INVESTIGATIONS OF MICRORNA REGULATION OF THE URIDINEDIPHOSPHATE
GLUCURONOSYLTRANSFERASE (UGT) GENE FAMILY OF ENZYMES

A Dissertation in

Genetics

By

Douglas F. Dluzen

© Douglas F. Dluzen

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

May 2014
The dissertation of Douglas F. Dluzen was reviewed and approved* by the following:

Philip Lazarus  
Boeing Professor and Chair  
Washington State University College of Pharmacy  
Dissertation Advisor  
Committee Co-Chair

Gavin P. Robertson  
Professor of Pharmacology, Pathology, and Dermatology  
Committee Co-chair

Faoud T. Ishmael  
Assistant Professor of Pulmonary Medicine

Jianming Hu  
Professor of Microbiology and Immunology

Curtis J. Omiecinski  
H. Thomas and Dorothy Willits Hallowell Chair  
Professor of Veterinary & Biomedical Sciences

David J. Spector  
Professor of Microbiology and Immunology

*Signatures are on file at the Graduate School
Abstract

The management and treatment of cancer and other diseases has been a central focus of biomedical research. Identifying novel mechanisms underlying patient response to cancer risk factors and therapeutic stratagems can offer clinicians new avenues for prevention and treatment. In particular, there is a large degree of inter-individual variability of drug metabolism within patient populations. The mechanisms governing human drug metabolism are very complex and the purpose of this dissertation work is to identify novel mechanisms contributing to this process.

The UDP-Glucuronosyltransferase (UGT) gene family codes for 19 unique protein isoforms responsible for the phase II metabolism of many endogenous and exogenous compounds. This includes >60% of all known drugs. The UGT enzymes attach a sugar moiety (glucuronic acid) to target substrates including endogenous compounds such as bilirubin and steroid hormones as well as drugs and environmental agents including tobacco specific nitrosamines (TSNAs), morphine, and chemotherapeutic compounds such as raloxifene. The UGTs exhibit substantial inter-individual variability with respect to their mRNA expression, protein expression, and enzymatic activity. This contributes to variability in patient response and toxicological issues associated with compounds and therapies metabolized by the UGTs. Understanding the mechanisms governing UGT expression and regulation can provide insight into patient response to disease risk factors and treatments. We hypothesized that microRNA (miRNA) contribute to the expression regulation of UGT protein family members.

miRNA are an endogenously expressed class of RNA species that do not code for proteins. miRNA generally bind to the 3’ untranslated region (UTR) of mRNA and help guide a large protein complex called the RNA-induced silencing complex (RISC) to
the target mRNA. RISC binding represses protein expression of the targeted mRNA via inhibition of protein translation and often causes mRNA degradation. Using in silico prediction analyses, we have identified miR-491-3p as an endogenous regulator of hepatic expression of UGT1A protein expression. Transfection of miR-491-3p mimics into HuH-7 cells significantly repressed UGT1A1 (P<0.001), UGT1A3 (P<0.05), and UGT1A6 (P<0.05) mRNA levels. This repression correlated with a significant reduction in the metabolism of raloxifene into raloxifene-6-glucuronide (ral-6-gluc; P<0.01) and raloxifene-4'-glucuronide (ral-4'-gluc; P<0.01). When endogenous miR-491-3p expression levels in HuH-7 cells are repressed, there is a significant increase (~80%; P<0.01) in UGT1A1 mRNA and a corresponding increase in glucuronidation of raloxifene into ral-6-gluc (50%, P<0.05) and ral-4'-gluc (22%, P<0.01). Knockdown of endogenous miR-491-3p in HepG2 cells did not significantly alter UGT1A1 mRNA levels, but did increase the formation of ral-6-gluc (50%, P<0.05) and ral-4'-gluc (34%, P<0.001).

Furthermore, a significant (P<0.05) increase in UGT1A3 and UGT1A6 mRNA levels was observed in miR-491-3p non-expressing versus expressing liver specimens and both UGT1A3 (P<0.05; r = -0.296) and UGT1A6 (P<0.01; r = -0.487) mRNA levels are inversely correlated with miR-491-3p expression levels in human liver. These results suggest that UGT1A isoforms are regulated by miR-491-3p. This is the first study to examine miRNA regulation of the UGT drug-metabolizing enzyme family.

The UGT2B gene family is comprised of six unique genes and protein isoforms. We have identified miR-216b as a novel epigenetic regulator of UGT2B protein expression. Luciferase reporter assays confirmed miR-216b conserved binding motifs within the 3’ UTRs of UGT2B4, 2B7, 2B10, and 2B15 that were identified using in silico prediction algorithms. The functionality of these binding sites were further confirmed using in vitro pull-down assays. Over-expression of miR-216b mimics in HuH-7 cells significantly repressed UGT2B7 (P<0.001) and UGT2B10 (P<0.01) mRNA levels and over-
expression of miR-216b in Hep3B cells decreased UGT2B4 ($P<0.001$) and UGT2B10 ($P<0.01$) mRNA levels. UGT2B7 protein levels were significantly repressed in a dose-dependent manner in the presence of increasing miR-216b concentrations in HuH-7 cells and this was correlated with decreasing levels of epirubicin glucuronide formation ($P<0.01$). Inhibition of endogenous miR-216b levels increased UGT2B7 mRNA levels in HuH-7 ($P<0.05$) and in Hep3B cells ($P<0.01$) and increased UGT2B7 epirubicin glucuronidation by approximately 40% and 25%, respectively ($P<0.05$). This is the first evidence indicating that miRNA genes can regulate members of the UGT2B gene family and that miR-216b regulation of UGT2B proteins may play a role in the metabolism of UGT2B substrates in human liver.

The results presented in this dissertation identify novel miRNA regulators of UGT protein expression. miR-491-3p and miR-216b directly regulate expression of UGT1As in the liver and UGT2B isoforms in cell lines, respectively, and contribute to the complex mechanisms involved in UGT-mediated drug metabolism. These studies could have important implications in inter-individual expression of UGT enzymes. This will be imperative in identifying predictors of patient response to clinical therapies.
# Table of Contents

LIST OF FIGURES .................................................................................................................. viii
LIST OF TABLES .................................................................................................................... x
LIST OF ABBREVIATIONS ....................................................................................................... xi
ACKNOWLEDGEMENTS .......................................................................................................... xiii

Chapter 1: REVIEW OF THE LITERATURE ................................................................................. 1
  1.1 General Background and Significant ............................................................................. 2
    1.1.1 Introduction to Cancer and Personalized Medicine .................................................... 2
    1.1.2 Cancer Epidemiology .................................................................................................. 3
  1.2 The Molecular Biology of Cancer .................................................................................. 3
    1.2.1 The Stages of Tumorigenesis .................................................................................... 4
    1.2.2 Epigenetics and Cancer .............................................................................................. 6
    1.2.3 Genetic Variation and Cancer Susceptibility ............................................................... 7
  1.3 Overview of Drug Metabolism ....................................................................................... 8
    1.3.1 Phase I Drug Metabolism ............................................................................................ 9
    1.3.2 Phase II Drug Metabolism .......................................................................................... 10
  1.4 The UDP-Glucuronosyltransferase (UGT) Gene Family ................................................ 11
    1.4.1 UGT Nomenclature, Organization, and Homology ..................................................... 12
    1.4.2 UGT Cellular Location, Protein Structure, and Protein Interactions ....................... 15
    1.4.3 UGT Tissue Expression .............................................................................................. 18
  1.5 Importance of UGTs and Metabolism ............................................................................ 20
    1.5.1 UGT Glucuronidation of Endogenous and Exogenous Compounds ......................... 20
    1.5.2 UGT Pharmacogenetics: Impact on Cancer Risk and Drug Metabolism ..................... 24
    1.5.3 Mechanisms of UGT Expression Regulation ............................................................ 28
      1.5.3.1.1 Alternative Splicing of the UGT Enzyme Family ................................................ 28
      1.5.3.1.2 Regulation of the UGT Enzymes via Tissue-Specific Transcription Factors ........ 31
      1.5.3.1.3 Regulation of UGT Enzymes via Ligand-Activated Transcription Factors .......... 33
    1.5.4 UGT Summary ............................................................................................................. 34
  1.6 Regulation of Gene Expression via MicroRNA ............................................................... 35
    1.6.1 MicroRNA Nomenclature, Structure, Biogenesis, and Function ............................... 36
    1.6.2 The miRNA Seed Sequence and Target Recognition ................................................. 41
    1.6.3 miRNA and Cancer .................................................................................................... 42
    1.6.4 miRNA and Drug Metabolism .................................................................................... 45
  1.7 AIMS and Central Hypothesis ....................................................................................... 49

Chapter 2: MICRORNA 491-3p Regulation of UGT1A protein expression in human liver ........ 50
  2.1 Abstract ............................................................................................................................. 51
  2.2 Introduction ....................................................................................................................... 52
  2.3 Materials and Methods .................................................................................................... 54
  2.4 Results ............................................................................................................................... 62
  2.5 Discussion ........................................................................................................................ 75
Chapter 3: miR-216b Regulation of UGT2B Family Enzymes

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Abstract</td>
<td>82</td>
</tr>
<tr>
<td>3.2 Introduction</td>
<td>83</td>
</tr>
<tr>
<td>3.3 Materials and Methods</td>
<td>85</td>
</tr>
<tr>
<td>3.4 Results</td>
<td>94</td>
</tr>
<tr>
<td>3.5 Discussion</td>
<td>110</td>
</tr>
</tbody>
</table>

Chapter 4: Future Directions and Final Considerations

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Future Directions</td>
<td>116</td>
</tr>
<tr>
<td>4.1.1 miR-491-3p and Cancer</td>
<td>116</td>
</tr>
<tr>
<td>4.1.2 miR-216b and Cancer</td>
<td>119</td>
</tr>
<tr>
<td>4.1.3 Transcriptional Induction of UGT-Regulating miRNAs</td>
<td>124</td>
</tr>
<tr>
<td>4.1.4 Environmental Induction of UGT-Regulating miRNAs</td>
<td>124</td>
</tr>
<tr>
<td>4.2 Final Considerations</td>
<td>126</td>
</tr>
</tbody>
</table>

Appendix: Investigations into MicroRNA Regulation of Members of the Hepatic Nuclear Factor Family

REFERENCES

128 153
LIST OF FIGURES

Figure 1.1. Overview of phase I and phase II drug metabolism
Figure 1.2. The UDP-Glucuronosyltransferase reaction mechanism
Figure 1.3. The genomic organization of the UGT1A gene locus
Figure 1.4. The Genomic organization of the UGT2 gene locus
Figure 1.5. Phylogenetic tree of the UGT gene family
Figure 1.6. Schematic of the hypothesized UGT active structure and ER localization
Figure 1.7. Representative outline of the major domains of the UGT enzymes
Figure 1.8 Schematic of the alternatively spliced exon 5 and 3’ UTR of the UGT1A gene locus
Figure 1.9 The nomenclature of miRNA genes
Figure 1.10 Schematic of the canonical pathway of miRNA biogenesis
Figure 1.11 Schematic of the proposed binding of both miR-491-3p and miR-125a-3p to the UGT1A1 3’ UTR
Figure 2.1. The UGT1A gene locus and potential miR-491-3p miRNA recognition elements (MRE) within the common UGT1A 3’ UTR
Figure 2.2. Tissue and cell line expression of miR-491-3p and UGT1A family members
Figure 2.3. Quantitative analysis of UGT1A mRNA and enzymatic activity in HuH-7 Cells
Figure 2.4. The effect of miR-491-3p inhibitor on UGT1A1 expression and activity in HuH-7 cells
Figure 2.5. The effect of miR-491-3p inhibitor on UGT1A1 expression and activity in HepG2 cells
Figure 2.6. Expression of UGT1A3 and UGT1A6 mRNA versus miR-491-3p in human liver specimens
Figure 2.7: UGT1A mRNA secondary structure as assessed by the RNA Fold program
Figure 3.1. In silico prediction and binding of miR-216b to UGT2B isoforms
Figure 3.2. Tissue and cell line expression of miR-216b
Figure 3.3. UGT2B 3’ UTR luciferase activity in the presence of miR-216b mimics
Figure 3.4 RISC and miR-216b association with UGT2B isoform 3’ UTRs

