# INVESTIGATING THE RESPONSE TO INDOLE DERIVATIVES THROUGH THE ARYL HYDROCARBON RECEPTOR AND OTHER SIGNALING PATHWAYS

An Undergraduate Research Scholars Thesis

by

ERIC FRIS

Submitted to the Undergraduate Research Scholars program
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by
Research Advisor: Dr. Arul Jayaraman

May 2016

Major: Chemical Engineering

# TABLE OF CONTENTS

# ABSTRACT

Investigating the Response of Indole Derivatives through the Aryl Hydrocarbon Receptor and other Signaling Pathways

Eric Fris
Department of Chemical Engineering
Texas A&M University

Research Advisor: Dr. Arul Jayaraman
Department of Chemical Engineering

Gastrointestinal tract bacteria can transform molecules into metabolites that can either be utilized by the mammalian host or can induce signaling in the host. One such metabolite, indole (and its derivatives), is generated from the amino acid tryptophan. Indole is small and can passively diffuse into the host cell. Indole and its derivatives such as indole-3-acetate (I3A) have been shown to activate the the aryl hydrocarbon receptor (Ahr), which, in turn, activates a range of signaling pathways. However, preliminary data from our lab suggest that indole derivatives also bind to extracellular cell surface receptors. This work proposes to investigate the balance between Ahr-mediated signaling and Ahr-independent signaling by indole derivatives. This project also seeks to develop a mathematical model to describe this mechanism, as well as test the general applicability of this response in different cell lines.

# DEDICATION

I would like to dedicate this thesis to the most influential professor and mentor in my undergraduate experience, Dr. Arul Jayaraman. Without his passion for his research and his desire to share in his passion with his students, I would not be nearly as prepared and excited for my journey to medical school.

# ACKNOWLEDGMENTS

I would like to thank the entire Jayaraman Lab Group for their help and support throughout this process. Their teaching and mentoring have proven invaluable. The papers published and the information provided to me by collaborators, including Dr. Safe and Dr. Alaniz, have also proven equally invaluable. Finally, I would like to acknowledge the faculty and staff associated with Undergraduate Research Scholars and those all those who work in the LAUNCH office to make this program an amazing experience.

# NOMENCLATURE

| | |
|---|---|
| Ahr | Aryl hydrocarbon receptor |
| TCDD | 2,3,7,8-tetrocholorodibenzo-p-dioxin, an environmental toxin |
| NF-$\kappa$B | a primary transcription factor for inflammatory responses; many upstream pathways eventually lead to it |
| Indole | An aromatic heterocycle composed of a benzene ring fused to an aromatic five-membered ring ($C_8H_7N$) |
| CYP1A1 | a gene encoding a protein in the Cytochrome P450 family that is known to be transcribed when a cell receives signals indicating toxins or inflammation |

# CHAPTER I

# INTRODUCTION

The interaction between mammalian host cells and the resident microbes in the GI tract is not well understood. Many metabolites used by the host cells are actually produced by bacterial symbionts [5]. For example, indole-3-acetate (I3A) is a metabolite that is derived from the amino acid tryptophan (Trp) by the microbiota and may have important antioxidant and anti-carcinogenic effects [1]. I3A is also known to activate the aryl hydrocarbon receptor (Ahr) in mammalian colon cells [2]. The Ahr receptor has broad substrate specificity and is activated by several ligands and has been shown to play an important role in inflammation and immune responses. Studies have proposed that the Ahr is an important target for anti-cancer treatment strategies [2].

**Examination of the compounds in the Ahr Pathway**

The above example underscores the importance of studying the role of bacterial metabolites in the human body. If metabolites largely produced and regulated by bacteria can regulate, modulate, and control the host response to inflammation and disease, it can lead to the development of new drugs and treatment strategies. The long-term goal of this work is to develop anti-inflammatory molecules based on GI tract microbiota metabolism. The specific goal of the proposed work is to investigate and model the signaling pathways that are activated by I3A .

*Accounting for diffusion*

Since I3A is a small planar molecule, it is likely to diffuse into the host cell. Recent work from our lab has shown that I3A activates the Ahr. Recent data also indicate that I3A conjugated to a large protein (bovine serum albumin; BSA) that cannot diffuse into the cell also induces in cellular responses, which suggest that this molecule is acting through an extracellular receptor to activate signaling and gene expression.

