding region complexed with DNA response elements showed that the structure was a distinctive two zinc finger domain, and revealed that the ER DBD made contact with four basepairs and organized a network of water molecules to stabilize the interaction with the estrogen response element (ERE) (41). The crystallized DBDs bound the DNA as a dimer, with direct contacts at the dimer interface. The symmetry that was seen explains the palindromic nature of the ERE half sites: AGGTCA NNN TGACCT. The D domain acts as a hinge region between the DNA binding domain and the ligand binding domain (E). The LBD is composed of 12 alpha helices that surround a central pocket for a hydrophobic ligand (42). The estrogen receptor is interesting in that it is capable of binding many different ligands that are produced endogenously, encountered in the environment, or produced as pharmaceuticals (43, 44). Upon binding of an agonist, helix 12 undergoes a conformational change that reveals a cleft in the protein capable of binding an LXXLL motif in coactivators (45), whereas antagonists act to position helix 12 such that coactivators can not bind (46). The LBD is also required for nuclear localization and dimerization, with the binding of ligand stabilizing dimers (47). The F domain has a modulatory function on dimerization, agonists, and antagonists (48).

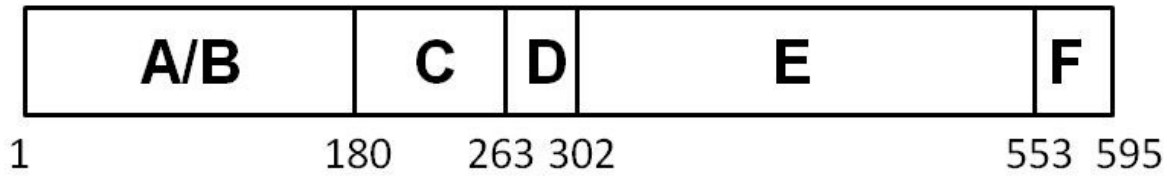


Figure 1.1. Estrogen receptor α domains. The amino acid positions and functional domains are indicated.

ERα regulates the expression of genes in response to estrogenic compounds, typified by 17β-estradiol, which bind to the ligand binding domain and cause a conformational change that recruits coactivators or corepressors. It can exert its effects in a genomic manner, by binding as a dimer to estrogen response elements or interacting with other transcription factors like AP-1, or in a non-genomic manner by affecting for example MAPK signaling pathways (49). Genome-wide studies have examined estrogen receptor binding sites and it was found they are usually

more than 1 kb away from a gene's transcriptional start site, and bind ERα to form long range loops of interacting chromatin regions coordinating gene expression (50, 51).

1.3 Engineering Transcription Factors

Much attention has been focused on protein engineering of transcription factors. One area has been the DNA binding domain since altering this region allows for control over the genes regulated. Evolution has produced multiple ways of binding DNA including the structure groups helix-turn-helix (eg LacI, homeodomain proteins), zinc coordinating (Cys_2His_2 zinc finger, nuclear hormone Cys_4 zinc finger, p53 loop-sheet-helix, GAL4), zippers (leucine zipper, helix-loop-helix), and others (52). Most classes of DBDs do not show an easily decipherable code between the amino acids in the protein and the sequence of DNA bound, the main exception being the Cys_2His_2 zinc finger (53, 54). This structure consists of a 30 amino acid module that coordinates a zinc ion between two cysteines and two histidines, with specific residues in a finger associating with a specific base, and multiple fingers arranged in a linear manner. The result is that one finger binds a triplet of DNA (with a fourth base overlapping the next finger's triplet), with multiple fingers binding longer linear sequences. This modular arrangement was seized upon by researchers to engineer sequence specific DNA binding proteins (55, 56, 57), just as evolution had seized upon it and turned it into the most common DBD in humans with over 700 examples (7). A recent DBD that also shows potential for engineering due to its simple recognition code is the TAL effector domain (58, 59).

Regulation of transcription factors has also been a target of engineering. Nuclear hormone ligand binding domains have been engineered to respond to new ligands to control gene expression in an orthogonal manner (60, 61, 62, 63), and have also been used as ligand activated switches to control enzymes (64). Temperature controlled self splicing inteins have been developed to control transcription factor function (65). The use of different natural or engineered activation domains has been investigated to increase the activity of a transcription factor (66, 67). Likewise, repressors have been investigated and developed into useful systems (68).

The approaches taken to engineer transcription factors have varied. Much research is based on rational design and the modularity of protein domains (69). An artificial transcription factor can

be constructed by fusing a DNA binding domain (either natural or engineered) with a known activation or repression domain, and optional regulatory domains. However there are often unexpected results from putting together separate components. Directed evolution can be a very useful tool for engineering both components and whole transcription factors, and can work in the absence of structural information (70).

1.4 Gene Switches and Endogenous Gene Regulation

The ability to regulate endogenous genes has applications in basic research as well as therapeutic relevance. A fundamental method of investigating gene function is to either knockout a gene, which involves a lot of time and expense, or to over-express an introduced copy which may not reproduce the natural splicing and regulation of the endogenous locus. By introducing factors that affect the natural locus a more flexible level of control can often be achieved. From a therapeutic perspective, introducing a transcription factor that affects the endogenous gene can produce a more natural mix of target gene splice variants (71).

Many endogenous loci have thus far been targeted for regulation (72). The first human loci successfully targeted by an artificial transcription factor was the *Erb-2* gene, commonly upregulated in cancer, which Barbas's group used a six-finger engineered DBD to both up- and down-regulate (73). Another cancer related gene, *MDR1* involved in multidrug resistance, was repressed using a five-finger DBD with the KRAB repressor (74). By targeting the erythropoietin gene, a hormone that controls red blood cell production and is therefore a therapeutic target for anemia, the importance of the chromatin structure was established with linker regions between nucleosomes allowing binding of the zinc finger (75). The vascular endothelial growth factor A, an essential regulator of blood vessel development and implicated in tumor growth, has been targeted by multiple groups by constitutively acting constructs (76, 77, 78, 79, 80) or by ligand activated constructs (81, 82). Other cancer related genes targeted include *IGF2* and *H19*, *bax*, checkpoint kinase 2, and mammary serine protease inhibitor (83, 84, 85, 86). The cellular differentiation genes *PPARγ*and *Oct4* have been successfully targeted (87, 88).

1.5 Estrogenic Compound Biosensors

Biosensors are detection systems composed of biological material that can be used to monitor for particular ligands or metabolites (89). They can act in a variety of ways such as inducing a reporter gene with a convenient assay, by allowing an increase in cell growth, or they may be a protein that generates its own signal upon ligand interaction. Ligand controlled transcription factors, like ERα, can be converted into biosensors for the activating ligand. This can be beneficial for investigating which ligands will activate a response, for engineering specificity to a new ligand, or for metabolic engineering. The ligand binding domain is the key to this activity, and it can be removed from its native context to engineer systems that function in different ways.

Multiple biosensors for nuclear hormone-like compounds have been developed that differ in the reporter used, and also in the host species (90). Assays have been developed for estradiol activity that are close to its natural activity, such as an increase in prepubertal mouse uterine weight or the growth of the estradiol responsive MCF-7 tumor cell line (91). While these have the benefit of being highly relevant physiologically, they take longer than other assays and are not useful from an engineering perspective. As ERα is a transcription factor, biosensors have been constructed simply by cloning a reporter gene such as β-galactosidase, luciferase, or EGFP under the control of estrogen receptor response elements and expressing these in yeast (92, 93, 94). Other studies have replaced the natural DBD with that of the yeast GAL4 transcription factor (95). Yeast one- or two-hybrid assays have been developed that rely on either a β-galactosidase reporter or a growth-based assay (96, 97, 98). Fluorescently tagged estrogen receptors or LBDs have been used to monitor ligand binding based on construct stabilization or Forster resonance energy transfer (99, 100, 101). Other systems have been developed based on complementation of split luciferase fragments expressed in mammalian cells (102), or ligand stimulated splicing out of inteins from a thymidylate synthase gene in bacteria (103).

1.6 Project Overview

The theme of this thesis was an investigation of tools and applications of transcription factor engineering, using the human estrogen receptor α as its basis. The work highlights the utility of the ERα structure from an engineering perspective and also reveals important limitations both of

the modular protein design approach, and of engineering gene switches using a system different from the intended application (yeast vs mammalian cells).

Chapter 2 describes the use of previously engineered ERα ligand binding domains to control gene expression in yeast and mammalian cells. The GAL4 DNA binding domain was used in conjunction with LBDs engineered to respond to the small molecules DHB and L9 in order to demonstrate ligand dependent transcriptional activation of fluorescent protein reporter genes in yeast. These gene switch constructs were also functional in mammalian cells but the addition of activation or repression domains led to ligand independent activity, highlighting the unpredictability of modular engineering. A less sensitive DHB responsive LBD mutant, 4S, was seen to perform well with the addition of activation domains. Constructs were designed to create logic gates within mammalian cells. These results showed that ligand binding domains engineered from a single scaffold could be used to independently control genes, and that conceivably any number of orthogonal gene switches could be created using one engineering approach.

Chapter 3 describes attempts to create a mammalian screening system for engineering gene switches. A protoplast fusion approach was taken which allows for library creation within *E. coli* and subsequent clonal transfer into mammalian cells in culture for screening. Despite multiple construct designs, a *VEGF-A* promoter induced green fluorescent protein reporter did not show significant induction by a gene switch. However, a reporter constructed under the control of a GAL4 responsive promoter was shown to be induced by constitutive or ligand dependent GAL4 DBD containing gene switches. After integrating this reporter into the genome of HeLa cells, a library of ligand binding domain mutants was screened by protoplast fusion, but known functional mutants were not recovered, indicating that the approach was not suitable for gene switch engineering. A fluorescence complementation ligand sensor developed in Chapter 5 was used to investigate protoplast fusion as a means of engineering solely the ligand binding domain, but this also failed to isolate known functional mutants. The previously established yeast two-hybrid screening system was used to isolate mutants that showed increased sensitivity to hexylresorcinol, but when these mutants were transferred into HeLa cells they exhibited a loss of sensitivity, again confirming the need for a well performing mammalian screening system.

Chapter 4 describes the creation of ligand controlled gene switches that activate the endogenous gene vascular endothelial growth factor A. The gene switch performance was seen to depend on its design, with the particular DNA binding domain, activation domain, and order of domains having an effect on both basal and induced expression levels, thus showing that even a rationally designed gene switch construct requires optimization that is not easily predictable. Using a *VEGF-A* promoter controlled luciferase reporter, up to 17 fold ligand-dependent induction was observed. The gene switch successfully stimulated a twofold induction of *VEGF-A* mRNA and secreted protein when introduced into HeLa cells in culture.

Chapter 5 describes the creation of an estrogenic compound biosensor. A series of constructs were made consisting of fragments of a split mVenus fluorescent protein fused at several different N terminal and C terminal positions flanking the ligand binding domain of the estrogen receptor. When expressed in HeLa cells, construct 6 (ERα 312-595) showed a nine fold increase in fluorescence in the presence of estrogen receptor agonists or antagonists. Construct 2 (ERα 281-549) discriminated between agonists and antagonists by showing a slight decrease in fluorescence in the presence of agonists while being induced by antagonists. The fluorescent signal increased over a period of 24 hours. Ligand titration curves performed with construct 6 showed a good correlation with the known relative binding affinities of the compound. The sensor could detect a number of compounds of interest due to their potential as environmental endocrine disruptors. The lack of a substrate requirement, the speed of signal development, the potential for high throughput assays, and the ability to distinguish agonists from antagonists make this an attractive sensor.

1.7 References

1. Brivanlou,A.H. and Darnell,J.E.,Jr. (2002) Signal transduction and the control of gene expression. *Science,* **295,** 813-818.

2. Lemon,B. and Tjian,R. (2000) Orchestrated response: A symphony of transcription factors for gene control. *Genes Dev.,* **14,** 2551-2569.

3. Furney,S.J., Higgins,D.G., Ouzounis,C.A. and Lopez-Bigas,N. (2006) Structural and functional properties of genes involved in human cancer. *BMC Genomics,* **7,** 3.

4. Egwuagu,C.E. (2009) STAT3 in CD4+ T helper cell differentiation and inflammatory diseases. *Cytokine,* **47,** 149-156.

5. Desvergne,B., Michalik,L. and Wahli,W. (2006) Transcriptional regulation of metabolism. *Physiol. Rev.,* **86,** 465-514.

6. King,M., Arnold,J.S., Shanske,A. and Morrow,B.E. (2006) T-genes and limb bud development. *Am. J. Med. Genet. A.,* **140,** 1407-1413.

7. Vaquerizas,J.M., Kummerfeld,S.K., Teichmann,S.A. and Luscombe,N.M. (2009) A census of human transcription factors: Function, expression and evolution. *Nat. Rev. Genet.,* **10,** 252-263.

8. Dubendorff,J.W. and Studier,F.W. (1991) Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *J. Mol. Biol.,* **219,** 45-59.

9. Guzman,L.M., Belin,D., Carson,M.J. and Beckwith,J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.,* **177,** 4121-4130.

10. Gossen,M. and Bujard,H. (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U. S. A.,* **89,** 5547-51.

11. Stieger,K., Belbellaa,B., Le Guiner,C., Moullier,P. and Rolling,F. (2009) In vivo gene regulation using tetracycline-regulatable systems. *Adv. Drug Deliv. Rev.,* **61,** 527-541.

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

13. Chien,C.T., Bartel,P.L., Sternglanz,R. and Fields,S. (1991) The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci. U. S. A.,* **88,** 9578-9582.

14. Galvao,T.C. and de Lorenzo,V. (2006) Transcriptional regulators a la carte: Engineering new effector specificities in bacterial regulatory proteins. *Curr. Opin. Biotechnol.,* **17,** 34-42.

15. Dietrich,J.A., McKee,A.E. and Keasling,J.D. (2010) High-throughput metabolic engineering: Advances in small-molecule screening and selection. *Annu. Rev. Biochem.,* **79,** 563-590.

16. Bibikova,M., Beumer,K., Trautman,J.K. and Carroll,D. (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science,* **300,** 764.

17. Shukla,V.K., Doyon,Y., Miller,J.C., DeKelver,R.C., Moehle,E.A., Worden,S.E., Mitchell,J.C., Arnold,N.L., Gopalan,S., Meng,X., et al. (2009) Precise genome modification in the crop species zea mays using zinc-finger nucleases. *Nature,* **459,** 437-441.

18. Townsend,J.A., Wright,D.A., Winfrey,R.J., Fu,F., Maeder,M.L., Joung,J.K. and Voytas,D.F. (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature,* **459,** 442-445.

19. Geurts,A.M., Cost,G.J., Freyvert,Y., Zeitler,B., Miller,J.C., Choi,V.M., Jenkins,S.S., Wood,A., Cui,X., Meng,X., et al. (2009) Knockout rats via embryo microinjection of zinc-finger nucleases. *Science,* **325,** 433.

20. Carbery,I.D., Ji,D., Harrington,A., Brown,V., Weinstein,E.J., Liaw,L. and Cui,X. (2010) Targeted genome modification in mice using zinc finger nucleases. *Genetics,* doi: 10.1534/genetics.110.117002 .

21. Ochiai,H., Fujita,K., Suzuki,K., Nishikawa,M., Shibata,T., Sakamoto,N. and Yamamoto,T. (2010) Targeted mutagenesis in the sea urchin embryo using zinc-finger nucleases. *Genes Cells,* **15,** 875-885.

22. Mehrtens,F., Kranz,H., Bednarek,P. and Weisshaar,B. (2005) The arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiol.,* **138,** 1083-1096.

23. Van Eenennaam,A.L., Li,G., Venkatramesh,M., Levering,C., Gong,X., Jamieson,A.C., Rebar,E.J., Shewmaker,C.K. and Case,C.C. (2004) Elevation of seed alpha-tocopherol levels using plant-based transcription factors targeted to an endogenous locus. *Metab. Eng.,* **6,** 101-108.

24. Holmes-Davis,R., Li,G., Jamieson,A.C., Rebar,E.J., Liu,Q., Kong,Y., Case,C.C. and Gregory,P.D. (2005) Gene regulation in planta by plant-derived engineered zinc finger protein transcription factors. *Plant Mol. Biol.,* **57,** 411-423.

25. Alper,H. and Stephanopoulos,G. (2007) Global transcription machinery engineering: A new approach for improving cellular phenotype. *Metab. Eng.,* **9,** 258-267.

26. Tang,S.Y., Fazelinia,H. and Cirino,P.C. (2008) AraC regulatory protein mutants with altered effector specificity. *J. Am. Chem. Soc.,* **130,** 5267-5271.

27. Alper,H., Moxley,J., Nevoigt,E., Fink,G.R. and Stephanopoulos,G. (2006) Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science,* **314,** 1565-1568.

28. Watanabe,M., Watanabe,D., Akao,T. and Shimoi,H. (2009) Overexpression of MSN2 in a sake yeast strain promotes ethanol tolerance and increases ethanol production in sake brewing. *J. Biosci. Bioeng.,* **107,** 516-518.

29. Takahashi,K. and Yamanaka,S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* **126,** 663-676.

30. Hakelien,A.M., Landsverk,H.B., Robl,J.M., Skalhegg,B.S. and Collas,P. (2002) Reprogramming fibroblasts to express T-cell functions using cell extracts. *Nat. Biotechnol.,* **20,** 460-466.

31. Raty,J.K., Pikkarainen,J.T., Wirth,T. and Yla-Herttuala,S. (2008) Gene therapy: The first approved gene-based medicines, molecular mechanisms and clinical indications. *Curr. Mol. Pharmacol.,* **1,** 13-23.

32. Whitfield,G.K., Jurutka,P.W., Haussler,C.A. and Haussler,M.R. (1999) Steroid hormone receptors: Evolution, ligands, and molecular basis of biologic function. *J. Cell. Biochem.,* **Suppl 32-33,** 110-122.

33. Huang,P., Chandra,V. and Rastinejad,F. (2010) Structural overview of the nuclear receptor superfamily: Insights into physiology and therapeutics. *Annu. Rev. Physiol.,* **72,** 247-272.

34. Heldring,N., Pike,A., Andersson,S., Matthews,J., Cheng,G., Hartman,J., Tujague,M., Strom,A., Treuter,E., Warner,M., et al. (2007) Estrogen receptors: How do they signal and what are their targets. *Physiol. Rev.,* **87,** 905-931.

35. Walter,P., Green,S., Greene,G., Krust,A., Bornert,J.M., Jeltsch,J.M., Staub,A., Jensen,E., Scrace,G. and Waterfield,M. (1985) Cloning of the human estrogen receptor cDNA. *Proc. Natl. Acad. Sci. U. S. A.,* **82,** 7889-7893.

36. Greene,G.L., Gilna,P., Waterfield,M., Baker,A., Hort,Y. and Shine,J. (1986) Sequence and expression of human estrogen receptor complementary DNA. *Science,* **231,** 1150-1154.

37. Kuiper,G.G., Enmark,E., Pelto-Huikko,M., Nilsson,S. and Gustafsson,J.A. (1996) Cloning of a novel receptor expressed in rat prostate and ovary. *Proc. Natl. Acad. Sci. U. S. A.,* **93,** 5925-5930.

38. Nilsson,S. and Gustafsson,J.A. (2010) Estrogen receptors: Their actions and functional roles in health and disease. In Bunce,C.M. and Campbell,M.J. (eds.), *Nuclear Receptors.* Springer, Netherlands, Vol. 8, pp. 91-91-141.

39. Lubahn,D.B., Moyer,J.S., Golding,T.S., Couse,J.F., Korach,K.S. and Smithies,O. (1993) Alteration of reproductive function but not prenatal sexual development after insertional disruption of the mouse estrogen receptor gene. *Proc. Natl. Acad. Sci. U. S. A.,* **90,** 11162-11166.

40. Chandra,V., Huang,P., Hamuro,Y., Raghuram,S., Wang,Y., Burris,T.P. and Rastinejad,F. (2008) Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. *Nature,* **456,** 350-356.

41. Schwabe,J.W., Chapman,L., Finch,J.T. and Rhodes,D. (1993) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: How receptors discriminate between their response elements. *Cell,* **75,** 567-578.

42. Wurtz,J.M., Bourguet,W., Renaud,J.P., Vivat,V., Chambon,P., Moras,D. and Gronemeyer,H. (1996) A canonical structure for the ligand-binding domain of nuclear receptors. *Nat. Struct. Biol.,* **3,** 87-94.

43. Kuiper,G.G., Carlsson,B., Grandien,K., Enmark,E., Haggblad,J., Nilsson,S. and Gustafsson,J.A. (1997) Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta. *Endocrinology,* **138,** 863-870.

44. Katzenellenbogen,J.A. and Muthyala,R. (2003) Interactions of exogenous endocrine active substances with nuclear receptors. *Pure Appl. Chem.,* **75,** 1797-1817.

45. Heery,D.M., Kalkhoven,E., Hoare,S. and Parker,M.G. (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature,* **387,** 733-736.

46. Brzozowski,A.M., Pike,A.C., Dauter,Z., Hubbard,R.E., Bonn,T., Engstrom,O., Ohman,L., Greene,G.L., Gustafsson,J.A. and Carlquist,M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature,* **389,** 753-758.

47. Tamrazi,A., Carlson,K.E., Daniels,J.R., Hurth,K.M. and Katzenellenbogen,J.A. (2002) Estrogen receptor dimerization: Ligand binding regulates dimer affinity and dimer dissociation rate. *Mol. Endocrinol.,* **16,** 2706-2719.

48. Koide,A., Zhao,C., Naganuma,M., Abrams,J., Deighton-Collins,S., Skafar,D.F. and Koide,S. (2007) Identification of regions within the F domain of the human estrogen receptor alpha that are important for modulating transactivation and protein-protein interactions. *Mol. Endocrinol.,* **21,** 829-842.

49. Bjornstrom,L. and Sjoberg,M. (2005) Mechanisms of estrogen receptor signaling: Convergence of genomic and nongenomic actions on target genes. *Mol. Endocrinol.,* **19,** 833-842.

50. Carroll,J.S., Meyer,C.A., Song,J., Li,W., Geistlinger,T.R., Eeckhoute,J., Brodsky,A.S., Keeton,E.K., Fertuck,K.C., Hall,G.F., et al. (2006) Genome-wide analysis of estrogen receptor binding sites. *Nat. Genet.,* **38,** 1289-1297.

51. Fullwood,M.J., Liu,M.H., Pan,Y.F., Liu,J., Xu,H., Mohamed,Y.B., Orlov,Y.L., Velkov,S., Ho,A., Mei,P.H., et al. (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature,* **462,** 58-64.

52. Luscombe,N.M., Austin,S.E., Berman,H.M. and Thornton,J.M. (2000) An overview of the structures of protein-DNA complexes. *Genome Biol.,* **1,** REVIEWS001.

53. Klug,A. (2010) The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu. Rev. Biochem.,* **79,** 213-231.

54. Miller,J., McLachlan,A.D. and Klug,A. (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from xenopus oocytes. *EMBO J.,* **4,** 1609-1614.

55. Choo,Y. and Klug,A. (1994) Toward a code for the interactions of zinc fingers with DNA: Selection of randomized fingers displayed on phage. *Proc. Natl. Acad. Sci. U. S. A.,* **91,** 11163-11167.

56. Segal,D.J., Dreier,B., Beerli,R.R. and Barbas,C.F.,3rd. (1999) Toward controlling gene expression at will: Selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc. Natl. Acad. Sci. U. S. A.,* **96,** 2758-2763.

57. Sera,T. and Uranga,C. (2002) Rational design of artificial zinc-finger proteins using a nondegenerate recognition code table. *Biochemistry,* **41,** 7074-7081.

58. Boch,J., Scholze,H., Schornack,S., Landgraf,A., Hahn,S., Kay,S., Lahaye,T., Nickstadt,A. and Bonas,U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science,* **326,** 1509-1512.

59. Moscou,M.J. and Bogdanove,A.J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science,* **326,** 1501.

60. Chockalingam,K., Chen,Z., Katzenellenbogen,J.A. and Zhao,H. (2005) Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. *Proc. Natl. Acad. Sci. U. S. A.,* **102,** 5691-6.

61. Tedesco,R., Thomas,J.A., Katzenellenbogen,B.S. and Katzenellenbogen,J.A. (2001) The estrogen receptor: A structure-based approach to the design of new specific hormone-receptor combinations. *Chem. Biol.,* **8,** 277-287.

62. Gallinari,P., Lahm,A., Koch,U., Paolini,C., Nardi,M.C., Roscilli,G., Kinzel,O., Fattori,D., Muraglia,E., Toniatti,C., et al. (2005) A functionally orthogonal estrogen receptor-based transcription switch specifically induced by a nonsteroid synthetic ligand. *Chemistry & Biology,* **12,** 883-93.

63. Schwimmer,L.J., Rohatgi,P., Azizi,B., Seley,K.L. and Doyle,D.F. (2004) Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules. *Proc. Natl. Acad. Sci. U. S. A.,* **101,** 14707-14712.

64. Skretas,G. and Wood,D.W. (2005) Regulation of protein activity with small-molecule-controlled inteins. *Protein Sci.,* **14,** 523-532.

65. Zeidler,M.P., Tan,C., Bellaiche,Y., Cherry,S., Hader,S., Gayko,U. and Perrimon,N. (2004) Temperature-sensitive control of protein activity by conditionally splicing inteins. *Nat. Biotechnol.,* **22,** 871-876.

66. Beerli,R.R., Segal,D.J., Dreier,B. and Barbas,C.F.,3rd. (1998) Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U. S. A.,* **95,** 14628-14633.

67. Karzenowski,D., Potter,D.W. and Padidam,M. (2005) Inducible control of transgene expression with ecdysone receptor: Gene switches with high sensitivity, robust expression, and reduced size. *BioTechniques,* **39,** 191-196.

68. Imhof,M.O., Chatellard,P. and Mermod,N. (2002) Comparative study and identification of potent eukaryotic transcriptional repressors in gene switch systems. *J. Biotechnol.,* **97,** 275-285.

69. Koide,S. (2009) Generation of new protein functions by nonhomologous combinations and rearrangements of domains and modules. *Curr. Opin. Biotechnol.,* **20,** 398-404.

70. Jackel,C., Kast,P. and Hilvert,D. (2008) Protein design by directed evolution. *Annu. Rev. Biophys.,* **37,** 153-173.

71. Rebar,E.J., Huang,Y., Hickey,R., Nath,A.K., Meoli,D., Nath,S., Chen,B., Xu,L., Liang,Y., Jamieson,A.C., et al. (2002) Induction of angiogenesis in a mouse model using engineered transcription factors. *Nat. Med.,* **8,** 1427-1432.

72. Sera,T. (2009) Zinc-finger-based artificial transcription factors and their applications. *Adv. Drug Deliv. Rev.,* **61,** 513-526.

73. Beerli,R.R., Dreier,B. and Barbas,C.F.,3rd. (2000) Positive and negative regulation of endogenous genes by designed transcription factors. *Proc. Natl. Acad. Sci. U. S. A.,* **97,** 1495-1500.

74. Bartsevich,V.V. and Juliano,R.L. (2000) Regulation of the MDR1 gene by transcriptional repressors selected using peptide combinatorial libraries. *Mol. Pharmacol.,* **58,** 1-10.

75. Zhang,L., Spratt,S.K., Liu,Q., Johnstone,B., Qi,H., Raschke,E.E., Jamieson,A.C., Rebar,E.J., Wolffe,A.P. and Case,C.C. (2000) Synthetic zinc finger transcription factor action at an endogenous chromosomal site. activation of the human erythropoietin gene. *J. Biol. Chem.,* **275,** 33850-33860.

76. Liu,P.Q., Rebar,E.J., Zhang,L., Liu,Q., Jamieson,A.C., Liang,Y., Qi,H., Li,P.X., Chen,B., Mendel,M.C., et al. (2001) Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. activation of vascular endothelial growth factor A. *J. Biol. Chem.,* **276,** 11323-11334.

77. Bae,K.H., Kwon,Y.D., Shin,H.C., Hwang,M.S., Ryu,E.H., Park,K.S., Yang,H.Y., Lee,D.K., Lee,Y., Park,J., et al. (2003) Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nat. Biotechnol.,* **21,** 275-280.

78. Tachikawa,K., Schroder,O., Frey,G., Briggs,S.P. and Sera,T. (2004) Regulation of the endogenous VEGF-A gene by exogenous designed regulatory proteins. *Proc. Natl. Acad. Sci. U. S. A.,* **101,** 15225-15230.

79. Kwon,H.S., Shin,H.C. and Kim,J.S. (2005) Suppression of vascular endothelial growth factor expression at the transcriptional and post-transcriptional levels. *Nucleic Acids Res.,* **33,** e74.

80. Maeder,M.L., Thibodeau-Beganny,S., Osiak,A., Wright,D.A., Anthony,R.M., Eichtinger,M., Jiang,T., Foley,J.E., Winfrey,R.J., Townsend,J.A., et al. (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell,* **31,** 294-301.

81. Pollock,R., Giel,M., Linher,K. and Clackson,T. (2002) Regulation of endogenous gene expression with a small-molecule dimerizer. *Nat. Biotechnol.,* **20,** 729-733.

82. Dent,C.L., Lau,G., Drake,E.A., Yoon,A., Case,C.C. and Gregory,P.D. (2007) Regulation of endogenous gene expression using small molecule-controlled engineered zinc-finger protein transcription factors. *Gene Ther.,* **14,** 1362-1369.

83. Jouvenot,Y., Ginjala,V., Zhang,L., Liu,P.Q., Oshimura,M., Feinberg,A.P., Wolffe,A.P., Ohlsson,R. and Gregory,P.D. (2003) Targeted regulation of imprinted genes by synthetic zinc-finger transcription factors. *Gene Ther.,* **10,** 513-522.

84. Falke,D., Fisher,M., Ye,D. and Juliano,R.L. (2003) Design of artificial transcription factors to selectively regulate the pro-apoptotic bax gene. *Nucleic Acids Res.,* **31,** e10.

85. Tan,S., Guschin,D., Davalos,A., Lee,Y.L., Snowden,A.W., Jouvenot,Y., Zhang,H.S., Howes,K., McNamara,A.R., Lai,A., et al. (2003) Zinc-finger protein-targeted gene regulation: Genomewide single-gene specificity. *Proc. Natl. Acad. Sci. U. S. A.,* **100,** 11997-12002.

86. Beltran,A., Parikh,S., Liu,Y., Cuevas,B.D., Johnson,G.L., Futscher,B.W. and Blancafort,P. (2007) Re-activation of a dormant tumor suppressor gene maspin by designed transcription factors. *Oncogene,* **26,** 2791-2798.

87. Ren,D., Collingwood,T.N., Rebar,E.J., Wolffe,A.P. and Camp,H.S. (2002) PPARgamma knockdown by engineered transcription factors: Exogenous PPARgamma2 but not PPARgamma1 reactivates adipogenesis. *Genes Dev.,* **16,** 27-32.

88. Bartsevich,V.V., Miller,J.C., Case,C.C. and Pabo,C.O. (2003) Engineered zinc finger proteins for controlling stem cell fate. *Stem Cells,* **21,** 632-637.

89. East,A.K., Mauchline,T.H. and Poole,P.S. (2008) Biosensors for ligand detection. *Adv. Appl. Microbiol.,* **64,** 137-166.

90. Gillies,A.R., Skretas,G. and Wood,D.W. (2008) Engineered systems for detection and discovery of nuclear hormone-like compounds. *Biotechnol. Prog.,* **24,** 8-16.

91. Soto,A.M., Sonnenschein,C., Chung,K.L., Fernandez,M.F., Olea,N. and Serrano,F.O. (1995) The E-SCREEN assay as a tool to identify estrogens: An update on estrogenic environmental pollutants. *Environ. Health Perspect.,* **103 Suppl 7,** 113-122.

92. Coldham,N.G., Dave,M., Sivapathasundaram,S., McDonnell,D.P., Connor,C. and Sauer,M.J. (1997) Evaluation of a recombinant yeast cell estrogen screening assay. *Environ. Health Perspect.,* **105,** 734-742.

93. Bovee,T.F., Helsdingen,R.J., Koks,P.D., Kuiper,H.A., Hoogenboom,R.L. and Keijer,J. (2004) Development of a rapid yeast estrogen bioassay, based on the expression of green fluorescent protein. *Gene,* **325,** 187-200.

94. Sanseverino,J., Gupta,R.K., Layton,A.C., Patterson,S.S., Ripp,S.A., Saidak,L., Simpson,M.L., Schultz,T.W. and Sayler,G.S. (2005) Use of saccharomyces cerevisiae BLYES expressing bacterial bioluminescence for rapid, sensitive detection of estrogenic compounds. *Appl. Environ. Microbiol.,* **71,** 4455-4460.

95. Wilkinson,J.M., Hayes,S., Thompson,D., Whitney,P. and Bi,K. (2008) Compound profiling using a panel of steroid hormone receptor cell-based assays. *J. Biomol. Screen.,* **13,** 755-765.

96. Chen,Z. and Zhao,H. (2003) A highly efficient and sensitive screening method for trans-activation activity of estrogen receptors. *Gene,* **306,** 127-134.

97. Chen,Z., Katzenellenbogen,B.S., Katzenellenbogen,J.A. and Zhao,H. (2004) Directed evolution of human estrogen receptor variants with significantly enhanced androgen specificity and affinity. *J. Biol. Chem.,* **279,** 33855-64.

98. Lee,H.S., Sasagawa,S., Kato,S., Fukuda,R., Horiuchi,H. and Ohta,A. (2006) Yeast two-hybrid detection systems that are highly sensitive to a certain kind of endocrine disruptors. *Biosci. Biotechnol. Biochem.,* **70,** 521-524.

99. Muddana,S.S. and Peterson,B.R. (2003) Fluorescent cellular sensors of steroid receptor ligands. *Chembiochem,* **4,** 848-855.

100. Umezawa,Y. (2005) Genetically encoded optical probes for imaging cellular signaling pathways. *Biosens. Bioelectron.,* **20,** 2504-2511.

101. De,S., Macara,I.G. and Lannigan,D.A. (2005) Novel biosensors for the detection of estrogen receptor ligands. *J. Steroid Biochem. Mol. Biol.,* **96,** 235-244.

102. Paulmurugan,R. and Gambhir,S.S. (2006) An intramolecular folding sensor for imaging estrogen receptor-ligand interactions. *Proc. Natl. Acad. Sci. U. S. A.,* **103,** 15883-15888.

103. Skretas,G. and Wood,D.W. (2005) A bacterial biosensor of endocrine modulators. *J. Mol. Biol.,* **349,** 464-474.

Chapter 2: Engineering Orthogonal Ligand Specificity in a Single Scaffold

2.1 Introduction

Gene regulation systems that provide temporal and spatial regulation of target gene expression in response to small molecule ligands (small molecule-dependent gene switches or circuits) are powerful tools for gene therapy, tissue engineering, metabolic engineering, and functional genomics (1, 2). Notably, small molecule-dependent gene switches were recently used to generate induced pluripotent stem cells (iPS) for regenerative medicine (3, 4). The need for orthogonal regulatory elements is seen from a surge in interest in creating gene circuits, inspired by their electrical counterparts (5).

Many different systems have been developed to regulate gene expression in mammalian cells (6). Some of these arose from prokaryotic gene regulation systems, while others are based on eukaryotic transcription factors. Most rely on the modularity of biological control systems through the linkage of specific DNA binding domains (DBD), activation domains (AD) such as VP16 from the herpes simplex virus (7) or p65 from the human NF-κB transcription factor (8), or repression domains such as the Kruppel-associated box (KRAB) from human zinc finger transcription factors (9), and may or may not be modulated by a ligand responsive domain (10).

An early and successful approach was the development of elements from the tetracycline resistance operon of *E. coli* (11). The initial system fused the activation domain of VP16 to the TetR protein producing the tTA construct which localizes an activating function to promoters containing the tetracycline operator sequence (*tetO*), to which tetR binds in the absence of the antibiotic tetracycline (12). This formed a tetracycline repressible system (TetOFF), since it resulted in gene expression in the absence of tetracycline that was turned off by its addition. Improvements were subsequently made including the use of the more effective tetracycline derivative doxycycline (13), mutation of the TetR domain resulting in the variant rtTA and the TetON system that functions by requiring doxycycline for DNA binding and gene expression (13), mutation to reduce basal activity of rtTA (14), engineering of the activation domain VP16 (15) or the use of the p65 activation domain (16), the addition of the KRAB repressor domain

(17), the addition of an androgen receptor ligand binding domain (18), and construction of autoregulatory systems (19).

Control systems have also been derived from other bacterial operons. These include response to the streptogramin group of antibiotics through the use of the repressor pristinamycin-induced protein and its DNA binding site from *E. coli* (20). Macrolide antibiotics, such as erythromycin, were used to control expression through an erythromycin dependent transactivator and its DNA operator sequence from *E. coli* (21). The *Pseudomonas putida* repressor CymR and operator *CuO* elements from the *p*-cymene degradation pathway were combined to give a cumate controlled system (22). Phage, bacterial, and mammalian components were combined to produce the coumermycin/novobiocin regulation system where the p65 activation domain is fused to the N-terminal domain of the *E. coli* gyrase B protein, which is in turn linked to the phage lambda repressor DNA binding domain. The presence of the antibiotic coumermycin from *Streptomyces* stimulates the dimerization of the gyrase B subunits and therefore also the DNA binding domains, which can then bind the lambda operator sequence in the controlled gene. Novobiocin prevents dimerization, turning the switch off (23). The quorum sensing system from *Streptomyces coelicolor* has also been converted to an expression control system regulated by the butyrolactone SCB1 (24).

Functional elements can be brought together by chemically induced dimerization. FKBP12 is protein that binds the immunosuppressant natural products FK506 and rapamycin (25). By using a dimeric form of FK506 called FK1012, two FKBP12 domains can be recruited and when fused to an activation domain and DNA binding domain this forms a ligand controlled gene expression system (26). Alternatively, the rapamycin-FKBP12 complex binds to FRAP (FKBP-rapamycin associated protein) (27), and by linking FKBP to a DNA binding domain and a FRAP fragment to p65, a rapamycin controlled switch was created (28). Alternative ligands for this system have been investigated to avoid the cell proliferation inhibitory effect of rapamycin (29).

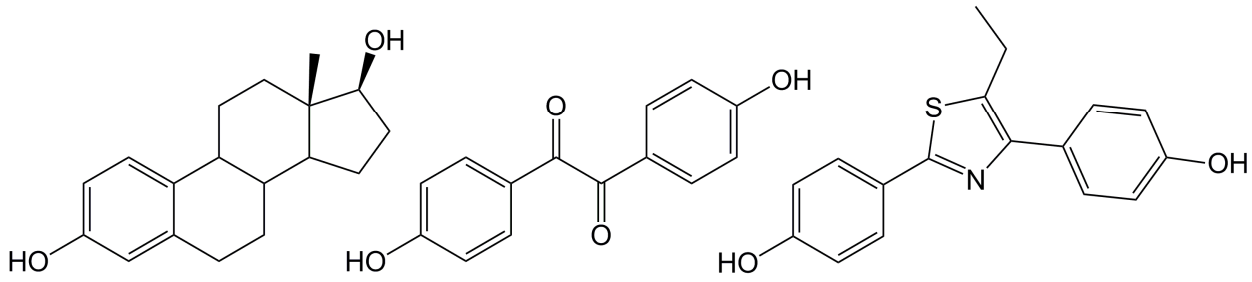
Other groups have taken inspiration from nuclear hormone receptors (NHRs) such as the progesterone, ecdysone, and estrogen receptors which offer desirable characteristics as gene switches for transcriptional control. These receptors are typically sequestered in the cytosol of

eukaryotic cells, becoming active only upon binding a small ligand, resulting in dimerization, translocation, and activation of promoters harboring specific responsive elements (30). With distinct domains for ligand binding, DNA binding and activation/repression functions, NHRs offer protein engineers the flexibility of creating chimerical transcriptional activators or repressors by modular design (31, 32). Access to organic synthesis of small molecule hormone-like compounds (33) makes these natural allosteric transcriptional switches attractive targets for engineering. Nuclear hormone receptors enable a wide range of target protein expression levels, tunable in a ligand dose-dependent manner. NHR ligand binding domains have also been combined with other proteins to enable posttranslational control of protein function (34). Specificity re-engineering approaches involving NHRs have typically involved mutating the ligand-binding domain (LBD) to a form that is not activated by the natural ligand, but is instead activated by a synthetic small molecule hormone-like compound inactive against the wild type receptor LBD. When using components from eukaryotes, it is important to prevent crosstalk between the natural and designed systems to prevent unintended activation by endogenous signaling pathways.

The antiprogestin system (35) is based on a progesterone receptor deletion mutant that binds to the progesterone antagonist mifepristone (RU486), but not to progesterone. The truncated ligand binding domain was fused to the GAL4 DNA binding domain and the VP16 activation domain creating a gene switch responsive to mifepristone (36). A longer region of the ligand binding domain, missing only the C-terminal 19 amino acids, was found to perform better (37), and the use of the p65 activation domain also improved performance (38).

The ecdysone system is based on the insect steroid receptor involved in *Drosophila* morphogenesis. The ecdysone receptor was initially shown to respond to certain ecdysteroids when expressed in mammalian cells (39), but was subsequently discovered to act as a heterodimer with the *Drosophila* ultraspiracle protein (40), and as a gene switch acted best when heterodimerized with the retinoid X receptor and replacing the N-terminal region of EcR with the VP16 activator (41). Further development of the system involved testing different insect ligand binding domains and their mutagenesis to respond to particular ligands, the optimization of activation domains, and the construction of homodimerizing variants (42, 43, 44, 45, 46).

