2022

EXPRESSION OF MOUSE FULL-LENGTH ARYL HYDROCARBON RECEPTOR AND HUMAN ARYL HYDROCARBON RECEPTOR LIGAND BINDING DOMAIN IN PICHIA PASTORIS

Yiyuan Wang
University of the Pacific

EXPRESSION OF MOUSE FULL-LENGTH ARYL HYDROCARBON RECEPTOR AND HUMAN ARYL HYDROCARBON RECEPTOR LIGAND BINDING DOMAIN IN PICHIA PASTORIS

By

Yiyuan Wang

A Dissertation Submitted to the

Graduate School

In Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

Thomas J. Long School of Pharmacy
Pharmaceutical and Chemical Sciences

University of the Pacific
Stockton, California

2022

EXPRESSION OF MOUSE FULL-LENGTH ARYL HYDROCARBON RECEPTOR AND HUMAN ARYL HYDROCARBON RECEPTOR LIGAND BINDING DOMAIN IN PICHIA PASTORIS

By

Yiyuan Wang

APPROVED BY:

Dissertation Advisor: William K. Chan, Pharm.D., Ph.D.

Committee Member: Miki Susanto Park, Ph.D.

Committee Member: Geoffrey Lin-Cereghino, Ph.D.

Department Chair: William K. Chan, Pharm.D., Ph.D.

EXPRESSION OF MOUSE FULL-LENGTH ARYL HYDROCARBON RECEPTOR AND HUMAN ARYL HYDROCARBON RECEPTOR LIGAND BINDING DOMAIN IN PICHIA PASTORIS

ABSTRACT

By Yiyuan Wang

University of the Pacific
2022

Aryl hydrocarbon receptor (AHR) is a ligand-activated transcription factor that regulates biological responses to planar aromatic hydrocarbon. AHR activates gene transcription by binding to its corresponding enhancer with its partner-aryl hydrocarbon receptor nuclear translocator (ARNT). In addition, this receptor has been shown to regulate xenobiotic-metabolizing enzymes such as cytochrome P450. AHR exists widely in body tissues and affects bioactivation of carcinogenic compounds, T cell differentiation, fatty acid synthesis, cell proliferation, hematopoietic stem cell differentiation, respiratory reactivity, and insulin sensitivity. Although the precise mechanism illustrating the endogenous AHR function remains unclear, there has been intense interest in exploring AHR as a potential target for the treatment of diseases such as cancer and autoimmune diseases.

It is known that mouse ahr d-allele possesses low ligand-binding affinity, whereas mouse *ahr b*-allele has a higher ligand-binding affinity. The *d*-allele functions more similarly to human AHR than the *b*-allele, which is most commonly studied. Human AHR can be rather difficult to study since it is relatively unstable and less sensitive to some ligands *in vitro*. Thus we generated a deletion construct which has the ligand-binding domain of human AHR and hoped that the expression yield could be increased.

Here, I present the process and the results of expressing the mouse full-length *b*-allele of AHR and the human AHR ligand binding domain (LBD, amino acids 108 to 400) in *Pichia pastoris*. A higher enrichment of the *b*-allele and LBD was observed in wild-type yeast (yJC100) strain when compared to the protease-deficient yeast (ySMD1163) strain. This observation was consistent with the increased copy number in the wild-type strain. Although the LBD transcript was detected in both the wild-type and protease-deficient strains, the LBD protein was only detected in the wild-type strain.

ACKNOWLEDGEMENTS

My gratitude goes to Dr. William Chan for his hours of patience and guidance while I navigated graduate school. He is a very responsible advisor and principal investigator. In addition, I thank him for his encouragement, patience, and guidance during my study. His effort and positive attitude will continue to influence me.

I would also like to thank Dr. Park, Dr. Lin-Cereghino for being my committee members. Their suggestions are very helpful and appreciated. Finally, I thank my friends and my family for helping me and keeping me accompanied through this time.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

AHR	aryl hydrocarbon receptor
ahrr	aryl hydrocarbon receptor repressor
AMPs	antimicrobial peptides
AOX	alcohol oxidase
ARA9	aryl hydrocarbon receptor-associated 9
ARNT	aryl hydrocarbon receptor nuclear translocator
bHLH	basic-helix-loop-helix
BMGY	Buffered glycerol-complex medium
BMMY	Buffered methanol-complex medium
bp	base pairs
CD4	cluster of differentiation 4
CYP1A1	cytochrome P450 1A1
CYP1B1	cytochrome P450 1B1
CYP450	cytochrome P450 enzymes
DCs	dendritic cells
DTT	dithiothreitol
DRE	dioxin response element
E1	elute 1
FT	flow-through
hAHR	human aryl hydrocarbon receptor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HIF-2R	hypoxia-inducible factor 2R
Hsp90	heat shock protein 90
IECs	intestinal epithelial cells
IELs	intraepithelial lymphocytes
IgG	immunoglobulin G
IL	interleukin
ILC3	Group 3 innate lymphoid cells
LBD	ligand binding domain
mAHR	mouse aryl hydrocarbon receptor
MBP	Maltose-binding protein
MOI	multiplicity of infection
mRNA	messenger RNA
p23	an HSP90-interacting protein
PAHs	polycyclic aromatic hydrocarbons
PAS	Per-ARNT-Sim
Q-rich	glutamine-rich
TCDD	2, 3, 7, 8-tetrachlorodibenzo-p-dioxin
Th	T-helper
Tregs	regulatory T cells
XRE	xenobiotics response elements
YNB	Yeast Nitrogen Base
YPD	Yeast extract Peptone Dextrose

CHAPTER 1: INTRODUCTION

1.1 The Aryl Hydrocarbon Receptor (AHR)

Aryl hydrocarbon receptor (AHR) is a protein transcribed from *ahr* gene, which is included in biological regulatory responses to planar aromatic hydrocarbons by specific ligands. Since it is a ligand-activated transcription factor, it activates the transcription of genes by binding to the corresponding heterodimer, which is composed of AHR and aryl hydrocarbon receptor nuclear translocator (ARNT) composition.[1,2]

In the canonical signaling pathway (XRE), the activation of AHR is due to the binding of exogenous or endogenous ligands, resulting in nuclear translocation and coordinated recruitment with ARNT protein to XRE induction. Without ligand binding, the AHR stayed in the cytoplasm and bounded with the chaperone proteins like HSP90, p23, and ARA9. After binding and activating with the corresponding ligand, it is speculated that AHR undergoes a conformational change in the PAS-A domain, which will help the nuclear translocation, chaperone protein dissociation, and the dimerization of AHR and ARNT through the PAS-A region.[3-5,7] Once it enters the nucleus, AHR dissociates from chaperone proteins and binds to ARNT to form a heterodimer protein complex. At the same time, the ligand-induced conformational changes in the PAS-A region will contribute to the AHR and ARNT dimerization.[6,7] The whole process of canonical signaling path is shown in Figure 1. The AHR signal will be inhibited by one of the following mechanisms: the first was the induction of CYP1A1 and the clearance of ligand metabolism; the second was the expression of AHR inhibitor (ahrr); the last was the degradation of AHR by 26S proteasome. Therefore, in many studies, we can study the mechanism of AHR by inhibiting the signal pathway of AHR.

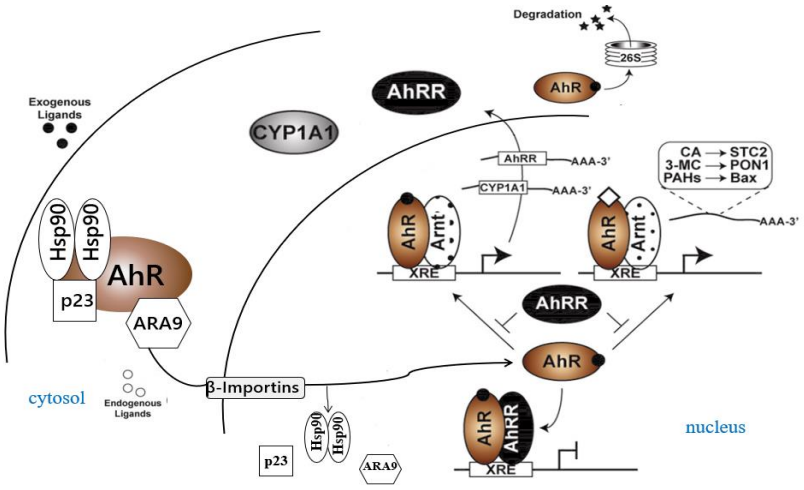

Figure 1. The Canonical Signaling Pathway (XRE) of AHR.

AHR is widely in various tissues of the human body. As a ligand-activated transcription factor, it is involved in many crucial exogenous cytochrome P450 enzymes. Also, it participates in many biological processes in the body, such as its biological activation of carcinogenic compounds, T cell differentiation, fatty acid synthesis, cell proliferation respiratory effects.[7-9] Researchers found that it not only mediates the toxicity of various environmental toxicants but also plays a vital role in cell proliferation, differentiation, and human immune regulation.[8-11] Although the exact mechanism of endogenous AHR function is not entirely understood, researchers have always been interested in using AHR as a potential target for the treatment of diseases such as cancer and autoimmune diseases.

AHR is involved in various essential cell cycle regulation and homeostasis signal transduction pathways so that AHR can be acted as a molecular target for cancer therapy. If combined with CYP1A1, it plays a role in the detoxification of environmental carcinogenesis. It prevents metabolic activation of mixtures of diets containing cancer activity, cancer progression,

or prevention based on the balance of activation/detoxification of pro-carcinogens and extrahepatic metabolism of natural dietary products.[13] Increased AHR expression in lung adenoma cells can increase intracellular oxidative stress and promote cell growth, suggesting that interfering with AHR expression may prevent the early development of lung adenomas.[14, 16] TCDD and a non-toxic AHR agonist induce the expression of the cell cycle regulation-dependent kinase inhibitor p21, causing cell growth inhibition in pancreatic cancer with high AHR expression, suggesting that increasing AHR expression may provide a therapeutic target for pancreatic cancer.[15]

The expression of AHR in most types of immune cells and the promoter regions of many genes regulating immune responses demonstrates the importance of this receptor in the immune process.[17] For example, tryptophan metabolism by microorganisms in the gut produces AHR agonists to support the development and maintenance of type 3 innate lymphocytes (ILC3) in the gut, while ILC3 as the innate immune system corresponds to the CD4 T of the adaptive immune system cells, the latter mainly produce IL-17 and IL-22. AHR signaling is also required for the maintenance of IL-22-producing intraepithelial lymphocytes (IELs), which are involved in mucosal wound healing and antimicrobial peptides (AMPs) production in intestinal epithelial cells (IECs).[19] The AhR-IL-22 signaling pathway in the gut plays a vital role in host defense against microbial pathogens and promotes disease tolerance to limit deleterious effects. Weak activation of AhR supports a pro-inflammatory response (Th17/22), whereas strong activation of AhR promotes the induction of tolerogenic dendritic cells (DCs) and regulatory T cells (Tregs).[18-21]

1.2 Functional Domain of Human AHR

AHR protein is a member of the basic helix-loop-helix/Per-Arnt-Sim (bHLH/PAS) family of transcription factors and contains multiple domains with important functions.[7,12] As shown in Figure 2, the common entity bHLH domain of multiple transcription factors are located at the N-terminal of the protein, which has two domains with unique functions and is highly conserved among the family. One is the basic region (b), which is involved in the binding of transcription factors and DNA. The second is the helix-loop-helix (HLH) region, which promotes protein-protein interaction.[22] The next PAS domain contains PAS-A and PAS-B, which are fragments of 200-350 amino acids and have high sequence similarity with PAS protein. In addition, the PAS domain supports specific secondary interactions with other proteins containing the PAS domain, such as AHR and ARNT, thus forming dimer and heteropolymer protein complexes.[23] The ligand-binding site of AHR is resided in the PAS-B domain and contains several conserved residues, which are essential for ligand binding. Finally, the glutamine-rich (Q-rich) domain is located in the C-terminal region of the protein and is involved in the recruitment and transactivation of coactivators.[24,25] We are interested in the properties of the AhR domain, especially the ligand binding domain (LBD), as shown in Figure 2.

AHR can bind and be activated by a structurally diverse array of both synthetic and natural compounds through the ligand binding domain. However, resolving the full-length three-dimensional structure of AhR has always been an arduous task. Despite the efforts of scientists to achieve this goal, the entire AhR protein does not yet have a complete structure. Until 2013, the first X-ray diffraction based AHR 3D structure of the PAS-A homodimer structure obtained from recombinant *E. coli* expression was reported for the mouse PAS-A domain (residues 110 to 267); the resolution is 2.55 Å (PDB ID: 4M4X).[26] Another two other AhR structures were

reported in 2017, containing multiple AhR domains and showing distinct interactions between AhR and its dimerization partner ARNT and its interaction with both DNA strands. The two structures reveal the complex formation between the bHLH and PAS-A domains of human AhR and their interactions with ARNT and DNA.[27-29] However, neither of these two subdomains is included in this architecture due to the observed high flexibility of the AhR PAS-B domain and transaction domain (C-side). Therefore, people usually developed LBD homology modeling since there aren't enough experimentally determined structures under the molecular level. With the availability of extensive structural information on homologous PAS-containing proteins, a reliable model of the mouse AhR PAS-B domain was developed by comparative modeling techniques.[30] The PAS domain structures of the functionally related hypoxia-inducible factor 2R (HIF-2R) and AhR nuclear transporter (ARNT) proteins, which share the highest degree of sequence identity and similarity with AhR, usually were selected as templates for the development of two Models. To confirm the characteristics of the modeling domains, the effect of point mutations at selected residue positions on TCDD binding to AhR and TCDD-dependent transformation and DNA binding was analyzed. Mutagenesis and functional analysis results are consistent with the proposed model and confirm that the cavity modeled inside the domain is involved in ligand binding. In addition, the physicochemical characterization of some residues and their mutants and the effect of mutagenesis on TCDD and DNA binding have also shown some critical features required for ligand binding and mAHR activation at the molecular level, thus enabling deep learning.[31-33] But more reliable information on the structural characteristics of the AhR LBD will only be obtained when the structures of homologous proteins with higher sequence identity to AhR become available.