*Bovine Serum Albumen (BSA) and indole*

Bovine Serum Albumen (BSA) is a very well-studied model protein in biochemistry and biology labs. Many scientists consider it and GFP (Green Fluorescent Protein) to be two of the most well-documented proteins. At roughly 580 amino acids and 66 kDa, this protein serves as a large stabilizing protein in cell cultures and media. Due to its large size and stability, it will be an effective "inert" protein to which indole can be conjugated [6].

*Stoichiometry*

Recent data in our lab suggests that as long as BSA is in low concentration compared to the total protein content present in the cell, it will not affect the signaling pathway of interest. Recent work in our lab also suggests that the stoichiometry of the synthesis of conjugated indole and BSA (I-BSA) is not 1:1 (i.e. more than one I3A molecule may attach to a single BSA). While this does not affect the general trends in our study (the relative expression of indole compounds and associated controls), this will become important later as we seek to accurately quantify and develop a model for I3A induced response.

**Experimental objectives and controls**

This information leads to several interesting questions. If I3A can produce similar changes in gene expression through a pathway other than the Ahr, what is this pathway and what is the receptor that it binds? Is this pathway independent of the Ahr pathway or is it interconnected with Ahr signaling? Where does the equilibrium lie in this model, and what causes it to favor a particular direction?

In order to properly investigate these questions, a background on experimental control must be established, including positive controls, known Ahr agonists and antagonists, and proper solvents. According to Cheng et al., TCDD (2,3,7,8-tetrachlorodibenzo-*p*-dioxin) is a well-studied agonist of Ahr that reliable induces a variety of responses, namely CYP1A1 [2]. It can be used as a reliable baseline positive control in CaCo2 cells because it induces a response that can be easily measured – CYP1A1 expression.

The following is a diagram of the pathway discussed in the previous paragraph. *Figure 1* illustrates a very simplified diagram of what is known about the Ahr-CYP1A1 mechanism.

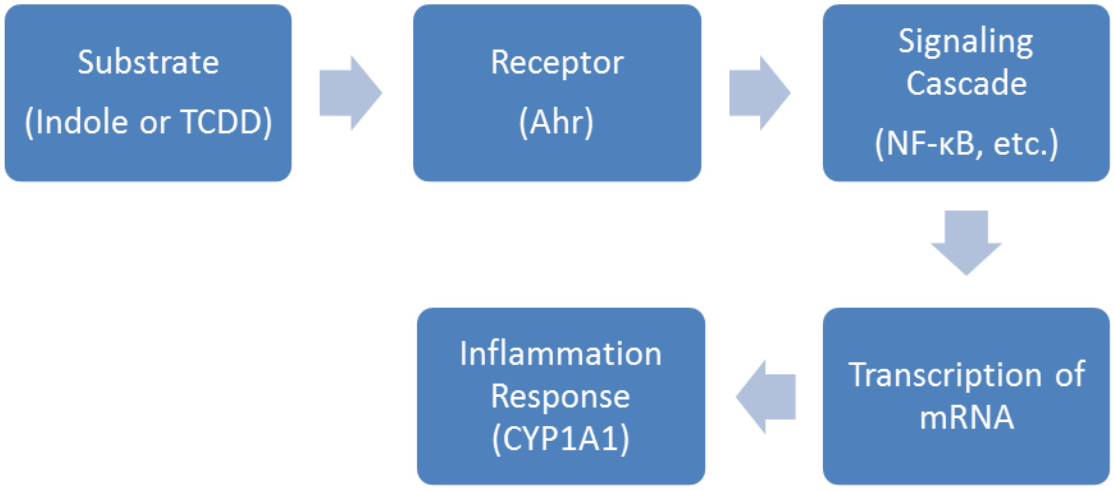


*Figure 1: An illustrative representation of the Ahr-CYP1A1 Pathway*

A proper solvent for these experiments is usually either DMF (dimethylformamide) or DMSO (dimethylsulfoxide). DMF will be used over DMSO to keep consistency with prior literature, in case there are solvent effects present.