The estrogen receptor has also been used as the basis for gene control. Initial constructs showed the feasibility of adding the VP16 activation domain and the use of the GAL4 DNA binding domain for use in mammalian and yeast systems (32, 47, 48, 49, 50). The use of the KRAB repression domain and other DNA binding domains has been explored (51, 52, 53, 54, 55, 56). Alteration of the ER-α LBD allows the use of ligands other than 17β estradiol (E_2), including 4-hydroxytamoxifen and other synthetic compounds (57, 58). Previous research in the Zhao lab by Dr. Karuppiah Chockalingam and Ka-Chun Lai produced variants sensitive to the compounds 4,4'-dihydroxybenzil (DHB) or 2,4-di(4-hydroxyphenyl)-5-ethylthiazole (L9) (Figure 2.1) by using sequential site-saturation mutagenesis and error prone PCR of the ligand binding pocket of the hERα-LBD (59, 60).



E_2 DHB L9

Figure 2.1. Chemical structures of E_2, DHB, and L9.

The ability to control gene expression through *cis* regulatory sequences in DNA and *trans* acting protein or RNA molecules has allowed the construction of many artificial gene circuits and synthetic biology devices. These have been pursued for reasons including the generation of fundamental knowledge regarding natural gene transcription and developmental networks (61), and the creation of systems that can be of use in industry or medicine (62). Beginning with pioneering work in *E. coli* (63, 64), synthetic regulatory systems such as logic gates, transcriptional cascades, toggle switches, and oscillators were subsequently developed in mammalian systems (65).

The goal of this chapter was to demonstrate that the engineered ligand/LBD pairs DHB/7S and L9/L7E (Figure 2.2) can be used to independently control genes in yeast and mammalian cells, and the results have been published (66). The orthogonal pairs were used to control expression of

fluorescent proteins in yeast. A synthetic biology approach was taken to show the functionality of these pairs in mammalian cells through the construction of gene circuit logic gates. Finally, it was shown that the engineered LBDs did not act predictably when fused to other functional domains. Portions of this chapter are reproduced with permission from: McLachlan MJ, Chockalingam K, Lai KC, and Zhao H. (2009) "Directed evolution of orthogonal ligand specificity in a single scaffold". *Angewandte Chemie International Edition* **48**:7783-7786. Copyright Wiley-VCH Verlag GmbH & Co. KGaA.

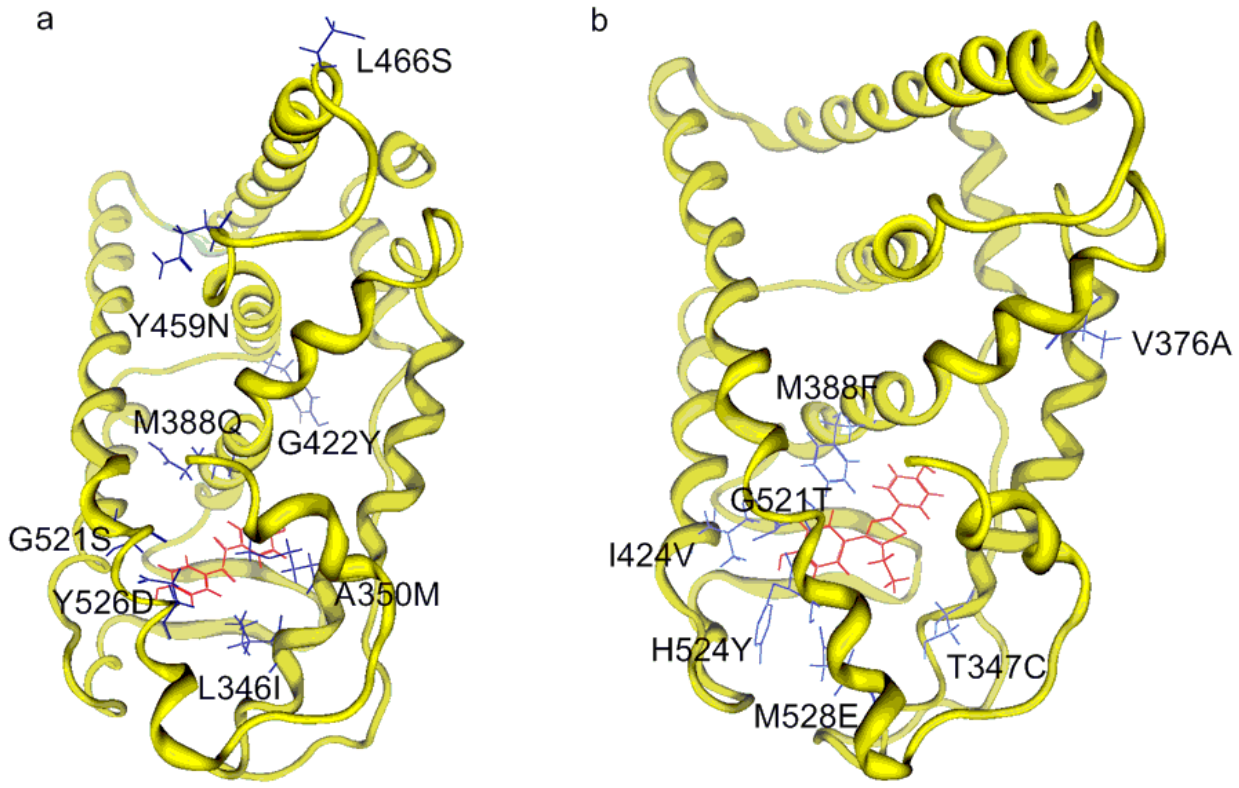


Figure 2.2. The locations of mutated sites are shown for mutant 7S (a) and L7E (b). The DHB and L9 ligands were docked into the cavity left by E_2 in the crystal structure of hERα-LBD-E_2 (PDB code 1GWR), as described elsewhere (59).

2.2 Results and Discussion

2.2.1 Orthogonality of Engineered Ligand Binding Domains

In order to determine the orthogonality of the DHB and L9 responsive ligand binding domains, the yeast two-hybrid system used for screening was employed to determine the sensitivity to E2, DHB, and L9 for the best mutant generated at each of the seven rounds of protein engineering (59, 60). The system functions by ligand induced activation of a gene switch construct consisting of the GAL4 DNA binding domain (amino acids 1-147) fused to ERα ligand binding domain variants (amino acids 312-595). This interacts with the steroid receptor coactivator 1 (SRC1) / GAL4 activation domain construct to transcribe a GAL-responsive histidine biosynthesis reporter gene, allowing growth of the yeast in media lacking histidine. Remarkable sensitivity and specificity were apparent with 7S/DHB and L7E/L9 pairs. While not quite as sensitive as the native estrogen receptor is for its ligand E_2, both 7S and L7E showed nanomolar sensitivity to their respective ligands (Figure 2.3 and Table 2.1). The mutant 7S was 850-fold more sensitive toward DHB than the WT, around 300,000-fold less sensitive to E_2, representing over 2.5×10^9 fold specificity shift, and only showed a response to L9 at micromolar concentrations. The mutant L7E was 3,400-fold more sensitive to L9 than the WT, and showed no response to E_2 or DHB at concentrations up to 10^{-5} M. Throughout the process of engineering, the DHB and L9 series of mutants were fully orthogonal to each other, exhibiting no crossreactivity at micromolar or lower ligand concentrations. (Book 4 pg8-28)

Table 2.1. The best mutant for each round of mutagenesis, the mutations it contained, and the point of half-maximal induction (EC50) for the ligands E2, DHB, and L9 from the yeast two-hybrid growth assay. The data is shown as the mean and standard error of the mean expressed in nM (nd = not detected) (66).

Clone	E2 EC50	DHB EC50	L9 EC50
WT	0.15 (0.044)	560 (170)	4800 (380)
1S (A350M)	0.71 (0.26)	10 (4.1)	46000 (18000)
2S (A350M, L346I)	33 (0.57)	3.5 (0.068)	39000 (5700)
3S (A350M, L346I, M388Q)	936 (280)	39 (2.8)	nd
4S (A350M, L346I, M388Q, G521S, Y526D)	nd	44 (4.4)	nd
5E (A350M, L346I, M388Q, G521S, Y526D, F461L, V560M)	nd	69 (19)	nd
5S (A350M, L346I, M388Q, G521S, Y526D, G442Y)	200000 (39000)	8.3 (1.6)	nd
6S (A350M, L346I, M388Q, G521S, Y526D, G442Y, L466S)	35000 (14000)	2.5 (0.30)	130000 (35000)
7S (A350M, L346I, M388Q, G521S, Y526D, G442Y, Y459N, L466S)	44000 (26000)	0.66 (0.16)	20000 (5200)
L1S (G521T)	350 (11)	30000 (3400)	370 (16)
L2S (G521T, H524Y)	nd	nd	300 (17)
L3S (G521T, H524Y, M388F)	nd	nd	46 (4.8)
L4S (G521T, H524Y, M388F, T347C)	nd	nd	36 (2.0)
L5S (G521T, H524Y, M388F, T347C, M528E)	nd	nd	3.7 (0.18)
L6S (G521T, H524Y, M388F, T347C, M528E, I424V)	nd	nd	4.1 (0.18)
L7E (G521T, H524Y, M388F, T347C, M528E, I424V, V376A)	nd	nd	1.4 (0.40)

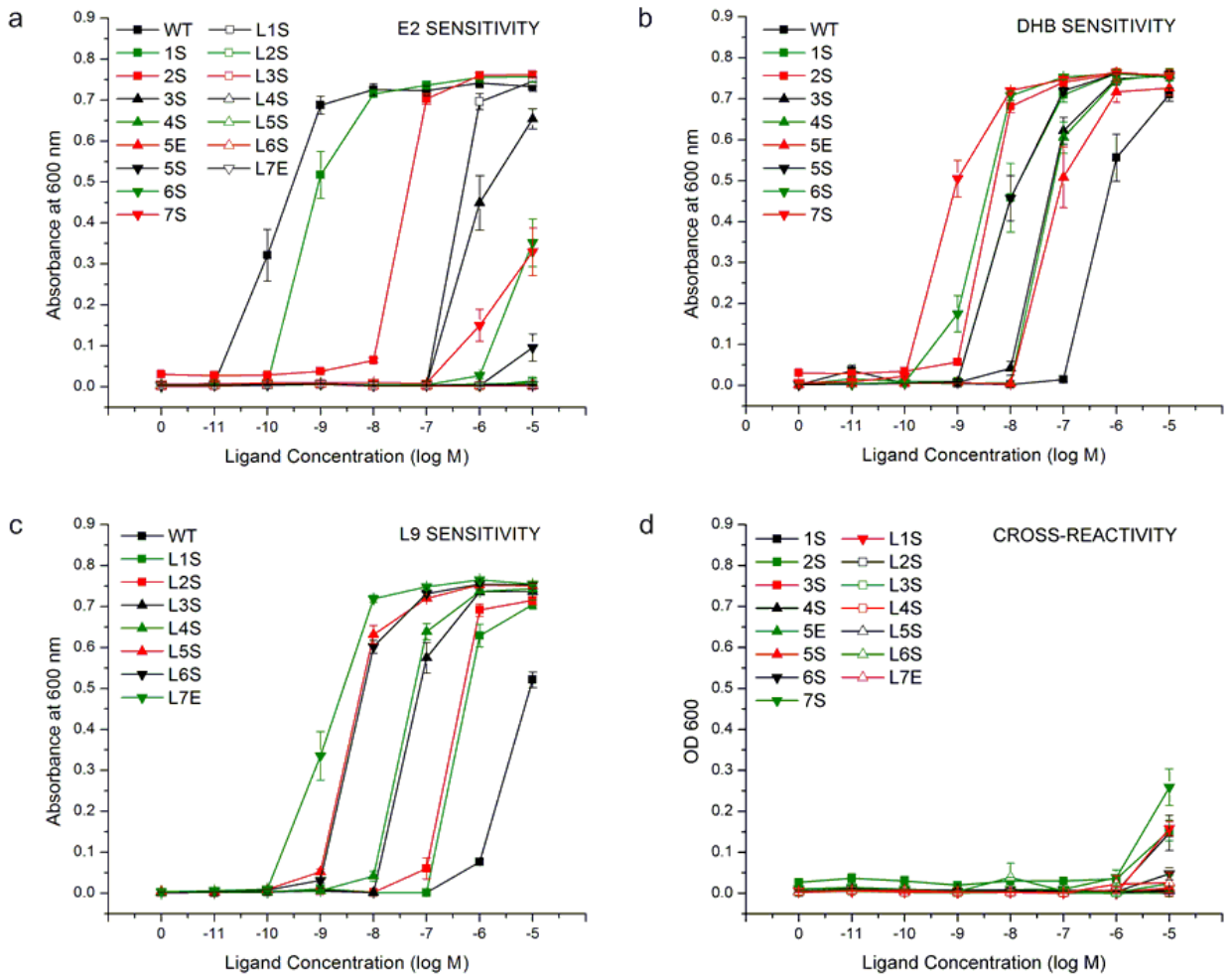


Figure 2.3. Sensitivity of gene switch mutants to hormone in a yeast two-hybrid growth assay. (a) Growth of mutants in estradiol. (b) Growth of DHB mutants with DHB ligand. (c) Growth of L9 mutants with L9 ligand. (d) Growth of L9 mutants with DHB, and DHB mutants with L9.

The engineering of ligand binding domains has proven to be a fruitful research area. Amongst the family of NHRs, the LBD of the human estrogen receptor α (hERα) has proven to be a particularly versatile platform for the creation of orthogonal ligand-receptor pairs (57, 58, 59, 67, 68, 69). This is perhaps not surprising given that the evolution of nuclear hormone receptors has followed a similar process, with mutations conferring sensitivity to new ligands deriving from an ancestral estrogen receptor gene (70). Early research showed that single residue changes could affect ligand binding of ERα, often discriminating between agonists and antagonists such as the G521R mutation which decreases E2 binding but retains binding of 4-hydroxytamoxifen (57, 71, 72). More extensive mutagenesis approaches with the specific goal of creating orthogonal

ligand/ligand binding domain pairs have been attempted (67, 68, 69, 73, 74). Of these efforts, and efforts to shift the ligand specificity of other NHR LBDs (75), the 7S/DHB and L7E/L9 pairs represent two of the best examples of ligand specificity reversal reported (66).

2.2.2 Control of Target Genes in Yeast

By adding the 7S/DHB and L7E/L9 pairs to the ERα/E2 pair, three orthogonal estrogen receptor α based gene switches were available in yeast (Figure 2.4a). As these had so far only been used to control the GAL4 induced HIS3 reporter gene, it was necessary to demonstrate that other genes could be controlled. The green fluorescent protein *GFP*, yellow fluorescent protein *YFP*, and a red fluorescent protein *mCherry* were chosen as target genes for their ease of detection. They were cloned under the control of the yeast *gal1* promoter which contains GAL4 response elements, hence enabling control by the GAL4-LBD gene switch. An alternative construct design was also investigated by cloning the fluorescent proteins under the $GAL4_{17mers(x3)}$-*TATACYC1* promoter that controls the β-galactosidase gene in the yeast two-hybrid strain YRG2. This promoter showed a high constitutive expression of the reporters in the absence of the gene switch and so was not investigated further.

Three reporter strains were constructed. These contained the wildtype switch with GFP, the 7S switch with mCherry, and the L7E switch with YFP. In the absence of ligand the strains showed little expression of the reporter, although some expression of YFP was visible. When these strains were exposed to each ligand, activation was observed only in the presence of the correct ligand. E2 stimulated expression of GFP, DHB stimulated expression of mCherry, and L9 stimulated expression of YFP (Figure 2.4b). (Book 4 pg 100)

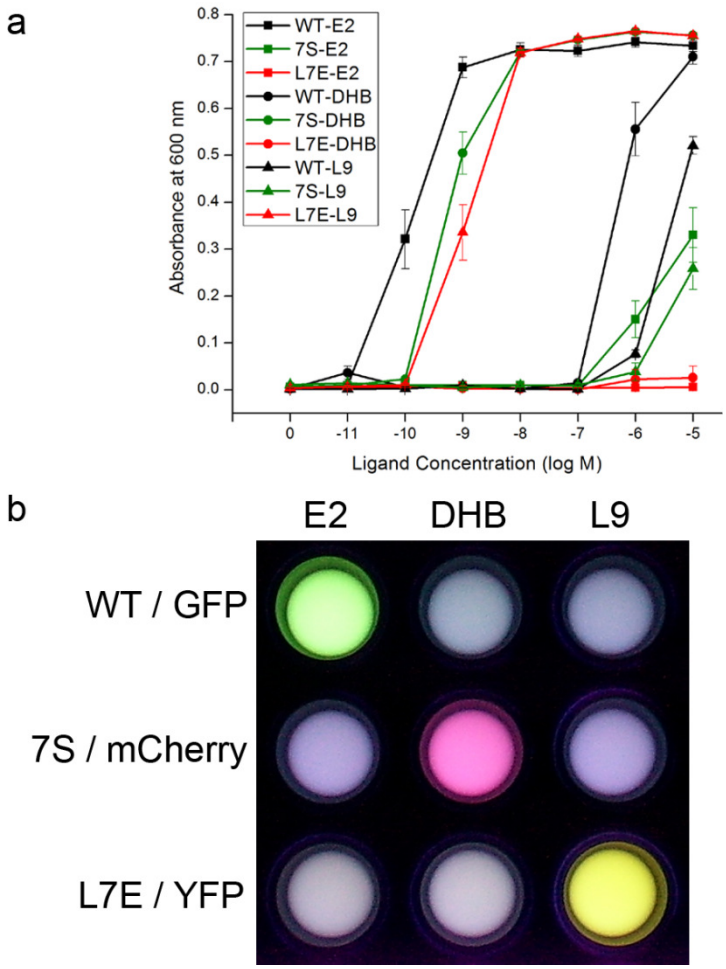


Figure 2.4. (a) Dose-response curves for WT, 7S, and L7E in the presence of E_2, DHB, or L9. (b) Gene switches controlling the expression of fluorescent proteins in yeast from a *Gal1* promoter. Row one contains GFP + WT, row two contains mCherry + 7S, row three contains YFP + L7E. The three columns contain 10^{-7} M E_2, DHB, or L9 from left to right.

2.2.3 Construction of Mammalian Gene Switches

The next goal was to examine the performance of the gene switches in mammalian cells. The GAL4-LBD constructs from yeast were cloned into the pCMV5 vector which allows strong constitutive expression in mammalian cells from the cytomegalovirus derived promoter. When expressed in HeLa cells with a GAL4 responsive luciferase reporter, 7S and L7E gave 11-fold and 33-fold induction, respectively, compared to a 67-fold induction from the wild type LBD (Figure 2.5a). The engineered switches showed no appreciable induction with either E2 or the non-specific ligand up to 1000 nM, while the wild type ligand binding domain construct showed

induction with both DHB and L9 ligands at 100 nM and above (Figure 2.5b). (Book2 pg281 – Book3 pg11)

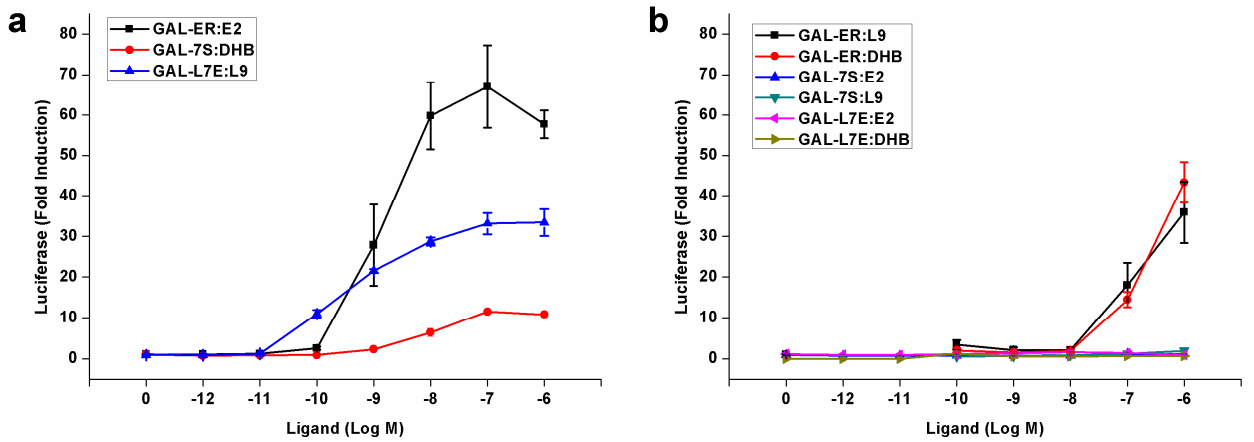


Figure 2.5. Induction of luciferase in HeLa cells by the GAL4-LBD format gene switches. (a) Gene switch activity in the presence of the target ligand. (b) Gene switch activity in the presence of non-target ligands.

2.2.4 Mammalian Gene Switch Domain Addition

The addition of other functional domains to the engineered LBDs would allow for targeting of different promoters, give more control over the activation, or allow for repression of genes. Attempts to use the 7th round L9 mutant ligand binding domain in different construct designs gave unanticipated results. To test its versatility, L7E was cloned into a construct containing the engineered zinc finger DNA binding domain N1 (whose binding target is GGG GTA GAA) and the activation domain VP64 (four copies of the VP16 core activation domain residues) (31). When expressed in HeLa cells with an N1 responsive luciferase reporter, a high basal activity and minimal induction was observed (Figure 2.6a). To examine if this effect was due to the N1 DBD, GAL4 DBD-containing constructs were used with the activators VP16 or VP64 (Figure 2.6b,c). In both cases the activation domain caused an increase in basal expression, resulting in only a few fold activation with the L9 ligand. Using different lengths of the C-terminal F domain of the ERα, a region that modulates the receptor function (76, 77), did not relieve the ligand independent expression (Figure 2.6b). Another activation domain, p65, was tested, but this too showed the same pattern whether it was present at the N or C terminus of the gene switch (Figure

2.6d). Thus changes in the DBD or activation domains resulted in a poor ability of L7E to confer ligand dependent activity. (Book 2 pg 171, 200, 231)

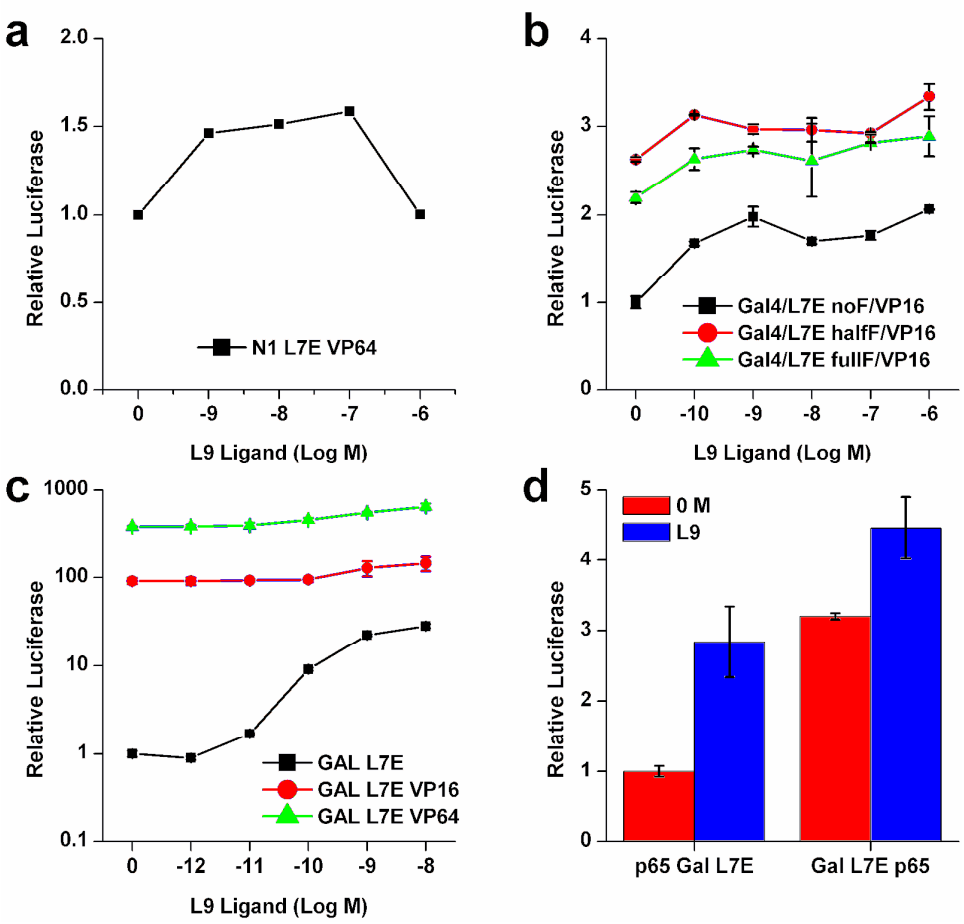


Figure 2.6. Performance of the L7E ligand binding domain in mammalian constructs. (a) L7E was cloned into a construct containing the N1 engineered zinc finger DNA binding domain and the VP64 activation domain and transfected with an N1 responsive luciferase reporter. (b) The effect of the F domain on constructs containing the VP16 activation domain. (c) The effect of adding VP16 or VP64 activation domains to the functional GAL-L7E construct. (d) The effect of adding the p65 activation domain at the N or C terminus. Data is the mean and standard deviation of duplicate samples.

During a set of unrelated experiments it was observed that the addition of a tag modified the performance of the gene switch. A codon optimized yellow fluorescent protein (EYFP), or the N terminal or C terminal half when split at position 155, was fused next to the p65 domain and reduced preferentially the basal level of reporter expression. This effect was seen both with the 4S DHB LBD and the L7E L9 LBD. When fused at the C terminus of p65-GAL-4S, EYFP

reduced basal expression of a luciferase reporter by 40 % and the induced expression only 20 % (Figure 2.7a). The fusion of the half tags reduced basal expression even further while decreasing the induced level by three fold. By reducing basal expression more than the induced expression, the result was a higher ratio of induction, with the YN C-terminal tag almost doubling the ratio to over 100 fold induction. The phenomenon was more pronounced when the tags were added at the N-terminus, with 600 fold induction observed with YN-p65-GAL-4S (Figure 2.7b). (Book 5 pg 251-283)

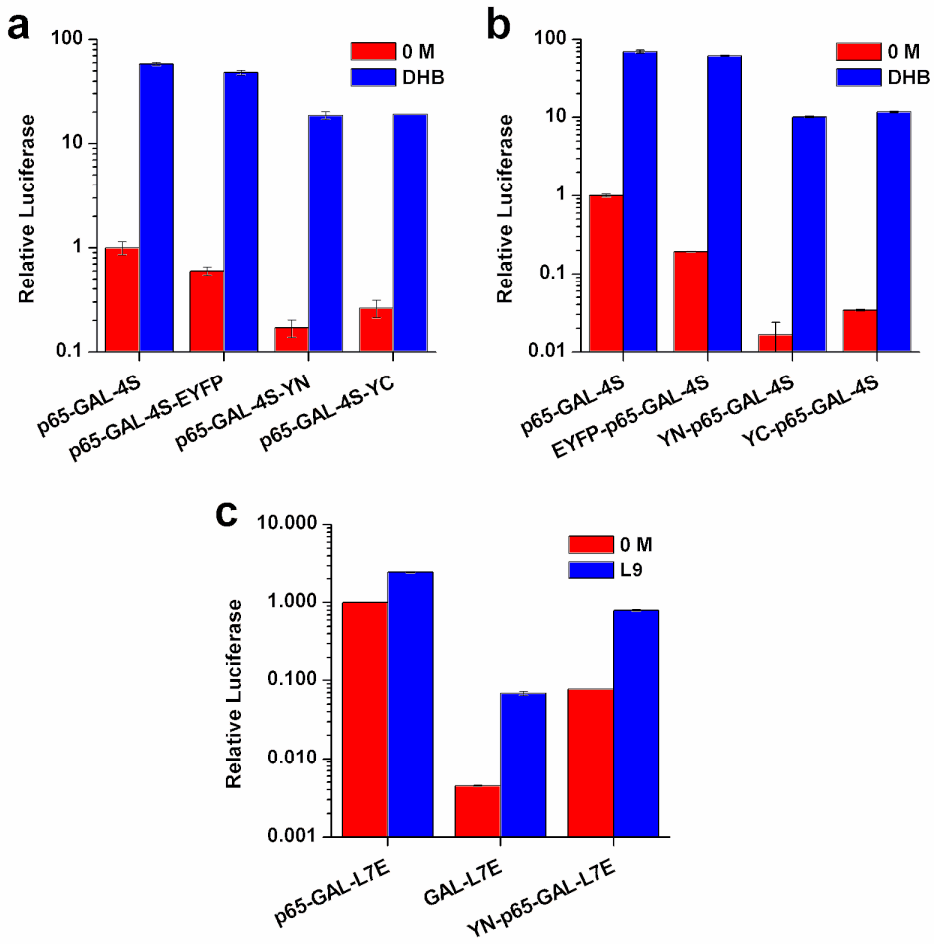


Figure 2.7. EYFP tag has an effect on gene switch performance. The EYFP gene, YN (EYFP residues 1-154), or YC (EYFP residues 155-239) were fused to the C-terminus (a) or N-terminus (b) of the p65-GAL-4S gene switch. (c) Recovery of some inducibility of the p65-GAL-L7E construct by the fusion of YN at its N-terminus. Samples show the induction of a *GalTATA*-luciferase reporter in the absence and presence of 1 µM DHB or L9 ligand, normalized such that the p65-GAL-LBD construct in the absence of ligand equaled 1. Data is the mean and standard deviation of duplicate samples.

To see if this tag could improve the ligand-dependent induction of the L7E LBD, YN was fused at the N-terminus of p65-GAL-L7E and compared to the well performing GAL-L7E (Figure 2.7c). A 10 fold induction ratio was observed, which was an improvement over the non-tagged construct even though it was lower than the ratio for the GAL-L7E and still showed a high basal activity. Fusion tags such as His_6, maltose binding protein, hemagglutinin (HA), or fluorescent proteins are commonly used in biochemistry to allow for easy purification of proteins (78), or visualization of their activity *in vivo*. While tags are often specifically used to improve expression or solubility of a protein (79), their effect on protein function can sometimes be overlooked, leading to misattribution of an effect to the protein of interest rather that the tag itself (80, 81). These results are a useful reminder that all components of a fusion protein construct should be examined, even those that are not expected to have a functional effect.

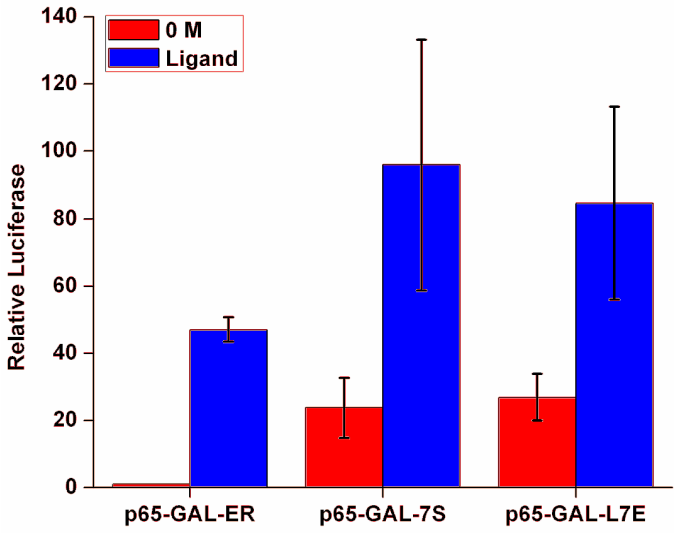


Figure 2.8. Luciferase induction in HeLa cells with the addition of p65 activation domain and GAL4 activation domain at the N-terminus of the LBDs of hERα, 7S, and L7E. The expression in the absence and presence of 1 μM ligand was normalized such that the 0 M value for p65GALER was equal to 1.

To determine whether the suboptimal performance was specific to the L7E engineered ligand binding domain, the 7th round DHB LBD 7S was also tested in different configurations. A 20-fold increase in basal expression was observed when using the NFκB p65 activation domain with either the 7S or L7E LBD, in contrast to its use with the wild type LBD (Figure 2.8). Although the presence of ligand did lead to an induction of luciferase, it was only by 4-fold. The use of a different activation domain, the strong viral VP16, indicated that the increase in basal expression

was not due to p65 but was caused by the engineered ligand binding domain itself. Since the 4th round DHB mutant had been observed to function well with additional domains (Figure 2.7) a series of constructs were made comprising the GAL4 DNA binding domain, the 4S to 7S DHB mutants, and VP16. The fourth round mutant 4S gave strong induction when measured in a HeLa cell based luciferase assay, but the 5S, 6S, and 7S variants showed increasing basal expression at each round (Figure 2.9). A similar panel of L9 variants was examined, containing the fourth to seventh round LBDs, and it was found that all showed high basal activity (data not shown). (Book4 pg121-127, 163-167)

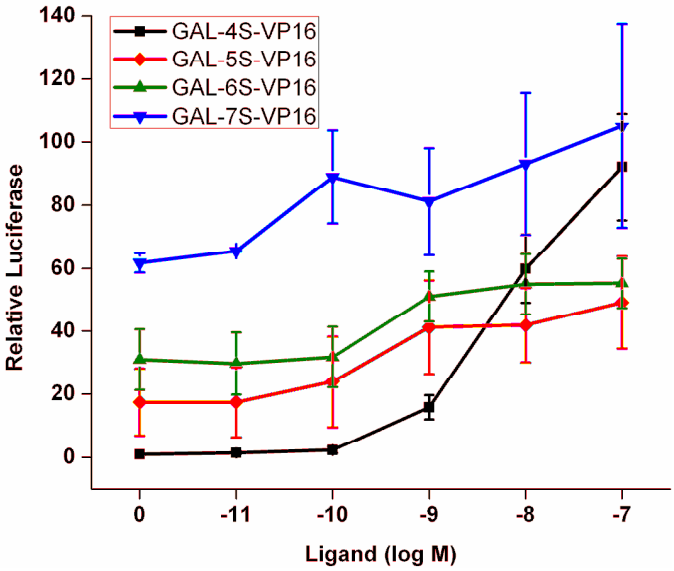


Figure 2.9. Luciferase induction in HeLa cells for DHB mutants 4S, 5S, 6S, and 7S with a C-terminal VP16 activation domain. The values were normalized such that the 0 M value for GAL4SVP16 was equal to 1.

To expand the use of the gene switch system, the ability to repress transcription was investigated through the use of the KRAB domain. A constitutively expressing reporter was constructed by inserting the strong cytomegalovirus immediate early promoter (*CMV*) upstream of the Gal response elements of *GalTATA*-Luciferase, giving a highly expressing construct that should be repressible through *trans* factors binding at the Gal elements. The KRAB domain was cloned at either the N or C terminus of GAL-LBD constructs and its activity on the reporter was monitored in HeLa cells (Figure 2.10). The KRAB domain was functional. However, repression was predominantly ligand independent with 85-95 % repression in the absence of ligand. Only 2-4 fold further repression was observed upon addition of ligand. (Book 2 pg 279, Book 3 pg 18)

Since the use of the GAL DBD with the KRAB domain was not ligand dependent, the use of the full length ERα was investigated. Previous research by de Haan *et al*. had shown that tagging both the N and C termini of ERα yielded an effective estrogen response element targeting repressor, despite some ligand independent activity (82). After fusing the KRAB domain at each end of the full length 7S variant, 4-fold repression in the absence of ligand was observed, but this was relieved by the introduction of the G400V (50) mutation to the ligand binding domain (Figure 2.11). This K7SK-G400V construct had 40 % repression without DHB, but a further 10 fold repression upon the addition of ligand.

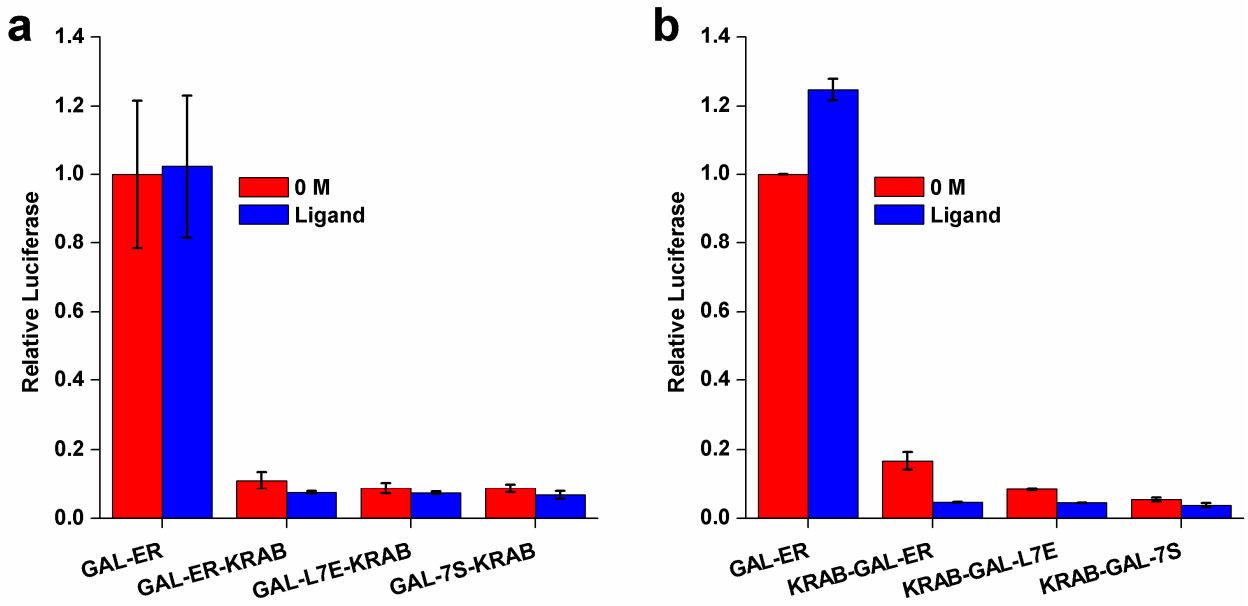


Figure 2.10. KRAB repression domain added to GAL-LBD gene switch constructs at the C terminus (a) or the N terminus (b). Constructs were expressed in HeLa cells with the constitutive but GAL responsive reporter CMVgalTATA-Luciferase in the presence or absence of 1 μM estradiol, L9, or DHB ligand.

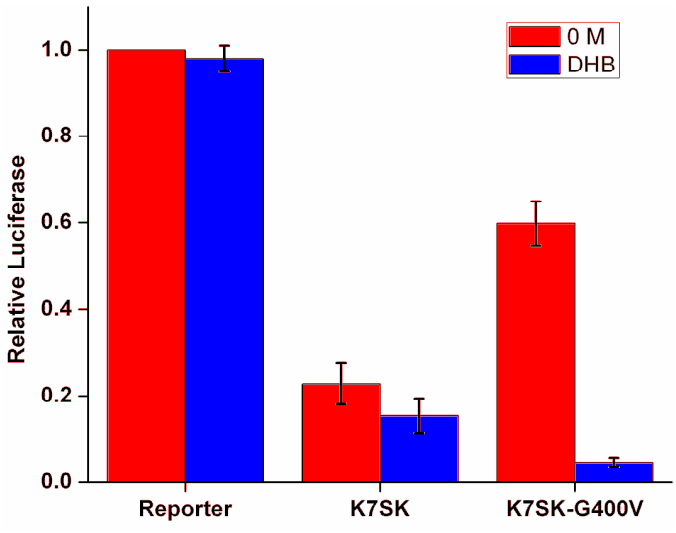


Figure 2.11. Luciferase repression in HeLa cells from switches with the KRAB repression domain attached to the 7S LBD and human ERα DNA binding domain, in the absence and presence of 1 μM DHB. The values were normalized such that the 0 M value for the *CMV-2ERE*-Luciferase reporter was equal to 1.

These results highlight the difficulty in engineering truly modular components for synthetic biology. While the 4S DHB responsive LBD showed flexibility in terms of the domains it could function with, the 7th round DHB and L9 LBDs were recalcitrant to context modifications. Due to unanticipated interactions between protein domains, one cannot simply cut and paste components, showing that the engineering approach should be as close as possible to the intended application. Failure rates of up to 76 % have been reported for the assembly of designed zinc finger modules (83), supporting the complexity of modular design shown here. A crystal structure of the full length estrogen receptor bound to both ligand and DNA would be useful in identifying inter domain contacts that could be implicated in the constitutive activity of the engineered LBDs once they are in the context of different domains. However, only isolated domains of the ERα including the LBD and DNA binding domain have been crystallized (84, 85). Indeed, the only intact nuclear receptor complex crystal structure, the PPAR-γ-RXR-α heterodimer bound to DNA, does show interactions between the PPAR-γ ligand binding domain and both DNA binding domains (86).

2.2.5 Construction of Logic Gates

To demonstrate the utility of the engineered ligand/receptor pairs for creation of genetic circuits in mammalian systems, the logic functions AND, OR, NAND, and NOR were implemented in HeLa cells using a luciferase reporter, Gal4 or estrogen receptor DNA binding domain promoter response elements (*GalRE* or *ERE*), wild type or engineered ligand binding domains, and the KRAB repressor domain. These were transiently transfected into HeLa cells in the presence of ligand and the luciferase gene expression was monitored.

