Isolation and Functional Characterization of a Dioxin-Inducible CYP1A Regulatory Region From Zebrafish (Danio rerio)

by

Gary T. ZeRuth

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy
Department of Biology
College of Arts and Sciences
University of South Florida

Major Professor: Richard S. Pollenz, Ph.D.
Brian T. Livingston, Ph.D.
Jessica L. Moore, Ph.D.
Robert L. Potter, Ph.D.

Date of Approval: April 11, 2008

Keywords: arylhydrocarbon receptor (AHR), cytochrome p4501A1 (CYP1A1), 2,3,7,8,-tetrachlorodibenzo-p-dioxin (TCDD), xenobiotic response element (XRE), gene regulation

© Copyright 2008, Gary T. ZeRuth
Dedication

Dedicated to Melissa S. ZeRuth, without whom I could never have come so far.
Acknowledgements

Acknowledgements are given to Jeff Yoder for help in the initial screening of the PAC library and helpful discussion of this work. Thanks are extended to Leonard Zon for providing the PAC library and arraying the pooled clones. Michael Carvan is recognized for providing p-1897Om1A3luc reporter vector. Jeannette Wentworth is acknowledged for technical help on the initial screening of the PAC library and the Southern blotting experiments. Special thanks to Dr. Richard S. Pollenz for mentoring and support. Acknowledgments are given to the Pollenz Lab members, Edward J. Dougherty and Sarah E. Wilson for consultation, assistance and support.
### Table of Contents

List of Figures iii

List of Acronyms v

Abstract vi

Chapter One: Introduction 1

Isolation of AHR and ARNT 1

bHLH-PAS Proteins 3

AHR Ligands 5

AHR-mediated Signaling 8

CYP1A1 Regulation 13

AHR Regulated Genes 22

TCDD-mediated Developmental Toxicity in Zebrafish 26

AHR and ARNT in Fishes 28

Chapter Two: Results 32

Isolation of the zfCYP1A upstream region 32

Induction of the zfCYP1A gene by TCDD 35

*In vitro* analyses of AHR/ARNT association with zfXREs 42

Functional characterization of individual XREs within the zfCYP1A regulatory region 47

The pattern of expression controlled by the zfCYP1A and mCYP1A1 regulatory regions varies in different cell lines 56

Analysis of the zebrafish CYP1A proximal promoter 60

Functional analysis of XRE flanking sequence 62

Identification of additional *cis*-regulatory regions which impact the induction of zfCYP1A 67

Chapter Three: Discussion of Impact, Relevance, and Future Direction 73

Chapter Four: Materials and Methods 87

References 97

Appendices 123

Appendix A: PAC 133 Sequencing data 123

Appendix B: Construct Maps 131

Appendix C: Additional Studies 147
Generation of destabilized EGFP and EYFP constructs 147
ARNT domain Swapping 153

About the Author

End Page
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Comparison of bHLH-PAS Protein Structures</td>
<td>6</td>
</tr>
<tr>
<td>1.2</td>
<td>Common AHR Ligands</td>
<td>9</td>
</tr>
<tr>
<td>1.3</td>
<td>Schematic Overview of AHR-mediated signaling</td>
<td>14</td>
</tr>
<tr>
<td>1.4</td>
<td>Model of CYP1A1 Regulation by AHR</td>
<td>21</td>
</tr>
<tr>
<td>2.1</td>
<td>Comparison of CYP1A Regulatory Regions from Different Species</td>
<td>34</td>
</tr>
<tr>
<td>2.2</td>
<td>Induction of the zfCYP1A gene by TCDD</td>
<td>36</td>
</tr>
<tr>
<td>2.3</td>
<td>Analysis of the Effect of Orientation and Position on TCDD-induced Luciferase Activity</td>
<td>39</td>
</tr>
<tr>
<td>2.4</td>
<td>Luciferase Reporter Analysis of the Ability of the Region Between -580 and +71 to Function as a Promoter</td>
<td>41</td>
</tr>
<tr>
<td>2.5</td>
<td>Analysis of TCDD-mediated Induction from Reporter Constructs Containing the zfCYP1A Promoter Region</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Between -580 and +71 and CYP1A Enhancer Regions from Different Species</td>
<td></td>
</tr>
<tr>
<td>2.6</td>
<td><em>In vitro</em> Analysis of the Association of zfAHR2 with XREs</td>
<td>45</td>
</tr>
<tr>
<td>2.7</td>
<td>Association of zfAHR2 and rtARNTb with zfXREs</td>
<td>46</td>
</tr>
<tr>
<td>2.8</td>
<td>Comparison of TCDD-induced Luciferase Activity Between the p-2699/+71 and p-2608/-2100Ur Constructs</td>
<td>49</td>
</tr>
<tr>
<td>2.9</td>
<td>Functional Analysis of zfXREs 1 and 3</td>
<td>51</td>
</tr>
<tr>
<td>2.10</td>
<td>Functional Analysis of zfXREs</td>
<td>53</td>
</tr>
<tr>
<td>2.11</td>
<td>Functional Analysis of zfXREs in Different Cell Lines</td>
<td>55</td>
</tr>
<tr>
<td>2.12</td>
<td>Comparison of XRE8 Function Between the p-2608/-2100Ur and p-2727/-2100Ur Constructs</td>
<td>57</td>
</tr>
<tr>
<td>2.13</td>
<td>Characterization of Mouse and Zebrashish CYP1A Promoter and Enhancer Regions in Various Cell Lines</td>
<td>58</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>AHR</td>
<td>arylhydrocarbon receptor</td>
<td></td>
</tr>
<tr>
<td>ARNT</td>
<td>arylhydrocarbon receptor nuclear translocator</td>
<td></td>
</tr>
<tr>
<td>XRE</td>
<td>xenobiotic response element</td>
<td></td>
</tr>
<tr>
<td>TCDD</td>
<td>2, 3, 7, 8 – tetrachlorodibenzo-p-dioxin</td>
<td></td>
</tr>
<tr>
<td>3MC</td>
<td>3-methylcholanthrene</td>
<td></td>
</tr>
<tr>
<td>BaP</td>
<td>benzo[a]pyrene</td>
<td></td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
<td></td>
</tr>
<tr>
<td>PAH</td>
<td>polycyclic aromatic hydrocarbon</td>
<td></td>
</tr>
<tr>
<td>HAH</td>
<td>halogenated aromatic hydrocarbon</td>
<td></td>
</tr>
<tr>
<td>CYP1A1</td>
<td>Cytochrome P4501A1</td>
<td></td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
<td></td>
</tr>
<tr>
<td>EMSA</td>
<td>electrophoretic mobility shift assay</td>
<td></td>
</tr>
<tr>
<td>siRNA</td>
<td>small interfering RNA</td>
<td></td>
</tr>
<tr>
<td>HNF-3α</td>
<td>hepatic nuclear factor 3α</td>
<td></td>
</tr>
<tr>
<td>CREB</td>
<td>cAMP response element binding protein</td>
<td></td>
</tr>
<tr>
<td>Sp1</td>
<td>stimulatory protein 1</td>
<td></td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
<td></td>
</tr>
<tr>
<td>XME</td>
<td>xenobiotic metabolizing enzyme</td>
<td></td>
</tr>
<tr>
<td>ARE</td>
<td>anti-oxidant response element</td>
<td></td>
</tr>
</tbody>
</table>
Isolation and Functional Characterization of a Dioxin-Inducible CYP1A Regulatory Region From Zebrafish (*Danio rerio*)

Gary T. ZeRuth

ABSTRACT

Cytochrome P4501A1 (CYP1A1) is a phase I bio-transformation enzyme involved in the metabolism of xenobiotics via the oxygenation of polycyclic aromatic hydrocarbons (PAHs) including the carcinogen, benzo(*a*)pyrene. Induction of the CYP1A1 gene is regulated at the transcriptional level and is ligand dependent with the prototypical 2,3,7,8,-tetrachlorodibenzo-*p*-dioxin (TCDD) being the most potent known inducer of CYP1A1 transcription. This process is mediated by the AHR/ARNT signaling pathway whereby ligand binds AHR in the cytoplasm allowing its translocation to the nucleus where it binds with its hertrodimericization partner, ARNT and subsequently binds DNA at cognate binding sites termed xenobiotic responsive elements (XREs) located in the 5’ flanking region of the CYP1A1 and other genes.

The zebrafish (*Danio rerio*) has recently become an important model system for the study of TCDD-mediated developmental toxicity due to their relative ease of maintaining and breeding, external fertilization, abundant transparent embryos, and sensitivity to TCDD similar to mammalian models. It is therefore essential to
characterize the molecular mechanisms of AHR mediated gene regulation in this organism.

The upstream flanking region of a putative CYP1A gene from zebrafish was identified by the screening of a PAC genomic library. Sequencing revealed a region which contains 8 putative core xenobiotic response elements (XREs) organized in two distinct clusters. The region between –580 to –187 contains XRE 1-3 while the region between –2608 to –2100 contains XRE 4-8. Only XRE 1, 3, 4, 7, and 8 exhibited TCDD-dependant association of AHR/ARNT complexes when evaluated by gel shift assays. The use of \textit{in vitro} mutagenesis and Luciferase reporter assays further showed that only XRE’s 4, 7, and 8 were capable of conveying TCDD-mediated gene induction. The role of nucleotides flanking the core XRE was investigated through the use of EMSA and reporter assays. Similar methods were employed on additional transcription factor binding sites identified by \textit{in silico} analyses revealing two sites conforming to an HNF-3α and CREB motif, respectively, which demonstrate importance to regulation of the gene.
Chapter One

Introduction

Isolation of AHR and ARNT

The aryl hydrocarbon hydroxylase, (AHH), was identified as being greatly varied in inducibility amongst different strains of mice in response to polycyclic aromatic hydrocarbons, (PAHs), and halogenated aromatic hydrocarbons, (HAHs) (Nebert and Gelboin, 1969; Poland et al., 1974). Through the use of classical murine genetics, it was further ascertained that these variations were controlled by a single gene locus termed $Ah$ (Green, 1973; Schmidt and Bradfield, 1996; Thomas and Hutton, 1973). The $Ah$ locus was later found to code for a receptor, the aryl hydrocarbon receptor (AHR), which is capable of binding ligand with high affinity leading to the subsequent induction of AHH (Poland et al., 1976). Polymorphisms in the $Ahr$ were found to be the cause of the variable inducibility of AHH between mouse strains and $[^{125}\text{I}]-$photoaffinity labeling led to the discovery of three alleles encoding for high affinity receptors designated $Ahr^{b-1}$, $Ahr^{b-2}$, and $Ahr^{b-3}$ and a single allele encoding the low affinity $Ahr^{d}$ allele. (Poland and Glover, 1990; Poland et al., 1987) Purification of the AHR and the development of antibodies specific to the receptor revealed that differences existed not only amongst mouse strains, but the molecular weight varied remarkably between species as well (Hahn et al., 1994; Poland and Glover, 1987; Poland et al., 1991). These innovations
additionally led to the generation of the first AHR cDNA and the revelation that the receptor was the second member of the basic-helix-loop-helix Per-ARNT-Sim (bHLH-PAS) family of proteins identified (Burbach et al., 1992; Ema et al., 1992).

Interestingly, the first member of the bHLH-PAS family isolated was the heterodimerization partner of the AHR, the aryl hydrocarbon nuclear translocator protein, ARNT, cloned approximately one year earlier (Hoffman et al., 1991). ARNT was isolated in an attempt to identify constituents of AHR signaling absent in an induction defective mouse hepatoma cell line containing normal concentrations of AHR capable of binding ligand but which failed to subsequently localize to the nucleus. (Hoffman et al., 1991; Legraverend et al., 1982) Due to the ability of ARNT to restore function to the loss-of-function mutants, the name aryl hydrocarbon nuclear translocator protein was actually a misnomer as it was later found that ARNT is not required for the nuclearization of liganded AHR but instead acts as a dimerization partner required for the binding of DNA. (Dolwick et al., 1993b; Elferink et al., 1990; Reyes et al., 1992) Indeed, studies by Pollenz et al. (1994) showed that AHR was capable of translocating to the nucleus following TCDD treatment in ARNT deficient cells and that ARNT was confined to the nucleus both prior to and following ligand exposure while unliganded AHR was primarily cytosolic. (Pollenz et al., 1994) The unveiling of the scheme of heterodimerization between AHR and ARNT laid the foundation for the identification and characterization of novel members of the emerging super-family of bHLH-PAS proteins.
bHLH-PAS Proteins

While other proteins had been previously identified which contain bHLH domains, including the well characterized MyoD, AHR and ARNT possess a domain adjacent to the bHLH which shows homology to the *Drosophila Period* (Per) and *Single minded* (Sim) genes and is termed Per-ARNT-Sim (PAS) after the identifying members (Burbach et al., 1992; Ema et al., 1992; Hoffman et al., 1991; Huang et al., 1993). The PAS region is typically composed of approximately 300 amino acids and is divided into PAS A and PAS B subdomains each consisting of 50 amino acid degenerate repeats (Burbach et al., 1992; Gu et al., 2000; Nambu et al., 1991). Having a role in dimerization, the PAS domain has been implicated in heterodimerization between Per and Sim and homodimerization of Per (Huang et al., 1993). In the case of AHR, the PAS domains serve as a binding site for chaperone proteins (Whitelaw et al., 1993) and ligand (Burbach et al., 1992; Dolwick et al., 1993a; Schmidt and Bradfield, 1996; Whitelaw et al., 1993) as well as assisting in DNA binding (Dolwick et al., 1993a) while AHR/ARNT dimerization occurs concertedly within the HLH and PAS domains (Reisz-Porszasz et al., 1994; Schmidt and Bradfield, 1996). In addition to the basic, DNA binding domain, the HLH, and the PAS domains, the majority of bHLH-PAS proteins contain a C-terminal transcriptional activation domain (TAD). Unlike the previous domains, however, which share sequence homology, the TADs lack conservation amongst members of the superfamily (Gu et al., 2000). A graphical depiction of the domain organization within selected bHLH-PAS proteins is shown in Figure 1.1.
The existence of transcriptional activation domains supports the fact that most bHLH-PAS proteins serve as transcriptional regulators which act upon target genes in order to elicit an adaptive response to environmental stimuli. (Furness et al., 2007; Kewley et al., 2004; Massari and Murre, 2000) In order to form active transcription factors, bHLH-PAS proteins are required to dimerize and are thus divided into two distinct classes. Class I proteins are capable of detecting environmental stimuli but must dimerize with a Class II protein in order to adopt an active, DNA binding form. While Class I proteins are incapable of homodimerization or dimerization with another Class I member, Class II bHLH-PAS proteins are far more promiscuous and confined to the nuclear compartment where they serve as master regulators capable of homo or heterodimerization. Some examples of Class I proteins include HIF1 and HIF2α (hypoxia inducible factors); regulators of the cellular response to hypoxia, (Tian et al., 1997; Wang et al., 1995; Wang and Semenza, 1995) SIM 1 and 2 (single minded proteins); involved in neurogenesis and mid-line development, (Ema et al., 1996; Probst et al., 1997) the circadian rhythm protein, Clock, (Gekakis et al., 1998; King et al., 1997) and the AHR (aryl hydrocarbon receptor); involved in xenobiotic metabolism. Class II receptors include ARNT and ARNT2 (aryl hydrocarbon receptor translocator proteins), BMAL1 and 2 (brain and muscle ARNT-like proteins), and Per. Figure 1.1 shows the domain structures of representative bHLH-PAS proteins. Of all the bHLH-PAS proteins, AHR and ARNT remain the best characterized members.

Unlike the Class I AHR which only binds ARNT, ARNT exhibits promiscuity by dimerizing with AHR, hypoxia inducible factors HIF1α and HIF2α, and Single mired
proteins Sim1 and Sim2 (Furness et al., 2007; Kewley et al., 2004). ARNT -/- mice die in utero at gestational day (GD) 10.5 due to a failure of the placenta to vascularize. Other defects observed include forebrain hypoplasia, placental hemorrhaging, visceral arch and neural tube abnormalities, and delayed rotation of the embryo (Kozak et al., 1997). Phenotypically similar to HIF1α knockout mice, (Ke and Costa, 2006) these findings implicate ARNT as a compulsory, HIF associated, developmental transcription factor.

ARNT2 is a close structural homolog of ARNT, bearing 57% sequence similarity in the mouse with divergence primarily within the COOH-terminus (Hirose et al., 1996). Furthermore, unlike the ubiquitously expressed ARNT, ARNT2 is expressed primarily in the CNS and kidneys in mice (Jain et al., 1998). ARNT2 -/- mice perish perinatally bearing a phenotypic resemblance to SIM-1 knock-out mice suggesting that ARNT2 may be the heterodimerization partner of SIM-1 required for neurogenesis. Data also supports the hypothesis that ARNT or ARNT2 may have overlapping function prior to embryonic day (ED) 8.5. While ARNT2 can form dimers with AHR and HIF1α, its primary function appears to be as a pairing partner for SIM (Jain et al., 1998).

**AHR Ligands**

Unlike the Class II ARNTs which primarily serve as dimerization partners for the Class I bHLH-PAS proteins, the AHR serves as a “sensor” of environmental cues; therefore ligand activation must ensue to initiate the pathway. Ligand binds the AHR in the form of a structurally diverse array of chemicals both natural and synthetic in nature which are capable of activating the receptor subsequently leading to the regulation of a
**Figure 1.1.** Comparison of bHLH-PAS Protein Structures

Fig. 1.1. **Comparison of bHLH-PAS Protein Structures.** Schematic overview of the domain structure of representative bHLH-PAS proteins. The domains are indicated. b: Basic region, HLH: Helix-loop helix domain, PAS A and PAS B: Per-ARNT-Sim motifs, TAD: Transactivation domain. The scale at the bottom represents number of amino acids.
battery of genes. Representative AHR ligands are shown in Figure 1.2. The most well characterized AHR ligands are the polycyclic aromatic hydrocarbons (PAHs) such as benzo[a]pyrene and 3-methylcholanthrene and the halogenated aromatic hydrocarbons (HAHs) represented by dibenzo-\(p\)-dioxins, biphenyls, and dibenzofurans. Environmental pollutants, exposure to these chemicals contributes to a broad spectrum of toxic and biological effects.

PAHs typically exist as the product of combustion and can lead to cancers due to the generation of ROS (reactive oxygen species) resulting from the metabolism of the insulting compound and consequent DNA and protein adduct formation and cellular damage (Gelboin, 1980). While HAHs also lead to toxicological responses in mice which include wasting syndrome, thymic involution, tumor promotion, teratogenicity, immunosuppression, reduced fertility, epidermal hyper and metaplasia, and death, (Poland and Knutson, 1982; Safe, 1990) unlike PAHs, the molecular mechanism behind these effects are not understood but most, if not all, of these responses rely on the AHR. The most potent known agonist of AHR is 2,3,7,8-tetrachlorodibenzo-\(p\)-dioxin, (dioxin, TCDD), which has an affinity for the receptor in the pM range versus PAHs which exhibit binding affinities in the nM to \(\mu\)M range. The toxic effects of TCDD are thought to stem from the regulation of genes targeted by the activated AHR as opposed to direct genotoxicity of the compound or its metabolic by-products. Interestingly, the severity of the response to TCDD depends on multiple factors including the type of cell, sex, age, and species exposed supporting a gene regulatory mechanism of toxicity over that of direct cellular damage.
To date, no endogenous AHR ligand has been identified however a number of naturally occurring, non-anthropogenic chemicals have been shown to bind the AHR and induce target genes, albeit much more weakly than TCDD. A variety of dietary plant derivatives consisting largely of flavonoids have been identified as weak AHR agonists/antagonists and may explain an evolutionary purpose for the AHR as an activator of xenobiotic metabolism enzymes (Denison and Nagy, 2003). The activation of AHR in the absence of exogenous ligand as well as numerous developmental defects in AHR null mice suggest that an endogenous, unidentified AHR ligand likely exists and may possibly be in the form of an indole, tetrapyrole, or an amino acid metabolite (Denison and Nagy, 2003). While natural or endogenous AHR ligands have yet to be confirmed, in regard to the elucidation of the molecular mechanisms behind AHR signaling and the impact on human health, activation by the classical PAHs and HAHs are relevant. The precise risk to human health is still unknown but TCDD was upgraded to a Group 1 “human carcinogen” in 1997 by the International Agency for Research on Cancer and remains the prototypical AHR ligand.

**AHR-mediated Signaling**

Ligand binding to the AHR is thought to be supported by the fact that the latent AHR exists as a complex with several other proteins consisting of the 90kD heat shock protein, Hsp90, the 23kD p23, and the Hepatitis B virus X-associated protein 2, XAP2 (Carver et al., 1998; Denis et al., 1988; Kazlauskas et al., 1999; Ma and Whitlock, 1997; Meyer et al., 1998; Perdew, 1988). Immunoprecipitation experiments from two independent laboratories first identified interactions between AHR and a dimer of the
**FIGURE 1.2. Common AHR Ligands**

2,3,7,8-tetrachlorodibenzo-\(p\)-dioxin

2,2’,4,4’,5-pentachlorobiphenyl

2,3,7,8-tetrachlorodibenzo-\(p\)-furan

benzo[a]pyrene

3-methylcholanthrene

\(\beta\)-naphthoflavone

Curcumin

Indole-3-carbinol

**Fig. 1.2. Common AHR ligands.** Chemical structures of representative AHR ligands and inducers. Classical halogenated aromatic hydrocarbons are shown at the top. Classical polycyclic aromatic hydrocarbons are shown in the middle. Naturally occurring dietary AHR ligands are shown at the bottom.
molecular chaperone, Hsp90 (Denis, 1988; Perdew, 1988). The role of Hsp90 is thought to be one in which it stabilizes the AHR in a conformation susceptible to ligand binding to the PAS domain of AHR (Antonsson et al., 1995; Carver et al., 1994; Coumailleau et al., 1995; Denis, 1988; Perdew, 1988; Pongratz et al., 1992) and to repress dimerization with ARNT and DNA binding in the absence of ligand and an additional unidentified event (Heid et al., 2000; McGuire et al., 1994; Pongratz et al., 1992). Indeed, it has been established that liganded AHR enters the nucleus along with Hsp90 but dimerization with ARNT and DNA binding require dissociation from the heat shock protein (Heid et al., 2000).