Figure 3.5. UGT2B mRNA expression in HuH-7 and Hep3B cells

Figure 3.6. UGT2B mRNA levels in HuH-7 and Hep3B cells with 50 nM miR-216b

Figure 3.7. UGT2B mRNA levels in HuH-7 and Hep3B cells with inhibited miR-216b

Figure 3.8. UGT2B7 protein expression and activity in the presence of miR-216b mimic or inhibitor

Figure 4.1 Caco-2 UGT1A8 and UGT1A10 mRNA levels

Figure 4.2 Hypothesized miRNA-transcriptional regulatory networks

Figure A1.1 Binding of miR-216b to the HNF4α 3’ UTR

Figure A1.2 Repression of HNF4α protein expression by miR-216b

Figure A1.3 Expression of HNF4α mRNA versus miR-216b in human liver samples

Figure A1.4 miRNA Regulation of HNF4α expression

Figure A1.5 Validation of in silico miRNA target predictions of HNF-family 3’ UTRs
LIST OF TABLES

Table 1.1. UGT activity against specific substrate classes

Table 1.2 List of known human DMEs with miRNA regulators

Table 1.3 List of known human DME transcription factors with miRNA regulators

Table 3.1 List of PCR primers used for UGT2B 3’ UTR PCR amplification and cloning, luciferase mutational analysis, and UGT2B 3’ UTR pull-down assays

Table A1.1 Summary of *in silico* Predictions identifying miRNAs targeting HNF 3’ UTRs
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3' UTR</td>
<td>3' untranslated region</td>
</tr>
<tr>
<td>ADH</td>
<td>alcohol dehydrogenase</td>
</tr>
<tr>
<td>Ago2</td>
<td>Argonaute 2</td>
</tr>
<tr>
<td>AhR</td>
<td>aryl hydrocarbon receptor</td>
</tr>
<tr>
<td>ALDH</td>
<td>aldehyde dehydrogenase</td>
</tr>
<tr>
<td>APC</td>
<td>adenomatous polyposis coli</td>
</tr>
<tr>
<td>ARE</td>
<td>antioxidant response element</td>
</tr>
<tr>
<td>ARNT</td>
<td>hydrocarbon receptor nuclear transport</td>
</tr>
<tr>
<td>BLD</td>
<td>below limit of detection</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CAR</td>
<td>constitutive androstane receptor</td>
</tr>
<tr>
<td>CLL</td>
<td>chronic lymphocytic leukemia</td>
</tr>
<tr>
<td>CNV</td>
<td>copy number variation</td>
</tr>
<tr>
<td>COMT</td>
<td>catechol O-methyl transferase</td>
</tr>
<tr>
<td>CSC</td>
<td>cancer Stem Cell</td>
</tr>
<tr>
<td>CYP450</td>
<td>cytochrome P450</td>
</tr>
<tr>
<td>DME</td>
<td>drug-metabolizing enzyme</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>EMT</td>
<td>epithelial-mesenchymal transition</td>
</tr>
<tr>
<td>ENCODE</td>
<td>Encyclopedia of DNA Elements</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>ERE</td>
<td>estrogen response element</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>FOXA1</td>
<td>forkhead box protein A1</td>
</tr>
<tr>
<td>FRET</td>
<td>fluorescence resonance energy transfer</td>
</tr>
<tr>
<td>FXR</td>
<td>farnesoid X receptor</td>
</tr>
<tr>
<td>GR</td>
<td>glucocorticoid receptor</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione-S-transferase</td>
</tr>
<tr>
<td>GWAS</td>
<td>genome-wide association study</td>
</tr>
<tr>
<td>HNF</td>
<td>hepatic nuclear receptor family</td>
</tr>
<tr>
<td>HNF1α</td>
<td>hepatic nuclear factor 1 alpha</td>
</tr>
<tr>
<td>kB</td>
<td>kilobase</td>
</tr>
<tr>
<td>IncRNA</td>
<td>long non-coding RNA</td>
</tr>
<tr>
<td>LXR</td>
<td>liver X receptor</td>
</tr>
<tr>
<td>MHY</td>
<td>mutY Homolog</td>
</tr>
<tr>
<td>miRNA</td>
<td>microRNA</td>
</tr>
<tr>
<td>MOI</td>
<td>multiplicity of infection</td>
</tr>
<tr>
<td>NAT</td>
<td>N-acetyltransferase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ncRNA</td>
<td>non-coding RNA</td>
</tr>
<tr>
<td>NH2</td>
<td>amine</td>
</tr>
<tr>
<td>NIH</td>
<td>National Institutes of Health</td>
</tr>
<tr>
<td>NNAL</td>
<td>4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol</td>
</tr>
<tr>
<td>NST</td>
<td>nucleotide sugar transport</td>
</tr>
<tr>
<td>NT</td>
<td>nucleotides</td>
</tr>
<tr>
<td>OH</td>
<td>hydroxyl</td>
</tr>
<tr>
<td>PAH</td>
<td>polycyclic aromatic hydrocarbon</td>
</tr>
<tr>
<td>piRNA</td>
<td>piwi-interacting RNA</td>
</tr>
<tr>
<td>PPAR</td>
<td>peroxisome proliferator-activated receptor</td>
</tr>
<tr>
<td>pre-miRNA</td>
<td>precursor microRNA</td>
</tr>
<tr>
<td>pri-miRNA</td>
<td>primary microRNA</td>
</tr>
<tr>
<td>PXR</td>
<td>pregnane X receptor</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>quantitative real-time PCR</td>
</tr>
<tr>
<td>RB</td>
<td>retinoblastoma protein</td>
</tr>
<tr>
<td>RISC</td>
<td>RNA-induced silencing complex</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>RXR</td>
<td>retinoid X receptor</td>
</tr>
<tr>
<td>SDM</td>
<td>site-directed mutagenesis</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>SULT</td>
<td>sulfotransferase</td>
</tr>
<tr>
<td>TCDD</td>
<td>2,3,7,8-tetrachlorodibenzo-p-dioxin</td>
</tr>
<tr>
<td>TGFβ</td>
<td>transforming growth factor beta</td>
</tr>
<tr>
<td>TNBC</td>
<td>triple negative breast cancer</td>
</tr>
<tr>
<td>TPMT</td>
<td>thiopurine S-methyl transferase</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>TSNA</td>
<td>tobacco-specific nitrosamine</td>
</tr>
<tr>
<td>UDPGA</td>
<td>UDP-glucuronic acid</td>
</tr>
<tr>
<td>UGT</td>
<td>UDP-glucuronosyltransferase</td>
</tr>
<tr>
<td>XRE</td>
<td>xenobiotic response element</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

It is very difficult to put into words the effort, patience, and sacrifice needed to obtain a doctorate in the life sciences. I would not have been able to conduct this research and present it here without the generous support and advice from many people. I would first and foremost like to think Dr. Philip Lazarus for going out on a limb to take on a student with no lab and supporting him through his studies to investigate a potentially risky but novel mechanism governing UGT expression. Thank you Phil for your mentorship and guidance over the years and providing and supporting me with the advice and growth necessary to complete this dissertation.

I would also like to thank my committee members for their invaluable input and support. I would especially like to thank Dr. Gavin Robertson for providing me a temporary home for the past year and for the resources necessary to complete my studies. I also want to thank Dr. Faoud Ishmael for his support and collaboration on my pull-down assay experiments. I want to take this opportunity to thank Dr. Dongxiao Sun, Dr. Gang Chen, Dr. Nathan Jones, Dr. Ryan Bushey, Dr. Andrea Primeau, and all other previous Lazarus lab members for their support and help while I worked on this project. I am very grateful for their cheerfulness and input during my struggles. I also want to thank those at Penn State University, too many to name, who have had a helping hand in my progress and research, especially within the Cancer Institute, Department of Pharmacology, Genetics Program, and core facilities.

I want to acknowledge several funding sources for making this research possible including the NIH, AstraZeneca, SwitchGear Genomics, and Sigma Aldrich. The grant money provided by AstraZeneca and Sigma Aldrich in particular provided the opportunity to finish this work. I will also take this moment to disclose that I have no financial or ethical conflicts of interest.
I want to take this moment to thank my wife Julia. I met Julia within the confines and walls of the same building I pursued these research interests, and without her love and support I know this would not have been possible. Thank you Julia for being so wonderful. To my family and friends, especially my sister, thank you for your support, love, and distractions.

I leave now with a quote from the scientist and passionate leader I am inspired by the most, who I would have very much enjoyed to have a chat over a beer with - Dr. Albert Einstein. “Nature shows us only the tail of the lion. But there is no doubt in my mind that the lion belongs with it even if he cannot reveal himself to the eye at all once...”
CHAPTER 1

REVIEW OF THE LITERATURE
1.1 General Background

1.1.1 Introduction to Cancer and Personalized Medicine

Cancer is a multifaceted disease that has claimed the lives of millions of individuals worldwide. Cancer results from the uninhibited growth and proliferation of normal and progenitor cells within the body. Chiefly, the cellular mechanisms responsible for cell maintenance, growth, normal apoptosis, and homeostasis within the surrounding microenvironment become dysfunctional, allowing for the onset and progression of cancer. Genetic mutation is a driver in cancer initiation and progression, and genetic mutations have been identified as markers of cancer risk. A major goal of biomedical research is the study of inherited genetic factors contributing to patient risk of cancer occurrence. Identifying genetic mutations and biomarkers can help researchers and clinicians determine a patient’s individual risk to cancer occurrence, as well as the best possible treatments for a particular diagnosis. This is a central tenant of personalized medicine.

Numerous environmental factors have been implicated in the occurrence of cancer within the general population, including: diet, environmental exposures, and other behavioral risks. Pharmacogenetics and pharmacogenomics is the study of how an individual’s genetic profile influences their response to exogenous drugs such as chemotherapeutic agents and cancer risk factors such as tobacco-related carcinogens. [1]. Pharmacogenetics has given researchers and clinicians a looking glass into the population in the hopes of identifying novel biomarkers and targets for therapeutic treatment. Additionally, technological advances in genome sequencing have allowed biomedical researchers to take a closer look at both the individual genome as well as the cancer genome. Taken together, it is hoped that these tools allow researchers to assess
both individual genetic risk to cancer occurrence and individual response to treatment so that clinicians can create a personalized medical plan.

1.1.2 Cancer Epidemiology

Cancer is a complex and deadly disease that impacts a significant portion of the global population, physically, emotionally, and financially. The National Institutes of Health (NIH) estimates that nearly 12 million Americans were living with some form of cancer in 2008 [2]. Cancer has been estimated to cost over $100 billion in direct medical treatment in 2007 [2]. In 2012, it was estimated that 1.6 million new cases of cancer were diagnosed in the United States and over 577,000 Americans will die from cancer [2]. Cancers of the prostate, lung and bronchus, and colon and rectum are estimated to account for 52% of all new cancer cases in men; whereas in women 53% of new cases will occur in the breast, lung and bronchus, and colon and rectum. Cancer of the lung and bronchus, prostate, breast, colon and rectum, and prostate were estimated to account for over 50% of all cancer deaths in men and women in 2012 [2].