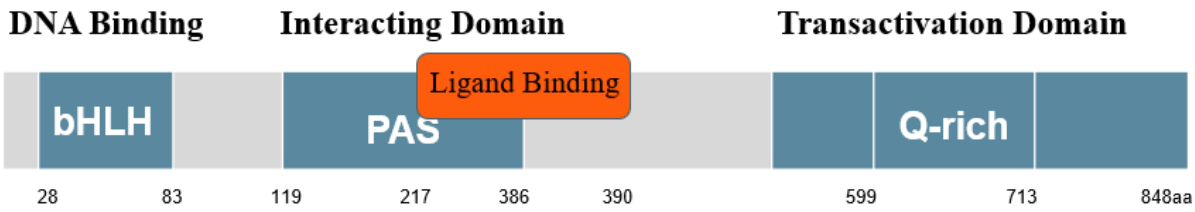

Figure 2. Human AHR functional domain

1.3 Similarity of Mouse AHR and Human AHR

As shown in Figure 3, the amino acid sequence of mouse AHR *b*-allele and human AHR has a high similarity. The amino acid sequence similarity is as high as 89% for the ligand binding domains. However, there are significant differences in dioxin reactivity among different species with structurally distinct AHRs. For example, studies have shown at least a 1000-fold difference in TCDD tolerance between the more sensitive guinea pig and tolerant hamster species. Since an alanine residue in the ligand binding domain (A375 corresponding to the V381 of hAHR), the mouse AHR *b*-allele ($mAHR^b$) has a ten-fold higher affinity for AHR ligands (TCDD) when it compared with the human AHR (hAHR). However, the affinity of mouse AHR *d*-allele for AHR ligand is similar to human AHR.[34,35]

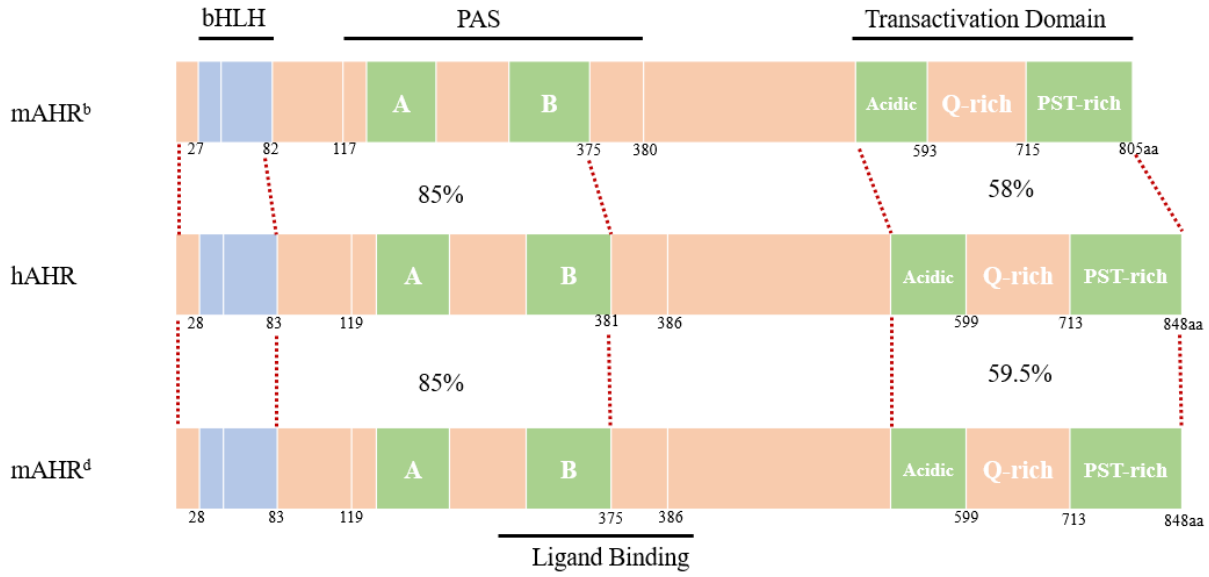

Figure 3. Comparison of human and mouse AHR protein structure[34]

Therefore, a truncated protein hAHR aa108-400 will be generated for studying the ligand binding domain of hAHR, which contains the PAS-B domain of human AHR and several conserved amino acid residues.

1.4 Different Expression Systems of AHR

Information for protein synthesis is stored in DNA and decoded through the process of transcription to produce messenger RNA (mRNA). The information encoded by the mRNA is then translated into protein. Transcription refers to the transfer of information from DNA to mRNA, while translation is the synthesis of proteins based on the sequences specified by the mRNA. At the same time, protein synthesis and regulation depend on functional requirements in the cell. Researchers often need to construct functional proteins of interest to study how specific proteins regulate biology. Given the size and complexity of proteins, chemical synthesis is not viable for this work. Also, living cells and their cellular machinery are often used as factories to

build and construct proteins based on genetic templates provided. DNA is easily synthesized or constructed in vitro using established recombinant DNA technologies compared with proteins. Thus, gene-specific DNA templates can be constructed to serve as templates for protein expression, with or without an additional reporter or affinity tag sequences. Proteins produced from this DNA template are called recombinant proteins. Strategies for recombinant protein expression include transfecting cells with a template-containing DNA vector and then culturing the cells so that they transcribe and translate the desired protein. The cells were then lysed to extract the expressed protein for subsequent purification. The choice of system depends on the type of protein, requirements for functional activity and desired yield. These expression systems are summarized below and include mammalian, insect, yeast, bacterial, etc. expression systems. Each system has advantages and challenges, choosing the right system for a specific application is vital for the successful expression of recombinant proteins.

According to different expression systems, the expression products of AHR are also different. Protein expression products are growing rapidly for expression systems such as *E. coli* and bacteria. However, their applications are limited by the inability to fold correctly and perform proper post-translational modification of mammalian proteins. However, due to its low expression cost, simplicity and convenience, short growth cycle, and high yield, it is still one of the essential expression systems of AHR. More than twenty years ago, our laboratory overexpressed the AHR protein in the baculovirus expression system. As a result, it expresses relatively high-yield recombinant proteins, and the expressed AHR is functional, which can bind to ligands, form dimers, and activate DRE signaling pathways. However, since baculovirus must be produced regularly through the low multiplicity of infection (MOI) of insect cells, cloned high-titer virus strains may cause baculovirus genome mutations, thereby reducing the yield of

recombinant proteins and possibly changing the nature of the recombinant proteins.[36,37] If we want long-term stable expression, it is clear that the baculovirus expression system does not provide stable and complete support. For mammalian expression systems, AHR can stabilize long-term expression and has post-translational modifications of proteins. However, due to its high cost, low protein yield, and long cultivation cycle. Therefore, it does not apply to my experimental project. Next, *Pichia pastoris* expression system will be introduced in detail and it also will be used in expressing AHR recombinant protein.

1.5 *Pichia pastoris* Expression System

Pichia pastoris is a eukaryote. It has the ability to modify post-translational recombinant proteins, such as producing biologically active molecules, folding correctly, forming disulfide bonds, and glycosylation.[38] At the same time, *Pichia pastoris* is a methylotrophic yeast that can metabolize methanol as its single carbon source. Methanol is oxidized to formaldehyde during methanol metabolism. However, the alcohol oxidase that catalyzes the oxidation of methanol has a poor affinity for oxygen. Therefore, in the process of metabolizing methanol, *Pichia pastoris* must produce a large number of enzymes to metabolize methanol. At the same time, promoters produced by regulatory enzymes are therefore used to drive the expression of heterologous proteins in Pichia pastoris. The two genes encoding alcohol oxidase in P*ichia pastoris* are AOX1 and AOX2. In general, most of the enzyme's activity depends on the AOX1 gene.[39-41] And its expression is primarily controlled by methanol. Therefore, the target gene inserted into the plasmid with the AOX1 promoter can be used to induce the expression of the desired heterologous protein. Following the published paper from our lab, the *Pichia pastoris* expression system can be developed to overexpress functional AHR proteins with higher yield

stably and lower cost compared to the baculovirus insect and mammalian cell expression system.[42]

To summarize, the hypotheses for the project are:

1. Purified mouse full-length *b*-allele of AHR will be enriched in Elute 1 (E1) after Pichia expression and metal affinity purification.

2. Higher copy number will be selected by higher Zeocin concentration and higher protein expression will be detected in higher copy number strain.

3. Purified aa108-400 human AHR will be enriched in Elute 1 (E1) after Pichia expression and metal affinity purification in both strains.

To test these hypotheses, there are three specific aims:

1. Separately clone a stable codon-optimized mouse *b*-allele AHR and human aa108-400 AHR genes, with alcohol oxidase (AOX) promoter, c-myc and 6-his tag at the end of genes, into *Pichia pastoris* genome to express mouse *b*-allele AHR and human aa108-400 AHR proteins.

2. Cobalt resin will be used to enrich the mouse *b*-allele AHR and human aa108-400 AHR proteins.

3. Relative copy numbers of mouse *b*-allele AHR and human aa108-400 AHR genes will be determined by qPCR.

CHAPTER 2: EXPRESSION MOUSE FULL-LENGTH B ALLELE AHR IN PICHIA PASTORIS

2.1 Successful pPICZ-B Mouse Full-length *b*-allele AHR Plasmid Construct Creation

2.1.1 pPICZ plasmid for cloning. Primers were designed to amplify the nucleotides sequence of mouse full-length *b*-allele AHR from the incorrect construct pPICZ-A mouse full-length *b*-allele AHR, which contains a stop codon. The pPICZ A, B, & C Pichia vectors usually are designed for simple cloning and selection, high-level expression, and rapid detection and purification of the recombinant protein. In addition, these vectors contain the Zeocin™ resistance gene to select integrated strains after electroporation directly. The recombinant protein is expressed as a fusion protein containing the c-myc epitope and a polyhistidine (6xHis) tag. The vector has multiple cloning sites to insert foreign genes and allows high-level methanol-induced expression of the target gene in *Pichia pastoris*. pPICZ A, B, and C Map as shown in Figure 4.

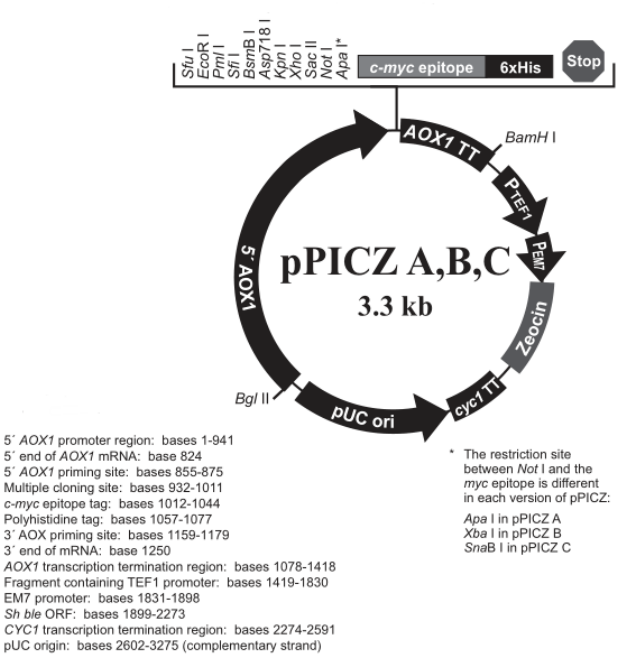

Figure 4. Map of pPICZ A, B, and C.

The corresponding restriction enzyme was used to cut the pPICZ-A mouse full-length *b*-allele AHR through EcoRI and Kpnl. Gel electrophoresis (Figure 5 a & b) shows the sequence of mouse AHR full-length *b*-allele is around 2.4kb through Polymerase Chain Reaction (PCR), which is consistent with the calculation result of our nucleotides.

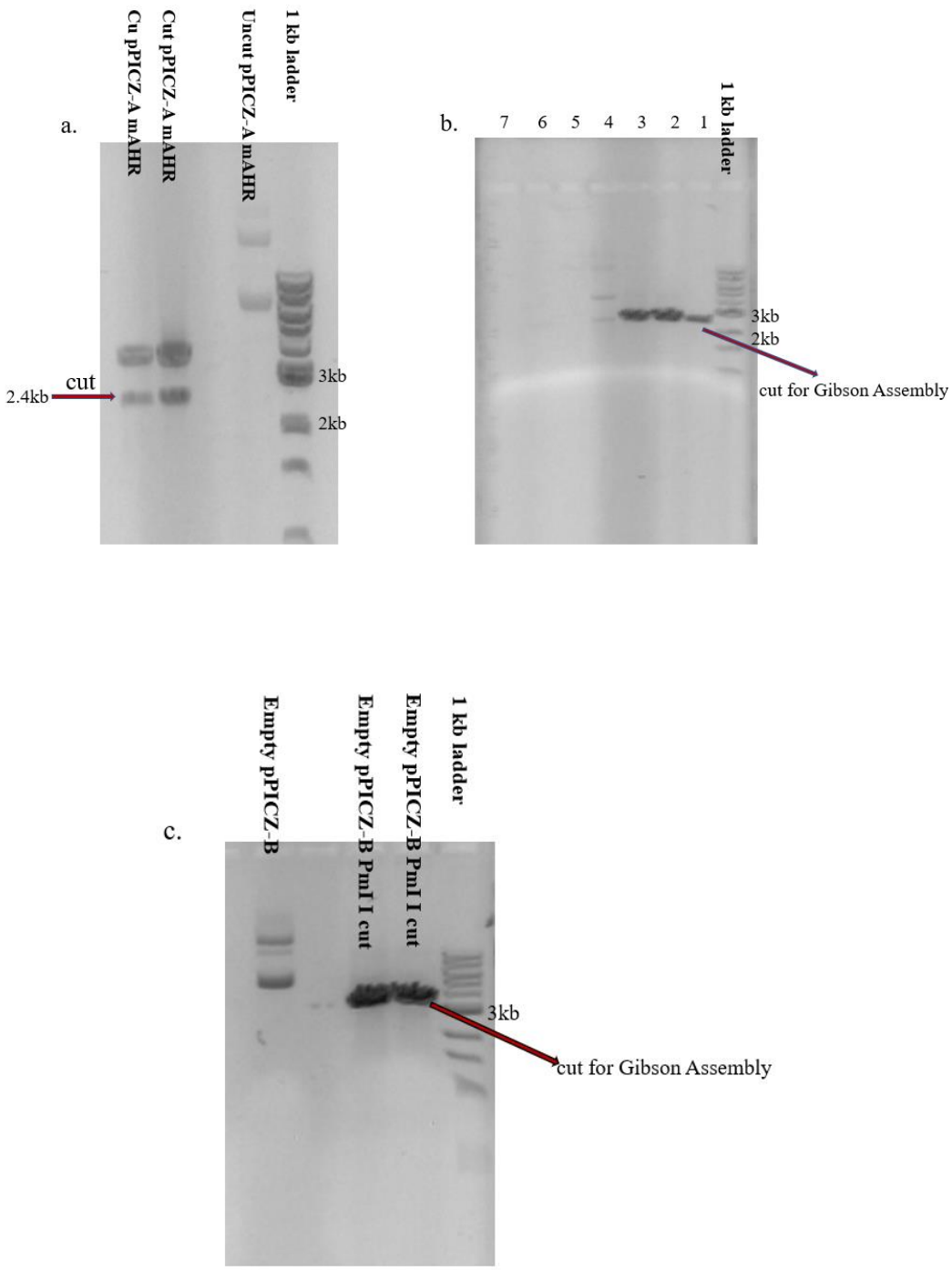

Figure 5. Construct pPICZ-B mouse full-length *b*-allele AHR plasmid. (a) Cutting pPICZ-A mouse full-length *b*-allele AHR through EcoRI and Kpnl. Gel electrophoresis shows the sequence of mouse full-length *b*-allele AHR is around 2.4kb. (b) Amplifying DNA sequence of mouse full-length *b*-allele AHR. Gel electrophoresis shows the appropriate size of amplification products. (1. 0.1ng DNA cut pPICZ-A full-length *b*-allele mouse AHR as template 2. 1ngDNA cut pPICZ-A mouse full-length *b*-allele AHR as template 3. 0.1ng DNA uncut pPICZ-A mouse full-length *b*-allele AHR as template 4. 1ng DNA uncut pPICZ-A mouse full-length *b*-allele

(Figure 5 Continued)

AHR as template 5. 0.1ng DNA cut pPICZ-A mouse full-length *b*-allele AHR as a template without PCR 6. 1ngDNA cut pPICZ-A mouse full-length *b*-allele AHR as a template without PCR 7. 1ngDNA uncut pPICZ-A mouse full-length b allele AHR as a template without PCR) (c) Linearized empty pPICZ-B plasmid through Pml I. Gel electrophoresis shows the appropriate size of the linearized empty pPICZ-B plasmid.

2.1.2 pPICZ-B mouse full-length *b*-allele AHR plasmid construct creation. Gibson Assembly allows the successful assembly of multiple DNA fragments, regardless of fragment length or end compatibility. This method has been successfully used to assemble oligonucleotides, DNA with varying overlaps (15-80 bp), and fragments hundreds of kilobases long and has been rapidly adopted by the synthetic biology community due to its ease of manipulation. Flexibility and applicability for use with large DNA constructs. Furthermore, restriction digestion of DNA fragments is not required after PCR. Meanwhile, the backbone vector can be digested or synthesized by PCR. Because it requires fewer steps and fewer reagents, it is cheaper and faster than traditional cloning protocols. Gibson Assembly allows the insertion of one or more DNA fragments into almost any location in a linearized vector and is independent of the presence of restriction sites in the specific sequence to be synthesized or cloned. Larger numbers of DNA fragments can be ligated with higher efficiency in a single reaction than traditional methods. The process of Gibson Assembly is shown in Figure 6. Figure 5 b and c show the DNA fragment and Linear vector, respectively.