The AND gate (Figure 2.12a) was constructed with a constitutively expressed GAL4-L7E switch, a *GalRE* controlled ER-7S switch, and an *ERE* controlled luciferase gene. Induction of luciferase was mainly observed when both L9 and DHB ligands were present (140-fold), although some leaky expression was observed when either ligand was present alone. The OR gate (Figure 2.12b) was constructed by using a luciferase reporter downstream of both GAL4 and estrogen receptor response elements, with constitutively expressed GAL4-L7E and ER-7S switches. This resulted in luciferase induction in the presence of L9 (100-fold), DHB (40-fold), or both (600-fold). The NAND gate (Figure 2.12c) was made with the strong constitutive *CMV* promoter with an inserted *ERE*, a constitutively expressed GAL4-L7E switch, and a GAL responsive KRAB-ER-KRAB G400V switch. The most repression of luciferase was observed in the desired L9 + E2 condition. However, this was only 4-fold with 2-fold repression observed in the presence of E2 alone. This indicates that presence of some leaky expression or ligand-independent repression. The NOR gate (Figure 2.12d) was comprised of constitutively expressed KRAB-ER-L7E-KRAB-G400V or KRAB-ER-7S-KRAB-G400V switches, with luciferase downstream of a constitutive cytomegalovirus (*CMV*) promoter with an inserted *ERE*. This showed 10-16 fold repression in the presence of L9, DHB, or both ligands. (Book4 pg 162)

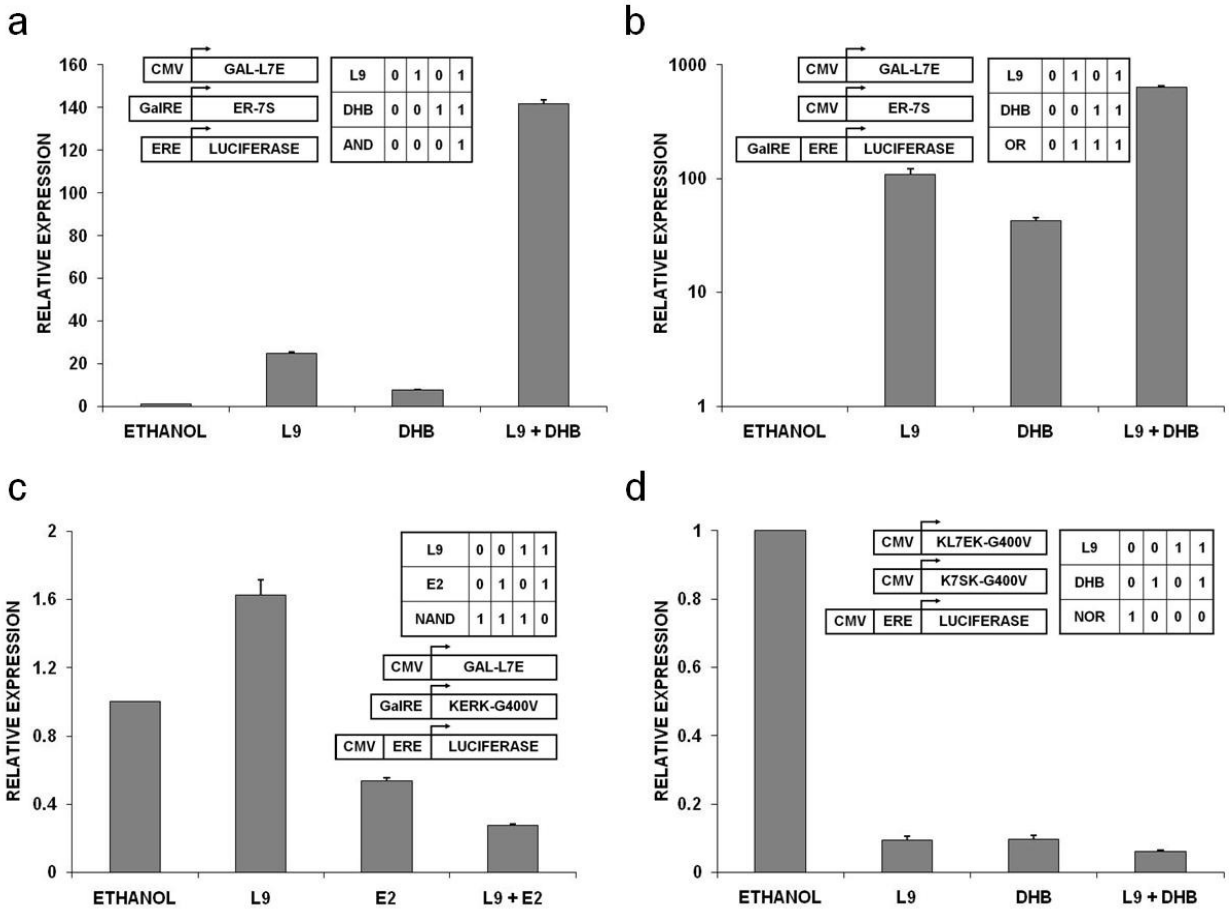


Figure 2.12. Logic gates in HeLa cells. The graphs show luciferase activity (mean, standard error) normalized to β-galactosidase expression, and with the value for the ethanol treatment taken as equal to 1. (a) AND gate with 10^{-8} M L9, DHB, or both. (b) OR gate with 10^{-8} M L9, DHB, or both. (c) NAND gate with 10^{-6} M L9, E2, or both. (d) NOR gate with 10^{-6} M L9, DHB, or both.

As seen from Figure 2.12, some of the logic gates were more successful than others. The NOR gate performed best, giving a consistent repression in each of the ligand conditions. This is likely due to both gene switches being of a similar format, and the presence of only one type of binding site in the reporter preventing any synergistic effects. While the OR gate showed high induction in the presence of any ligand, there was a large difference between the single ligands and the DHB + L9 condition. The synergistic effect observed could be due to both the *GalRE* and *ERE* being bound, increasing the recruitment of the transcriptional machinery. Synergistic action at promoters is a well established concept whereby binding to a promoter element is enhanced by dimerization, cooperative binding of transcription factors, or heterosynergy between bound transcription factors and the basal transcriptional machinery (87, 88, 89). Promoter engineering

could be attempted to create interference between the binding sites, such as interspersing the Gal and estrogen response elements, in order to reduce the synergy. Other synthetic logic gates in mammalian cells have also shown leaky expression and synergy. In the creation of BioLogic gates from the Fussenegger group (90), tetracycline, streptogramin, macrolide, and butyrolactone control systems were used to create multiple logic functions. Their NAND gate showed repression with single inputs, and their OR gate showed a two-fold difference in induction between single and double inputs.

More extensive synthetic devices have been engineered in mammalian cells. An oscillating signal generator was constructed using destabilized fluorescent proteins controlled by tetracycline and pristinamycin elements (91). Linkage of a synthetic circuit to an endogenous metabolite was demonstrated by controlling the level of uric acid in urate oxidase deficient mice (92). Intercellular signaling was demonstrated through the use of acetaldehyde as a signaling molecule (93). These two examples are of interest because the controlling signal is generated *in vivo*. Future directions are likely to involve generating tighter control over expression, more interfaces with endogenous signals, better engineering of DNA binding domains to target any gene, and an output that affects cell morphology or development (94, 95).

2.3 Conclusions and Outlook

In summary, a single scaffold was used in an effective and repeatable manner for the generation of two orthogonal ligand-receptor pairs, which had virtually a complete reversal of ligand specificity. The three pairs (E2/ERα, DHB/7S, and L9/L7E) were shown to independently control target genes in yeast and despite a lack of receptor based logic gates (5, 65), they were successfully combined into higher order functions as exemplified by the creation of logic gates in mammalian cells. It is envisioned that any number of further pairs could be generated in the same way for the development of new small molecule regulated gene expression systems.

However, important limitations of the modular engineering approach were revealed in these experiments. Although the GAL4-LBD construct that was used for engineering in yeast performed well in mammalian cells, the most sensitive DHB and L9 responsive ligand binding domains were not transferable to other construct designs without the presence of high basal

activity. These effects were neither predictable from the yeast screening system nor from the corresponding ERα construct, and suggests that the screening system used for engineering gene switches should be as close to the intended application as possible.

The different ligand controlled gene expression systems developed have different advantages and disadvantages. The tetracycline-based system has enjoyed wide usage and is generally considered to have few adverse effects (96), but the non-human elements have been shown to generate an immune response, especially in primates, that reduces the system's effectiveness (97, 98). Another limitation with the tetracycline-based system is that tetracycline and its related compounds will accumulate in developing bone and teeth, so they are contraindicated during pregnancy and childhood (99). The VP16 activation domain used in this and other systems has been shown to inhibit gene expression when present at high levels (100, 101), likely because it sequesters factors involved in transcription.

The nuclear hormone receptor systems have the benefit of using mainly human-derived components, which should reduce immunogenicity if the system is to be used for gene therapy. The exception is the ecdysone receptor since its origins lie in insect proteins, and this could also interact with the retinoid X receptor in mammalian systems. The progesterone system suffers from detrimental public perception due to its ligand, mifepristone/RU486, being an abortive agent even though it would be used at a much lower concentration than needed for this effect. Some of the systems are limited to a single ligand or group of related ligands, making it necessary to use different systems to control multiple genes simultaneously. The receptor systems that allow engineering of the ligand binding domain could use the same basic format but with different LBDs for this purpose.

The benefits of the engineered estrogen receptor-based system include the use of only human-derived protein domains, an engineering approach that allows tailoring of the system to respond to drugs having favorable kinetics and pharmacology, and the ability to use different ligands to control multiple targets independently. An ideal gene switch system would be immune-system friendly (which may be more likely with mammalian components), flexible in the promoters that it can target by using different DNA binding domains, and be expandable to new inducers.

A number of future directions are possible. The ligands used here are not ideal for gene switch systems since they are neither readily available nor approved for use in humans. Approved compounds with known pharmacology could be used as new targets for specificity engineering. New screening methods could be developed to avoid unexpected results from switching from a yeast based environment to a mammalian one.

2.4 Materials and Methods

2.4.1 Chemicals, Media, and Reagents

DNA polymerase, restriction enzymes, and T4 DNA liagse were from New England Biolabs (Beverly, MA). Chemicals were from Sigma-Aldrich (Carlsbad, CA). Media components were from BD Biosciences (San Jose, CA), and premixed synthetic complete amino acid dropout mixes were from MP Biomedicals (Irvine, CA) or Sigma-Aldrich (Carlsbad, CA). Cell culture media were from the UIUC cell media facility (Urbana, IL). The ligands DHB and L9 were synthesized by Professor John Katzenellenbogen's laboratory (Urbana, IL). DNA isolation and miniprep kits were from Qiagen (Valencia, CA), and yeast miniprep kits were from Zymo Research (Orange, CA).

2.4.2 Yeast Two-Hybrid Ligand Titration

Frozen stocks of the DHB and L9 series mutants were streaked on synthetic complete medium minus leucine / tryptophan (SC-LW) plates and grown for two days. A colony was picked and grown overnight to saturation in SC-LW media and then diluted to an OD of 0.002 in synthetic complete medium minus leucine / tryptophan / histidine (SC-LWH). 190 µL of the diluted yeast culture was placed into wells of a flat bottomed 96-well plate, to which was added 10 µL of ligand diluted 50-fold in SC-LWH media to give a final ligand dilution of 1000-fold. Ligand concentrations of E2, DHB, and L9 of 10^{-11} M to 10^{-5} M were tested, as well as the absence of any ligand. Plates were grown at 30 °C in a box containing wet paper towels to minimize evaporation. After 26-28 hours of growth, each well was resuspended and the optical density at 600 nm was measured using a Spectramax 190 microplate reader (Molecular Devices, Sunnyvale CA).

2.4.3 Fluorescent Yeast

Plasmids were transformed into yeast strain YM4271 (*MATa*, *ura3-52*, *his3-200*, *ade2-101*, *lys2-801*, *leu2-3, 112*, *trp1-901*, *tyr1-501*, *gal4-Δ512*, *gal80-Δ538*, *ade5::hisG*) (Clontech, Mountain View, CA) by the lithium acetate method (102) in various combinations, and selected for on synthetic complete medium minus leucine / tryptophan / uracil (SC-LWU) plates. The GFP reporter was transformed in conjunction with the yeast two-hybrid plasmids pBD-Gal4-WT and pGAD424-SRC1. The mCherry reporter was transformed with pBD-Gal4-7S and pGAD424-SRC1. The YFP was transformed with pBD-Gal4-L7E and pGAD424-SRC1. The strains were grown overnight in SC-LWU media at 30 °C and 200 μL was inoculated into 2 mL yeast extract peptone adenine dextrose (YPAD) media with 10^{-7} M ligand and grown overnight. The cells were pelleted, washed with phosphate buffered saline (PBS), resuspended in 50 μL PBS, and transferred to a 96-well plate. The plate was placed on a UV illuminator (Spectroline, Westbury NY) producing 435 nm light and photographed.

2.4.4 Cell Culture and Transformation

HeLa cells were grown in minimal essential medium (MEM) plus 1 mM sodium pyruvate, and 10% fetal bovine serum (FBS) (UIUC Cell Media Facility, Urbana, IL) at 37 °C with 5% carbon dioxide. When cells were 80% confluent, they were trypsinized and split into 24-well plates with MEM media plus 1 mM sodium pyruvate, and 5% charcoal dextran stripped calf serum (UIUC Cell Media Facility). Cells were grown for 24 hours until they were over 90% confluent and transfected using 1.5 μL lipofectamine 2000 (Invitrogen, Carlsbad CA), 100 ng β-galactosidase expression plasmid, 690 ng luciferase reporter plasmid, and 10 ng of the relevant gene switch plasmids per well. After four hours, the media was changed to MEM media plus 1 mM sodium pyruvate, and 5% charcoal dextran stripped calf serum, plus E_2, DHB, or L9 ligand. Cells were incubated for 24 hours, then lysed and assayed for luciferase activity using the Luciferase Assay System (Promega, Madison WI). Luciferase levels were normalized to β-galactosidase expression.

2.4.5 Cloning of Expression Constructs

Standard molecular techniques were used to create the constructs used in these experiments. For the mammalian constructs the *ERE* controlled luciferase construct (2ERE-pS2-pGL3-Luc) has been used previously (59). Gal response elements or the constitutive expression element *CMV* were added by blunt ligation upstream of the *ERE* element. The p65 domain was cloned from a plasmid gifted from the Hillen laboratory. The N1 and KRAB domains were cloned from plasmids gifted from the Barbas laboratory. Constitutively expressed gene switches were cloned between the KpnI and BamHI sites of pCMV5. Inducible gene switches were cloned by replacing the luciferase gene under the relevant promoter. For the yeast fluorescent protein constructs, the *gal1* promoter was amplified by PCR from genomic DNA and cloned upstream of *GFP*, *mCherry*, or *YFP*. The constructs discussed in this chapter are shown in Table 2.2.

Table 2.2. Constructs used in this work.

HZ	Genotype	Plasmid	Details
2391	YRG2	pGAD424-SRC	*MATa ura3-52 his3-200 ade2-101 lys2-801 trp1-901 leu2-3 112 gal4-542 gal80-538 LYS2::UASGAL1-TATA GAL1-HIS3 URA3::UASGAL4 17mers(x3)-TATACYC1-lacZ*
886	YRG2-SRC	ER pBD Gal4-CAM	
887	YRG2-SRC	1S pBD Gal4-CAM	
888	YRG2-SRC	2S pBD Gal4-CAM	
889	YRG2-SRC	3S pBD Gal4-CAM	
890	YRG2-SRC	4S pBD Gal4-CAM	
891	YRG2-SRC	5E pBD Gal4-CAM	
892	YRG2-SRC	5S pBD Gal4-CAM	
893	YRG2-SRC	6S pBD Gal4-CAM	
894	YRG2-SRC	7S pBD Gal4-CAM	
895	YRG2-SRC	L1S pBD Gal4-CAM	
896	YRG2-SRC	L2S pBD Gal4-CAM	
897	YRG2-SRC	L3S pBD Gal4-CAM	
898	YRG2-SRC	L4S pBD Gal4-CAM	
899	YRG2-SRC	L5S pBD Gal4-CAM	
900	YRG2-SRC	L6S pBD Gal4-CAM	
901	YRG2-SRC	L7E pBD Gal4-CAM	
1034	YM4271		*MATa, ura3-52, his3-200, ade2-101, lys2-801, leu2-3, 112, trp1-901, tyr1-501, gal4-Δ512, gal80-Δ538, ade5::hisG*
1734	YM4271	gal1p-GFP-pRS426, pGAD424SRC, ER-pBD-Gal4-CAM	
1739	YM4271	gal1p-YFP-pRS426, pGAD424SRC, L7E-pBD-Gal4-CAM	
1740	YM4271	gal1p-mCherry-pRS426, pGAD424SRC, 7S-pBD-Gal4-CAM	
3124	DH5alpha	N1-L7E-VP64 pcDNA3	

Table 2.2 (cont). Constructs used in this work

HZ	Genotype	Plasmid	Details
620	DH5alpha	CMVgalTATA-Luciferase pGL3	
621	DH5alpha	Gal-L7E pCMV5	
622	DH5alpha	Gal-ER pCMV5	
623	DH5alpha	Gal-7S pCMV5	
624	DH5alpha	Gal-L7E-KRAB pCMV5	
625	DH5alpha	Gal-ER-KRAB pCMV5	
626	DH5alpha	Gal-7S-KRAB pCMV5	
631	DH5alpha	Gal-L7E-VP64 pCMV5	
634	DH5alpha	Gal-L7E half F-VP16 pCMV5	
635	DH5alpha	Gal-L7E full F –VP16 pCMV5	
636	DH5alpha	Gal-L7E no F -VP16 pCMV5	
638	DH5alpha	KRAB-Gal-L7E pCMV5	
639	DH5alpha	KRAB-Gal-ER pCMV5	
640	DH5alpha	KRAB-Gal-7S pCMV5	
641	DH5alpha	p65GalL7E pCMV5	
642	DH5alpha	p65GalER pCMV5	
643	DH5alpha	p65Gal7S pCMV5	
644	DH5alpha	Gal-L7E-p65 pCMV5	
3125	DH5alpha	p65-Gal-4S-EYFP pCMV5	
3126	DH5alpha	p65-Gal-4S-YN pCMV5	
3127	DH5alpha	p65-Gal-4S-YC pCMV5	
3128	DH5alpha	EYFP-p65-Gal-4S pCMV5	
3129	DH5alpha	YN-p65-Gal-4S pCMV5	
3130	DH5alpha	YC-p65-Gal-4S pCMV5	
3131	DH5alpha	YN-p65-Gal-L7E	
2392	DH5alpha	Gal4SVP16 pCMV5	
2393	DH5alpha	Gal5SVP16 pCMV5	
2394	DH5alpha	Gal6SVP16 pCMV5	
2395	DH5alpha	Gal7SVP16 pCMV5	
1162	DH5alpha	KRAB-ER(1-595)-KRAB pCMV5	
1163	DH5alpha	KRAB-7S(1-595)-KRAB pCMV5	
1164	DH5alpha	KRAB-L7E(1-595)-KRAB pCMV5	
648	DH5alpha	2ERE-pS2-Luciferase pGL2	
1321	DH5alpha	CMV-2ERE-pS2-Luciferase pGL2	
1115	DH5alpha	galTATA-ERE-pS2-Luciferase pGL3	
2398	DH5alpha	galTATA-Full7S(1-595) pGL3	
2399	DH5alpha	Full7S(1-595) pCMV5	
2400	DH5alpha	galTATA-KRAB-ER(1-595 G400V)-KRAB pCMV5	

2.5 References

1. Fussenegger,M. (2001) The impact of mammalian gene regulation concepts on functional genomic research, metabolic engineering, and advanced gene therapies. *Biotechnol. Progr.,* **17,** 1-51.

2. Toniatti,C., Bujard,H., Cortese,R. and Ciliberto,G. (2004) Gene therapy progress and prospects: Transcription regulatory systems. *Gene Ther.,* **11,** 649-57.

3. Wernig,M., Lengner,C.J., Hanna,J., Lodato,M.A., Steine,E., Foreman,R., Staerk,J., Markoulaki,S. and Jaenisch,R. (2008) A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. *Nat. Biotechnol.,* **26,** 916-924.

4. Markoulaki,S., Hanna,J., Beard,C., Carey,B.W., Cheng,A.W., Lengner,C.J., Dausman,J.A., Fu,D., Gao,Q., Wu,S., et al. (2009) Transgenic mice with defined combinations of drug-inducible reprogramming factors. *Nat. Biotechnol.,* **27,** 169-171.

5. Sayut,D.J., Kambam,P.K. and Sun,L. (2007) Engineering and applications of genetic circuits. *Mol. BioSyst.,* **3,** 835-40.

6. Weber,W. and Fussenegger,M. (2007) Novel gene switches. *Handbook of Experimental Pharmacology,* 73-105.

7. Cousens,D.J., Greaves,R., Goding,C.R. and O'Hare,P. (1989) The C-terminal 79 amino acids of the herpes simplex virus regulatory protein, Vmw65, efficiently activate transcription in yeast and mammalian cells in chimeric DNA-binding proteins. *EMBO J.,* **8,** 2337-2342.

8. Schmitz,M.L. and Baeuerle,P.A. (1991) The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B. *EMBO J.,* **10,** 3805-3817.

9. Margolin,J.F., Friedman,J.R., Meyer,W.K., Vissing,H., Thiesen,H.J. and Rauscher,F.J.,3rd. (1994) Kruppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. U. S. A.,* **91,** 4509-4513.

10. Ansari,A.Z. and Mapp,A.K. (2002) Modular design of artificial transcription factors. *Curr. Opin. Chem. Biol.,* **6,** 765-772.

11. Stieger,K., Belbellaa,B., Le Guiner,C., Moullier,P. and Rolling,F. (2009) In vivo gene regulation using tetracycline-regulatable systems. *Adv. Drug Deliv. Rev.,* **61,** 527-541.

12. Gossen,M. and Bujard,H. (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U. S. A.,* **89,** 5547-51.

13. Gossen,M., Freundlieb,S., Bender,G., Muller,G., Hillen,W. and Bujard,H. (1995) Transcriptional activation by tetracyclines in mammalian cells. *Science,* **268,** 1766-1769.

14. Urlinger,S., Baron,U., Thellmann,M., Hasan,M.T., Bujard,H. and Hillen,W. (2000) Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *Proc. Natl. Acad. Sci. U. S. A.,* **97,** 7963-7968.

15. Baron,U., Gossen,M. and Bujard,H. (1997) Tetracycline-controlled transcription in eukaryotes: Novel transactivators with graded transactivation potential. *Nucleic Acids Res.,* **25,** 2723-2729.

16. Urlinger,S., Helbl,V., Guthmann,J., Pook,E., Grimm,S. and Hillen,W. (2000) The p65 domain from NF-kappaB is an efficient human activator in the tetracycline-regulatable gene expression system. *Gene,* **247,** 103-110.

17. Deuschle,U., Meyer,W.K. and Thiesen,H.J. (1995) Tetracycline-reversible silencing of eukaryotic promoters. *Mol. Cell. Biol.,* **15,** 1907-1914.

18. Anastassiadis,K., Kim,J., Daigle,N., Sprengel,R., Scholer,H.R. and Stewart,A.F. (2002) A predictable ligand regulated expression strategy for stably integrated transgenes in mammalian cells in culture. *Gene,* **298,** 159-172.

19. Shockett,P., Difilippantonio,M., Hellman,N. and Schatz,D.G. (1995) A modified tetracycline-regulated system provides autoregulatory, inducible gene expression in cultured cells and transgenic mice. *Proc. Natl. Acad. Sci. U. S. A.,* **92,** 6522-6526.

20. Fussenegger,M., Morris,R.P., Fux,C., Rimann,M., von Stockar,B., Thompson,C.J. and Bailey,J.E. (2000) Streptogramin-based gene regulation systems for mammalian cells. *Nat. Biotechnol.,* **18,** 1203-1208.

21. Weber,W., Fux,C., Daoud-El Baba,M., Keller,B., Weber,C.C., Kramer,B.P., Heinzen,C., Aubel,D., Bailey,J.E. and Fussenegger,M. (2002) Macrolide-based transgene control in mammalian cells and mice. *Nat. Biotechnol.,* **20,** 901-907.

22. Mullick,A., Xu,Y., Warren,R., Koutroumanis,M., Guilbault,C., Broussau,S., Malenfant,F., Bourget,L., Lamoureux,L., Lo,R., et al. (2006) The cumate gene-switch: A system for regulated expression in mammalian cells. *BMC Biotechnol.,* **6,** 43.

23. Zhao,H.F., Boyd,J., Jolicoeur,N. and Shen,S.H. (2003) A coumermycin/novobiocin-regulated gene expression system. *Hum. Gene Ther.,* **14,** 1619-1629.

24. Weber,W., Malphettes,L., de Jesus,M., Schoenmakers,R., El-Baba,M.D., Spielmann,M., Keller,B., Weber,C.C., van de Wetering,P., Aubel,D., et al. (2005) Engineered streptomyces quorum-sensing components enable inducible siRNA-mediated translation control in mammalian cells and adjustable transcription control in mice. *J. Gene Med.,* **7,** 518-525.

25. Standaert,R.F., Galat,A., Verdine,G.L. and Schreiber,S.L. (1990) Molecular cloning and overexpression of the human FK506-binding protein FKBP. *Nature,* **346,** 671-674.

26. Ho,S.N., Biggar,S.R., Spencer,D.M., Schreiber,S.L. and Crabtree,G.R. (1996) Dimeric ligands define a role for transcriptional activation domains in reinitiation. *Nature,* **382,** 822-826.

27. Brown,E.J., Albers,M.W., Shin,T.B., Ichikawa,K., Keith,C.T., Lane,W.S. and Schreiber,S.L. (1994) A mammalian protein targeted by G1-arresting rapamycin-receptor complex. *Nature,* **369,** 756-758.

28. Rivera,V.M., Clackson,T., Natesan,S., Pollock,R., Amara,J.F., Keenan,T., Magari,S.R., Phillips,T., Courage,N.L., F.,C.,Jr, et al. (1996) A humanized system for pharmacologic control of gene expression. *Nat. Med.,* **2,** 1028-1032.

29. Liberles,S.D., Diver,S.T., Austin,D.J. and Schreiber,S.L. (1997) Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen. *Proc. Natl. Acad. Sci. U. S. A.,* **94,** 7825-7830.

30. Aranda,A. and Pascual,A. (2001) Nuclear hormone receptors and gene expression. *Physiol. Rev.,* **81,** 1269-304.

31. Beerli,R.R., Schopfer,U., Dreier,B. and Barbas,C.F.,3rd. (2000) Chemically regulated zinc finger transcription factors. *J. Biol. Chem.,* **275,** 32617-27.

32. Braselmann,S., Graninger,P. and Busslinger,M. (1993) A selective transcriptional induction system for mammalian cells based on Gal4-estrogen receptor fusion proteins. *Proc. Natl. Acad. Sci. U. S. A.,* **90,** 1657-61.

33. Fink,B.E., Mortensen,D.S., Stauffer,S.R., Aron,Z.D. and Katzenellenbogen,J.A. (1999) Novel structural templates for estrogen-receptor ligands and prospects for combinatorial synthesis of estrogens. *Chemistry & Biology,* **6,** 205-19.

34. Picard,D. (2000) Posttranslational regulation of proteins by fusions to steroid-binding domains. *Methods Enzymol.,* **327,** 385-401.

35. Nordstrom,J.L. (2003) The antiprogestin-dependent GeneSwitch system for regulated gene therapy. *Steroids,* **68,** 1085-1094.

36. Wang,Y., O'Malley,B.W.,Jr, Tsai,S.Y. and O'Malley,B.W. (1994) A regulatory system for use in gene transfer. *Proc. Natl. Acad. Sci. U. S. A.,* **91,** 8180-8184.

37. Wang,Y., Xu,J., Pierson,T., O'Malley,B.W. and Tsai,S.Y. (1997) Positive and negative regulation of gene expression in eukaryotic cells with an inducible transcriptional regulator. *Gene Ther.,* **4,** 432-441.

38. Burcin,M.M., Schiedner,G., Kochanek,S., Tsai,S.Y. and O'Malley,B.W. (1999) Adenovirus-mediated regulable target gene expression in vivo. *Proc. Natl. Acad. Sci. U. S. A.,* **96,** 355-60.

39. Christopherson,K.S., Mark,M.R., Bajaj,V. and Godowski,P.J. (1992) Ecdysteroid-dependent regulation of genes in mammalian cells by a drosophila ecdysone receptor and chimeric transactivators. *Proc. Natl. Acad. Sci. U. S. A.,* **89,** 6314-6318.

40. Yao,T.P., Segraves,W.A., Oro,A.E., McKeown,M. and Evans,R.M. (1992) Drosophila ultraspiracle modulates ecdysone receptor function via heterodimer formation. *Cell,* **71,** 63-72.

41. No,D., Yao,T.P. and Evans,R.M. (1996) Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. U. S. A.,* **93,** 3346-3351.

42. Kumar,M.B., Fujimoto,T., Potter,D.W., Deng,Q. and Palli,S.R. (2002) A single point mutation in ecdysone receptor leads to increased ligand specificity: Implications for gene switch applications. *Proc. Natl. Acad. Sci. U. S. A.,* **99,** 14710-14715.

43. Kumar,M.B., Potter,D.W., Hormann,R.E., Edwards,A., Tice,C.M., Smith,H.C., Dipietro,M.A., Polley,M., Lawless,M., Wolohan,P.R., et al. (2004) Highly flexible ligand binding pocket of ecdysone receptor: A single amino acid change leads to discrimination between two groups of nonsteroidal ecdysone agonists. *J. Biol. Chem.,* **279,** 27211-27218.

44. Karzenowski,D., Potter,D.W. and Padidam,M. (2005) Inducible control of transgene expression with ecdysone receptor: Gene switches with high sensitivity, robust expression, and reduced size. *BioTechniques,* **39,** 191-196.

45. Panguluri,S.K., Kumar,P. and Palli,S.R. (2006) Functional characterization of ecdysone receptor gene switches in mammalian cells. *FEBS J.,* **273,** 5550-5563.

46. Tavva,V.S., Palli,S.R., Dinkins,R.D. and Collins,G.B. (2008) Improvement of a monopartite ecdysone receptor gene switch and demonstration of its utility in regulation of transgene expression in plants. *FEBS J.,* **275,** 2161-2176.

47. Elliston,J.F., Tsai,S.Y., O'Malley,B.W. and Tsai,M.J. (1990) Superactive estrogen receptors. potent activators of gene expression. *J. Biol. Chem.,* **265,** 11517-11521.

48. Louvion,J.F., Havaux-Copf,B. and Picard,D. (1993) Fusion of GAL4-VP16 to a steroid-binding domain provides a tool for gratuitous induction of galactose-responsive genes in yeast. *Gene,* **131,** 129-134.

49. Lee,H.S., Aumais,J. and White,J.H. (1996) Hormone-dependent transactivation by estrogen receptor chimeras that do not interact with hsp90. evidence for transcriptional repressors. *J. Biol. Chem.,* **271,** 25727-25730.

50. Aumais,J.P., Lee,H.S., Lin,R. and White,J.H. (1997) Selective interaction of hsp90 with an estrogen receptor ligand-binding domain containing a point mutation. *J. Biol. Chem.,* **272,** 12229-12235.

51. Ma,Z.Q., Tsai,M.J. and Tsai,S.Y. (1999) Suppression of gene expression by tethering KRAB domain to promoter of ER target genes. *J. Steroid Biochem. Mol. Biol.,* **69,** 155-163.

52. Lou,J., Cao,W., Bernardin,F., Ayyanathan,K., RauscherIII,F.J. and Friedman,A.D. (2000) Exogenous cdk4 overcomes reduced cdk4 RNA and inhibition of G1 progression in hematopoietic cells expressing a dominant-negative CBF - a model for overcoming inhibition of proliferation by CBF oncoproteins. *Oncogene,* **19,** 2695-2703.

53. Xu,L., Zerby,D., Huang,Y., Ji,H., Nyanguile,O.F., de los Angeles,J.E. and Kadan,M.J. (2001) A versatile framework for the design of ligand-dependent, transgene-specific transcription factors. *Mol. Ther.,* **3,** 262-73.

54. Ayyanathan,K., Fredericks,W.J., Berking,C., Herlyn,M., Balakrishnan,C., Gunther,E. and Rauscher,F.J.,3rd. (2000) Hormone-dependent tumor regression in vivo by an inducible transcriptional repressor directed at the PAX3-FKHR oncogene. *Cancer Res.,* **60,** 5803-5814.

55. Wang,Q.F. and Friedman,A.D. (2002) CCAAT/enhancer-binding proteins are required for granulopoiesis independent of their induction of the granulocyte colony-stimulating factor receptor. *Blood,* **99,** 2776-2785.

56. Roscilli,G., Rinaudo,C.D., Cimino,M., Sporeno,E., Lamartina,S., Ciliberto,G. and Toniatti,C. (2002) Long-term and tight control of gene expression in mouse skeletal muscle by a new hybrid human transcription factor. *Mol. Ther.,* **6,** 653-63.

57. Whelan,J. and Miller,N. (1996) Generation of estrogen receptor mutants with altered ligand specificity for use in establishing a regulatable gene expression system. *J. Steroid Biochem. Mol. Biol.,* **58,** 3-12.

58. Littlewood,T.D., Hancock,D.C., Danielian,P.S., Parker,M.G. and Evan,G.I. (1995) A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. *Nucleic Acids Res.,* **23,** 1686-1690.

59. Chockalingam,K., Chen,Z., Katzenellenbogen,J.A. and Zhao,H. (2005) Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. *Proc. Natl. Acad. Sci. U. S. A.,* **102,** 5691-6.

60. Chockalingam,K. (2006) Engineering estrogen receptor-based gene switches and a superoxide dismutase for therapeutic applications. Ph.D. Thesis. University of Illinois at Urbana-Champaign.

61. Mukherji,S. and van Oudenaarden,A. (2009) Synthetic biology: Understanding biological design from synthetic circuits. *Nat. Rev. Genet.,* **10,** 859-871.

62. Aubel,D. and Fussenegger,M. (2010) Mammalian synthetic biology--from tools to therapies. *Bioessays,* **32,** 332-345.

63. Elowitz,M.B. and Leibler,S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature,* **403,** 335-338.

64. Gardner,T.S., Cantor,C.R. and Collins,J.J. (2000) Construction of a genetic toggle switch in escherichia coli. *Nature,* **403,** 339-342.

65. Greber,D. and Fussenegger,M. (2007) Mammalian synthetic biology: Engineering of sophisticated gene networks. *J. Biotechnol.,* **130,** 329-45.

66. McLachlan,M.J., Chockalingam,K., Lai,K.C. and Zhao,H. (2009) Directed evolution of orthogonal ligand specificity in a single scaffold. *Angew. Chem. Int. Ed.,* **48,** 7783-7786.

67. Gallinari,P., Lahm,A., Koch,U., Paolini,C., Nardi,M.C., Roscilli,G., Kinzel,O., Fattori,D., Muraglia,E., Toniatti,C., et al. (2005) A functionally orthogonal estrogen receptor-based transcription switch specifically induced by a nonsteroid synthetic ligand. *Chemistry & Biology,* **12,** 883-93.

68. Islam,K.M., Dilcher,M., Thurow,C., Vock,C., Krimmelbein,I.K., Tietze,L.F., Gonzalez,V., Zhao,H. and Gatz,C. (2009) Directed evolution of estrogen receptor proteins with altered ligand-binding specificities. *Protein Eng. Des. Sel.,* **22,** 45-52.

69. Kinzel,O., Fattori,D., Muraglia,E., Gallinari,P., Nardi,M.C., Paolini,C., Roscilli,G., Toniatti,C., Gonzalez Paz,O., Laufer,R., et al. (2006) A structure-guided approach to an orthogonal estrogen-receptor-based gene switch activated by ligands suitable for in vivo studies. *J. Med. Chem.,* **49,** 5404-7.

70. Thornton,J.W. (2001) Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. *Proc. Natl. Acad. Sci. U. S. A.,* **98,** 5671-5676.

71. Pakdel,F. and Katzenellenbogen,B.S. (1992) Human estrogen receptor mutants with altered estrogen and antiestrogen ligand discrimination. *J. Biol. Chem.,* **267,** 3429-3437.

72. Danielian,P.S., White,R., Hoare,S.A., Fawell,S.E. and Parker,M.G. (1993) Identification of residues in the estrogen receptor that confer differential sensitivity to estrogen and hydroxytamoxifen. *Mol. Endocrinol.,* **7,** 232-40.

73. Miller,N. and Whelan,J. (1998) Random mutagenesis of human estrogen receptor ligand binding domain identifies mutations that decrease sensitivity to estradiol and increase sensitivity to a diphenol indene-ol compound: Basis for a regulatable expression system. *J. Steroid Biochem. Mol. Biol.,* **64,** 129-135.

74. Tedesco,R., Thomas,J.A., Katzenellenbogen,B.S. and Katzenellenbogen,J.A. (2001) The estrogen receptor: A structure-based approach to the design of new specific hormone-receptor combinations. *Chem. Biol.,* **8,** 277-287.

75. Schwimmer,L.J., Rohatgi,P., Azizi,B., Seley,K.L. and Doyle,D.F. (2004) Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules. *Proc. Natl. Acad. Sci. U. S. A.,* **101,** 14707-14712.

76. Nichols,M., Rientjes,J.M. and Stewart,A.F. (1998) Different positioning of the ligand-binding domain helix 12 and the F domain of the estrogen receptor accounts for functional differences between agonists and antagonists. *EMBO J.,* **17,** 765-773.

77. Koide,A., Zhao,C., Naganuma,M., Abrams,J., Deighton-Collins,S., Skafar,D.F. and Koide,S. (2007) Identification of regions within the F domain of the human estrogen receptor alpha that are important for modulating transactivation and protein-protein interactions. *Mol. Endocrinol.,* **21,** 829-842.

78. Terpe,K. (2003) Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.,* **60,** 523-533.

79. Waugh,D.S. (2005) Making the most of affinity tags. *Trends Biotechnol.,* **23,** 316-320.

80. Brothers,S.P., Janovick,J.A. and Conn,P.M. (2003) Unexpected effects of epitope and chimeric tags on gonadotropin-releasing hormone receptors: Implications for understanding the molecular etiology of hypogonadotropic hypogonadism. *J. Clin. Endocrinol. Metab.,* **88,** 6107-6112.

81. Wu,J. and Filutowicz,M. (1999) Hexahistidine (His6)-tag dependent protein dimerization: A cautionary tale. *Acta Biochim. Pol.,* **46,** 591-599.

82. de Haan,G., Chusacultanachai,S., Mao,C., Katzenellenbogen,B.S. and Shapiro,D.J. (2000) Estrogen receptor-KRAB chimeras are potent ligand-dependent repressors of estrogen-regulated gene expression. *J. Biol. Chem.,* **275,** 13493-501.

83. Ramirez,C.L., Foley,J.E., Wright,D.A., Muller-Lerch,F., Rahman,S.H., Cornu,T.I., Winfrey,R.J., Sander,J.D., Fu,F., Townsend,J.A., et al. (2008) Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat. Methods,* **5,** 374-375.

84. Schwabe,J.W., Chapman,L., Finch,J.T. and Rhodes,D. (1993) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: How receptors discriminate between their response elements. *Cell,* **75,** 567-578.

85. Tanenbaum,D.M., Wang,Y., Williams,S.P. and Sigler,P.B. (1998) Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. *Proc. Natl. Acad. Sci. U. S. A.,* **95,** 5998-6003.

86. Chandra,V., Huang,P., Hamuro,Y., Raghuram,S., Wang,Y., Burris,T.P. and Rastinejad,F. (2008) Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. *Nature,* **456,** 350-356.

87. Georges,A.B., Benayoun,B.A., Caburet,S. and Veitia,R.A. (2010) Generic binding sites, generic DNA-binding domains: Where does specific promoter recognition come from? *FASEB J.,* **24,** 346-356.

88. Carey,M., Kolman,J., Katz,D.A., Gradoville,L., Barberis,L. and Miller,G. (1992) Transcriptional synergy by the epstein-barr virus transactivator ZEBRA. *J. Virol.,* **66,** 4803-4813.

89. Barkhem,T., Haldosen,L.A., Gustafsson,J.A. and Nilsson,S. (2002) Transcriptional synergism on the pS2 gene promoter between a p160 coactivator and estrogen receptor-alpha depends on the coactivator subtype, the type of estrogen response element, and the promoter context. *Mol. Endocrinol.,* **16,** 2571-2581.

90. Kramer,B.P., Fischer,C. and Fussenegger,M. (2004) BioLogic gates enable logical transcription control in mammalian cells. *Biotechnol. Bioeng.,* **87,** 478-484.

91. Tigges,M., Marquez-Lago,T.T., Stelling,J. and Fussenegger,M. (2009) A tunable synthetic mammalian oscillator. *Nature,* **457,** 309-312.

92. Kemmer,C., Gitzinger,M., Daoud-El Baba,M., Djonov,V., Stelling,J. and Fussenegger,M. (2010) Self-sufficient control of urate homeostasis in mice by a synthetic circuit. *Nat. Biotechnol.,* **28,** 355-360.

93. Weber,W., Daoud-El Baba,M. and Fussenegger,M. (2007) Synthetic ecosystems based on airborne inter- and intrakingdom communication. *Proc. Natl. Acad. Sci. U. S. A.,* **104,** 10435-10440.

94. Lu,T.K., Khalil,A.S. and Collins,J.J. (2009) Next-generation synthetic gene networks. *Nat. Biotechnol.,* **27,** 1139-1150.

95. Haynes,K.A. and Silver,P.A. (2009) Eukaryotic systems broaden the scope of synthetic biology. *J. Cell Biol.,* **187,** 589-596.

96. Reboredo,M., Kramer,M.G., Smerdou,C., Prieto,J. and De Las Rivas,J. (2008) Transcriptomic effects of tet-on and mifepristone-inducible systems in mouse liver. *Hum. Gene Ther.,* **19,** 1233-1247.