In addition to Hsp90, the latent AHR is also complexed with a molecule of the FKBP52 immunophilin-like, Hepatitis B virus X-associated protein (XAP2, Ara9, AIP1) (Carver and Bradfield, 1997; Ma and Whitlock, 1997). Originally it was ascertained that XAP2 maintained the cytoplasmic localization of AHR, thus enhancing its ability to be activated by ligand (Bell and Poland, 2000; Berg and Pongratz, 2002; Carver et al., 1998; Kazlauskas et al., 2000; LaPres et al., 2000; Ma and Whitlock, 1997; Meyer and Perdew, 1999; Meyer et al., 2000; Petrulis et al., 2003) however, recent studies have shown that these observations may be specific to the Ah<sup>b-1</sup> allele and that other species do not exhibit association with XAP2 to the level observed with the b-1 receptor nor does XAP2 maintain a cytoplasmic localization, but instead merely inhibits nucleocytoplasmic shuttling by interfering with the association of nuclear import receptors (Pollenz and Dougherty, 2005; Pollenz et al., 2006). The lack of necessity for XAP2 in AHR
signaling is supported by the fact that AHR exhibits normal functioning in *Saccharomyces cerevisiae* which does not possess a homolog of XAP2 (Gu et al., 2000).

The third protein known to exist in the heterotetrameric latent AHR complex is p23 which is known to interact with Hsp90 in other systems (Chadli et al., 2000; Grenert et al., 1997; Sullivan et al., 1997). Studies involving the yeast homologs of Hsp90 and p23, Hsp82 and Sba1 respectively, suggest that p23 blocks the ATPase activity of Hsp90, stabilizing the Hsp90-AHR interaction (Cox and Miller, 2004). Earlier studies confirmed this role for p23 when loss the protein from the AHR-Hsp90 complex resulted in ligand-independent interaction between AHR and ARNT. The addition of molybdate, a chemical known to stabilize p23-Hsp90 interactions, restored normal function suggesting that a role of p23 is to stabilize the Hsp90-AHR latent complex (Kazlauskas et al., 1999). The association of p23 with Hsp90 along with the observation that the AHR may associate with p60 and Hip indicates that the latent AHR complex may be similar to that seen in other steroid hormones (Nair et al., 1996).

The proposed model for AHR signaling then follows that a small, hydrophobic ligand, typified by TCDD, passes through the plasma membrane where it binds to the latent AHR complex within the PAS domain of AHR between amino acids 232-334 (Burbach et al., 1992). The binding of ligand is supported by the folding of the AHR PAS domain into a favorable lignd binding conformation due to its association with Hsp90 (Gu et al., 2000). Binding of ligand presumably causes a conformational change both displacing p23 and allowing for the nuclear translocation of the complex. Within the nucleus, it is suspected that AHR, due in part to the displacement of p23 by ligand,
can form a dimer with ARNT, possibly after phosphorylation of ARNT (Chen and Tukey, 1996; Long et al., 1998). AHR-ARNT dimerization leads to the dissociation of the molecular chaperone proteins and the formation of an active, DNA binding complex. The activated AHR-ARNT recognizes cognate enhancer sequences termed xenobiotic response elements (XREs) located within the regulatory region of target genes. The XRE core consensus sequence is defined as: 5’-(T/G)NGCGTG-3’ whereby the basic region of ARNT binds to 5’-GTG and the AHR basic region binds the remaining nucleotides (Bacsi et al., 1995; Denison et al., 1988; Fisher et al., 1990; Hapgood et al., 1989; Lusska et al., 1993; Swanson et al., 1995) and the specificity of binding may be controlled, in part, by the PAS domain of AHR (Dolwick et al., 1993b). Binding of AHR-ARNT to the XREs of target genes results in gene regulatory events which are largely dependent on the COOH-terminal TAD domains of AHR and ARNT (Jain et al., 1994; Whitelaw et al., 1994).

The endpoint of ligand-mediated AHR signaling is the degradation of the AHR. Indeed, studies have shown that AHR is rapidly depleted in both cell culture and animals following TCDD exposure, (Giannone et al., 1995; Ma et al., 2000; Pollenz, 1996; Pollenz, 2002; Pollenz et al., 1998; Roman et al., 1998) the degradation event is connected to the nuclear localization of the AHR, (Song and Pollenz, 2002) and likely occurs via the 26S proteosome (Davarinos and Pollenz, 1999; Ma and Baldwin, 2000; Wentworth et al., 2004). The mechanism by which the AHR is degraded is yet unknown but it is suspected that ubiquitination is involved considering most proteins targeted to the proteosome are ubiquitinated. To date, no evidence of ubiquitination has been identified.
CYP1A1 Regulation

A member of the P450 family of cytochromes, the aryl hydrocarbon hydroxylase or Cytochrome P₄₅₀ 1A1 (CYP1A1) is the most well characterized target of ligand-activated AHR and has become the prototype for the study of AHR-mediated signaling. The product of the inducible CYP1A1 locus is a heme-thiolate monooxygenase responsible for the metabolism of lipophilic aromatic hydrocarbons. These xenobiotic metabolizing enzymes receive electrons from NADPH-P450 reductase which activates an oxygen capable of being inserted into a specific substrate or group of substrates. In the case of CYP1A1, these substrates are planar aromatic hydrocarbons and the monooxygenation event of CYP1A1 opens the benzene rings of the PAH allowing for its subsequent metabolism. Ligands of the AHR, PAHs then induce their own metabolism. Due to the toxic 2,3,7,8-chlorination of TCDD and similar HAHs, these persistent chemicals are poorly metabolized by xenobiotic metabolism enzymes yet induce transcription of XMEs via the activation of AHR nonetheless.

Early experiments revealed that TCDD-mediated CYP1A1 induction did not occur in AHR or ARNT deficient cells (Jones et al., 1986) and was a primary response which occurs in the absence of protein synthesis (Whitlock, 1999). Further investigation demonstrated that TCDD-mediated induction of the gene was controlled by cis-regulatory elements contained within ~500bp of the 5’ upstream region of the gene in mice (Jones et al., 1985) which functioned as an enhancer up or downstream of an MMTV promoter regardless of orientation (Jones et al., 1986). Several features of the enhancer were
Figure 1.3. Schematic Overview of AHR-mediated Signaling. Diagram showing ligand-dependant aryl hydrocarbon receptor signaling. Ligand, represented by TCDD, enters the cell and binds to the latent, cytoplasmic AHR complex. The activated AHR translocates to the nucleus where it dimerizes with ARNT, dissociates from the chaperone proteins, and binds DNA at XREs located upstream of target genes. AHR is targeted for degradation following DNA binding, likely via the 26S proteosome.
subsequently characterized. Using various truncations of the rat CYP1A1 enhancer region upstream of a CAT reporter gene driven by an SV-40 promoter and gel retardation assays, two regions composed of 15 nucleotides which were required for gene induction following treatment with 3-MC were isolated. Comparisons of the two regions revealed a 5 bp sequence, GCGTG, common between the two required regions which were designated “xenobiotic responsive elements” or “XREs” (Fujisawa-Sehara et al., 1987). Experiments with the murine CYP1A1 enhancer region expanded the protein-DNA interaction to a core sequence, 5’-T(A/T)GCGTG-3’. This sequence was specifically recognized by AHR/ARNT heteromers but showed that binding of the 7bp core sequence alone by the liganded receptor was not sufficient to drive activation of the downstream gene; indicating that the sequence flanking the XRE core is essential for transactivation (Denison et al., 1988; Hapgood et al., 1989). Subsequent studies attempting to characterize the role flanking nucleotides play in binding AHR/ARNT in vitro and in transactivating a downstream reporter produced a modified consensus: 5’-
(T/G)\textsubscript{1}YG\textsubscript{2}C\textsubscript{3}G\textsubscript{4}T\textsubscript{5}G\textsubscript{6}(A/C)\textsubscript{7}(C/G)\textsubscript{8}(A/T)\textsubscript{9} -3’ (Lusska et al., 1993; Shen and Whitlock, 1992; Swanson et al., 1995). Indeed, Shen and Whitlock showed that either a C at position 1, G at position 7, A at position 8, or G at position 9 abolished XRE function in CAT reporters (Shen and Whitlock, 1992). These observations are in agreement with studies done by Swanson et al (Swanson et al., 1995). Significantly, experiments which placed tandem repeats of an XRE upstream of an SV-40 controlled CAT gene showed a 4-fold increase in both induced and constitutive CAT activity in constructs containing two XREs versus a single XRE (Fujisawa-Sehara et al., 1987). It is also of importance that either linker scanning mutants, whereby the core XRE was replaced with unrelated
DNA of the same length, or truncations of the enhancer revealed that the elimination of any one of four mouse XREs resulted in a 25% decrease in TCDD induced activity. Removal of all four XREs eliminated responsiveness to TCDD (Fisher et al., 1990). Additionally, another important component of the CYP1A1 enhancer, a GC-box was identified between -952 and -943 upstream of the mouse CYP1A1 transcriptional start site which was capable of being bound in vitro by Sp1 (or a related factor) and when removed, produced a 5-fold decrease in reporter activity. Interestingly, this site is incapable of function in the absence of bound XREs (Fisher et al., 1990).

The enhancer region defined above is non-functional unless linked to a functional transcriptional promoter (Jones et al., 1986; Neuhold et al., 1989). Likewise, the murine CYP1A1 promoter contains several regulatory elements which fail to function in the absence of the enhancer. A TATAAA box is located at position -30 upstream of the transcriptional start site, a proximal and a distal CTF/NF1 site are located at positions -59 and -136 respectively, and a G-box is located at position -130 (Jones and Whitlock, 1990). The TATAAA box is an essential component of the CYP1A1 promoter as mutation of this sequence reduces gene activation by >80%. The distal CTF/NF1 site and the G-box appear to bind a functionally equivalent protein as elimination of either one has no effect on gene activity; however, when both are eliminated a 50% reduction in inducible activity is observed. Furthermore, footprinting experiments reveal that mutation of the G-box shifts protection at that site toward the distal CTF/NF1 site (Jones and Whitlock, 1990). The proximal CTF/NF1 site (identified by others as a BTE site in the rat CYP1A1 promoter) also contributes significantly to promoter function (Jones and
Whitlock, 1990; Yanagida et al., 1990). A number of proteins have been shown capable
of binding this sequence including Sp1, Gut-enriched Kruppel-like factor (GKLF), and
BTEB1, 3, and 4 (Imataka et al. 1992, Sogawa et al. 1993, Zhang et al. 1998, Shields et
al. 1998, Kaczynski et al. 2001, Kaczynski et al. 2002). While Sp1 binding leads to
enhanced activation of the gene, (Yanagida et al. 1990, Kobayashi et al. 1996) GKLF and
BTEB1, 3, and 4 binding leads to transcriptional repression of CYP1A1 (Imataka et al.,
1992; Kaczynski et al., 2002; Sogawa et al., 1993; Zhang et al., 1998b). Kaczynski et al.
have proposed a model whereby the constitutively expressed BTEB proteins repress
CYP1A1 activity by competing with Sp1 at the BTE site but this model has yet to be
confirmed.

In the absence of ligand, the CYP1A1 enhancer/promoter is inactive and assumes
a nucleosomal configuration (Wu and Whitlock, 1992; Wu and Whitlock, 1993). Studies
using DNaseI protection and LMPCR have revealed that binding of AHR/ARNT
heteromers at the XREs within the CYP1A1 enhancer results in a disruption of chromatin
structure localized to approximately 180bp surrounding the XRE, followed by the loss of
the nucleosome at the promoter. The relaxation of the nucleosomal promoter allows the
binding of TBP, NF1, and general transcription factors initiating transcription (Ko et al.,
1997; Morgan and Whitlock, 1992; Okino and Whitlock, 1995). As hundreds of base
pairs remain in a nucleosomal configuration between the XRE containing enhancer and
the promoter, the possibility of direct communication between the two regions is
unlikely. Additional studies determined that the loss of the nucleosome at the promoter
was the result of communication via the TAD of AHR bound at the enhancer (Ko et al.,
1996; Ko et al., 1997). Interestingly, the TAD containing C-terminus of AHR appears to have no effect on the nucleosomal re-arrangement of the enhancer implying that its effect on the promoter involves the recruitment or binding to other factors involved in stabilizing the promoter chromatin (Ko et al., 1996). A likely model is one in which AHR/ARNT binds at the enhancer and associates with a complex of other proteins involved in both chromatin remodeling and stabilization of the general transcription factors at the promoter.

A number of other proteins have been implicated in the regulation of CYP1A1 which may act to either remodel chromatin or stabilize the transcriptional machinery at the promoter. It has previously been established that Sp1 interacts \textit{in vivo} with the HLH-PAS domain of AHR/ARNT via its zinc finger domain (Kobayashi et al., 1996). Interaction between these two proteins lends to a potential model in which a DNA loop is formed by binding of Sp1 at the promoter to AHR/ARNT heteromer at the enhancer. Furthermore, AHR or ARNT have been shown to interact \textit{in vitro} with the general transcription factors TFIIB, IIF, and TBP (Rowlands et al., 1996; Swanson and Yang, 1998) supporting the role of AHR/ARNT in stabilizing the general transcription complex.

Other observed interactions include the HAT co-activator, CBP, interacting at the transactivation domain of ARNT (Kobayashi et al., 1997) and RIP-140, retinoblastoma protein (Rb), Nedd8, and promyelocytic leukemia nuclear bodies (PML) interacting with AHR (Fujii-Kuriyama and Mimura, 2005; Hankinson, 2005), all of which have been shown to enhance reporter gene expression. Recent studies have also demonstrated the involvement of the p-160 HAT coactivators, SRC-1 (NCoA-1), NCoA-2 (GRIP-1, TIF-
2), and p/CIP (AIB, ACTR) in mediating TCDD-dependent CYP1A1 expression. ChIP assays and real time PCR reveal that all three proteins associate with the CYP1A1 enhancer region \textit{in vivo} within 15 minutes of TCDD treatment, while antibodies specific to each reduce XRE-driven expression of reporter genes (Beischlag et al., 2002; Fujii-Kuriyama and Mimura, 2005; Hankinson, 2005; Hestermann and Brown, 2003; Kumar and Perdew, 1999). Overexpression of the coactivators enhances reporter gene activity and shows that all three are capable of interacting with AHR while SRC-1 and NCoA-2 interact with ARNT (Beischlag et al., 2002). Further studies need to be performed to elucidate the precise roles the p-160 family of receptors play in AHR-dependent signaling.

Brahma/SW12-related Gene 1 Protein (Brg-1) is the ATPase subunit of certain ATP-dependent chromatin remodeling complexes and has been shown to associate with the TAD of AHR (Wang and Hankinson, 2002). Overexpression of exogenous Brg-1 enhanced expression of XRE–driven reporters in Hepa-1 cells and restored endogenous CYP1A1 activity in Brg-1 deficient cells when co-expressed with SRC-1 while an ATPase deficient Brg-1 mutant failed to do so. Finally, ChIP analysis demonstrated that Brg-1 associates with the mouse CYP1A1 enhancer region in a TCDD and ARNT dependent manner implicating its role in AHR-mediated induction of the gene (Fujii-Kuriyama and Mimura, 2005; Hankinson, 2005; Wang and Hankinson, 2002).

The TRAP/DRIP/ARC mediator complex has also been shown to be involved in AHR-mediated regulation of CYP1A1. ChIP analyses show that two sub-units of the mediator complex, Med220 and CDK8 associate with the murine CYP1A1 enhancer
shortly after binding of AHR/ARNT and p/CIP (10-30 min). RNAi experiments revealed that depletion of endogenous Med220 resulted in inhibition of endogenous CYP1A1 induction following treatment with TCDD (Wang et al., 2004). While Med220 and CDK8 have been shown to associate with the CYP1A1 enhancer *in vivo*, other subunits of the mediator complex have previously been shown to bind to the general transcriptional complex suggesting a role for mediator to bridge the enhancer and promoter regions of CYP1A1 (Malik and Roeder, 2000).

A hypothetical model for CYP1A1 regulation based on the current literature is shown in Figure 1.4. Additional interactions are currently being evaluated. Indeed, the estrogen receptor alpha (ERα) has recently been in the spotlight for its possible role as yet another cofactor. While an overwhelming battery of coactivators have been implicated in AHR-mediated signaling, it is important to consider that many have suggested roles based on over-expression of proteins to levels far exceeding what a cell would experience in a normal, physiological setting or *in vitro* interactions of artificially expressed proteins. Furthermore, even as new resources such as ChIP analyses and quantitative real-time PCR arise as valuable tools in assessing the proteins involved within this pathway, these procedures are not flawless and could possibly lead to the false implication of elements which, in reality, are not involved. Care must be taken in the analysis of future studies to avoid this. Even as a large number of factors are currently suggested to regulate TCDD-mediated gene induction, it is probable that more will arise.
Fig. 1.4. Model of CYP1A1 Regulation by AHR. Hypothetical schematic overview of AHR-mediated transcriptional regulation of CYP1A1. Numbers indicate proposed events chronologically. Transcription factors and co-activators are shown bound to DNA and proteins respectively. See text and references for additional details.
AHR Regulated Genes

In addition to CYP1A1, ligand activated AHR leads to the induction of a battery of other genes such as Glutathione S-transferase Ya (Gst-Ya), Uridine Diphosphosphate-Glucuronosyl transferase (UGT1A1), Aldehyde dehydrogenase (ALDH), Quinone oxidoreductase (NQO1), and additional members of the P450 family of cytochrome monooxygenases. CYP1A2 and CYP1B1 are both phase I biotransformation enzymes known to be regulated by AHR. CYP1B1 is expressed constitutively in extrahepatic tissues such as the mammary, ovary, and prostate (Shimada et al., 1996; Sutter et al., 1994) and has been implicated in the bioactivation of benzo[a]pyrene and other procarcinogens. Importantly, CYP1B1 has been shown to be regulated by the AHR and contains three XREs within a 190bp span of its promoter region (Tang et al., 1996). Apart from AHR-mediated inducibility and a role in PAH metabolism, CYP1B1 exhibits more differences than similarities to CYP1A1. CYP1B1 is largely expressed in tissues originating from the mesenchyme while CYP1A1 is expressed ubiquitously. CYP1B1 is constitutively expressed while CYP1A1 generally shows little to no basal activity in the absence of liganded AHR. Differences also exist structurally in that CYP1A1 consists of seven exons, like most other P450s, while CYP1B1 consists of only three exons. While the regulation of CYP1B1 is poorly understood, it is known that the mouse and human CYP1B1 promoters lack a TATA box, CTF/NF1 sites, or BTE sites as are found in the CYP1A1 regulatory region. Instead, the gene is under the control of a TATA-like sequence located at position -27 relative to the transcriptional start site and a series of Sp1 sites located within the proximal promoter (Wo et al., 1997). Intriguingly,
AHR/ARNT has been shown to bind at only one of the three CYP1B1 XREs while the other two are bound by a complex of proteins termed anomalous complex or anC (Eltom et al., 1999; Zhang et al., 1998a). The anC, which binds specifically due to two nucleotides flanking the consensus XRE sequence, likely functions to inhibit maximal CYP1B1 induction in the presence of high levels of activated AHR/ARNT by competing for binding with AHR/ARNT at the 5’ XRE. On the contrary, anC is also likely responsible for the constitutive activity of Cyp1b1 by synergistically activating the XRE bound to AHR/ARNT in the presence of very low levels of activated AHR (Zhang et al., 2003). The specific proteins which make up the anomalous complex or the exact molecular mechanisms behind the regulation of CYP1B1 are currently unclear.

CYP1A2 is involved in aromatic amine metabolism and the metabolism of a number of drugs including caffeine and theophylline. CYP1A2 is inducible by PAHs but unlike CYP1A1, is expressed constitutively and predominately in the liver. The CYP1A2 gene, in mammals, is on the same chromosome and orientated in a head-to-head fashion with CYP1A1, separated by approximately 23kb. Two regions were identified within the ~2.5kb upstream region of the CYP1A2 gene which are essential for 3-MC-mediated transcriptional activation (Quattrochi et al., 1994). One of the identified regions, termed X1, contains an XRE-like sequence which weakly associates with AHR in the presence of 3-MC. Elimination of this region resulted in an approximately 50% reduction in activity. The second identified region, termed X2, did not associate with AHR in vitro but may play a role in 3-MC-mediated induction due to a putative AP1 site (Quattrochi et al., 1994). Recent studies using a dual reporter vector under control of the 23kb region between human CYP1A1 and CYP1A2 shows that an XRE cluster near the CYP1A1
transcriptional start site may work bi-directionally to regulate CYP1A2 (Ueda et al., 2006). Further studies need to be performed to elucidate the mechanisms behind CYP1A2 regulation.