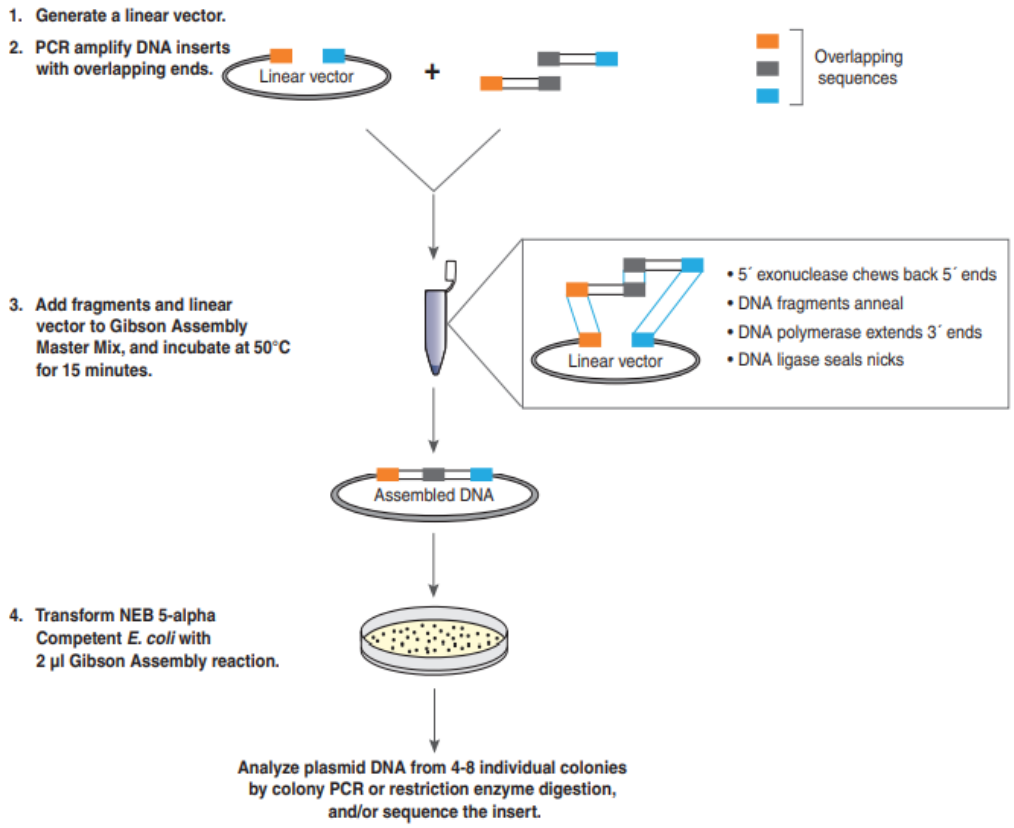

Figure 6. Overview of the Gibson Assembly Cloning Method

In order to get enough construct plasmid, we inserted the construction to BL-21 strain through electroporation as shown in Figure 7.

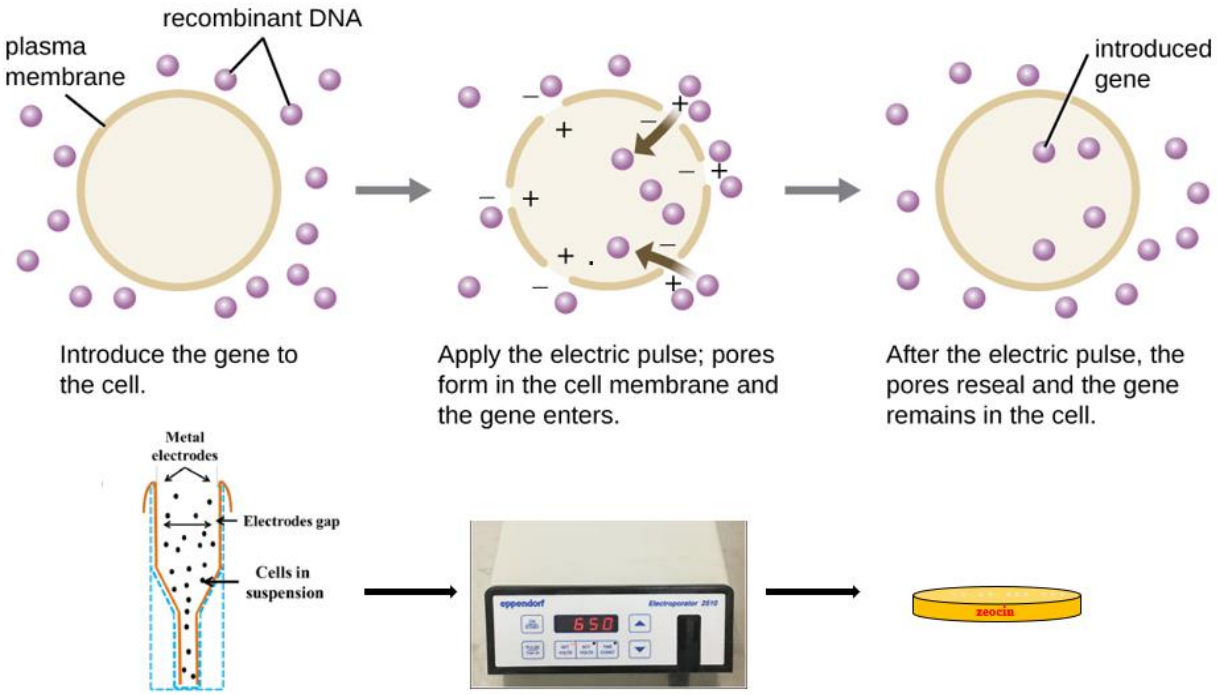

Figure 7. Electroporation Transformation

2.1.3 Check pPICZ-B mouse *b*-allele AHR plasmid construct. We choose the Stu I cut the plasmid since there are only two restriction enzyme sites on the pPICZ-A or pPICZ-B mouse full-length *b*-allele AHR plasmid construct. One is on the sequence of mAHR^{b}, and another is on the vector. Stu I located on nucleotides 1566 (mAHR^{b} sequence) and 4951 (vector) of pPICZ-A mouse full-length *b*-allele AHR construct, which has another 270 bp on the vector end compared to pPICZ-B mouse full-length *b*-allele AHR construct. As shown in Figure 8, we will see two bands for each construct after double digestion with the above information. For the pPICZ-A mouse full-length *b*-allele AHR construct, the band will be 3.4 kb and 2.5 kb. Also, for the pPICZ-B mouse full-length *b*-allele AHR construct, the band will be 3.4 kb and 2.3 kb.

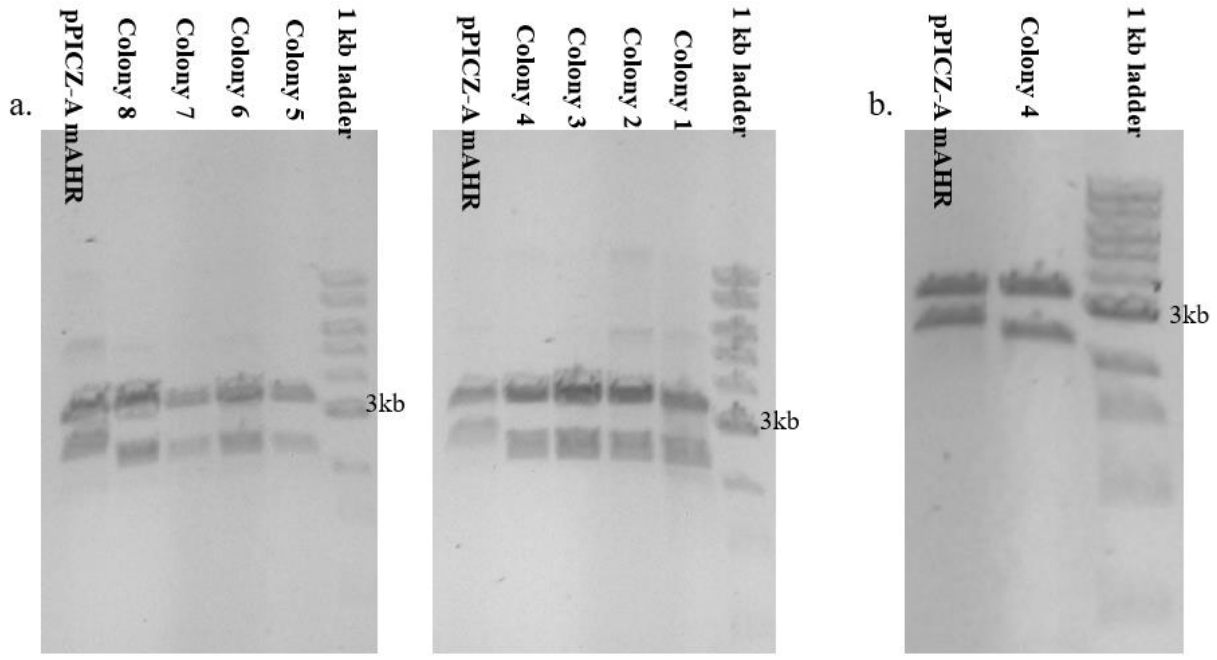

Figure 8. Detection of construct plasmid (pPICZ-B mAHR^{b}) after electroporation. (a) Double digestion of eight selected colonies through Stu I. The homemade solution is used to extract the constructed plasmid. (b) Double digestion of colony 4 by Stu l. The miniprep kit is used to extract the constructed plasmid.

2.2 Transformation Into Yeast Strains

To promote integration, we linearize the pPICZ-B mouse *b*-allele AHR construct at the Sac I enzyme site within the 5'AOX1 region. Then transform the pPICZ-B mouse full-length *b*-allele AHR plasmid construct into both wild-type (yJC100) strain and protease-deficient (ySMD1163) strain through electroporation. Figure 9 shows the linearized plasmid. We also can check the result of transformation easily by colony PCR as shown in Figure 10.

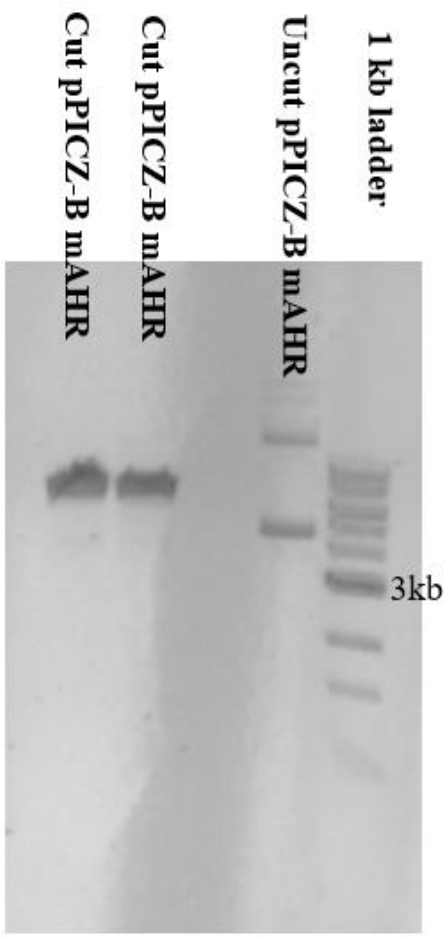

Figure 9. Linearizing pPICZ-B-mAHR[b] through Sac I.

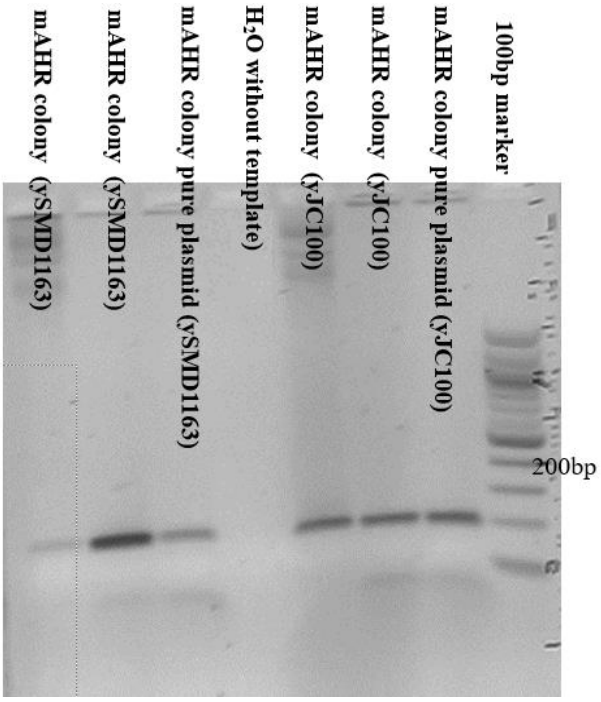

Figure 10. Detection of recombinant plasmid (pPICZ-B mAHR^{b}). After electroporation of the pPICZ-B mouse full-length *b*-allele AHR into two types of yeast strains. We use specific primers to detect target DNA by colony PCR.

2.3 Mouse Full-length *b*-allele AHR Induction Time Determination

In an effort to identify an optimal induction time for recombinant protein, we did a time course for 72 hours of both yeast strains. Normalization expression volume amount of protein for Coomassie blue stain and Western blot, we can see the highest protein level in 3 hours induction time of WT strain (Figure 11), 6 hours induction time of protease-deficient strain (Figure 12). We will choose the 6 hours as idea induction time for further study protein expression.

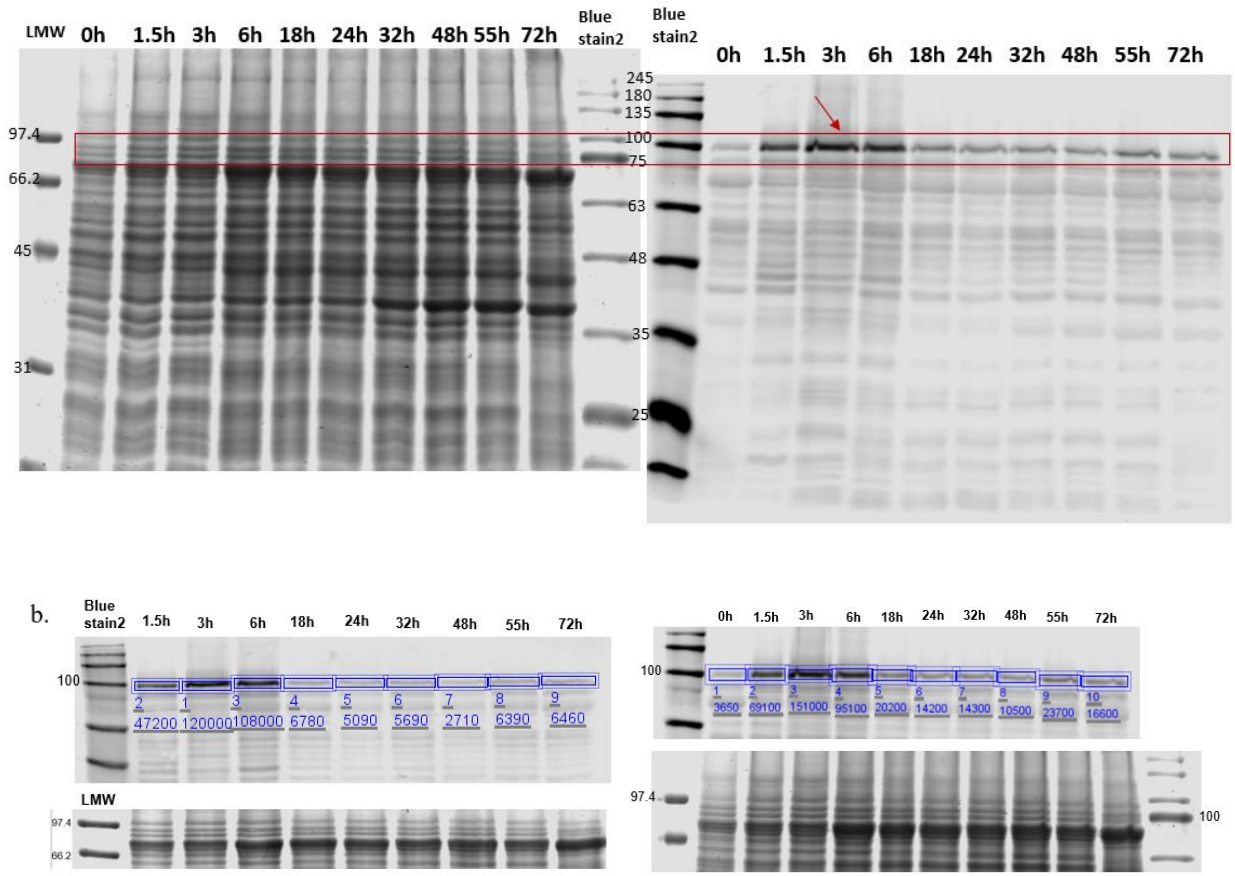

Figure 11. Induction time course of mouse full-length *b*-allele AHR in yJC100. (a) Collect different time point cultures and harvest protein. (Left) Coomassie blue stain (Right) Western blot. Use SA210 antibody to detect mouse full-length *b*-allele AHR. (b) Same process as (a). (Left) Load 10µl sample for each line. (Right) Load 8µl sample for each line.