97. Latta-Mahieu,M., Rolland,M., Caillet,C., Wang,M., Kennel,P., Mahfouz,I., Loquet,I., Dedieu,J.F., Mahfoudi,A., Trannoy,E., et al. (2002) Gene transfer of a chimeric trans-activator is immunogenic and results in short-lived transgene expression. *Hum. Gene Ther.,* **13,** 1611-1620.

98. Lena,A.M., Giannetti,P., Sporeno,E., Ciliberto,G. and Savino,R. (2005) Immune responses against tetracycline-dependent transactivators affect long-term expression of mouse erythropoietin delivered by a helper-dependent adenoviral vector. *J. Gene Med.,* **7,** 1086-1096.

99. Smilack,J.D. (1999) The tetracyclines. *Mayo Clin. Proc.,* **74,** 727-729.

100. Triezenberg,S.J., Kingsbury,R.C. and McKnight,S.L. (1988) Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev.,* **2,** 718-729.

101. Gilbert,D.M., Heery,D.M., Losson,R., Chambon,P. and Lemoine,Y. (1993) Estradiol-inducible squelching and cell growth arrest by a chimeric VP16-estrogen receptor expressed in saccharomyces cerevisiae: Suppression by an allele of PDR1. *Mol. Cell. Biol.,* **13,** 462-472.

102. Schiestl,R.H. and Gietz,R.D. (1989) High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr. Genet.,* **16,** 339-346.

Chapter 3: Mammalian Screening System for Gene Switch Engineering

3.1. Introduction

The power of directed evolution as a protein engineering tool lies in its ability to link a large library of genotypes to individual phenotypes that can be screened or selected for. While there are a number of *in vitro* systems that can manage this task such as SELEX (1) or emulsion encapsulation (2), the simplest and most flexible method is to isolate each gene variant from the others within its own living host cell. One of the fundamental enabling technologies of the molecular biology revolution was the ability to ligate a gene into a small piece of DNA called a plasmid, and efficiently insert this into a bacterium and allow the cellular machinery to transcribe the gene into RNA and translate the RNA into a protein (3, 4, 5). Using these techniques and a screening system it is possible to isolate DNA sequences or proteins of interest from a library of variants.

The characteristics of the bacteria *E. coli* and the yeast *S. cerevisiae* allow for straightforward library screening. Gene libraries can easily be constructed through restriction digestion and ligation of DNA into plasmids, followed by high efficiency transformation into bacteria by either heat shock or electroporation. While the transformation of yeast with ligation products is inefficient, libraries of up to 10^{10} plasmids can be constructed by introducing DNA fragments and taking advantage of the *in vivo* homologous recombination mechanisms active in yeast (6, 7). The fast growth of both bacteria and yeast speeds the library screening process, and the plasmids are stably maintained through the use of selection markers, allowing for the isolation of individual clones. While multiple plasmids can be transformed into either *E. coli* or *S. cerevisiae* (8, 9), these can be reduced by lowering the amount of DNA transformed or the effects of the individual plasmids can be separated by plasmid isolation from promising clones followed by retransformation.

However there are many applications where it is necessary to screen a library in mammalian cells. For screens affecting mammalian specific genes or signaling pathways, a mammalian cell based assay is usually required. More generally, many proteins require post-translational

modifications for correct function and these may vary between species. For example there are mammalian specific patterns of glycosylation that are essential for immune system function (10).

Screening in mammalian cells presents multiple challenges. Standard methods of introducing DNA into mammalian cells in culture include calcium phosphate precipitation, electroporation, or complexing the DNA with a cationic lipid (lipofection) (11). The amount of DNA used for these processes is much greater than that used for bacterial and yeast transformation, and results in large numbers of plasmids introduced into each mammalian cell, up to 10^5 copies per cell under standard lipofection conditions for example (12). The large number of plasmids is in part required because once in the cytoplasm only a small fraction will actually reach the nucleus in order to be transcribed. This is not appropriate for library screening as the effect from any individual plasmid will be masked by the others also in the cell. Plasmids are also not usually stably maintained in mammalian cells, and will be steadily degraded after transfection. This, combined with the long cell cycle length of around 24 hours, means that individual clones cannot be isolated apart from rare events resulting from plasmid integration into the mammalian genome and requiring many weeks of selection.

One method that does allow for mammalian screening involves the use of retroviral libraries. Retroviruses are a class of virus whose genome is a single stranded RNA molecule, that upon infection of a eukaryotic host is reverse transcribed into a DNA molecule that then integrates into the genome (13). Knowledge of the retroviral lifecycle and genome structure allowed the creation of engineered vectors that allow for the integration of a desired gene into the genome. By replacing the viral genes with a desired gene and using an engineered viral packaging cell line, virus is produced that allows for the integration of a gene of interest but that does not have the ability to further replicate. This system allows for the multiplicity of infection to be controlled such that a single integration event per mammalian cell is achievable. Retroviral approaches have been used for many applications including screening cDNA expression libraries for genes causing a particular phenotype (14), genome-scale RNAi gene knockdown screens (15), enzyme engineering (16), and engineering of zinc finger artificial transcription factors (17).

Another approach for mammalian library screening is to use a technique called protoplast fusion, where whole cells are stimulated to fuse together to form a hybrid cell. Spontaneous cell hybridization has been observed in mammalian cell culture experiments (18) but many organisms have a strong wall around their cell membrane that would prevent such occurrences. Protoplasts are the result of removing the rigid cell wall from bacterial, fungal, or plant cells, leaving the cellular contents surrounded simply by the lipid bilayer membrane (19). Research showed the feasibility of intra- and inter-species protoplast fusion methods in plants (20), and the stimulatory effect of the addition of high concentrations of polyethylene glycol (PEG) (21). The method was demonstrated in bacteria (22, 23), and was developed as a method to introduce genetic material from *E. coli* into mammalian cells in a clonal manner (24). In addition to its use in strain development and basic research, protoplast fusion has been used to clone genes from cDNA libraries (25, 26), and to screen libraries for active variants (27, 28).

The established method for engineering ligand binding domain controlled gene switches used in the Zhao laboratory was a yeast two-hybrid strategy (29). This relies on the transcription of a GAL4 response element controlled histidine biosynthesis gene, allowing growth of the yeast in media lacking histidine. Transcription is promoted by the recruitment of a GAL4DBD-ERαLBD construct associated in a ligand dependent manner with a steroid receptor coactivator 1 / GAL4AD fusion construct. The level of reporter gene production, and hence growth, is determined by how well the ligand of interest binds to the particular LBD mutant present in the cell. A saturation mutagenesis approach was systematically used to identify variants at amino acid positions known to affect ligand binding.

As discussed in Chapter 2, the ligand/LBD pairs L9/L7E and DHB/7S are not ideal for use as a gene switch in mammalian systems. The ligands are not approved for use in humans, preventing any downstream gene therapy applications, and are not easily available commercially. Also, the final engineered ligand binding domains showed poor performance with alternate DNA binding domains and activation domains. Partly to address some of these issues, Victor Gonzalez from the Zhao and Katzenellenbogen laboratories had used the yeast two-hybrid screening strategy (29) to engineer the ERα LBD for increased specificity towards hexylresorcinol. The mutations found were H524P, M421G, M343C, L391V, G521T, F337L, and L479R with the most sensitive

variant being the fourth round mutant with an EC_{50} of 36.6 nM to hexylresorcinol and 800 nM to estradiol, with the sixth round mutant showing an EC_{50} of 91.1 nM to hexylresorcinol and no response to estradiol up to 10^{-5} M (30).

Hexylresorcinol was chosen as a target ligand for a number of reasons. The compound (Figure 3.1) is a doubly hydroxylated benzene with a 6-carbon alkyl chain attached. When present at levels of 10 μM it has a low level of agonistic activity toward the estrogen receptor α ligand binding domain when measured by a yeast two-hybrid assay (Figure 3.2), although estrogenic effects at lower concentrations have been reported (31). This initial agonist activity allows the yeast two-hybrid screening approach to be used to engineer increased sensitivity. The compound has been recognized for many decades as both non-toxic to humans and possessing antimicrobial activities (32). It shows anesthetic activity and is an ingredient in throat lozenges (33). By acting as an inhibitor of the enzyme polyphenol oxidase, it can be used to prevent browning of fruit and crustaceans in the food industry (34, 35). The low toxicity and low activity toward the estrogen receptor make this compound an attractive ligand for use as a gene switch activator.

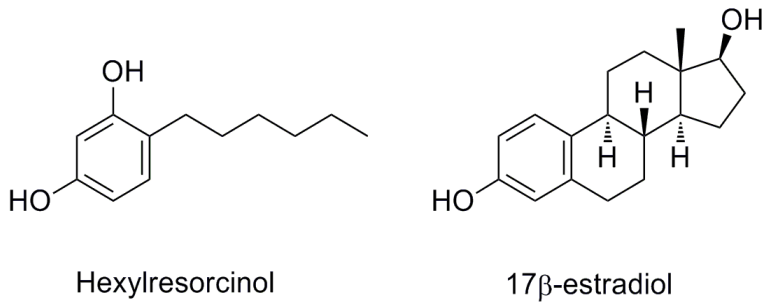


Figure 3.1. Hexylresorcinol and 17β-estradiol ligands.

A number of results had suggested that a mammalian screening method could have advantages over the yeast two-hybrid approach. The yeast system is based on the well characterized GAL4 DNA binding domain and response elements, whereas it is desirable to use the gene switch to target any given promoter. The need to change the DBD after engineering ligand specificity opens the door to unappreciated LBD-DBD interdomain contacts changing the performance of the gene switch. The mammalian intracellular milieu consists of many proteins not present in yeast that could interact in unknown ways with the gene switch. From a functional perspective, as seen in Chapter 1, the engineered ligand binding domains do not always perform as

anticipated upon transfer into mammalian cells. It was also observed that transfer of the hexylresorcinol engineered LBDs into mammalian gene switches gave a reduced sensitivity to ligand (Figure 3.3). If the engineering of the gene switch could take place in a context as close as possible to the desired use, unexpected results could be avoided. This would mean engineering in mammalian cells, and starting with the complete construct design including activation domain and DNA binding domain to target the gene of interest.

The goal of this chapter was therefore to investigate the utility of protoplast fusion as a means of screening in mammalian cells, and attempt to engineer a hexylresorcinol responsive gene switch targeting the *VEGF-A* gene (see Chapter 4). The previously engineered hexylresorcinol responsive ligand binding domains were shown to have a reduced sensitivity when cloned and expressed in mammalian cells. The protoplast method was shown to be capable of introducing plasmid into HeLa cells, and plasmid could be recovered back from them. A GAL response element controlled *EGFP* gene was created and integrated into HeLa cells and could be induced by gene switches containing GAL DNA binding domains. The *VEGF-A* promoter did not exhibit reliable induction of the *EGFP* gene, despite multiple constructs and positive feedback loops being tested. Although some initial results showed promise for the protoplast fusion screening using the GAL-responsive cell line, under library screening conditions the system was not able to recover a known active mutant. A fluorescent sensor construct was tested using protoplast fusion but this also was unable to retrieve the known mutant, leading to the determination that gene switch engineering using this approach was not feasible. Returning to the yeast two-hybrid screening system, three rounds of engineering were performed and yielded variants that showed increased sensitivity to hexylresorcinol, but again when these mutants were cloned into the mammalian gene switches a loss of sensitivity was observed.

3.2. Results and Discussion

3.2.1. Rationale for Developing a Mammalian Screening System

The utility of the hexylresorcinol/LBD pairs was evaluated for use as a gene switch. The sensitivity of the 4th, 5th, and 6th round mutants (named C2-4S, C2-5S, and C2-6E) was confirmed using the yeast growth assay (Figure 3.2). The LBDs were recloned into mammalian expression constructs in three forms: within the context of the full length estrogen receptor, within the gene switch context p65-Gal-LBD, and also directly recloned without modification. When analyzed for their ability to show ligand induced transcription from a luciferase reporter gene in HeLa cells, they revealed a loss of sensitivity (Figure 3.3). In all cases at least a ten-fold reduction in sensitivity to hexylresorcinol was observed, showing that these LBDs were not suitable for use as a gene switch. These findings spurred the consideration of alternative approaches to engineering gene switches for use in mammalian cells. (Book 4 pg 169-205)

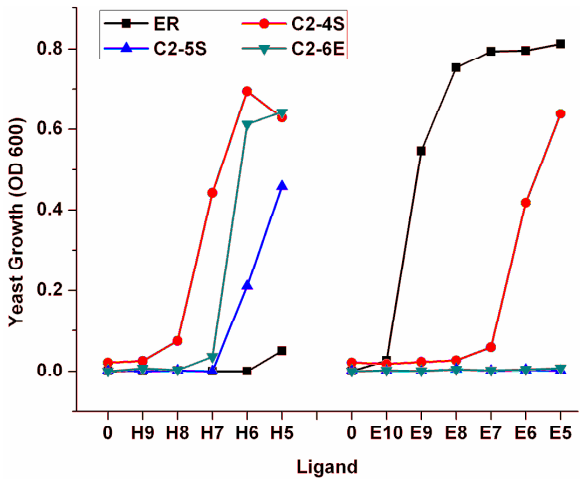


Figure 3.2. Response of engineered ligand binding domains to hexylresorcinol. Yeast strains containing the 4th, 5th, and 6th round engineered mutants were grown in the presence of hexylresorcinol (0 = no ligand, H9-H5 are 10^{-9} M to 10^{-5} M concentrations of hexylresorcinol) or estradiol (0 = no ligand, E10-E5 are 10^{-10} M to 10^{-5} M concentrations of estradiol). Growth was monitored by determining the optical density at 600 nm.

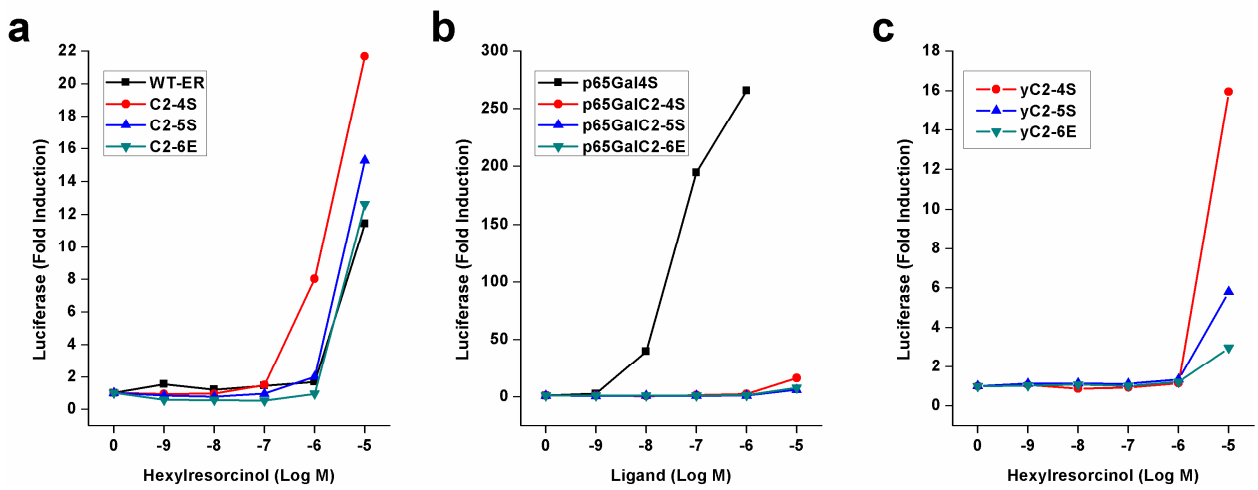


Figure 3.3. Loss of sensitivity of hexylresorcinol engineered LBDs in mammalian HeLa cells. (a) LBDs cloned into the full length ERα context activating a pS2-ERE-luciferase reporter in response to hexylresorcinol. (b) LBDs cloned with a p65 activation domain and Gal4 DNA binding domain activating a GalTATA-luciferase reporter. Shown for comparison is the 4th round DHB mutant (p65Gal4S) induced by the DHB ligand. (c) The yeast engineering constructs recloned without modification activating a GalTATA-luciferase reporter.

3.2.2. Protoplast Fusion as a Means of Introducing Plasmids Into Mammalian Cells

An engineering approach using mammalian cells was conceived (Figure 3.4). The desired gene switch would be cloned in *E. coli*, allowing for the creation of libraries of LBD variants. These would be grown in culture, converted to protoplasts, and fused to a HeLa cell line containing an integrated *EGFP* gene under the control of the promoter of interest. Expression of the gene switch would follow the transfer of the plasmid into the nucleus, and its activity would be stimulated by the addition of ligand. The gene switch would cause expression of the *EGFP* gene, allowing isolation of that HeLa cell by fluorescence activated cell sorting (FACS). The plasmid encoding the active gene switch would then be isolated from the HeLa cell, transformed back into *E. coli* for verification or for a further round of protoplast fusion. This approach would have an advantage over a retroviral based library since it would have fewer steps (no retrovirus production step after cloning the library) and the active clones would be directly recovered as plasmids from the HeLa cells, as opposed to the PCR recovery required to retrieve a chromosomally integrated construct. The method was investigated with the assistance of Jing Liang (Zhao Laboratory).

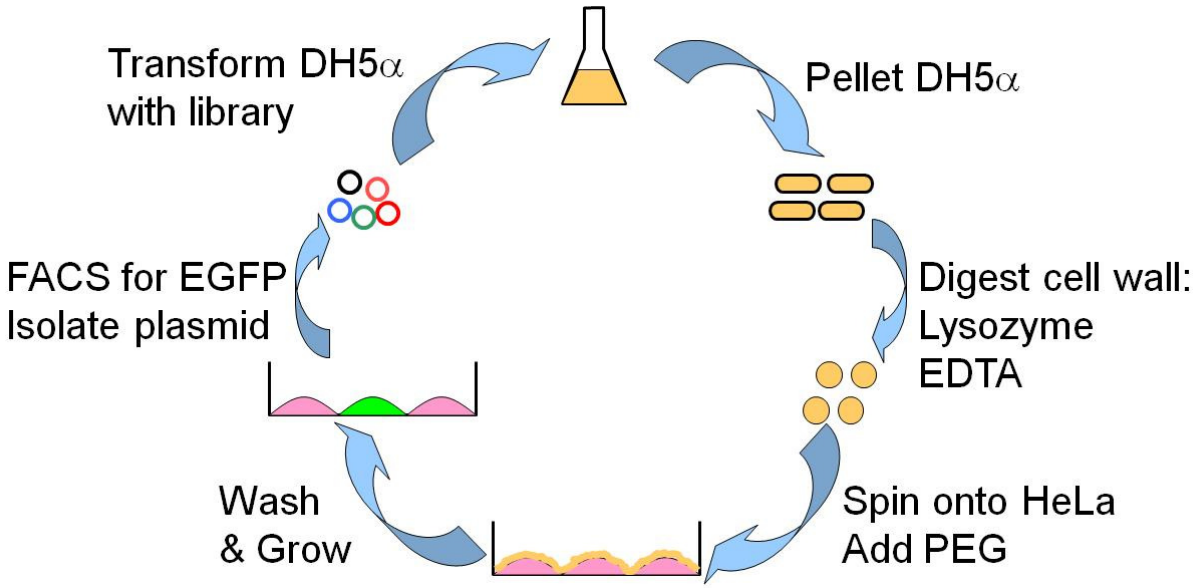


Figure 3.4. Overview of the proposed mammalian based protoplast fusion screening method for gene switch engineering.

As a first step toward developing protoplast fusion for our library screening system, it was necessary to show that it could introduce plasmid into the HeLa cells efficiently. The codon optimized fluorescent protein EGFP was cloned into the expression vector pCMV5, and shown by lipofection to be expressed in HeLa cells (Figure 3.5a). The percentage of fluorescent cells was dependent on the amount of plasmid introduced, and was typically between 50-80%. After some optimization, protoplast fusion was used to reliably introduce the EGFP plasmid into the HeLa cells (Figure 3.5b). Here the proportion of fluorescent cells was typically 30-40% of the population. One well of a 6-well plate contains approximately one million HeLa cells at confluent growth. If protoplast fusion is performed at 50% confluency, one 6-well plate should give around one million successfully transformed cells which is judged sufficient for library screening. (Book 5 pg 121, 175-183)

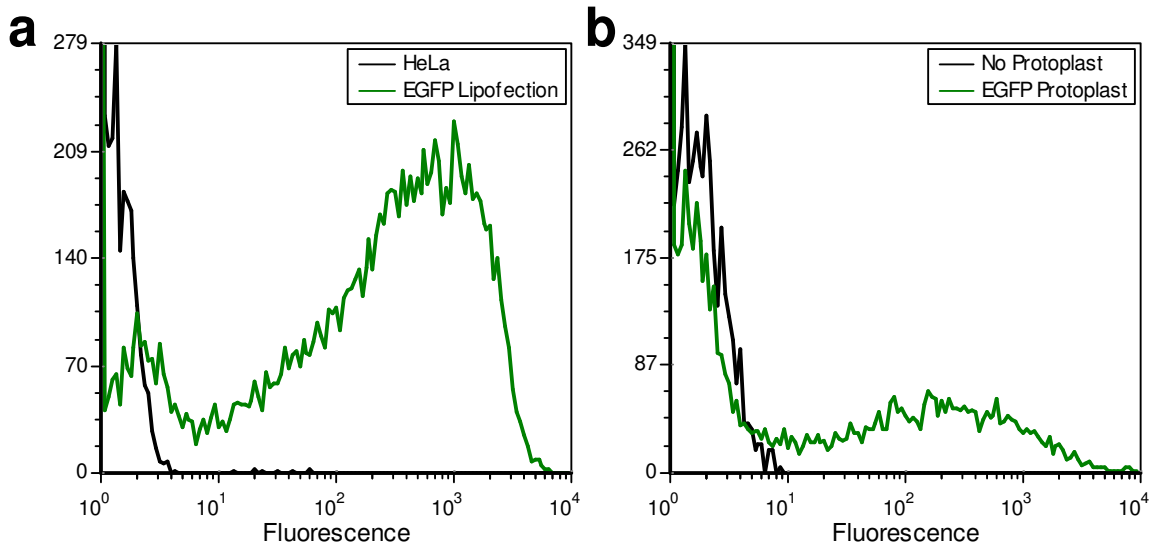


Figure 3.5. Expression of EGFP in HeLa cells monitored by flow cytometry. (a) EGFP plasmid introduced via lipofection. (b) EGFP plasmid introduced by protoplast fusion.

3.2.3. Construction of an EGFP Integrated HeLa Reporter Cell Line

Having shown that EGFP can be successfully introduced into HeLa and monitored by flow cytometry, the next requirement of the screening system was to construct a stable cell line that contained an integrated *EGFP* reporter gene. Two reporter genes were constructed in parallel. As a control, a *GalTATA*-EGFP construct was made to take advantage of the well performing and highly inducible synthetic promoter. As an endogenous target, we had chosen to regulate the Vascular Endothelial Growth Factor A gene so a *VEGF-A*-EGFP construct was also made. Testing of these reporter genes was performed by lipofection prior to the lengthy process of integration.

The *GalTATA*-EGFP reporter performed well. As shown in Figure 3.6a, the reporter exhibited only a slight basal expression that increased fluorescence above the level of the β-galactosidase non-fluorescent control. When the Gal targeting gene switch p65-GAL-4S was also present, there was no induction in the absence of DHB ligand, but around a ten-fold increase in the number of fluorescent cells when in the presence of DHB, and a 50-fold increase in the population fluorescence.

The *VEGF-A*-EGFP reporter did not show desirable induction characteristics (Figure 3.6b). A number of conditions were tested for their ability to activate this reporter. A constitutively active transcription factor, nlsVZ8p65, was constructed using an engineered zinc finger DNA binding domain called VZ8 (36) in conjunction with a nuclear localization sequence (nls) and the activation domain from the transcription factor p65. A ligand controlled gene switch was also created, p65(451-551)VZ8-4S, with a shorter region of the p65 activation domain and the addition of the 4th round DHB activated ERα LBD. As a non gene switch control, the addition of 100 μM cobalt chloride was used since this compound mimics the effect of hypoxia, leading to VEGF-A induction. The ligand-dependent gene switch had little effect on the expression of EGFP, and the constitutive switch caused only a slight increase. The cobalt chloride treatment gave the highest induction of EGFP, but this was only a 4 or 5 fold increase in either fluorescent cells or population fluorescence.

Since the *GalTATA* reporter performed well, it was stably inserted into the HeLa genome by retroviral integration. The *GalTATA*-EGFP segment was cloned into the retroviral vector pLHCX, lipofected into the packaging line GP2-293, and virus particles were collected and exposed to HeLa cells. Stable transformants were isolated by selection with hygromycin, and with the help of Jing Liang, were expanded and screened for reporter gene induction. A number of independent cell lines were isolated that showed low basal fluorescence and good induction upon the introduction of a constitutively acting artificial transcription factor GALp65 by either lipofection (Figure 3.7a) or protoplast fusion (Figure 3.7b). (Book 5 pg158-220)

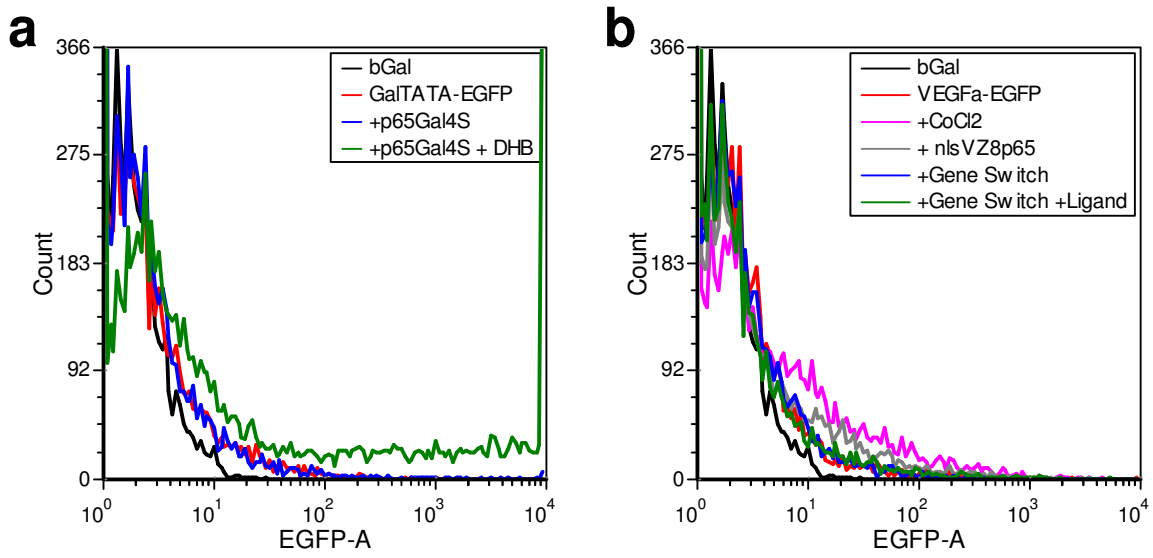


Figure 3.6. Activity of EGFP reporters introduced to HeLa cells by lipofection. (a) HeLa cells were transfected with the non-fluorescent β-galactosidase (bGal), the *GalTATA*-EGFP reporter plasmid, the reporter plasmid plus the DHB controlled gene switch p65Gal4S, or the reporter plus gene switch plus DHB ligand. (b) HeLa cells transfected with the β-galactosidase gene, the *VEGF-A*-EGFP reporter gene, the reporter gene plus cobalt chloride ($CoCl_2$), the reporter gene plus the gene switch p65(451-551)VZ8-4S, or the reporter gene plus the gene switch plus DHB.

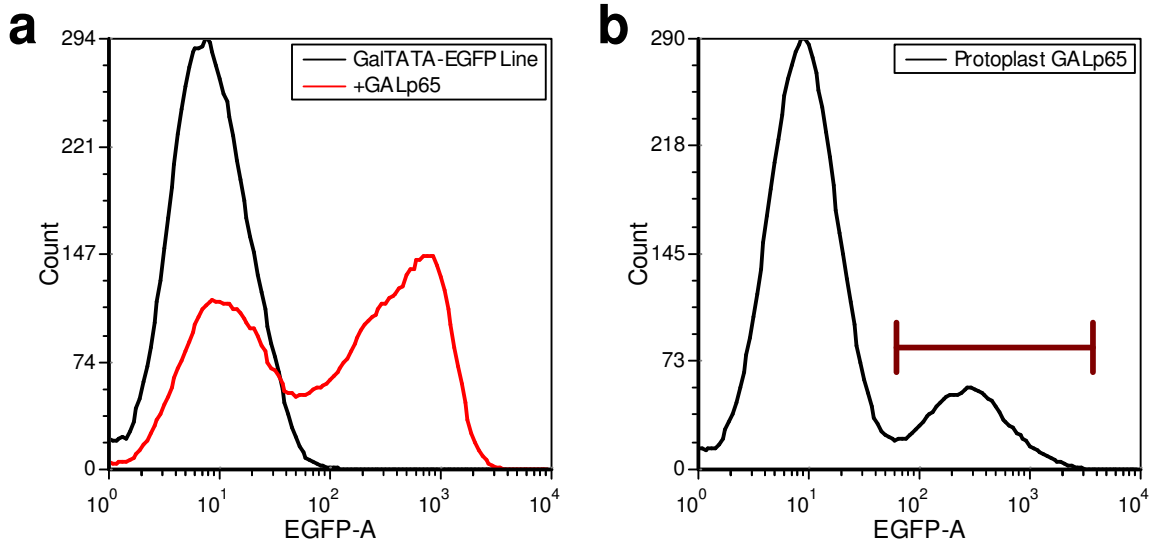


Figure 3.7. Induction of fluorescence in the stable *GalTATA*-EGFP HeLa cell line. (a) The *GalTATA*-EGFP cell line was lipofected with the β-galactosidase gene (*GalTATA*-EGFP Line) or with the constitutively acting GALp65 construct. (b) Induction of fluorescence within the *GalTATA*-EGFP cell line upon introduction of GALp65 by protoplast fusion.

3.2.4. VEGF-A Reporter Constructs

A number of different reporter constructs were investigated to find a functional VEGF-A reporter that could be used for mammalian screening. Due to the low induction of fluorescent signal from the *VEGF-A* promoter, several constructs were made to test the effect of signal amplification, as summarized in Table 3.1. The cascade design takes advantage of the well performing *GalTATA*-EGFP reporter by attempting to use a ligand controlled VEGF-A targeting gene switch to induce expression of the GALp65 construct from a VEGF-A promoter. *VEGF-A*-mVenus-nlsVZ8p65 forms a VEGF autoinduction loop by using a fluorescent mVenus tag attached to a constitutive *VEGF-A* targeting construct (nlsVZ8p65) under control of a *VEGF-A* promoter. *VEGF-A*-mVenus-GALp65 adds a fluorescently tagged transcription factor to the cascade scheme, *GalTATA*-mVenus-galp65 forms a fluorescent autoinduction loop via the *GalTATA* promoter. When expressed in HeLa cells and analyzed by flow cytometry, none of the combinations tested gave desirable results. The cascade system showed induction by cobalt chloride but not the gene switch (Figure 3.8). The *GalTATA*-mVenus-GALp65 gave a lower signal than the gtEGFP reporter and a decrease when ligand was added. Neither *VEGF-A*-mVenus-NLSvz8p65 nor *VEGF-A*-mVenus-GALp65 conditions gave a fluorescent signal, suggesting that either transcription or translation of the N-terminal fluorescent protein from the *VEGF-A* promoter is not effective.

To test transcription from the *VEGF-A* reporters, reverse transcription PCR was performed. RNA was isolated from HeLa cells lipofected with the gene switch p65GAL4S and the reporters *GalTATA*-pGL3 (luciferase), *VEGF-A*-pGL3 (luciferase), *GalTATA*-EGFP, and *VEGF-A*-EGFP. RT-PCR was performed to measure the mRNA levels for the luciferase and EGFP, and both cases showed good induction from the *GalTATA* promoter but very low levels without noticeable induction for the *VEGF-A* promoter (Figure 3.9). This suggests that the difficulties in developing a good reporter are partly due to a low level of transcription induced from the *VEGF-A* promoter.(Book 5 pg 250 – Book 6 pg 123)

Table 3.1. *VEGF-A* reporter amplification strategy plasmids

Cascade	VEGFa Autoinduction	Fluorescent GALp65	GalTATA Autoinduction
P65-4S-nlsVZ8-pCMV5 VEGFa-GALp65 GalTATA-EGFP	P65-4S-nlsVZ8-pCMV5 VEGFa- mVenus-nlsVZ8p65	P65-4S-nlsVZ8-pCMV5 VEGFa- mVenus-GALp65 GalTATA-EGFP	P65-4S-nlsVZ8-pCMV5 VEGFa-GALp65 GalTATA-mVenus -GALp65

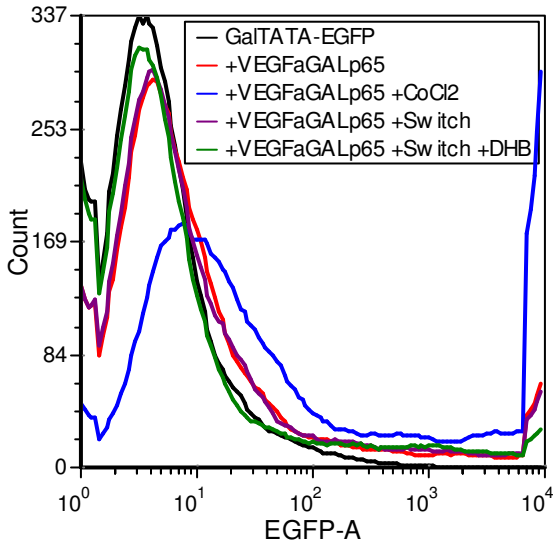


Figure 3.8. Cascade system for VEGFa reporter induction.

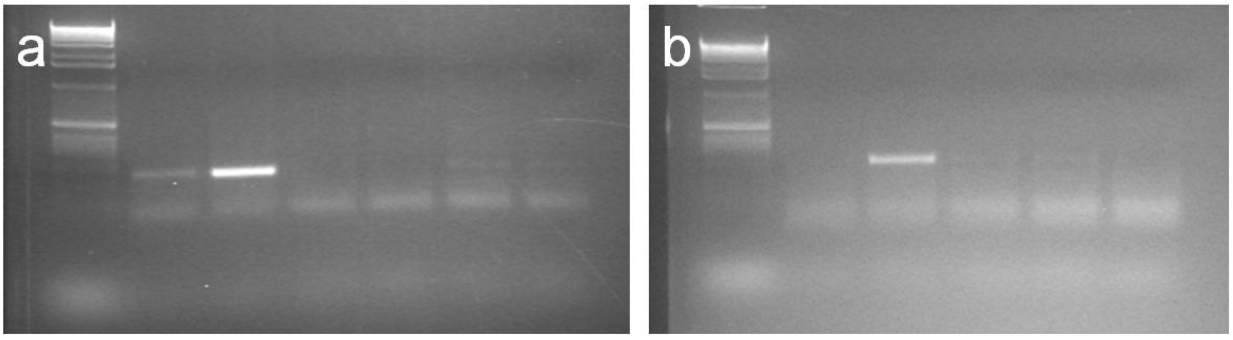


Figure 3.9. RT-PCR from reporter constructs. (a) EGFP mRNA, lane 1 is a DNA ladder, lanes 2 and 3 are uninduced and induced expression from *GalTATA*-EGFP, lanes 4 and 5 are negative controls, lanes 6 and 7 are uninduced and induced expression from *VEGF-A*-EGFP. (b) Luciferase mRNA, lane 1 is a DNA ladder, lanes 2 and 3 are uninduced and induced expression from *GalTATA*-pGL3, lanes 4 and 5 are uninduced and induced expression from *VEGF-A*-pGL3.

3.2.5. Library Screening Using Protoplast Fusion in the *GalTATA*-EGFP Cell Line

Due to the lack of a highly functional *VEGF-A* reporter construct, it was decided to attempt screening using protoplast fusion with the *GalTATA*-EGFP integrated cell line. This required the use of the GAL4 DNA binding domain gene switch. As an initial test of the screening process, I attempted to enrich the wild type LBD present in a 1:10 library ratio of p65GALER:p65GAL4S. As seen in Figure 3.10, the protoplast fusion was successful with the constitutively acting control plasmid GALp65, giving an obvious 20 % of the population with increased fluorescence. The ligand controlled gene switch gave a long low tail with about 5 % of cells having increased fluorescence upon 100 nM DHB induction. The p65GALER construct was responsible for a small increase in fluorescence upon the addition of 100 nM estradiol. The top 4000 fluorescent cells were collected and their plasmids were isolated and transformed into *E. coli*. Colonies were recovered at a level of 1 colony per 4-5 fluorescent HeLa cells collected by FACS. Plasmids were isolated from ten colonies and digested with a diagnostic restriction enzyme BsrGI. All clones produced a digestion pattern characteristic of either the ER or 4S construct, and revealed that the DHB exposed cells contained a ratio of one ER and nine 4S LBDs, which was the same as the initial library. Of the clones isolated from the estradiol exposed cells, two contained the ER LBD, indicating a slight enrichment from the initial library.

This experiment was repeated, and cells were collected by FACS into a highly fluorescent pool (500 cells), or a moderate fluorescent pool (1000 cells). When isolated clones were analyzed by restriction digestion, unsorted cells gave 0/10 ER LBDs, the moderate pool gave 2/10 ER, and the highly fluorescent cells gave 4/10 ER. These results indicated that successful enrichment was possible from a low complexity library.

To determine the effectiveness of the screening approach under more realistic conditions, I attempted to isolate the first round DHB mutant known from the yeast two-hybrid engineering (A350M, see Chapter 2). A saturation mutagenesis library was constructed at position 350 in the gene switch template p65GAL-LBD. This plasmid library was diluted with an excess of plasmid p65GAL-ER, containing the wild type ligand binding domain such that the A350M mutant was expected to be present at a ratio of 1:1000 of the entire library. Protoplasts were made, fused to the *GalTATA*-EGFP cells, and exposed to 50 nM DHB for two days. The cells underwent FACS

and the top 1 % of cells was collected, and plasmids were isolated from these 2000 cells. Of four clones retransformed into *E. coli*, one gave poor sequence, and the remaining three were wild type. A second round of enrichment using protoplast fusion was performed using the plasmid library isolated from the first round. Again, four clones were sequenced but only wild type sequence was observed. These results showed that under realistic library conditions, the protoplast fusion method did not effectively enrich functional gene switch clones. (Book 6 pg 157 – Book 7 pg 5, 102-126)

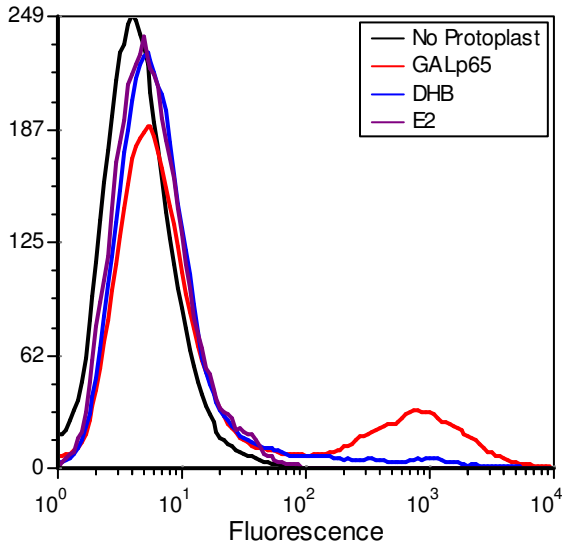


Figure 3.10. Initial trial of protoplast fusion based screening in *GalTATA*-EGFP cell line. Shown is the flow cytometry data of the unfused cell line, and protoplast fusion conditions with the constitutive GALp65, the p65GALER/p65GAL4S library induced by DHB, and the p65GALER/p65GAL4S library induced with estradiol (E2).

3.2.6. Library Screening Using Protoplast Fusion With a Fluorescent Sensor Construct

The fluorescent sensor construct discussed in Chapter 5 was investigated as a screening system for mammalian cells. These constructs consist of the ERα ligand binding domain (amino acids 312-595) flanked by a split fluorescent protein (mVenus), and show an increase in fluorescence upon the binding of a ligand. Although the construct is not a gene switch (it has no DNA binding domain to influence gene transcription), it was reasoned that it may provide an alternative means of engineering the ligand binding domain to respond to new ligands, by selecting for variants that show an increased fluorescence with the ligand of interest.