CYP2S1 was identified in 2001 as the only member of the novel, 2S, family of cytochrome P450s (Rivera et al., 2002; Rylander et al., 2001). CYP2S is a member of the CYP2 family which like Cyp2a5 and CYP2A6, from mouse and human respectively, is inducible by dioxin (Gokhale et al., 1997; Rivera et al., 2002). Expression of CYP2S is similar to that observed for CYP1B1, being prominent in epithelial tissues including skin, trachea, lung, and intestine (Rylander et al., 2001; Saarikoski et al., 2005). CYP2S1 also resembles CYP1B1 in that it lacks a TATA box within the promoter (Rivera et al., 2007). Several XREs were identified within the 5.2kb region upstream of the translational start codon of the mouse Cyp2s1 gene; however studies using reporter vectors containing combinations of mutated XREs show that TCDD-mediated induction is attributed only to a region containing three overlapping XREs between -393 and -408 (Rivera et al., 2007). AHR/ARNT was capable of binding all three of the overlapping XREs in a ligand dependant manner as shown by EMSA. Each of the three XREs was also able to induce a reporter gene in cell culture; however simultaneous mutation of any two of the three trimeric XREs severely reduced TCDD responsiveness to near control levels as did mutation of all three (Rivera et al., 2007). A regulatory region containing a series of overlapping XREs such as exhibited in CYP2S1 has not been previously identified. Interestingly, the mouse region which contains the trimeric XREs also was found to contain three overlapping HREs of which at least one binds HIF1α/ARNT and
is responsive to hypoxia. (Rivera et al., 2007) The human Cyp2s1 promoter was also found to contain two overlapping XREs and two overlapping HREs and is responsive to both dioxin and hypoxia (Rivera et al., 2002). Elucidation of the molecular mechanisms of CYP2S1 regulation is not only important because CYP2S1 may be partially important for the toxic effects of PAHs and dioxin, but additionally may help gain a better understanding of AHR-mediated signaling.

While AHR additionally regulates a number of phase II xenobiotic metabolizing enzymes which contain XREs within their regulatory regions, these genes also typically contain antioxidant response elements (ARE) which are bound by nuclear factor erythroid 2 p45-related factor (Nrf2), the product of another target gene of AHR (Kohle and Bock, 2007; Miao et al., 2005). In addition to activation by AHR, Nrf2 can be activated by reactive oxygen species resulting from phase I XME metabolism of PAHs (Kohle and Bock, 2007; Marchand et al., 2004). Although functional AREs and XREs have not been identified in the regulatory regions of all phase II genes known to be regulated by AHR, analysis of NQO1 in AHR (-/-) and Nrf2 (-/-) null mice showed that TCDD-inducible expression required both AHR and Nrf2 (Kohle and Bock, 2007; Ma et al., 2004). These findings led to a model whereby phase II XMEs may be regulated directly by AHR binding to XREs, by coordinate binding of AHR and Nrf2, or by AREs being bound by Nrf2 which is itself regulated by AHR and ROS. These mechanisms of cross-talk have yet to be confirmed but may be important in gaining a better understanding of the complexities of AHR-mediated signaling.
Cyp1a1 I (-/-) and cyp1a2 (-/-) null mice have shown relatively little protection against the toxic effects of TCDD when compared to the extensive protection exhibited by Ahr (-/-) null mice (Bunger et al., 2003; Gonzalez and Fernandez-Salguero, 1998; Smith et al., 2001; Uno et al., 2004) suggesting that induction of these genes may play little role in mediating TCDD toxicity. Microarray studies have identified numerous genes which are putatively regulated by AHR and have shown that the scope of genes may be far outside that of xenobiotic metabolizing enzymes including genes involved in reproduction, growth and development, cell cycle control, and differentiation (Tijet et al., 2006; Yoon et al., 2006). Recent advances in technology such as microarray analyses may help identify the genes responsible for the toxic response to dioxin, yet it is imperative that an understanding of the molecular mechanisms of AHR gene regulation be gained through the characterization of well-defined AHR targets.

TCDD-mediated Developmental Toxicity in Zebrafish

In zebrafish, embryonic dioxin exposure leads to a series of fairly well characterized developmental defects which include disruption of erythropoiesis, altered regional blood flow, craniofacial malformation, impaired lower jaw development, apoptosis and local circulation failure in the dorsal midbrain, edema, retarded development, and death (Antkiewicz et al., 2005; Belair et al., 2001; Dong et al., 2002; Henry et al., 1997; Hill et al., 2003; Teraoka et al., 2002). The defects associated with TCDD toxicity are exhibited between 48-120 hours postfertilization (hpf) with a reduction in the number of myocytes, reduced blood flow, and a change in the morphology of pronephric glomerulus being the earliest observed defects (Carney et al.,
The reduced peripheral blood flow and reduction in the number of myocytes which occur at 48 hpf as well as a change in the morphology of the heart are the earliest onset of cardiovascular dysfunction which ensues throughout the first 120 hpf (Antkiewicz et al., 2005; Carney et al., 2006; Henry et al., 1997; Teraoka et al., 2002). The morphological changes in the zebrafish heart following TCDD exposure are primarily due to the blockage of the common cardinal vein from migrating dorsally toward the heart between 72 and 96 hpf (Antkiewicz et al., 2005; Bello et al., 2004; Carney et al., 2006). Heart morphology is further affected by an aberration of the normal looping of the heart which occurs concurrently to defective remodeling of the common cardinal vein (Antkiewicz et al., 2005; Carney et al., 2006; Chen et al., 1997). These events lead to a heart which is mis-positioned and has an elongated atrium and a compact ventricle, although it is unclear whether these are direct effects of AHR regulation or secondary to a decrease in cardiac output (Antkiewicz et al., 2005; Carney et al., 2006).

Osmoregulatory defects are also observed in zebrafish embryos exposed to dioxin. Edema is observed in the pericardium and yolk sac at 72 and 96 hpf respectively (Belair et al., 2001; Dong et al., 2002; Henry et al., 1997). Given that the gills do not play a role in osmoregulation until after 96 hpf (Rombough, 2002) and studies have shown that the pronephric kidney is not affected by TCDD prior to the onset of edema, (Hill et al., 2004) the likely cause of edema is linked to skin permeability and/or the circulatory defects defined above.

Additional adverse effects of TCDD on zebrafish development include inhibited growth of the cartilage which forms the lower jaw, increased apoptosis in the dorsal
midbrain, a reduced brain volume, and a reduction of definitive erythrocytes (Belair et al., 2001; Henry et al., 1997; Teraoka et al., 2002). The molecular mechanisms behind these defects are largely unknown but it is not likely that cranio-facial malformations or a failure of primitive erythrocytes to switch to definitive erythrocytes is secondary to cardiac dysfunction. The role AHR plays in mediating these effects will be a significant focus for studies in years to come.

**AHR and ARNT in Fishes**

Although polymorphisms are apparent in fish as they are in mammals, most fish exhibit multiple AHR genes whereas only a single gene is present in mammals. These multiple products are most likely the result of gene duplication events which occurred throughout the course of evolutionary history (Hahn, 2002). In zebrafish (*Danio rerio*) two known Ah receptors were isolated and designated zfAHR1 and zfAHR2 (Andreasen et al., 2002; Tanguay et al., 1999). Interestingly, while the zfAHR1 is the ortholog of the mammalian receptors, it is the zfAHR2 which is responsible for the TCDD-mediated AHR activity in this species (Andreasen et al., 2002; Prasch et al., 2003). Studies by Andreasen et al. show that zfAHR2 is expressed relatively ubiquitously while zfAHR1 is limited to expression primarily in the liver (Andreasen et al., 2002). The functional differences between the two receptors likely lies in the ligand binding domains and transactivation domains of the proteins as shown by experiments using zfAHR1/zfAHR2 chimeras (Andreasen et al., 2002). Like other piscine AHRs, zfAHR2 exhibits great similarity to mammalian AHR in the b-HLH-PAS domains but lacks the Q-rich region in the C-terminus essential for mammalian transactivation (Hahn, 2002; Tanguay et al.,
1999). Studies by Kumar et al. however, suggest that the requirement of this Glutamine rich region, in humans at least, may be limited to a single hydrophobic residue (Leu-678) indicating that abundant glutamines may not be required *per se* for transactivation and fish AHRs may possess the necessary hydrophobic residue (Kumar and Perdew, 1999). Furthermore, while the Q-rich domain was deemed necessary for hAHR transactivation, in mice the Q-rich region of the AHR enhanced the transactivation ability but was not required (Jain et al., 1994; Sogawa et al., 1995).

Only a single ARNT has been identified in most teleost fishes. Two splice variants of an ARNT1 homolog have been isolated in rainbow trout (*Onchorhynchus mykiss*) and designated rtARNTa and rtARNTb (Pollenz et al., 1996). These proteins are identical over the first 533 amino acids which include the b-HLH and PAS domains but diverge in the carboxyl end due to an additional 373 bp sequence in rtARNTb which causes a frame shift in the product. While both rtARNTa and rtARNTb are capable of binding AHR *in vitro*, only rtARNTb appears able to facilitate transactivation of CYP1A1, likely due to the inefficiency of rtARNTa to bind DNA. Pollenz et al. also showed that rtARNTa is capable of behaving as a dominant negative inhibitor of rtARNTb mediated gene induction; however, while rtARNTb is expressed ubiquitously, rtARNTa is expressed at much lower levels and restricted in its distribution (Pollenz et al., 1996). Interestingly, a single ARNT was isolated from the Atlantic killifish (*Fundulus heteroclitus*) as well; however, phylogenetic analyses revealed that, unlike rtARNT, the protein was a homolog of mammalian ARNT2 (Powell et al., 1999). In zebrafish, the three alternatively spliced ARNTs originally identified by Tanguay et al.
were also homologous to mammalian ARNT2 and were designated zfARNT2a, b, and c (Tanguay et al., 2000). zfARNT2b was shown to bind zfAHR2 in vitro and could moderately induce XRE-driven reporters in COS-7 cells, yet zfARNT2 morphants and zfARNT2 -/- embryos still exhibited the same endpoints of TCDD toxicity observed in wild type fish (Prasch et al., 2004) suggesting that zfARNT2 is not the ARNT involved in mediating TCDD toxicity. Three additional alternatively spliced ARNTs were subsequently isolated from the zebrafish and found to be homologous to the rtARNTb and mammalian ARNT1. Designated zfARNT1a, b, and c, these three proteins were found to be expressed continuously throughout the timecourse critical for TCDD-mediated developmental toxicity, albeit at considerably lower levels than zfARNT2. Furthermore, zfARNT1b and c were capable of forming dimers with zfAHR2, capable of binding DNA in vitro, and inducing XRE-driven reporter constructs. Most importantly, zfARNT1 morphants showed protection against three of the endpoints of TCDD toxicity: pericardial edema, reduced blood flow, and reduced lower jaw growth (Prasch et al., 2006). It is important to note that there may be a species specific difference in XRE recognition sequences by AHR and ARNT. Tanguay et al. reported that zfAHR2/rtARNTb dimers failed to bind the murine XRE containing the core sequence 5’ –TTGCGTG- 3’ but actively bound the rainbow trout XRE containing the sequence 5’ –TAGCGTG- 3’ (Tanguay et al., 1999). Upon isolation of zfARNT1, Prasch et al. showed binding of zfAHR2/zfARNT2b and c dimers to the same murine XRE indicated above (Prasch et al., 2006). Surprisingly, mouse AHR/rtARNTb dimers bound both murine and rainbow trout XRE containing oligonucleotides (Tanguay et al., 1999). Further research
will have to be done to elucidate the potential differences in DNA recognition between species.
Isolation of the zfCYP1A upstream region

To obtain the sequence of the zfCYP1A1 promoter/enhancer, zfCYP1A1 cDNA sequence was used to query the Sanger zebrafish genome database. No sequence could be recovered that contained significant identity to zfCYP1A1 so a zebrafish genomic PAC library (Amemiya and Zon, 1999) was screened using primers specific to the 5’-untranslated region (5’-UTR) and most 5’-region of the open reading frame (ORF) of the zfCYP1A1 cDNA. Two PACs, designated #133 and #150, containing putative zfCYP1A1 genes were identified. Restriction enzyme digestion and Southern analysis of the two PACs identified identical bands that hybridized to the zfCYP1A1 cDNA probe. A 0.5 kb and 2.5kb HindIII fragment as well as a 12 kb SpeI fragment were subcloned from PAC #150 and sequenced. The 2.5 kb HindIII fragment contained the putative ATG start codon and the first 134 bp of coding sequence that showed 100% identity to the zfCYP1A1 cDNA in Genbank (accession #BC094977). The fragment also contained 14 bp of the 5’UTR with 100% identity to the zf-CYP1A1 cDNA. After nucleotide 14, the sequence showed minimal identity to zfCYP1A1 cDNA but contained a putative splice acceptor site at the region when identity was lost. The remaining sequence within the 2.5 kb HindIII fragment showed no identity to the CYP1A1 cDNA and contained substantial
regions of repetitive DNA that were 80%–90% AT-rich. These results indicate that the zfCYP1A1 gene contains an intron within the 5’UTR sequence. This is consistent with the structure of other CYP1A1 genes (Carvan et al., 1999; Kubota et al., 1991; Powell et al., 2004; Sogawa et al., 1986). The size of the intron is estimated to be >2500 bp, but the entire sequence could not be obtained due to the high levels of AT-rich regions, numerous regions of repetitive DNA, and a lack of unique restriction sites to allow subcloning of smaller fragments. Thus, to obtain the remaining 5’UTR and its 5’-flanking region, the SpeI fragment was sequenced with an oligonucleotide complementary to the missing portion of the 5’UTR. This approach identified the remaining 74 bp of the 5’UTR and putative CAAT and TATA boxes. Subsequent sequence analysis identified the splice donor site for intron 1 and a putative promoter/enhancer that spanned a region 2629 bp upstream from the transcription start site. Several elements were identified within the isolated region which are characteristic of previously characterized CYP1A regulatory regions. A putative TATA box is located at position -31 relative to the transcriptional start site which has been designated +1. Additionally, two CTF/NF1 sites were identified at positions -53 and -438 and an Sp1 site at -2474. These sites have all been implicated in the regulation of CYP1A1 in other organisms (Fisher et al., 1990; Jones and Whitlock, 1990; Yanagida et al., 1990). Importantly, eight putative XREs were identified which conform to the consensus sequence: T/GNGCGTG. These XREs were designated 1-8 with XRE1 being nearest the TATA box. Thus, the results indicate that zfCYP1A promoter/enhancer
Figure 2.1. Comparison of CYP1A Regulatory Regions from Different Species

Figure 2.1. Comparison of CYP1A regulatory regions from different species. The location of XREs, indicated by rectangles, are shown in relation to the putative transcriptional start site. Functional XREs are shaded. The TATA box is represented by a shaded circle. CTF/NF1 or BTE sites are indicated by triangles. Shaded triangles have been previously functionally characterized. Small shaded squares indicate putative HNF-3 sites. Zf = zebrafish; rt = rainbow trout; fh = Fundulus heteroclitus; r = rat; m = mouse; h = human.
contains numerous XREs and other consensus regulatory sequences and bears an overall structure that is similar to the mouse, rat, and trout. A comparison of the zebrafish region to other characterized CYP1A regulatory regions is shown in Figure 2.1.

**Induction of the zfCYP1A gene by TCDD**

Since zebrafish have undergone a gene duplication event, they can contain multiple copies of genes as well as pseudogenes that are nonfunctional (Postlethwait et al., 1998; Woods et al., 2000). Therefore it was pertinent to verify that the isolated zfCYP1A gene could support gene regulation and was inducible by TCDD. To address this question, total RNA was isolated from ZFL cells that were treated with vehicle or TCDD for 6 h. RT-PCR was carried out to amplify CYP1A and actin mRNA. To confirm that expression was from the identified zfCYP1A promoter, the CYP1A primers were complementary to the 5’UTR and ORF and designed to be on either side of the first intron. Thus, the expected band from the amplification of CYP1A mRNA was 217 bp whereas a band that was generated by amplifying genomic DNA would be >3 kb. The results show that a band of 217 bp was weakly visible in the untreated ZFL but was dramatically elevated in the presence of TCDD. (Fig. 2.2) Amplification of actin shows that the changes in the level of CYP1A are not related to differences in the level of RNA used in the assay. Thus, the data are consistent with previous studies that have identified TCDD-inducible CYP1A in zebrafish (Henry et al., 2001; Miranda et al., 1993) and support the hypothesis that the identified zfCYP1A gene is indeed inducible by TCDD.
Figure 2.2. Induction of the zfCYP1A gene by TCDD

A. The location of the forward and reverse primers used for reverse transcriptase PCR are indicated with arrows and shown in relation to the zfCYP1A gene. B. Triplicate samples of were exposed to 2nM TCDD (TC-6) or 0.05% DMSO (0) for 6 hours and total RNA was prepared. mRNA was amplified following reverse transcription with PCR primers specific to β-Actin (357 bp) or CYP1A (217 bp). PCR products were visualized in a 2% agarose gel containing ethidium bromide and exposed to UV light. Specific markers of 357, 323, and 200 bp are indicated.
While the previous experiments suggest that the identified *zfCYP1A* gene was inducible by TCDD, it was still important to show that the 5’ flanking region contained specific regions that conferred TCDD-responsiveness. Unfortunately, due to the extreme AT-rich content and high level of repetitive DNA that is present between -600 and -2000, it was not possible to generate a construct containing the full 2600 bp of the promoter/enhancer. Thus, PCR was utilized to amplify the region between -580 and -187 (containing XREs 1-3) and the region between -2608 and -2100 (containing XREs 4-8). These fragments were ligated in the forward and reverse orientations both upstream and downstream of the SV-40 promoter in the pGL3promoter vector. TCDD mediated induction of luciferase activity was then evaluated in the mouse Hepa-1 cell line. The use of the mouse Hepa-1 line and not the zebrafish ZFL line for these studies was based on the ability to grow large numbers of cells and the ability to obtain high levels of transfection efficiency that facilitated the analysis of the luciferase and β-galactosidase activities. The results shown in Figure 2.3 reveal that the -2608 to -2100 fragment confers TCDD responsiveness to the SV40 promoter regardless of its location or orientation. However, the maximal induction (approximately 40-fold) was observed in the -2608/-2100Rup construct in which the 506bp *zfCYP1A* fragment was placed in a reverse orientation upstream of the SV40 promoter. In this context, the magnitude of the response and the overall level of induction were approximately five-fold higher than when the fragment was inserted in the forward orientation (-2608/-2100Fup).

Interestingly, when the -2608 to -2100 fragment was placed downstream of the SV40 promoter in either the forward or reverse orientation (-2608/-2100Fdown and -2608/-
2100Rdown), it was also capable of inducing significant levels of luciferase activity in a TCDD-dependent manner. However, in this context, the orientation of the zfCYP1A fragment made no difference in the overall level of the response although the magnitude of the response was 5-fold less than that observed for the -2608/-2100Rup.

In contrast to the results with the -2608 to-2100 fragment, the constructs containing the proximal region between -580 to -187 (-580/-187Fup and -580/-187Fdown) did not exhibit elevations in luciferase activity in the presence of TCDD. Even when placed in the reverse orientation, the proximal fragment failed to confer TCDD-responsiveness to the reporter gene construct (Fig. 2.3). This finding is intriguing since the region between -580 to -187 contains three putative XREs. It was important then to assess the function of XREs 1-3 in the context of their native promoter. PCR was used to amplify the region between -580 and +71 and the resulting fragment was ligated into pGL3Basic to generate the -580/+71Basic construct. This construct was transfected into Hepa-1 cells and treated with TCDD. The results in Figure 2.4A show that -580/+71 exhibits high levels of constitutive activity which is approximately 35-fold higher than naked vector however failed to convey TCDD-mediated induction. These results confirm the observations made on the -580/+71Fup and -580/+71Fdown constructs and additionally validates the function of this region as a functional promoter. As a positive control, Hepa-1 cells were transfected with p-1897Om1A3luc that contains the full length trout CYP1A3 promoter/enhancer. This construct was induced approximately threefold in the Hepa-1 line. To determine whether the lack of TCDD-responsiveness by the -580/+71Basic construct was related to the analysis in a murine background, the construct...
Figure 2.3. Analysis of the Effect of Orientation and Position on TCDD-induced Luciferase Activity

The indicated reporter constructs as well as pSV-β-galactosidase were transfected into Hepa-1 cells and treated with either 2nM TCDD or 0.05% DMSO for seven hours. Luciferase activity was measured with a Turner Instruments luminometer. β-Galactosidase levels were measured by spectrophotometry (OD420). Normalization was carried out by dividing the relative luciferase levels for each sample by the corresponding level of β-galactosidase. White bars represent DMSO treated cells while black bars indicate cells treated with TCDD. Bars represent the mean ± SE of three independent samples. * indicates statistically significant from vehicle treated controls. P<0.001
was transfected into the zebrafish liver cell line, ZFL. This line contains zfAHR2 and is responsive to AHR ligands (Miranda et al., 1993; Woods et al., 2000). The results show that that -1897Om1A3luc is responsive to TCDD in the ZFL cells and was induced approximately threefold (Fig. 2.4B) as was observed in Hepa-1 cells, (Fig. 2.4A) whereas -580/+71Basic exhibits elevated activity well above the parental vector, but is still not responsive to TCDD. Thus, these studies show that the region of the zfCYP1A gene containing defined TATA and CAAT boxes (-580 to +71) can function to promote gene expression in both mouse and zebrafish backgrounds. The region which confers TCDD responsiveness, however, appears to be located between -2100 and -2608.