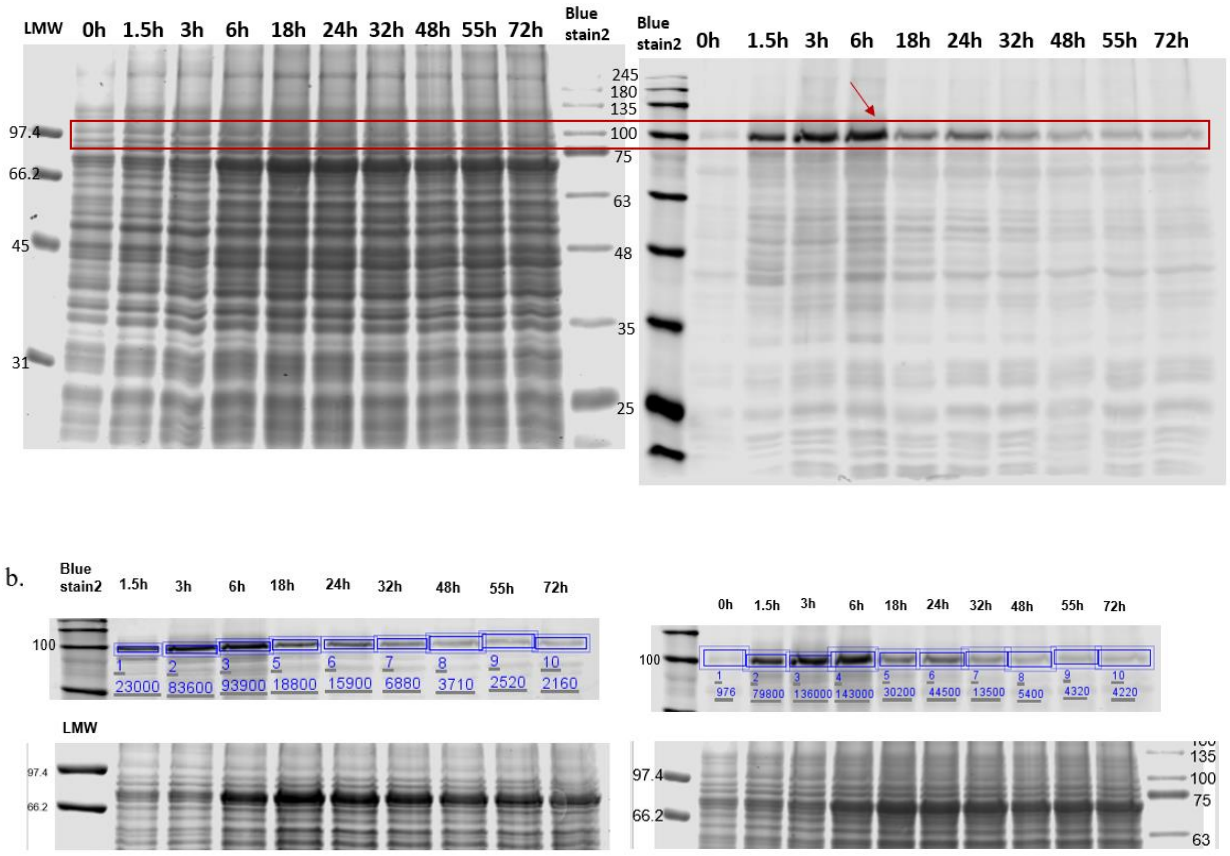

Figure 12. Induction time course of mouse full-length *b*-allele AHR in ySMD1163. (a) Collect different time point cultures and harvest protein. (Left) Coomassie blue stain (Right) Western blot. Use SA210 antibody to detect mouse full-length *b*-allele AHR. (b) Same process as (a). (Left) Load 10µl sample for each line. (Right) Load 8µl sample for each line.

2.4 Mouse AHR Expression and Purification

The Pichia expressed mouse full-length *b*-allele AHR protein was detectable by Western blot but not by Coomassie Staining (Figure 13). The amount of mouse full-length *b*-allele AHR protein is such less amount that it cannot be seen on Coomassie Staining. However, the antibody is so sensitive that we can see a strong band on the membrane of Western blot. We can see there has a significant band of E1 in both Pichia strains (Figure 13). After three times dependent experiments, we can see a higher amount of Load and E1 in the WT strain compared with the protease-deficient strain (Figure 14.).

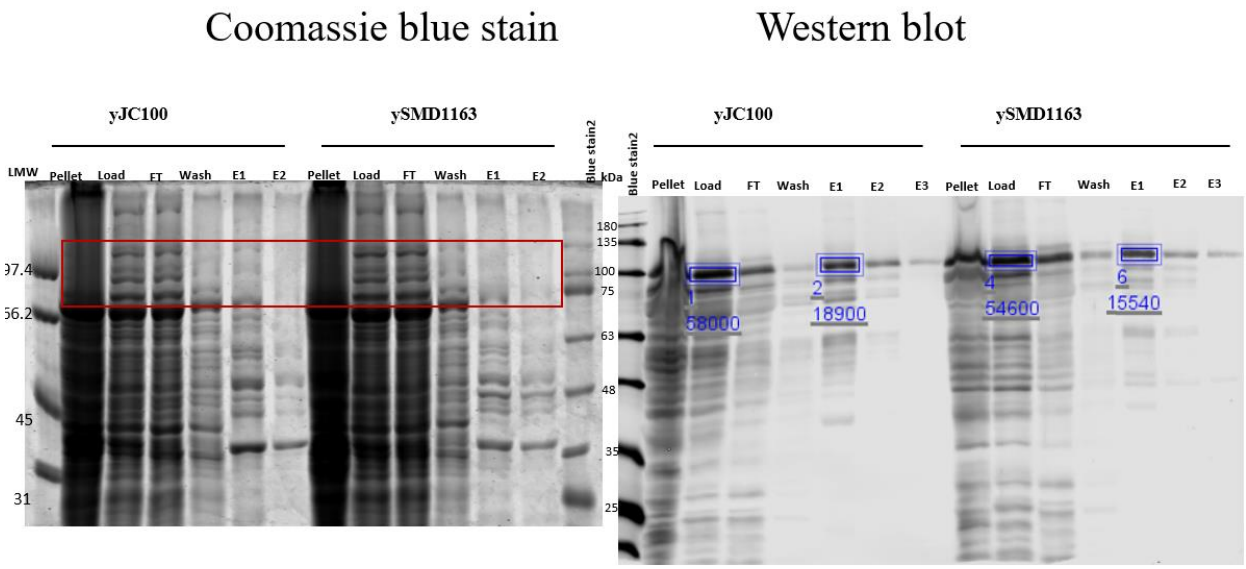

Figure 13. Expression and purification of mouse full-length *b*-allele AHR in two types of yeast strains. Use 2.6 ml load for purification. Collect flow-through (FT), wash, elute 1(E1), E2, and analysis. Use 10µl for Coomassie blue stain analysis and 8µl for western blot. SA210 antibody to detect mouse full-length *b*-allele AHR. (Left) Coomassie blue stain (Right) Western blot.

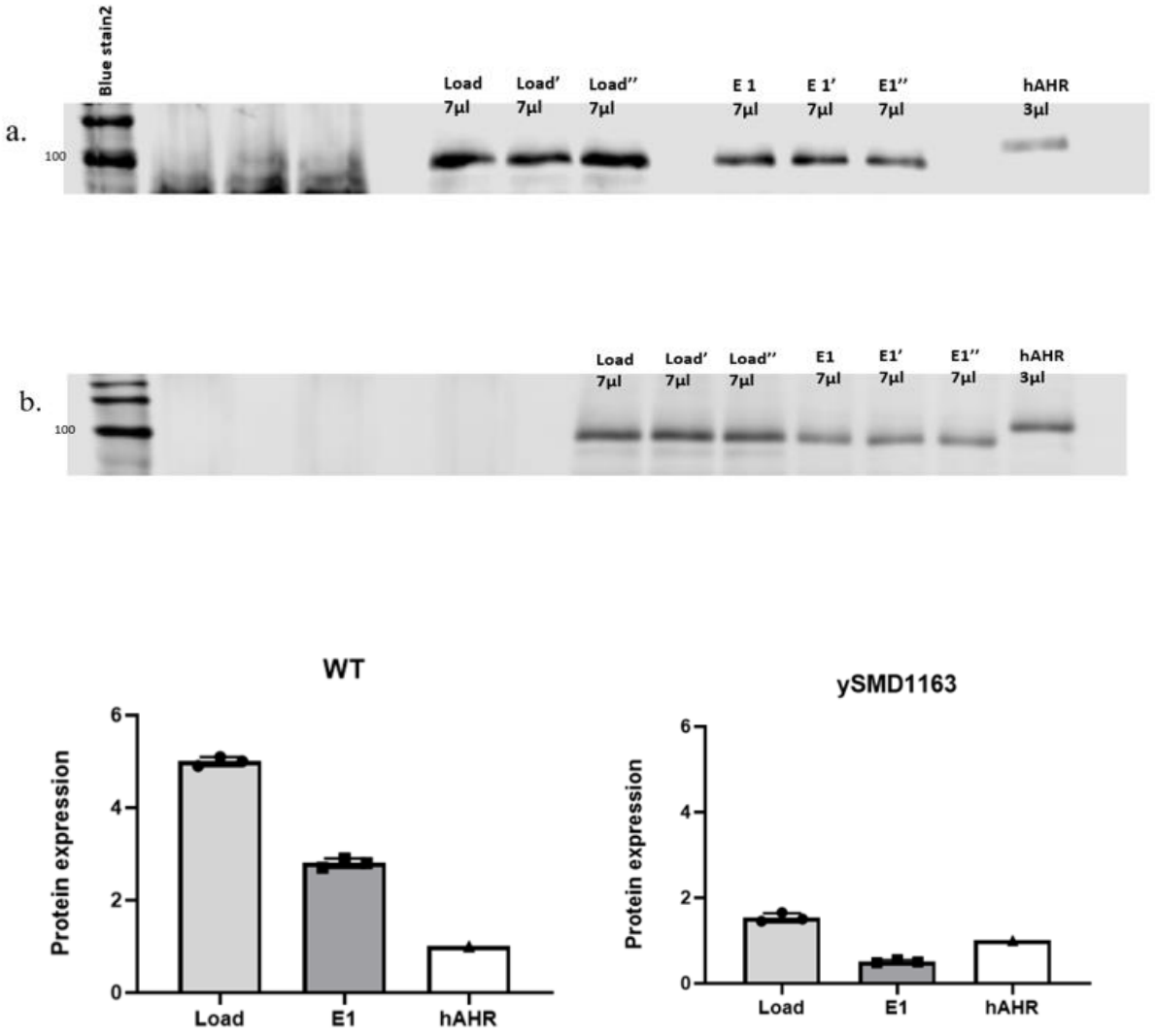

Figure 14. Expression and purification of mouse full-length *b*-allele AHR in two types of yeast strains. Repeat the whole process three times. Use 7µl Load and E1, 3µl hAHR for Western blot. SA210 antibody to detect mouse full-length *b*-allele AHR. (a) yJC100 strain (b) ySMD1163 strain.

2.5 Relative Copy Number of Mouse Full-length *b*-allele AHR

We use the different concentrations of Zeocin (from 100µg/µl to 1000µg/µl) to select the transformed yeast strain. We assume that the strain will obtain much resistance when integrated more copies into the genome. Mouse full-length *b*-allele AHR in WT (yJC100) strain can grow

up to 1000 μg/μl Zeocin YPD plate, but the mouse full-length *b*-allele AHR in protease-deficient (ySMD1163) only can grow on 500μg/ μl zeocin YPD plate. I use the actin as the reference gene to plot the curve of Log (Cq) vs. template amount. Since we already know actin has one copy in yeast, we can conclude the relative copy number for yJC100 (~ 4 copies) and ySMD1163 (~ 3.2 copies), which is consistent with the higher Zeocin concentration will select the strain with higher copies.

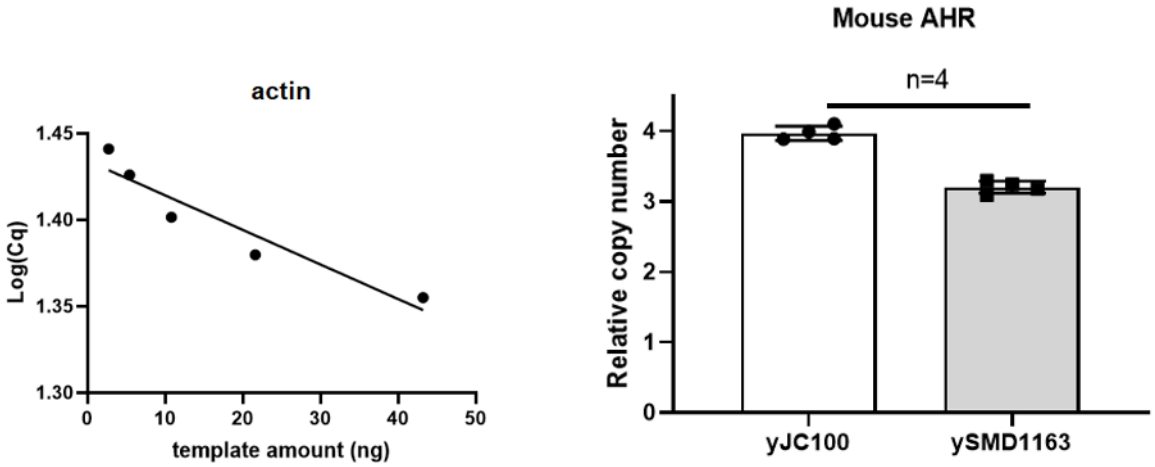

Figure 15. Relative copy number of mouse full-length *b*-allele AHR in WT (yJC100) and protease-deficient (ySMD1163) strain.

CHAPTER 3: EXPRESSION LIGAND BINDING DOMAIN OF HUMAN AHR IN PICHIA PASTORIS

3.1 Successful pPICZ-B aa108-400 of Human AHR Plasmid Construct Creation

3.1.1 pPICZ-B aa108-400 of human AHR plasmid construct creation. Primers were designed to amplify the nucleotides sequence of aa108-400 of human AHR from the full sequence human AHR, which is inserted into the pPICZ-B vector. Gibson Assembly method was used to ligate the pPICZ-B linearized vector (Figure 16 a) to aa108-400 hAHR DNA sequence (Figure 16 b).