1.2 The Molecular Biology of Cancer

Cancer of nearly all tissues is classically defined by uncontrollable cellular division and proliferation of genetically mutated cells [3]. Cancer cells exhibit six (and oftentimes more) biological capabilities leading to tumorigenesis: inhibition of growth suppressors, uncontrolled cellular proliferation, resisting programmed cell death, the capacity for sustained deoxyribonucleic acid (DNA) replication, induction of angiogenesis, and potential for tissue invasion and metastasis [3]. Somatic mutations controlling these capabilities can allow for cancer cells to proliferate through cell cycle
checkpoints and promote a tumor microenvironment that sustains the growth of the tumor mass [4].

1.2.1 The Stages of Tumorigenesis

The progression from normal to tumorigenic phenotypes can be broken down into three distinct stages: initiation, promotion, and progression [5]. These stages all involve step-wise somatic genetic mutations and clonal expansion of a specific mutational advantage; cancer cells thus experience a natural selective process, not unlike evolution, leading to the progression of a robust cancer phenotype and the possible onset of metastasis [6-8]. Cancerous tumors are also maintained by a small subset of cancer stem cells (CSCs), driving clonal expansion and often surviving initial rounds of drug treatment. CSCs are believed to be a primary reason for cancer relapse within patients in remission, seeding new rounds of cancer cell proliferation and tumor growth [9-11].

The multi-step process of initiation, promotion, and progression is promoted by mutations in the cellular genetic material. Initiation is the first insult to a cell’s DNA, in which a chemical, biological, or physical agent interacts with DNA resulting in an irreversible mutation [12]. This physical alteration of DNA can result in mutations within tumor-suppressor genes, such as the retinoblastoma-association (RB) gene, and can alter cellular cycle checkpoints and growth regulation signaling pathways. This allows mutated cells to enter into the promotion phase of tumorigenesis. Promotion is the stage in which cellular growth signaling and gene expression is altered and actively promotes the tumorigenic phenotype [3, 12]. The RB gene is a classic example of the “two-hit” hypothesis, in which both alleles of a tumor-suppressor gene must be mutated for the initiation of tumorigenesis [13, 14]. Knudson’s “two-hit” hypothesis also differentiates
between two mutational events occurring within one somatic tissue, or one mutational event carried through the germ-line and one event resulting in somatic tissue from environmental insult [13]. In either case, tumorigenesis can result. Important tumor-suppressors can also be found to be haplo-insufficient, in which there are insufficient expression levels of critical tumor suppressor genes required to maintain the homeostasis of normal cellular activity [15]. In these cases, a single mutation in only one of the alleles of tumor suppressor genes such as p53, PTEN, p27, and BRCA1 is enough to promote cells to becoming tumorigenic [15-17].

Cancer promotion involves the reversible change in cellular gene expression, proliferative signaling, and/or apoptotic signaling via some inducing agent [18]. These inducing agents, whether endogenous (such as hormonal signaling) or exogenous (such as carcinogens found in cigarette smoke) often interact with both the initiated cells and the surrounding tissue to promote tumorigenesis [19-21]. Promotion can also be driven by the activation of oncogenes – mutated and/or over-expressed proteins driving cellular function from normal to cancerous phenotypes [5, 22]. Src kinase exhibits elevated expression in numerous cancer tissues and tumors, and can contribute to the progression of cells to an invasive phenotype via associations with proliferation factors, angiogenic factors, and other related pathways [23, 24]. Increases in Ras, B-Raf, and the Akt/PKB pathway have also been implicated as oncogenes via their activation of downstream proliferation and growth pathways [3, 25].

The contribution of DNA mutation, the activation of oncogenes, and loss of tumor-suppressive factors promote tumorigenesis and eventual metastasis of cancerous cells. An important pathway in normal development, called epithelial-mesenchymal transition (EMT), is turned on and used by cancer cells to metastasize from the primary tumor and invade surrounding tissues and systems, eventually developing into
secondary tumors [26]. Mutations in transforming growth factor beta (TGFβ) have been implicated in the EMT of cancer cells in which downstream targets of TGFβ, such as the Smad family of proteins, increase cellular growth and the ability of the mutated cell to invade surrounding tissues [27, 28]. EMT is hallmarked by the aberrant expression and transcriptional regulation of numerous cell-cell anchors, i.e. repression of E-cadherin, and allowing cancer cells to migrate from the primary tumor site [29]. The underlying mechanisms and causes of EMT initiation in cancer are still under investigation.

1.2.2 Epigenetics and Cancer

For decades most of the DNA present in the human genome was labeled ‘junk’ and/or ‘selfish’, with no discernible function and presumed to be the leftovers of human evolution [30]. Recent research has indicated that the ‘junk’ DNA of the human genome provides important structural, regulatory, and heritable forms of gene expression governance. These dynamic regulatory mechanisms, often without any change in the underlying genomic sequence, encompass what is now called the epigenome [31]. The Encyclopedia of DNA Elements (ENCODE) project has now confirmed that >80% of all the human genome has a biological role [32]. Epigenetic mechanisms of gene regulation include: chromatin remodeling and silencing, histone-tail modification, X-chromosome inactivation, and expression of non-protein coding ribonucleic acid (ncRNA) species, including small RNAs such as microRNA (miRNA) [31, 33]. These methods of gene regulation are essential to normal growth and development and are often altered in disease states [32].

Epigenetic regulation of gene expression can be disrupted in cancer [32]. DNA hypomethylation has been observed at CpG islands across the genome in cancer cells, altering gene expression in those cells and contributing to their sustainability and
malignant phenotype [34, 35]. DNA hypermethylation has been observed in tumor samples correlating with repressed expression of tumor suppressors such as p15 and p16 [36]. Global promoter methylation of oncogenes and tumor suppressors can serve as biomarkers for aberrant gene regulation in tumor tissues, and may ultimately provide researchers avenues for new drug therapies [37]. For example, hypermethylation of the BRCA1 tumor-suppressor is correlated with low expression of pRB and p16 in triple negative breast cancer (TNBC) cells, providing insight into the mechanisms of TNBC formation and regulation [38]. Expression of ncRNAs are implicated in tumor biology: the human non-coding RNA MALAT-1 gene, encoding a 8-kilobase (kb) RNA transcript, is upregulated in human liver and lung cancers and contributes to the regulation of genes associated with cancer cell migration and metastasis [39, 40]. The expression and function of miRNA genes is also associated with cancer and drug response, and will be discussed in greater detail later in this chapter. Overall, epigenetic regulation of gene expression is critical to the functionality of normal growth and development and continued investigation in this field of biology may provide potential avenues for both cancer drug targets and/or biomarkers for cancer risk factors or therapeutic response.

1.2.3 Genetic Variation and Cancer Susceptibility

Genetic variability in the population can predispose individuals to certain types of cancer. For example, individuals who have inherited mutations and/or single nucleotide polymorphisms (SNPs) within the BRCA1 or BRCA2 DNA repair genes are at high risk to develop breast, ovarian, and other cancers [41]. Germ line mutations within the DNA repair enzyme adenomatous polyposis coli (APC) and mutY Homolog (MHY) genes predispose individuals to the development of colorectal cancer and cancers of the stomach [42, 43]. Over the last decade, researchers have identified numerous other genetic variants linked to the risk of cancer development using genome-wide association
studies (GWAS). These studies incorporate the genetic variation of thousands of individuals to identify novel, high-risk genomic loci associated with the development and/or progression of cancer and other diseases, including breast cancer, non-small cell lung cancer, and many others [44-46]. These high-throughput studies can help uncover novel markers for cancer susceptibility as well as possible therapeutic targets.

Genetic variation has been identified in drug metabolizing enzymes (DMEs) such as the cytochrome P450s (CYP450s) and the UDP-glucuronosyltransferases (UGTs) and are associated with altered enzymatic activity and/or cancer risk [47, 48]. Some polymorphisms have been linked to social behavior and response to environmental influences; eg genetic polymorphisms in the CYP2A6 gene, responsible for the inactivation of nicotine to cotinine, are linked to smoking behavior and nicotine clearance [49]. A functional missense polymorphism in the UGT2B10 gene, responsible for the metabolism of numerous tobacco-specific nitrosamines (TSNAs), exhibits decreased activity against TSNAs such as 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) and may increase individual susceptibility to tobacco-related cancers [50, 51]. Identifying novel polymorphisms and regulating mechanisms of DMEs has been an on-going effort to uncover novel targets for cancer therapeutics and biomarkers predicted of drug efficacy [52]. This review will now focus on DMEs, including their broad scope and relevancy to cancer biology and therapy.

1.3 Drug Metabolism

Drug metabolism occurs primarily in the liver, although metabolism can occur extra-hepatically as well, and most metabolism occurs in two phases; phase I drug metabolism generally unveils a functional group within a substrate that causes an
increase in substrate polarity, and phase II metabolism conjugates an additional moiety (eg glucuronic acid) which generally leads to metabolic inactivation and readies the substrate for excretion from the body (Fig. 1.1). Polymorphisms and mutations in any step of this process can affect patient response to drug regimes. Individual variability of gene expression and genetic variants affecting enzymatic function of DMEs has been documented in patients and contributes to patient-specific toxicities [53-55]. Many studies have focused on identifying novel pharmacogenetic markers predictive of DME activity and patient drug responses [56, 57].

1.3.1 Phase I Drug Metabolism

Phase I drug metabolism increases the solubility of the parent substrates and unveils a functional group within the substrate via oxidation and/or reduction reactions. These groups include hydroxyl, sulfide, and amine groups (Fig. 1.1) The CYP450s are a super-family of enzymes consisting of 57 actively transcribed genes which are responsible for more than 80% of phase I metabolism and act upon more than 90% of all drugs [53, 55]. The CYPs exhibit high endogenous expression and functional variability within the human population [58]. This variability can be attributed to many factors, including hundreds of identified SNPs located within CYP genes [55, 59]. For example, CY2D6 alone has over 80 polymorphisms which impart a wide range of functional affects, from hyper-activity to no activity [60]. Other enzyme families contribute to phase I metabolism, including alcohol dehydrogenases (ADHs), aldehyde dehydrogenases (ALDHs), peroxidases, and carbonyl (NADPH) reductases.
Figure 1.1: Overview of phase I and phase II drug metabolism. Substrates (S), generally hydrophobic in nature, undergo a functional transformation via phase I enzymes to unmask a function group (e.g., –OH, SH, or NH₂). Phase II enzymes conjugate a polar moiety (e.g., glucuronic acid, acetyl group) which inactivates the substrate, making it more hydrophilic and readily excreted from the body. Highlighted in red are the phase II UGTs and glucuronidated substrates which are the focus of this dissertation.

1.3.2 Phase II Drug Metabolism

Following phase I metabolism, phase II drug metabolism inactivates the parent substrates by attaching a polar moiety such as a sulfuryl or glucuronide group (Fig. 1.1). This conjugated product is more hydrophilic than the lipid-soluble parent substrate and can be readily excreted from the body in the urine, feces, or bile [54]. There are several enzyme families responsible for the majority of phase II metabolism, with the largest contributing metabolic family, the UGTs, responsible for the phase II metabolism of
1.4 The UDP-Glucuronosyltransferase Gene Family.

The UGTs are phase II metabolic enzymes responsible for the glucuronidation of numerous endogenous and exogenous compounds. UGTs, specifically enzymes within the UGT 1A, 2B, and 2A sub-families, catalyze the conjugation of UDP-glucuronic acid (UDPGA) to a functional motif generally unmasked in phase I metabolism (Fig. 1.2) [48, 65, 66].