In order to determine whether or not the zebrafish CYP1A enhancer region was able to convey a similar TCDD-responsiveness to the endogenous zfCYP1A promoter as was observed with the SV-40 promoter of the pGL3promoter vector, the -2100/-2608 region was ligated in either the forward or reverse orientation upstream of the -580/+71 region and transfected into Hepa-1 cells. These constructs were designated p-2608/-2100Uf or p-2608/-2100Ur, respectively. Interestingly, the results show that while the region driving the SV-40 promoter exhibited an approximately threefold greater induction in the reverse versus the forward orientation, (Fig. 2.3) in the context of the native promoter the difference was markedly less yielding approximately 18-fold induction in the forward orientation and 22-fold induction in the reverse orientation. (Fig. 2.5) Despite this difference, the results show that the -2100/-2608 portion of the zfCYP1A regulatory region is capable of conveying TCDD-mediated induction to its endogenous promoter. As positive controls, the mouse CYP1A1 regulatory region
Figure 2.4. Luciferase Reporter Analysis of the Ability of the region between -580 and +71 to Function as a Promoter

A) Hepa-1 cells

B) ZFL cells

Figure 2.4. Luciferase reporter analysis of the ability of the region between -580 and +71 to function as a promoter. The indicated reporter constructs as well as pSV-β-galactosidase were transfected into Hepa-1 cells (A) or ZFL cells (B) and treated with 2nM TCDD or 0.05% DMSO for 7 hours. Luciferase activity was measured with a Turner Instruments luminometer. β-Galactosidase levels were measured by spectrophotometry (OD420). Normalization was carried out by dividing the relative luciferase levels for each sample by the corresponding level of β-galactosidase . White bars represent DMSO treated cells while black bars indicate cells treated with TCDD. Bars represent the mean ± SE of three independent samples. * indicates statistically significant from vehicle treated controls. P<0.001
between -1315 and -819 and the rainbow trout CYP1A3 regulatory region between -1897 and -1392 which have been previously shown to exhibit TCDD-responsiveness, (Carvan et al., 1999; Jones et al., 1986) were ligated upstream of the zf-580/+71 region and transfected into Hepa-1 cells. While these constructs respectively yielded approximately ten and threefold levels of induction, the magnitude of the response by the zf-2100/-2608 region was far greater indicating that zfCYP1A is highly responsive to TCDD in Hepa-1 cells.

**In vitro analyses of AHR/ARNT association with zfXREs**

Since the fragments containing XREs 1-3 did not confer TCDD-responsiveness, it was of interest to determine whether they represented bona-fide regions that actually bound to AHR/ARNT complexes. To carry out these studies it was first important to establish the conditions for the detection of zfAHR2/ARNT complexes *in vitro*. Previous studies carried out in COS-1 cells have shown that the zfAHR2 can drive reporter gene expression in the presence of zfARNT2b. (Abnet et al., 1999; Tanguay et al., 2000) However, recent studies have challenged these findings by showing that zfARNT2b does not support TCDD-mediated responses *in vivo* (Prasch et al., 2004; Prasch et al., 2003). Thus, studies were carried out using rtARNTb and zfARNT2b since rtARNTb has been shown to form a functional dimer with mAHR in vitro (Necela and Pollenz, 1999; Necela and Pollenz, 2001; Pollenz et al., 1996). For these studies, zfAHR2, zfARNT2b, and rtARNTb were synthesized in reticulocyte lysates, incubated with TCDD, and analyzed by electrophoretic mobility shift assays (EMSA). The results shown in Figure 2.6 reveal that complexes containing zfAHR2 and rtARNT2b produce a specific shift that is
Figure 2.5. Analysis of TCDD-mediated Induction from Reporter Constructs Containing the zfCYP1A Promoter Region Between -580 and +71 and CYP1A Enhancer Regions from Different Species

**A.** The indicated reporter constructs as well as pSV-β-galactosidase were transfected into Hepa-1 cells and treated with 2nM TCDD or 0.05% DMSO for 7 hours. Luciferase activity was measured with a Turner Instruments luminometer. β-Galactosidase levels were measured by spectrophotometry (OD420). Normalization was carried out by dividing the relative luciferase levels for each sample by the corresponding level of β-galactosidase. White bars represent DMSO treated cells while black bars indicate cells treated with TCDD. **B.** Fold induction was determined by dividing nRLU of vehicle treated cells by dioxin treated cells. Bars represent the mean ± SE of three independent samples.
TCDD dependent and can be competed with antibodies specific to the zfAHR2 protein. In contrast, no shift was detected in the samples activated with zfAHR2 and zfARNT2b. These results are consistent with the hypothesis that TCDD-mediated signaling does not utilize ARNT2 proteins, but occurs through dimers with AHR and ARNT1 proteins (Prasch et al., 2004; Prasch et al., 2003).

To assess the functionality of the XREs present in the zfCYP1A promoter, duplex oligonucleotides were prepared that contained the core XRE as well as 6–7 nucleotides of flanking sequence. Each XRE was then evaluated for binding to zfAHR2/rtARNTb complexes by EMSA (Fig. 2.7). The results show that only five of the eight XREs associate with zfAHR2/rtARNTb dimers in a TCDD-dependent manner. XRE3, XRE7, and XRE8 showed the most intense shifts, while XRE1 and XRE 4 associated with AHR/ARNT dimers in a TCDD-dependent manner, but showed slightly less intensity. Since all XREs were labeled to the same specific activity, the reduced intensity of the shifted bands likely represents a reduced level of affinity between the AHR/ARNT dimer and the XRE. XRE2 and XRE6 showed no detectable shifts, while XRE5 showed a very weak shift after prolonged exposure of the film (data not shown). Identical results were obtained when the zfXREs were evaluated in the presence of mouse AHR/ARNT complexes. To verify that the lack of binding by XRE2, XRE5, and XRE6 was not due to the use of zfAHR2 and rtARNTb, studies were repeated using zfAHR2 and zfARNT2b. However, the use of zfAHR2/zfARNT2b heterodimers also failed to produce a detectable shift (data not shown). Thus, these results show that only a subset of the XREs are
Fig. 2.6. In vitro analysis of the association of zfAHR2 with XREs. A. The indicated proteins were expressed in vitro, resolved by SDS-PAGE, and blotted to nitrocellulose. Blots were stained with zf-4 IgG (1.0 μg/mL), rt-84 IgG (1.0 μg/mL), or anti-ARNT2 antibodies (1:250) followed by GAR-HRP or RAG-HRP IgG (1:10,000). Reactivity was visualized by ECL. B. Equal amounts of zfAHR2 were mixed with equal amounts of either zfARNT2b or rtARNTb and incubated with TCDD (16 nM) or DMSO (1.0%) for 2 h at 30°C. Samples were mixed with $^{32}$P-labeled mXRE in the presence or absence of the indicated antibodies and resolved on 5% acrylamide/0.5% TBE gels, dried, and exposed to film.
Equal amounts of zfAHR2 were mixed with equal amounts of rtARNTb and incubated with TCDD (16 nM) or DMSO (1.0%) for 2 h at 30°C. Samples were mixed with the indicated $^{32}$P-labeled zfXREs and resolved on 5% acrylamide/0.5% TBE gels, dried, and exposed to film.
functional in binding to AHR/ARNT in vitro and that there are slight differences in their affinities for AHR/ARNT dimers. In addition, the results suggest that the inability for the -580 to -187 fragment to respond to TCDD in the reporter gene studies is not due to the inability of the XREs to associate with AHR/ARNT complexes.

**Functional characterization of individual XREs within the zfCYP1A regulatory region**

Due to the 1.7kb AT-rich region which separates the proximal (-580 to +71) and distal (-2100 to -2608) cluster of XREs in the zfCYP1A regulatory region, PCR and PCR-dependant site directed mutagenesis could not be performed to amplify the complete 2.7kb fragment. Thus, the previous characterization of the zfCYP1A regulatory region utilized constructs containing only the regions between -2100 to -2608 and -580 to +71. Before site directed mutagenesis could be employed on the p-2608/−2100Ur construct in order to determine the functionality of individual XREs, it was imperative to determine whether its ability to drive a luciferase reporter was representative of the full length promoter/enhancer. Thus, a full-length construct was generated by cutting the zfCYP1A promoter/enhancer region from the original PAC clone and ligating it into pGL3. The full-length p-2699/+71 and p-2608/−2100Ur constructs were then transfected into Hepa-1 cells and the level of TCDD-induced luciferase activity quantified. It can be observed in Figure 2.8 that the full-length p-2699/+71 was highly inducible by TCDD and averaged approximately 20-fold induction over control treated cells. This is consistent with the p-2608/−2100r construct that averaged approximately 32-fold induction. The p-2699/+71 construct did exhibit reduced levels of total RLU in both
control and TCDD exposed cells by comparison to the p-2608/−2100Ur construct, but the difference between the two constructs may have been due in part to a significantly higher level of transfection efficiency for p-2608/−2100Ur than the larger, AT-rich p-2699/+71. Therefore, these studies indicate that both the p-2608/−2100Ur and the p-2699/+71 construct are highly responsive to TCDD exposure and validate the use of p-2608/−2100Ur for the analysis of the regulation of the zfCYP1A gene.

Previous experiments have shown that the region between -580 and +71 which contains XREs 1-3 is incapable of conveying TCDD-mediated gene induction (Fig. 2.4) even though XRE1 and XRE3 are able to bind AHR/ARNT in vitro in a TCDD-dependant manner (Fig. 2.7). To determine whether XRE1 and XRE3 contribute to maximal induction by acting in concert with XREs in the distal cluster, in vitro mutagenesis was employed on p-2608/-2100Ur in order to render XREs 1 and 3 non-functional, alone or in combination. The mutants, designated p-2608/-2100Ur(-1), p-2608/-2100Ur(-3), or p-2608/-2100Ur(-1-3), or the non-mutated control were transfected into Hepa-1 cells and assayed for luciferase activity. The results seen in Figure 2.9 indicate that the loss of XRE1 or XRE3 does not significantly affect the overall levels of maximum gene induction by TCDD. Considering the fact that these XREs are capable of binding AHR/ARNT in vitro, it is intriguing that XRE1 and 3 do not play an apparent role in the regulation of the downstream gene. It should be noted that the distance of the elements from the transcriptional start site is likely not the reason behind the functionality of these XREs, as they failed to drive an SV-40 promoter when placed within the same position as the functional, distal region of XREs. While it is possible that activated
Figure 2.8. Comparison of TCDD-induced Luciferase Activity Between the p-2699/+71 and p-2608/-2100Ur Constructs

Hepa-1 cells were transfected with either p-2699/+71 or p-2608/-2100Ur and treated with TCDD (2 nM) or DMSO (0.05%) for 15 h. Luciferase activity was measured using a Turner Instruments luminometer and \( \beta \)-galactosidase activity was determined by spectrophotometry (OD 420). (A) The normalized relative luciferase units are shown for the indicated plasmids. Open bars represent control treated cells while black bars indicate TCDD treated cells. Each bar = mean ± S.E. from four independent experiments. (B) The fold induction is shown for the indicated plasmids. Fold induction was determined by dividing the normalized RLU of samples treated with TCDD by the normalized RLU of control treated samples presented in (A).
AHR/ARNT is capable of binding these elements *in vitro* but not *in vivo*, the reason for this would likely be due to chromatin condensation within the region containing XREs 1 and 3 or the inaccessibility of the elements due to interference by other proteins binding nearby as opposed to differences in intracellular binding affinities versus those observed by EMSA. Further experimentation will be required to elucidate the reasons for the lack of function of these XREs.

Contrary to what was observed for the XREs within the proximal region, the distal region between -2100 and -2608 containing XREs 4 through 8, is capable of conveying TCDD-mediated gene induction (Figs. 2.3 and 2.5). Additionally, EMSA has shown that XREs 4, 7, and 8 are capable of being bound by AHR/ARNT in a TCDD dependant fashion (Fig. 2.7). Therefore, *in vitro* mutagenesis was also performed on these three XREs in order to determine their individual contributions to maximal gene induction. Transfection into Hepa-1 cells and subsequent luciferase analysis shows that individual mutation of XRE4, XRE7, or XRE8 resulted in significant reductions in both raw levels of luciferase activity as well as fold-change of induction (Fig. 2.10). While the transcriptional activity of each XRE does not appear to be equivalent by assessing the degree of reduction observed by individual XRE mutations, this inequality is more apparent when more than one XRE is mutated. Indeed, while elimination of any one of the active XREs (4, 7, or 8), resulted in a 30–50% decrease in TCDD-mediated luciferase activity, XRE4 or XRE7 alone supported approximately 25% of the maximal luciferase induction in the absence of additional functional XREs while XRE8 alone showed minimal activity above p-2608/-2100Ur(-478) or the p-580/+71Basic control. Therefore,
**Figure 2.9. Functional Analysis of zfXREs 1 and 3**

Fig. 2.9. Functional analysis of zfXREs 1 and 3. Hepa-1 cells were transfected with the indicated plasmids and treated with TCDD (2 nM) or DMSO (0.05%) for 15 h. Luciferase activity was measured using a Turner Instruments luminometer and β-galactosidase activity was determined by spectrophotometry (OD 420). (A) The normalized relative luciferase units are shown for the indicated plasmids. Open bars represent control treated cells while black bars indicate TCDD treated cells. Each bar = mean ± S.E. from three independent experiments. (B) The fold induction is shown for the indicated plasmids. Fold induction was determined by dividing the normalized RLU of samples treated with TCDD by the normalized RLU of control treated samples presented in (A).
these results suggest that in this model system, each of the regulatory sequences does not provide the same level of regulation to the CYP1A gene as suggested by Fisher et al. (Fisher et al., 1990). To verify that the previous results were not due to the analysis of the zfCYP1A promoter in a mouse cell line, studies were repeated using the zebrafish ZFL liver cell line. This line contains zfAHR2 as well as zfARNT1 and zfARNT2 and is capable of supporting TCDD-mediated gene regulation (Carvan et al., 2000; Miranda et al., 1993; Pollenz and Dougherty, 2005; Wentworth et al., 2004; ZeRuth and Pollenz, 2005). In comparison to the mammalian cell lines, transfection of the ZFL cells was much less efficient and the cells exhibited a higher level of basal activity that resulted in lower levels of fold induction in several constructs. Nevertheless, the results in the ZFL line (Fig. 2.11A) showed several similarities to the results in the Hepa-1 line (Fig. 2.10). First, single mutations of XRE4 or XRE7 resulted in a significant reduction in TCDD-inducible luciferase activity. Second, mutation of XRE8 did not affect the luciferase induction as significantly as loss of XRE4 or XRE7. Thus, as in the Hepa-1 cells, the contribution of each XRE4 and XRE7 to the induction of the luciferase reporter was much more prominent than XRE8. To compare the trend of the results across the different cell lines used in the studies, the results were scaled with the overall fold induction of wild type p-2608/-2100Ur construct set at 100%. The fold induction yielded by constructs harboring XRE mutations was plotted as a percentage of the maximal induction (Fig. 2.11B). The results show that there is a similar trend of the various constructs when analyzed in Hepa-1 or zebrafish cells with mutation of XRE4 or XRE7 having a more dramatic impact than mutation of XRE8. The graph also contains results from studies completed in the HepG2 line that is of human origin. In this cell line, the
Figure 2.10. Functional Analysis of zfXREs

The indicated plasmids were transfected into Hepa-1 cells as described for Fig. 2.9. (A) The normalized relative luciferase units are shown for the indicated plasmids. Open bars represent control treated cells while black bars indicate TCDD treated cells. Each bar = mean±S.E. from five independent experiments. Schematics of the constructs are shown to the left of their respective bars. Open rectangles indicate XREs. Shaded rectangles indicate mutated XREs. The TATAA box and luciferase cassette are shown. (B) The fold induction is shown for the results presented in (A) and was determined by dividing the normalized RLU of samples treated with TCDD by the normalized RLU of control treated samples. a. Statistically different from p-2608/-2100Ur (p < .001). b. Statistically different from p-2608/-2100Ur(-7) (p < .01). c. Statistically different from all constructs except p-2608/-2100Ur(-478) and p-580/+71Basic (p < .001). d. Statistically different from p-2608/-2100Ur(-8) (p < .001). e. Statistically different from all constructs except p-2608/-2100Ur(-478) and p-580/+71Basic (p < .001). f. Statistically different from all constructs except p-2608/-2100Ur(-4-7) and p-2608/-2100Ur(-478) (p < .001).
basal level of luciferase activity in controls was low for all the constructs and this resulted in very high levels of fold-change in the presence of TCDD. It can be observed that the overall trend obtained with the various constructs in the HepG2 line follows that observed in the Hepa-1 and ZFL cells despite the fact that that overall fold induction varies dramatically between the cell lines.

In order to determine whether the inability of XRE8 to function alone was due to the truncation of the fragment immediately upstream of the core sequence, a new construct was made extending the 5’ end of the insert up to -2727 and designated p-2727/-2100Ur. XREs 4, 7, and 8 were mutated individually or in combination as done previously with p-2608/-2100Ur and evaluated for luciferase activity in the Hepa-1 cell line. The results in Figure 2.12 show that p-2727/-2100Ur exhibited approximately 60% higher levels of basal and induced luciferase activity over p-2608/-2100Ur while the overall fold-induction increased by approximately 17% between the constructs. Considering the differences in the level of activity and fold-induction is consistent between p-2608/-2100Ur and p-2727/-2100Ur equivalent mutants, it can be inferred that the region between -2727 and -2608 contributes to the overall transcriptional activity of the construct however does not change the functionality of XRE8 or its ability to regulate the gene in the absence of additional functional XREs. The molecular mechanism that underlies the inability of XRE8 to function alone is not presently clear but it is possible that XRE8 is inaccessible prior to AHR/ARNT binding at XRE 4 and 7 or that AHR/ARNT binding to XRE8 cannot effectively recruit transcriptional coactivators. Indeed, chromatin relaxation and/or DNA bending has been shown to occur following
Figure 2.11. Functional Analysis of zfXREs in Different Cell Lines

Fig. 2.11. Functional analysis of zfXREs in different cell lines. (A) The indicated plasmids were transfected into ZFL cells and treated with TCDD (2 nM) or DMSO (0.05%) for 15 h. Luciferase activity was measured using a Turner Instruments luminometer and β-galactosidase activity was determined by spectrophotometry (OD 420). The normalized relative luciferase units are shown for the indicated plasmids. Open bars represent control treated cells while black bars indicate TCDD treated cells. Schematics of the constructs are shown to the left of their respective bars as described for Fig. 2.10A. Each bar = mean±S.E. from three independent experiments. The number in parentheses indicates the fold-change for TCDD samples compared to DMSO treated controls. a. Statistically different from p-2608/−2100Ur (p < .001). b. Statistically different from paired control. c. Statistically different from p-2608/−2100Ur(−4) and p-2608/−2100Ur(−7) (p < .005). (B) Scaled line graph showing the TCDD-induced luciferase activity as a percentage of the wild type p-2608/−2100Ur construct. Data were derived from Hepa-1 cells (circles), ZFL cells (triangles), or HepG2 cells (squares).
AHR/ARNT binding at an XRE (Elferink and Whitlock, 1990; Okino and Whitlock, 1995) and this may impact protein–protein interactions or may prevent access to XRE8 prior to AHR/ARNT binding at XREs 4 or 7. In addition, a putative XF-1 site overlaps XRE8 and this could also contribute to its lack of function in the absence of binding at XRE4 and XRE7 as XF-1 binding has been previously been observed in the mouse CYP1A1 enhancer (Saatcioglu et al., 1990). Since ChIP assays cannot distinguish binding to enhancer regions that are on the same fragment of amplified DNA, and the sequential association of AHR/ARNT with XREs in any gene has not been resolve. The importance of these findings will require additional analysis.

The pattern of expression controlled by the zfCYPIA and mCYPIA1 regulatory region varies in different cell lines

During the course of the studies in this report, it was noted that the overall level of gene induction of different reporter constructs varied dramatically when tested in different cell lines. To formally investigate this observation, four different reporter constructs containing regions from the zebrafish of mouse CYPIA1 promoter/enhancer were evaluated in seven different cell lines. The two constructs derived from zebrafish were the full-length p-2699/+71 construct (ZFL) and p-2727/−2100Ur (ZFA). The mouse constructs included one containing the region from −1674/+47 from the mouse CYPIA1 promoter/enhancer (MFL), or the region spanning −1316/−819 ligated upstream of the zebrafish CYPIA promoter region, −580/+71 (MMA). The cell lines utilized for these studies and their tissue of origin are detailed in Fig. 2.13A. Interestingly, two distinct patterns of induction were observed. The results show that the constructs
Fig. 2.12. Comparison of XRE8 function between the p-2608/-2100Ur and p-2727/-2100Ur constructs. The indicated plasmids were transfected into Hepa-1 cells and treated with TCDD (2 nM) or DMSO (0.05%) for 15 h. Luciferase activity was measured using a Turner Instruments luminometer and β-galactosidase activity was determined by spectrophotometry (OD 420). The normalized relative luciferase units are shown for the indicated plasmids. Open bars represent control treated cells while black bars indicate TCDD treated cells. Each bar = mean±S.E. from three independent experiments. The number in parentheses indicates the fold-change for TCDD samples compared to DMSO treated controls. * indicates percent of -2608/-2100Ur fold induction. # indicates percent of -2727/-2100Ur fold induction.
Figure 2.13. Characterization of Mouse and Zebrafish CYP1A Promoter and Enhancer Regions in Various Cell Lines

(A) Species and tissue of origin, AHR allele, and presence or absence of ARNT2 for each of the cell lines. (B and C) Fold induction of luciferase when the designated constructs were transfected into the indicated cell lines and treated with TCDD (2nM). Each data point is the mean of at least three different samples. ZFL = p-2699/+71, ZFA = p-2608/-2100Ur, MFL = mouse-1647/+57, MMA = Mm-1315/-819Uf.