# CHAPTER II

# METHODS

## Conjugation of indole

Indole (including I3A and derivatives) can passively diffuse into host cells and activate the Ahr. Activation of the Ahr leads to the expression of several target genes, including the cytochrome P450 1A1 gene (CYP1A1). To eliminate activation of CYP1A1 gene expression due to I3A diffusing and activating the Ahr, I3A will be conjugated to BSA, (I-BSA). Since BSA is sufficiently large (~67 KDa), I-BSA cannot diffuse into the cell and uptake, and therefore any interaction with the Ahr must require active transport or signaling by binding to an extracellular receptor.

## Cell lines

Young adult mouse colonocyte (YAMC) cells and the CaCo2 human epithelial cell line will be used as the host cell lines in these experiments, based on ongoing research and prior experience in this lab. BMDM (bone marrow derived macrophages) may be used at a later time to generalize the results to more cell types.

CaCo2 and YAMC cells will be cultured in standard corning plates with media commonly used for the cell line. For experiments cells will be plated on 24 or 96 well plates, where 4 different combinations will be tested: a positive control (TCDD which is well established to induce CYP1A1 gene expression), a negative control (no treatment), I3A, and I-BSA. The extent of CYP1A1 gene expression will be determined using RTq-PCR (Real-Time Quantitative Reverse

Transcription PCR), which determines the relative expression of CYP1A1 relative to a housekeeping control gene (HRPT) [3]. The experiments will be carried out for different time points so that the uptake of these molecules and the activation of target genes can be determined.

**A note about lab safety**

The lab protocol requires Biosafety Level 2 (BL2) and Blood Borne Pathogen (BBP) training. I and all researchers in this lab have completed such training. Furthermore, this protocol is based upon similar research in the same lab, and the safety and ethics concerning it are therefore justified.

**Protocols**

The methods for this project require several distinct parts. In addition to the methods developed by our lab, in several cases methods will be used directly from the protocols or manuals associated with certain equipment (e.g. the RTq-PCR machine). This section will outline in more detail the specific protocols followed for this project. When necessary, equipment protocols will be referenced, though not delineated here.

*Cell culturing and passaging*

To maintain a healthy supply of cells, plates of YAMC and CaCo-2 cells must be properly maintained. To passage cells from an existing cell line, the following procedure is used. Media and Trypsin EDTA (to cleave bonds to the plate) are warmed in a 37°C water bath. Spent media is aspirated from the plate and rinsed with phosphate buffer solution (PBS). The trypsin is added and the plate incubated at 37°C. Fresh media is added and pipetted up and down to detach the

cells from the plate, then cells and media are transferred to a falcon tube. The falcon tube is centrifuged and the supernatant discarded. The pellet is resuspended in 1 mL of media. 10 µL of media and cells are pipetted into a microcentrifuge tube and mixed with a staining mixture. The mixture is transferred to a hemocytometer, where cells can be counted. Cell density is calculated, and between 8.25*10^5 and 1.25*10^6 cells are seeded onto a new petri dish. The new dish is labeled and incubated for 3-7 days until 80% confluent.

*Seeding Cells with Substrate*

Cells are seeded into two 96-well plates. A plating scheme, shown below in *Table 1*, is developed, which includes a control row, DMF solvent only row, indole at 100 µL (I100), and indole at 1000 µL (I1k). Ignoring triplicates of identical wells, there are effectively four columns. These four columns account for the RTq-PCR reference gene and target gene, as well as the positive control, TCDD, and its absence. Therefore, any given column contains either HRPT (housekeeping gene) or CYP1A1 (target gene) and either TCDD or absence of TCDD. The plate is then covered and centrifuged for 2 minutes and incubated for 16 hours.

**The Indole experiment plating scheme**

*Table 1: TCDD+ treatments are underlined; target gene wells are marked in **bold**.*

| P1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | CTRL | CTRL | CTRL | CTRL | CTRL | CTRL | **CTRL** | **CTRL** | **CTRL** | **CTRL** | **CTRL** | **CTRL** |
| B | DMF | DMF | DMF | DMF | DMF | DMF | **DMF** | **DMF** | **DMF** | **DMF** | **DMF** | **DMF** |
| C | I100 | I100 | I100 | I100 | I100 | I100 | **I100** | **I100** | **I100** | **I100** | **I100** | **I100** |
| D | I1k | I1k | I1k | I1k | I1k | I1k | **I1k** | **I1k** | **I1k** | **I1k** | **I1k** | **I1k** |
| E | | | | | | | | | | | | |
| F | | | | | | | | | | | | |
| G | | | | | | | | | | | | |
| H | | | | | | | | | | | | |