A number of mutant ligand binding domains were cloned into the sensor construct format, to analyze whether different fluorescent sensor LBDs could recapitulate the binding properties as seen in yeast or mammalian cells. As seen in Figure 3.11a, the 4S DHB responsive mutant shows induction of fluorescence in the presence of 10^{-7} M DHB, but no response to a similar level of estradiol. This replicates the dose response for this mutant in the yeast two-hybrid assay (see Figure 2.3b for comparison). Interestingly, when the 7th round DHB and L9 responsive mutants (7S and L7E) were cloned as fluorescent sensors, they show the high basal activity and low induction that were observed with their gene switch activity in mammalian cells (Figure 3.11b). The sensor constructs therefore convey useful information about the possible performance of the LBDs for gene switch purposes, and screening by protoplast fusion was performed. (Book 7 pg 85-126)

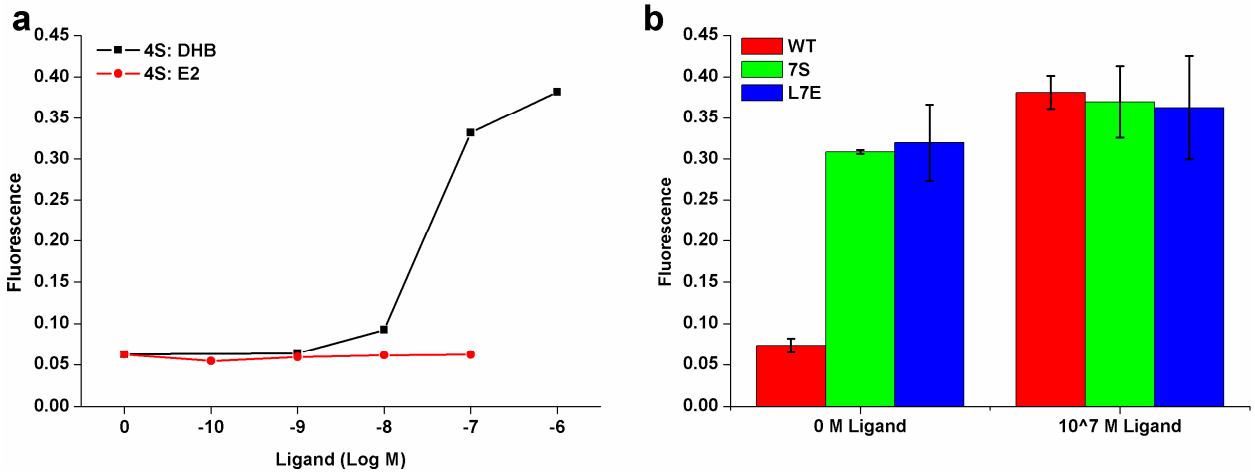


Figure 3.11. Performance of fluorescent sensors containing engineered ligand binding domains. (a) Sensor containing the 4S DHB mutant with a DHB and estradiol ligand titration. (b) Sensor constructs containing the wild type ERα LBD, or the 7th round DHB (7S) or L9 (L7E) mutants in the absence or presence of ligand.

Again it was decided to try to isolate the mutant A350M, known to show increased sensitivity to DHB. Using the wild type sensor construct as a scaffold, saturation mutagenesis libraries were made for the 19 sites targeted for our gene switch engineering studies (including position 350). Protoplasts were made for either an all position library or only position 350 library, and these were fused into HeLa cells, treated with DHB, and underwent FACS (Figure 3.12a). Plasmids isolated from the top 1 % of cells were recovered into DH5α cells and four colonies were

miniprepped and sequenced. None of the colonies were A350M, but 2 of the four colonies from the 350 library were A350L, as was one of the four colonies from the all 19 position library.

The A350L mutant was cloned into the p65Gal-LBD gene switch context and tested for its ability to activate a luciferase reporter gene in mammalian cells. As seen in Figure 3.13, the A350L mutant did show an increase in DHB sensitivity. So the protoplast fusion method successfully pulled out a functional mutant. However, the sensitivity of the A350M mutant was higher so this would have been expected to be isolated in the screen. There are more leucines coded for by the NNS saturation mutagenesis codon (CUC, CUG, and UUG) than the single AUG methionine, so library bias may have contributed to the failure to retrieve the expected mutant. Also, only a small number of clones were sequenced in this initial trial.

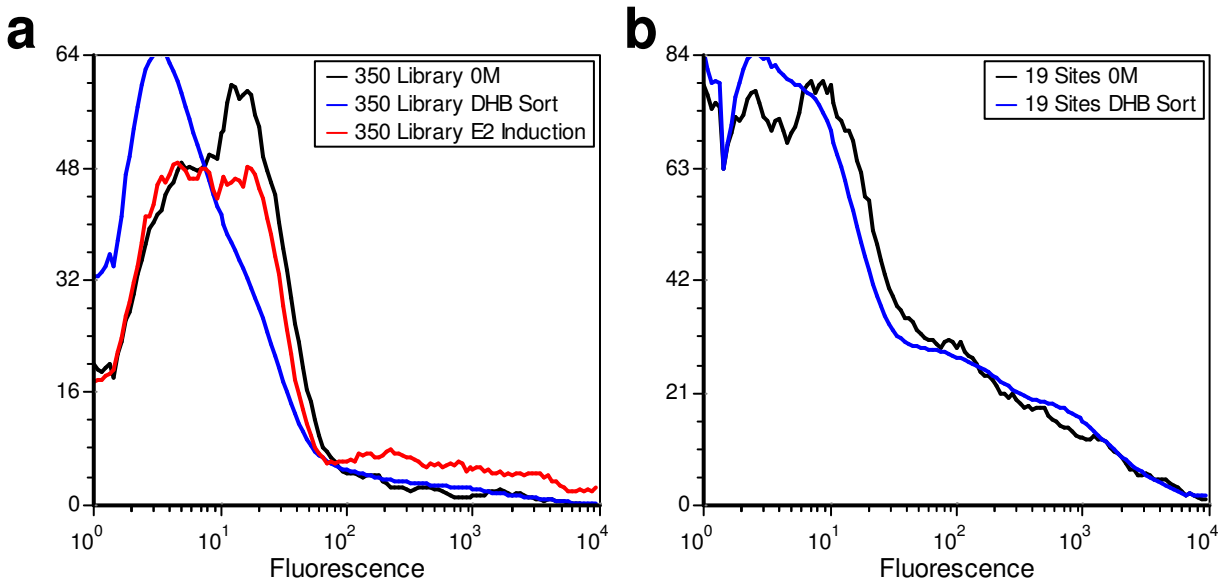


Figure 3.12. FACS on the sensor protoplast fusion libraries. (a) The first protoplast fusion screening. For clarity, only shown is the data for the position 350 library. The three curves show the uninduced fused HeLa cells, induction with 100 nM estradiol, and induction with 50 nM DHB. (b) The second sensor construct protoplast fusion using a library of all 19 sites with uninduced, and induction with 10 nM DHB. In both trials the top 1% of cells were collected for plasmid isolation.

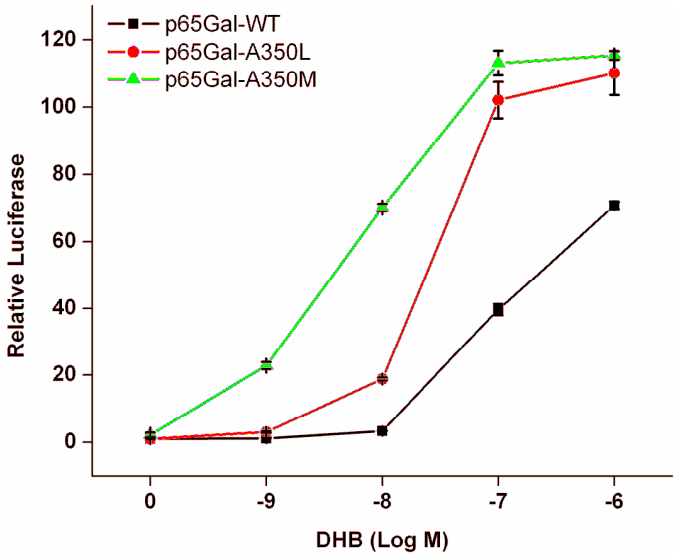


Figure 3.13. Luciferase reporter gene activation by the A350 mutants cloned into the mammalian gene switch context p65Gal-LBD. Data is normalized such that the no ligand WT value equaled 1, and is the average and standard deviation of duplicate samples.

Encouraged by the isolation of a mutant more sensitive to DHB than the wild type LBD, another protoplast fusion attempt was performed using the fluorescent sensor constructs (Figure 3.12b). A high level of basal fluorescence was observed and a lack of induction with DHB ligand. Although the top 1% of cells were collected, none of the 10 plasmids sequenced from these were either the A350M or A350L mutants previously identified.

Although some promising results were obtained, the protoplast fusion screening system was determined to be unreliable for gene switch engineering. Under library screening conditions, both the gene switch approach and the fluorescent sensor construct approach failed to retrieve the known A350M mutant responsive to DHB. For the gene switch approach, a number of steps are necessary for induction. First, the plasmid must be introduced to the reporter line, enough copies need to reach the nucleus to allow transcription of the gene switch for subsequent translation. The gene switch then requires activation by ligand and induction of the reporter EGFP gene to allow isolation of the HeLa cell by FACS. At this point, enough of the plasmid must still remain in the HeLa cell for isolation. At each of these steps, there is a loss of efficiency that reduces the performance of the system. The sensor construct allows for a direct readout of fluorescence, saving one transcription and translation step, however, significant basal signal was observed and a distinct induced population was not evident. Protoplast fusion has been used previously to

successfully screen libraries of arginine rich RNA binding proteins (27), and for internal ribosome entry sites (28). In both of these examples, the elements under selection were not ligand activated, as is the case of the gene switch. The requirement for ligand activation does decrease the number of induced cells, so this was an extra hurdle that the gene switch engineering faced with protoplast fusion. The amount of plasmid mutation after recovery from the HeLa cells was higher than expected based on previous reports, suggesting only around 1 % of plasmids recovered from mammalian cells were defective (37). All these factors seemed to push the method beyond its limits, resulting in poor performance.

3.2.7. Yeast Two-Hybrid Screening for the Hexylresorcinol Ligand

After it was determined that the protoplast fusion method was not feasible for gene switch engineering, the yeast two-hybrid method was resumed in another attempt to generate a hexylresorcinol responsive ERα LBD. Three rounds of stepwise site saturation mutagenesis were applied to 19 sites of the wild type hERα-LBD (WT): 343, 346, 347, 349, 350, 383, 384, 387, 388, 391, 404, 421, 424, 425, 428, 521, 524, 525, and 528. The results from yeast are shown in Figure 3.14 and Table 3.2. Round 1 produced the mutant G524T, round 2 yielded A350M, and the third round gave L384M. Each round produced an increase in the sensitivity to hexylresorcinol, a 94-fold improvement at round 3 over the wild type starting point, although the third round mutant achieved this sensitivity at a cost of increased ligand independent basal activity. Sensitivity to estradiol decreased at the first and second rounds, but then increased at the third round.

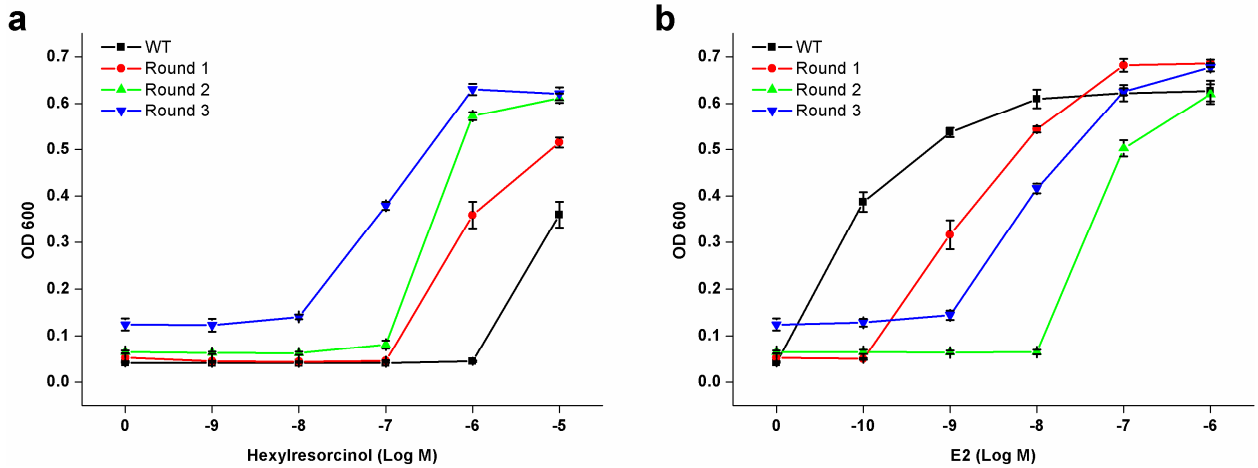


Figure 3.14. Growth of yeast strains carrying the engineered LBDs in response to (a) hexylresorcinol, and (b) estradiol (E2).

Table 3.2. Sensitivity of the mutants to hexylresorcinol and estradiol. The values are the mean and standard error of the mean expressed in nM based on the yeast growth curves.

Clone		Hexylresorcinol EC50	E2 EC50
WT		9300 (780)	0.075 (0.0063)
Round 1	(G524T)	1100 (240)	1.7 (0.25)
Round 2	(G524T/A350M)	410 (37)	78 (7.9)
Round 3	(G524T/A350M/L384M)	99 (5.7)	9.1 (0.32)

At this point, the engineered ligand binding domains were cloned into the gene switch context p65-GAL-LBD for testing in HeLa cells. Luciferase assays were performed to measure the induction of a *GalTATA*-Luciferase reporter gene and the results are shown in Figure 3.15. The second round mutant was seen to be the most sensitive to hexylresorcinol. However, a decrease in sensitivity as compared to the yeast growth curve was observed, as had been the case with the original variants engineered by Victor Gonzalez. It was not until micromolar concentrations of hexylresorcinol were present that any induction was observed, with only a four-fold increase observed for the best mutant. At 10 µM, the round 2 mutant gave over twenty-fold induction. However, at this high concentration of ligand, even the wild type LBD responded with a nine-fold induction (Figure 3.15a). Consistent with a lowered sensitivity in the HeLa cells, the response to estradiol was also reduced (Figure 3.15b) (Book 7 pg 130-158, Book 8 pg 1-56)

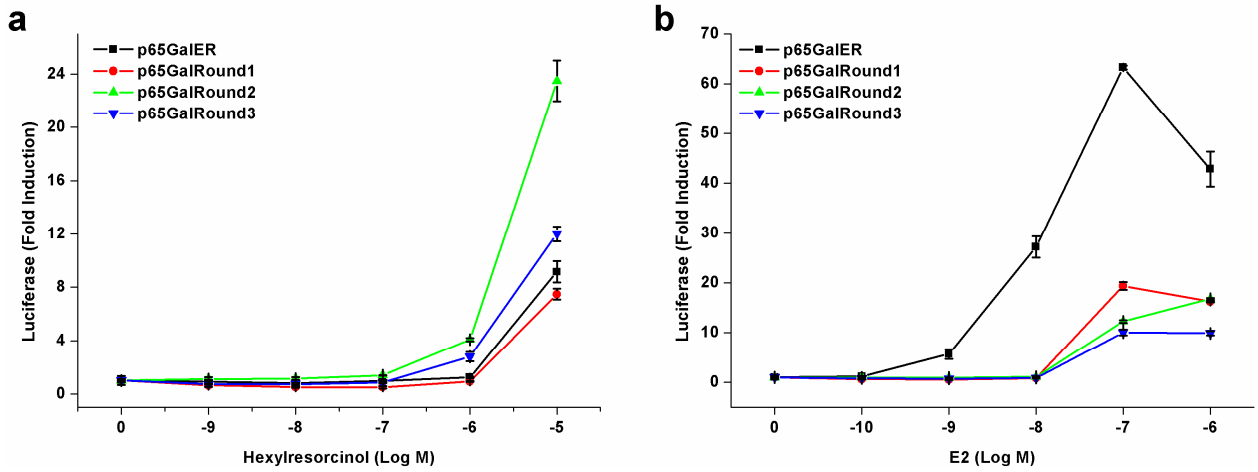


Figure 3.15. Mammalian luciferase using the engineered hexylresorcinol LBDs cloned with p65 activation domain and the Gal4 DNA binding domain. Shown is the response to (a) hexylresorcinol, and (b) estradiol.

3.3. Conclusions and Outlook

Despite individual steps of the protoplast fusion method were performing adequately, the overall method was not feasible for engineering gene switches in mammalian cells. The low level of induction was seen to be a problem, both in comparing the highly inducible synthetic *GalTATA* promoter to the endogenous *VEGF-A* promoter and also when comparing the activity of a constitutive switch to one containing a ligand binding domain. While not appropriate for the current engineering goals, protoplast fusion could conceivably be used for directed evolution of other targets that require expression in mammalian cells.

The loss of sensitivity that was observed with the newly engineered hexylresorcinol responsive LBDs upon their transfer into mammalian cells highlights that a functional mammalian screening method is still desirable. A retroviral screening system could be investigated. While this would involve a few extra steps for the creation of the library, once integrated into mammalian cells, the gene switch would be stable for as long a screening time as is needed. This could have the advantage of allowing an initial FACS screen to remove constitutively active mutants, followed by a second screen for ligand inducible mutants.

3.4. Materials and Methods

3.4.1. Reagents

DNA polymerase, restriction enzymes, and T4 DNA ligase were from New England Biolabs (Beverly, MA). Chemicals were from Sigma-Aldrich (Carlsbad, CA). Media components were from BD Biosciences (San Jose, CA), and premixed synthetic complete amino acid dropout mixes were from MP Biomedicals (Irvine, CA) or Sigma-Aldrich (Carlsbad, CA). Cell culture media were from the UIUC cell media facility (Urbana, IL). Lipofectamine 2000 was from Invitrogen (Carlsbad, CA). The DHB ligand was synthesized by Victor Gonzalez (Urbana, IL). DNA isolation and miniprep kits were from Qiagen (Valencia, CA), and yeast miniprep kits were from Zymo Research (Orange, CA).

3.4.2. Cloning of Constructs

Standard molecular biology techniques were used to create the constructs used in these experiments (Table 3.3). The *VEGF-A* promoter was amplified by PCR from human genomic

DNA and cloned into the luciferase vector pGL3 (Promega, Madison WI) using the KpnI and NcoI sites. Gene switches were cloned into the mammalian expression vector pCMV5.

Table 3.3. Constructs used in this work

HZ	Genotype	Plasmid	Details
3243	DH5alpha	C2-4S Full ER pCMV5	
3244	DH5alpha	C2-5S Full ER pCMV5	
3245	DH5alpha	C2-6E Full ER pCMV5	
3246	DH5alpha	p65GALC2-4S pCMV5	
3247	DH5alpha	p65GALC2-5S pCMV5	
3248	DH5alpha	p65GALC2-6E pCMV5	
3249	DH5alpha	GALC2-4S pCMV5	
3250	DH5alpha	GALC2-5S pCMV5	
3251	DH5alpha	GALC2-6E pCMV5	
1766	DH5alpha	EGFP pCMV5	
2017	DH5alpha	GalTATA-EGFP pGL3	
2020	DH5alpha	VEGFa-EGFP pGL3	
3252	DH5alpha	p65GAL4S pCMV5	
2021	DH5alpha	nlsVZ8p65 pCMV5	
1875	DH5alpha	p65(451-551)VZ8-4S pCMV5	
3253	DH5alpha	GALp65 pCMV5	
613	DH5alpha	GalTATA pGL3	
3254	DH5alpha	P65-4S-nlsVZ8-pCMV5	
2024	DH5alpha	VEGFa-GALp65	
3255	DH5alpha	VEGFa- mVenus-nlsVZ8p65	
3256	DH5alpha	VEGFa- mVenus-GALp65	
3257	DH5alpha	GalTATA-mVenus -GALp65	
642	DH5alpha	p65GALER pCMV5	
2243	DH5alpha	Vn-ER312-595-Vc pCMV5	
3240	DH5alpha	Vn-4S-Vc pCMV5	
3241	DH5alpha	Vn-7S-Vc pCMV5	
3242	DH5alpha	Vn-L7E-Vc pCMV5	
3258	DH5alpha	p65GAL-A350L pCMV5	
3259	DH5alpha	p65GAL-A350M pCMV5	
3260	DH5alpha	p65GAL-Round1 pCMV5	G524T
3261	DH5alpha	p65GAL-Round2 pCMV5	G524T/A350M
3262	DH5alpha	p65GAL-Round3 pCMV5	G524T/A350M/L384M
886	YRG2-SRC	ER pBD Gal4-CAM	
3263	YRG2-SRC	Round1 pBD Gal4-CAM	G524T
3264	YRG2-SRC	Round2 pBD Gal4-CAM	G524T/A350M
3265	YRG2-SRC	Round3 pBD Gal4-CAM	G524T/A350M/L384M

3.4.3. Cell Culture

HeLa cells were grown in minimal essential medium (MEM) plus 1 mM sodium pyruvate, and 10% fetal bovine serum (FBS) (UIUC Cell Media Facility, Urbana, IL) at 37 °C with 5% carbon dioxide. When cells were 80% confluent, they were trypsinized and split into 24-well plates with MEM media plus 1 mM sodium pyruvate, and 5% charcoal dextran stripped calf serum (UIUC Cell Media Facility). Cells were grown for 24 hours until they were over 90% confluent and transfected using 1.5 μL lipofectamine 2000 (Invitrogen, Carlsbad CA), 100 ng β-galactosidase gene expression plasmid, 690 ng luciferase reporter gene expression plasmid, and 10 ng of the relevant gene switch plasmids per well. After four hours, the media was changed to the MEM media plus 1 mM sodium pyruvate, and 5% charcoal dextran stripped calf serum, plus estradiol or hexylresorcinol ligand. Cells were incubated for a further 24 hours, then lysed and assayed for the luciferase activity using the Luciferase Assay System (Promega, Madison WI). Luciferase levels were normalized to β-galactosidase expression.

3.4.4. Stable Cell Line Creation

The procedures outlined in the "Retroviral gene expression and user manual", protocol PT3132-1 (Clontech, Mountain View, CA) were followed. The *GalTATA*-EGFP and *VEGF-A*-EGFP constructs were cloned into the retroviral vector pLHCX (Clontech, Mountain View, CA). 4 μg of the plasmid was lipofected using lipofectamine 2000 with 4 μg of pVSVG plasmid into the retroviral packaging line GP2-293 (grown in DMEM/high glucose/2 mM L-glutamine/1 mM sodium pyruvate/10 % FBS) in a 60 mm plate. After five hours, the media was replaced and the cells were incubated for 48-72 hours and then the media containing retrovirus was collected, filtered, and stored at -80 °C. Different amounts of retrovirus, from 20 μL to 1500 μL, were added to HeLa cells in the presence of 8 μg/mL polybrene and incubated for 24 hours. The media was changed and cells incubated for 24 hours before changing the media and including 200 μg/mL hygromycin. Cells were grown under antibiotic selection for two weeks, after which colonies were isolated with cloning rings, trypsinized, and transferred to individual wells of a 24 well plate. Cell lines were expanded under selection into wells of a 12 well plate, then a 6 well plate, then into a T25 flask. At this point, cells were assayed for reporter gene activity and well performing cell lines were expanded for use and storage in liquid nitrogen.

3.4.5. Protoplast Fusion

Protoplast fusion procedures were adapted from published methods (27, 38). Generally, 50 ng of expression plasmid was electroporated into DH5α *E. coli* cells. The cells were incubated at 37 °C in Luria Broth (LB) media for 1 h and then transferred into a flask containing 50 mL LB and 100 μg/mL ampicillin. The flask was incubated at 37 °C with shaking for 4-6 h until the optical density at 600 nm reached 0.8 at which point chloramphenicol was added to a concentration of 200 μg/mL. After a 12-16 h incubation to amplify the plasmid copy number, the cells were pelleted by centrifugation and resuspended in 3 mL of ice-cold 20 % sucrose / 50 mM Tris-HCl pH8. 600 μL of 10 mg/mL lysozyme in 0.25 M Tris-HCl pH 8 was added to the cell solution and incubated on ice for 5 min. 1200 μL of 0.25 M EDTA pH 8 was added, followed by a 5 min incubation on ice. 1200 μL of 50 mM Tris-HCl pH 8 was added and incubated for at 37 °C for 10 min to generate the protoplasts. At this point, 1 μL of the solution was placed on a microscope slide, covered with a cover slip, and observed by phase contrast light microscopy. A good preparation had at least 80 % individual spherical protoplasts visible. If the population showed a high proportion of oblong bacteria (indicating intact cell walls), the solution was incubated for another 5-10 min at 37 °C until most were converted into protoplasts. The protoplast solution was resuspended by adding 20 mL of ice-cold 10 % sucrose / 10 mM $MgCl_2$ in MEM media, slowly and with gentle mixing to prevent osmotic shock and lysing of the protoplasts.

Media was removed from 6-well plates containing HeLa cells at approximately 80 % confluency. Between 1-3 mL of protoplast solution was added to each well and the plate was centrifuged at 1000 *g* for 10 min to lay the protoplasts onto the HeLa cells. The supernatant was carefully removed by aspiration, and 1 mL of 50 % polyethylene glycol 1000 (average molecular weight 1000) was added to the side of the well. Plates were incubated for 3 min at room temperature and the PEG was removed by aspiration. Two gentle washes were performed to remove the PEG but leave the protoplast layer intact, using 2 mL of MEM without serum added to each well followed by aspiration. 2 mL of MEM / 5% Charcoal Dextran treated Calf Serum / Penicillin / Streptomycin / Kanamycin was added per well and plates were placed at 37 °C in a humidified incubator with 5 % CO_2 for 4-6 h (during which most of the protoplast layer would lift off the

HeLa cells), then the media was replaced and ligand was added. After 24-48 h incubation, the media was removed, cells were washed with 2 mL PBS and collected by trypsinization.

3.4.6. Flow Cytometry and FACS

HeLa cells were exposed to a solution of trypsin / EDTA for 30-60 s until they began to exhibit a rounded morphology. The trypsin / EDTA was removed by aspiration and the cells were incubated for another 1-2 min and resuspended in MEM + 5% Charcoal Dextran stripped Calf Serum. Samples were pelleted by centrifugation for 5 min at 800 *g*, media was removed by aspiration, and cells were resuspended in 300-400 µL PBS / 5 mM EDTA. For monitoring purposes, cells were measured on a BD LSRII flow cytometer (Becton Dickinson, Franklin Lakes, NJ) using standard GFP/FITC filter sets. Analysis was performed using FCS Express 3 software (De Novo Software, Los Angeles, CA). Events were gated on the region corresponding to single whole cells and the mean fluorescence for 10,000 cells was recorded. For isolation of positively expressing HeLa cells by fluorescence activated cell sorting (FACS), samples were analyzed on a FACSAria II flow cytometer (BD Biosciences, Franklin Lakes, NJ). Cells that expressed above the level of a negative control, or the top 1 % of cells, were collected into a tube containing MEM + 5% Charcoal Dextran stripped Calf Serum and 35000 non transformed HeLa cells to act as a carrier.

Plasmid-containing HeLa cells were pelleted and resuspended in 50 µL of 10 mM tris-HCl with 0.2 mg/mL tRNA. 50 µL of Qiagen miniprep buffer P2 was added to lyse the cells, followed by 70 µL of buffer P3. The solution was centrifuged at 13,000 rpm for 5 min and the supernatant was transferred to a new tube. An equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol was added and vortexed. After centrifuging for 1 min, the top aqueous layer was transferred to a new tube and an equal volume of chloroform was added, vortexed, centrifuged for 1 min, and the top layer removed to a new tube. 1 µL of 20 mg/mL glycogen was added, followed by one volume of isopropanol. The tube was mixed and incubated to -80 °C for 30 min and then centrifuged for 10 min. The supernatant was removed and the pellet washed with 200 µL 70 % ethanol. The tube was centrifuged for a further five min, the supernatant was removed, and the pellet was air dried. The isolated plasmid was resuspended in 10 µL 10 mM tris-HCl, 2 µL was electroporated into DH5α *E. coli* cells which were plated on LB + Ampicillin plates.

3.4.7. Yeast Two-hybrid Screening

Screening was performed as described previously (29). The first and second round of screening was performed at a concentration of 10^{-6} M hexylresorcinol, while the third round used 10^{-7} M. For the first round, saturation mutagenesis was applied individually to 19 sites of the wild type hERα-LBD: 343, 346, 347, 349, 350, 383, 384, 387, 388, 391, 404, 421, 424, 425, 428, 521, 524, 525, and 528. For the second round, the first mutant G525T was used as a template and the remaining 18 sites were mutagenized. For the third round, the mutant G524T/A350M was used as a template and the remaining 17 sites were mutagenized.

3.5. References

1. Tuerk,C. and Gold,L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science,* **249,** 505-510.

2. Tawfik,D.S. and Griffiths,A.D. (1998) Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.,* **16,** 652-656.

3. Jackson,D.A., Symons,R.H. and Berg,P. (1972) Biochemical method for inserting new genetic information into DNA of simian virus 40: Circular SV40 DNA molecules containing lambda phage genes and the galactose operon of escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.,* **69,** 2904-2909.

4. Lobban,P.E. and Kaiser,A.D. (1973) Enzymatic end-to end joining of DNA molecules. *J. Mol. Biol.,* **78,** 453-471.

5. Cohen,S.N., Chang,A.C., Boyer,H.W. and Helling,R.B. (1973) Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. U. S. A.,* **70,** 3240-3244.

6. Ma,H., Kunes,S., Schatz,P.J. and Botstein,D. (1987) Plasmid construction by homologous recombination in yeast. *Gene,* **58,** 201-216.

7. Benatuil,L., Perez,J.M., Belk,J. and Hsieh,C.M. (2010) An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Eng. Des. Sel.,* **23,** 155-159.

8. Goldsmith,M., Kiss,C., Bradbury,A.R. and Tawfik,D.S. (2007) Avoiding and controlling double transformation artifacts. *Protein Eng. Des. Sel.,* **20,** 315-318.

9. Scanlon,T.C., Gray,E.C. and Griswold,K.E. (2009) Quantifying and resolving multiple vector transformants in S. cerevisiae plasmid libraries. *BMC Biotechnol.,* **9,** 95.

10. Marth,J.D. and Grewal,P.K. (2008) Mammalian glycosylation in immunity. *Nat. Rev. Immunol.,* **8,** 874-887.

11. Kingston,R.E. (2003) Introduction of DNA into mammalian cells. *Curr. Protoc. Mol. Biol.,* **Chapter 9**.

12. Tseng,W.C., Haselton,F.R. and Giorgio,T.D. (1997) Transfection by cationic liposomes using simultaneous single cell measurements of plasmid delivery and transgene expression. *J. Biol. Chem.,* **272,** 25641-25647.

13. Buchschacher,G.L.,Jr. (2001) Introduction to retroviruses and retroviral vectors. *Somat. Cell Mol. Genet.,* **26,** 1-11.

14. Kitamura,T., Onishi,M., Kinoshita,S., Shibuya,A., Miyajima,A. and Nolan,G.P. (1995) Efficient screening of retroviral cDNA expression libraries. *Proc. Natl. Acad. Sci. U. S. A.,* **92,** 9146-9150.

15. Root,D.E., Hacohen,N., Hahn,W.C., Lander,E.S. and Sabatini,D.M. (2006) Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat. Methods,* **3,** 715-719.

16. Chen,K.C., Wu,C.H., Chang,C.Y., Lu,W.C., Tseng,Q., Prijovich,Z.M., Schechinger,W., Liaw,Y.C., Leu,Y.L. and Roffler,S.R. (2008) Directed evolution of a lysosomal enzyme with enhanced activity at neutral pH by mammalian cell-surface display. *Chem. Biol.,* **15,** 1277-1286.

17. Tschulena,U., Peterson,K.R., Gonzalez,B., Fedosyuk,H. and Barbas,C.F.,3rd. (2009) Positive selection of DNA-protein interactions in mammalian cells through phenotypic coupling with retrovirus production. *Nat. Struct. Mol. Biol.,* **16,** 1195-1199.

18. Sorieul,S. and Ephrussi,B. (1961) Karyological demonstration of hybridization of mammalian cells in vitro. *Nature,* **190,** 653-654.

19. Weibull,C. (1958) Bacterial protoplasts. *Annu. Rev. Microbiol.,* **12,** 1-26.

20. Power,J.B., Cummins,S.E. and Cocking,E.C. (1970) Fusion of isolated plant protoplasts. *Nature,* **225,** 1016-1018.

21. Kao,K.N. and Michayluk,M.R. (1974) A method for high-frequency intergeneric fusion of plant protoplasts. *Planta,* **115,** 355-367.

22. Fodor,K. and Alfoldi,L. (1976) Fusion of protoplasts of bacillus megaterium. *Proc. Natl. Acad. Sci. U. S. A.,* **73,** 2147-2150.

23. Schaeffer,P., Cami,B. and Hotchkiss,R.D. (1976) Fusion of bacterial protoplasts. *Proc. Natl. Acad. Sci. U. S. A.,* **73,** 2151-2155.

24. Schaffner,W. (1980) Direct transfer of cloned genes from bacteria to mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.,* **77,** 2163-2167.

25. Seed,B. and Aruffo,A. (1987) Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. *Proc. Natl. Acad. Sci. U. S. A.,* **84,** 3365-3369.

26. Caporale,L.H., Chartrain,N., Tocci,M. and DeHaven,P. (1990) Protoplast fusion in microtiter plates for expression cloning in mammalian cells: Demonstration of feasibility using membrane-bound alkaline phosphatase as a reporter enzyme. *Gene,* **87,** 285-289.

27. Tan,R. and Frankel,A.D. (1998) A novel glutamine-RNA interaction identified by screening libraries in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.,* **95,** 4247-4252.

28. Venkatesan,A. and Dasgupta,A. (2001) Novel fluorescence-based screen to identify small synthetic internal ribosome entry site elements. *Mol. Cell. Biol.,* **21,** 2826-2837.

29. Chockalingam,K., Chen,Z., Katzenellenbogen,J.A. and Zhao,H. (2005) Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. *Proc. Natl. Acad. Sci. U. S. A.,* **102,** 5691-6.

30. Gonzalez,V. (2008) Development of a novel ligand-receptor pair through protein engineering of the human estrogen receptor. MS Thesis. University of Illinois at Urbana-Champaign.

31. Amadasi,A., Mozzarelli,A., Meda,C., Maggi,A. and Cozzini,P. (2009) Identification of xenoestrogens in food additives by an integrated in silico and in vitro approach. *Chem. Res. Toxicol.,* **22,** 52-63.

32. Leonard,V. (1925) The significance of hexylresorcinol and its homologues in relation to the problem of internal antisepsis. *Science,* **62,** 408-412.

33. Buchholz,V., Leuwer,M., Ahrens,J., Foadi,N., Krampfl,K. and Haeseler,G. (2009) Topical antiseptics for the treatment of sore throat block voltage-gated neuronal sodium channels in a local anaesthetic-like manner. *Naunyn Schmiedebergs Arch. Pharmacol.,* **380,** 161-168.

34. Buta,J.G., Moline,H.E., Spaulding,D.W. and Wang,C.Y. (1999) Extending storage life of fresh-cut apples using natural products and their derivatives. *J. Agric. Food Chem.,* **47,** 1-6.

35. Frankos,V.H., Schmitt,D.F., Haws,L.C., McEvily,A.J., Iyengar,R., Miller,S.A., Munro,I.C., Clydesdale,F.M., Forbes,A.L. and Sauer,R.M. (1991) Generally recognized as safe (GRAS) evaluation of 4-hexylresorcinol for use as a processing aid for prevention of melanosis in shrimp. *Regul. Toxicol. Pharmacol.,* **14,** 202-212.

36. Liu,P.Q., Rebar,E.J., Zhang,L., Liu,Q., Jamieson,A.C., Liang,Y., Qi,H., Li,P.X., Chen,B., Mendel,M.C., et al. (2001) Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. activation of vascular endothelial growth factor A. *J. Biol. Chem.,* **276,** 11323-11334.

37. Razzaque,A., Mizusawa,H. and Seidman,M.M. (1983) Rearrangement and mutagenesis of a shuttle vector plasmid after passage in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.,* **80,** 3010-3014.

38. Hollenbaugh,D., Aruffo,A., Jones,B. and Linsley,P. (2003) Use of monoclonal antibodies for expression cloning. *Curr. Protoc. Mol. Biol.,* **Chapter 6,** Unit 6.11.

Chapter 4: Gene Switch Targeting of the Endogenous *VEGF-A* Gene

4.1 Introduction

The promise of a well designed gene switch system is that it can be used to regulate not just artificial reporter genes, but any desired gene at its endogenous genome locus. This allows for specific induction or repression of a desired gene for research purposes or for gene therapy applications. A key component to gene targeting is the creation of specific DNA binding domains to effectively target the gene switch to the gene of interest. Ideally the DNA binding domain would be specific to a single site in the genome to avoid off target effects. The majority of research in this area has used the Cys_2His_2-type zinc finger domain due to its perceived modularity, with one finger interacting with three base pairs of DNA and tandem arrays of fingers binding longer regions (1, 2). Different approaches have been taken to engineer a DBD to a specific site including the use of phage display to engineer specific individual fingers, a bacterial two-hybrid screen, rational design, or a combination of these (3, 4, 5, 6, 7).

The gene that was chosen for targeting in this thesis was Vascular Endothelial Growth Factor A (*VEGF-A*). VEGF-A plays a critical role in the development of blood vessels and the lymphatic system (8) and was identified by a number of groups as a factor that stimulates endothelial cell growth and increases permeability of the vascular system (9, 10, 11). By stimulating the formation of blood vessels during development, it plays an essential role, with loss of even one allele of this gene leading to embryonic lethality (12). In addition to its role in normal development, it is involved in wound healing, regulating hematopoietic stem cells, growth of the uterine lining, as well as in the development of tumors which require vascularization in order to grow. The ability to regulate VEGF-A levels has obvious clinical relevance with upregulation leading to better wound healing and treatment of ischemia, while downregulation could help potentially aid in the treatment of cancer and diabetic blindness (13, 14, 15).

The natural regulation of *VEGF-A* is a multifaceted process occurring at both transcriptional and post-transcriptional levels. The gene consists of 8 exons with multiple variants being generated through alternative splicing of the transcribed mRNA (16). The promoter region contains binding sites for transcription factors including the Sp family, AP1, the estrogen and progesterone

receptors, and the hypoxia inducible factor (HIF). The binding of HIF-1 to the promoter is a main contributor to *VEGF-A* transcription and promotes the growth of vasculature into areas of hypoxia (17). The main transcriptional start site is located 1038 bp upstream of the traditional AUG start codon, however there is an alternate downstream transcriptional start, two cap-independent internal ribosome entry sequences (IRES) (18, 19), and some isoforms generated through a CUG codon 499 bp downstream of the main transcriptional start (20). At the post-transcriptional level, hypoxia acts to stabilize the *VEGF-A* mRNA through elements that bind to instability elements in the 3' untranslated region (21, 22), while translation can be inhibited through the IFN-γ-activated inhibition of the translation complex (22).

Due to the clinical relevance of altering VEGF-A levels, a number of groups have engineered artificial transcription factors to target the gene. A series of papers from Sangamo Biosciences showed that artificial zinc finger domains designed to bind to accessible regions of chromatin in the promoter could efficiently induce VEGF-A secretion in HEK293 cells by addition of the activation domain VP16 or p65, could repress expression by addition of a histone methyltransferase domain, functioned in a mouse model to stimulate angiogenesis and speed wound healing, and could be controlled using a modified progesterone responsive ligand binding domain (23, 24, 25, 26). One of the DBDs from this research was subsequently used to create a rapamycin inducible dimerizing gene switch (27). DNA binding domains designed from endogenous zinc fingers were used by Kim's group to successfully induce or repress VEGF-A expression in HEK293 or in human tumor cells xenografted into mice (28, 29, 30). Sera's group has also used constitutively acting gene switches to induce VEGF-A in HEK293 cells (31, 32).

The goal of this chapter was to target the endogenous *VEGF-A* gene with an engineered ligand controlled gene switch, and identify design considerations that influence the performance. By using previously engineered DNA binding domains and the engineered DHB responsive ligand binding domain 4S, constructs were created that showed up to 17 fold induction of a *VEGF-A* promoter controlled luciferase reporter. Construct performance depended on the particular DBD used, the activation domain used, and the arrangement of the functional domains. When introduced into HeLa cells, the constructs resulted in successful ligand dependent VEGF-A induction of a few fold. A higher level of basal expression was observed when HEK293 cells

were assayed, and minimal induction was seen. Studies are continuing in order to identify the reasons for the discrepancy with previously published reports using HEK293.

4.2 Results and Discussion

4.2.1 Cloning of a Luciferase Reporter and Initial Gene Switch Constructs

To begin engineering a gene switch capable of influencing the endogenous *VEGF-A* gene, a luciferase reporter was cloned to enable the characterization of constructs. A 3 kb length of the *VEGF-A* promoter / 5' untranslated region was obtained by PCR from human genomic DNA and cloned into the pGL3 luciferase reporter vector. The cloned reporter has the first codon of the luciferase gene fused at the main translational start site of *VEGF-A*. Some initial gene switch constructs were made to test induction of the reporter. Liu et al. of Sangamo Biosciences had previously designed multiple zinc finger DNA binding domains targeted to *VEGF-A* (23), and one of these was used here. The VZ-8 (called herein VZ8) binds to the target GGGGAGGAT(c) which is 8 base pairs upstream from the main transcriptional start site of the *VEGF-A* gene. This domain was cloned with either the wild type ERα or engineered DHB responsive 4S ligand binding domain, and the presence or absence of p65 activation domain variants. As seen in Figure 4.1, all constructs showed ligand dependent induction of luciferase, indicating that both the reporter and the gene switches were functional. The addition of the full p65 activation domain (amino acids 282-551) increased the basal expression, while use of a reduced length of p65 (amino acids 451-551) showed low basal expression and a ten-fold induction for the construct p65(451-551)VZ8-4S. From these few constructs tested, it can be seen that the design of a gene switch can be optimized using factors such as the activation domain and the order of domains.