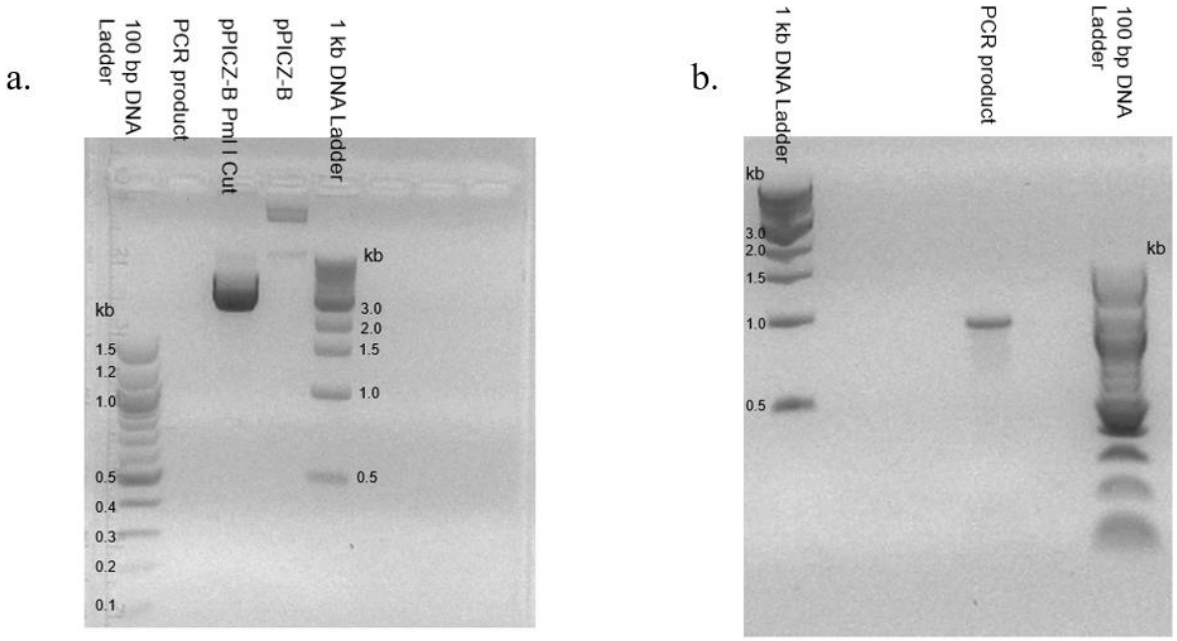

Figure 16. Construct pPICZ-B aa108-400 of human AHR plasmid. (a) Linearized empty pPICZ-B plasmid through Pml I. Gel electrophoresis shows the appropriate size of the linearized empty pPICZ-B plasmid. PCR doesn't work because the template of PCR (PL301 pPICZ-B-hAHR) is not good enough to amplify. (b) Use a new pPICZ-B-hAHR extraction plasmid as a template. Gel electrophoresis shows the PCR product (aa108-400 of hAHR)

3.1.2 Check pPICZ-B aa108-400 of human AHR plasmid construct. We choose eight colonies of pPICZ-B aa108-400 of human AHR after electroporation into the BL-21 strain. EcoR I and Xho I will be used to do the double digestion of eight selected colonies. Gel electrophoresis is shown in Figure 17.

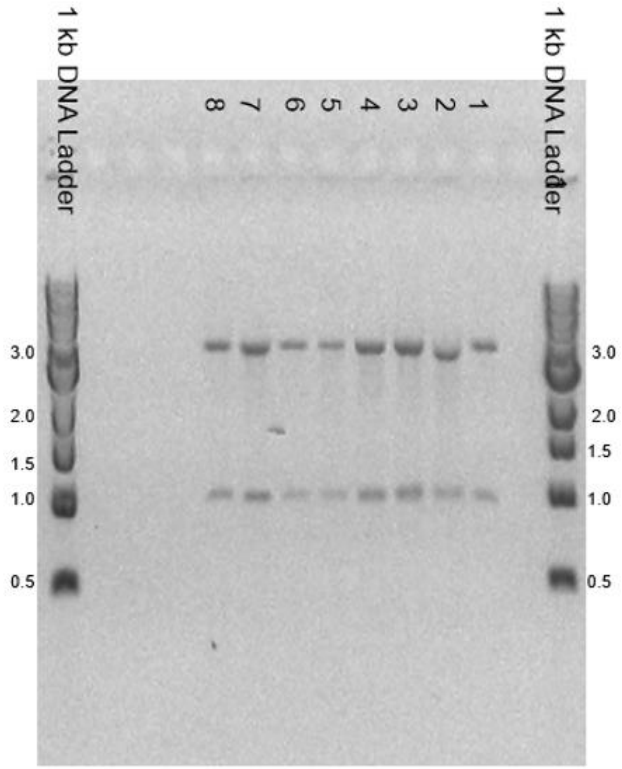

Figure 17. Detection of recombinant plasmid (pPICZ-B aa108-400 of hAHR) after electroporation. Double digestion of eight selected colonies. The homemade solution is used to extract the plasmid.

3.2 Transformation Into Yeast Strains

Same as 2.2, the pPICZ-B aa108-400 hAHR plasmid construct was transformed into both wild-type (yJC100) strain and protease-deficient (ySMD1163) strain through electroporation.

Figure 18 shows the linearized plasmid. The result of transformation will be detected by colony PCR as shown in Figure 19.

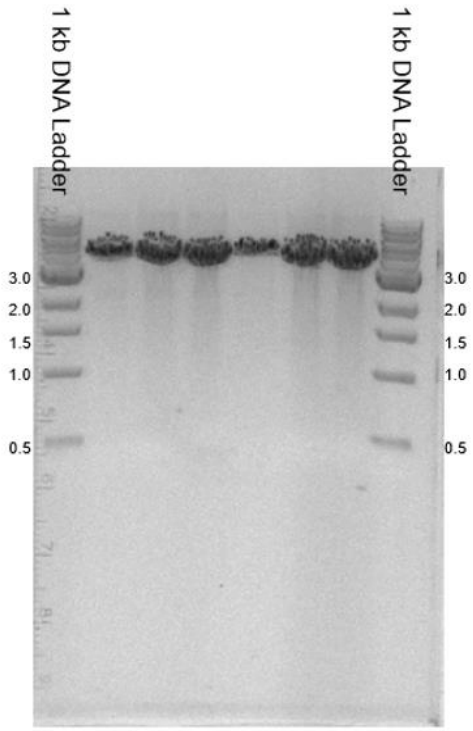

Figure 18. Linearizing pPICZ-B aa108-400 of hAHR through Sac I.

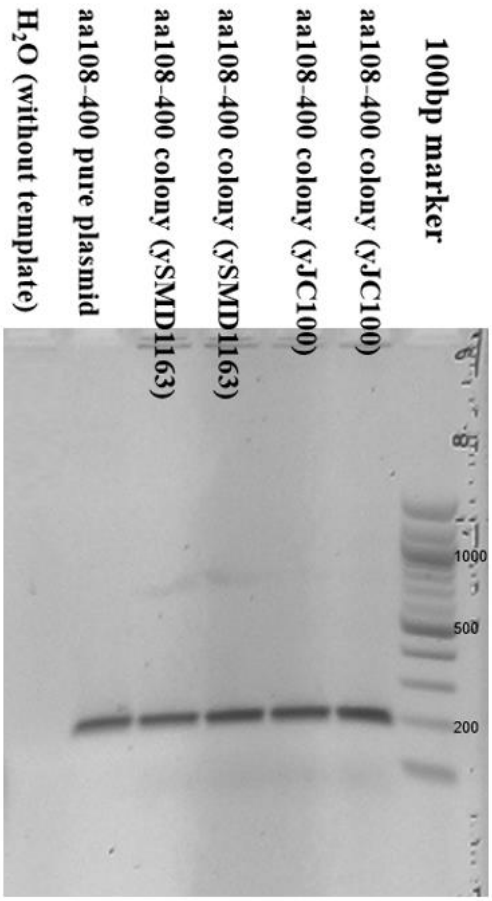

Figure 19. Detection of recombinant plasmid (pPICZ-B aa108-400 of hAHR). After electroporation of the pPICZ-B aa108-400 of hAHR into two yeast strains. We use the specific primer to detect target DNA by colony PCR.

3.3 Expression and Purification of hAHR aa108-400.

3.3.1 Use SA210 to detect the hAHR aa108-400 in Western blot. SA210 is a recombinant protein corresponding to aa 1-402 of the N-terminal AHR. So we need to know whether SA210 picks the aa108-400 AHR band since we lost aa 1-107 of AHR. Here we use MBP aa108-400 fusion protein to whether the results. The Maltose-binding protein (MBP) is one of the most popular fusion partners to produce recombinant proteins in bacterial cells. MBP allows one to use a simple capture affinity step on amylose-agarose columns, resulting in a protein that is often 70-90% pure.[43] Figure 20 shows the SA210 can pick up the MBP aa108-400.

We also know that our lab did the full sequence AHR expression and purification in the ySMD1163 strain. After repeating Dr. Yujuan. Zheng's experiment, we got the results for Figure 21.

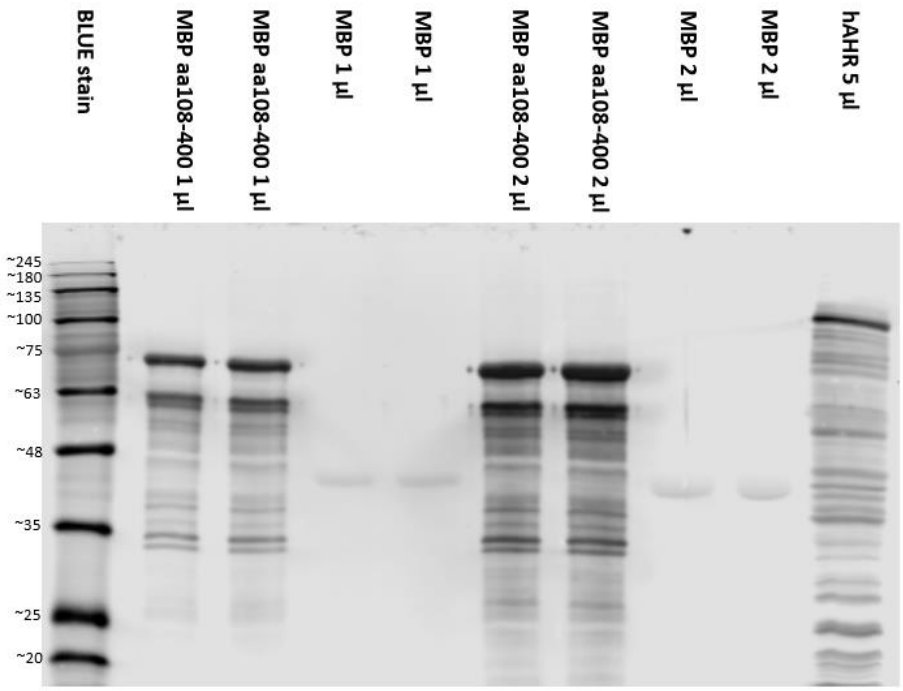

Figure 20. SA210 antibody can be used to detect MBP aa108-400. We use hAHR as positive and WT MBP as a negative control to see whether SA210 has good specificity to aa108-400 of hAHR.

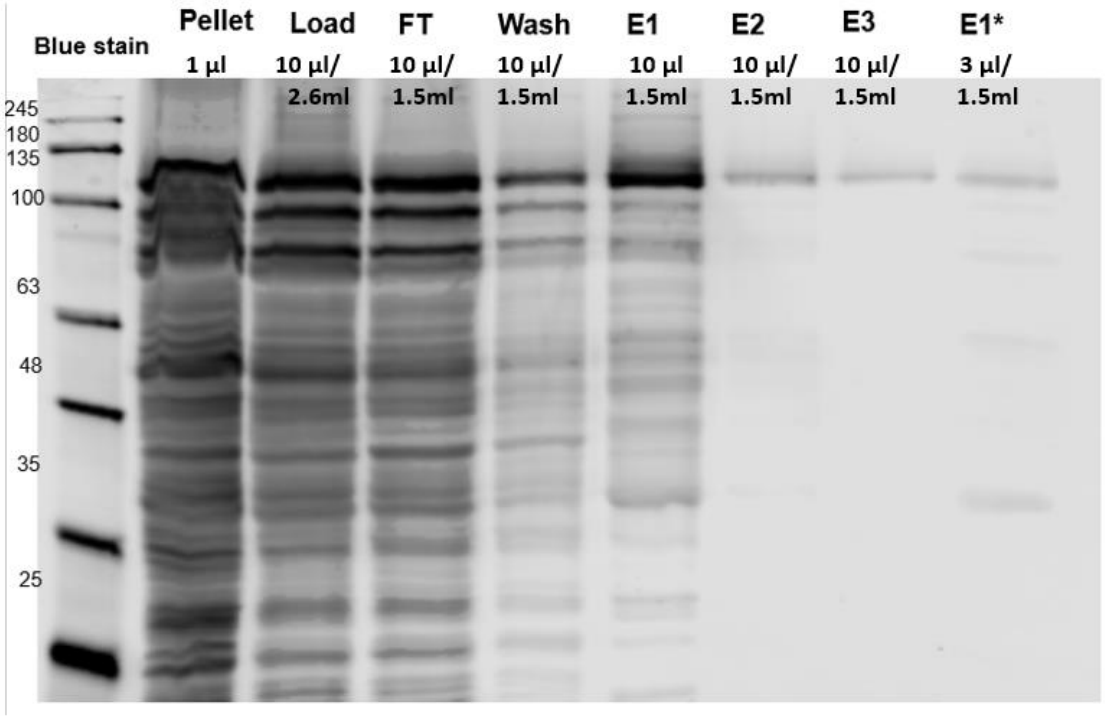

Figure 21. Expression and purification of human AHR in ySMD1163. Use 2.6 ml load for purification. Collect flow-through (FT), wash, elute 1(E1), E2, E3, and analysis. E1*(pi-pPICZ-B hAHR in 50% glycerol 6/20/17) as a positive control to see human AHR express success.

Unfortunately, we cannot find the significant band clearly when we use SA210 to purify aa108-400 AHR in both strains, as shown in Figure 22. To further study the reasons, we first compare the different Load and E1 in different strains. Figure 23 shows the strong band for mouse AHR and full-length hAHR but not aa108-400 hAHR.