*Seeding cells – indole-BSA*

The protocol for seeding wells for Indole-BSA conjugations is similar to the previous section. The only difference is the use of a conjugated Indole-BSA reagent rather than pure indole. (This conjugation was carried out in large batches by other members in our lab who were using it for other experiments.) Each row, shown in *Table 2*, consists of a similar reagent scheme as *Table 1*: a control row, solvent control row, and Indole-BSA row at 100µM (CTRL, DMF, I-BSA respectively).

**The Indole-BSA experiment plating scheme**

*Table 2: TCDD+ treatments are underlined; target gene wells are marked in **bold**.*

| P1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | CTRL | CTRL | CTRL | CTRL | CTRL | CTRL | **CTRL** | **CTRL** | **CTRL** | **CTRL** | **CTRL** | **CTRL** |
| B | DMF | DMF | DMF | DMF | DMF | DMF | **DMF** | **DMF** | **DMF** | **DMF** | **DMF** | **DMF** |
| C | IB100 | IB100 | IB100 | IB100 | I100 | I100 | **IB100** | **I100** | **I100** | **IB100** | **I100** | **I100** |
| D | | | | | | | | | | | | |
| E | | | | | | | | | | | | |
| F | | | | | | | | | | | | |
| G | | | | | | | | | | | | |
| H | | | | | | | | | | | | |

*Isolating RNA*

Intracellular reactions have occurred, but there is no way to quantify them unless separation and purification methods are used to isolate the reaction products. One of the most straightforward ways to analyze the results of adding controls, antagonists, and agonists is to isolate and quantify the mRNA transcripts.

A protocol is used for this section that uses proprietary reagents and mixes to separate mRNA from the remaining components of the cell, including DNA. The steps will not be outlined here, although emphasis will be placed on the following information: RNA is highly susceptible to degradation. Human saliva contains RNase, which is one way in which cells degrade RNA once it is no longer needed. This means that great care must be taken when working in this section.

Isopropyl alcohol and bleach must be used on the lab bench, and the lab worker must not talk or open his/her mouth while working on this section of the procedure.

*Quantifying RNA*

The Nanodrop is used to quantify the amount of RNA present from each well. Exactly 1.5 µL from each sample are placed on the drop receiver. The concentration is recorded. Additionally, for each sample, the 260A/280A and 260A/230A absorbance wavelength ratios are taken. These numbers give information on the purity of the RNA sample and any presence of DNA.

*Amplification using RTq-PCR*

Quantitative Real-Time Polymerase Chain Reaction is used to determine levels of expression. Given the concentrations from the quantification step, the appropriate volume from each well can be loaded into the RTq-PCR machine to start with the same amounts of RNA. The PCR machine is run using its standard protocols.

Using the LightCycler 96 program, RTq-PCR data is recorded and saved. The data is exported and analyzed using ddCq analysis. The final data is in the form of fold changes over control (e.g. the number 20 indicates a twentyfold increase in expression over the normalized control value).

# CHAPTER III

# RESULTS

There are two essential and distinct sets of data to consider. The first experiment ignored I-BSA and simply considered the knockdown effects of indole on TCDD. The second experiment compared the knockdown of indole and indole conjugated to BSA (I-BSA).

## The Indole experiments

The following graph is developed from LightCycler 96 RTq-PCR data. The machine provides Cq values, which are based on preset levels of fluorescence that indicate how much RNA is present. Then ΔΔCq values are calculated. The first delta represents the difference between CYP1A1, the expressed gene, and HRPT, the housekeeping gene. This normalizes the mRNA expression to the base level of a gene whose transcript levels do not change with the introduction of TCDD. The second delta represents the difference between treated (TCDD+) cells and untreated (TCDD) cells. This comparison is as standard as finding a change in temperature, volume, pressure, etc. where a relative value is more meaningful than an absolute one.

The fold changes are normalized to the control values, such that the fold change for the control wells is 1. Therefore, a fold change of 20, for example, represents a twenty-fold increase in mRNA expression over the control conditions.