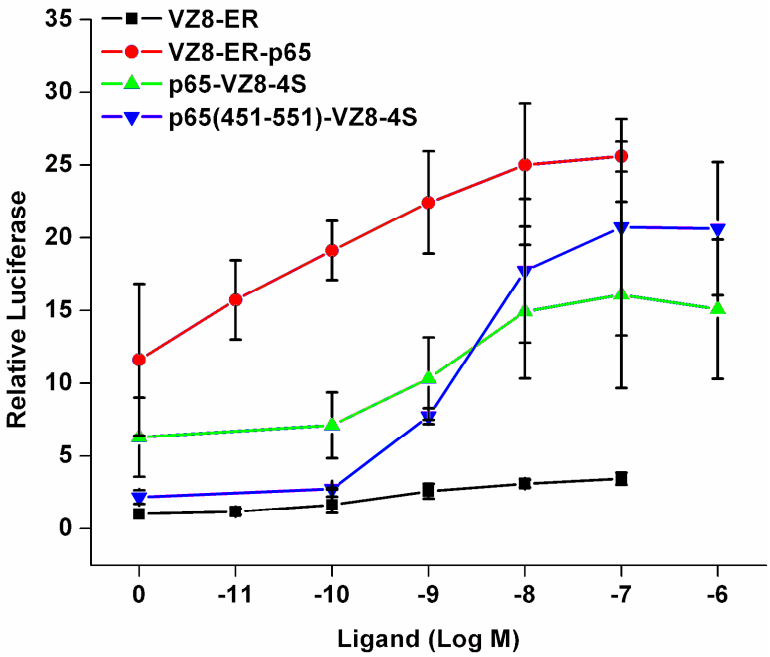


Figure 4.1. Induction of the luciferase reporter gene by *VEGF-A* targeting gene switches. Data is normalized such that the 0 M point for VZ8ER equaled 1. ER gene switches were exposed to differing levels of estradiol, while the 4S switches were exposed to DHB. 10 ng of gene switch, 690 ng of the luciferase reporter expression plasmid, and 100 ng of the β-galactosidase expression plasmid were lipofected into HeLa cells.

The level of induction from the *VEGF-A* targeting constructs was markedly reduced compared to the GAL4 DBD-containing switches of previous chapters. This could be due to many factors including expression and solubility of the gene switch, affinity of the switch for the promoter, the presence of only one binding site in the *VEGF-A* promoter compared to the multiple sites in the synthetic *GalTATA* promoter, or dimerization of the gene switches. The effect of gene dosage was examined with a constitutively acting construct, nlsVZ8p65, containing a nuclear localization sequence fused to the DNA binding and activation domains. It was seen that a higher amount of gene switch plasmid transfection resulted in a higher induction of the luciferase reporter (Figure 4.2). Subsequent experiments typically used 100 ng of gene switch expression plasmid per well of a 24 well plate.

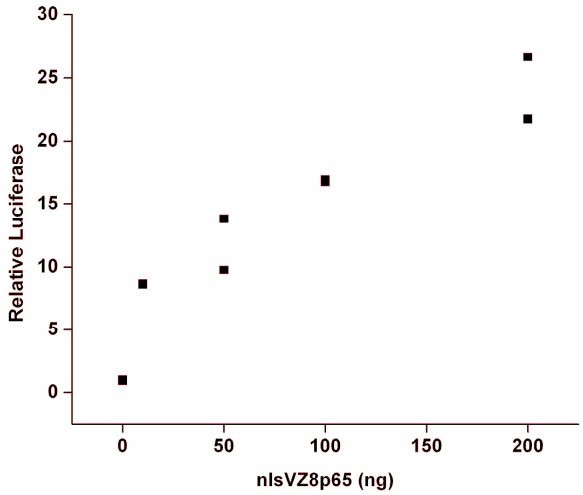


Figure 4.2. Effect of gene switch dosage on luciferase reporter induction.

4.2.2 Construct Optimization

To begin optimization of the gene switch for targeting *VEGF-A*, the effect of the activation domain was investigated. A number of activation domains are commonly used with engineered transcription factors and include the yeast GAL4, VP16 *Herpes simplex* viral protein, and the p65 subunit of the NF-kB transcription factor. While VP16 had been used in some of the previous constructs discussed in this thesis, an engineered version consisting of four copies of the VP16 minimal activation domain, called VP64, has been reported to have higher activity (33). The VP64 and p65 activation domains were cloned at the C-terminus of the GAL4 DNA binding domain and tested for their ability to induce transcription from the *GalTATA*-luciferase reporter. Both showed activity, with GAL-VP64 giving 80 fold induction and GAL-p65 showing over 11000 fold induction (Figure 4.3). The performance of an activation domain in engineered transcription factors varies depending on both the cell type and the particular construct design (23, 34, 35). Since the p65 construct gave the highest activity in the current system and is generally well-performing, this was used in further construct designs. Also for the current gene switch application, human components are preferred as they should have the lowest chance of stimulating an adverse immune response.

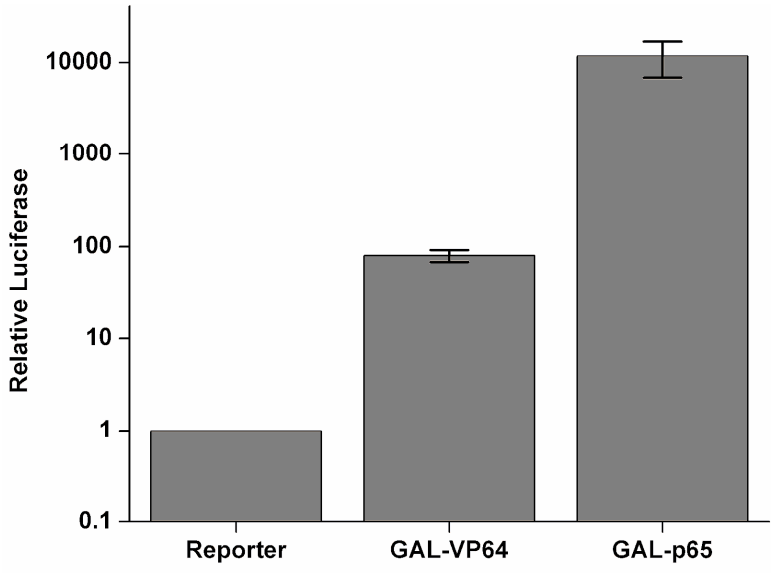


Figure 4.3. Comparison of VP64 and p65 activation domain activities. A *GalTATA*-Luciferase reporter was transfected into HeLa cells in the absence or presence of constructs containing the GAL4 DNA binding domain and either the VP64 or p65 activation domain. The data is the mean and standard error of of two independent experiments and is normalized to the reporter only condition.

Naturally evolved activation domains, including that of p65, are often composed of sub-regions that each possesses activation characteristics. The human p65 protein is 551 amino acids long and the activation domain has been localized to the C-terminal region (36, 37, 38) and functions through binding to the general transcription factors TFIIB and TATA-binding protein (39). A large portion of the activation activity can be reconstituted by solely the last 30-40 amino acids, with additional regions of activity within 416-458 and 458-521 (37). To test the effect of these regions on the gene switch, different lengths of the p65 gene were cloned at the N-terminus of the VZ8-4S to create a construct that was induced by DHB and targeted the *VEGF-A* promoter. The p65 region affected both the basal and maximal level of expression (Figure 4.4). The longest region used (282-551) gave the second highest expression level but exhibited the highest basal level, thus giving a poor performance in this construct. Each of the individually identified activation regions (amino acids 417-459, 460-523, and 510-551) from Moore et al. (37) conferred a similar low basal level, and ligand dependent induction of 6-12 fold. The best region tested here was 451-551 which showed around 17-fold induction with ligand. These results are consistent with previous studies using the GAL-p65 constructs that showed that including residues upstream of about position 414 led to lower expression (36).

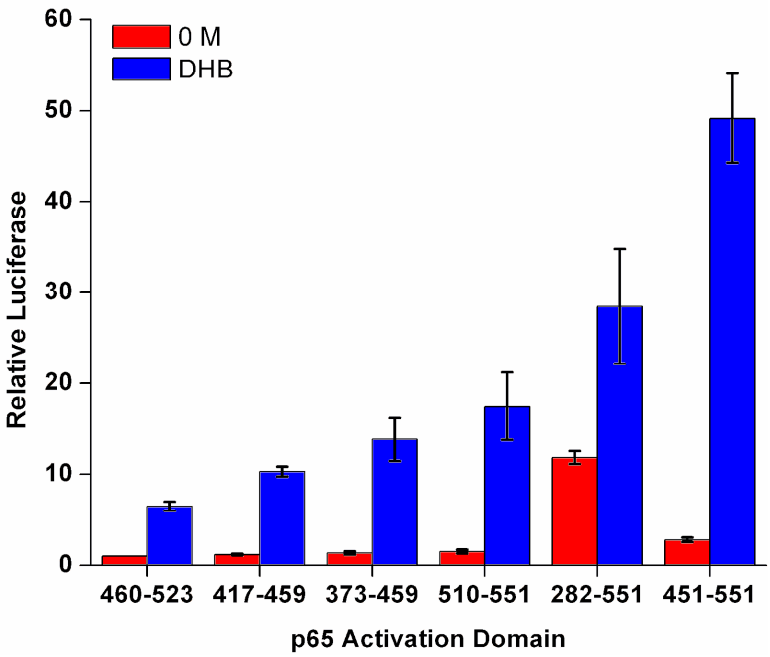


Figure 4.4. Effect of sub-regions of the p65 activation domain. Different regions of p65 (as indicated by their amino acid positions) were cloned into the gene switch p65-VZ8-4S and activation of a *VEGF-A* controlled luciferase reporter was monitored in the absence (0 M) or presence of 10^{-6} M DHB. Data is normalized such that the lowest value equaled 1.

The effect of adding interdomain linkers was investigated, since linkers can influence the expression, folding, and function of proteins. Four linkers (1-GGSGT, 2-GGSGTGSGSG, 3-FDRNRPS, and 4-MELAQRN) were selected from analyses of natural and artificial interdomain linkers (40, 41). Linkers 1 and 2 are glycine/serine-rich regions of 5 or 10 amino acids that are typical of the regions used in protein engineering to link two separate domains and are expected to allow flexibility of the domains. Linker 3 is an extended unstructured region, while linker 4 is a short helical region. These were inserted between the domains of the construct p65-VZ8-4S. As seen in Figure 4.5, the linkers did not influence the performance of the gene switch to a large extent. All constructs were functional and showed an induction of 5-10 fold.

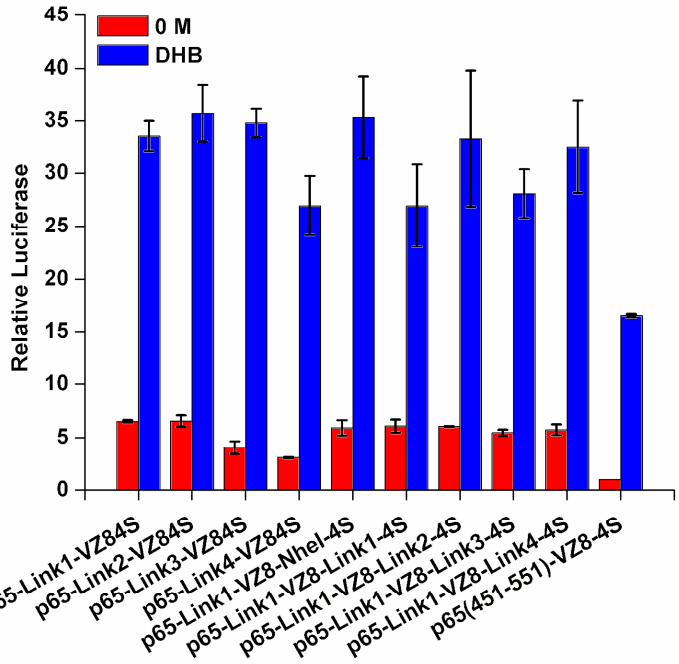


Figure 4.5. Effect of linkers between the functional domains. Induction of a *VEGF-A*-Luciferase reporter by constructs differing in the linker region between functional domains. Data is the mean and standard error of the mean of two independent experiments and is normalized by setting the lowest value to 1.

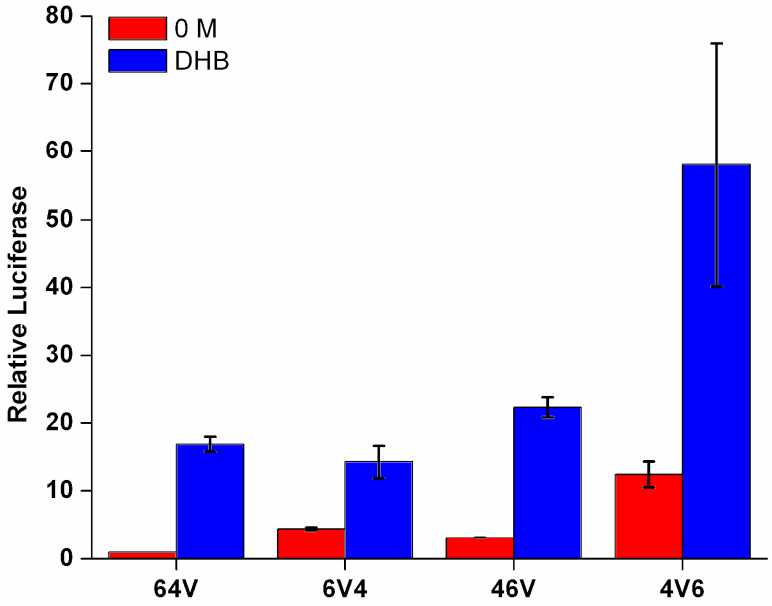


Figure 4.6. Domain reordering. Induction of a *VEGF-A*-Luciferase reporter in HeLa cells by constructs that differ by the arrangement of the functional domains p65 (6), VZ8 (V), and 4S (4). Data is the mean and standard error of the mean of two independent experiments and is normalized by setting the lowest value to 1.

The effect of domain order was examined by creating constructs with the order of the functional domains rearranged. A difference of around ten fold was observed in basal activity and around three fold for the level of induced expression. The best induction ratio was observed for the construct p65-4S-VZ8 with almost 17 fold activation (Figure 4.6). The arrangement of domains has been recognized as a significant modulator of gene switch performance (42).

To improve the performance of the gene switches, different published DNA binding domains were investigated. The DBD used so far for targeting the *VEGF-A* gene, VZ8, consists of three zinc fingers and was engineered by Liu et al. (23). Another active DBD from this paper was VZ434 which binds to the sequence GGGGGTGAC, 434 bases downstream of the main transcriptional start site. This was constructed and cloned into the constitutive nlsVZ434p65 and ligand inducible p65(451-551)-VZ434-4S gene switches and compared with the corresponding VZ8 constructs in their ability to activate the *VEGF-A*-Luciferase reporter. The original VZ8 constructs both performed better than those with VZ434, as seen in Figure 4.7. The nlsVZ8p65 switch gave 17 fold induction compared to only 11 fold induction from nlsVZ434p65. However, both of these were lower than another DBD in the F435p construct which gave 40 fold induction. This construct contains a three finger DNA binding domain that was engineered by comparison to natural human zinc fingers, along with a hemagglutinin tag, nuclear localization sequence, and p65 activation domain (28). F435p binds to the sequence GGGNGGGGA which is present on the reverse strand of the *VEGF-A* promoter at positions -90 and -391 and on the forward strand at position -1345 upstream of the main transcriptional start site. The DHB inducible VZ8 construct gave 14 fold induction compared to 4 fold for the VZ434.

The differences observed with the three published DBDs could be due to a number of reasons. The benefit of F435 over the other two DBDs could be because of the multiple binding sites in the promoter, which is known to increase the transcriptional activity (43). This is supported by the observation that the initial reported induction of F435p was only 4 fold with a *VEGF-A*-Luciferase construct that extended to promoter position -950 (28), while a 40 fold induction was seen here using a promoter containing the additional upstream binding site. There is also the possibility that the HA tag on F435p is somehow improving performance. VZ434 was reported

to stimulate *VEGF-A* transcription to a greater extent than VZ8 in HEK 293 cells while the reverse was observed here in HeLa cells which could indicate a cell-type specific property.

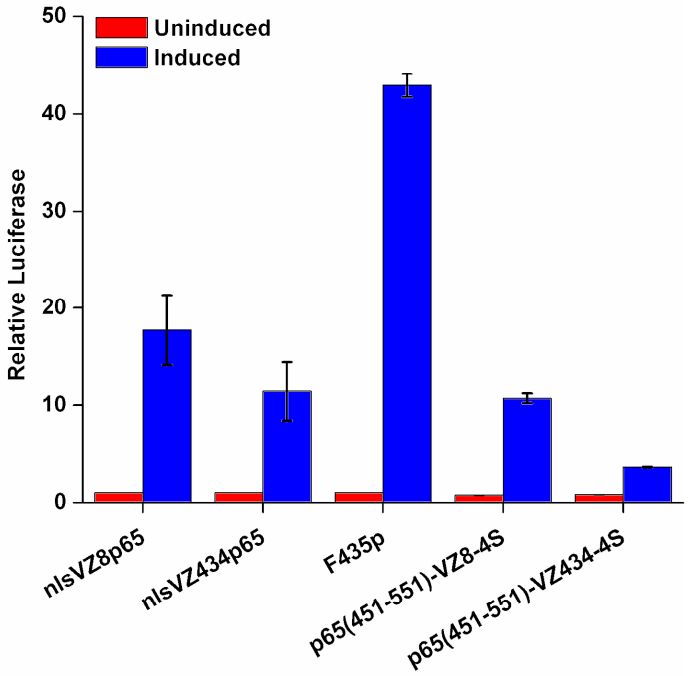


Figure 4.7. Effect of different DNA binding domains on induction of a *VEGF-A*-Luciferase reporter. For the constitutive switches (nlsVZ8p65, nlsVZ434p65, F435p) the uninduced value is for the reporter only and the induced value is for the presence of the switch. For the ligand inducible switches uninduced and induced are the values for the absence and presence of 1 μM DHB. Data is normalized to the reporter only value and is the mean and standard error of the mean of at least two samples.

To determine if increasing the DBD to six fingers, capable of targeting an 18 bp sequence, would improve performance additional fingers were selected and constructed. A new zinc finger DNA binding domain, VZ3, was designed to bind downstream of VZ8. The targeted site is *GGGGAGGAT* c GCGGAGGCT, with the VZ8 sequence italicized and the VZ3 site underlined. Based on zinc finger designs from Sangamo Biosciences (6) three individual fingers were chosen to bind GCGGAGGCT. The nucleotide sequence bound by each finger and the amino acids in the finger (-1 to position 6 of the recognition site) were: F1 GCT (QSSDLTR), F2 GAG (RSDNLAR), F3 GCG (RSDDLTR). These were cloned into the scaffold of the N1 DBD from the Barbas group (33) and assembled into the constructs p65-4S-VZ3, p65-4S-VZ38, and p65-4S-VZ3gggs8 which has a GGGS amino acid linker between the two sets of three fingers. All constructs were functional as indicated by the luciferase assay (Figure 4.8). The VZ3 domain led

to 8.5 fold induction of the *VEGF-A* luciferase reporter and was thus functional but less active than the VZ8 domain. When VZ3 and VZ8 were fused to form the six finger constructs, they showed lower levels of basal and induced expression, but showed a 14-15 fold induction. The new constructs did not, however, outperform the p65(451-551)-VZ8-4S gene switch.

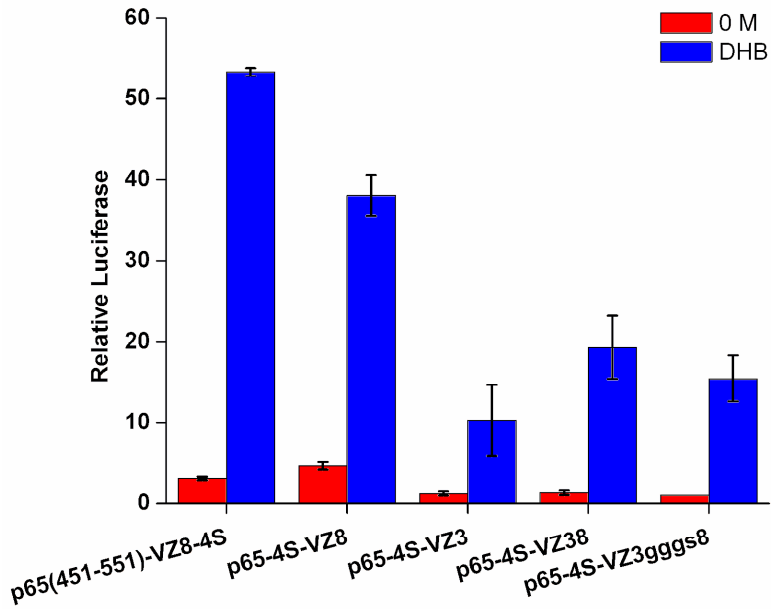


Figure 4.8. Creation of an extended DNA binding domain. A comparison of luciferase reporter induction by VZ8, the newly engineered VZ3, and two six finger combined domains.

4.2.3 Gene Switch Expression and Localization

To determine the expression of the *VEGF-A* targeting gene switch, it was investigated using western blotting. The constructs were seen to be expressed at similar levels with VZ8 containing constructs and GAL4 constructs both expressed, as well as the wild type ERα (Figure 4.9). The localization of a representative gene switch was examined by confocal fluorescence microscopy. The construct p65GAL4S was tagged with the yellow fluorescent protein at its C terminus and its expression in HeLa was visualized by confocal microscopy (Figure 4.10). By staining the cells with the DNA stain DAPI, the nuclei were visualized and the gene switch was observed in comparison. In the absence of DHB, the gene switch seemed to be localized around the nucleus or throughout the cytoplasm, while the presence of DHB seemed to stimulate a translocation into the nucleus. This is the expected mode of action for the gene switch as the construct needs to be present in the nucleus to stimulate transcription.

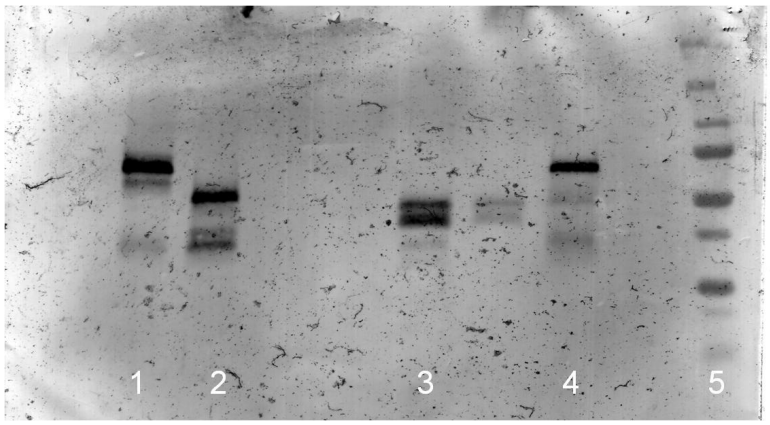


Figure 4.9. Western blot of gene switch expression. 100 ng of plasmid was transfected into HeLa cells and protein was isolated after 24 hours. Lane 1: p65(451-551)-VZ8-4S, lane 2:VZ8-ER, lane 3:GAL-L7E, lane 4: ERα, lane 5: molecular marker.

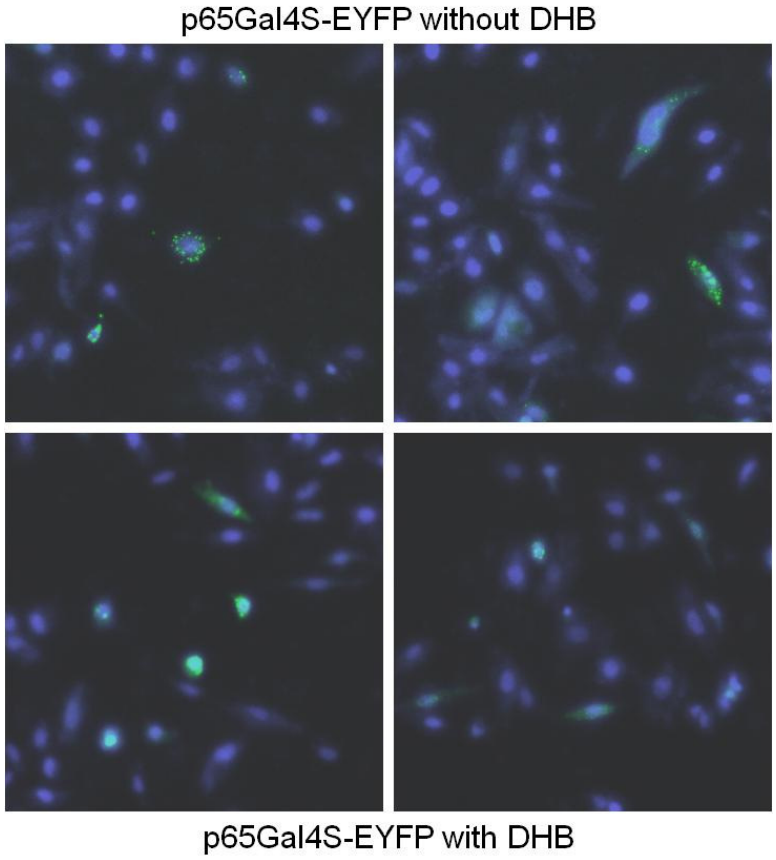


Figure 4.10. Confocal fluorescence microscopy of a gene switch expressed in HeLa cells in the presence and absence of its ligand DHB. HeLa cells were transfected with a yellow fluorescent protein (EYFP, colored green in the figure) tagged p65GAL4S switch, fixed with formaldehyde, stained with DAPI (blue), and examined by microscopy. The top images are in the absence of DHB, and the bottom images are in the presence of DHB.

4.2.4 Bimolecular Fluorescence Complementation Analysis of Gene Switch Activity

The wild type estrogen receptor alpha functions through dimerization. It is expected that the GAL4-LBD containing switches would also dimerize as GAL4 response element is bound by a homodimer of the activation domain. It is unknown whether the engineered zinc finger DNA binding domains act as a dimer, as the rationale for their use is to be able to recognize any given DNA sequence and not be constrained by a symmetrical binding site. The role of dimerization was proposed to be investigated through the use of bimolecular fluorescence complementation (BiFC). This technique relies on the ability of the green fluorescent protein and its derivatives to be cleaved at certain positions into two inactive fragments which can then reassociate to form a fluorescent signal. If these two fragments are fused to proteins that interact, the process can be detected by an increase in fluorescence. Many studies have used this technique to monitor protein-protein interactions *in vivo* (44).

Constructs were made that contained either the N terminal or C terminal half of the yellow fluorescent protein EYFP, split at residue 155, fused to either the wild type ERα or a representative gene switch p65-GAL-4S. When coexpressed in HeLa cells and analyzed by flow cytometry, no increase over the non-fluorescent control was observed for the ERα constructs (Figure 4.11a), or for p65GAL4S constructs despite a fluorescent signal visible for the full length EYFP tagged switches (Figure 4.11b). Since another yellow fluorescent protein variant, Venus, had been shown to have better performance in BiFC applications (45), constructs were recloned using Venus fragments. When coexpressed and analyzed, the constructs gave a fluorescent signal, but it was mainly dependent simply on the level of plasmid transfected and showed very little of the expected ligand dependent signal (Figure 4.12).

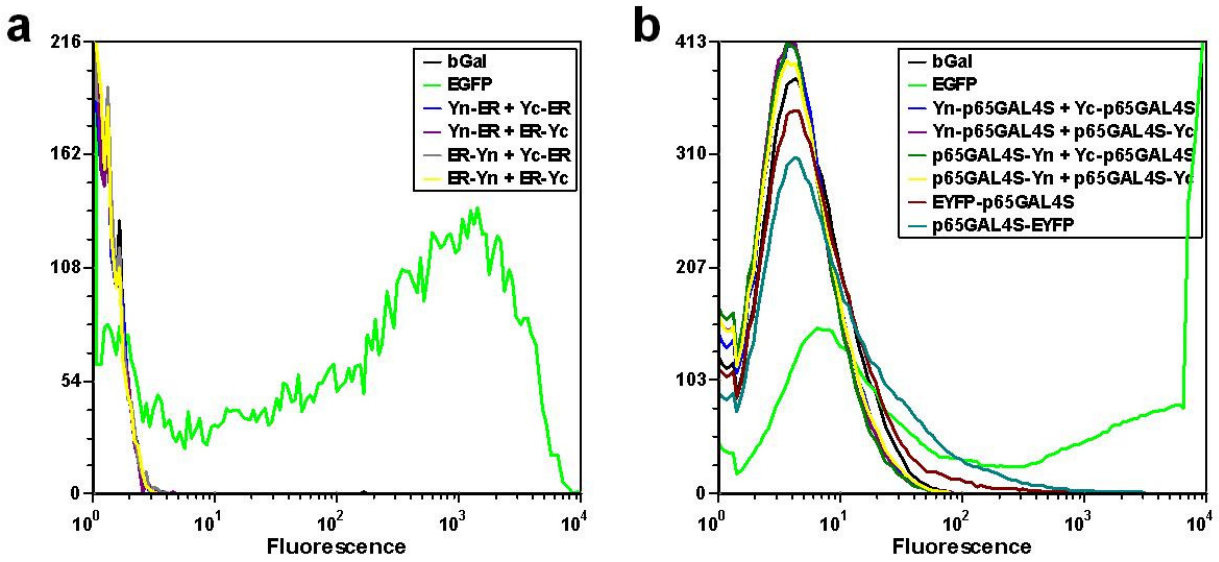


Figure 4.11. EYFP fragment tagged gene switch constructs. (a) ERα tagged at the N or C terminus with the N or C terminal fragment of EYFP. (b) p65GAL4S tagged constructs. HeLa cells were transfected, incubated 2 days, and then exposed to estradiol or DHB ligand 2 hours before being analyzed by flow cytometry.

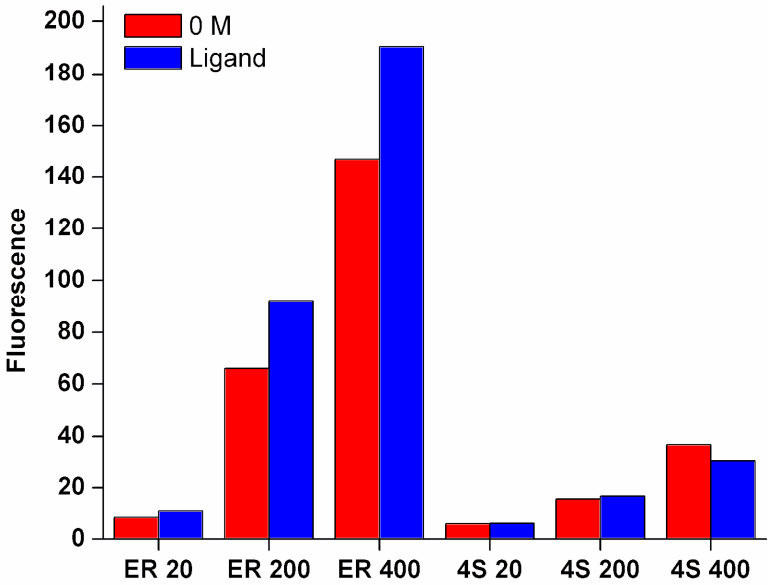


Figure 4.12. Bimolecular fluorescence complementation of mVenus fragment tagged constructs. ER samples are Vn-ER coexpressed with ER-Vc at 20, 200, or 400 ng each construct per well of a 12 well plate. 4S samples are p65-GAL-4S-Vn and p65-GAL-4S-Vc coexpressed at 20, 200, or 400 ng. Samples were expressed in the absence of ligand, or in the presence of 10 nM estradiol (ER) or 1000 nM DHB (4S).

These results suggested that the BiFC approach was not suitable for the investigation of dimerization of the gene switches. The signal that was generated was more dependent on the fluorescent protein fragments used than the gene switch, and the results of the Venus constructs suggest a non-specific interaction. For these reasons, the approach was not pursued further. However, a fluorescent complementation ligand sensor was successfully created and is discussed in Chapter 5.

4.2.5 Endogenous *VEGF-A* Targeting: Quantitative RT-PCR and ELISA

The ultimate goal of the gene switch research was to successfully alter the expression of the endogenous *VEGF-A* gene. As such, the gene switch constructs were transfected into HeLa cells and the secretion of VEGF-A into the culture media was measured by ELISA (Figure 4.13). The hypoxia mimic cobalt chloride stimulated over a 3.5 fold induction of VEGF-A secretion into the media, while the constitutive switch nlsVZ8p65 caused an increase in secretion of over 2.5 fold. The introduction of the DHB induced p65(451-551)-VZ8-4S gave a slight increase in basal production of VEGF-A, but the addition of DHB led to twice the amount of expression compared to the β-galactosidase control. These induction ratios are lower than those reported for HEK293 cells, which for the original constitutive VZ8 construct was 8 fold and around 40 fold induction was seen with a VZ434 construct and p65 activation domain (23). The F435p construct was reported to induce VEGF-A secretion 21 fold in HEK293 in transient transfection (28), while stable cell lines containing the rapamycin induced dimerizing switch with the VZ8 DBD showed 25-1000 fold induction (27). It is possible that HeLa cells are less inducible than HEK293, however, when I transfected HEK293 cells, a large basal expression of around 500 pg/mL was seen with little induction in response to cobalt chloride or gene switch presence. This could be due to differences in the transfection reagent used, or culturing conditions. Both Liu *et al*. and Bae *et al.* used Lipofectamine (Invitrogen, Carlsbad CA) for their VEGF-A induction assays (23, 28), however this reagent was replaced with a new formulation, Lipofectamine 2000, which was used in the current experiments. A subsequent paper from Sangamo Biosciences that used Lipofectamine 2000 also showed a high basal expression of 600 pg/mL in HEK293 (24). Experiments are ongoing to identify the cause of this discrepancy.

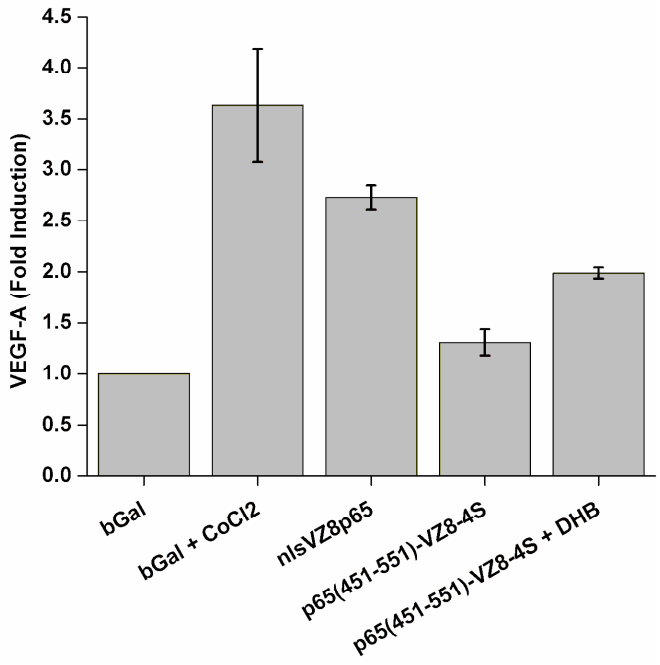


Figure 4.13. Induction of endogenous VEGF-A secretion by the action of a gene switch.

When the induction of *VEGF-A* mRNA was monitored, an increase due to the action of the gene switch was observed (Figure 4.14). As monitored in HeLa cells by quantitative reverse transcription PCR, the cobalt chloride hypoxia mimic led to a 12 fold induction of mRNA, with both the constitutive switch nlsVZ8p65 and the p65(451-551)-VZ8-4S induced by DHB giving just over twofold induction. This is somewhat less than expected based on published data from HEK293 experiments, which showed 8 fold induction with a constitutive VZ8 construct (23). However, it is similar to a number of other functional constructs (28).

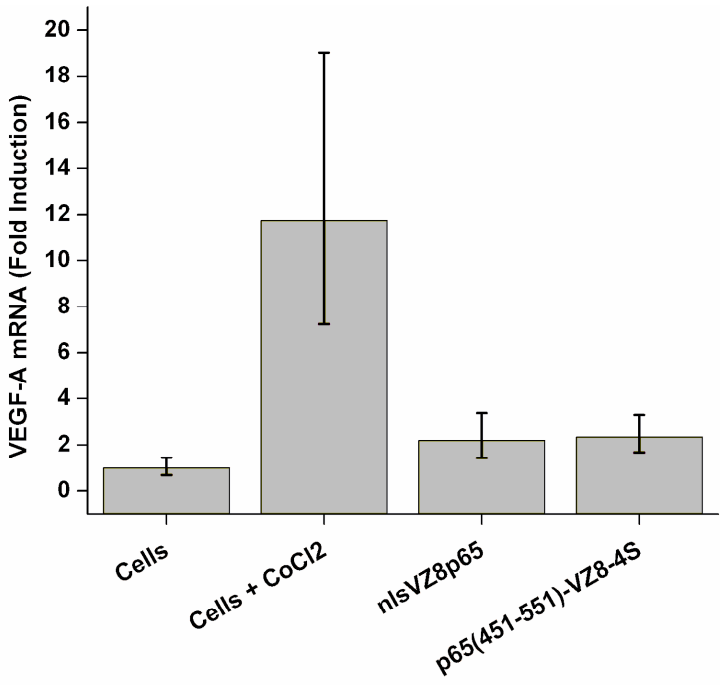


Figure 4.14. *VEGF-A* mRNA induction in HeLa cells monitored by quantitative RT-PCR. Samples are cells in the absence or presence of 100 μM cobalt chloride, or cells transfected with the constitutive nlsVZ8p65 or the inducible switch p65(451-551)-VZ8-p65 in the presence of 1 μM DHB.

4.3 Conclusions and Outlook

A ligand dependent gene switch was successfully used to upregulate the endogenous vascular endothelial growth factor A gene in cell culture. A number of factors were shown to influence the basal and induced levels of transcription from the gene switch including the DNA binding domain and activation domain that was used, and the order of arrangement of the domains in a switch. Further optimization of the gene switch could be performed to analyze whether each of these factors are independent, or if they interact in unpredictable ways within a construct or in different cell types. The difference in endogenous VEGF-A production in HEK293 cells between the current work and previously published reports remains to be solved, and experiments are ongoing to address this.

The low level of induction observed with the reporter studies show the difficulties in regulating a natural promoter as compared to a model promoter. This could be partly due to the lack of the post-transcriptional control with the luciferase reporter that is functioning with the actual *VEGF-*

A gene. These low induction levels are also likely to have contributed to the difficulties encountered in creating a mammalian screening system as discussed in Chapter 3.

4.4 Materials and Methods

4.4.1 Reagents

DNA polymerase, restriction enzymes, and T4 DNA ligase were from New England Biolabs (Beverly, MA). Chemicals were from Sigma-Aldrich (Carlsbad, CA). Cell culture media were from the UIUC cell media facility (Urbana, IL) or Invitrogen (Carlsbad, CA). DHB ligand was synthesized by Victor Gonzalez. DNA miniprep and RNA isolation kits were from Qiagen (Valencia, CA). The VEGF-A ELISA kit was from PeproTech (Rocky Hill, NJ). One component ABTS peroxidase substrate was from KPL (Gaithersburg, MD). Lipofectamine 2000 was from Invitrogen (Carlsbad, CA).

4.4.2 Cloning of Constructs

Standard molecular techniques were used for cloning the constructs used (see Table 4.1). The *VEGF-A* luciferase reporter was made by cloning a PCR product from human genomic DNA covering the region -1978 to + 1038 of the *VEGF-A* promoter between the KpnI and NcoI sites of pGL3-Basic (Promega, Madison WI). Gene switches were constructed by PCR and restriction digestion. F435p was a gift from Changkyu Oh (Toolgen, Korea).

4.4.3 Cell Culture and Transfection

HeLa cells were grown in minimal essential medium (MEM) / 1 mM sodium pyruvate / 10% fetal bovine serum (FBS) in a humidified incubator at 37 °C with 5% carbon dioxide. When cells were 80-90 % confluent, they were trypsinized and split into 12-24 well plates with MEM media / 1 mM sodium pyruvate / 5% charcoal dextran treated calf serum (CDCS). Cells were grown for 24 hours until they were over 90% confluent and transfected using lipofectamine 2000. For the gene switch optimization experiments, 100 ng of gene switch was transfected into HeLa with 100 ng β-galactosidase control vector and 600 ng *VEGF-A*-Luciferase reporter. Media was changed after 4-6 hours, ligand was added, and cells were incubated for 24 hours before lysis and detection of luciferase activity.

Table 4.1. Constructs used in this work.