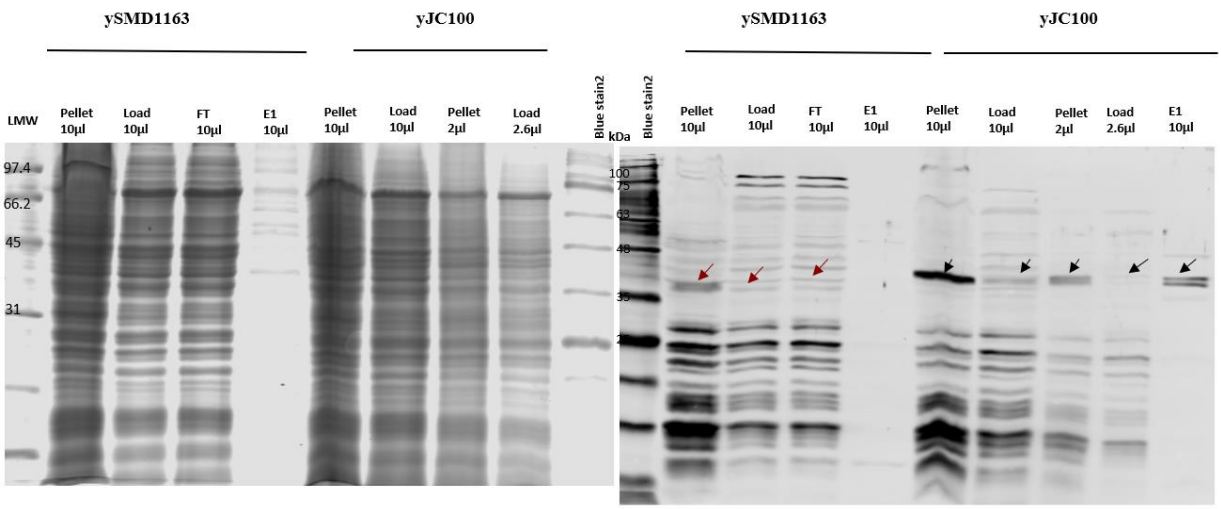

Figure 22. Expression and purification of aa108-400 of hAHR in two types of yeast strains. Use 2.6 ml load for purification. Collect flow-through (FT), wash, elute 1(E1), E2, and analysis. Use 15µl for Coomassie blue stain and 10µl for western blot. SA210 antibody to detect 108-400 of hAHR. (Left) Coomassie blue stain (Right) Western blot.

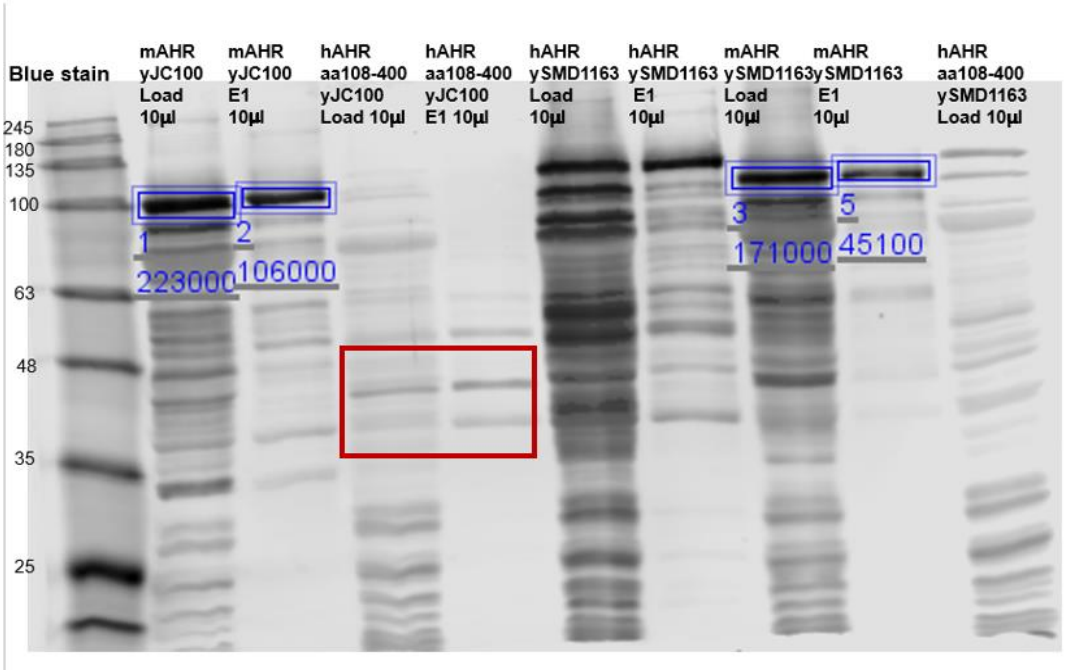

Figure 23. Comparison of different load and E1 through Western blot. SA210 antibody can be used to detect aa108-400 of hAHR and mouse AHR.

3.3.2 hAHR aa108-400 expression and purification. We know the pPICZ-B aa108-400 hAHR integrates into the yeast genome since we do see the band in Colony PCR. Dr. Lin-Cereghino provided us with another antibody, which is an anti c-myc antibody from the mouse. Thank you for his advice, and we can see the strong bands of WT strain on Western blot, as shown in Figure 24. But not any bands in ySMD1163. We did three times dependent purification for aa108-400 hAHR in WT strain to see the relationship between the Load and E1, as shown in Figure 25.

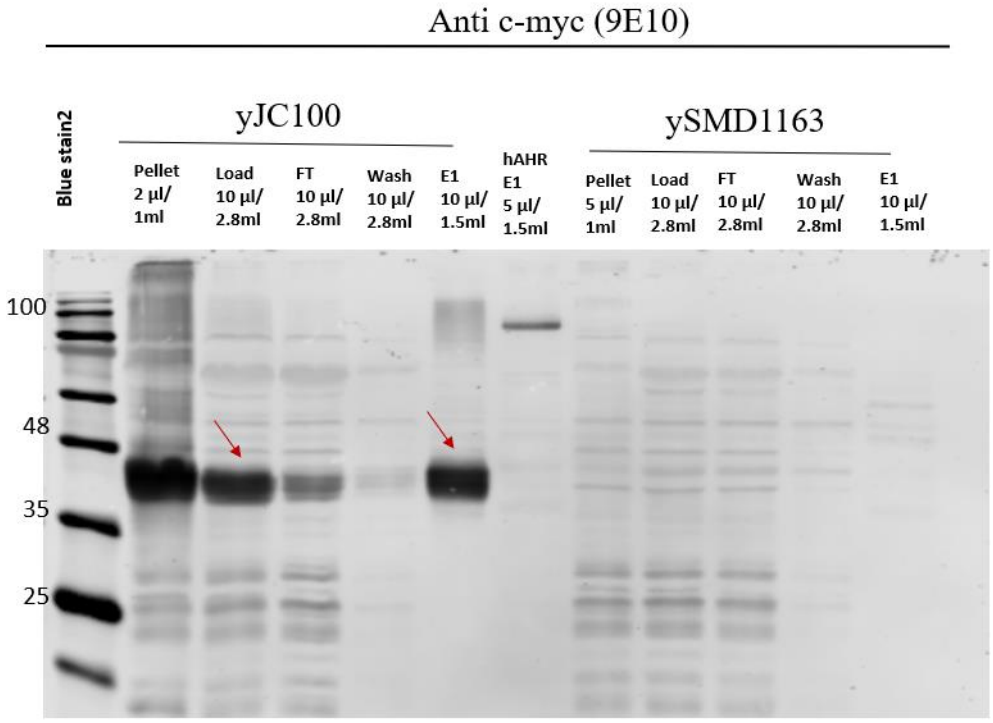

Figure 24. Expression and purification of aa108-400 of hAHR in two strains. Use 2.8 ml load for purification. Collect flow-through (FT), wash, elute 1(E1), E2, E3, and analysis. E1*(pi-pPICZ-B hAHR) as a positive control to see human AHR express success. We use the c-myc antibody to detect protein expression.

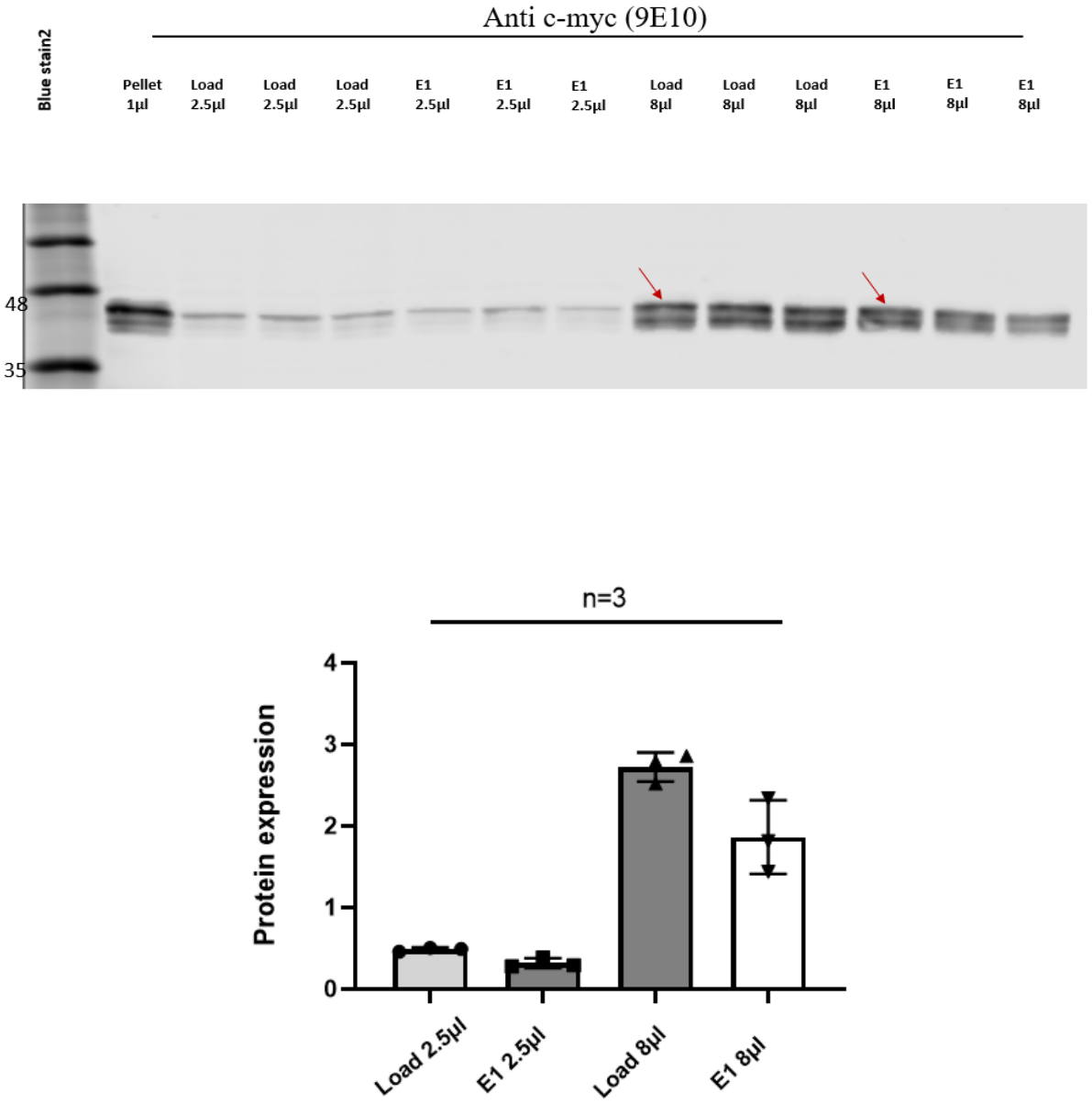

Figure 25. Expression and purification of aa108-400 of hAHR in yJC100 (WT) yeast stains. Repeat the whole process three times. Use the c-myc antibody to detect Pellet, Load, and E1 in Western blot.

3.4 Relative Copy Number and Fold Change of aa108-400 hAHR

We use the different concentrations of Zeocin (from 100μg/μl to 1000μg/μl) to select the transformed yeast strain. We assume that the strain will obtain much resistance when integrated more copies into the genome as previously. The aa108-400 hAHR in WT (yJC100) strain can grow up to 1000μg/μl Zeocin YPD plate, but the aa108-400 hAHR in protease-deficient (ySMD1163) only can grow on 500μg/μl Zeocin YPD plate. The full-length hAHR was selected by 100μg/μl before. I use the actin as the reference gene to plot the curve of Log (Cq) vs. template amount. Since we already know actin has one copy in yeast, we can conclude the relative copy number for yJC100 (~ 4 copies) and ySMD1163 (~ 2.9 copies) of the aa 108-400 hAHR, which is consistent with the higher Zeocin concentration will select the strain with higher copies. But we still see full-length hAHR has a similar relative copy number with the aa108-400 hAHR in ySMD1163 (Figure 26). We also checked the messenger RNA level for aa108-400 in both strains since no recombinant protein (aa108-400 hAHR) was expressed in the ySMD1163 strain. The result shows that the aa108-400 hAHR in the ySMD1163 has a higher mRNA level than the WT strain, even though no recombinant protein was expressed (Figure 27).

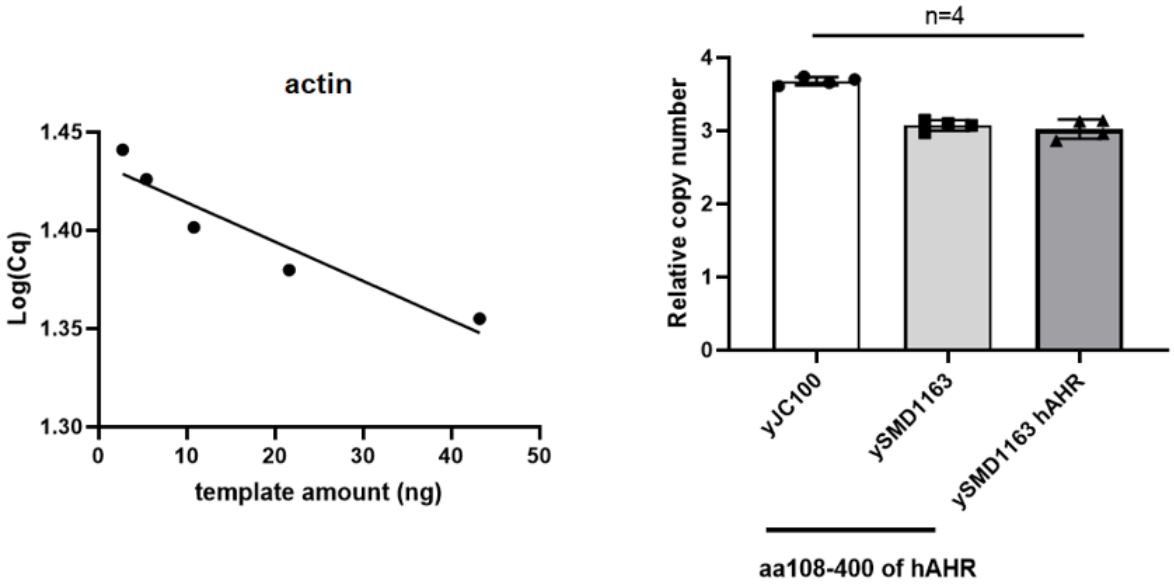

Figure 26. Relative copy number of aa108-400 AHR in WT (yJC100) and protease-deficient (ySMD1163) strain, together with full-length hAHR.

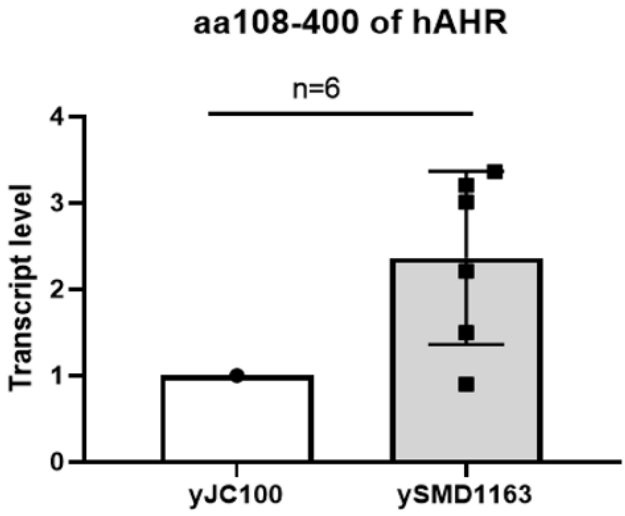

Figure 27. The mRNA levels of aa108-400 AHR in WT (yJC100) and protease-deficient (ySMD1163) strain.

CHAPTER 4: DISCUSSION

The AhR was discovered in 1976 by Dr. Poland as a dioxin-binding protein.[44] AHR has been shown to mediate many pathways involved in our immune systems, body function, and others. Our lab studied AHR for many years, including its ligand, chaperon, signal pathway, and expression system.