**Figure 1.2: The UDP-glucuronosyltransferase reaction mechanism.**

UGTs conjugate glucuronic acid from the donor molecular UDPGA to a functional group that is typically unmasked during phase I metabolism. Glucuronic acid is conjugated to a substrate’s –NH2, -OH, or –COOH group leaving a hydrophilic glucuronidated substrate and UDP by-product.
The addition of a glucuronide to the target substrate increases the water solubility of the target substrate, thereby promoting the excretion of the glucuronidated metabolite [54]. The UGTs commonly conjugate UDPGA to free hydroxyl (-OH) or amine (-NH$_2$) groups within the target substrate, with the resulting products called an O-glucuronide or N-glucuronide, respectively. The majority of these UGT enzymes exhibit O-glucuronidation activity, however N-glucuronidation can be conjugated by UGT1A4 and UGT2B10 [67-69]. UGT3 family members use a different sugar donor, in this case transferring N-acetylglucosamine from the donor UDP-N-acetylg glucosamine [70, 71]. UGT8 functions as a galactosyltransferase in the synthesis of glycerolipid and uses UDP galactose as the sugar donor [70, 72].

1.4.1 UGT Nomenclature, Genomic Organization, and Homology

The UGT gene super-family is organized into four families: UGT1, UGT2, UGT3, and UGT8. The UGT1 family consists of nine functional enzymatic isoforms: UGT1A1, 1A3, 1A4, 1A5, 1A6, 1A7, 1A8, 1A9, and 1A10, as well as four transcribed pseudogenes, 1A2, 1A11, 1A12, and 1A13 [73]. The UGT1 family is located on chromosome 2q37 and spans over 200 kb of genomic DNA. The UGT1A family is organized as 13 unique promoters and first exons that are alternatively spliced to shared exons 2 to 5. The common fifth exon also contains a shared 3’ untranslated region (3’ UTR) [66, 73]. Figure 1.3 details the local organization of the UGT1A isoforms, and exemplifies the alternative splicing of UGTs 1A8, 1A6, and 1A1 to the common exons 2 to 5.

The UGT2 gene family is comprised of two sub-families: the UGT2A gene family consisting of UGTs 2A1, 2A1, and 2A3; and the UGT2B family consisting of the seven functional enzymes 2B4, 2B7, 2B10, 2B11, 2B15, 2B17, and UGT2B28, along with five
transcribed pseudogenes 2B24, 2B25, 2B26, 2B27, and 2B29 [66, 73]. The UGT2 gene locus is located on chromosome 4q13 and span almost 1.5 megabases (MB) of genomic DNA. All of the UGT2B genes, and UGT2A3, contain individual promoters and are encoded as individual 6-exon genes, with unique 3’ UTRs [73, 74].

Figure 1.3: The genomic organization of the UGT1A gene locus. Each UGT1A isoform can be alternatively spliced to common exons 2-5, including a common 3’ UTR. Pseudogenes are in grey boxes. Figure not drawn to genomic scale.

UGTs 2A1 and 2A2 share unique first exons alternatively spliced to common exons 2 to 6 [75]. Thus, the UGT2A gene family represents a structural hybrid of the UGT1 and UGT2 gene families. Figure 1.4 details the organization of the UGT2 gene locus, including all pseudogenes. The UGT3 gene family consists of two unique genes, UGTs 3A1 and 3A2, and is located on chromosome 5p13. The last member of the UGT gene family, UGT8, is also one of the least studied, and is located on chromosome 4q26 [73].
Figure 1.4: The genomic organization of the UGT2 gene locus. Each functional UGT2B enzyme is encoded by a unique 6-exon gene (white boxes). The UGT2B pseudogenes are non-functional (grey boxes). The UGT2A family members are shown in brown boxes. UGT2A1 and 2A2 contain alternatively spliced first exons to shared exons 2-6. Figure not drawn to genomic scale.

The UGT enzyme family exhibits a high degree of sequence homology among isoforms [65, 73, 74, 76]. Figure 1.5 represents a phylogenetic alignment of all UGT family member DNA sequences. The UGT1A family shares at least 60% DNA sequence homology among the 1A isoforms, with the majority of sequence differences occurring within the first exon [73]. The UGT2B gene family members share over 70% DNA sequence homology, with a majority of differences found in the first exons of each 2B gene member [73]. The UGT2B gene family is believed to have evolved over multiple genomic duplication events, with UGT2B4, 2B15, and 2B17 being the ancestral UGT2B isoforms (Figure 1.5) [77]. There is close to 40% DNA sequence homology among all UGT family isoforms and families [54].
1.4.2 UGT Cellular Localization, Protein Structure, and Protein Interactions

The UGT proteins are bound to the membrane of the endoplasmic reticulum (ER) with the enzymatic domains located within the lumen of the ER. Figure 1.6 provides a hypothesized structure of the UGT proteins exemplifying the N-terminal amino acids located within the ER lumen and C-terminal cytosolic tail [76, 78, 79]. UGT enzymes are divided into several basic domains: the N-terminal domain contains the ER retention signal and the amino acid structure responsible for each isoforms’ substrate specificity; the UDPGA binding domain conserved in almost all of the UGT isoforms; and the transmembrane domain containing the cytosolic tail (Figure 1.7). The ER retention signal is eventually cleaved from the mature protein during final processing and localization [78]. The UGT substrate binding domain and UDPGA binding domain are linked together.
by a ~30 amino acid bridge partially responsible for the tertiary structure of the enzymes [80].

![Diagram of UGT active structure and ER localization](image)

**Figure 1.6: Schematic of the hypothesized UGT active structure and ER localization.** UGT enzymes are anchored to the membrane of the endoplasmic reticulum by a C-terminal transmembrane domain. The C-terminal domain also includes a cytosolic tail. The UGT enzymatic domains and N-terminus are located within the lumen of the ER.

Several members of the UGT gene family contain post-translational modifications. UGTs 1A6, 1A9, 2B7, 2B15, 2B17 are N-glycosylated and glycosylation may be required for proper folding of the UGT1A9 enzyme [81]. UGTs 1A1, 2B7, and 2B15 are phosphorylated by Src kinase at several amino acids and these modifications are also required for enzymatic function [82, 83]. For instance, mutational analysis of threonine residues in the N-terminus of UGT1A1 (positions T-75 and T-112) identifies these phosphorylated residues as critical to UGT1A1’s functionality [82]. This is not surprising considering that many of the amino acids in the substrate binding domain of the UGT proteins are essential for substrate specificity and binding. Histidine-39 is conserved in all UGT1A and UGT2B family members except UGT1A4 and 2B10; and is required for glucuronidation reactions involving proton transfer [84]. UGT1A4 and 2B10 contain a proline or leucine, respectively, at the corresponding position to histidine-39 that is important for their glucuronidation activities against tertiary amine substrates [67, 84].
UGT family members may also interact with one another to form homodimers or heterodimers. UGT1A1 can homodimerize in the ER membrane and this interaction is pH-dependent [85]. Fluorescence resonance energy transfer (FRET) analysis in COS cells identified the capabilities of all UGTs in the 1A family to homodimerize and UGT1A1’s ability to heterodimerize with all 1A family members with the exception of UGT1A5 [86, 87]. UGT2B7 has also been shown to homodimerize and the same study identified an 18-amino acid sequence motif that is essential for this interaction and which is present in all other UGT2B isoforms [88]. The interactions between UGT family members can impact their enzymatic activity. For example, when UGT1A1 is co-expressed with either UGT1A4 or 1A6 there is a significant increase in the maximal enzymatic rate ($V_{\text{max}}$) of bilirubin O-glucuronidation (bilirubin is exclusively glucuronidated by UGT1A1) [89]; whereas there is a decrease in the $V_{\text{max}}$ of estradiol-3-O-glucuronidation when 1A1 is co-expressed with UGT1A9 [89].

The UGT enzymes also rely on additional proteins to transport their substrates into and out of the ER. Members of the nucleotide sugar transporter (NST) gene family, including UDP-galactose transporter-related isozyme 1 and human UGTrel7, transport
UDPGA and UDP-N-acetylgalactosamine into the ER lumen for use by the UGT family members [90]. This process is independent of UGT protein expression [90]. Liver microsomes from the Gunn rat did not have impaired UDPGA uptake in the presence of inactive UGT1A enzymes [91]. Gunn rats, which contain a coding mutation in exon four of the UGT1A gene locus, do not express functional UGT1A enzymes [92]. The export of glucuronidated product out of the ER lumen is also dependent on additional enzymes, including the multi-drug resistant proteins isoforms 1 to 4 [90, 93-95].

1.4.3 UGT Tissue Expression

The UGT mRNAs and proteins are expressed in a wide-range of tissues in the human body. The tissue with the highest overall UGT expression is the liver, the primary organ of metabolism [96]. Using quantitative real-time, reverse transcription PCR (qRT-PCR) and immunoblotting, UGT mRNA and protein expression has been identified in numerous extra-hepatic tissues involved in metabolism including the kidney, stomach, small intestine, colon, jejunum, and spleen [97, 98]. Tissues of the aero-digestive tract, including the esophagus, mouth, trachea, larynx, olfactory epithelium, and lung also express UGT family members, albeit at lower levels than in the liver [52, 97, 98]. Additionally, UGT mRNA and protein expression can be found in steroidogenic tissues including the breast, thyroid, thymus, prostate, testis, adrenal glands, heart, ovaries, placenta, cervix, bladder, uterus, and endometrium, as well as brain, cerebellum, and adipose tissues [96, 98, 99]. While there is wide-spread tissue expression of the UGT genes, it should be noted that not all UGT family members, especially of the UGT1A and 2B sub-families, are expressed in all of the tissues listed above. There is also considerable variability of UGT isoform mRNA expression within each tissue. In the liver
for example, mRNA levels vary by at least 100-fold for each of the UGT2B genes (Jones et al. *DMD, In press*) and UGT1A1 mRNA levels can vary 9-fold, UGT1A4 mRNA can vary 28-fold, and UGT1A6 mRNA can vary 22-fold [96].

UGTs 1A7, 1A8, and 1A10 are extra-hepatic enzymes, primarily expressed in digestive tract tissues including the intestine and colon [100, 101]. UGT1A5 mRNA levels in liver are very low, with slightly higher expression in the intestine [102]. Quantitative analysis of UGT1A and 2B expression in 25 human livers identified UGT2B4 as the highest expressed mRNA, followed by UGT2B7, UGT2B10, UGT1A1, and UGT1A6. Additionally, the combined mRNA expression levels of the UGT2Bs in the liver is close to 70% of all total UGT mRNA [96]. UGT2B4 mRNA levels are nearly 60% of the total UGT2B mRNA in the liver, followed by UGT2B7 (~30%), UGT2B10 (~8%), UGT2B17 (~2%), UGT2B15 (~1%), and UGT2B11 (~0.5%). There was no UGT2B28 expression detected in the liver or in other extra-hepatic metabolic tissues expressing UGT2B isoforms including the pancreas or lung (Jones et al., *DMD, In press*). UGT2B4, UGT2B7, UGT2B10, and UGT2B15 mRNA levels are also significantly correlated with one another in the human liver (Jones et al., *DMD, In press*).