HZ	Genotype	Plasmid
1676	DH5alpha	VZ8-ER pCMV5
1677	DH5alpha	VZ8-ER-p65 pCMV5
2021	DH5alpha	nls-VZ8-p65 pCMV5
3079	DH5alpha	F435p
3266	DH5alpha	GAL-VP64 pCMV5
3253	DH5alpha	GAL-p65 pCMV5
1679	DH5alpha	VEGFa-Luciferase pGL3
613	DH5alpha	GalTATA-Luciferase pGL3
1874	DH5alpha	p65(510-551)-VZ8-4S pCMV5
1875	DH5alpha	p65(451-551)-VZ8-4S pCMV5
1876	DH5alpha	p65(282-551)-VZ8-4S pCMV5
1877	DH5alpha	p65(417-459)-VZ8-4S pCMV5
1878	DH5alpha	p65(373-459)-VZ8-4S pCMV5
1879	DH5alpha	p65(460-523)-VZ8-4S pCMV5
3267	DH5alpha	p65-Link1-VZ8-4S pCMV5
3268	DH5alpha	p65-Link2-VZ8-4S pCMV5
3269	DH5alpha	p65-Link3-VZ8-4S pCMV5
3270	DH5alpha	p65-Link4-VZ8-4S pCMV5
3271	DH5alpha	p65-Link1-VZ8-NheI-4S pCMV5
3272	DH5alpha	p65-Link1-VZ8-Link1-4S pCMV5
3273	DH5alpha	p65-Link1-VZ8-Link2-4S pCMV5
3274	DH5alpha	p65-Link1-VZ8-Link3-4S pCMV5
3275	DH5alpha	p65-Link1-VZ8-Link4-4S pCMV5
3276	DH5alpha	p65-4S-VZ8 pCMV5
3277	DH5alpha	4S-p65-VZ8 pCMV5
3278	DH5alpha	4S-VZ8-p65 pCMV5
3279	DH5alpha	nls-VZ434-p65 pCMV5
3280	DH5alpha	p65(451-551)-VZ434-4S pCMV5
3281	DH5alpha	p65-4S-VZ3 pCMV5
3282	DH5alpha	p65-4S-VZ38 pCMV5
3283	DH5alpha	p65-4S-VZ3gggs8 pCMV5
3125	DH5alpha	p65-GAL-4S-EYFP pCMV5
3126	DH5alpha	p65-GAL-4S-Yn pCMV5
3127	DH5alpha	p65-GAL-4S-Yc pCMV5
3128	DH5alpha	EYFP-p65-GAL-4S pCMV5
3129	DH5alpha	Yn-p65-GAL-4S pCMV5
3130	DH5alpha	Yc-p65-GAL-4S pCMV5
3284	DH5alpha	Yn-ER pCMV5
3285	DH5alpha	Yc-ER-pCMV5
3286	DH5alpha	ER-Yn pCMV5
3287	DH5alpha	ER-Yc pCMV5
3288	DH5alpha	Vn-ER pCMV5
3289	DH5alpha	ER-Vc pCMV5
3290	DH5alpha	p65-GAL-4S-Vn pCMV5
3291	DH5alpha	p65-GAL-4S-Vc pCMV5

4.4.4 Quantitative RT-PCR

HeLa cells were transfected in a 6 well plate with 400 ng gene switch along with 3600 ng empty pCMV5 vector using Lipofectamine 2000. 1 μM DHB was added to samples containing the inducible switch. Untransfected cells were incubated in the absence or presence of 100 μM cobalt chloride. After 1 day of incubation, cells were trypsinized and collected. RNA was isolated using an RNeasy kit (Qiagen, Valencia CA), and treated with DNA-free (Ambion, Austin TX) to remove genomic DNA contamination. Quantitative RT-PCR was performed using a HotStartIT SYBR Green mastermix (USB, Cleveland OH) in a MiniOpticon thermal cycler (Bio-Rad, Hercules CA). *VEGF-A* was amplified with primers VEGFA-mRNA-F (GTG CAT TGG AGC CTT GCC TGG) and VEGFA-mRNA-R (ACT CGA TCT CAT CAG GGT ACT C) and normalized to β-actin with primers bACTIN-F (GCA CAG AGC CTC GCC TT) and bACTIN-R (GTT GTC GAC GAC GAG CG).

4.4.5 ELISA

The human VEGF ELISA development kit (PeproTech, Rocky Hill NJ) was used for the determination of VEGF-A secreted into the media of cell culture experiments. The capture antibody was diluted to 0.5 μg/mL with PBS and 100 μL was added per well of a MaxiSorp 96 well plate (Nunc, Rochester NY) for incubation at room temperature overnight. The liquid was removed and wells were washed four times with 300 μL of wash buffer (0.05 % Tween-20 in PBS). 300 μL of blocking reagent, consisting of 1 % bovine serum albumin (BSA) in PBS, was added per well and the plate was incubated for 1 hour. Dilutions of the supplied human VEGF standard were made at concentrations of 0, 50, 100, 200, 400, 600, 800, and 1000 pg/mL in dilution buffer (0.05 % Tween-20, 0.1 % BSA in PBS). The blocking reagent was removed and wells were washed four times with 300 μL of wash buffer then 100 μL of standard or media sample was added to duplicate or triplicate wells, followed by a 2 hour incubation. Wells were washed and 100 μL detection antibody (diluted in dilution buffer to 0.25 μg/mL) was added and incubated for 2 hours. After washing, the avidin-HRP conjugate was diluted 2000 fold and 100 mL was added per well and incubated 30 min. The plate was washed and 100 μL of one component ABTS peroxidase substrate was added and incubated. At 10 minute intervals the plate was scanned at 405 nm with correction at 650 nm using a Spectramax 190 microplate

reader (Molecular Devices, Sunnyvale CA). Data was used from a scan time when the corrected OD was less than 0.2 for the zero samples or 1.2 for the 1000 pg/mL sample.

4.4.6 Western Blotting

HeLa cells were split into a 24 well plate and grown overnight until almost confluent. Cells were transfected using lipofectamine 2000 with 100 ng of gene switch plasmid and 450 ng empty pCMV5 vector. After 5 hours of exposure to the lipofection mixture the media was replaced with fresh MEM / 5 % CDCS, and cells were incubated 24 hours. Cells were collected by trypsinization, resuspended in 100 µL PBS and lysed by freezing at -80 °C. The sample was thawed, pelleted to remove debris, and 0.5 µL of 100 mM phenylmethylsulfonyl fluoride was added to inhibit protein degradation. 5 µL of the lysate was added to an equal volume of 4x Laemmli buffer, boiled for 5 min, and centrifuged a full speed for 1 min. The sample was loaded on a 20% SDS-PAGE gel with a 4 % stack and underwent electrophoresis. The protein was transferred to a PVDF membrane via electroblotting and the membrane was blocked overnight at 4 °C. The F10 antibody to ERα (Santa Cruz Biotechnology, Santa Cruz CA) was diluted 1/1000 and incubated with the membrane for 1 hour. The membrane was washed three times and the secondary antibody incubation was performed with a goat antimouse antibody for 30 min. The blot was washed three times and developed using Western Blue stabilized substrate for alkaline phosphatase (Promega, Madison WI).

4.4.7 Confocal Microscopy

Hela cells were grown in MEM on glass cover slips in 6 well plates and transfected when they were 50 % confluent. 1000 ng gene switch with 3000 ng empty pCMV5 vector were lipofected using 7.5 µL Lipofectamine 2000 per well. The media was replaced after 6 hours and 1 µM DHB was added to the ligand condition. Cells were grown for 24 hours then fixed. Cover slips were rinsed with PBS, then incubated with 0.3 % Triton X-100 / 4 % formaldehyde in PBS for 7.5 min, rinsed with PBS then incubated with 4 % formaldehyde in PBS for another 7.5 min, followed by three five minute washes with PBS. Cover slips were stored in PBS at 4 °C until use. Cells were stained with 1 µg/mL diamidino-2-phenylindole (DAPI) for 5 min, followed by a PBS wash. Cover slips were placed cell side down on a microscope slide and the corners were

fixed in place using clear nail varnish. Images were collected on a Zeiss confocal LSM510 microscope using a FITC filter set for the EYFP tagged switch, and a bright field view for DAPI.

4.5 References

1. Pavletich,N.P. and Pabo,C.O. (1991) Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1 A. *Science,* **252,** 809-817.

2. Sera,T. (2009) Zinc-finger-based artificial transcription factors and their applications. *Adv. Drug Deliv. Rev.,* **61,** 513-526.

3. Choo,Y. and Klug,A. (1994) Toward a code for the interactions of zinc fingers with DNA: Selection of randomized fingers displayed on phage. *Proc. Natl. Acad. Sci. U. S. A.,* **91,** 11163-11167.

4. Segal,D.J., Dreier,B., Beerli,R.R. and Barbas,C.F.,3rd. (1999) Toward controlling gene expression at will: Selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc. Natl. Acad. Sci. U. S. A.,* **96,** 2758-2763.

5. Sera,T. and Uranga,C. (2002) Rational design of artificial zinc-finger proteins using a nondegenerate recognition code table. *Biochemistry,* **41,** 7074-7081.

6. Liu,Q., Xia,Z., Zhong,X. and Case,C.C. (2002) Validated zinc finger protein designs for all 16 GNN DNA triplet targets. *J. Biol. Chem.,* **277,** 3850-3856.

7. Maeder,M.L., Thibodeau-Beganny,S., Osiak,A., Wright,D.A., Anthony,R.M., Eichtinger,M., Jiang,T., Foley,J.E., Winfrey,R.J., Townsend,J.A., et al. (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell,* **31,** 294-301.

8. Birk,D.M., Barbato,J., Mureebe,L. and Chaer,R.A. (2008) Current insights on the biology and clinical aspects of VEGF regulation. *Vasc. Endovascular Surg.,* **42,** 517-530.

9. Connolly,D.T., Heuvelman,D.M., Nelson,R., Olander,J.V., Eppley,B.L., Delfino,J.J., Siegel,N.R., Leimgruber,R.M. and Feder,J. (1989) Tumor vascular permeability factor stimulates endothelial cell growth and angiogenesis. *J. Clin. Invest.,* **84,** 1470-1478.

10. Ferrara,N. and Henzel,W.J. (1989) Pituitary follicular cells secrete a novel heparin-binding growth factor specific for vascular endothelial cells. *Biochem. Biophys. Res. Commun.,* **161,** 851-858.

11. Gospodarowicz,D., Abraham,J.A. and Schilling,J. (1989) Isolation and characterization of a vascular endothelial cell mitogen produced by pituitary-derived folliculo stellate cells. *Proc. Natl. Acad. Sci. U. S. A.,* **86,** 7311-7315.

12. Ferrara,N., Carver-Moore,K., Chen,H., Dowd,M., Lu,L., O'Shea,K.S., Powell-Braxton,L., Hillan,K.J. and Moore,M.W. (1996) Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene. *Nature,* **380,** 439-442.

13. Bao,P., Kodra,A., Tomic-Canic,M., Golinko,M.S., Ehrlich,H.P. and Brem,H. (2009) The role of vascular endothelial growth factor in wound healing. *J. Surg. Res.,* **153,** 347-358.

14. Ferrara,N. (2005) VEGF as a therapeutic target in cancer. *Oncology,* **69 Suppl 3,** 11-16.

15. Yancopoulos,G.D., Davis,S., Gale,N.W., Rudge,J.S., Wiegand,S.J. and Holash,J. (2000) Vascular-specific growth factors and blood vessel formation. *Nature,* **407,** 242-248.

16. Tischer,E., Mitchell,R., Hartman,T., Silva,M., Gospodarowicz,D., Fiddes,J.C. and Abraham,J.A. (1991) The human gene for vascular endothelial growth factor. multiple protein forms are encoded through alternative exon splicing. *J. Biol. Chem.,* **266,** 11947-11954.

17. Pages,G. and Pouyssegur,J. (2005) Transcriptional regulation of the vascular endothelial growth factor gene--a concert of activating factors. *Cardiovasc. Res.,* **65,** 564-573.

18. Akiri,G., Nahari,D., Finkelstein,Y., Le,S.Y., Elroy-Stein,O. and Levi,B.Z. (1998) Regulation of vascular endothelial growth factor (VEGF) expression is mediated by internal initiation of translation and alternative initiation of transcription. *Oncogene,* **17,** 227-236.

19. Bastide,A., Karaa,Z., Bornes,S., Hieblot,C., Lacazette,E., Prats,H. and Touriol,C. (2008) An upstream open reading frame within an IRES controls expression of a specific VEGF-A isoform. *Nucleic Acids Res.,* **36,** 2434-2445.

20. Meiron,M., Anunu,R., Scheinman,E.J., Hashmueli,S. and Levi,B.Z. (2001) New isoforms of VEGF are translated from alternative initiation CUG codons located in its 5'UTR. *Biochem. Biophys. Res. Commun.,* **282,** 1053-1060.

21. Levy,A.P., Levy,N.S. and Goldberg,M.A. (1996) Post-transcriptional regulation of vascular endothelial growth factor by hypoxia. *J. Biol. Chem.,* **271,** 2746-2753.

22. Ray,P.S., Jia,J., Yao,P., Majumder,M., Hatzoglou,M. and Fox,P.L. (2009) A stress-responsive RNA switch regulates VEGFA expression. *Nature,* **457,** 915-919.

23. Liu,P.Q., Rebar,E.J., Zhang,L., Liu,Q., Jamieson,A.C., Liang,Y., Qi,H., Li,P.X., Chen,B., Mendel,M.C., et al. (2001) Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. activation of vascular endothelial growth factor A. *J. Biol. Chem.,* **276,** 11323-11334.

24. Snowden,A.W., Gregory,P.D., Case,C.C. and Pabo,C.O. (2002) Gene-specific targeting of H3K9 methylation is sufficient for initiating repression in vivo. *Curr. Biol.,* **12,** 2159-2166.

25. Rebar,E.J., Huang,Y., Hickey,R., Nath,A.K., Meoli,D., Nath,S., Chen,B., Xu,L., Liang,Y., Jamieson,A.C., et al. (2002) Induction of angiogenesis in a mouse model using engineered transcription factors. *Nat. Med.,* **8,** 1427-1432.

26. Dent,C.L., Lau,G., Drake,E.A., Yoon,A., Case,C.C. and Gregory,P.D. (2007) Regulation of endogenous gene expression using small molecule-controlled engineered zinc-finger protein transcription factors. *Gene Ther.,* **14,** 1362-1369.

27. Pollock,R., Giel,M., Linher,K. and Clackson,T. (2002) Regulation of endogenous gene expression with a small-molecule dimerizer. *Nat. Biotechnol.,* **20,** 729-733.

28. Bae,K.H., Kwon,Y.D., Shin,H.C., Hwang,M.S., Ryu,E.H., Park,K.S., Yang,H.Y., Lee,D.K., Lee,Y., Park,J., et al. (2003) Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nat. Biotechnol.,* **21,** 275-280.

29. Kwon,H.S., Shin,H.C. and Kim,J.S. (2005) Suppression of vascular endothelial growth factor expression at the transcriptional and post-transcriptional levels. *Nucleic Acids Res.,* **33,** e74.

30. Yun,C.O., Shin,H.C., Kim,T.D., Yoon,W.H., Kang,Y.A., Kwon,H.S., Kim,S.K. and Kim,J.S. (2008) Transduction of artificial transcriptional regulatory proteins into human cells. *Nucleic Acids Res.,* **36,** e103.

31. Tachikawa,K., Schroder,O., Frey,G., Briggs,S.P. and Sera,T. (2004) Regulation of the endogenous VEGF-A gene by exogenous designed regulatory proteins. *Proc. Natl. Acad. Sci. U. S. A.,* **101,** 15225-15230.

32. Mori,T., Sasaki,J., Saito,Y., Aoyama,Y. and Sera,T. (2010) Self-propagating artificial transcription factors to enhance upregulation of target genes. *Bioorg. Med. Chem. Lett.,* **20,** 3479-3481.

33. Beerli,R.R., Segal,D.J., Dreier,B. and Barbas,C.F.,3rd. (1998) Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U. S. A.,* **95,** 14628-14633.

34. Weber,W., Kramer,B.P., Fux,C., Keller,B. and Fussenegger,M. (2002) Novel promoter/transactivator configurations for macrolide- and streptogramin-responsive transgene expression in mammalian cells. *J. Gene Med.,* **4,** 676-686.

35. Karzenowski,D., Potter,D.W. and Padidam,M. (2005) Inducible control of transgene expression with ecdysone receptor: Gene switches with high sensitivity, robust expression, and reduced size. *BioTechniques,* **39,** 191-196.

36. Ballard,D.W., Dixon,E.P., Peffer,N.J., Bogerd,H., Doerre,S., Stein,B. and Greene,W.C. (1992) The 65-kDa subunit of human NF-kappa B functions as a potent transcriptional activator and a target for v-rel-mediated repression. *Proc. Natl. Acad. Sci. U. S. A.,* **89,** 1875-1879.

37. Moore,P.A., Ruben,S.M. and Rosen,C.A. (1993) Conservation of transcriptional activation functions of the NF-kappa B p50 and p65 subunits in mammalian cells and saccharomyces cerevisiae. *Mol. Cell. Biol.,* **13,** 1666-1674.

38. Schmitz,M.L. and Baeuerle,P.A. (1991) The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B. *EMBO J.,* **10,** 3805-3817.

39. Schmitz,M.L., Stelzer,G., Altmann,H., Meisterernst,M. and Baeuerle,P.A. (1995) Interaction of the COOH-terminal transactivation domain of p65 NF-kappa B with TATA-binding protein, transcription factor IIB, and coactivators. *J. Biol. Chem.,* **270,** 7219-7226.

40. Argos,P. (1990) An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *J. Mol. Biol.,* **211,** 943-958.

41. George,R.A. and Heringa,J. (2002) An analysis of protein domain linkers: Their classification and role in protein folding. *Protein Eng.,* **15,** 871-879.

42. Beerli,R.R., Schopfer,U., Dreier,B. and Barbas,C.F.,3rd. (2000) Chemically regulated zinc finger transcription factors. *J. Biol. Chem.,* **275,** 32617-27.

43. Georges,A.B., Benayoun,B.A., Caburet,S. and Veitia,R.A. (2010) Generic binding sites, generic DNA-binding domains: Where does specific promoter recognition come from? *FASEB J.,* **24,** 346-356.

44. Kerppola,T.K. (2008) Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu. Rev. Biophys.,* **37,** 465-487.

45. Shyu,Y.J., Liu,H., Deng,X. and Hu,C.D. (2006) Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions. *BioTechniques,* **40,** 61-66.

Chapter 5: Development of a Fluorescent Biosensor for Detection of Estrogenic Compounds

5.1 Introduction

Small molecules play essential roles in biology as metabolites, signaling compounds, and allosteric regulators of protein function and as such it is desirable to detect their presence and monitor their effects. As a member of the nuclear hormone receptor superfamily, the estrogen receptor alpha (ERα) acts as a ligand-regulated transcription factor that is involved in a wide range of physiological processes such as mammary gland development, fertility, bone growth and maintenance, and metabolism (1). Its canonical ligand is 17β-estradiol (E_2) which binds the ligand binding domain (LBD) and induces a conformational change in the position of helix 12 that stimulates dimerization, the recruitment of coactivators or corepressors, and produces an effect on gene transcription (2). ERα is also bound and influenced by a multitude of other steroid hormones and pharmaceuticals (Figure 5.1), as well as natural products and industrial chemicals (Figure 5.2) (3, 4, 5). As a result there has been much research into developing methods of identifying compounds that interact with ERα, which can be used to measure their concentrations, and also elucidate the conformational changes they induce.

The ligands capable of binding to the ERα ligand binding pocket, and the structural mechanisms underlying binding, have been extensively investigated (6, 7). E_2 is a four-ringed, hydrophobic, steroid molecule with hydroxyl groups (the A ring phenolic hydroxyl and the D ring hydroxyl) that form important contacts within the ligand binding pocket. Compounds with similar structures, especially those maintaining the phenolic A ring, can show measurable binding to the receptor. The primary endogenous compounds are 17β-estradiol, estriol, and estrone (along with its sulfated form) which can be metabolized by hydroxylation leading to the formation of 2-hydroxy and 4-hydroxy estradiol (8). Diethylstilbestrol was an early synthetic estrogen drug whose use was discontinued after it was discovered to promote developmental defects and cancer following *in utero* exposure (9). Other pharmaceuticals that interact with the estrogen receptor include 4-hydroxytamoxifen, raloxifene, dienestrol, norethindrone, and ICI 182780.

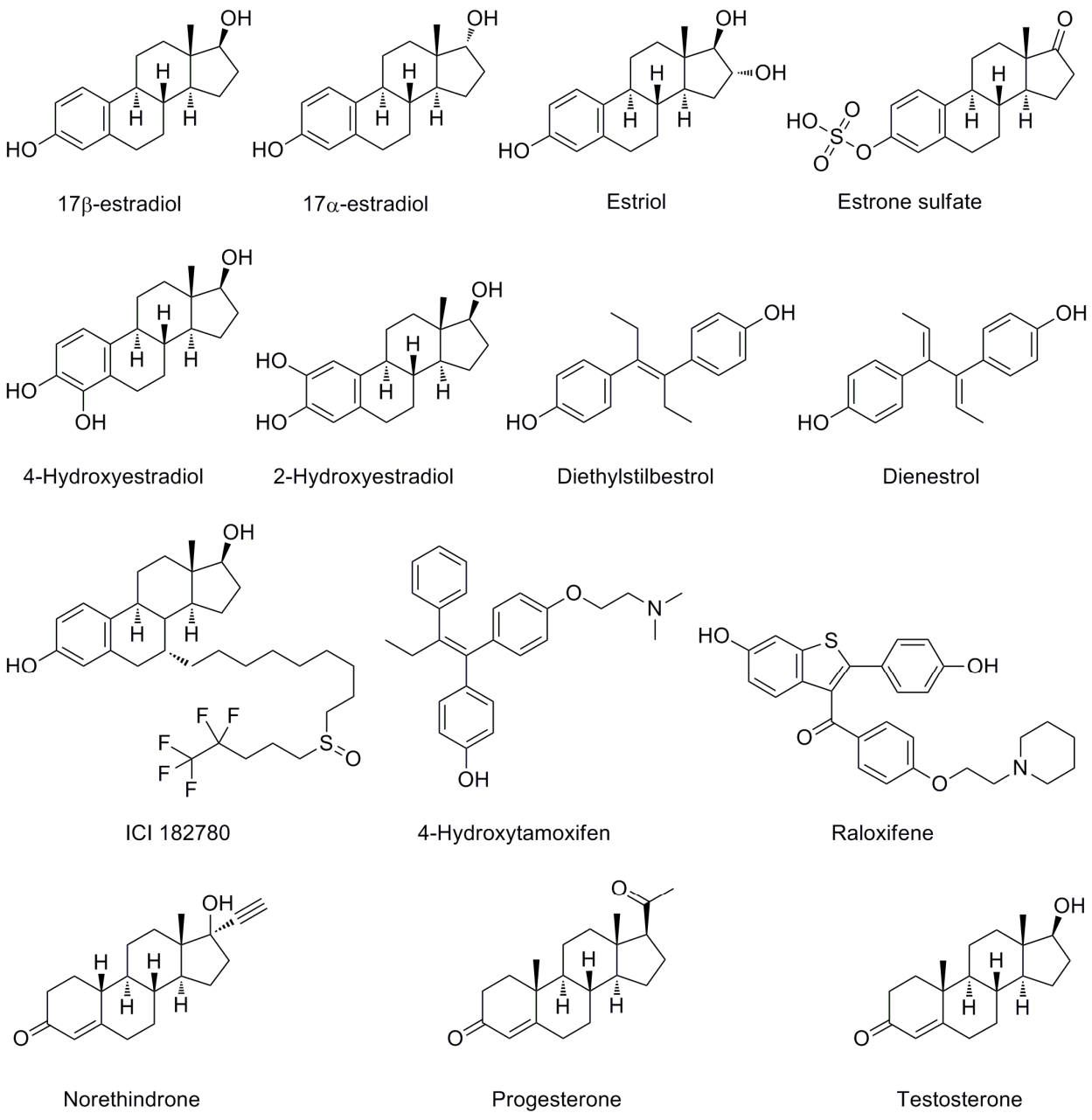


Figure 5.1. Endogenous estrogens, pharmaceuticals, and steroid hormones.

Some natural products and industrial chemicals exhibit estrogenic activity (Figure 5.2). Phytoestrogens such as genistein and daidzein from soy, and resveratrol from grapes, have been proposed to have beneficial health effects via an interaction with the estrogen receptor (10). Alternatively, there is a growing concern over the effect of organic compounds produced by the chemical industry. While a single phenol group does not show estrogenic activation, larger organic compounds such as bisphenol A, and the detergent byproducts tert-octylphenol and

nonylphenol do show such effects (5). It is of great interest to develop easy ways of measuring the estrogenic effect of the many compounds people come into contact with during their daily lives.

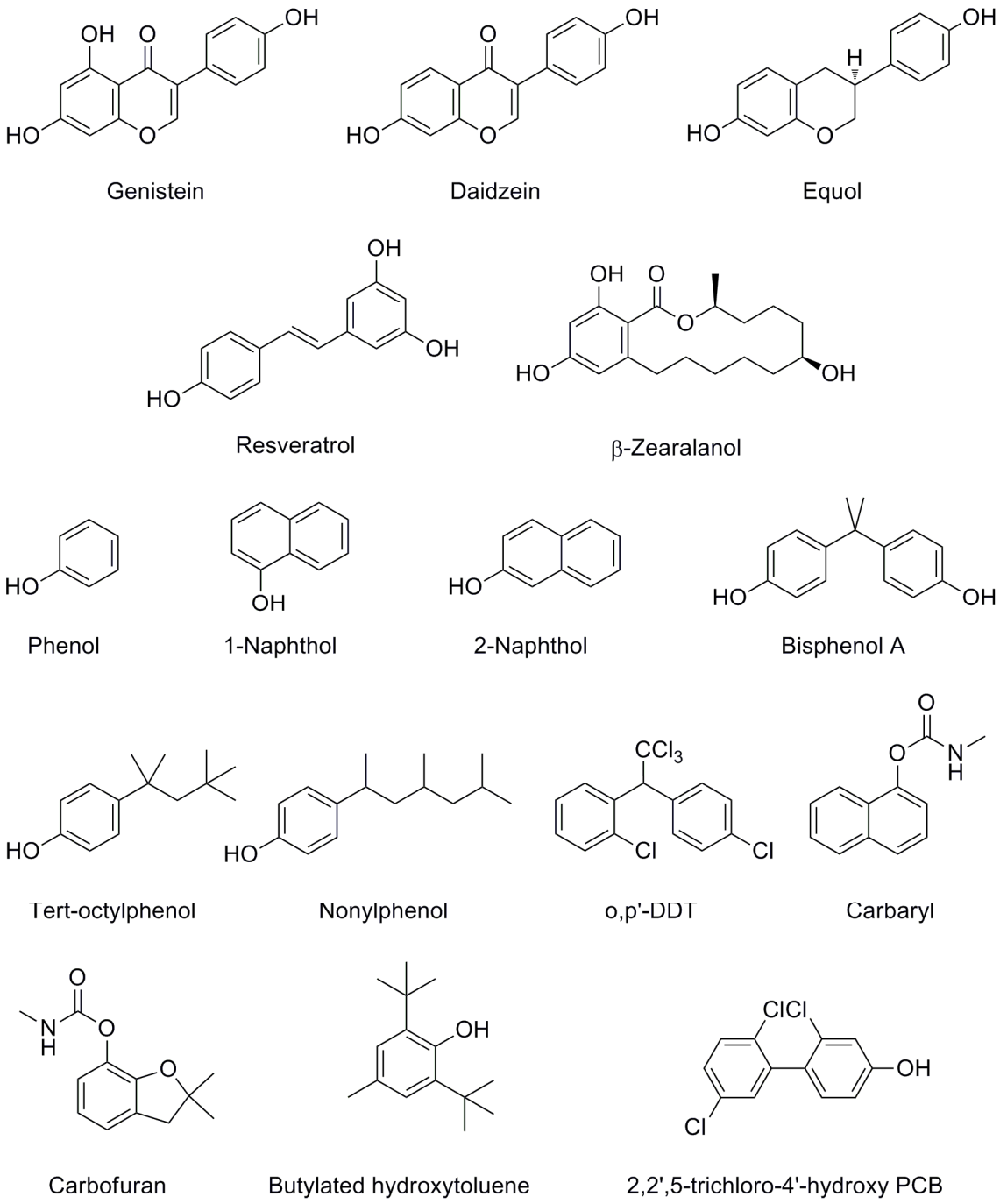


Figure 5.2. Natural products and industrial chemicals discussed in this chapter, some of which are estrogenic.

Many techniques have been developed to monitor small molecules and their binding to proteins. These include biophysical techniques such as X-ray crystallography, nuclear magnetic resonance (NMR), microcalorimetry, mass spectrometry, and surface plasmon resonance which are reviewed elsewhere (11). Biosensors are a method of detection where the sensor is composed of biological material, acting either *in vivo* or *in vitro*. These can generally be classified as either systems that stimulate the induction of an easily measured reporter gene, or molecular systems that directly generate a signal themselves (12).

Reporter genes have been used extensively both in fundamental studies of transcription and also as sensors (12, 13). They are popular due to their ease of measurement and their ability to be generalized for detection of different ligands, relying on the fusion of the reporter gene to a ligand-regulated promoter that results in transcriptional activation in the presence of the target ligand. Common examples include the genes chloramphenicol acetyl transferase, β-galactosidase, bacterial luciferase, firefly luciferase, uroporphyrinogen III methyltransferase, β-glucuronidase, and autofluorescent proteins such as the green fluorescent protein. These reporter gene assays are typically performed within easily grown hosts such as *Escherichia coli*, *Saccharomyces cerevisiae*, or mammalian cells in culture.

Molecular biosensors are engineered proteins that contain segments of fluorescent proteins (14, 15), bioluminescent proteins (16, 17), or enzymatic domains, or consist of purified protein that is monitored through optical or other means. Purified protein can be incubated with either a radiolabelled or fluorescently modified ligand, whose displacement can be monitored under the influence of the test ligand. Many applications have been developed using Fluorescence Resonance Energy Transfer (FRET), which can be used in living cells (18). This involves the nonradiative transfer of energy from an excited donor (such as cyan fluorescent protein) to an acceptor fluorophore (such as yellow fluorescent protein) over short distances of up to 10 nm and can therefore be used to monitor protein interactions or conformational changes within a protein. Bioluminescence resonance energy transfer (BRET) is similar but the donor is a luciferase enzyme that catalyzes the oxidation of a coelenterazine molecule (19).

Protein complementation assays are based on the interesting ability shown by a number of proteins to be cleavable at particular positions to produce two inactive fragments that are capable of reassociation to reproduce the functional protein (20). Examples include β-lactamase, β-galactosidase, dihydrofolate reductase, luciferase, and fluorescent proteins. Bimolecular fluorescence complementation (BiFC) is a recently developed method for analyzing protein interactions based on the reconstitution of a functional fluorescent protein from non fluorescent fragments (21, 22). The reassociation is stimulated by co-localization of the fragments via fusion to two interacting proteins, thus forming the foundation of an assay for protein-protein interaction (Figure 5.3).

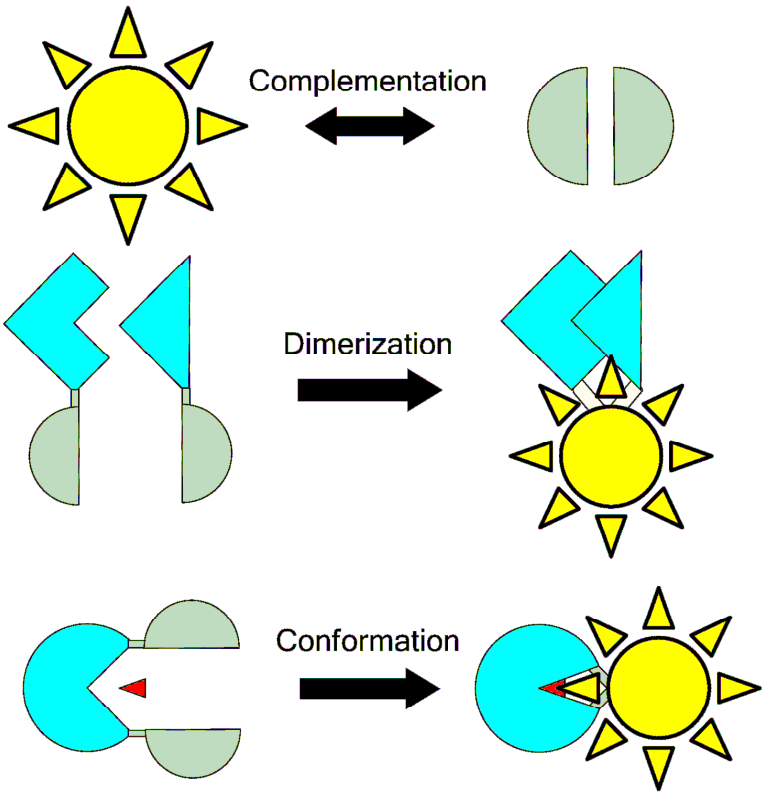


Figure 5.3. Protein complementation assays. The top panel shows the splitting of a fluorescent or luminescent protein into non active fragments. The middle panel shows the reassociation of these fragments when they are brought into contact through attachment to dimerizing proteins. The third panel shows the reconstitution of fluorescence or luminescence via a ligand stimulated conformational change in a protein linked to both fragments.

Many biosensors and bioassays have been developed for detection of (or regulation via) estrogenic compounds (23, 24). Systems using reporter genes include fluorescent proteins such as green fluorescent protein (GFP) (25, 26, 27), β-galactosidase (28, 29, 30, 31, 32), β-lactamase (33), bacterial luciferase (34), and the yeast HIS3 enzyme (35). Competitive binding to ERα has been monitored with purified protein and ^{3}H labeled estradiol (36), or fluorescent ligands (37, 38). The native biological effects of estrogenic compounds have been monitored via gene transcription (39), growth of breast cancer MCF-7 cells (40), and induction of the vitellogenin protein in fish (41). Fluorescence assays have been developed by linking the ligand binding domain of ERα to fluorescent proteins and monitoring the stabilization of fluorescence by ligand binding (42), or monitoring the ligand controlled position of helix 12 by using FRET (43, 44, 45). Protein complementation assays using luciferase have been established (46, 47, 48, 49), and a conceptually similar approach using removal of an interfering protein intein to regenerate a selectable thymidilate synthase enzyme was also successful (50, 51, 52, 53).

The goal of this chapter was to construct and determine the utility of an estrogenic compound biosensor based on the technique of fluorescence complementation. A series of constructs were made consisting of different lengths of the ERα LBD flanked by two halves of the split fluorescent protein mVenus. When expressed in mammalian cells, certain of these constructs showed an increase in fluorescence in the presence of ligands known to bind the estrogen receptor, and some distinguished agonists from antagonists. The fluorescent induction began within an hour of ligand addition, and increased over 24 h of incubation. A good correlation between the expected relative binding affinities and the measured EC_{50} values of different compounds was seen, and the sensor allowed for the detection of environmental compounds of interest.

5.2 Results and Discussion

5.2.1 Cloning of mVenus and Sensor Constructs

To create the estrogenic compound biosensor, it was decided to use the fluorescent protein Venus since it is bright, shows fast maturation at 37 °C (54), and has been shown to perform well in bimolecular fluorescence complementation applications (55). The mutations F46L, F64L, M153T, V163A, and S175G were introduced into the EYFP plasmid to convert it to Venus, as well as the A206K mutation which inhibits the weak dimerization activity of GFP variants (56), thus producing the monomeric mVenus construct. This was then split at position 155 to produce the two non fluorescent fragments Vn (amino acids 1-154) and Vc (amino acid 155-239). Transfection studies where the plasmid constructs were introduced into HeLa cells and analyzed by flow cytometry confirmed that mVenus was expressed and fluorescent. The two fragments Vn and Vc, were not fluorescent when expressed independently but when co-expressed at high levels they exhibited spontaneous association to give a fluorescent signal (Figure 5.4).

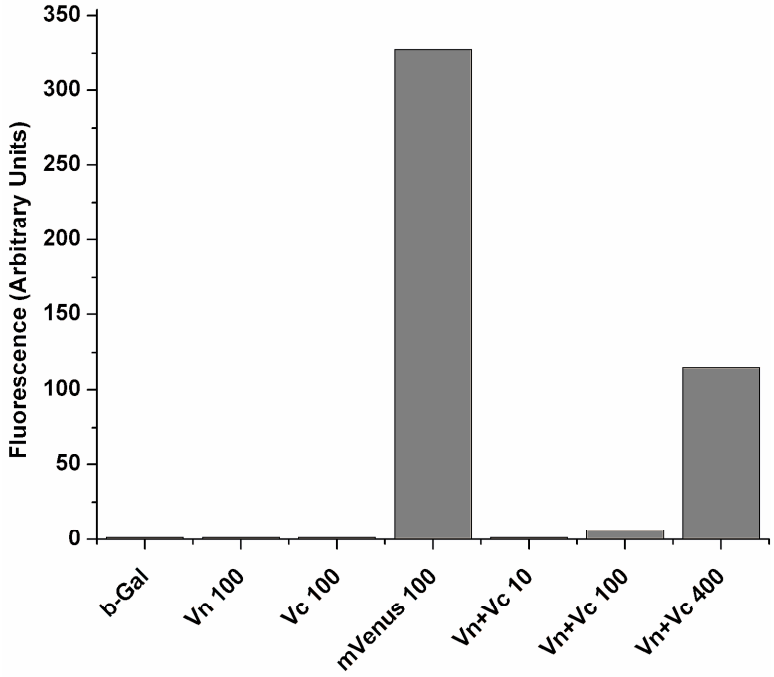


Figure 5.4. Fluorescence of mVenus and fragments in HeLa cells. Plasmid was lipofected into HeLa cells (along with β-galactosidase plasmid to bring the total DNA to 800 ng) in 24-well plates. Samples from left to right are β-galactosidase, 100 ng Vn, 100 ng Vc, 100 ng mVenus, 10 ng of Vn and Vc, 100 ng of Vn and Vc, 400 ng of Vn and Vc.

It was reasoned that by fusing the non fluorescent fragments to the ligand binding domain of ERα, the conformational change upon ligand binding would bring the fragments close enough to complement each other. This approach was analogous to that used for previously reported FRET and luciferase complementation sensors. A series of constructs (referred to in this chapter as constructs 1 to 6) were made consisting of the ligand binding domain of the estrogen receptor α flanked by Vn and Vc (Figure 5.5). The two starting points for the LBD were amino acid positions 281 and 312, and the ending points were 532, 549, or 595 giving six different constructs. Position 281 will include most of the hinge region upstream of the LBD and has been used as the fusion site in other sensors (43, 46), while position 312 is a few residues downstream of the LBD starting position and has been used successfully for the gene switches discussed in other chapters of this work. Position 532 was used as the end of the LBD for the sensor made by Muddana and Peterson and removes helix 12 (42), ending at 549 will remove the less understood F domain and was used as the fusion point for Paulmurugan and Ghambir's luciferase complementation construct (46), as was the C-terminus of the receptor at position 595. (Book 6 pg 55-74, 104-107, 125-146; Book 7 pg9-41)

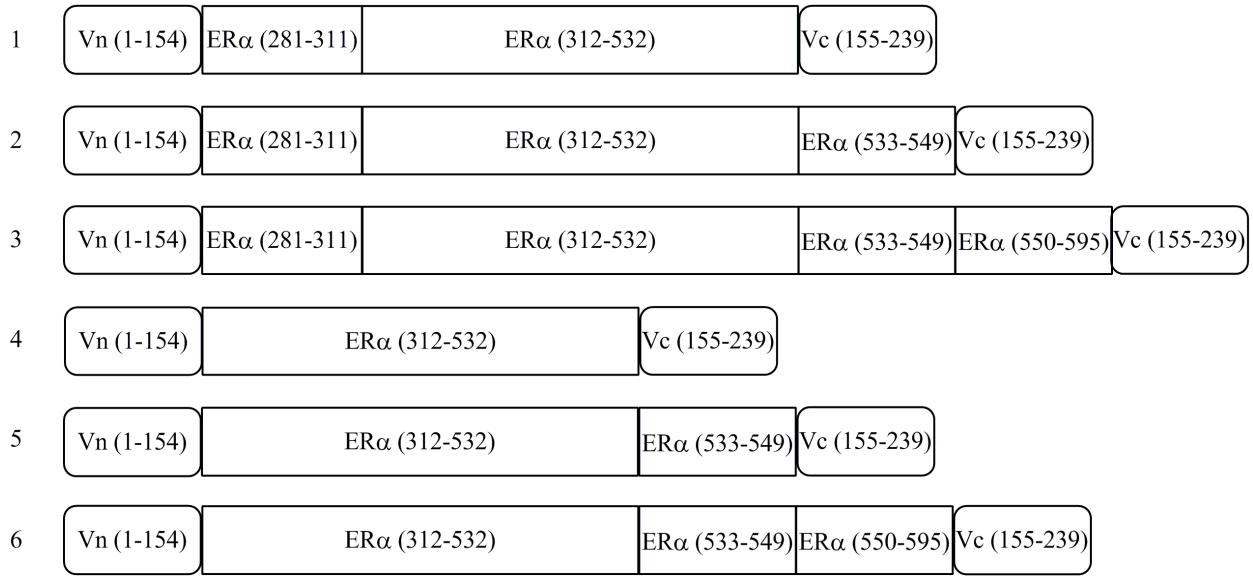


Figure 5.5. Sensor constructs 1 to 6. The sensors differ in the lengths of the estrogen receptor region.

5.2.2 Agonist and Antagonist Compound Evaluation

Transfection of the six constructs into HeLa cells followed by exposure to 100 nM of known agonists (17β-estradiol, diethylstilbestrol, estriol, or β-zearalanol), the antagonist ICI 182780, or the selective estrogen receptor modulators (SERMs, whose agonist or antagonist effect depends on the tissue type (57)), 4-hydroxytamoxifen, and raloxifene revealed that each construct had different properties (Figure 5.6). Constructs 1 and 4 both had low basal fluorescence in the absence of ligands and showed minimal induction in their presence. The commonality between these constructs is their ending at position 532, which cleaves off helix 12 that is responsible for the ligand-dependent conformational change that normally activates the receptor. Construct 3 also performed poorly, having a high basal fluorescence and only a slight increase upon the addition of ligand. Construct 6 had a low basal fluorescence and showed 9-fold induction with both agonists and antagonists. Constructs 2 and 5 seemed to exhibit some distinction between agonists and antagonists, a property that could be useful for pharmaceutical screening purposes. Construct 2 had high basal expression and a 15-30% decrease when in the presence of agonists, but an increase of 50-90% in the presence of antagonists. Construct 5 had a low basal signal with strong induction of 16-fold in the presence of the three antagonists, with a variable but lower induction to the agonists. The differences in basal fluorescence between the constructs could result from different spatial positioning of the two halves of mVenus, with closer arrangements allowing complementation in the absence of ligand that is enhanced or decreased by the conformational change upon agonist or antagonist binding. (Book 8 pg 26-43)

The properties of fluorescence complementation used in this system of sensor compare favorably to previously reported sensors. It has the benefit of requiring no exogenous substrates to produce a signal, as opposed to the luciferase complementation sensor (46). The signal generated is simply an increase in fluorescence (similar to Muddana and Peterson's sensor (42)) which is more robust and does not suffer from crosstalk that may occur with FRET assays due to absorption of energy by the acceptor fluorophore from the excitation laser instead of the donor fluorophore. The sensor is expressed in mammalian cells, as opposed to yeast or bacterial, meaning the transport of ligands into the cell is more appropriate.