The first human AHR overexpression in baculovirus was more than thirty years ago. Researchers in our lab also use baculovirus to express the AHR. For example, we use a baculovirus system to express human AHR and ARNT.[37] Other scientists also use the baculovirus system to express rat AHR to study its ligand binding.[36] Due to the nature of the insect cells, it requires more times infection by cloned baculovirus to continually express recombinant proteins with high titer viral stock since baculovirus will be consumed for a long time expression.[36] This will increase the possibility of the mutations for the baculovirus genome over time, which might change the characteristic of the recombinant protein or lower the amount of protein expression. We also use the bacterial expression system to express many deletion constructs of human AHR. Bacteria usually express a high yield amount of AHR with economic experiments, but it lacks post-translation. Our lab usually studies the function and signaling pathway through the mammalian cells because of its time-consuming and expensive cost. Therefore, my project focuses on the optimized Pichia expression system, which is relatively reliable and performs much eukaryotic post-translational modification. In addition, the transformed yeast cells, which can be made into glycerol stocks and stored at - 80°C, can be continuously used to produce the recombinant proteins.

Previous researchers in our lab had successfully expressed full-length hAHR in the ySMD1163 strain. Unfortunately, the amount of soluble human AHR is still meager, and it is hard to do further studies like ligand binding, x-ray crystal structure, and new ligand findings. Therefore, we created the truncated protein of the AHR ligand-binding domain to see whether we could increase the expression yield. The Pichia expression system makes it a better alternative to obtain functional AHR for mechanistic studies. We also express mouse full-length b allele of AHR since the affinity is different between mouse and human.

4.1 Expression of mouse full-length *b*-allele AHR in Pichia pastoris

AhR, a ligand-activated transcription factor, is formed in a stable complex by two molecules, heat shock protein 90 (Hsp90), immunophilin-like AhR interacting protein (AIP, also known as XAP2 and ARA9), and p23. The interaction of Hsp90 with human AhR is less stable and requires the presence of molybdate in the buffer to stabilize the receptor/Hsp90 complex.[45-47] That is, mouse AHR is more stable compared with human AHR, which is the reason for expressing mouse AHR. We already know the amino acid sequence of mouse AHR and human AHR has high similarities. However, there are significant differences in gene regulation, recruitment of coactivators, and ligand affinity in $mAHR^{b}$ with hAHR.

We hypothesis that purified mouse full-length *b*-allele of AHR will enrich in Elute 1 (E1) after Pichia expression and metal affinity purification. After the mouse full-length *b*-allele of AHR transformed into both WT (yJC100) and protease-deficient (ySMD1163) strains. We see the enrichment of mouse full-length *b*-allele of AHR in both strains, but the yield is not as we expect high compared with the full-length human AHR in the ySMD1163 strain (Figure 22). I also found it hard to make good quality ySMD1163 competent cells when using the traditional method. Thanks to Dr. Lin-Cereghino for teaching me how to make fresh competent cells for

both yeast strains. After the electroporation, the results of colony and qPCR told us that the mouse full-length *b*-allele AHR integrated into the yeast genome successfully. Besides, the higher Zeocin concentration selected a higher copy number, which is consistent with the protein expression result (WT strain has a higher yield of protein expression). Generally, we would see a more elevated amount of protein expression in protease-deficient since the protein cannot be degraded when it has the same copy number. I would expect to see higher protein expression yield when we selected the WT strain (mouse full-length *b*-allele AHR) to only grow on 500 μg/μl Zeocin but not 1000 μg/μl.

4.2 Expression of aa108-400 Human AHR (LBD) in *Pichia pastoris*

The aa108-400 human AHR contains the ligand binding domain (aa 217-390) of hAHR. The ligand binding site of AHR is contained within the PAS-B domain and contains several conserved residues critical for ligand binding. Modulation of the AHR function is primarily conferred through interaction at the ligand binding domain using agonists and antagonists. And why I choose the ligand binding domain when creating the truncated protein? There have several reasons: 1) The function of this domain can be tested by ligand binding; 2) If we got enough protein, we could have the crystal structure of ligand binding domain; 3) Study the binding affinity. Our lab also expresses MBP aa108-400 of human AHR in *E.coli* successfully.

We hypothesis that purified aa108-400 human AHR will enrich in Elute 1 (E1) after Pichia expression and metal affinity purification in both strains. Unfortunately, we do not see an interesting recombinant protein in the ySMD1163 strain. The results of colony PCR and qPCR show that the aa108-400 hAHR successfully transforms into the yeast genome. Therefore, in order to see what happened, we measured the mRNA level for both strains. The result shows that the aa108-400 hAHR in the ySMD1163 has a higher mRNA level than the WT strain, even

though no recombinant protein was expressed. Therefore, we assume that aa108-400 of human AHR (LBD) cannot be used as a template to generate protein, or the transcriptome is less stable in the ySMD1163 strain. I was wondering why aa108-400 hAHR cannot be expressed in ySMD1163. How about choosing another vector to insert the cDNA sequence into the yeast genome? The idea result is to express an abundant amount of ligand binding domain to be able to generate a crystal structure for x-ray crystallography and test the function by ligand binding.

4.3 Conclusion and Future Work

4.3.1 Conclusion For mouse full-length *b*-allele of AHR expression, we can conclude: 1) the higher enrichment E1 of mouse full-length *b*-allele AHR in WT strain; 2) higher copy number leads to higher amount soluble mouse full-length *b*-allele AHR protein in WT strain.

For aa 108-400 hAHR expression, we can conclude: 1) aa108-400 of human AHR (LBD) can be expressed and purify in the WT strain; 2) aa108-400 of human AHR (LBD) cannot be used as a template to generate protein, or the transcriptome is less stable in the ySMD1163 strain.

4.3.2 Future work Further modification of this Pichia expression system is needed to obtain a more abundant and purer recombinant protein. And the function of mouse *b*-allele AHR and aa108-400 hAHR is needed to be studied.

CHAPTER 5: MATERIALS AND METHODS

5.1 Material

All chemicals and ingredients to prepare bacteria and yeast media (LB, YPD, BMGY, and BMMY) were purchased from Thermo Fisher Scientific (Hampton, NH) or Sigma Aldrich (St. Louis, MO). All oligonucleotides, Zeocin, otherwise specified, were purchased from Thermo Fisher Scientific (Hampton, NH). All restriction enzymes were purchased from New England Biolabs (Ipswich, MA). Anti-AHR SA210 rabbit IgG was purchased from Enzo Life Sciences (Farmingdale, NY). Anti-c-myc mouse IgG was purchased from Santa Cruz Biotechnology (Dallas, TX, USA). IRDye800 donkey secondary IgG was purchased from LI-COR (Lincoln, NE). IRDye680-conjugated DRE was purchased from Integrated DNA Technologies (San Diego, CA). The Gene Art codon-optimized human AHR sequence (GA-hAHR) was purchased from Life 34 Technologies (Thermo Fisher Scientific, Grand Island, NY). Cobalt agarose beads were purchased from Gold Biotechnology (St. Louis, MO). The Gibson assembly kit was purchased from New England Bioabs (Ipswich, MA). The Colony PCR master mix was purchased from Lucigen Corporation (Middleton WI). The pPICZ-B plasmid and the Pichia strains (ySMD1163 and yJC100) were gifts from Dr. Geoff Lin-Cereghino (University of the Pacific). Western blot analyses were performed using an LI-COR Odyssey imaging system (Lincoln, NE). Electroporation was performed using an Eppendorf 2510 electroporator (Edison, NJ). Cells pellets were homogenized using Bullet Blender Storm 24 (Troy, NY). Media and culture plates were made by following the recipes from “Pichia Expression Kit Version L” and are briefly introduced as below. YPD (Yeast extract Peptone Dextrose) plate: 1% yeast extract, 2% peptone, 2% dextrose, 2% Agar. BMGY (Buffered Glycerol-complex Medium): 1%

yeast extract, 2% peptone, 10% 1 M potassium phosphate buffer at pH 6.0, 1.34% YNB (Yeast Nitrogen Base), 4×10-5 % biotin, 1% glycerol. BMMY is the same as BMGY except for replacing 1% glycerol with 0.5% methanol. ZymoPURE II Plasmid Maxiprep kit and Direct-zol RNA Miniprep kit were purchased from Zymo Research (Irvine, CA, USA). MMLV high-performance reverse transcriptase was purchased from Epicentre (Madison, WI, USA). iTaq Universal SYBR Green Supermix was 103 purchased from Bio-Rad (Hercules, CA, USA).

5.2 Methods

5.2.1 Polymerase chain reaction (PCR). Mouse AHR (pPICZ-A mouse AHR) and full-length human AHR (pPICZ-B hAHR) were obtained from the Dr. William Chan lab. Mouse AHR template from two restriction enzyme reactions (10 µl pPICZ-A mouse AHR plasmid, 196ng/ µl; 1 µl EcoR I; 1 µl Knp I; 2 µl 10X BSA; 2 µl Reaction Buffer; 2 µl 10X RNaseA; 2 µl dH_2O, q.s. 20 µl). Gently tap to mix, then put in the 37℃ oven overnight. PCR was performed using a Q5 DNA polymerase reaction (5 µl 5X reaction buffer; 2 µl 2.5 mMdNTPs; 1.3 µl 10uM forward primer OL881; 1.3 µl 10 uM reverse primer OL882; 1 µl 0.1ng mouse AHR template; 0.3 µl Q5 DNA polymerase; 14.1 µl dH_2O, q.s. 25 µl) in: (Hot Start) 98℃ for the 30s; 30 cycles of 98℃ for 10s, 55℃ for 15s, 72℃ 2 minutes 30s; then 72℃ for 2 minutes, hold at 4℃ indefinitely. The aa108-400 human AHR was obtained from pPICZ-B hAHR by Polymerase Chain Reaction. PCR was performed the same as mouse AHR except for the primers and extension time (forward primer OL835, reverse primer OL836; extension time 60s). Primer's information is shown in Table 1.

Table 1

Primer information for PCR, Colony PCR, RT-PCR.

Oligo (OL) Number	Primer Sequence
OL835	pPICZ-B-pohAHR aa108-400 Forward primer 5'- taattattcgaaacgaggaattcac atg GAG GGT TTG AAC TTG CAA GAG GGT G -3'
OL836	pPICZ-B-pohAHR aa108-400 Reverse primer 5'- ga tcc gag acg gcc ggc tgg gcc ac T AGT GTT TCT CTT TCT CAA GTG CTC -3'
OL881	pPICZ-B-mAHR b1 Forward primer 5'- taattattcgaaacgaggaattcac atg TCC TCC GGT GCT AAC ATC ACT TAC G -3'
OL882	pPICZ-B- mAHR b1 Reverse primer 5'- ga tcc gag acg gcc ggc tgg gcc ac A GGA CTG GAC CTT AGA CAA GAA AGC -3'
OL896	Forward primer for detecting mAHR. (Locate at the beginning part). 5'- TCCTCCGGTGCTAACATCACTTAC-3'
OL897	Reverse primer for detecting mAHR. (Locate at the beginning + 200bp part). 5'- GGACAACTTGTCCAACT-3'
OL898	Forward primer for detecting hAHR aa108-400 and hAHR. (Locate at the beginning of hAHR aa108-400). 5'-GAGGGTTTGAACTTGCAAGA -3'
OL899	Reverse primer for detecting hAHR aa108-400 and hAHR. (Locate at the beginning of hAHR aa108-400 + 200bp part). 5'- TGTCTCTGGAACTCAGCTCT -3'
OL900	Forward primer for detecting hAHR aa108-400 and mouse AHR. 5'-ccagagacaattgcactgggc-3'
OL901	Reverse primer for detecting hAHR aa108-400 and mouse AHR. 5'- ggtggctgcaatggagtagcga-3'
OL902	Forward primer for Pichia strain, use actin as reference. Copy number determination by RT-PCR. Use with OL903. 5'-CCTGAGGCTTTGTTCCACCCATCT-3'

(Table 1 Continued)

OL903	Reverse primer for Pichia strain, use actin as reference. 5'-GGAACATAGTAGTACCACCGGACATAACGA-3'

Note. All primers were optimized for Pichia expression.

5.2.2 Gibson assembly reaction. Both cDNA (mouse AHR and aa108-400 human AHR) were cloned into Pml I site of pPICZ-B plasmid to generate pPICZ-cDNA plasmid by Gibson assembly with PCR product of Pml I overhang primer. The cloned plasmid was used to express a recombinant protein with c-myc and 6-His tags at the C-terminus. Gibson Assembly Reaction was performed (1 μl 0.02pmol pPICZ-B Pml I; 9 μl 0.07pmol mouse AHR or aa108-400 hAHR Pml I overhang; 2X Gibson assembly master mix) in thermocycler: 50℃ for 20 minutes, then 4℃ afterward.

5.2.3 Electro transformation. Dilute 2 μl Gibson mix with 4 μl Nuclease-free water (1: 3) mix well, then use 1 μl of the diluted mixture for electroporation. The pPICZ-B mouse AHR and pPICZ-B aa108-400 human AHR construct were propagated in BL21 cells to generate a sufficient amount, respectively. Then the construct was linearized with Invitrogen recommended Sac I enzyme, which only cuts at the plasmid sequence once but not the insert before being transformed into the freshly prepared competent ySMD1163 Pichia pastoris strain (a protease deficient Pichia strain) by electroporation using an Eppendorf 2510 electroporator, according to the manufacturer's recommendation. The preparation of competent ySMD1163 cells and yJC100 followed by Dr. Geoff Lin-Cereghino's lab protocol. In order to select the Pichia cells, which containing the codon-optimized mouse AHR or aa108-400 human AHR gene and stably integrated into the yeast genome, we incubated the Pichia cells in 1 mL of 1 M

sorbitol at 30℃ shaking for one hour, followed by the addition of 1 mL of YPD media. Then the resulting culture was plated onto a YPD plate containing Zeocin (100 μg/μl ~ 1000 μg/μl), followed by 30℃incubation in a dark place for 2 ~ 4 days until the transformed yeast colonies were formed.

5.2.4 Colony PCR. Colonies were then streaked to obtain individual colonies and grown on another YPD plate containing Zeocin at 30 ℃ for 2 ~ 4 days. The newly formed colonies were validated by colony PCR to confirm the presence of the gene by using appropriate primers. Use OL896 and OL897 to detect the mouse AHR, OL898 and OL899 to detect the aa108-400 hAHR (Table 1).

5.2.5 Protein expression and purification. For a 100 mL Pichia expressed mouse AHR or aa108-400 hAHR purification, the freshly prepared colony was inoculated in BMGY media and incubated at 29℃ with rotation at 250 rpm until OD_{600} reached 2.5 absorbance units. A cell suspension was centrifuged at 1,500 g for 5 min at room temperature. The resulting pellet was washed twice with BMMY (50 mL) and then resuspended in 250 mL BMMY with 0.5% methanol in a 1 L baffled flask to induce the mouse AHR expression for 6 h with rotation at 250 rpm at 29℃. The cell suspension was then centrifuged at 3,100 g for 5 min at 4℃. The resulting cell pellet was washed with 20 mL of cold PBS and then centrifuged at 12,000 g for 5 min at 4℃. Resuspend the pellet (~6ml) with 5X pellet volume (~30ml) of lysis buffer (25mM HEPES, pH 7.4, 0.3M KCl, 10% glycerol, and 300 μl 100X Gold protease inhibitor cocktail GB108-2). Then aliquot lysate into 1.5ml Eppendorf tubes specifically for Bullet Blender (ZrOBO5 beads ~ 0.2ml; Speed setting at 8 for 5 minutes, cool down for 5 minutes, and repeat once). Spin at 16,000g for 20 minutes at 4℃, collect the supernatant. The resulting supernatant

(Load), which has the Pichia expressed protein, was collected as the starting material for metal-affinity purification using cobalt resin. To purify the recombinant protein, 1 mL of cobalt resin column was preconditioned by dH_2O and lysis buffer, 2.6 ~ 2.8 ml Load (Pichia cell lysate supernatant) was poured onto the 1 mL cobalt resin column at 4℃. The Load was passed through the column 5 times and collected the Flow-through (FT) at the last time. Then wash the column with 2.5 mL lysis buffer and collect the culture as Wash 1. Another 5 ml elution buffer (add 0.1702 imidazole, the final imidazole concentration is 0.5M) with lysis buffer containing 20 mM imidazole was applied onto purification. To elute the protein bound to the resin, 5 mL of an elution buffer containing was applied onto the column, and the eluate was collected. The majority of the recombinant protein was captured in the first 1.5 mL of the eluates (E1). Elute 2 (E2) and elute 3 (E3) are also collected as follows.