UGT2B4 is extra-hepatically expressed at low levels (compared to liver) in the pancreas, lung, larynx, esophagus, tonsil, mouth and tongue (Jones et al., *DMD, In press*). UGT2B7 mRNA expression is detectable at low levels in extra-hepatic tissues including the pancreas, esophagus and lung and UGT2B10 mRNA levels are high in the tonsil and tongue, but these levels are lower overall than in the liver. UGT2B10 is also expressed in the pancreas but at low levels compared to liver (Jones et al., *DMD, In press*). UGT2B15 and UGT2B17 expression is higher than other UGTs in the breast and prostate, and UGT2B17 is expressed at high levels in the lung and pancreas [64] (Jones et al., *DMD, In press*). UGT2B15 and UGT2B17 are heavily involved in the metabolism
of endogenous hormones such as estrogens and androgens which target tissues of the breast and prostate. Extra-hepatic expression of UGT2B genes is highest in the tonsil, in which expressions levels were similar to that of normal liver (Jones et al, DMD, In press).

UGT2A1 mRNA expression is not found in the liver but in extra-hepatic tissues including the larynx, tonsil, colon, lung, and trachea and UGT2A2 is also extra-hepatically expressed in tissues including the breast, trachea, kidney, and larynx [52, 75, 103]. UGT2A3 mRNA is highly expressed in the liver and is expressed extra-hepatically in the colon, pancreas, lung, tonsil, kidney, trachea, and larynx [103].

The UGT gene family also exhibits hepatic and extra-hepatic expression of alternative mRNA splice variants and proteins. These variants, along with their tissue expression and impact on UGT regulation, will be discussed in greater detail later in this review.

1.5 Importance of UGTs and Metabolism

1.5.1 UGT Glucuronidation of Endogenous and Exogenous Compounds

The UGT enzyme family influences and regulates the metabolism of important endogenous and exogenous compounds. Their expression and enzymatic function contributes to the balance of endogenous steroids and hormones, bile acids, and toxic by-products found in the body, such as bilirubin [69, 104]. Additionally, the UGTs are particularly important in the metabolic clearance of carcinogens and environmental toxins, as well as up to 70% of all known drugs, consumed by humans [54, 61].

The endogenous and exogenous substrates glucuronidated by the UGT gene family are diversified across a broad range of substrate classes, including simple and
complex phenol substrates, primary, secondary, and tertiary amines, heterocyclic amines, and steroids of varying carbon-chain length [69]. Updated from Tukey et al., Table 1.1 summarizes the UGT isoforms responsible for the metabolism of specific substrate classes [69]. Table 1.1 also illustrates the significant overlap of UGT-substrate specificity within several substrate classes. For example, UGT1A1 can metabolize phenols, flavones, coumarins, and C_{18} steroids. But C_{18} steroids can also be glucuronidated by UGTs 1A3, 1A4, 2B4, 2B7, and 2B17, depending on the particular substrate. UGTs 1A4 and 2B10 are the only UGTs responsible for the metabolism of secondary and tertiary amines [67, 69].

<table>
<thead>
<tr>
<th>Substrate Class</th>
<th>UGTs Responsible for Metabolism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple phenols</td>
<td>1A1, 1A3, 1A4, 1A5, 1A6, 1A7, 1A8, 1A9, 1A10, 2B4, 2B7, 2A1, 2A2, 2B15, 2B17</td>
</tr>
<tr>
<td>Complex phenols</td>
<td>1A1, 1A3, 1A4, 1A6, 1A7, 1A8, 1A9, 1A10, 2B4, 2B7, 2A1, 2A2, 2B15, 2B17</td>
</tr>
<tr>
<td>Aliphatic alcohols</td>
<td>1A4, 1A9, 2B7, 2B15, 2A1</td>
</tr>
<tr>
<td>Anthraquinones/flavones</td>
<td>1A1, 1A3, 1A7, 1A8, 1A9, 1A10, 2B15, 2A1</td>
</tr>
<tr>
<td>Courmarins</td>
<td>1A1, 1A3, 1A6, 1A7, 1A8, 1A9, 1A10, 2B15, 2A1</td>
</tr>
<tr>
<td>Bile acids</td>
<td>1A3, 2B4, 2B7, 2A2, 2A3</td>
</tr>
<tr>
<td>Carboxylic acids</td>
<td>1A3, 1A9, 2A1</td>
</tr>
<tr>
<td>Primary amines</td>
<td>1A3, 1A4, 1A6, 1A9, 2B10</td>
</tr>
<tr>
<td>Secondary amines</td>
<td>1A4, 2B10</td>
</tr>
<tr>
<td>Tertiary amines</td>
<td>1A4, 2B10</td>
</tr>
<tr>
<td>Heterocyclic amines</td>
<td>1A3, 1A6, 1A7, 1A8, 1A9, 1A10</td>
</tr>
<tr>
<td>Opioids</td>
<td>1A3, 1A8, 2B7, 2A1</td>
</tr>
<tr>
<td>C18 steroids</td>
<td>1A1, 1A3, 1A4, 1A7, 1A8, 1A9, 1A10, 2B7, 2B15, 2A1, 2A2</td>
</tr>
<tr>
<td>C19 steroids</td>
<td>1A4, 2B15, 2B17, 2A1</td>
</tr>
<tr>
<td>C20 steroids</td>
<td>1A4, 2B17, 2A1</td>
</tr>
</tbody>
</table>

Table 1.1: UGT activity against specific substrate classes. This table is adopted from Tukey and Strassburg [69]. Additionally, activity data is incorporated for UGTs 1A5, 2B10, 2A1, 2A2, and 2A3 [51, 75, 102, 105]. Many of the UGTs also metabolize isoform-specific substrates.
Many of the UGTs also metabolize isoform-specific substrates. A classic example of this specificity is the metabolism of the heme breakdown product bilirubin, a toxic, endogenous compound that can cause hyper-bilirubinemia and jaundice if not probably excreted from the body [106]. UGT1A1 is the sole enzyme responsible for the glucuronidation of bilirubin [106]. There are few UGT isoforms other than UGT1A1 that are solely responsible for the metabolism of an individual endogenous compound [104]. One example is UGT1A6, which is the only UGT isoform to glucuronidate the neurotransmitter serotonin [107]. Several exogenous compounds are also glucuronidated exclusively by specific UGTs; UGT1A1 and the chemotherapeutic agent etoposide [108, 109]; UGT2B10 and UGT2B17 and the nicotinic metabolite trans-3'-hydroxycotinine [110]; UGT2B7 and the chemotherapeutic agent epirubicin [111, 112], to name a few.

UGT substrate-specificity can be driven by the stereochemistry of the substrate. For example, estradiol, an endogenous estrogen hormone, is a diastereomer at the 3-hydroxyl and 17-hydroxyl positions. UGT2B4, UGT2B7, and UGT2B17 selectively glucuronidate estradiol at the 17- hydroxyl position, but UGT2B4 only conjugates a glucuronide to epiestradiol (17α-estradiol) whereas UGT2B17 exclusively glucuronidates β-estradiol (17β-estradiol) [113]. UGT2B7 is capable of conjugating both diastereomers [113, 114]. Several UGTs are involved in the metabolism of dibenzo [def,p] chrysene (DB[a,l]P) enantiomers. DB[a,l]P is a carcinogen found in cigarette smoke. UGTs 1A1, 1A4, 1A7, 1A8, 1A9, 1A10, and 2B7 are capable of glucuronidating either one or both of the DB[a,l]P-trans-11,12-diol enantiomers. The enantiomers of DB[a,l]P-trans-11,12-diol (DB[a,l]P-(+)-trans-11S,12S-diol and DB[a,l]P-(-)-trans-11R,12R-diol) are pro-carcinogenic [115]. The metabolism of both DB[a,l]P enantiomers is specific to UGT-expression patterns in the respiratory-tract tissues and plays a role in the first-pass
metabolism of this carcinogen [115]. In another example, raloxifene, a chemotherapeutic reagent used in the prevention and treatment of breast cancer, is metabolized by UGT1A1 and 1A8 at both the 6-β- and 4′-β- positions; whereas UGT1A10 only conjugates a glucuronide at the 4′-β- position [116].

There is considerable variability with respect to glucuronidation of UGT substrates. Serotonin glucuronidation varies >13 to 40-fold in human liver microsomes and UGT1A6 expression influences the glucuronidation and therefore the levels of serotonin in the body [107, 117]. UGT1A9 protein expression levels in human liver microsomes can vary >17-fold and contributes to patient variability in the glucuronidation of mycophenolic acid and propofol [118]. Interindividual variability in glucuronidation of other UGT-substrates has been observed in human liver samples, including >15-fold variability in the glucuronidation of acetaminophen (via UGTs 1A1, 1A6, and 1A9); 235-fold variability in the glucuronidation of anastrozole (via UGT1A4); 70-and-50-fold variability in the glucuronidation of NNAL-N-glucuronidation (UGT2B7), respectively; and >10-fold variability in the glucuronidation of fenofibric acid (via UGTs 1A3, 1A6, 1A9, and 2B7) [119-122].

Variability in glucuronidation in the general population is not surprising given the observed interindividual variability in UGT protein levels in human liver tissue samples. In a panel of 25 human liver microsomes, protein levels of UGTs 1A1, 1A4, 1A6, and 2B7 were variable 6-, 13-, 2-, and 4-fold respectively and protein levels of 1A4, 1A6, and 2B7 were not correlated with their respective mRNA levels [96]. UGT1A1 and 1A9 protein levels are not correlated with their mRNA levels in human livers as well [123, 124].

UGT proteins can also become deregulated in cancer and this may also play a role in individual pharmacological response to drug treatments. UGT1A protein isoforms,
including 1A1, 1A3, 1A4, and 1A9 exhibit decreased expression in liver and colon cancer tissue as compared with normal tissues [125, 126]. UGT1A10 and 2B7 mRNA levels are significantly decreased in breast cancer tumors as compared to normal controls, altering metabolism of 4-hydroxylated estrone in breast cancer cells [127]. Polymorphisms in the coding region and promoter elements of the UGTs can contribute to UGT protein expression and are very influential in patient response to drug treatments, as well as environmental cancer risk factors.

1.5.2 UGT Pharmacogenetics: Impact on Cancer Risk and Drug Metabolism

Because the UGTs are so heavily involved in the metabolism of hormones, carcinogens, and drugs, numerous research efforts have focused on studying the pharmacogenetics of UGT metabolism. Common polymorphisms within UGT1A1’s TATA box promoter element results in a ~70% reduction of UGT1A1 protein expression and activity and in increased bilirubin concentrations in the body, causing Gilbert’s disease [128]. This variant is transcribed by the UGT1A1*28 allele and contains seven TA repeats in the TATA box promoter element of the UGT1A1 promoter as opposed to the six TA repeats found in the wild-type promoter [129-131]. The UGT1A1*28 allelic frequency is ~40% prevalent in the population [132]. The UGT1A1*28 allele has been associated with toxicity in cancer patients undergoing chemotherapeutic treatment with irinotecan. Irinotecan is metabolized into the active metabolite SN-38, which is glucuronidated by UGT1A1 [133]. Patients with the UGT1A1*28/*28 genotype are eight times more likely to have hematologic toxicities when on a high dosage of irinotecan and have over three-fold decreased UGT1A1 enzymatic activity compared to wild-type [57, 129]. Because of the important role of the UGT1A1 *28 allele in irinotecan metabolism a
Food and Drug Administration (FDA) approved genotyping test is used by clinicians to identify UGT1A1*28 alleles in the patient population taking irinotecan to help tailor irinotecan drug regimens [134, 135].