The fold changes are graphed in the figure below. As expected, the fold change in CYP1A1 expression for TCDD- cells is approximately one, whereas the fold change for TCDD+ cells is

much greater than one. Compared to the control, I100 indole experiences knockdown, and I1k experiences even greater knockdown.

**The Indole experiment – normalized fold increases**

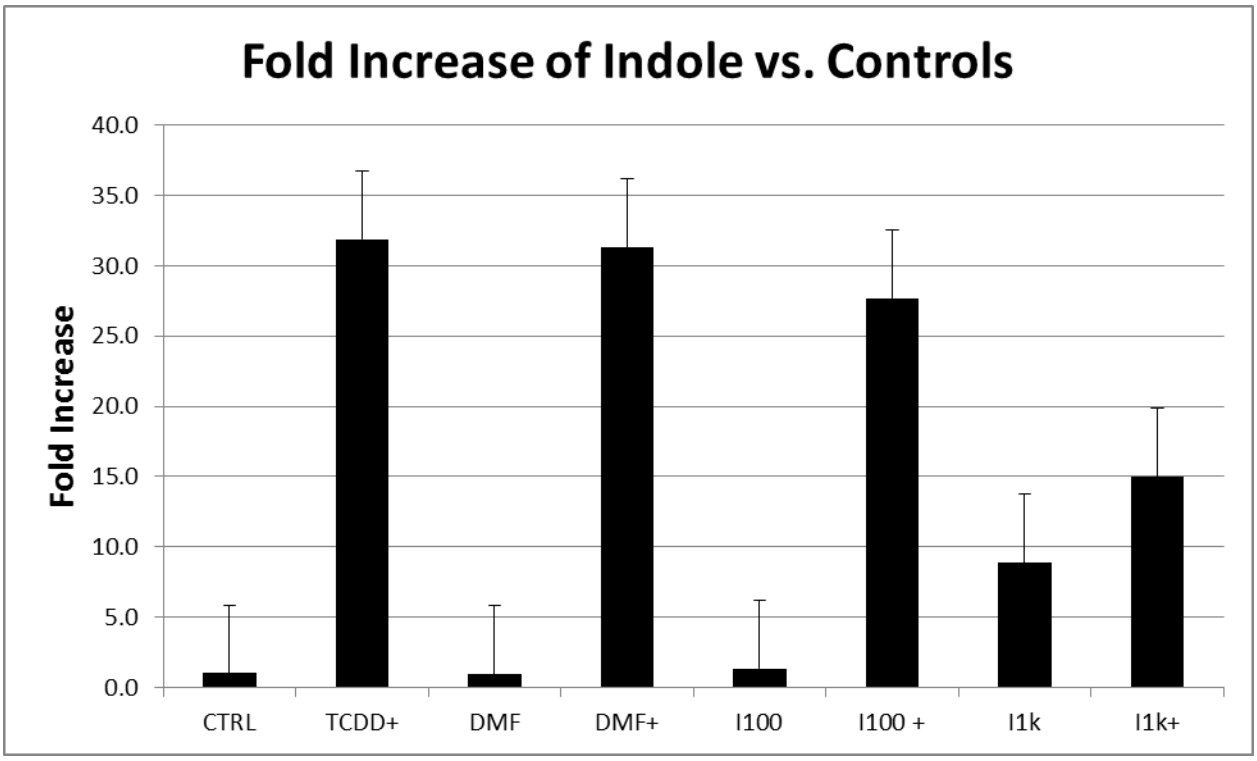


*Figure 2: The results for I1k+are statistically significant to 95% confidence.*

As evident in *Figure 2*, there is significant variation in each of the bars (as shown by the error bars). However, to 95% confidence there is significance in the difference between I1k+ and TCDD+ or DMF+. This means that there is a significant correlation between the presence of indole and the occurrence of a knockdown of inflammatory response. Because we, the researchers, are supplying the indole, there is time dependence to the correlation. Therefore, we can safely and accurately deduce causation.

This graph does not indicate a statistically significant difference between I100 and DMF+ or TCDD+. However, previous tests in our lab have shown significance here, albeit with a much smaller knockdown than I1k. This seems to suggest that perhaps indole at the lower concentration is near the minimum concentration sufficient to induce a knockdown response.