5.2.6 Coomassie blue staining and Western blot analysis. Coomassie Blue staining and Western blot analysis of SDS-PAGE were used to determine the presence and purity of the purified proteins. The general Western blot analysis using a near-infrared detection method. Mix the lysate sample 1:1 (v/v) with 1x treatment buffer. Vortex completely and heat at 95°C for 1.5-3 min (depending on the volume of the samples). Briefly centrifuge to collect the samples at the bottom, which is ready for SDS-PAGE. Minimize loading time to reduce the diffusion of samples to avoid distortion of bands. Choose the percentage of the gel depending on the molecular weight of target proteins. Pre-wet the filter paper and nitrocellulose membrane before wet transfer. Proteins on the gel were transferred to a nitrocellulose membrane at 300 mA for 120 minutes at 4°C, followed by a blocking step incubation in a blocking buffer (PBS containing 5% BSA) at room temperature for 1 h. Subsequently, the membranes were treated with primary and secondary antibodies. Wash 5 minutes for 5 times with 1x PBST before proceeding to

secondary antibody incubation. Anti-AHR SA210 (1: 5,000) and anti-c-myc (1: 5,000) antibodies were used as primary antibodies diluted in blocking buffer and incubated overnight at 4℃. Secondary antibody incubation (LI-COR IRDye 800 donkey anti-rabbit IgG or donkey anti-mouse, 1: 10,000) was performed in the blocking buffer for 2 h at room temperature. Wash the membrane as previously, and the last time wash with cold PBS to lower the background signal. The LI-COR Odyssey imaging system was used for detecting the results.

5.2.7 RNA extraction. Lyse cells with TRI reagent. Proceed to RNA Purification. Add an equal volume of ethanol (95- 160 100%) directly to one volume sample in TRI. Mix well by inverting and vortex. Load the mixture into a column in a collection tube. Centrifuge 30 sec. Transfer the column into a new collection tube and perform in-column DNase I treatment. Wash the column with 400 µL RNA Wash Buffer. Centrifuge 30 sec and discard the flow through. Prepare DNase I reaction mix in a clean tube: 5 µL DNase I (6 U/µL), 75 µL DNA Digestion Buffer for each tube. Mix by gentle inversion. Add 80 µL of the mix per column directly to the column matrix. Incubate at room temperature for 15 min and then centrifuge for 30 sec. Add 400 µL Direct-zol RNA Prewash to the column and centrifuge for 30 sec. Discard the flow-through and repeat this step. Add 700 µL RNA Wash Buffer to the column and centrifuge for 2 min to ensure complete removal of the wash buffer. Discard the flow-through and transfer the column to an RNase-free tube. Add 50 µL DNase/RNase-free water directly to the column matrix and centrifuge for 1 min. Quantify with nanodrop.

5.2.8 Reverse transcription. Denature the RNA samples and anneal the primers. For each reaction, combine the following components on ice in a RNase-free 0.2 mL PCR tube: 10 µL total reaction volume = 0.4 µL random primer + 100 pg-1 µg (use 0.5-1 µg for lower Cq number) + RNase-free water. Incubate at 65 °C for 2 min and chill on ice (set as 4 °C) for at

least 1 min in the thermocycler. Briefly centrifuge and combine with the following components on ice for one reaction (total volume would be 20 µL): 3 µL RNase-free water, 2 µL 10x MMLV buffer, 2 µL 100 mM DTT, 0.5 µL ScriptGuard RNase inhibitor, 2 µL dNTP mix (2.5 mM), 0.5 µL MMLV high performance reverse transcriptase. Use the PCR thermocycler to 161 perform the protocol: 10 min room temperature followed by 37 °C 60 min; terminate the reaction by heating at 85 °C for 5 min and then 4 °C for at least 1 min. Centrifuge briefly. The cDNA can be used immediately for real-time PCR or stored at -20 °C for future use.

5.2.9 Quantitative PCR (qPCR). qPCR was performed with 1ul of cDNA from reverse transcription solution, 10 µL of Bio-Rad iTaq SYBR green Supermix, and 0.8 pmol sequence-specific primers using a Bio-Rad CFX Connect real-time PCR machine with the following protocol: 40 cycles of 90 °C for 10 s/60 °C for 1 min with fluorescence readings taken at 60 °C.

5.3.0 Statistical analysis. GraphPad Prism 9 software (La Jolla, CA, USA) was utilized for statistical analysis.

REFERENCES

Stejskalova L, Dvorak Z, Pavek P. (2011) Endogenous and exogenous ligands of aryl hydrocarbon receptor: current state of art. Curr Drug Metab. 12(2), 198-212.

Linh P. Nguyen and Christopher A. Bradfield. (2008) The Search for Endogenous Activators of the Aryl Hydrocarbon Receptor. Chem. Res. Toxicol. 21, 102-116102.

Wilson, C. L., and Safe, S. (1998) Mechanisms of ligand-induced aryl hydrocarbon receptor-mediated biochemical and toxic responses. Toxicol. Pathol. 26, 657-671.

Barouki R, Coumoul X, Fernandez-Salgueroc P. (2007) The aryl hydrocarbon receptor, more than a xenobiotic-interacting protein. Febs Letters. 581(19), 3608-3615.

Berg P, Pongratz I. (2002). Two parallel pathways mediate cytoplasmic localization of the dioxin (aryl hydrocarbon) receptor. Biological Chemistry. 277(35), 32310-32319.

Zheng Y, Xie J, Huang X, et al. (2016) Binding studies using Pichia pastoris expressed human aryl hydrocarbon receptor and aryl hydrocarbon receptor nuclear translocator proteins. Protein Expr Purif. 122, 72-81.

Xie J, Huang X, Park MS, et al. (2014) Differential suppression of the aryl hydrocarbon receptor nuclear translocator-dependent function by an aryl hydrocarbon receptor PAS-A-derived inhibitory molecule. Biochem Pharmacol. 88(2), 253-265.

Bruno Lamas, Jane M. Natividad & Harry Sokol. (2018) Aryl hydrocarbon receptor and intestinal immunity. Mucosal Immunology. Vol. 11, 1024-1038.

Lucie Larigot, Ludmila Juriceka, Julien Dairoub, Xavier Coumoula. (2018) AhR signaling pathways and regulatory functions. Biochimie Open 7, 1-9.

MacDonald, C. J., Ciolino, H. P. & Yeh, G. C. (2001) Dibenzoylmethane modulates aryl hydrocarbon receptor function and expression of cytochromes P50 1A1, 1A2, and 1B1. Cancer. Res. 61, 3919-3924.

Brigitta Stockinger, Paola Di Meglio, Manolis Gialitakis, and João H. Duarte. (2014) The Aryl Hydrocarbon Receptor: Multitasking in the Immune System. Annu Rev Immunol. Vol. 32, 403-432.

K M Burbach, A Poland, and C A Bradfield. (1992) Cloning of the Ah-receptor cDNA reveals a distinctive ligand-activated transcription factor. Proc Natl Acad Sci. 89 (17), 8185-8189.

Androutsopoulos, V.P., A.M.Tsatsakis, et al. (2009) Cytochrome P450 CYP1A1: wider roles in cancer progression and prevention. BMC Cancer 9: 187.

Chang, J.T., H.Chang, et al. (2007) Requirement of aryl hydrocarbon receptor overexpression for CYP1B1 up-regulation and cell growth in human lung adenocarcinomas. Clin Cancer Res. 13 (1): 38-45.

Koliopanos, A., J.Kleeff, et al. (2002) Increased aryl hydrocarbon receptor expression offers a potential therapeutic target for pancreatic cancer. Oncogene 21 (39): 6059-6070.

J.E. Vorontsova, et al. (2018) Aryl-Hydrocarbon Receptor as a Potential Target for Anticancer Therapy. Biomedical Chemistry, 13 (1): 36-54.

Sun, Y.V., Boverhof, D.R., et al. (2004) Comparative analysis of dioxin response elements in human, mouse and rat genomic sequences. Nucleic Acids Res, 32: 4512-4523.

Conghui Zhu, Qunhui Xie, et al. (2014) The role of AHR in autoimmune regulation and its potential as a therapeutic target against CD4 T cell mediated inflammatory disorder. Int Mol Sci, 15 (6): 10116-35.

Cervantes-Barragan L., Colonna M. (2018) AHR signaling in the development and function of intestinal immune cells and beyond. Semin.Immunopathol. 40 (4): 371-77.

Ehrlich A.K. et al. (2018) TCDD, FICZ, and other high affinity AhR ligands dose-dependently determine the fate of $CD4^+$ T cell differentiation. Toxicol. Sci. 161 (2): 310-20.

Boule L.A. et al. (2018) Aryl hydrocarbon receptor signaling modulates antiviral immune responses: ligand metabolism rather than chemical source is the stronger predictor of outcome. Sci. Rep. 8: 1826.

Bert N. Fukunaga, Markus R. Probst, Suzanne Reisz-Porszasz, and Oliver Hankinson. (1995) Identification of Functional Domains of the Aryl Hydrocarbon Receptor. Biological Chemistry. Vol. 270, No. 49, 29270-29278.

Jones S (2004). An overview of the basic helix-loop-helix proteins. Genome Biology. 5 (6), 226.

Coumailleau P, Poellinger L, Gustafsson JA, Whitelaw ML (1995). Definition of a minimal domain of the dioxin receptor that is associated with Hsp90 and maintains wild type ligand binding affinity and specificity. Biological Chemistry, 270 (42), 25291-300.

Goryo K, Suzuki A, Del Carpio CA, Siizaki K, Kuriyama E, Mikami Y, Kinoshita K, Yasumoto K, Rannug A, Miyamoto A, Fujii-Kuriyama Y, Sogawa K (2007). Identification of amino acid residues in the Ah receptor involved in ligand binding. Biochemical and Biophysical Research Communications. 354 (2), 396-402.

Dalei Wu, Nalini Potluri, Youngchang Kim, and Fraydoon Rastinejad. (2013) Structure and Dimerization Properties of the Aryl Hydrocarbon Receptor PAS-A Domain. Molecular and Cellular Biology. 33, 4346-4356.

J. V. De Souza, S. Reznikov, R. Zhu, A.K. Bronowska (2019). Druggability assessment of mammalian Per-Arnt-Sim [PAS] domains using computational approaches, Medchemcomm. 10: 1126-1137.

S.H. Seok, W. Lee, L. Jiang, K. Molugu, A. Zheng, Y. Li, S. Park, C.A. Bradfield, Y. Xing (2017). Structural hierarchy controlling dimerization and target DNA recognition in the AHR transcriptional complex, Proc. Natl. Acad. Sci. U. S. A. 114, 5431-5436.

K.W. Schulte, E. Green, A. Wilz, M. Platten, O. Daumke (2017). Structural Basis for Aryl Hydrocarbon Receptor-Mediated Gene Activation, Structure. 25, 1025-1033.

Murray, I. A., Reen, R. K., Leathery, N., Ramadoss, P., Bonati, L., Gonzalez, F. J., Peters, J. M., and Perdew, G. H. (2005) Evidence that ligand binding is a key determinant of Ah receptor-mediated transcriptional activity, Arch. Biochem. Biophys. 442, 59-71.

Crosson, S., and Moffat, K. (2001) Structure of a flavin-binding plant photoreceptor domain: Insights into light-mediated signal transduction, Proc. Natl. Acad. Sci. U.S.A. 98, 2995-3000.

Bonati, L., Pandini, A., Pitea, D., Song, Y., and Denison, M. S. (2002) New insights into the structure of the mAhR ligand binding domain, Organohalogen Compd. 59, 429-432.

Alessandro Pandini, Michael S. Denison, Yujuan Song, Anatoly A. Soshilov, and Laura Bonati. (2007) Structural and Functional Characterization of the Aryl Hydrocarbon Receptor Ligand Binding Domain by Homology Modeling and Mutational Analysis. Biochemistry 2007, 46, 696-708.

Colin A. Flaveny and Gary H. Perdew. (2009)Transgenic Humanized AHR Mouse Reveals Differences between Human and Mouse AHR Ligand Selectivity. Mol Cell Pharmacol. 1(3), 119-123.

K M Dolwick, J V Schmidt, L A Carver, H I Swanson and C A Bradfield. (1993) Cloning and expression of a human Ah receptor cDNA. Molecular Pharmacology. 44, 911-917.

Ming Qi Fan, Alex R. Bell, David R Bell, et al. (2009) Recombinant expression of Aryl Hydrocarbon Receptor for quantitative ligand-binding analysis. Anal Biochem. 384(2), 279-287.

W K Chan, R Chu, S Jain, J K Reddy and C A Bradfield. (1994) Baculovirus expression of the Ah receptor and Ah receptor nuclear translocator. Evidence for additional dioxin responsive element-binding species and factors required for signaling. Biological Chemistry. Vol. 269, No. 42, 26464-26471.

Higgins D R. (2001) Overview of protein expression in Pichia pastoris. Curr Protoc Protein Sci. Chapter 5: Unit5-7.

Higgins D R, Cregg J M. (1998) Introduction to Pichia pastoris. Methods Mol Biol. 103, 1-15.

Lundblad R L. (1999) Glycosylation in Pichia pastoris. Biotechnol Appl Biochem. 30, 191-2.

Bretthauer R K, Castellino F J. (1999) Glycosylation of Pichia pastoris-derived proteins. Biotechnol Appl Biochem. 30, 193-200.

Yujuan Zheng, JinghangXie, Xin Huang, Jin Donga, Miki S. Parka, William K. Chan. (2016) Binding studies using Pichia pastoris expressed human aryl hydrocarbon receptor and aryl hydrocarbon receptor nuclear translocator proteins. Protein Expression and Purification. Vol 122, 72-81.

Lebendiker M, Danieli T. (2011) Purification of proteins fused to maltose-binding protein. Methods. Mol Biol. 681: 281-93.

Poland A, Glover E, Kende AS. (1976) Stereospecific, high affinity binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin by hepatic cytosol. Evidence that the binding species is receptor for induction of aryl hydrocarbon hydroxylase. J Biol Chem. 25; 251(16): 4936-46.

Colin Flaveny, Rashmeet K. Reen, Ann Kusnadi, Gary H. Perdew. (2008) The mouse and human Ah receptor differ in recognition of LXXLL motifs. Archives of Biochemistry and Biophysics, 471(2): 215-223.

John B. Hogenesch, William K. Chan, Victoria H. Jackiw, R. Clark Brown, Yi-Zhong Gu, Marilyn Pray-Grant, Gary H. Perdew, Christopher A. Bradfield. (1997) Characterization of a Subset of the Basic-Helix-Loop-Helix-PAS Superfamily That Interacts with Components of the Dioxin Signaling Pathway. Journal of Biological Chemistry, 272 (13): 8581-8593.

Manchester, D. K., Gordon, S. K., Golas, C. L., Roberts, E. A., & Okey, A. B. (1987) Ah receptor in human placenta: stabilization by molybdate and characterization of binding of 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin, 3-methylcholanthrene, and benzo (a) pyrene. Cancer research,47(18), 4861-4868.