A majority of the UGT family members contain promoter and coding polymorphisms that influence expression and/or enzymatic activity. Three non-synonymous polymorphisms within the UGT1A6 gene (amino acid changes S7A, T181A, R184S) contribute to the observed >40-fold glucuronidation variability of UGT1A6 substrates in human liver microsomes. The UGT1A6R184S variant exhibits a two-fold decrease in the UGT1A6 K_m value, and both the UGT1A6S7A and UGT1A6T181A variants showed a two-fold increase in V_max [117]. These differences affect overall UGT activity [117]. Two UGT1A9 promoter-polymorphisms are correlated with higher UGT1A9 glucuronidation of mycophenolic acid and propofol in human liver samples as a result of altered transcriptional regulation [118]. Investigations in the Lazarus laboratory reveal that the UGT1A4L48V (UGT1A4*3) variant exhibits a ~4-fold increase in the V_max/K_m for the formation of both olanzapine glucuronides and contribute to interindividual variability in olanzapine metabolism [136]. The UGT1A4*3 and UGT1A1*28 variants contribute to the altered metabolism of clozapine in human liver samples, and the UGT1A4*3 variant was predictive of clozapine 5-N-glucuronide formation [137]. UGT1A4 is the only UGT enzyme that conjugates a glucuronide moiety to tamoxifen's amine group [138]. The UGT1A4*3 variant plays a role in interindividual response to tamoxifen therapy. UGT1A4*3 over-expressing cells exhibit higher levels of glucuronidation against tamoxifen and both isomers of 4-OH-tamoxifen [138]. UGTs 1A10, 1A8, and 2B7 exhibit the highest O-glucuronidation of 4-OH-tamoxifen and endoxifen. Tamoxifen metabolism was not significantly affected by the UGT1A8A173G or UGT1A10Q139K polymorphisms, presumably due to the structurally conserved amino acid changes in the variant alleles.
[139]. However, UGT2B7 exhibited the highest overall activity against tamoxifen in human liver microsomes [139-141]. A non-synonymous polymorphism within codon 268 of UGT2B7 replaces the wild-type histidine residue with a tyrosine [142]. UGT2B7$^{H268Y}$ variants had an over two-fold decrease in activity against 4-OH-tamoxifen [143]. Patients with a homozygous genotype of the UGT2B7$^{H268Y}$ variant exhibit ~30% reduction in trans 4-OH-tamoxifen in liver microsomes prepared from human liver tissue samples [143]. These investigations suggest that even single mutational events may impact patient response to drug therapy.

Variants of the UGT-enzyme family can also influence the risk individuals have of developing diseases such as cancer after exposure to environmental and dietary carcinogens. The UGT1A4$^{P24T}$ variant exhibits a significantly higher level of N-glucuronidation of NNAL, a very potent pro-carcinogenic found in tobacco smoke. The UGT2B7$^{H268Y}$ variant displays the opposite effect, with liver samples of individuals with the variant genotype having significantly lower activity against NNAL [122]. There was a significant decrease in the metabolism of several carcinogenic polycyclic aromatic hydrocarbons (PAHs) between cell lines over-expressing wild-type UGT2A1 and the codon 75 lysine to arginine UGT2A1$^{L75R}$ variant [52]. This may impact the localized metabolism of tobacco-specific carcinogens in tissues of the aero-digestive tract and influence cancer risk in smokers [52].

Another example of a functional polymorphism is the low-prevalent cysteine to glycine amino acid mutation at codon 183 in UGT1A9. UGT1A9$^{C183G}$ variants have reduced activities in a substrate-dependent manner, and this polymorphism primarily impairs UGT1A9’s ability to form a homo-dimer, affecting enzymatic activity [144]. Some polymorphisms within the UGT gene family can also eliminate the overall function of the enzyme. For example, the codon 67 polymorphism in the UGT2B10 gene changes the
wild-type aspartic acid to a tyrosine residue (the UGT2B10<sup>D67Y</sup> variant is known as the *2 allele). Liver microsomes from individuals heterozygous for the UGT2B10*2 polymorphic allele exhibited 21% to 31% lower glucuronidation of nicotine and cotinine, respectively. Homozygous *2/*2 individuals had 5-to16-fold lower UGT2B10 activity levels against nicotine and cotinine, respectively; suggesting that the *2 allele drastically reduces UGT2B10 activity [145]. The UGT2B10<sup>D67Y</sup> variant is associated with reduced levels of NNAL-N-glucuronidation. Liver samples either heterozygous or homozygous for the *2 allele exhibited a 1.8-to-12-fold reduction in NNAL-N-glucuronide levels [50]. UGT2B10*2 is associated with reduced levels of nicotine-glucuronide and cotinine-glucuronide in smokers’ urine by 42% and 48%, respectively [146]. Together, these observations support the hypothesis that the codon 67 polymorphism in UGT2B10 may influence smoking behavior due to its influence on the metabolism of nicotine itself. With an allelic prevalence of 10% in Caucasians, this polymorphism could also serve as a biomarker for smoking habit and cancer risk [146].

Copy-number variation (CNV) and deletion polymorphisms also play a role in UGT pharmacogenetics. UGT2B17 and UGT2B28 are two of the most commonly deleted genes in the human genome, with both copies of the entire UGT2B17 gene deleted in 10% of Caucasians [147, 148]. UGT2B17 and UGT2B28 are involved in the metabolism of human sex steroids and hormones, and the UGT2B17 gene deletion has been correlated with osteoporosis and other metabolic bone diseases [149, 150]. UGT2B17 plays a role in the detoxification of carcinogens found in cigarette smoke and the UGT2B17 gene deletion has been linked to decreased glucuronidation activity against nicotine, cotinine, 3-hydroxy-cotinine, and NNAL [146, 147]. UGT2B17 exhibits sex-specific expression and activity in humans, and is expressed up to 4-fold higher in men than in women, with corresponding activity levels in human liver samples [99]. The
UGT2B17 gene deletion is associated with an increase in risk of lung cancer and related to a decrease in NNAL glucuronidation levels in women [147].

Polymorphisms and deletions within the UGT genes can serve as biomarkers predictive of UGT functionality [129]. However, many UGT SNPs do not account for the high degree of interindividual variability observed with UGT expression [118]. For example, it was previously mentioned that three non-synonymous SNPs in the UGT1A6 gene can influence activity of UGT1A6 against different substrates; but it was estimated that these polymorphisms contribute to only 15 to 20% of the 40-fold variability shown by UGT1A6 glucuronidation in human liver microsomes, indicating other factors are involved in the regulation of UGT enzymatic activity [117]. Several examples of the factors that contribute to the expression and functional variability of the UGT enzyme family in both normal and cancerous tissues will now be discussed.

1.5.3 Mechanisms of UGT Expression Regulation

1.5.3.1 Alternative Splicing of the UGT Enzyme Family. The UGT gene locus codes not only for the wild-type UGT protein isoforms, but also inactive UGT protein isoforms that can negatively affect wild-type activity. As mentioned previously, the UGT1A gene locus consists of nine functional and alternative 1st exons that are spliced to common exons 2 to 5, giving rise to the wild-type variant 1 (v1) mRNA transcript (Figure. 1.3) [151]. The UGT1A family also contains an alternative exon 5, called exon 5b, that can be spliced to exon 4 of the UGT1A genes, giving rise to the variant 2 (v2) mRNA transcript. Additionally, the wild-type exon 5 (exon 5a) can be spliced to the 3'-end of the v2 transcript, resulting in a third mRNA transcript called variant 3 (v3). These alternatively spliced mRNAs create an additional 18 total transcripts in addition to the wild-type UGT1A mRNAs (v1 transcripts; Fig. 1.8) [125, 151-153]. The v2 and v3 UGT1A mRNA
transcripts code for UGT isoforms that contain a premature stop codon and create a truncated and inactive UGT protein isoform (UGT1A_i2) which lacks the last 90 amino acids present in the wild-type isoforms (UGT1A_i1; Fig. 1.8) [125].

Figure 1.8: Representative schematic of the alternatively spliced exon 5 and 3’ UTRs of the UGT1A gene locus. The UGT1A variant 1 (v1) transcript codes for the functional, wild-type UGT1A_i1 protein. The UGT1A v2 and v3 transcripts code for a non-functional UGT1A_i2 protein containing either exon 5b of the UGT1A gene locus of exon 5b followed by exon 5a.

The v2/v3 mRNA transcripts are expressed at 16-, 17-, 57-, and 29-fold lower levels than the v1 mRNA for UGT1A1, 1A4, 1A6, and 1A9, respectively, in human liver [151]. Additionally, the UGT1A1 v1 and v2/v3 transcripts are each positively correlated with glucuronidation activity in human liver microsomes, suggesting that the v2/v3 transcripts are involved with the enzymatic activity of the UGT1A1_i1 protein [151]. Targeted knock-down of the v2/v3 transcripts in HepG2 cells exhibited an increase of nearly 80% in raloxifene metabolism by UGT1A1_i1 protein. Although the v2/v3 transcripts are expressed at higher levels in HepG2 cells than in normal liver, these results suggest that the UGT1A1_i2 protein coded for by the v2/v3 transcripts have a
mechanistic effect on UGT1A1_i1 enzymatic activity. It is thought that protein-protein
interactions between the wild-type and i2 variants causes the reduced enzymatic activity
of the i1 UGT isoforms in a dominant-negative manner [125]. Co-expression studies of
UGT1A family members with their i2 variant revealed up to an 80% decrease in
glucuronidation activity of the UGT1A_i1 enzymes [152].

Several members of the UGT2B gene family are also regulated by inactive i2
protein isoforms. Extensive alternative splicing has been identified for the UGT2B4 gene
locus, giving rise to 7 additional UGT2B4 transcripts. These variant UGT2B4 mRNAs
code for several inactive UGT2B4_i2 protein isoforms which were found to negatively
regulate glucuronidation activity in a dominant-negative manner when co-expressed with
the UGT2B4_i1 protein [154]. There are 22 additional UGT2B7 mRNA transcripts that
have been identified containing alternative 3’ ends and coding for inactive UGT2B7_i2
proteins with no detectable glucuronidation capability. As is the case for UGT2B4,
UGT2B7_i2 proteins inhibit wild-type UGT2B7 enzymatic activity in a dominant-negative
manner [155]. There have been no studies identifying novel mRNA transcripts for
UGT2B10 or other UGT2B genes at this time.

The UGT2A family members are also subject to the regulation of alternative
splicing, with downstream implications in the functionality of the wild-type isoforms. A
novel slicing event involving the deletion of the entire third exon of UGT2A1 has been
identified and codes for the UGT2A1_i2 protein [75]. The UGT2A1_i2 protein variant
lacks enzymatic activity [75]. Inactive UGT2A1_i2 proteins can dimerize with UGT2A1_i1
and significantly reduce the V_{max} of UGT2A1_i1 activity against compounds such as 1-
hydroxy(OH)-pyrene and 3-OH-benzo[a]pyrene [75], thereby playing a role in the
detoxification of these compounds. The UGT2A1 mRNA transcript containing the exon 3
deletion is expressed in several tissues including the lung, larynx, colon, and tonsil and
may thus have an impact on endogenous wild-type UGT2A1 activity in these tissues [75].

There is also a UGT2A2 mRNA variant that contains a complete exon 3 deletion and the mRNA levels for this variant are expressed at levels 25-50% of the wild-type UGT2A2 mRNA in UGT2A2-expressing tissues [103]. The UGT2A2_i2 protein does not have enzymatic activity and co-expression of the UGT2A2_i2 with UGT2A2_i1 in HEK293 cells reduced UGT2A2_i1 enzymatic activity in a dominant-negative manner [103]. There have been no studies examining UGT2A3 mRNA variants to date.