**Fig. 2.13. Characterization of mouse and zebrafish CYP1A promoter and enhancer regions in various cell lines.** (A) Species and tissue of origin, AHR allele, and presence or absence of ARNT2 for each of the cell lines. (B and C) Fold induction of luciferase when the designated constructs were transfected into the indicated cell lines and treated with TCDD (2nM). Each data point is the mean of at least three different samples. ZFL = p-2699/+71, ZFA = p-2608/-2100Ur, MFL = mouse-1647/+57, MMA = Mm-1315/-819Uf.
containing the zebrafish CYP1A promoter/enhancer (ZFL and ZFA), were more responsive in the Hepa-1, B19, and A498 cell lines than the constructs containing the mouse CYP1A1 regions (Fig. 2.13 B and C). In contrast, the hRPE, TCM, C2C12, and HepG2 cell lines significantly favor the full-length mouse promoter/enhancer over the constructs containing the zebrafish regions. The data presented in Figure 2.13B also supports previously detailed results which show that the zebrafish region -2100 to -2608 conveys an approximately 2-fold higher level of induction than does the mouse CYP1A1 region between -819 and -1316. It was previously detailed in this report that p-2608/-2100Ur construct exhibited increased levels of luciferase activity and fold-induction over the full-length p-2699/+71 construct. The data presented in Figure 2.13 supports these findings although it is interesting to note that the increase is significantly more dramatic in the cell lines which favor the zebrafish constructs (Fig 2.13B) versus those which favor the full-length mouse construct (Fig 2.13C). Intriguingly, the contrary is observed in Fig. 2.13C wherein the full-length mouse construct, MFL, produced considerably greater levels of induction than both p-2727/-2100Ur and p-2699/+71. It is also important to recognize that the mouse enhancer region between -819 and -1316 yielded significantly less activity when driving the zebrafish CYP1A promoter (MMA) than the full-length MFL in these four cell lines. The molecular basis for the differences in response of the various reporters in the different cell lines is currently unclear. However, there does not appear to be a correlation to the level or species of AHR protein, the tissue type, or level of expression of ARNT2. The findings presented above suggest that the differences observed between the transcriptional activities in various cell lines may be due, in part, to interactions between the enhancer and promoter regions. It is possible that differences in
the expression, binding affinities, or transactivation domains of cell specific factors may contribute to the observed results.

**Analysis of the zebrafish CYP1A Proximal Promoter**

Since the different levels of response exhibited between the mouse and zebrafish regulatory regions may be due, in part, to their respective proximal promoters, it was of interest to characterize the zebrafish CYP1A promoter region. To do this, successive truncations of the zebrafish CYP1A promoter region were cloned upstream of pGL3 Basic, transfected into Hepa-1 cells treated with either DMSO or TCDD, and assayed for luciferase activity. The results shown in Figure 2.14 indicate that up to three regions are required for maximal promoter function while one region may have an inhibitory effect on transcription. The loss of the region between -580 and -490 resulted in a 40% decrease in activity while further truncation down to -439 results in an additional 10% loss. Transcription Element Search Software (TESS) identified two Sp1 binding sites and a CTF/NF1 binding site within these regions which may be responsible for the decrease in activity. Interestingly, loss of the region between -439 and -398 caused an increase in activity back to the levels yielded by p-580/+71 Basic suggesting that an inhibitory element may reside within this region. *In silico* analysis identified an Sp1 site overlapping a USF1 site within these 40 bases which are possibly involved in the observed inhibition. Further truncation down to -206 resulted in another 50% decrease in activity which may be due to the loss of putative ERα and HNF-3γ sites located within the lost region. CTF/NF1 sites and G-boxes have both been implicated in the control of
Fig. 2.14. Analysis of the zebrafish CYP1A proximal promoter. The indicated plasmids were transfected into Hepa-1 cells and treated with TCDD (2 nM) or DMSO (0.05%) for 15 h. Luciferase activity was measured using a Turner Instruments luminometer and $\beta$-galactosidase activity was determined by spectrophotometry (OD 420). The normalized relative luciferase units are shown for the indicated plasmids. Open bars represent control treated cells while black bars indicate TCDD treated cells. Each bar = mean±S.E. from three independent experiments. The schematics to the left represent the portions of the promoter contained within the respective construct. Shaded circles = TATA box; Rectangles = CTF/NF1 sites; Triangles = ER$\alpha$ sites; Diamonds = Sp1 sites.
the mouse CYP1A1 promoter as well as the existence of a possible inhibitory region. (Jones and Whitlock, 1990) Further studies will have to be performed to characterize the importance of these elements in the zebrafish CYP1A promoter.

**Functional analysis of XRE flanking sequence**

Previous results suggest that just the presence of a core 5’-GCGTG XRE sequence is not sufficient to ascribe function to a putative XRE in vivo (Denison et al., 1988). For example, zfXRE5 has the same orientation and core sequence as zfXRE4, but does not associate with AHR/ARNT dimers in vitro and does not appear to participate in AHR-mediated regulation of zfCYP1A in cell culture. Thus, it was of interest to determine whether nucleotides flanking XRE5 contribute to the lack of function of this sequence in cell culture and in vitro. To gain some insight into this question, the sequences of all eight zfXREs as well as the six XREs present in the mouse CYP1A1 were aligned and compared (Fig. 2.14). As previously detailed by Swanson et al. (Swanson et al., 1995) and others, it can be observed that all XREs contain the 5’GCGTG core at positions −2 through +3, however, those XREs with defined activity in vivo also show consensus residues at positions 4, 5, 6 and 8. In contrast, XRE1, XRE2, XRE3, XRE5 and XRE6 as well as mouse XRE C that lack function in vivo, do not fit the consensus at residues 6 and 8. Thus, in vitro mutagenesis was used to change T >A at position 6 and T >G at position 8 in XRE5 so that it more resembled XRE4 (termed XRE5 > 4). In addition, the converse changes were made in XRE4 to convert it to XRE5 (termed XRE4 > 5). To assess whether the changes affected the ability of AHR/ARNT dimers to associate with the sequences in vitro, EMSA was utilized. The results in Figure
### Fig. 2.15. Comparison of mouse and zebrafish XRE flanking regions.
Core XREs and flanking sequences from zebrafish CYP1A (XREs: 1–8) and mouse CYP1A1 (XREs: A–F) were aligned. Shaded region indicates conserved nucleotides. Nucleotide positions are numbered below and ability of the XRE to bind in vitro or function in cell culture is indicated to the right, Y: yes, N: no. Nucleotides in non-functional XREs which diverge from conserved bases are circled.

<table>
<thead>
<tr>
<th>XRE</th>
<th>Nucleotide Sequence</th>
<th>Bind In Vitro</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CATGTAATGTCGCGTGCT</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>GAGTTTGCATGGCGCGCT</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>TGTATGAGTGTCGCTGA</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>4</td>
<td>GGTAAAGCCTCGCGTGA</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>5</td>
<td>TGAGAGTGTCGCTGCTC</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>6</td>
<td>AGTTATCGTGCTGCTC</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>7</td>
<td>CAAAGACAGCTCGCGCT</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>8</td>
<td>TTGGCCACGCTGCTGGAGG</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>A</td>
<td>AGTCCAGCTCGCGTGAGA</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>B</td>
<td>GGAACCTCGCTGCTGCCA</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>C</td>
<td>TCTGAGGCTAGCGTGCCT</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>D</td>
<td>GCCGCGGAGGTTGGCTGA</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>E</td>
<td>TCCCCCAGCTAACGCTGAC</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>F</td>
<td>GGGCGGCGGTTGGCGATG</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Model</td>
<td>NCGCGTGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consensus</td>
<td>TNGCGTGC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.15A confirm that XRE5 does not associate with AHR/ARNT dimers with the same efficiency as zfXRE4 and the XRE4 > 5 mutations at positions 6 and 8 do not appear to affect binding in vitro. In contrast, when the binding of XRE5 > 4 is compared to that of XRE5, there is an approximate 50% increase in association with AHR/ARNT. It was next pertinent to determine whether the various nucleotide changes results in increased or decreased ability to drive luciferase expression in cell culture. Thus, the various constructs were introduced into Hepa-1 cells and the overall level of TCDD-induced luciferase activity compared. Interestingly, mutation of XRE4 > 5 slightly reduced the level of luciferase induction in comparison to p-2608/−2100Ur(−78), but the level of change was not significant over several experiments. (Fig. 2.15B) Likewise, mutation of XRE5 > 4, resulted in a slight elevation of the luciferase activity and overall level of fold induction, but the elevation was never significant. Thus, the XRE5 > 4 mutation functions like the XRE8, in that it can associate with AHR·ARNT dimers in vitro, but does not appear to support higher levels of TCDD-mediated induction of luciferase in the absence of other XREs. Collectively these results suggest that nucleotides within the 3’-flanking sequence of the core XRE can influence association of AHR·ARNT in vitro, but changes in culture may be too subtle to detect.

The role of nucleotides flanking the core XRE motif in binding of AHR/ARNT dimers has been evaluated by a number of labs (Shen and Whitlock, 1992; Swanson et al., 1995). For example, Swanson et al. (Swanson et al., 1995) have shown that binding in vitro does not occur when a G is located at position 4. In addition, Shen and Whitlock (Shen and Whitlock, 1992) have shown that an A or C must be present at position 4 and a
Figure 2.16. Effect of Mutating Nucleotides Flanking XREs on AHR/ARNT Binding

**Binding In Vitro**

![Figure 2.16](image)

**Fig. 2.16. Effect of mutating nucleotides flanking XREs on AHR/ARNT binding in vitro.** (A) EMSA autoradiograph shows the ability of *in vitro* translated zebrafish AHR2 and rainbow trout ARNTb to bind synthetic double-stranded oligonucleotides containing zebrafish XRE4 (4), zebrafish XRE5 (5), XRE4 with mutations at positions 6 and 8 (4 > 5), or XRE5 with mutations at positions 6 and 8 (5 > 4). Arrow indicates the AHR·ARNT·DNA complex. (B) The indicated plasmids were transfected into Hepa-1 cells and treated with TCDD (2 nM) or DMSO (0.05%) for 15 h. Luciferase activity was measured using a Turner Instruments luminometer and β-galactosidase activity was determined by spectrophotometry (OD 420). The normalized relative luciferase units are shown for the indicated plasmids. Open bars represent control treated cells while black bars indicate TCDD treated cells. Each bar = mean±S.E. from three independent samples. The number in parentheses indicates the fold-change for TCDD samples compared to DMSO treated controls. a. Statistically different from p-2608/−2100Ur (*p* < .001).
G or C must be present at position 5 for function in cell culture. These investigators also suggested that a G abolishes function at position 6 (Shen and Whitlock, 1992). These rules partially explain why a number of the putative zfXREs are non-functional, as zfXREs 1, 2, 3, and 6, have T or G at position 4, and all have a T at position 6 (as does the non-functional mouse XRE C). However, XRE5 is also non-functional yet matches the consensus at positions 4 and 5. Importantly, when the residues at positions 6 and 8 were changed to those found in the functional XRE4, there was a significant increase in the binding to AHR/ARNT dimers *in vitro*, although the binding did not reach the level associated with wild type XRE4. (Fig. 2.15A) However, when these same mutations were made at XRE5 in the p-2608/–2100Ur reporter construct, there was a slight, but not significant elevation in TCDD-mediated induction of luciferase activity compared to controls. (Fig. 2.15B) Thus, it appears that bases at positions 6 and 8 may play a role in DNA binding *in vitro*, but the changes do not support function in the model system used to assess activity. This hypothesis is supported by the correlate studies that changed the bases at positions 6 and 8 in the functional XRE4 so they mimicked those of XRE5. In this case, the changes did not affect either the *in vitro* binding or activation in cells. Indeed, others have suggested that function is abolished when an A is located at position 5 (Shen and Whitlock, 1992), however, zebrafish XRE4 binds AHR/ARNT and is functional yet has an A at position 5. Thus, to truly assess the function of individual XREs, it will be necessary to determine binding to specific sites at the endogenous gene.
Identification of additional \textit{cis}-regulatory regions which impact the induction of \textit{zfCYP1A}

Since not all of the putative XREs in the \textit{zfCYP1A} regulatory region were functional and since the pattern of expression varied in different cell lines, it was of interest to determine whether additional \textit{cis}-regulatory regions were involved in the TCDD-mediated induction of the CYP1A gene. Analysis of the \textit{zfCYP1A} enhancer by Transcription Element Search Software (TESS) revealed an extensive list of putative binding sites within the distal 500 bp enhancer region which contains XREs 4-8. To pare down the number of sites to a more manageable list, TESS was used to evaluate the mouse CYP1A1 enhancer and identify common sites that shared similar relative positions within the enhancers of both genes. A summary of sites identified is shown in Figure 2.16. To begin to assess the function of the various sites, \textit{in vitro} mutagenesis was used to modify an Sp1 site located at $-2474$, HNF-3 sites located next to XRE7 and XRE5, an AP2 site, and a CREB site. Each site was mutated within the wild type p-$2608/-2100$ Ur construct and evaluated for TCDD-inducible activity in the Hepa-1 cell line. Interestingly, mutation of the AP2, Sp1 and proximal HNF3 sites did not cause a change in either the magnitude or level of induction (Fig. 2.17). In contrast, mutation of the putative CREB binding site caused a modest reduction in the both the basal and TCDD-induced luciferase activity. Similarly, mutation of the HNF-3 binding site located at position $-2547$ caused an even more dramatic reduction in both the basal and TCDD-induced luciferase activity that was reduced 10-fold below the level of the p-$580/+71$ construct. This finding is intriguing as it suggests that the removal of the site acts to
Fig. 2.17. Putative transcription factor binding sites within the zfCYP1A regulatory region that are common to both mouse and zebrafish. Putative transcription factor binding sites identified by TESS within the zfCYP1A enhancer that are common to both mouse and zebrafish. Specific binding regions are boxed and labeled. Binding sites that were targeted for mutagenesis studies are double boxed.
Fig. 2.18. Analysis of cis-regulatory elements within the zfCYP1A enhancer.
The indicated plasmids were transfected into Hepa-1 cells and treated with TCDD (2 nM) or DMSO (0.05%) for 15 h. Luciferase activity was measured using a Turner Instruments luminometer and β-galactosidase activity was determined by spectrophotometry (OD 420). Open bars represent control treated cells while black bars indicate TCDD treated cells. *Bar is statistically different from p-2608/−2100Ur. The transcription factor denoted in the construct name indicates that the specific transcription factor binding site has been mutated in that construct. –HNF-3(5) and –HNF-3(7) indicates HNF-3 binding sites flanking XRE 5 and XRE 7, respectively. The fold induction is shown in the graph on the right and was determined by dividing the normalized RLU of samples treated with TCDD by the normalized RLU of control treated samples. *Statistically different from p-2608/−2100Ur (p < .001).
repress the overall activity of the construct. Indeed, HNF-3 has been implicated in chromatin remodeling functions (Roux et al., 1995), nucleosome positioning (Shim et al., 1998), and has a supportive role in xenobiotic-mediated transcriptional regulation (Bombail et al., 2004; Rodriguez-Antona et al., 2003). It is quite possible that in the absence of protein binding to this site, the nucleosome is shifted toward the promoter and thereby blocks access of the transcriptional machinery from assembling. Another possibility is that mutation of this site causes a bending of the DNA such that the enhancer itself contacts the promoter blocking both assembly of transcriptional machinery as well as AHR/ARNT from accessing the XREs. Furthermore, HNF-3 has been shown to work cooperatively with C/EBP (Christoffels et al., 1998) and NF1/CTF (Jackson et al., 1993) in regulating the carbamolyphosphate synthetase I (CPS) and serum albumin genes, respectively. Since both C/EBP and NF1/CTF binding sites can be found in the zfCYP1A promoter/enhancer as determined by in silico analysis, it is interesting to speculate that such cooperative interactions with HNF-3 occur in this system as well.

To determine what effect the HNF-3 mutation had on a construct containing the full-length zfCYP1A regulatory region, the p-2608/+71(-HNF-3(7)) construct in which the PCR amplified region between -2100 and -2608 from p-2608/-2100Ur(-HNF-3(7)) was used to replace the same region of the p-2699/+71 construct was transfected into Hepa-1 cells and evaluated for luciferase activity. The results in Figure 2.18 indicate that contrary to the mutation in the p-2608/-2100Ur construct which resulted in ablation of both basal and inducible activity, the mutation in the p-2608/+71 construct resulted in a 6 and 8-fold increase in basal and inducible activity, respectively. The overall fold-induction between p-2608/+71 and p-2608/+71(-HNF-3(7)), however was consistent.
Figure 2.19. Effect of HNF-3 Mutation on the Full-length p-2699/+71 Construct

Fig. 2.19. Effect of HNF-3 mutation on the full-length p-2699/+71 construct. The indicated plasmids were transfected into Hepa-1 cells and treated with TCDD (2 nM) or DMSO (0.05%) for 15 h. Luciferase activity was measured using a Turner Instruments luminometer and β-galactosidase activity was determined by spectrophotometry (OD 420). Open bars represent control treated cells while black bars indicate TCDD treated cells. Each bar = mean±S.E. from three independent samples. The fold induction was determined by dividing the normalized RLU of samples treated with TCDD by the normalized RLU of control treated samples.
This data supports a role for HNF-3 in maintaining the position of the nucleosome as, if
the loss of the site in the context of p-2608/-2100Ur results in a shifting of the
nucleosome 3'-ward such that it blocks assembly at the promoter then the loss of the site
in the full-length construct may shift the nucleosome into the 1.7kb AT-rich region which
lacks any transcription factor binding sites. Positioning of the nucleosome within this
region may open up the upstream enhancer region leading to the enhanced activity
observed. While the mechanisms behind the observations at this site are unclear, the
results clearly suggest that binding of AHR/ARNT alone is likely not sufficient to obtain
a full transcriptional response and other binding proteins are required. This may explain
the large differences in overall level of luciferase induction of the various reporters in the
different cell lines and the need to assess gene regulation in the proper cellular context.
Future studies should be designed to assess the binding of additional factors to the
CYP1A enhancer in genomic DNA. The large distance between the enhancer and
promoter regions of the zebrafish CYP1A gene make this ideal for procedures such as \textit{in}
vivo footprinting and ChIP analyses as binding at the two regions should be easily
differentiated. It is anticipated that the future analysis of the \textit{zfCYP1A} gene will provide
significant information on how the AHR mediates gene regulation in both aquatic and
mammalian organisms.
Chapter 3

Discussion of Impact, Relevance, and Future Direction

The data presented herein confirms the isolation of the upstream regulatory region for a dioxin-inducible CYP1A gene in zebrafish. The identified region, which was sequenced approximately 2.8 kb upstream of the transcriptional start site, contains eight XREs and several other elements indicative of previously characterized CYP1A regulatory regions including a TATA box positioned at -31. Sequence analysis revealed that the XREs were organized into two distinct clusters. The proximal cluster of XREs (XREs 1-3) was incapable of conveying TCDD-mediated induction to a luciferase reporter gene while the distal cluster, containing XREs 4-8, enhanced luciferase expression in the presence of dioxin more than 20 fold. As determined by electrophoretic mobility shift assays, only a sub-set of the eight XREs were capable of binding AHR/ARNT \textit{in vitro}. Surprisingly, XRE1 and XRE3 were bound by AHR/ARNT \textit{in vitro} in a TCDD-dependent fashion despite their inability to drive luciferase expression in cell culture. The reason for the lack of function exhibited by these XREs is currently unclear. Considering the region between -580 and -187 containing XREs 1-3 failed to convey induction when cloned immediately upstream of an SV-40 promoter or in the context of its native promoter, it is unlikely that the lack of functionality is promoter specific. Additionally, when the region containing XREs 4-8 was cloned in the same
position upstream of an SV-40 promoter, it was capable of conveying significant levels of induction indicating that the distance from the promoter is also not responsible the inability of XRE1 and 3 to function. It is possible that XREs 1 and 3 are capable of binding protein *in vitro* but fail to function in cell culture due to chromatin condensation causing inaccessibility of the binding sites. To assess this possibility, tandem repeats of XRE1 or XRE3 along with several flanking nucleotides could be cloned upstream of a luciferase cassette driven by an SV-40 promoter and analyzed for TCDD-mediated luciferase induction. In such a scenario, the absence of additional native DNA which may recruit histones or influence the structure of the DNA will allow for the assessment of the XRE alone to enhance transcription. Others have characterized the mouse CYP1A1 XREs and found that XREC was also capable of binding AHR/ARNT *in vitro* but failed to function *in vivo* (Lusska et al., 1993). In these studies, tandem repeats of each of the six XREs were placed upstream of an MMTV promoter controlled chloramphenicol acetyltransferase cassette and even in such a context, XREC still failed to drive CAT expression. This suggests that in the mouse, chromatin inhibition of the XRE is likely not responsible for the lack of function. Another possibility for the failure of the XREs to contribute to gene induction may lay in the nucleotides flanking the core XRE sequence. Previous experiments have shown that an A is required at position 6 (Fig. 2.14) for XRE function in the mouse but not for *in vitro* binding (Lusska et al., 1993; Shen and Whitlock, 1992). These results would help explain the lack of function of XRE1 and XREC, but not XRE3 which contains an A at that position. It is important to note however that XREs 1 and 3 also diverge from the consensus suggested in Figure 2.14 at positions 4 and 8 as well. Although it has been previously shown through the use
of ligation mediated PCR that the AHR/ARNT complex likely only interacts with three or four guanine residues within the core XRE sequence, (Wu and Whitlock, 1993) flanking nucleotides may affect the ability or degree to which the DNA bends or may reinforce the stability of the bound AHR/ARNT complex and recruited cofactors. Utilizing \textit{in vitro} mutagenesis to alter the nucleotides at positions 4, 6, and 8 in XRE 1 and XRE3 within the p-580/+71 construct could shed light on the importance of these flanking nucleotides on XRE functionality. Furthermore, \textit{in vivo} footprinting and ligation mediated PCR should be performed on the zebrafish CYP1A regulatory region to determine whether or not protein interacts with the XREs in intact cells.