The Ahr signaling pathway by definition follows enzyme kinetics. The exact model is unknown, and may be hyperbolic, sigmoidal, or another relationship. Regardless, any system that follows enzyme kinetics has a minimum value for substrate concentration to be effective, and a maximum value that indicates saturation of the enzyme or receptor. Thus it is reasonable to conclude that indole at 100 µM may be near the minimum effective substrate concentration for the given conditions. While not a fundamental part of this study, analysis of such an activity/substrate concentration curve is applicable to future studies.

**The Indole-BSA experiments**

The indole-BSA scheme underwent the same analysis as the previous indole scheme. RTq-PCR was used to analyze the results, and the same procedure transformed the raw data into a meaningful form.

**The Indole-BSA experiment – normalized fold increases**

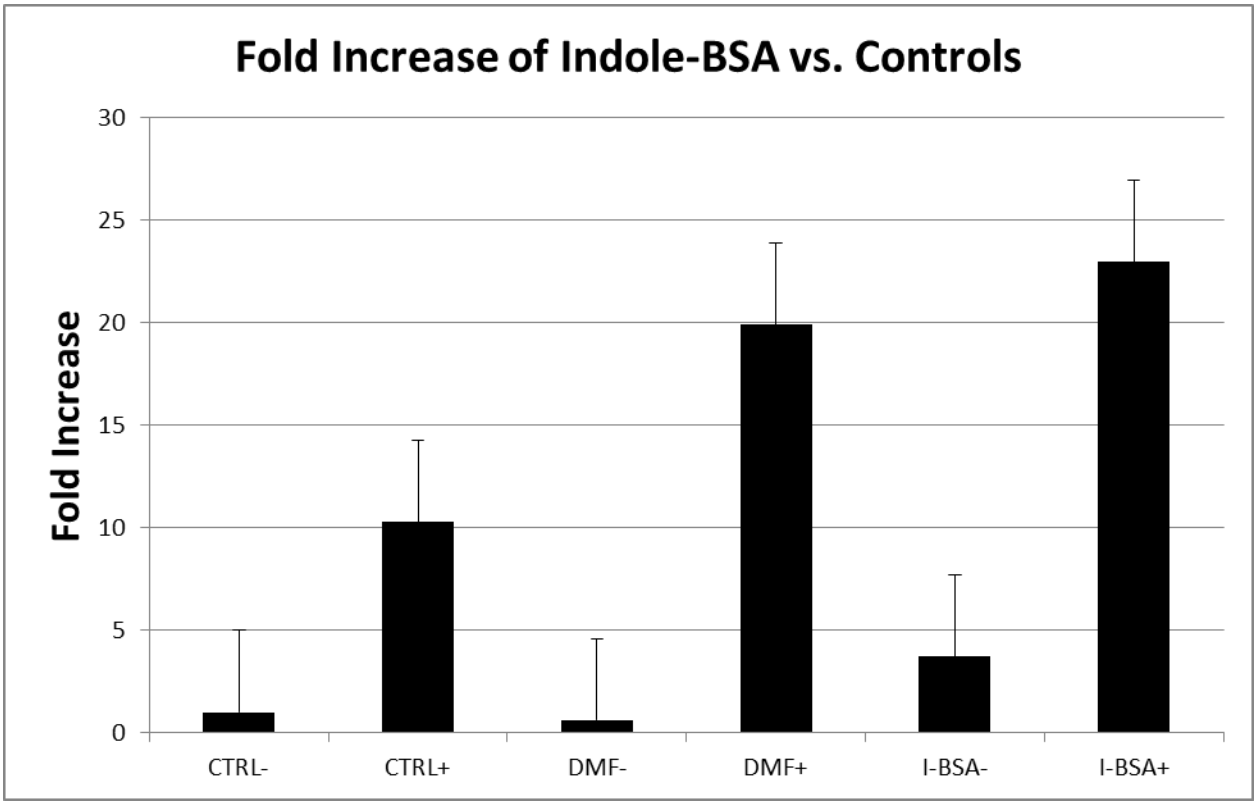


*Figure 3: No results are statistically significant.*

As shown in *Figure 3*, the results are more complicated than the results of the indole only experiments. The TCDD- conditions for the control, solvent control (DMF), and I-BSA are all relatively consistent. However, analysis of the TCDD+ results proves more difficult. The DMF control, expected to be the same expression as the normal control CTRL, actually has a significantly higher response. It is unclear what may be causing such a response.