1.5.3.2 Regulation of UGT Enzymes via Tissue-Specific Transcription Factors. A broad range of transcriptional activators and repressors are responsible for the tissue-specific and ligand-specific expression of the UGT enzyme family. Members of the UGT1A and 2B gene families have unique promoters for each isoform which are responsible for tissue-specific UGT expression [73]. Hepatic nuclear factor 1 alpha (HNF-1α) is a liver-enriched transcription factor which regulates all of the UGT1 and UGT2B genes [156-158]. HNF-1α binds to the palindrome consensus sequence GTTAATNATTAAC in the promoters and proximal promoter elements of the UGT genes, including UGT1A1, 1A3, 1A4, 2B7, and 2B17 [157, 158]. Mutational analysis of the HNF-1α binding site within the UGT1A1 promoter region indicates that HNF-1α is essential for the basal in vitro transcriptional levels of UGT1A1 [159]. HNF-1α is a member of the hepatic nuclear receptor family (HNF), containing members: HNF1α, HNF1β, HNF3α, HNF3β, HNF3γ, HNF4α, HNF4γ, HNF6α, and HNF6β. These tissue-specific transcription factors are primarily responsible for gene regulation in the liver, but can also control UGT gene expression in pancreas, prostate, and colon tissues [157]. Additionally, HNF-1α
promoter-binding is required for enhanced transcriptional activation of UGT1A1 via the aryl hydrocarbon receptor (AhR), constitutive androstane receptor (CAR), pregnane X receptor (PXR), and glucocorticoid receptor (GC) [159, 160]. HNF-1α promoter binding can be enhanced via its interaction with several other transcription factors. For example, HNF-1α can bind with the intestine-specific transcription factor Cdx2 to regulate activation of the extra-hepatic UGT1A8 and UGT1A10 in the intestines [158, 161].

HNF-4α is another master regulator of liver gene expression and differentiation, and is involved with the expression of several UGT family members, including UGT1A1, UGT1A3, UGT1A9, and UGT2B15 in the liver [157, 158, 162]. HNF-4α mRNA levels are correlated with UGT1A1 mRNA levels in human liver [163]. Transcriptional activation of UGT1A9 in the liver via HNF-4α is dependent on co-activation with HNF-1α and both HNF-1α and HNF-4α mRNA levels correlate with UGT1A6 and UGT1A9 mRNA levels in human livers [164, 165]. However, the promoter binding element for HNF-4α is lacking in the promoters of the extra-hepatic UGT1A7, 1A8, and 1A10 enzymes and this contributes to the tissue-specific expression of these extra-hepatic UGTs [165].

These results coincide with the findings that the promoters of the UGT1A family can be methylated to alter expression in a tissue-specific manner [77, 166]. For example, cytosine methylation of CpG sites within the promoter of UGT1A1 can vary 25 to 41% in human livers, and UGT1A1 expression is switched off by methylation of its promoter in colon cancers [166, 167]. Additionally, hypermethylation of the UGT1A10 promoter causes lack of 1A10 expression in the liver; whereas in UGT1A10-expressing colon and intestine tissue the 1A10 promoter is hypo-methylated [158]. This indicates that both HNF-1α and HNF-4α are important liver and intestinal-specific regulators of UGT expression and their binding sites in the promoter elements of the UGT genes are sensitive to epigenetic regulation. Several other transcription factors are important in the
tissue-specific expression of the UGT enzymes. Along with Cdx2, transcription factor Sp1 controls UGT expression in tissues of the gastrointestinal tract [161]. UGT2B15 and UGT2B17 basal expression levels are maintained via the transcription factor fork head box protein A1 (FOXA1) in prostate cancer cell lines. FOXA1 is also known as HNF-3α [168].

1.5.3.3 Regulation of UGT Enzymes via Ligand-Activated Transcription Factors. Ligand-activated transcription factors regulate UGT expression in a tissue-specific manner, as well as in response to external stimuli. Several members of the nuclear receptor super-family of ligand-activated transcription factors regulate expression of the UGT enzymes, including PXR, liver X receptor (LXR), farnesoid X receptor (FXR), retinoid X receptor (RXR), peroxisome proliferator-activated receptor (PPAR), and CAR [157, 158, 169, 170]. Several members of the nuclear receptor family, including CAR and PXR, bind to ligands in the cytoplasm and transfer to the nucleus. Upon entering the nucleus, CAR and PXR can each heterodimerize with RXR and bind to the promoter of several UGT genes, including members of the 1A family [157, 171]. Bilirubin can bind to CAR and directly activate UGT1A1 expression in the liver; an example of direct induction of 1A1 expression in response to fluctuating bilirubin levels [170, 172]. PXR can also be activated by endogenous compounds such as lithocholic acid and drugs such as rifampicin. When bound by its ligand, PXR can activate transcription of UGT1A1, 1A3, 1A4, and 1A6 [173, 174]. LXR is an important regulator of cholesterol homeostasis in the body and can induce UGT1A3 transcription, thereby regulating downstream bile acid metabolism [175]. The bile acid chenodeoxycholic acid is an agonist of FXR, leading to downstream transcriptional activation of the bile acid-metabolizing UGT2B4, but also to a decrease in UGT2B7 expression in colon cancer cells [176, 177]. PPAR is targeting by
endogenous eicosanoids and bile acids, thereby translocation to the nucleus and activating UGT1A9 and UGT2B4 gene expression in the liver (two UGTS responsible for metabolism of bile acids) [157, 176, 178].

AhR has several agonists, including bilirubin, benzo[a]pyrene, and 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) [157, 169]. The AhR protein binds to several known xenobiotic response elements (XRE) in the promoters of UGT1A1, 1A3, 1A4, 1A6, and 1A9 and induces transcription with its binding partner, the hydrocarbon receptor nuclear transporter (ARNT) [157, 179, 180]. Reactive oxygen-species (ROS) also promote the recruitment of the transcription factor Nrf2 to the nucleus and its binding to antioxidant response elements (ARE) in the promoter of UGT1A1 in the liver and colon [181]. Responses to tissue-specific stimuli have also been uncovered for the hormone metabolizing UGT2B15 and UGT2B17. 17β-estradiol can bind to the estrogen receptor and bind to an estrogen receptor element (ERE) located in promoters of both UGT2B15 and 2B17, thereby up-regulating their expression in breast cancer cells [182].

1.5.4 UGT Summary

Given the complex regulatory mechanisms governing UGT expression and activity, it is important to continue to elucidate UGT metabolism in both healthy and diseased individuals. Interindividual variability plays an important role in individual response to environmental cancer risk factors, as well as treatments for diseases such as cancer, schizophrenia, and other metabolic disorders. For several of the UGT1A and 2B family members there is little to no correlation between mRNA and protein levels, as previously detailed. This suggests a possible mechanism of post-transcriptional regulation affecting UGT protein expression levels and ultimately enzymatic response to endogenous and exogenous compounds. We hypothesized that miRNA contribute to the
regulation of UGT family members and the remaining portions of this review will introduce the function and impact of miRNA regulation.

1.6 Regulation of Gene Expression via miRNA

miRNA are an endogenous, gene-encoded, single-stranded RNA species ~22 nucleotides in length. They are a subclass of ncRNA which consist of several highly expressed, endogenous RNA classes that are gene-encoded but are not translated into proteins. Classes of ncRNA include miRNA, piwi-interacting RNA (piRNA), transfer RNA (tRNA), long non-coding RNA (lncRNA), and other RNA species serving various regulatory and enzymatic functions within the cell [183]. miRNA post-transcriptionally regulate protein expression by binding to the 3’ UTR of target mRNA, resulting in decreased mRNA levels and/or inhibition of protein translation [183]. Mature miRNA strands are associated with the RNA-induced silencing complex (RISC), a large conglomerate of RNA-binding proteins and associated structural components that are guided to the target mRNA via the single-stranded miRNA [184]. The mechanism of miRNA regulation was first identified in Caenorhabditis elegans in which a small RNA nucleotide called lin-4 was found to bind the 3’ UTR of lin-14 mRNA and reduce protein expression. This method of gene regulation is an important step in the tissue development of C. elegans [185]. miRNA regulation has now been identified in most domains of the tree of life, including plants as well as viruses [186, 187]. miRNA are an ancient form of gene regulation and signaling that are involved in nearly every aspect of biology, including: tissue development and homeostasis [188, 189], oncogenesis [190, 191], apoptosis [192], cell cycle modulation [193], cell signaling [194, 195], and stem cell maintenance [196-198], among others.
1.6.1 MicroRNA Nomenclature, Structure, Biogenesis, and Function

A nomenclature system has been established to identify and name miRNA genes. Figure 1.9 outlines four different miRNA genes. Each gene is processed into a mature, single-stranded miRNA. The passenger, or *star strand, is removed from the immature miRNA duplex during miRNA processing and degraded, leaving the mature strand, however the passenger strand has been known to have function against mRNA genes (Fig. 1.9, E) [199]. Mature miRNA genes can come from either arm of a single, precursor miRNA duplex. For instance, miR-491-5p matures from the 5’ arm of the miR-491 duplex (Fig. 1.9, A), whereas miR-491-3p derives from the 3’ arm of the stem-loop structure (Fig. 1.9, B). miRNA genes can have paralogs that generally differ by only a few base pairs in sequence. These paralogs are identified as separate, yet sequence-related miRNA genes using lettered suffixes; for example, miR-216a (Fig. 1.9, C) and miR-216b (Fig. 1.9, D) [200].
Figure 1.9: The nomenclature of miRNA genes. A and B, mature single-stranded miRNA genes can be processed from both the 5’ arm and 3’ arm of the same miRNA gene stem-loop, giving rise to miR-5p or miR-3p miRNAs (orange rectangles). C, some miRNA genes have paralogs, in which cases there is only a single nucleotide or two difference in mature sequence (orange and blue squares). Each gene locus is identified as the miRNA gene number followed by a letter suffix to differentiate the unique miRNA gene (for example, miR-216a and miR-216b). E, in every case, the duplex is processed into a single-stranded mature miRNA with the passenger strand (‘) of the duplex being degraded.

miRNA genes are transcribed in the nucleus by RNA polymerase II and the resulting transcripts are capped at the 5’-end and polyadenylated at the 3’-end (Fig. 1.10) [183, 201]. miRNA genes can code for an individual miRNA transcript or as a polycistronic message containing several miRNA genes transcribed within a single RNA transcript. Polycistronic miRNA transcripts can contain several different or related microRNA genes from a single genomic locus [202, 203]. Rarely, miRNA genes can be found within the introns or 3’ UTRs of protein-coding genes, and are spliced out and processed as mature miRNA once the mRNA they come from is processed by the spliceosome [204]. The freshly transcribed miRNA gene is called the primary miRNA (pri-miRNA) and contains a canonical hairpin stem within the middle of the transcript that is ~70 base pairs (bp) long with a terminal, single-stranded loop at the center [183]. Polycistronic pri-miRNA transcripts contain a unique stem-loop structure for each miRNA species it contains. The pri-miRNA stem-loop structure is recognized by the nuclear RNAse enzymes Drosha and DGCR8, which bind to the 5’ and 3’ bases of the stem-loop and catalyze the cleavage of the 5’ and 3’ flanking sequences, resulting in the precursor miRNA (pre-miRNA) transcript [183, 204]. The pre-miRNA is then exported into the cytoplasm via exportin-5 in a GTP-dependent process, where it is further processed by the RNAse Dicer and associated TRBP protein. Dicer recognizes double-stranded RNA species and will catalyze the cleavage of the terminal loop on the end of the stem loop.
The remaining double-stranded RNA bound by Dicer and TRBP is recognized by Argonaute 2 (Ago2) and recruited into RISC [204].