Nucleotides flanking the core XRE may also be implicated in the inability of XREs 5 and 6 to bind AHR/ARNT \textit{in vitro} or function \textit{in vivo}. Inaccessibility of the XREs due to chromatin structure is unlikely in the case of XREs 5 and 6 since it has been previously established that AHR/ARNT binding at an XRE results in chromatin relaxation spanning approximately 200 bases around the XRE (Okino and Whitlock, 1995). Knowing that activated AHR likely binds to both XRE4 and XRE7, XREs 5 and 6 should be free of any chromatin related constraints in the presence of TCDD. Chromatin interference can further be ruled out by the fact that AHR/ARNT does not bind double stranded oligonucleotides containing XREs 5 and 6 \textit{in vitro}. Failure to bind \textit{in vitro} using \textit{in vitro} translated AHR and ARNT also suggests that inhibition is not being caused by competition or hindrances from additional proteins binding nearby the XRE. The observations stated above suggest that any constraints set upon AHR/ARNT-XRE interactions exist within the DNA sequence itself. Like XREs 1 and 3, mentioned
previously, XREs 5 and 6 both possess a T at position 6 (Fig. 2.14). A T at this position, however, has been implicated in loss of XRE function but has not been shown to affect protein binding (Lusska et al., 1993). The alignment in Figure 2.14 shows that XRE6 also diverges from functional XREs at positions 4, 5, and 8 while XRE5 diverges only at positions 6 and 8. Experiments described previously in this work suggest that mutations of nucleotides at positions 6 and 8 can enhance the ability of XRE5 to bind protein in vitro however the enhanced binding was not to the level observed by functional XREs. Furthermore, the reverse mutations made to the functional XRE4 did not appear to affect the level of binding observed in EMSA. These findings suggest that additional flanking nucleotides may be involved in protein binding and is supported by the fact that statistically significant changes in luciferase activity were not observed when the same mutations were made in reporter constructs. Swanson et al. have previously characterized XRE flanking nucleotides but their studies only examined in vitro interactions using synthetically generated oligonucleotides (Swanson et al., 1995). Wu and Whitlock have also characterized the role of flanking nucleotides in mouse XREs but their studies did not examine nucleotides extending beyond four bases around the core XRE (Wu and Whitlock, 1993). Utilizing a combination of synthetically generated double stranded oligonucleotides for EMSA and in vitro mutagenesis to alter flanking nucleotides of XRE5 alone and in combination until they are identical to a functional XRE may shed light on the precise nucleotide combination required for a functional XRE. Species specificity could likewise be tested by performing the same experiments on mouse CYP1A1 XREs and by using mouse AHR and mouse ARNT1 in EMSA experiments. Determination of specific nucleotide sequences required for AHR function
in transcriptional regulation could be greatly beneficial in the search for novel AHR target genes which contain *bona fide*, functional XREs within their regulatory regions.

It is also of interest to note that the XRE5 exhibits weak binding to AHR/ARNT in EMSA experiments after prolonged exposure. Characterization of the mouse CYP1B1 promoter has previously revealed that three of the five XREs located therein bound a non-AHR/ARNT complex which ran identically to AHR/ARNT on EMSA (Zhang et al., 1998a). These XREs diverged from AHR/ARNT binding XREs within the same region only at positions 6 and 7. While XRE5 in the zfCYP1A regulatory region does not share the same nucleotide sequence at these two positions, it would be of interest to repeat the EMSA of XRE5 with anti-AHR antibodies to ensure that the observed binding is indeed AHR/ARNT and not another protein present in the reticulocyte.

XRE4, XRE7, and XRE8 were capable of both binding *in vitro* and functioning to induce luciferase in TCDD treated cells transfected with luciferase reporter constructs. While all three of these XREs appear to be necessary for maximal induction, their contributions do not appear to be equal. When luciferase reporter assays were carried out using p-2608/-2100Ur constructs bearing mutations at each of the individual XREs, it was observed that the loss of XRE4 resulted in an approximately 50% reduction in activity compared to the non-mutated control while loss of XRE7 yielded a 70% loss of activity and the mutation of XRE8 only decreased activity by approximately 30%. These findings are contrary to previous studies which suggested that each of the mouse XREs contribute equally to maximal induction (Fisher et al., 1990). Later studies by Lusska et al. showed that when duplicate copies of each of the six mouse XREs, A-F, were placed
upstream of an MMTV promoter driven CAT reporter, XREC failed to respond to TCDD and XREA exhibited a considerable reduction in responsiveness compared to XREs B, D, E, and F (Lusska et al., 1993). Unfortunately, the experiments utilized by the investigators in these studies did not consider the enhancer elements within their native environment nor did they evaluate cooperation between the XREs. Mutations of individual mouse XREs in the context of the full-length mouse CYP1A1 regulatory region have not been made, thus it is unknown what contributions mouse XREA and XREC make toward maximal induction of that gene. The results presented in this report, however, show that in the absence of additional functional XREs, XREs 4 and 7 can contribute 25-30% of the induced activity observed in the non-mutated p-2608/-2100Ur while XRE8 is incapable of functioning alone. This is the first time it has been shown that an XRE regulating a cytochrome P-450 was required for maximal gene induction but incapable of functioning in the absence of additional functional XREs. The reasons behind this observation are currently unknown. It is possible that the sequence flanking the XRE is responsible for the inability of XRE8 to function alone although this is unlikely due to the fact that XRE only diverges from XRE7 at positions 4, 7, and 8. The nucleotides found at these three positions within XRE8 are also found at the same positions in other functional mouse and zebrafish XREs. It is additionally possible that XRE8 is inaccessible to AHR/ARNT due to the presence of chromatin prior to chromatin relaxation caused by AHR/ARNT binding at XRE4 or XRE7. In order to test these hypotheses, duplicate copies of XRE8 should be cloned upstream of an SV-40 or MMTV promoter driven luciferase promoter and assayed for TCDD-mediated induction. Additional inhibitory factors such as chromatin should not be present in such a construct.
and if the XRE sequence is sufficient to regulate transcriptional activity, TCDD should mediate induction of the reporter. Alternately, mutation of the nucleotides at positions 4, 7, and 8 within XRE8 in the context of p-2608/-2100Ur(-47) such that they are identical to XRE7 should shed light on the importance of nucleotide specificity. The converse mutations should also be made to XRE7 in the context of p-2608/-2100Ur(-48). If sequence is important, the XRE7 mutant should lose functionality while XRE8 should gain the ability to function alone. Finally, analysis of the p-2608/-2100Ur XRE mutants should be revisited in Hepa-1 cells treated with a histone deacetylase inhibitor such as trichostatin-A. These studies should lessen or eliminate any inhibition chromatin has on the ability of AHR/ARNT to bind XREs within the reporter constructs.

Another observation made throughout the course of these studies is that XRE reporter constructs behave differently in different cell lines. When various reporter constructs containing regulatory regions from zebrafish CYP1A and mouse CYP1A1 were transfected into various cell lines from different species and tissues, a distinct pattern of transcriptional regulation was observed. For instance, when the zebrafish p-2727/+71 construct was transfected into the human A498 kidney cell line it yielded approximately 20-fold induction when treated with TCDD. By comparison the mouse p-1647/+57 construct yielded less than 10-fold induction under the same conditions. Surprisingly however, in the C2C12 mouse myoblast cell line, transfection of p-2727/+71 led to only an approximately 10-fold induction of reporter while a 60-fold induction was observed in cells transfected with the mouse p-1647/+57. This overall trend was observed in several cell lines with the zebrafish constructs yielding considerably higher
levels of induction in the Hepa-1, A498, and B19 cells while the mouse constructs were favored in the C2C12, TCM, HepG2 and hRPE cell lines. The reasons behind this observation are unclear, however it does not appear to correlate with the AHR allele, the species or tissue of origin, or the presence or absence of ARNT2. Importantly, the data suggests that the differences observed may be dependent upon the proximal promoter.

Each of the cell lines were also transfected with the pMm-1315/-819Uf construct which contains the indicated mouse CYP1A1 enhancer region immediately upstream of the zebrafish CYP1A promoter region, -580/+71. In the Hepa-1, B19, and A498 cells which appear to favor the zebrafish full-length construct over the mouse, pMm-1315/-819Uf yielded approximately 20% less induction than the full length mouse, p-1647/+57, in the same cell lines. These results are expected considering the full-length mouse construct contains two additional functional XREs not present between -1315 and -819 which should result in a 20% reduction if all 5 XREs contribute equally as reported previously. (Fisher et al., 1990) In the cell lines which appear to favor the mouse constructs, however, pMm-1315/819Uf exhibits >50% reduced fold-induction compared to p-1647/+57. This difference is most noticeable in the C2C12 and HepG2 cell lines which exhibit greater overall levels of activity. To better understand the role the proximal promoter plays in these observed differences, additional constructs should be made which contain the zebrafish -2608/-2100 enhancer region immediately upstream of the mouse proximal promoter region and the mouse enhancer region between -1315 and -819 upstream of the mouse proximal promoter. A comparison of the luciferase activity exhibited by these constructs as well as p-2608/-2100Ur and pMm-1315/-819Uf should confirm the importance of the promoter region in begetting these differences.
The mouse CYP1A1 promoter region has been previously characterized by others (Jones and Whitlock, 1990). The results of these studies revealed that the mouse proximal promoter region contains a TATA box located at position -30, two CTF/NF1-like sites located at positions -59 and -136, and a G-box located at position -130. Mutation analyses showed that loss of either the TATA-box or the proximal CTF/NF1 sequence resulted in an 80% decrease in promoter function. Additionally, simultaneous loss of the distal CTF/NF1 site and the G-box resulted in a 50% decrease of promoter function while individual mutations of these elements had little or no effect. Truncation experiments also suggest that there may exist inhibitory elements located between positions -419 and -246. Transcription Element Search Software (TESS) was used to identify putative transcription factor binding sites within the zebrafish CYP1A proximal promoter region. A TATA-box was identified at position -31 along with several pertinent sites including putative CTF/NF1 sites, G-boxes and Sp1 sites. A proximal CTF/NF1 site was identified by TESS at position -54 and it is likely that the zebrafish promoter is similar to the mouse CYP1A1 promoter and also contains a distal CTF/NF1 site and a G-box which are required for maximal promoter function. Luciferase reporter vectors containing successive truncations of the zebrafish proximal promoter suggest that two distal CTF/NF1 sites and two distal Sp1-like sites may be important for the transcriptional activity of the downstream gene. Additionally, the data suggests that another element lies between bases -398 and -206 which is also required for full promoter activity. TESS identified a putative ERα binding site at position -330 and a putative HNF-3 binding site at position -225 both of which have previously been implicated in the
transcriptional regulation of cytochrome containing xenobiotic metabolism enzymes (Bombail et al., 2004; Gibson et al., 2002; Matthews et al., 2005; Rodriguez-Antona et al., 2003). Finally, the data suggests that an inhibitory element may exist between bases -439 and -411. An Sp1 site was identified overlapping a USF1 site within this region by in silico analysis and may have a role in transcriptional inhibition of the gene.

Overall, the zebrafish CYP1A proximal promoter greatly resembles the previously characterized mouse CYP1A1 promoter. Many additional studies still need to be performed to better characterize this region, however. In vitro mutagenesis needs to be utilized to mutate target binding sites within the promoter region both alone and in combination to assess the function of the sites since truncation analyses alter the physical characteristics of the DNA and cannot, therefore, be reliable. Mutations should also be carried out in the context of p-2608/-2100Ur and assayed in TCDD treated cells to obtain greater levels of overall activity and thus, more substantial differences between mutants and control samples. EMSA should be performed to identify the ability of the pertinent elements to bind protein in whole cell lysates and nuclear extracts from both DMSO and TCDD treated cells. It is possible that the different patterns of regulation observed between the constructs containing mouse and zebrafish elements is due to differential binding to elements within the proximal promoters. In order to assess this, EMSA can be performed using oligonucleotides containing elements of interest from both mouse CYP1A1 and zebrafish CYP1A and nuclear extracts from each of the seven cell lines used previously to determine differences in relative protein/DNA binding affinities between the cell lines. DNAse in vitro footprinting should be performed as well using
nuclear extracts from the seven cell lines of interest to further characterize differences in protein binding at the two promoters. *In vivo* footprinting and LMPCR should ideally be performed in ZFL cells to confirm protein binding at the zfCYP1A promoter in intact cells.

TESS was also used to identify putative transcription factor binding sites which were common between the mouse CYP1A1 and zebrafish CYP1A distal enhancers. *In vitro* mutagenesis of selected sites within p-2608/-2100Ur and subsequent luciferase assays identified both a CREB and HNF-3 sites which were important to the transcriptional activity of the downstream gene. Previously, no additional proteins have been implicated in binding at the enhancer other than AHR/ARNT and Sp1 (Fisher et al., 1990) however mutation of the CREB binding site results in modest reductions of both basal and induced reporter activity. Even more surprising was the finding that mutation of an HNF-3 site located at position -2547 completely ablates both basal and inducible transcriptional activity. This finding suggests that loss of the putative HNF-3 site somehow inhibits the assembly of the transcriptional machinery at the promoter in the presence or absence of TCDD. To investigate the role of the site in more detail, the mutation of the HNF-3 site was assessed in the context of the full-length, p-2608/+71 construct. Unexpectedly, mutation of the HNF-3 site in this construct resulted in a significant 6-fold increase in both basal and induced activity while only slightly elevating the overall fold induction. The conflicting results obtained from the experiments using the two different constructs led to the formulation of a hypothesis that HNF-3, or an HNF-3-like protein is responsible for the positioning of a nucleosome at the CYP1A
enhancer. In such a scenario, it can be envisioned that the nucleosome is shifted 3’-ward such that in the case of p-2608/-2100Ur, the nucleosome blocks assembly at the promoter while in p-2608/+71, the nucleosome is shifted into the 1 kb AT-rich region located between the proximal promoter and the enhancer region, opening up enhancer and leading to super-induction. Indeed, HNF-3α has been implicated by others in the positioning of the nucleosome (Shim et al., 1998) and in chromatin remodeling, (Roux et al., 1995) supporting this hypothesis. Fascinatingly, two similar binding motifs can be found in the mouse CYP1A1 regulatory region; both of which exist centrally within 200 bp spans known to undertake a nucleosomal configuration. HNF-3 is known to bind nucleosomal DNA (Shim et al., 1998) and this may explain why these sites may have been protected from DNaseI digestion in footprints of the mouse CYP1A1 promoter and enhancer regions (Okino and Whitlock, 1995; Shen and Whitlock, 1992). To test this hypothesis, reporter assays using p-2608/-2100Ur and p-2608/+71 HNF-3 mutants should be repeated in cells treated with a chromatin inhibitor such as Trichostatin A. If the nucleosome is responsible for the effects observed in HNF-3 mutants then Trichostatin-A treatment should reduce or eliminate the inhibition observed in p-2608/-2100Ur HNF-3 mutants while super-induction should be observed in the non-mutated p-2608/+71. Additionally, EMSA using oligonucleotides containing the putative HNF-3 binding site and Hepa-1 nuclear extracts could be utilized to confirm binding at the site in vitro while super-shifting with antibodies against HNF-3α could confirm that is the protein responsible for binding the element. These experiments may not be successful, however, since the forkhead region of HNF-3α interacts directly with histones which would be
absent in an EMSA. In vivo footprinting could also be utilized to detect the nucleosome in ZFL cells treated with control or HNF-3α siRNA.

Overall, the structure and function of the zebrafish CYP1A regulatory region appears to be similar to that which has been reported for other piscine and mammalian CYP1As. Several novel findings have been identified for this model organism, however. Within this report, it has been shown that there exists an inequality of the contribution toward maximal induction exhibited by the functional XREs within this regulatory region which contradicts what has been previously established in the mouse. Furthermore, previous studies have not identified an XRE which is required for maximal induction yet which fails to function in the absence of additional functional XREs as was observed for the zebrafish XRE8. These findings are significant in that they support a better understanding of the mechanisms of AHR/ARNT-XRE binding and subsequent gene regulation. Also within this report, it has been established that the degree of regulation mediated by CYP1A regulatory regions from different organisms is inconsistent between cell lines. The data suggests that these inconsistencies may be due, at least in part, to selective binding of proteins at the promoter. This novel finding may help elucidate how AHR target genes are differentially expressed in various tissues and organisms; especially as pertains to the application of findings obtained using cell culture and model organisms to human health. Finally, a putative HNF-3 binding site was identified within the zebrafish CYP1A enhancer region which may have a role in nucleosome positioning based on mutational analyses. Chromatin reorganization is a fundamental aspect of CYP1A regulation and an understanding the molecular mechanisms of nucleosome
arrangement in this regulatory region is essential to discerning the AHR-mediated control of this gene.

The characterization of the zebrafish CYP1A regulatory region was an important step toward using *Danio* as a model organism for the study of AHR-mediated signaling and TCDD-toxicity. Apart from the numerous benefits this organism offers for developmental studies, the zebrafish CYP1A gene may be useful for future studies of AHR-mediated regulation of target genes. Importantly, the distance between the promoter and enhancer regions in zebrafish CYP1A is far greater than observed in previously characterized organisms, thus making it ideal for procedures such as ChIP assays in that binding can be easily differentiated between the two regions. Additionally, the zebrafish can be used as a bio-detector of environmental pollutants such as TCDD by the creation of transgenic fish which can produce a fluorescent signal in the presence of such chemicals at a dose dangerous to human health. Such methods would be far more accurate and reliable than bio-chemical testing which is often inaccurate and incapable of knowing precisely the doses which may pose a threat to human health.
Chapter 4

Materials and Methods

Materials

TCDD (98% stated chemical purity) was obtained from Radian Corp. (Austin, TX) or Cambridge Isotope Laboratories and was solubilized in dimethylsulfoxide (DMSO).

Buffers

PBS is 0.8% NaCl, 0.02% KCl, 0.14% Na2HPO4, 0.02% KH2PO4, pH 7.4. Lysis buffer is 60 mM Tris, pH 6.8, 2% SDS, 15% glycerol, 2 mM EDTA, 5 mM EGTA, 10 mM DTT, 5% NP-40, 20 mM sodium molybdate, 0.005% bromphenol blue. TBS is 50 mM Tris, 150 mM NaCl, pH 7.5. TTBS is 50 mM Tris, 0.2% Tween 20, 150 mM NaCl, pH 7.5. TTBS_ is 50 mM Tris, 0.5% Tween 20, 300 mM NaCl, pH 7.5. BLOTTO is 5% dry milk in TTBS. Gel Shift Buffer is 50 mM Hepes, pH 7.5, 15 mM MgCl2, 50% glycerol.

Cells and growth conditions

Wild-type Hepa-1c1c7 (Hepa-1), were a generous gift from Dr. James Whitlock, Jr. (Department of Pharmacology, Stanford University). These cells were propagated in DMEM supplemented with 10% fetal bovine serum (FBS). Stable cell lines expressing the Ahb-2 AHR in the Hepa-1 background (B19) were propagated as detailed previously.
(Pollenz and Dougherty, 2005). ZFL cells were obtained from ATCC (Manassas, VA) and propagated in a 0.5/0.35/0.15 mixture of L15:DMEM:Ham-F12 supplemented with bovine insulin (10 mg/L), EGF (20μg/L) and heat-inactivated FBS (5%) at 28 °C. Human HepG2, human ARPE-19, human A498 cells, mouse TCM cells and mouse C2C12 cells were purchased from ATCC and propagated as detailed by the manufacturer. All cells were passaged at 1-week intervals and used in experiments during a 3-month period at approximately 70% confluence. For treatment regimens, TCDD was administered directly into growth media for the indicated incubation times. DMSO was used as the vehicle control and the final concentration present in the culture media was between 0.05 and 0.1%.