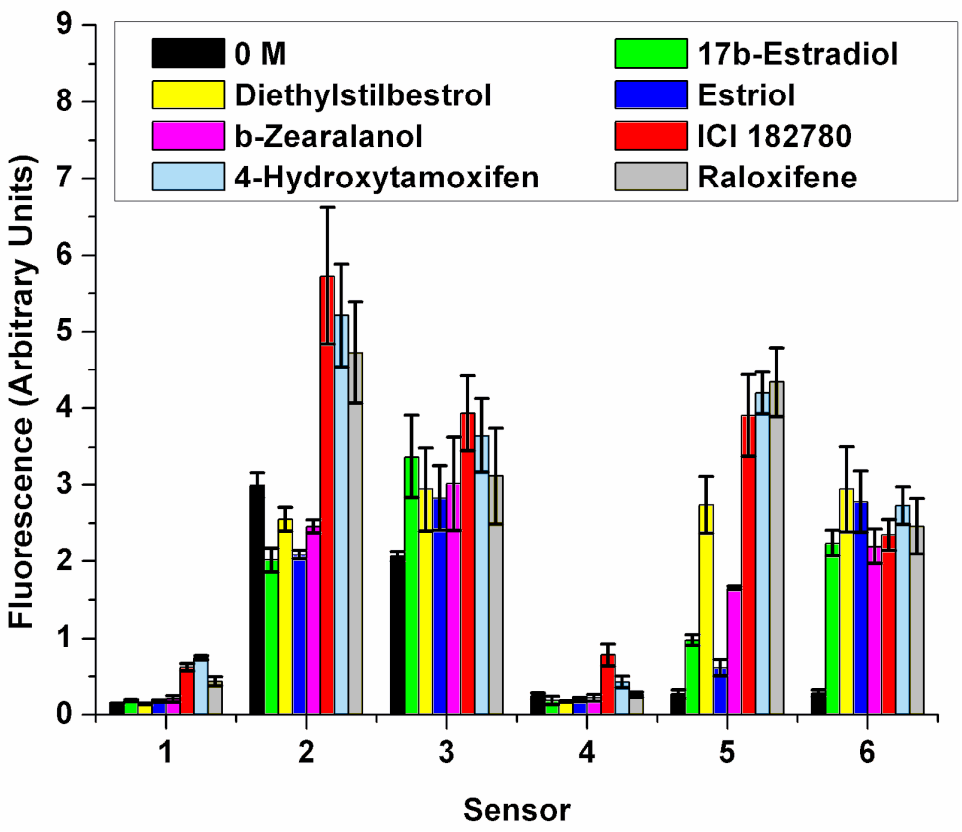


Figure 5.6. Sensor construct performance with agonists and antagonists. All constructs (1-6) were transiently transfected into HeLa cells and exposed to ethanol vehicle or 100 nM of the agonists E_2, diethylstilbestrol, estriol, or β-zearalanol, or the antagonist/ SERMs ICI 182780, 4-hydroxytamoxifen, or raloxifene. Data shown is the mean and standard error of the mean for two or more independent experiments.

5.2.3 Fluorescence Signal Generation Time Course

One of the important properties of an assay is the length of time it takes to give a positive signal. For this reason, the time course of signal generation for the fluorescence complementation biosensor was determined. Construct 6 was chosen for study as it showed the best induction properties. The sensor was transfected into HeLa cells and exposed to 100 nM 17β-estradiol for up to 24 h before the cells were fixed in formaldehyde for analysis by flow cytometry. The results show a 25 % ligand dependent induction within 1 h, 2 fold induction by 4 hours, and 9 fold induction with a 24 h incubation in the presence of ligand (Figure 5.7). (Book 8, pg 14-32)

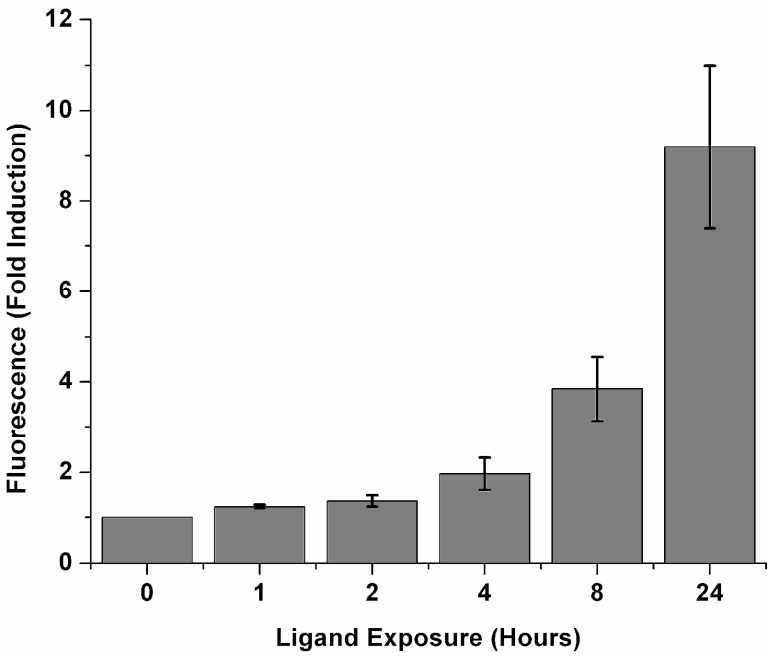


Figure 5.7. Time course of fluorescence signal generation of construct 6. HeLa cells were transiently transfected with construct 6 and exposed to 100 nM 17β-estradiol or ethanol vehicle for up to 24 hours, followed by formaldehyde fixation and flow cytometry. Data is the mean fold induction between ligand and no ligand conditions at each time point, with the standard error of the mean for three independent experiments.

The time for signal generation is intermediate compared to other reported biosensors. The fluorescence complementation approach has an inherent delay after ligand binding due to the length of time required for generation of the Venus fluorophore; however this is on the order of tens of minutes (54). The signal from a FRET sensor is very fast since the fluorophores are already formed, so the time taken to measure a change in emission ratio reflects only the time for ligand binding and the conformation change of helix 12 and can be detected within 4-20 minutes (44). Split luciferase systems in general are also capable of signal changes on the timescale of minutes (17), however the estrogen biosensor by Paulmurugan *et al* using this system (46) took between 6 and 12 hours to give a measurable signal, with maximal induction after 24 h which is similar to the sensor described here. The 24 hours required for maximal induction of the fluorescence and luciferase complementation signal suggest that these assays are not simply monitoring a conformation change. It is possible that the constructs are stabilized by ligand binding and the length of time reflects the growing accumulation of newly synthesized sensor. The fluorescence complementation assay is faster than growth based assays for estrogenic

compounds such as the E-SCREEN method which monitors growth of the human cell line MCF-7 after six days (40).

5.2.4 Vn-ER312-595-Vc Ligand Titration

As construct 6 was seen to perform well, it was decided to test its sensitivity to a number of different ligands. The ligand binding curves are shown in Figure 5.8, and the calculated concentration at half maximal induction (EC_{50}) is shown in Table 5.1. The EC_{50} values were plotted against known relative binding affinities (RBA) for the compounds (Carlson and Katzenellenbogen, unpublished data; and (3)) (Figure 5.9). (Book 8 pg 1-20)

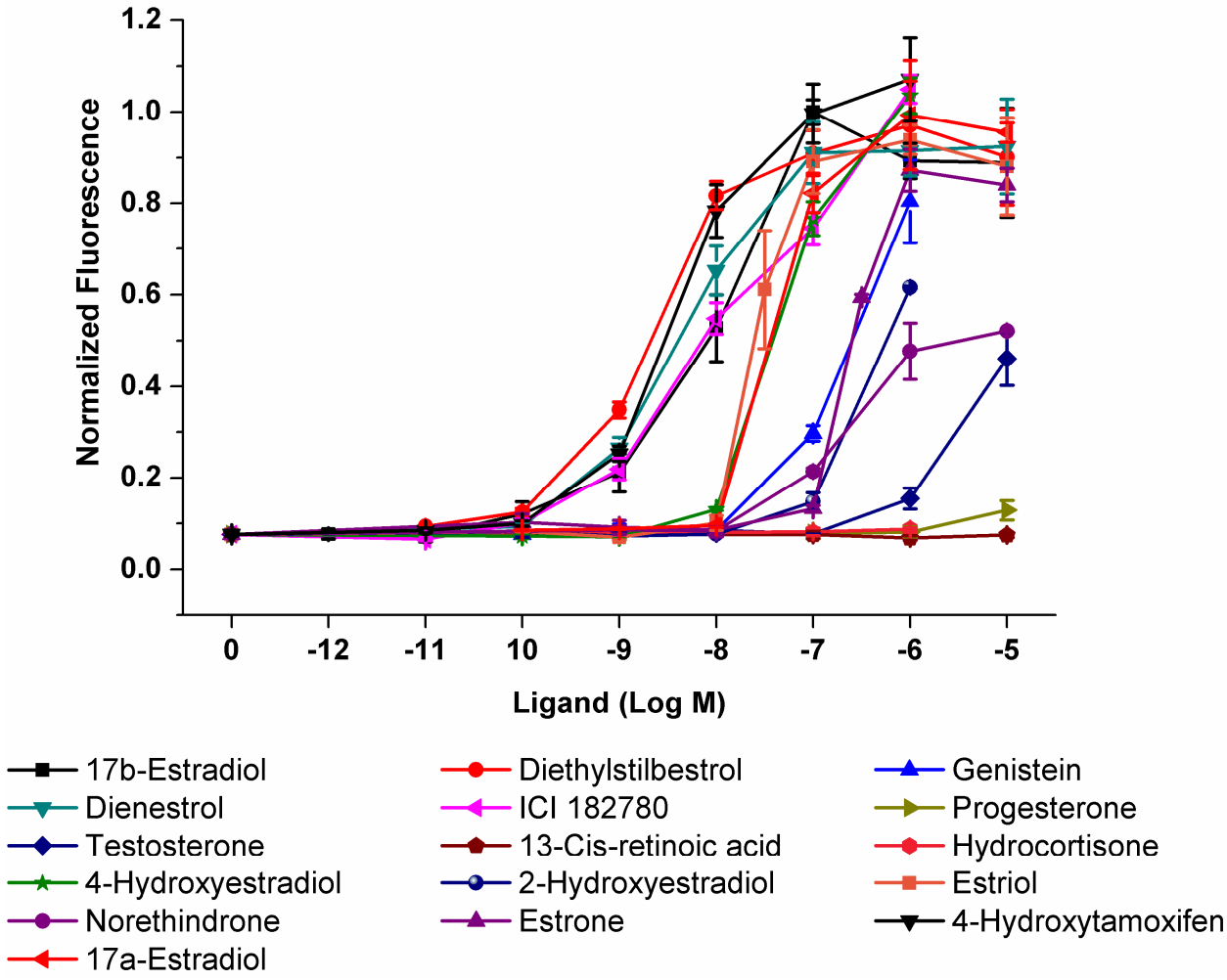


Figure 5.8. Ligand titration of compounds with construct 6.

Table 5.1. Concentration of ligand required to reach half maximal induction of construct 6 (nd = not determined).

Ligand	EC_{50} ± SEM (nM)
17β-Estradiol	8.1 ± 1.5
Diethylstilbestrol	2.2 ± 0.37
Genistein	220 ± 11
Dienestrol	4.4 ± 0.87
ICI182780	31 ± 13
Progesterone	nd
13-Cis retinoic acid	nd
Hydrocortisone	nd
Testosterone	14000 ± 3700
4-Hydroxyestradiol	58 ± 4.7
2-Hydroxyestradiol	550 ± 100
Estriol	36 ± 11
Norethindrone	880 ± 260
Estrone sulfate	230 ± 27
4-Hydroxytamoxifen	2.5 ± 0.73
17α-Estradiol	51 ± 1.7

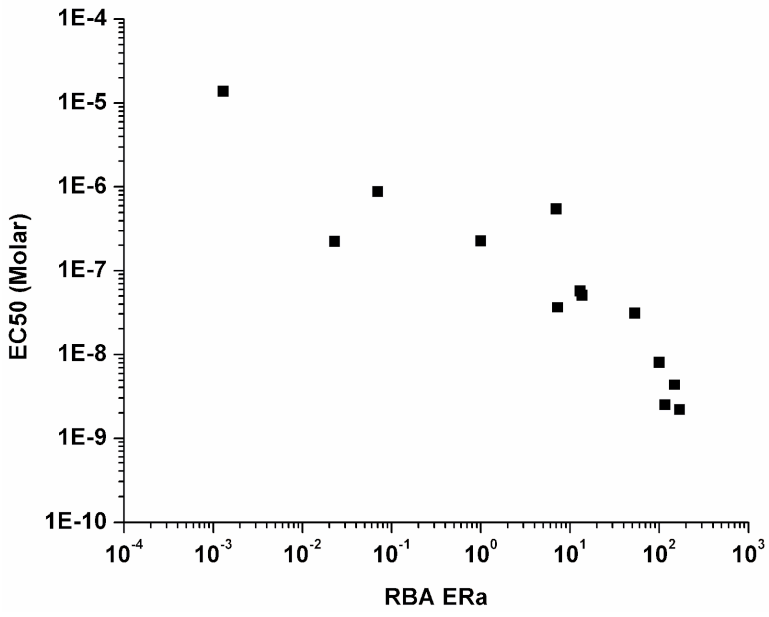


Figure 5.9. Scatter plot of RBA against EC_{50} determined by construct 6. The RBA of 17β-estradiol is set to be 100.

There was a reasonable correlation between the results obtained here and previously determined binding data, as can be seen in Figure 5.9. The sensitivity toward 17β-estradiol was lower than some measurements, although the value depends on the assay used ranging from 6 pM when

monitoring MCF-7 cell growth, up to nM values (41). The 8 nM EC50 was similar to that reported with FRET biosensors (43, 44). The results showed that binding of diethylstilbestrol, dienestrol, and 4-hydroxytamoxifen was higher than 17β-estradiol, as expected. Progesterone, 13-cis retinoic acid, and hydrocortisone were not expected to bind and indeed gave minimal response. Some low affinity ligands showed detectable binding, such as testosterone at micromolar and higher, and genistein at 100 nM and higher which was not detected by some sensors (46).

5.2.5 Detection of Exogenous Endocrine Active Substances

Having shown that construct 6 functioned well with a range of compounds, it was desired to test the binding of some environmental compounds. Figure 5.10 shows the results for compounds reported to be negative for estrogenic activity and also those expected to produce a response. Phenol, 1-naphthol, and 2-naphthol are commonly used in organic synthesis and the plastic industry, butylated hydroxytoluene is an antioxidant, while carbofuran and carbaryl are insecticides. All of these have organic ring structures but do not show estrogenic activity (40). When tested here at 0.01 mM, none showed an increase in fluorescence with construct 6; however the naphthols and carbaryl showed a small increase in fluorescence at 0.1 mM. Phytoestrogens are compounds derived from plants and can be consumed in amounts sufficient to produce plasma concentrations up to 10 μM (10). Daidzein is found in soy and is acted on by microorganisms in the gut to produce equol, both of which caused an increase in fluorescence at micromolar concentrations. Resveratrol is found in grapes and has been proposed to have many beneficial health effects but only weakly activated the sensor at 10 μM. The *Fusarium* fungus can grow on corn and other crops and produces zearalanone and its derivative β-zearalanol which is strongly estrogenic (5), as can be seen in Figure 5.10b. The nonionic detergent byproducts tert-octylphenol and nonylphenol were both detected at sub-micromolar levels, along with bisphenol A. The pesticide o,p'DDT and a PCB compound also displayed activation beginning at micromolar levels. (Book 8 pg 32-43)

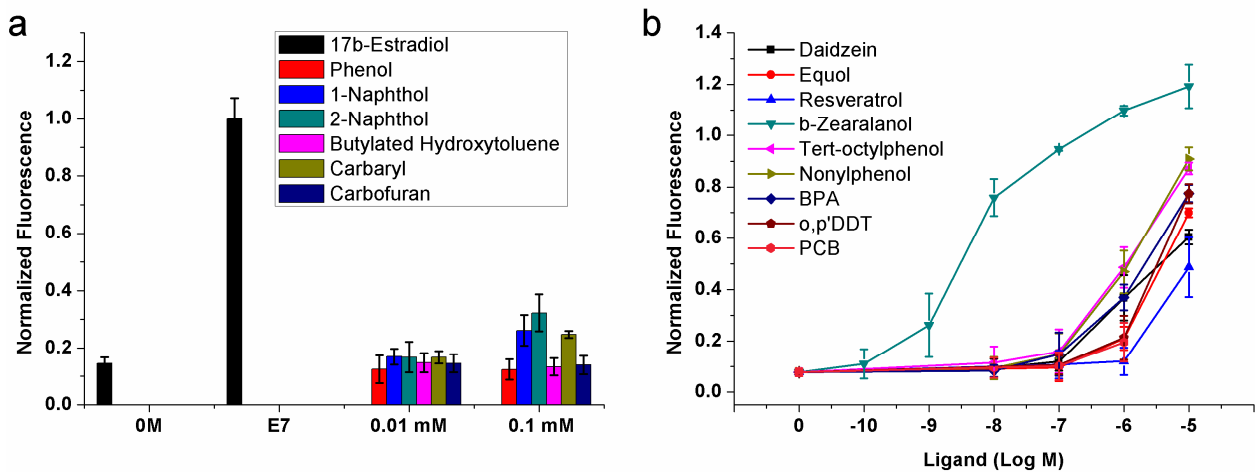


Figure 5.10. Detection of environmental chemicals using construct 6. (a) Compounds reported to not show estrogenic activity were tested at 0.01 mM and 0.1 mM, along with vehicle (0 M) and 100 nM estradiol (E7) as controls. (b) Compounds reported to show estrogenic activity: daidzein, equol, resveratrol, β-zearalanol, tert-octylphenol, nonylphenol, bisphenol A (BPA), op'DDT, and 2,4,6 trichloro 4-hydroxy PCB (PCB) were tested at concentrations up to 10 µM. Data was normalized to the 100 nM estradiol condition and is shown as the mean and standard error of the mean.

5.3 Conclusions and Outlook

The fluorescent complementation biosensor for estrogenic compounds described here proved to perform well. It was capable of detecting the binding of ligands whose affinity varied from nanomolar to tens of micromolar. Construct 6 could detect both agonists and antagonists from a range of sources, while construct 2 could distinguish antagonists from agonists. These properties could be useful for screening programs aiming to identify novel drugs targeting the ERα. Future directions could include further optimization of the construct in terms of linker regions and the best boundaries of the LBD region. For screening purposes, the sensor could be stably integrated into a human cell line such as HEK 293 to avoid the need for transfection.

5.4 Materials and Methods

5.4.1 Reagents

All DNA polymerases, restriction enzymes, and T4 DNA ligase were from New England Biolabs (Ipswich, MA). Cell media, fetal bovine serum, and charcoal dextran treated calf serum were obtained from the University of Illinois cell media facility (Urbana, IL). Opti-MEM media and lipofectamine 2000 were from Invitrogen (Carlsbad, CA). Carbamyl and carbofuran were from Chem Service (West Chester, PA), and raloxifene was from Tocris Bioscience (St Louis, MO). Other chemicals were from Sigma (St Louis, MO).

5.4.2 Cloning of Constructs

Unless otherwise stated, standard molecular biology techniques were used. A monomeric version of Venus fluorescent protein was constructed by mutagenesis of EYFP. Primers mVenus-Asmbl F1 (AACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAA GATCCGC), R2 (GGTAGTGGTCGGCGAGCTGCACGCCGCCGTCCTCGATGTTGTGGCG GATCTTGAAGTTGG), F3 (TCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACG GCCCCGTGCTGCTGCCCGACA), and R4 (CTTTGCTCAGCTTGGACTGGTAGCTCAGGT AGTGGTTGTCGGGCAGCAGCAC) were assembled using overlap extension and the product was used as a megaprimer on EYFP plasmid to extend to the C-terminus. This C-terminal fragment was used as a megaprimer with mVenus F64L-For (AAGCTGCCCGTGCCCTGGCCC ACCCTCGTGACCACCCTGGGCTAC) on EYFP. The N-terminal fragment was created by PCR with primers Cherry-For-KpnI (AAAAAAGGTACCATGGTGAGCAAGGGCGAGGAG) and mVenus F46L-Rev (CCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATCAGCTTCAG GGT), and was then used with the 64 C-terminal fragment to fill in the complete mVenus with EYFP as a template. The gene was digested with KpnI and BamHI and ligated into the vector pCMV5. The mVenus gene was split into fragments 1-154 and 155-239 by PCR of Vn with Cherry-For-KpnI and mVn154-Stop-Rev-BamHI (TTTTTTGGATCCTCAGGCGGTGATATA GACGTTGTGGCT), and PCR of Vc with EYFP155-For-KpnI (AAAAAAGGTACCATGGCC GACAAGCAGAAGAAC)and EGFP-Rev-BamHI (TTTTTTGGATCCTTACTTGTACAGCTC GTCCATGCCG), both of which were cloned into pCMV5. For cloning of the sensor constructs, the Vn-SalI-Vc-pCMV5 vector was constructed to insert the SalI site into mVenus-pCMV5 with Cherry-For-KpnI + mVenus154noStopRevSalI (AAAAAAGTCGACGGCGGTGATATAGAC

GTTGTGGCT); and mVenus155-SalI-For (AAAAAAGTCGACATGGACAAGCAGAAGAAC GGCATCAAGGCCAA) + EGFP-Rev-BamHI. The estrogen receptor α ligand binding domain regions were amplified by PCR with combinations of the forward primer ER281-For-SalIonly (AAAAAAGTCGACATGGGGTCTGCTGGAGACATGAGAGCT) or ER312-For-SalI (AAAA AAGTCGACGCCGACCAGATGGTCAGTGC), and reverse primer ER532-noStop-Rev-SalI (AAAAAAGTCGACGTTCTTGCACTTCATGCTGTACAGATG), ER549-noStop-Rev-SalI (A AAAAAGTCGACTAGGCGGTGGGCGTCCAGCAT), or ER595-noStop-Rev-SalI (AAAAAA GTCGACGACTGTGGCAGGGAAACCCTCT), and were cloned into the SalI site of Vn-SalI-Vc-pCMV5. See Table 5.2 for constructs used here.

Table 5.2. Constructs used in this work.

HZ	Genotype	Plasmid
2170	DH5alpha	mVenus pCMV5
2171	DH5alpha	Vn pCMV5
2172	DH5alpha	Vc pCMV5
3235	DH5alpha	Vn-ER281-532-Vc pCMV5
3236	DH5alpha	Vn-ER281-549-Vc pCMV5
3237	DH5alpha	Vn-ER281-595-Vc pCMV5
3238	DH5alpha	Vn-ER312-532-Vc pCMV5
3239	DH5alpha	Vn-ER312-549-Vc pCMV5
2243	DH5alpha	Vn-ER312-595-Vc pCMV5

5.4.3 Cell Culture and Transfection

HeLa cells were grown in minimal essential medium (MEM) / 1 mM sodium pyruvate / 10% fetal bovine serum (FBS) in a humidified incubator at 37 °C with 5% carbon dioxide. When cells were 80-90 % confluent, they were trypsinized and split into 12-well plates with 1 mL per well of MEM media / 1 mM sodium pyruvate / 5% charcoal dextran treated calf serum (CDCS). Cells were grown for 24 hours until they were over 90% confluent and transfected using lipofectamine 2000. Modifications were made to the standard protocol to minimize cell toxicity and optimize induction levels from sensor constructs. Typically, 1.5 μL lipofectamine 2000 was resuspended in 48.5 μL OptiMEM and incubated at room temperature for 10 min, then 50 ng of the sensor plasmid with 750 ng of the plasmid encoding β-galactosidase resuspended in 50 μL Opti-MEM was added and the mixture was incubated for 30 minutes and added to a well of HeLa cells. Cells were grown for 18 h after which the media was changed to MEM media / 1 mM sodium pyruvate / 5% CDCS plus the desired ligand. After a further 30 hours of growth, cells were analyzed by flow cytometry.

5.4.4 Flow Cytometry

Adherent cells were exposed to a solution of trypsin/EDTA for 30-60 s until they began to exhibit a rounded morphology. The trypsin/EDTA was removed by aspiration and the cells were incubated for another 1-2 min and resuspended in MEM + 5% CDCS. Samples were pelleted by centrifugation for 5 min at 800 ×*g*, media was removed by aspiration, and cells were resuspended in 300-400 µL PBS/5 mM EDTA. Cells were measured on a BD LSRII flow cytometer (Becton Dickinson, Franklin Lakes, NJ) using standard GFP/FITC filter sets. Analysis was performed using FCS Express 3 software (De Novo Software, Los Angeles, CA). Events were gated on the region corresponding to single whole cells and the mean fluorescence for 10000 cells was recorded.

5.4.5 Agonist and Antagonist Discrimination

Constructs 1-6 (ligand binding domain regions 281-532, 281-549, 281-595, 312-532, 312-549, 312-595) were transiently transfected by lipofection into HeLa cells. Samples were treated with no ligand, agonists (17β estradiol, diethylstilbestrol, estriol, or β-zearalanol), or antagonists / selective estrogen receptor modulators (ICI 182780, 4-hydroxytamoxifen, or raloxifene). Flow cytometry was performed and the average population fluorescence determined. Data was normalized by dividing by the sum of the six no ligand conditions.

5.4.6 Time Course of Fluorescence Signal Generation

Construct 6 (Vn-ER312-595-Vc) was transiently transfected by lipofection into HeLa cells. After 18 hours, the media was changed and 1 µL of 10^{-4} M 17β estradiol or ethanol was added to give a final ligand concentration of 10^{-7} M or 0 M. After 0, 1, 2, 4, 8, or 24 h of ligand exposure, cells were trypsinized, pelleted, and resuspended in 200 µl PBS and 1 mL of 3.6 % formaldehyde in PBS was added. Cells were fixed for 10 minutes, after which they were pelleted, resuspended in 300-400 µL PBS/5 mM EDTA, and stored at 4 °C until analysis by flow cytometry.

5.4.7 Ligand Titration

Construct 6 (Vn-ER312-595-Vc) was transiently transfected by lipofection into HeLa cells. After 18 hours, the media was changed and varying concentrations of the following ligands were

added: 17β-estradiol, diethylstilbestrol, genistein, dienestrol, ICI182780, progesterone, testosterone, 4-hydroxyestradiol, 2-hydroxyestradiol, estriol, norethindrone, estrone sulfate, 4-hydroxytamoxifen, and 17α-estradiol. Cells were analyzed by flow cytometry and the average population fluorescence was determined and normalized such that the 10^{-7} M 17β-estradiol value was equal to 1. Ligand titration curves were plotted using OriginPro 8 (OriginLab Corporation, Northampton, MA), which were fitted to determine the concentrations of half-maximal induction (EC_{50}). Relative binding affinities for the compounds were determined by Kathryn Carlson using recombinant ERα (Carlson and Katzenellenbogen, unpublished data).

5.4.8 Environmental Compound Detection

Construct 6 (Vn-ER312-595-Vc) was transiently transfected by lipofection into HeLa cells. After 18 hours, the media was changed and varying concentrations of the following ligands were added: daidzein, equol, resveratrol, β-zearalenol, p-tert-octylphenol, nonylphenol, bisphenol A, ortho-para'-DDT, 2,4,6 trichloro 4-hydroxy PCB, phenol, 1-naphthol, 2-naphthol, carbaryl, or carbofuran. After 30 hours, cells were analyzed by flow cytometry.

5.5 References

1. Nilsson,S. and Gustafsson,J.A. (2010) Estrogen receptors: Their actions and functional roles in health and disease. In Bunce,C.M. and Campbell,M.J. (eds.), *Nuclear Receptors.* Springer, Netherlands, Vol. 8, pp. 91-91-141.

2. Huang,P., Chandra,V. and Rastinejad,F. (2010) Structural overview of the nuclear receptor superfamily: Insights into physiology and therapeutics. *Annu. Rev. Physiol.,* **72,** 247-272.

3. Kuiper,G.G., Carlsson,B., Grandien,K., Enmark,E., Haggblad,J., Nilsson,S. and Gustafsson,J.A. (1997) Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta. *Endocrinology,* **138,** 863-870.

4. Kuiper,G.G., Lemmen,J.G., Carlsson,B., Corton,J.C., Safe,S.H., van der Saag,P.T., van der Burg,B. and Gustafsson,J.A. (1998) Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinology,* **139,** 4252-4263.

5. Katzenellenbogen,J.A. and Muthyala,R. (2003) Interactions of exogenous endocrine active substances with nuclear receptors. *Pure Appl. Chem.,* **75,** 1797-1817.

6. Anstead,G.M., Carlson,K.E. and Katzenellenbogen,J.A. (1997) The estradiol pharmacophore: Ligand structure-estrogen receptor binding affinity relationships and a model for the receptor binding site. *Steroids,* **62,** 268-303.

7. Pike,A.C. (2006) Lessons learnt from structural studies of the oestrogen receptor. *Best Pract. Res. Clin. Endocrinol. Metab.,* **20,** 1-14.

8. Gruber,C.J., Tschugguel,W., Schneeberger,C. and Huber,J.C. (2002) Production and actions of estrogens. *N. Engl. J. Med.,* **346,** 340-352.

9. Newbold,R.R., Padilla-Banks,E. and Jefferson,W.N. (2006) Adverse effects of the model environmental estrogen diethylstilbestrol are transmitted to subsequent generations. *Endocrinology,* **147,** S11-7.

10. Moutsatsou,P. (2007) The spectrum of phytoestrogens in nature: Our knowledge is expanding. *Hormones (Athens),* **6,** 173-193.

11. Renaud,J.P. and Delsuc,M.A. (2009) Biophysical techniques for ligand screening and drug design. *Curr. Opin. Pharmacol.,* **9,** 622-628.

12. East,A.K., Mauchline,T.H. and Poole,P.S. (2008) Biosensors for ligand detection. *Adv. Appl. Microbiol.,* **64,** 137-166.

13. Daunert,S., Barrett,G., Feliciano,J.S., Shetty,R.S., Shrestha,S. and Smith-Spencer,W. (2000) Genetically engineered whole-cell sensing systems: Coupling biological recognition with reporter genes. *Chem. Rev.,* **100,** 2705-2738.

14. Frommer,W.B., Davidson,M.W. and Campbell,R.E. (2009) Genetically encoded biosensors based on engineered fluorescent proteins. *Chem. Soc. Rev.,* **38,** 2833-2841.

15. Altschuh,D., Oncul,S. and Demchenko,A.P. (2006) Fluorescence sensing of intermolecular interactions and development of direct molecular biosensors. *J. Mol. Recognit.,* **19,** 459-477.

16. Binkowski,B., Fan,F. and Wood,K. (2009) Engineered luciferases for molecular sensing in living cells. *Curr. Opin. Biotechnol.,* **20,** 14-18.

17. Fan,F., Binkowski,B.F., Butler,B.L., Stecha,P.F., Lewis,M.K. and Wood,K.V. (2008) Novel genetically encoded biosensors using firefly luciferase. *ACS Chem. Biol.,* **3,** 346-351.

18. Ciruela,F. (2008) Fluorescence-based methods in the study of protein-protein interactions in living cells. *Curr. Opin. Biotechnol.,* **19,** 338-343.

19. Xu,Y., Piston,D.W. and Johnson,C.H. (1999) A bioluminescence resonance energy transfer (BRET) system: Application to interacting circadian clock proteins. *Proc. Natl. Acad. Sci. U. S. A.,* **96,** 151-156.

20. Morell,M., Ventura,S. and Aviles,F.X. (2009) Protein complementation assays: Approaches for the in vivo analysis of protein interactions. *FEBS Lett.,* **583,** 1684-1691.

21. Hu,C.D., Chinenov,Y. and Kerppola,T.K. (2002) Visualization of interactions among bZIP and rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell,* **9,** 789-798.

22. Kerppola,T.K. (2008) Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu. Rev. Biophys.,* **37,** 465-487.

23. Gillies,A.R., Skretas,G. and Wood,D.W. (2008) Engineered systems for detection and discovery of nuclear hormone-like compounds. *Biotechnol. Prog.,* **24,** 8-16.

24. Zacharewski,T. (1997) In vitro bioassays for assessing estrogenic substances. *Environ. Sci. Technol.,* **31,** 613-623.

25. Bovee,T.F., Helsdingen,R.J., Koks,P.D., Kuiper,H.A., Hoogenboom,R.L. and Keijer,J. (2004) Development of a rapid yeast estrogen bioassay, based on the expression of green fluorescent protein. *Gene,* **325,** 187-200.

26. Sievernich,A., Wildt,L. and Lichtenberg-Frate,H. (2004) In vitro bioactivity of 17alpha-estradiol. *J. Steroid Biochem. Mol. Biol.,* **92,** 455-463.

27. Hasenbrink,G., Sievernich,A., Wildt,L., Ludwig,J. and Lichtenberg-Frate,H. (2006) Estrogenic effects of natural and synthetic compounds including tibolone assessed in saccharomyces cerevisiae expressing the human estrogen alpha and beta receptors. *FASEB J.,* **20,** 1552-1554.

28. Routledge,E.J. and Sumpter,J.P. (1996) Estrogenic activity of surfactants and some of their degradation products assessed using a recombinant yeast screen. *Environ. Toxicol. Chem.,* **15,** 241-248.

29. Gao,C.Y. and Pinkham,J.L. (2000) Tightly regulated, beta-estradiol dose-dependent expression system for yeast. *BioTechniques,* **29,** 1226-1231.

30. Fu,K.Y., Chen,C.Y. and Chang,W.M. (2007) Application of a yeast estrogen screen in non-biomarker species varicorhinus barbatulus fish with two estrogen receptor subtypes to assess xenoestrogens. *Toxicol. in. Vitro.,* **21,** 604-612.

31. Lee,H.S., Sasagawa,S., Kato,S., Fukuda,R., Horiuchi,H. and Ohta,A. (2006) Yeast two-hybrid detection systems that are highly sensitive to a certain kind of endocrine disruptors. *Biosci. Biotechnol. Biochem.,* **70,** 521-524.

32. Coldham,N.G., Dave,M., Sivapathasundaram,S., McDonnell,D.P., Connor,C. and Sauer,M.J. (1997) Evaluation of a recombinant yeast cell estrogen screening assay. *Environ. Health Perspect.,* **105,** 734-742.

33. Wilkinson,J.M., Hayes,S., Thompson,D., Whitney,P. and Bi,K. (2008) Compound profiling using a panel of steroid hormone receptor cell-based assays. *J. Biomol. Screen.,* **13,** 755-765.

34. Sanseverino,J., Gupta,R.K., Layton,A.C., Patterson,S.S., Ripp,S.A., Saidak,L., Simpson,M.L., Schultz,T.W. and Sayler,G.S. (2005) Use of saccharomyces cerevisiae BLYES expressing bacterial bioluminescence for rapid, sensitive detection of estrogenic compounds. *Appl. Environ. Microbiol.,* **71,** 4455-4460.

35. Chen,Z. and Zhao,H. (2003) A highly efficient and sensitive screening method for trans-activation activity of estrogen receptors. *Gene,* **306,** 127-134.

36. Clark,J.H., Anderson,J.N. and Peck,E.J.,Jr. (1975) (3H) estradiol exchange assay for the determination of nuclear receptor-estrogen complex. *Methods Enzymol.,* **36,** 283-286.

37. Hwang,K.J., Carlson,K.E., Anstead,G.M. and Katzenellenbogen,J.A. (1992) Donor-acceptor tetrahydrochrysenes, inherently fluorescent, high-affinity ligands for the estrogen receptor: Binding and fluorescence characteristics and fluorometric assay of receptor. *Biochemistry,* **31,** 11536-11545.

38. Bolger,R., Wiese,T.E., Ervin,K., Nestich,S. and Checovich,W. (1998) Rapid screening of environmental chemicals for estrogen receptor binding capacity. *Environ. Health Perspect.,* **106,** 551-557.

39. Boehme,K., Simon,S. and Mueller,S.O. (2009) Gene expression profiling in ishikawa cells: A fingerprint for estrogen active compounds. *Toxicol. Appl. Pharmacol.,* **236,** 85-96.

40. Soto,A.M., Sonnenschein,C., Chung,K.L., Fernandez,M.F., Olea,N. and Serrano,F.O. (1995) The E-SCREEN assay as a tool to identify estrogens: An update on estrogenic environmental pollutants. *Environ. Health Perspect.,* **103 Suppl 7,** 113-122.

41. Olsen,C.M., Meussen-Elholm,E.T., Hongslo,J.K., Stenersen,J. and Tollefsen,K.E. (2005) Estrogenic effects of environmental chemicals: An interspecies comparison. *Comp. Biochem. Physiol. C. Toxicol. Pharmacol.,* **141,** 267-274.

42. Muddana,S.S. and Peterson,B.R. (2003) Fluorescent cellular sensors of steroid receptor ligands. *Chembiochem,* **4,** 848-855.

43. De,S., Macara,I.G. and Lannigan,D.A. (2005) Novel biosensors for the detection of estrogen receptor ligands. *J. Steroid Biochem. Mol. Biol.,* **96,** 235-244.

44. Umezawa,Y. (2005) Genetically encoded optical probes for imaging cellular signaling pathways. *Biosens. Bioelectron.,* **20,** 2504-2511.

45. Padron,A., Li,L., Kofoed,E.M. and Schaufele,F. (2007) Ligand-selective interdomain conformations of estrogen receptor-alpha. *Mol. Endocrinol.,* **21,** 49-61.

46. Paulmurugan,R. and Gambhir,S.S. (2006) An intramolecular folding sensor for imaging estrogen receptor-ligand interactions. *Proc. Natl. Acad. Sci. U. S. A.,* **103,** 15883-15888.

47. Paulmurugan,R., Tamrazi,A., Katzenellenbogen,J.A., Katzenellenbogen,B.S. and Gambhir,S.S. (2008) A human estrogen receptor (ER)alpha mutation with differential responsiveness to nonsteroidal ligands: Novel approaches for studying mechanism of ER action. *Mol. Endocrinol.,* **22,** 1552-1564.

48. Kim,S.B., Umezawa,Y., Kanno,K.A. and Tao,H. (2008) An integrated-molecule-format multicolor probe for monitoring multiple activities of a bioactive small molecule. *ACS Chem. Biol.,* **3,** 359-372.

49. Kim,S.B., Sato,M. and Tao,H. (2008) Circularly permutated bioluminescent probes for illuminating ligand-activated protein dynamics. *Bioconjug. Chem.,* **19,** 2480-2486.

50. Skretas,G. and Wood,D.W. (2005) Regulation of protein activity with small-molecule-controlled inteins. *Protein Sci.,* **14,** 523-532.

51. Skretas,G. and Wood,D.W. (2005) A bacterial biosensor of endocrine modulators. *J. Mol. Biol.,* **349,** 464-474.

52. Skretas,G. and Wood,D.W. (2005) Rapid detection of subtype-selective nuclear hormone receptor binding with bacterial genetic selection. *Appl. Environ. Microbiol.,* **71,** 8995-8997.

53. Skretas,G., Meligova,A.K., Villalonga-Barber,C., Mitsiou,D.J., Alexis,M.N., Micha-Screttas,M., Steele,B.R., Screttas,C.G. and Wood,D.W. (2007) Engineered chimeric enzymes as tools for drug discovery: Generating reliable bacterial screens for the detection, discovery, and assessment of estrogen receptor modulators. *J. Am. Chem. Soc.,* **129,** 8443-8457.

54. Nagai,T., Ibata,K., Park,E.S., Kubota,M., Mikoshiba,K. and Miyawaki,A. (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.,* **20,** 87-90.

55. Shyu,Y.J., Liu,H., Deng,X. and Hu,C.D. (2006) Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions. *BioTechniques,* **40,** 61-66.

56. Zacharias,D.A., Violin,J.D., Newton,A.C. and Tsien,R.Y. (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science,* **296,** 913-916.

57. Robertson,J.F. (2004) Selective oestrogen receptor modulators/new antioestrogens: A clinical perspective. *Cancer Treat. Rev.,* **30,** 695-706.

Author's Biography

Michael McLachlan was born and raised in Christchurch, New Zealand. Following graduation from Shirley Boys' High School he enrolled at Victoria University of Wellington. He completed his Bachelor's degree in Genetics and Molecular Biology in 1995, and his BSc Honours (First Class) in 1996 with a thesis on repetitive elements in the housefly genome. He worked with Dr Helen Davey at AgResearch in Hamilton, New Zealand studying STAT5B knockout mice and other transgenic mouse models. In 2004 he enrolled in the Biophysics and Computational Biology program at the University of Illinois at Urbana-Champaign. He performed his doctoral research in protein engineering in the laboratory of Professor Huimin Zhao. Following completion of his doctoral degree he will join Cellular Dynamics International in Madison, Wisconsin.