The I-BSA response, expected to be lower than the controls, is actually higher than the CTRL response (to statistical significance). However, it is not statistically different from the DMF control. Clearly, some sort of confounding variable is present. If I-BSA is ineffective, no

change relative to the controls would be expected.  But instead a higher response is present. Potential causes, confounding variables, and solutions will be discussed in the next section.

# CHAPTER IV

# CONCLUSION

This study has demonstrated that indole is capable of downregulating an inflammatory response by acting as an antagonist to the Ahr receptor. Indole was successfully taken up by the YAMC cells, and the measured response – *CYP1A1* expression – was successfully measured and determined to be statistically significant.

However, despite evidence supporting the hypothesis of an extracellular pathway for indole activation, the Indole-BSA experiment was unable to support this hypothesis. One might be tempted to hastily conclude that no such pathway exists, or that indole is incapable of extracellular regulation. But the results obtained cannot conclusively reject our hypothesis, because the data is of poor quality. There is no consistent trend that can refute our hypothesis for the following reasons.

During tests, our lab experienced frequent cell death in the lines treated with Indole-BSA. This did not occur in the cells treated only with indole. This severely limited the number of trials performed, leading to a small data set. Whereas the indole experiments were consistently reproduced, the Indole-BSA experiments were rather sporadic.

Secondly, other research and prior lab experiments suggest not rejecting our hypothesis outright. BSA is likely not truly an "inert" molecule and is affecting the pathway in some manner. At 67

kDa per molecule, even micromolar concentrations of BSA are putting large quantities of protein in the media. This is likely affecting cell surface receptors and signaling.

Finally, the Indole-BSA conjugation created in our lab contains uncertainties as well. The chemistry utilized for this conjugation involves linking indole to the amino acid lysine. BSA has approximately a dozen lysine residues that are located in a favorable position for indole conjugation. This means that any given BSA molecule may have multiple indoles conjugated, and some BSA molecules may have none. This inconsistency certainly results in a less refined experimental setup.

Given this knowledge, the goal is to continue to refine this experiment and eliminate confounding variables associated with BSA and other factors. Our lab is currently researching other molecules to be used for an indole conjugation. We hope to prove that indole plays an extracellular role, or at least confidently disprove this hypothesis.

Until then, the indole-Ahr pathway can continue to be investigated and discovered, for it holds promise in unlocking the secrets of bacteria-metabolite mediated responses. It is remarkable that a collection of bacteria that reside as symbionts in the mammalian gut can play such a vital role in immunological and inflammatory responses. Perhaps one day this knowledge will be used to treat diseases, and doctors will treat the microbiota as much as they treat the human tissues.

# REFERENCES


[1]2015. [Online]. Available: 1. https://www.ncbi.nlm.nih.gov/pubmed/21161399. [Accessed: 08- Sep- 2015].

[2]Y. Cheng, U. Jin, C. Allred, A. Jayaraman, R. Chapkin and S. Safe, 'Aryl Hydrocarbon Receptor Activity of Tryptophan Metabolites in Young Adult Mouse Colonocytes', Drug Metabolism and Disposition, vol. 43, no. 10, pp. 1536-1543, 2015.

[3] Ncbi.nlm.nih.gov, 'Real-Time Quantitative Reverse Transcription PCR', 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/probe/docs/techqpcr/. [Accessed: 04- Sep- 2015].

[4] Ncbi.nlm.nih.gov, 'CYP1A1 cytochrome P450, family 1, subfamily A, polypeptide 1 [Homo sapiens (human)] - Gene - NCBI', 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/gene/1543. [Accessed: 04- Sep- 2015].

[5]e. Sridharan GV, 'Prediction and quantification of bioactive microbiota metabolites in the mouse gut. - PubMed - NCBI', Ncbi.nlm.nih.gov, 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/25411059. [Accessed: 20- Sep- 2015].

[6] Sigmaaldrich.com, 'Bovine Serum Albumin lyophilized powder, BioReagent, suitable for cell culture, ≥96% (agarose gel electrophoresis) | Sigma-Aldrich', 2015. [Online]. Available: http://www.sigmaaldrich.com/catalog/product/sigma/a9418?lang=en&region=US. [Accessed: 12- Nov- 2015].