**Antibodies**

Specific antibodies against the zfAHR2 (zf-4) and rtARNTb (rt-84) are identical to those described previously. (Pollenz et al., 1996; Wentworth et al., 2004) All antibodies are affinity-purified IgG fractions. Antibodies specific to the zfARNT2b were purchased from Santa Cruz (Santa Cruz, CA). For Western blot analysis, goat-anti-rabbit antibodies conjugated to horseradish peroxidase (GAR-HRP) or rabbit anti-goat antibodies conjugated to horseradish peroxidase (RAG-HRP) were utilized (Jackson Immunoresearch, West Grove, PA).
Isolation of PAC clones containing the zfCYP1A gene

Oligonucleotide primers:

zfCYP1A1-UTR-upstream: 5’CTGGAAAGTATCCACTCGATCG3’

zfCYP1A1-ORF-downstream: 5’CCAGGACATTTCCGATAATCGG3’

were generated to the 5’UTR and ORF of the putative zfCYP1A mRNA (GenBank accession #BC094977). These primers were used to screen superpools of zebrafish genomic PACs, (Amemiya and Zon, 1999) by PCR. Two superpools displayed positive PCR products when visualized on a 2% agarose gel. The PAC 133 and 150 were then robotically dotted onto nitrocellulose filters and screened by colony hybridization using the 312 bp zfCYP1A1 cDNA fragment. One positive clone was isolated from each of the two superpools and termed #133 and #150. Southern blotting was used to identify fragments corresponding to the CYP1A gene, and these were subcloned into pBluescript SK- (Stratagene, Madison, WI) and sequenced. Sequence analysis and alignments were carried out using Lasergene software (DNAStar, Madison, WI). Analysis of regulatory elements was carried out using Transcription Element Search Software (TESS) freeware available through the University of Pennsylvania (http://www.cbil.upenn.edu/ tess). The sequence of the putative zfCYP1A promoter region (~2710 ± 65) that includes a portion of the 5’UTR and the splice site for intron 1 has been entered into GenBank (Accession# DQ182546).
**Generation of reporter constructs and site-directed mutagenesis**

PCR was used to amplify the indicated regions of the *zfCYP1A* promoter/enhancer from subcloned *SpeI* fragment of PAC #150. The PCR fragments were then ligated into pGL3promoter (Promega), to generate p-2608/-2100Fup, p-2608/-2100Rup, p-2608/-2100Fdown, p-2608/-2100Rdown, p-580/-187Fup, and p-580/-187F or ligated into pGL3Basic (Promega) to generate p-580/+71Basic, p-2608/-2100Uf, p-2608/-2100Ur, or p-2727/-2100Ur. Orientation was determined by restriction analysis. To generate constructs containing the full-length promoter/enhancer, the 12kb *SpeI* fragment was digested with *SacI* to yield a 2.7 kb fragment containing all of the previously examined XREs but terminating at −42 and missing the TATA box and transcriptional start site. To deal with this issue, the fragment was ligated into pSK- (Stratagene) and then cut out with *SacI* and *KpnI*. This fragment was ligated into the *SacI* and *KpnI* sites of p-2727/−2100Ur that had been cut to remove all but the +71 to −42 portion of the −2727/−2100Ur fragment. The full-length promoter/enhancer construct was termed p-2699/+71. The reporter vector containing the full promoter region from the mouse and the full promoter region from rainbow trout (p-1897Om1A3luc) were generous gifts from Dr. Michael Carvan (University of Wisconsin—Milwaukee) (Carvan et al., 1999). To generate Om-1897/-1392Uf, Om-1897/-1392Ur, Mm-1315/-819Uf, and Mm-1315/-819Ur, the
indicated regions were amplified by PCR and ligated into pGL3Basic (Promega) upstream of the -580/+71 region of the zfCYP1A.

Site-directed mutagenesis of specific cis regions of the zfCYP1A was carried out on the appropriate parental vectors using the Quickchange II XL site-directed mutagenesis kit using the manufacturer’s protocol (Stratagene). The primer sets used are listed with the specific base changes indicated in bold.

XRE1: 5’-CCATGTATGTGTGAGTGTGTACACATAC

5’- GTATGTACACACTCACACACATACATGG.

XRE3: 5’-CTCTCATTCACTCACACTCATACAC

5’-GTGTATGAGATGTGTGAGTGTGAATGAGAG.

XRE4: 5’-CACACCTTTGCCACTGATGCTTTACCTGTTGC

5’-GCAACAGGTAAGCAGCATCAGTGCAAAGGTTG.

XRE7: 5’-CAGGTGCGCGCAGCTCGATGCTGTGATC

5’-GATCAAACAGCATCGAGTGCAGCGCAGCT.

XRE8: 5’-CCTCCTCAGCTCACTCAACGTGGCGCAATC

5’-GATTGCGCAGTGTGAGTGCTGGAGGAGAG.

Sp1: 5’-CTTCCCATAAACCAACCGCAGAACAAC

5’-GTTTGTTCTGCAGGGTTTTATGGGAG.
HNF-3(7): 5’-GCACGCGATGCTTGAGGATCATGTTTATCGTAGC  
5’-GCTACGATAAACTGATCCACAGCATCGCGTGCC.

HNF-3(5): 5’-CGCACGCACATACTCTCACAC  
5’-GTGTGAGAGTATGTGCATGC.

AP2: 5’-CCAATCTTTTAAAGCGCTACAGGTC  
5’-GCACCTGTAGCGCTGTTAAAGGTTG.

CREB: 5’-CGACGGCCACGCGCTATACCCCATTCTGC  
5’-GCAGAATGGGATATAAGCGCGTGCGCGTCG.

XRE4(6/8): 5’-CACACACACACACTTATGACAGCGATG  
5’-CATCGCGTGCATAAGGTGTGGTGCTG.

XRE5(6): 5’-GCAGCGGTTTACCACATCGCAGCAGCAC  
5’-GTGTGCGTGCGATGTTGAACCGCTGC.

XRE5(8): 5’-GCAGCGGTTTACCACATCGCAGCAC  
5’-GTGTGCGTGCGATGTTGAACCGCTGC.
Transfection and reporter assays

Approximately $2 \times 10^5$ cells were plated into 35mm culture dishes and incubated at 37°C for 16–24 h. Transfection cocktails were set up so that multiple dishes could be transfected with aliquots of the same sample. This was accomplished by mixing the reporter plasmid and pSV-β-galactosidase in OptiMEM media (Gibco) and then incubating the mixture with either LipofectAMINE™ or LipofectAMINE 2000™ (Gibco) as specified by the manufacturer. Aliquots were then applied to the appropriate plates and after 24 h; the cells were exposed to TCDD or DMSO (0.05%) for 6–16 h. When multiple reporter plasmids were to be used in an experiment, the concentration of DNA utilized was the same and was verified by OD260 readings as well as agarose gel electrophoresis. Cells were harvested from plates by scraping directly into 200–400μl Reporter Lysis Buffer as specified by the manufacture (Promega). Luciferase activity was measured using identical sample volumes for 30 s in a Turner Instruments luminometer. β-Galactosidase activity was measured using the β-galactosidase assay kit as specified by the manufacturer (Promega). Typically, each data point was evaluated in triplicate and luciferase values were normalized by dividing the relative luciferase units (RLU) of each sample by the corresponding level of β-galactosidase activity (OD420 reading). In most experiments, the relative transfection efficiency for all plasmids was similar in a given cell line, although the level of transfection efficiency across cell lines was different. Results are presented as normalized luciferase units (raw) as well as fold-change between control and treated samples.
In vitro *expression of protein*

Recombinant protein was produced from zfAHR2, rtARNTb, and zfARNT2b expression plasmids using the TNT™ Coupled Rabbit Reticulocyte Lysate Kit essentially as detailed by the manufacturer (Promega). Upon completion of the 90 min reaction, samples were either combined with an equal volume of 2X gel sample buffer and boiled for 5 min, or stored at -80°C for use in functional studies. The actual concentration of protein expressed in each reaction is estimated to be 6ng/μL, based on previous studies.

In vitro *activation of AHR:ARNT complexes and electrophoretic mobility shift assay*

The following oligonucleotides containing the core zfXRE sequences (*bold*) and 5’GG3’ overhangs (*underlined*) were synthesized by IDT (Coralville, IA).

zfXRE1: 5’GGTGTAACACACGCACACATAC3’

zfXRE2: 5’GGGGAAACCCACGCCATGCAAA3’

zfXRE3: 5’GGATGAGTGTGCGTGTAATGA3’

zfXRE4: 5’GGTAAAGCATCGCGTGCAAGG3’

zfXRE5: 5’GGAGAGTGTGTGCGTGCGTTT3’

zfXRE6: 5’GGTGTAACACACGCTACGATAAA3’

zfXRE7: 5’GAACAGCATCGCGTGCGCGCA3’

zfXRE8: 5’GGGGCCACGTTGCAGCTGG3’
Duplex DNA was produced by mixing each oligonucleotide with its complementary strand, heating to 95°C and cooling to 25°C. Duplex DNAs were labeled with $[^{32}\text{P}]d\text{CTP}$ by Klenow fill in. *In vitro* binding assays and EMSA were carried out by combining approximately 25 ng of *in vitro* translated AHR and ARNT protein with 60 μL of MENG and incubated for 2 h at 30°C in the presence of TCDD (16 nM) or DMSO (0.5%). 14 μL of the activated sample was then incubated at 2°C for 15 min in 1X gel shift buffer supplemented with KCl (80 mM) and polydIdC (0.1mg/mL). 4 ng of the labeled zfXRE were added to each sample and incubated an additional 15 min at 22°C. The samples were resolved on 5% acrylamide/0.5% TBE gels, dried, and exposed to film.

*Western blot analysis and quantification of protein*

Protein samples were resolved by denaturing electrophoresis on discontinuous polyacrylamide slab gels (SDS-PAGE) and electrophoretically transferred to nitrocellulose. Immunochemical staining was carried out with varying concentrations of primary antibody in BLOTTO buffer supplemented with DL-histidine (20 mM) for 1–2 h at 22°C. Blots were washed with three changes of TTBS+ for a total of 45 min. The blot was then incubated in BLOTTO buffer containing a secondary antibody for 1 h at 22°C and washed in 3 changes of TTBS+ as above. Bands were visualized with the enhanced chemiluminescence (ECL) kit as specified by the manufacturer (Amersham, Arlington Hts, IL). Multiple exposures of each set of samples were produced.

*Statistical analysis*
Normalized RLU values were compared by ANOVA and Tukey-Kramer multiple comparison tests using InStat software (GraphPad Software Inc. San Diego, CA). Results are presented as mean ±SE. A probability value of ± 0.05 was considered significant.
References


Denis M, Cuthill, S., Wikstrom, A. C., Poellinger, L., Gustafsson, J. A. (1988) Association of the dioxin receptor with the Mr 90,000 heat shock protein: a
structural kinship with the glucocorticoid receptor. *Biochem Biophys Res Commun* 155:7.


Elferink CJ and Whitlock JP, Jr. (1990) 2,3,7,8-tetrachlorodibenzo-


Ma G, Kinneer K, Bi Y, Chan JY and W KY (2004) Induction of NAD(P)H:quinone oxidoreductase by 2,3,7,8-tetrachlorodibenzo-p-dioxin requires CNC (cap 'n' collar) basic leucine zipper factor Nrf2 (nuclear factor erythroid 2-related factor 2) cross-interaction between AhR (aryl hydrocarbon receptor) and Nrf2 signal transduction. *Biochem J* **377**:10.


Poland A, Glover E and Kende AS (1976) Stereospecific, high affinity binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin by hepatic cytosol. Evidence that the binding species


Appendix A

PAC 133 Sequencing Data

In order to obtain the sequence of the putative CYP1A regulatory region located within the fragments obtained from PAC 133 via restriction enzyme digestion, the 12kb SpeI fragment and the 2.5kb and 0.5kb HindIII fragments were subcloned into pBluescript ks- and sequenced. The initial sequencing was carried out using the T7 and T3 primers native to the vector. Subsequent primers were designed against the sequence retrieved from the initial sequencing (Figure AA1). Since the completion of this work, the zfCYP1A gene has been identified in the Sanger zebrafish genome project database. The zfCYP1A has been assigned to chromosome 18. The closest neighboring genes are the transient receptor potential cation channel, subfamily M, member 7 (trpm7), approximately 150 kb 5’-ward of CYP1A and fibroblast growth factor 7 (fgf7), approximately 50kb 3’-ward. The sequence obtained from the sequencing outlined above follows in this appendix. Directly sequenced nucleotides are capitalized. Sequence obtained in silico from other sources is represented by lowercase characters. Introns are italicized. Splice donor/acceptor sites are bold faced. The TATA box is underlined. SpeI and HindIII restriction enzyme sites are bold faced. XREs are bold and underlined. Numbering is relative to the transcriptional start site such that -1 is immediately 5’ of the transcriptional start site and +1 is immediately 3’.
Fig. AA1. Schematic overview of sequencing strategy. Visual representation of the sequencing strategy employed on the subcloned constructs containing the region upstream of the zfCYP1A. The 12kb SpeI fragment as well as the 2.5kb and 0.5kb HindIII fragments were subcloned into pBluescript ks- and sequenced using T7 and T3 primers. Subsequent primers were generated based upon the sequence retrieved from the initial sequencing. Primers and direction are indicated by arrows. The shaded region indicates the relative location of the HindIII fragments within the larger SpeI fragment. The positions of the 5’ untranslated region, the translational start codon, open reading frame, and splice donor/acceptor sites are labeled.
Appendix A continued

>Danio rerio Cytochrome P4501A 5’ Regulatory Region, 5’-UTR, and Intron 1

-3359  gttcacatacattgtagaatgcataatactctgtctatcacaacaaaaaagtcagctcaacagttgaaac
-3294  tggcagatttagactagcctagatagttctaccacaaatatagctgtgtttttctacatacatagactaaag
-3229  taactaanaatattgatgctagctgtatgtgtgtaatgtgctaatagtaaatctatatctgtgtaatgca
-3164  gatacactataatttgctatcctctcctctc tcagaaaaacactgaaacctaaaccataatttcatta
-3099  taaattagatgtacattataatatacatacatatactttgtacatgtgtataggctttacttattctatct
-3034  ctgagggaaaaatattttatatatgatatatattttagtatgtcaatatatttattatatatatatattat
-2969  tcaatgtgtaatatattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
> *Danio rerio* Cytochrome P4501A Intron 1 partial sequence and Exon 1 partial sequence
Appendix A continued

```
TTAATGACACGACAGGTGCGGAACACACACCTTCTCAGTATCGTTCATCTAAGAAGTGTTACT
GTAGACTCTCTGCTCCTGCGCATTCAAAATTTTCTTCTGTGCAATTATTTTGCTGAAGCTT
GTAACTTGATAT
```

Danio rerio Cytochrome P4501A Exon 1 partial sequence

```
AAGCTT
```

> Danio rerio Cytochrome P4501A Exon 1 partial sequence

![Diagram of ORF and T7/T3 primers with HindIII restriction site]

```
AAGCTTCGCCGCCGAGCCTCTGCGGATATCTGGAATAATCGGAAAAATCGGACACACCA
TTGAGTCTGAGCGCCATGATGAGTGTGCTATGCGCGGCTTTTTCGACTTACAGTATCGTC
CTGTTGTCGATCTCAAGGGAAATGTGATGTGATCCACGACACGTCCTCGGCTAAACAC
```
Appendix A continued

> *Danio rerio* Cytochrome P4501A Intron 2 partial sequence and Exon 2 partial sequence

TGGGGTGTTACGGTTTGTGTGAGCTTATTAGGAACATTCATTTCATTCTATTTCGGCTTAGTTGCTCTT
TTATTATATCGGGTGTTACGGTTTGTGTGAGCTTATTAGGAACATTCATTTCATTCTATTTCGGCTTAGT
TGCTCTTCTTTTTTTCTGCGTTACAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATG
CAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGA
GGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGC
GAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGA
GGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGC
GAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGA
GGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGC
GAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGA
GGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGC
GAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGA
GGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGC
GAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGA
GGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGC
GAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGAGGCGAGCTGGAATGGCGA

128
tccgacaaatgaaatgacaagtttagattaattcatcattaaaaacaataacaaaaaagctaatttaaattaaataaacaatagaaactttatatttttaacagcccaagttttttgtttaaaatatatttttttttttaaatggactattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
GAATGTGATCCACATGTAATACGCAGCGGCTGGTTTCAATATGCAACATTTTTTTTCCATGTGCC
TGATTTTTACTTTGTGTACAGATTTGTAATATCTACTTTTTTGTATTTATCAAAGTACTAACTCAGAT
TTTTATTTCCCTTTGAGATATCGGTGTCTGTGACTCGGAAACTCTTCGTCAGTGCTCATAATTGTA
TTATGAAAGGAATCTCTGAGATCTACTCTGCTATAATTAAATTCATGCTAACATGGTCTCAGATG
TTGTGTGGCATAGCATATTGATGTTGGCAAAAATAAATAAATAAACTTCTACTGCTG
Appendix B

Construct Maps

This appendix contains maps of constructs used in the experiments outlined in this report. The names and size of the constructs in base pairs are indicated. Pertinent restriction enzyme sites are noted along with their numerical positions. Bold faced numbers represent the position of the indicated DNA insert relative to the transcriptional start site. The diagrams below the maps represent the portion(s) of the regulatory region within the construct. XREs are shown as shaded rectangles. The XRE designation and its position relative to the transcriptional start site are indicated.
Appendix B continued

Figure AB1. p-2608/-2100Fup
Figure AB2. p-2608/-2100Rup

p-2628/-2100Rup
5518 bp
Appendix B continued

Figure AB3.  p-2608/-2100Fdown
Figure AB4. p-2608/-2100Rdown

p-2608/-2100Rdown
5518 bp
Figure AB5. p-580/-187Fup

p--580/-187Fup

5403 bp

KpnI (5) -580

ScaI (404) -187

MluI - 408
NheI - 414
SmaI - 421
XhoII - 425
BglII - 429

HindIII (638)

Ncol (671)

SV40

Luciferase

Amplifiable

sV40 pA

Sall (2595)

BamHI (2589)

XbaI (2327)
Figure AB6. p-580/-187Fdown

p-580/-187Fdown
5403 bp

Appendix B continued

137
Figure AB7. P-580/+71Basic

p-580/+71Basic
5469 bp
Appendix B continued

Figure AB8. p-2608/-2100Uf

5977 bp

-2608
-2100
-580
+71

KpnI (5)
SacI (513)
MluI (517)
NheI (523)
HindIII (1174)

-2608/-2100Uf

p-2608/-2100Uf

5977 bp

KpnI
SacI
NheI
HindIII

Amphoterin

Ampicillin

SV40 A

Luciferase

BamHI (3163)

BamHI (2901)

SacI (1061)

SacI (3169)

NcoI (1245)

-2608
-2100
-580
+71

KpnI
SacI
NheI
HindIII

139
Appendix B continued

Figure AB9. p-2608/-2100Ur

p-2608/-2100Uf
5977 bp

Argcillin

Syn. PA

SV40 PA

Luciferase

KpnI (5)-2100

Sacl (513)-2608

MluI (517)

Nhel (523)-580

Sacl (1061)

HindIII (1174)+71

Ncol (1245)

Sacl (3169)

BamHI (3163)

XbaI (2901)

KpnI

SacI

NheI

HindIII

580

-2608

-2100

-2100 -2608-580 +71
Appendix B continued

**Figure AB10. p-Om-1897/-1392Uf**

![Diagram of p-Om-1897/-1392Uf plasmid](image-url)
Appendix B continued

Figure AB11. p-Om-1897/-1392Ur

5974 bp
Appendix B continued

Figure AB12. p-Mm—1315/-819Uf
Appendix B continued

Figure AB13. p-Mm-1315/-819Ur

![Diagram of p-Mm-1315/-819Ur showing restriction enzyme sites and annotations.](image-url)

**p-Mm-1315/-819Ur**

5965 bp

**Key Enzymes and Sites:**
- **KpnI (5)** -819
- **SacI (510)** -1315
- **MluI (511)**
- **NheI (517)** -580
- **HindIII (1168)** +71

**Other Markers:**
- **Sall (3163)**
- **BamHI (3160)**
- **KpnI**
- **SacI**
- **NheI**
- **HindIII**

**Annotations:**
- **Ampicillin**
- **Luciferase**
- **SV40 pA**
- **Xbal (2898)**
- **XRE1**
- **XRE2**
- **XRE3**
- **XREC**
- **XREE**
- **XREF**
- **XREF2**
- **XREF3**
- **XREF4**

**Restriction Sites:**
- **KpnI**
- **SacI**
- **NheI**
- **HindIII**

**DNA Segments:**
- **-819**
- **-1315-580**
- **-42 +71**
Figure AB14. p–2699/+71
Figure AB15. p—2608/+71

Appendix B continued
Appendix C

Additional Studies

**Generation of destabilized EGFP and EYFP constructs**

In order to gain a better understanding of the timecourse of AHR-induced genes, constructs were generated which contain a gene encoding a destabilized enhanced green fluorescent protein (dEGFP) under control of the zfCYP1A regulatory region. The purpose of destabilizing the EYFP was to ensure that the fluorescent signal produced is not persistent and can be representative of transcriptional activity over time without protein accumulation from earlier transcriptional activity. To produce a destabilized EGFP, the degradation domain of the mouse ornithine decarboxylase gene was fused to the EGFP coding sequence. Additionally, a 3x nuclear localization signal was fused to the amino terminus of EGFP to generate a nuclear fluorescent signal which would be easier to view *via* fluorescent microscopy. PCR was utilized to make the constructs such that primers were synthesized which amplified both the 3xNLS and the dEGFP with overhangs respectively complementary to each other. An additional PCR cycle was run using both products as template, the forward NLS primer, and the reverse dEGFP primer. A non-destabilized EGFP was also generated by using a reverse EYFP primer which precedes the ornithine decarboxylase degradation domain and includes a stop codon. The
Appendix C continued

3xNLS template was purchased as a double stranded oligonucleotide. The pd2EGFP vector (Clontech) was used as a template for the dEGFP.

The primers used are listed below. Indicated restriction enzyme sites are underlined.

3xNLS double stranded oligonucleotide:

5′ - CCGAGATCCGGGATCCACCTTTTCTTTTTTTTGGGTCCACCCTTTTTTTTTTTT
3′ - GGCTCTAGGCCACCTAGGGTGAGAAAAGAAAAACCCCAGGTGGGGAAAAA

GGGTCAACCCGCTTCTTCTTAGGATCCATGGTGGGATATCTGAC -3′
CCCAGTTGGGCGAAGAAGAATCCTAGGTACCACCCTATAGACTG -5′

P1 FOR 3XNLS EcoRV – 5′ - ATCGTCAGATATCCACCATGGATCC

P2 REV 3XNLS -5′ - CAGCTCCTCGGCCCTTGCTCACCATCCGAGATCCGGTGGATCCCAGCTCC

P3 FOR GFP – 5′ - AGAGTGGGATCCACCGGATCTCGGATGGTGAGCAAGGGCGAGGAGCTG

P4 REV GFP XbaI – 5′ - CCTCCATCTAGACTACATGTGAGAGCTC

After creating the recombinant GFP inserts they were ligated into the multiple cloning site of pcDNA 3.1 (Clontech) and termed p-EGFP and p-dEGFP. Transfection of the constructs into Hepa1c1c7 cells, however failed to produce fluorescence. The reason for the failure of the constructs to work is unknown. To allow for the problem lying within the GFP, additional constructs were made containing the coding sequence for an enhanced yellow fluorescent protein (EYFP) known to work previously. These constructs were assembled modularly such that PCR was utilized to amplify the ODC and NLS from p-dEGFP with primers containing restriction enzyme sites integrated into them. EYFP was amplified from p-EYFP-C1 (Clontech) using primers with
Fig. AC1. Strategy for Generating EYFP Constructs. Schematic representation of the EYFP inserts used to generate EYFP constructs. EYFP = Enhanced yellow fluorescent protein. NLS = 3x nuclear localization signal. ODC = ornithine decarboxylase degradation domain. Restriction sites and the PCR primers within which they are integrated are indicated.
corresponding restriction enzyme sites. To ensure the lack of function exhibited by p-dEGFP was not due to the ornithine decarboxylase cassette or the nuclear localization sequence, the new constructs were assembled containing only the EYFP cassette, (p-EYFP), the EYFP cassette with the ODC degradation domain, (p-dEYFP), the EYFP cassette with the 3xNLS, (p-NLSEYFP), and the EYFP with both the ODC and NLS, (p-NLSdEYFP). A schematic overview of the modular assembly of the EYFP constructs is shown in Figure AC1. The primers used for the creation of the EYFP constructs follow.

3xNLS double stranded oligonucleotide:

3xNLS - 5’- CTAGCCCCACCATG CCA CCT AAG AAA AAA AGA AAG GTT
3’- GGGTGGGTAC GGT GGA TTC TTT TTT TCT TTC CAA

GAA GAT CCTGGTAC-3’
CTT CTA GGAC-5’

P1 FOR 3xNLS NheI – 5’- ATCGTCAGCTAGCCACCATGGATCC
EYFP IN NheI – 5’- TTTTGCTAGCCCACCATGGTGAGCAAGGGC
P4 REV EYFP HindIII – 5’- TTTTAAGCTTTCACTTGTACAGCTCGT
EYFP OUT FOR ODC – 5’- TTTTAAGCTTTGTAAGCTCGTC
ODC IN HindIII – 5’- TTTTAAGCTTAGCCATGGCTTC
ODC OUT BamHI – 5’- TTTTGATCCCTACACATATGCCT

The appropriate modules were ligated into pcDNA 3.1 and the constructs were once again transfected into Hepa1c1c7 cells. Fluorescence microscopy revealed that a
Fig. AC2. Construct maps of p-EYFP and p-dEYFP. Maps of the p-EYFP and p-dEYFP plasmids. The size of the constructs in base pairs is shown. Pertinent restriction enzyme sites are indicated. EYFP = enhanced yellow fluorescent protein. NLS = 3x nuclear localization signal. ODC = ornithine decarboxylase degradation domain.
fluorescent signal was produced by all constructs. The cells transfected with p-NLSEYFP and p-NLSdEYFP however, did not appear to localize to the nucleus and was dispersed throughout the cells. Treatment of transfected cells with cycloheximide additionally did not show a difference in the rate of degradation between p-EYFP and p-dEYFP as expected.

To test whether or not the EYFP could be controlled by an AHR dependant promoter, the recombinant EYFP inserts were cloned downstream of the zebrafish CYP1A enhancer region (-2628/-2100) and the zfCYP1A promoter (-580/+71). Maps of these constructs can be seen in Figure AC2. These constructs, termed p-2628/-2100UrEYFP, p-2628/-2100UrdEYFP, p-2628/-2100UrNLSEYFP, and p-2628/-2100UrNLSdEYFP were then transfected into Hepa1c1c7 which were treated with vehicle dimethylsulfoxide or TCDD. The cells produced a fluorescent signal, however the signal was produced even in the absence of TCDD to the same levels as TCDD treated cells. The reason for the constitutive activity is unknown but may be due to moderate basal levels exhibited by luciferase constructs bearing the same regulatory region. As seen in previous experiments with p-EYFP, p-dEYFP, and p-NLSEYFP, the signal did not localize to the nucleus or appear to degrade at an accelerated rate when treated with cycloheximide. Sequencing of the constructs revealed that the inserts had no aberrations. Further work will have to be done to design constructs capable of being used to address the timecourse of AHR-induced genes.
Appendix C continued

**ARNT Domain Swapping**

To assess the role the bHLH, PAS, and TAD domains of ARNT and ARNT2 have on their function, the coding regions of the human ARNT and mouse ARNT2 genes were analyzed and a strategy to insert restriction enzyme sites between domains was devised. PCR was utilized to alter nucleotides within the coding region to create unique restriction sites which did not alter the amino acids being encoded. The nucleotide and corresponding amino acid sequences of the human and mouse ARNT and ARNT2 are listed below. Generated restriction enzyme sites are bold faced. Mutated nucleotides are capitalized and underlined. A schematic overview of the domains and the restriction enzyme sites can be seen in Figure AC3.

>`C57/BL6 Mouse ARNT

```
 IST M A A T T A N P E W T S D V P S L
 atctgaccatgctggcgactacacgtaaaccgaaatgacatcagatgtacctgcgt
 62
 G P T I A S N G P G P G I Q G G G A V V
 63
gtccccacattctcttgaaaccccttgaccttggttacagtggagagctgttgta
 122
 Q R A I K R R S G L D F D D D E V E N T
 123
cagaggtcatataagcagctcgcgcttgatttttgatgatgagattgatgaaagat
 182
 N M R L C D D D Q M C N D K E R F A R S
 183
 aatatttttgatgtggatctgactactggaatgactggttggctggtcaggtgc
 242
 D D E Q S S A D K E R L A R E N H S E I
 243
 gatgatgagcagctcgcgataaaagagagacagctcgcagaaatcatagtgaata
 302
 E R R R N K M T A Y I T E L S D M V P
 303
 gaacctgacagtggacagggaagagcatgtaaccacagtacagctcagacatgtacct
 362
 T C S A L A R K P D R L T I L R M A V S
 363
 acatgtgaagctttgttgtaccttacagacagctgacgctgtctctctctctctct
 422
 H N K S L R G T G N T S T D G S Y K P S
 423
 cacatgagctctgtggaggggaactggcaacccatctactgtgcctcataacaggtcca
 482
 F L T D Q E L K H L L I E A A D G F L F
 483
 ttcctcactgactgacagggaacatgataactactacacgacgctgacatgtacctgcgt
 542
 I V S C T G R V V V D S V T P V L
 543
 attgtcactctttggacagggtgtgcttcttctctactagctcgtgagctggtctgtttg
 602
 N Q P Q S E W P G ST L Y D Q V H P D D
 603
 aacccagccacagtctgtgttgccggaacagtacagctgtgtacaggtgccagagat
 662
 V D K L R E Q L S T S E N A L T G R V L
 663
 gttgataaaaaacttccagagcaactcctatcatacgcagaaatgccctaaacagggcgggtctctg
 722
 D L K T G V K K E G Q Q S S M R M C M
 723
 gatgtcaagactggaaacgaatgaaaagagggcagctatccagctagttgtgatgtgtctctagctcgtggtccagt
 782
 G S R R S F I R N R C P R N G L G S V K
 783
 ggctccacagactgctctctctgcgcccaagagggtgtggtactactctctagctcgtcgacccctgt
 842
 S M N R L S F I R N R C P R N G L G S V K
 843
ttcagatagactagctgacggcttggagaagactcagagatcagacagcttgctctgctgag
 902
 E G E P H F V V V H C T G Y I K A W P P
```
Appendix C continued

903    gaaggagaacccctatcttttggtgtacactccgtcagctcacttggtgcccatcacacagccatcagctgcac
962    acC    ggT    tacatcaaggcctggccacca
963    V  A  I  G  R  L  Q  V  T  S  S  P  N  C  T  D  M  S  N  I
1022   gcaggtgtctccctcccagatgatgacccagaggctggccaggggagcaaattctgccta
1023   V  A  I  G  R  L  Q  V  T  S  S  P  N  C  T  D  M  S  N  I
1082   gtggccattggcaggctgcaggtaactagttctcccaactgtacagacatgagtaacatt
1083   C  Q  P  T  E  F  I  S  R  H  N  I  E  G  I  F  T  F  V  D
1142   tgtcagccaacagagttcatctcccgacacaacattgaagggatattcacttttgtagac
1143   H  R  C  V  A  T  V  G  Y  Q  P  Q  E  L  L  G  K  N  I  V
1202   gaattttgtcatcctgaagaccaacaacttctaagagacagctttcagcaggtggtgaaa
1203   L  K  G  Q  V  L  S  V  M  F  R  S  K  T  R  E  W  L
1262   ttaaaaggtcaggctgtccgtcgctgattcctggatctagctcaagaccgagagatggctg
1263   W  M  R  S  F  T  Q  N  P  Y  S  D  E  I  E  Y  I
1322   gatgggctggccagctataatcattcccaggtttctgtccagcctgtggcaagtgcagga
1323   S  E  H  S  K  P  L  E  K  S  E  G  L  F  A  Q  D  R  D  P
1382   tggatgagaacgagctcctttaccttccaaaacccttattcagatgaaattgagtatatt
1383   I  C  T  N  T  N  V  K  N  S  S  Q  E  P  R  P  T  L  S  N
1442   atctgccacaccaacacatgtgaacacctccagcagcctacgtcctccacac
1443   T  I  P  R  S  Q  L  G  P  T  A  N  L  S  L  E  M  G  T
1502   cacgtgcaccatacTagAcaacagcagcagcagcagcacagagaactggatatggtaccaggaaga
1503   D  G  L  A  S  Y  N  H  S  Q  V  S  V  Q  V  S  A  G
1562   gatgggctggccagctataatcattcccaggtttctgtccagcctgtggcaagtgcagga
1622   S  E  H  S  K  P  L  E  K  S  E  G  L  F  A  Q  D  R  D  P
1682   tggatgagaacgagctcctttaccttccaaaacccttattcagatgaaattgagtatatt
1742   I  C  T  N  T  N  V  K  N  S  S  Q  E  P  R  P  T  L  S  N
1802   gatgggctggccagctataatcattcccaggtttctgtccagcctgtggcaagtgcagga
1862   S  E  H  S  K  P  L  E  K  S  E  G  L  F  A  Q  D  R  D  P
1922   tggatgagaacgagctcctttaccttccaaaacccttattcagatgaaattgagtatatt
1982   I  C  T  N  T  N  V  K  N  S  S  Q  E  P  R  P  T  L  S  N
2042   tggatgagaacgagctcctttaccttccaaaacccttattcagatgaaattgagtatatt
2102   I  C  T  N  T  N  V  K  N  S  S  Q  E  P  R  P  T  L  S  N
2162   gatgggctggccagctataatcattcccaggtttctgtccagcctgtggcaagtgcagga
2222   S  E  H  S  K  P  L  E  K  S  E  G  L  F  A  Q  D  R  D  P
2282   tggatgagaacgagctcctttaccttccaaaacccttattcagatgaaattgagtatatt
2342   I  C  T  N  T  N  V  K  N  S  S  Q  E  P  R  P  T  L  S  N
2402   gatgggctggccagctataatcattcccaggtttctgtccagcctgtggcaagtgcagga
2462   S  E  H  S  K  P  L  E  K  S  E  G  L  F  A  Q  D  R  D  P
2522   gatgggctggccagctataatcattcccaggtttctgtccagcctgtggcaagtgcagga
2582   S  E  H  S  K  P  L  E  K  S  E  G  L  F  A  Q  D  R  D  P
2622
## Appendix C continued

### C57/BL6 Mouse ARNT2

<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>PA L R S P L G K M A T P A A V N P P E</td>
</tr>
<tr>
<td>62</td>
<td>ccggcgcctcgctctcccgcggcaagatggcaaccccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>122</td>
<td>Q V R M A G A M P A R G K K R S G M D</td>
</tr>
<tr>
<td>182</td>
<td>F D D E D G P S K F S R E N H S E I</td>
</tr>
<tr>
<td>242</td>
<td>T T C A L R K P D K L T I L R M A V S</td>
</tr>
<tr>
<td>302</td>
<td>acctgcaagtgcctcgtccttcgcggcagatggcaaccccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>362</td>
<td>H M K S M R G T N K S T D G A Y K P S</td>
</tr>
<tr>
<td>422</td>
<td>ttctcgcctcgctctcccgcggcaagatggcaaccccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>482</td>
<td>V Q P S E W F G S T L Y E Q V H P D D</td>
</tr>
<tr>
<td>542</td>
<td>aaccacgcacactcagagtttgggagacgtttatagcagagttgacccctccggag</td>
</tr>
<tr>
<td>602</td>
<td>V E K L R E Q L C T S E N S I T G R L</td>
</tr>
<tr>
<td>662</td>
<td>gtgcgagactgcaaggttcctgtgagggacgtttatagcagagttgacccctccggag</td>
</tr>
<tr>
<td>722</td>
<td>gprsrsficrmrcgcnaplldhl</td>
</tr>
<tr>
<td>782</td>
<td>ggctctcgccctctctcgctctgtgaggttgatgggaagctcctcttgacccctccggag</td>
</tr>
<tr>
<td>842</td>
<td>p l n r i t t m k r k r f r n g l p v k</td>
</tr>
<tr>
<td>902</td>
<td>gaagagagagagcctgcctctgtgaggttgatgggaagctcctcttgacccctccggag</td>
</tr>
<tr>
<td>962</td>
<td>gcgaagagacctctcctgcctctgtgaggttgatgggaagctcctcttgacccctccggag</td>
</tr>
<tr>
<td>1022</td>
<td>gcttcgcctcgctctcccgcggcaagatggcaaccccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>1082</td>
<td>prcisvigyqppoaldllkgkdril</td>
</tr>
<tr>
<td>1142</td>
<td>cccagatgcactagtgtgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>1202</td>
<td>gcgaagagacctctcctgcctctgtgaggttgatgggaagctcctcttgacccctccggag</td>
</tr>
<tr>
<td>1262</td>
<td>gatgggctgtcgtcatatgacttatctcaggtcccagtacccaacctacccgct</td>
</tr>
<tr>
<td>1322</td>
<td>ggtgttcacgaggccgggaagtctgtgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>1382</td>
<td>ecfhpedqshlresfqqvvk</td>
</tr>
<tr>
<td>1442</td>
<td>ggttctcagagggcagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>1502</td>
<td>ggaattttgccaccctccgagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>1562</td>
<td>gccttcggaattggatcaagccacccttcctgcctctgtgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>1622</td>
<td>gcatctcgggcaaggccttcagctcttccgtggtccatgtccagtgagtgaatgacattcag</td>
</tr>
<tr>
<td>1682</td>
<td>cactcggggaaggccttcagctcttccgtggtccatgtccagtgagtgaatgacattcag</td>
</tr>
<tr>
<td>1742</td>
<td>tctgtgcttcagagggcagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>1802</td>
<td>ggttctcagagggcagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>1862</td>
<td>ggaattttgccaccctccgagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>1922</td>
<td>gatgggctgtcgtcatatgacttatctcaggtcccagtacccaacctacccgct</td>
</tr>
<tr>
<td>1982</td>
<td>ggtgttcacgaggccgggaagtctgtgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2042</td>
<td>ggaattttgccaccctccgagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2102</td>
<td>gatgggctgtcgtcatatgacttatctcaggtcccagtacccaacctacccgct</td>
</tr>
<tr>
<td>2162</td>
<td>ggtgttcacgaggccgggaagtctgtgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2222</td>
<td>ggaattttgccaccctccgagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2282</td>
<td>gatgggctgtcgtcatatgacttatctcaggtcccagtacccaacctacccgct</td>
</tr>
<tr>
<td>2342</td>
<td>ggtgttcacgaggccgggaagtctgtgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2402</td>
<td>ggaattttgccaccctccgagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2462</td>
<td>gatgggctgtcgtcatatgacttatctcaggtcccagtacccaacctacccgct</td>
</tr>
<tr>
<td>2522</td>
<td>ggtgttcacgaggccgggaagtctgtgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2582</td>
<td>ggaattttgccaccctccgagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2642</td>
<td>gatgggctgtcgtcatatgacttatctcaggtcccagtacccaacctacccgct</td>
</tr>
<tr>
<td>2702</td>
<td>ggtgttcacgaggccgggaagtctgtgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2762</td>
<td>ggaattttgccaccctccgagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2822</td>
<td>gatgggctgtcgtcatatgacttatctcaggtcccagtacccaacctacccgct</td>
</tr>
<tr>
<td>2882</td>
<td>ggtgttcacgaggccgggaagtctgtgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>2942</td>
<td>ggaattttgccaccctccgagagttgaggtggtacctgacccctccggccgccgtcaaccctccggag</td>
</tr>
<tr>
<td>3002</td>
<td>gatgggctgtcgtcatatgacttatctcaggtcccagtacccaacctacccgct</td>
</tr>
</tbody>
</table>

...
Appendix C continued

R  H  G  S  D  L  P  H  K  Q  E  E  A  S  D  R  N  S  R
2282  cgtcacggtttctgacctaccaacataaaacaggaagaagaccagtgacccggaacagctctagg
N  N  K  S  E  *  K  C  P  L  Y  Y  Q  K  I  W  A  W  P  K
2342  aataacaagtcaatagataagttgtgttttatattaccgaaataatatgggcttggtcctaag
S  L  S  P  N  L  P  G  S  F  P  T  K  H  P  I  L  R  S  H
2402  tcgtctgtctcttaacctgcccgggtctatctcccacacaaagccccatataagagaacat
E  P  P  G  H  S  P  F  L  *  P  S  G  V  W  G  N  L  R  R
2462  gagccacccgctgacattccacatccctcctcttgacattcttgagcttggtggcaacattaaaggg
H  T  V  V  Q  A  H  F  Q  A  *  V  S  L  A  F  V  A  K
2522  caccacaacaggtgtgagcagcacacattccacacagctaggtgtcctctgctttttgtgctcct
A  S  V  M  V  N  N  R  P  G  S  V  G  H  *  P  *  K  W  Q
2582  gctagtgttatgctcaacaacagccaggtgtctgtgtggctgactgactgatgctgatgctgctg
N  G  G  F  T  G  C  A  G  A  G  R  L  A  S  S  N  N  L  S
2642  aatggaggttccacaggtctgccggagcaggcagctggctctccatcataaaatctcagt
F  L  *  K  K  K  E  R  K  R  F  H  K  Q  V  S  V  D  S  L
2702  ttccttttaaaaagaaaaagaaaaagatattctataagcaggtgtcagtgacagttta
S  T  *  P  F  L  F  L  L  M  D  V  N  C  A  V  D  K  S  F
2762  agtaccttaacctttctttctttctttctttatggtgatgtaactgtctgtgattaaatatttt
V  F  L  E  C  S  L  *  L  T  V  I  K  S  V  V  Y  M  C  N
2822  gtattttcttgagttttcttctctttactaaacagatttaaagctgttctgtgatgtgtaac
*  C  N  C  L  L  K  F  H  Y  N  K  N  D  F  A  L  K  K  K
2882  taatgtaacgctccttttaaaatctctactacacatatnagatctt tgctctgaaaaaaa
K  K
2942  aaaaaaa  2948
Figure AC3. Schematic of ARNT Domain Swapping

**Fig. AC3. Schematic of ARNT domain swapping.**  Schematic representation of the restriction enzyme sites generated within ARNT and ARNT2 to create interchangeable domains. The basic,helix-loop-helix, PAS A, PAS B, and transactivation domains are shaded. Numbers in parentheses represent the nucleotide position of the indicated restriction enzyme site in the context of the cDNA.
Figure AC4. Construct Maps of p-hARNT-ID and p-mARNT2-ID

Maps of the p-hARNT-ID and p-mARNT2-ID plasmids. The size of the constructs in base pairs is shown. Pertinent restriction enzyme sites along with their positions are indicated. bHLH = basic helix-loop-helix domain. PAS A and PAS B = Per ARNT Sim domains A and B. TAD = transactivation domain.
Figure AC5. Restriction Enzyme Digestion of p-mARNT2-ID

![Image of agarose gel with fragments labeled](image)

**Figure AC5. Restriction enzyme digestion of p-mARNT2-ID.** 1% agarose gel containing ethidium bromide was loaded with p-mARNT2 digested with the indicated restriction enzymes and run at 100V for one hour. Left lane contains lambda HindIII/EcoRI marker. Fragment sizes are shown. The labels at the bottom indicate the domain removed by the restriction digestion.
Appendix C continued

After generating the domains with the specified mutations *via* PCR, the fragments were digested with the appropriate enzymes and subcloned sequentially into pcDNA 3.1 (Clontech). The resultant constructs, named p-hARNT-ID and p-mARNT2-ID (Figure AC4) were then digested with restriction enzymes and visualized in an agarose gel to ensure the proper sized fragments were produced. A representative photo of the agarose gel can be observed in Figure AC5. Next the constructs were used to produce *in vitro* transcribed/translated proteins, however the products were ~200kD in size, far smaller than the expected size. Sequence data did not reveal any early stop codons. The reason for the improperly sized proteins is unclear and will require further investigation.
About the Author

Gary T. ZeRuth was born February 23rd, 1976 in Phillipsburg, NJ. He spent the early part of his childhood near Easton, PA from which he moved in 1985 to the Tampa Bay area of Florida. Gary received his Bachelor’s Degree in Biology from the University of South Florida in Tampa, FL and went on to pursue his Ph. D. from the same institution. In addition to the sciences, Gary has an avid love of history and has performed in a number of historical festivals, plays, and dinner theaters. He also has great interests in languages, classic literature, and the visual arts. Gary married an artist, Melissa, in 1999. Hoping to eventually establish his own laboratory and work in academia, Gary has accepted a post doctoral research position at the National Institutes of Health.