RISC’s main protein component is Ago2 - a member of the Argonaute family of proteins that has a PAZ domain capable of binding RNA sequences and an RNaseH domain for RNA cleavage [183, 206]. Once the processed miRNA duplex is loaded into RISC, the miRNA passenger strand must be removed to allow the mature miRNA gene to function. It is thought that thermodynamic stability of the RNA duplex results in the removal and eventual degradation of the passenger * strand [207]. The mature miRNA strand has a less stable base pair on the 5’ end of the RNA duplex. Ago2 recognizes this energetic difference with the other end of the miRNA duplex and initiates the removal of the * strand [204, 207]. However, it has been shown that both strands of a miRNA gene can be functional at the same time in the same cell [199]. GW182 protein family members are also an integral part of RISC and necessary for miRNA-mediated protein repression. The GW182 family members are important in disrupting the normal ribosomal translation-machinery while RISC is bound to its target mRNA [208, 209].

For instance, rarely, miRNA and RISC can bind to the 5’ UTR of target mRNAs, resulting in an increase in protein expression from the target mRNA [210]. Additionally, miRNA can interact directly with genomic DNA as a transcriptional regulator and can also function as RISC-independent decoys for RNA-binding proteins [211]. These regulatory functions will not be discussed in detail in this review. The predominant role of miRNA is to guide RISC to target mRNA and bind to the 3’ UTR, thereby causing a reduction in protein expression. The two major mechanisms of miRNA interaction causing reduced protein expression are: 1) a reduction in overall translation of the target mRNA, and 2) a decrease in the mRNA levels of the miRNA target [212, 213]. Both play a role in reducing protein output, where translational inhibition occurs quickly and causes
eventual sequestration or destabilization of target mRNA [214]. Experimental evidence has identified that sequestration of mRNA targets in P-bodies, mRNA decay, mRNA cleavage, and mRNA transcriptional inhibition - all miRNA-induced mechanisms capable of affecting target mRNA location and function [215]. miRNA and RISC can also inhibit translational initiation by binding to the 5-methyl cap of mRNA, inhibit translation elongation machinery, support co-translational degradation of nascent peptides, and cause premature termination of the mRNA codon sequences by inhibiting ribosomal movement [215]. More than likely it is a combination of these affects that are contributing to the translational inhibition of target mRNAs.
Figure 1.10: Schematic of the canonical pathway of miRNA biogenesis. miRNA genes are transcribed via RNA polymerase II as either a single gene or part or a polycistronic miRNA-transcript known as the pri-miRNA. The pri-miRNA is cleaved of its 5’ and 3’ flanking sequences by the RNAses Drosha and DGCR8 in the nucleus, leaving a stem-loop precursor ~70 bp in length that is known as the pre-miRNA. The pre-miRNA is exported out of the nucleus via exportin-5 and further processed into a double-stranded RNA duplex by the RNAse Dicer. Ago2, GW182, and other associated RISC components assemble on the RNA duplex and remove the passenger strand, leaving the mature miRNA. RISC will then bind to the 3’ UTR of target mRNA, directed via the mature miRNA sequence, and repress protein expression. There are several mechanisms in which miRNA can influence protein expression.

The resulting effects on protein expression are varied. miRNA can act as either a genetic “switch” or as a “fine-tuner” of protein expression levels. These two mechanisms exemplify the dynamic robustness of miRNA regulation [213]. For instance, in some cases miRNA regulation accounts for only mild (less than two-fold) protein repression and one miRNA gene can mildly repress protein expression from hundreds of different mRNAs [216]. It has also been observed that for proteins reduced by more than a third in overall expression levels, mRNA destabilization was a primary mechanism of reducing protein translation [214]. mRNA destabilization is also the major mechanism of action observed for transcripts that are highly expressed [212]. In a way, miRNA regulation can be thought of as similar to enzyme kinetics and/or titration curves. For instance, as a genetic ‘switch’, enough mRNA levels are needed to accumulate within the cell to overcome the binding of already present miRNA genes. This phenomenon is seen in stem cell maintenance with miR-145. miR-145 expression is highly upregulated during stem cell differentiation, and while levels of miR-145 accumulate, the levels of miR-145 target genes Sox-2 and Oct-4 are reduced and eventually turned-off. This regulatory mechanism switches off the machinery responsible for stem cell self-renewal [198].
When miRNA act as a ‘fine-tuner’ of protein expression, target mRNA levels can vary dramatically dependent upon the concentration of miRNA present. Low-levels of target mRNA can have dramatic effects on protein production, but when mRNA targets are highly-expressed the affect can be more subtle dependent on the miRNA expression levels [213].

### 1.6.2 The miRNA Seed Sequence and Target Recognition

miRNA target recognition is driven by a canonical ‘seed’ sequence. The seed sequence is recognized as nucleotides 2 to 7 from the 5’ end of the mature miRNA that binds in perfect Watson-Crick complementation to a target mRNA. The complementation and binding of the miRNA seed sequence to its target is necessary and sufficient to repress protein expression, even if there is only one copy of the miRNA bound to the mRNA 3’ UTR [217-219]. Binding of the 3’ end of the miRNA to its target is less stringent, in some cases there can be a high degree of complementation to the target, and in some cases 3’-end binding is nonexistent. Figure 1.11 exemplifies two miRNA

\[
491-3p: 3' \text{caucUUCUCCUUA}\text{--GAACGUUUU}
\]

\[
\text{UGT1A1: 5'} \text{uauuAGGAAAACUUUGCAUAa}
\]

\[
125a-3p: 3' \text{ccgaGGUUCUCCGAGGGCAC}
\]

\[
\text{UGT1A1: 5'} \text{gcagCCACG-ACCUCACCUGg}
\]

**Figure 1.11**: Schematic of the proposed binding of both miR-491-3p and miR-125a-3p to the UGT1A1 mRNA. Each miRNA has a strong seven to eight bp ‘seed’ sequence (circled) supported by variable complementarity at the 3’ end of the miRNA strand. The ‘seed’ sequence is an essential factor for miRNA target recognition.
genes, miR-491-3p and miR-125a-3p, and their predicted hybridization to the UGT1A1 3’ UTR. In both cases the essential seed sequence is present (Fig. 1.11, red circles) with variable downstream binding at the 3’ end of the mRNA. Binding at the 5’ end of the miRNA is essential for proper functioning no matter the strength of the 3’-end complementation [217]. Thus, due to sequence similarities and homology in the genome, a single miRNA transcript can regulate dozens to hundreds of different mRNA transcripts within a single cell. Additionally, mRNA transcripts have the potential to be bound by either multiple copies of a single miRNA gene at separate binding sites within the 3’UTR, or to be bound simultaneously by several different miRNA genes [213, 217, 219].

Because of the importance of the seed sequence in miRNA target recognition, several in silico algorithms are available to help identify possible miRNA genes that can bind to a gene of interest. TargetScan, miRanda, and PicTar are examples of algorithms that weigh several criteria to identify potential miRNA gene regulators against a known target mRNA and 3’ UTR sequence [220-222]. Watson-Crick complementation, the folding free-energy (ΔG) of the miRNA:mRNA hybridization, accessibility of the binding site within the 3’ UTR, site conservation among species, and abundance of binding sites along one 3’ UTR are all criteria weighed by several of these prediction programs to help identify true positive interactions [219, 223]. The in silico predictions require experimental validation to confirm true positives.

1.6.3 miRNA and Cancer

miRNA regulation is critically linked to oncogenesis and cancer progression. For example, the miR-15a/16-1 miRNA gene cluster is frequently deleted in patients with B cell chronic lymphocytic leukemia (CLL) and when re-expressed in leukemic cells, both miR-15a and miR-16 induce apoptosis [190]. miR-15a and miR-16 regulate expression
of the anti-apoptotic factor Bcl-2, which is over-expressed in CLL patients when the miR-15a/16-1 gene cluster is deleted [224]. This was the first study identifying the tumor-suppressive role of miRNA in cancer. Since, it has also been discovered that miRNA genes can have oncogenic effects as well, in which case they are known as oncomirs. Classic examples of oncomirs are several miRNAs transcribed from the miR-17-92 gene cluster on chromosome 13. The miR-17-92 cluster consists of six miRNA genes transcribed as one long polycistronic message and the entire locus is frequently over-expressed in several cancers, including solid tumors, lung carcinoma, hematopoietic cancers, and B-cell lymphomas [203]. Considering there are six unique miRNA genes derived from this locus, over-expression of this gene cluster results in a myriad of phenotypes, including; enhanced cellular proliferation, promotion of angiogenesis, inhibited differentiation, and enhanced cancer cell survival. These effects are also cell-type dependent [203]. In lung cancer cells, inhibited expression of miR-17-5p and miR-20a, via anti-sense oligonucleotides, induced apoptosis and helped identify these two miRNA genes as potential targets for miRNA-based drug therapies [192]. Intriguingly, some miRNA genes can act as both oncomirs and tumor suppressors depending on the cellular context. miR-29a is often downregulated in cancer because it can target several oncogenes such as MCL1 and CDK6. When miR-29a is over-expressed in acute myeloid leukemia cells it induces apoptosis [191, 225]. However, miR-29a over-expression in mice B-CLL can induce B cell growth and expansion by inhibiting PXDN expression, thereby reducing cell adhesion for B-cells and promoting a cancerous state [191, 225].

Polymorphisms have been identified in miRNA genes and can affect several aspects of their functionality, however >90% of all miRNA genes do not contain any known SNP [226]. For example, polymorphisms can alter the processing of the pre-
miRNA into the mature form [227]. A SNP located in the precursor of miR-125a can block the processing of miR-125a into the mature form, presumably by creating an altered secondary structure within the precursor that cannot be recognized by the miRNA-processing machinery [227]. Polymorphisms within mature miRNA genes are much more rare, but miRs-125a, 627, and 662 have SNPs within their seed sequences. Seed sequence polymorphisms can alter the target recognition of the mature miRNA, potentially causing aberrant regulation of off-target mRNAs [226]. SNPs in miRNA genes have found to be associated with breast cancer risk and tumorigenesis as well as risk for cancers of the head and neck [228-230]. Polymorphisms in the 3’ UTRs of miRNA-target genes are more common and frequently result in altered miRNA regulation in cancer and other diseases. These mutations within the mRNA can create and eliminate miRNA-binding sites, altering regulation of the polymorphic mRNA [231, 232]. 3’ UTR SNPs have been associated with altered miRNA regulation in sporadic colorectal cancer, and polymorphisms in miRNA-processing genes have been implicated in renal, esophageal, colon, and head and neck cancers [233-237]. Epigenetic regulation of miRNA expression in cancer has also been identified. The promoters of the tumor-suppressive miRNAs miR-148a and miR-34b are hypermethylated in metastatic cancers and re-expression of both miRNA in metastatic cancer cells reduced tumor growth and inhibited metastatic potential [238].

The clinical application of miRNA as a biomarker is steadily growing in use. miRNA expression signatures have been identified and associated with several factors of cancer biology and treatment. miRNome analysis of 540 samples of solid cancer tumors from breast, lung, prostate, and colon identified several over-expressed miRNA genes in all tumors studied and current research is validating these genes as potential drug targets [239]. An expression signature of 13 miRNA genes can differentiate normal
versus malignant B cells in patients with CLL and this signature is associated with disease prognosis and course [240].

miRNA expression signatures are proving to be a useful tool for oncologists to predict disease course as well as identify better therapeautic approaches. miRNA located in plasma serve as an effective avenue to identify novel biomarkers and expression signatures for a variety of cancers. For example, a panel of four miRNA genes (miR-100, miR-125b, miR-138, and miR-768-3p) are predictive of thyroid cancer malignancy and may be used as a diagnostic tool by clinicians to identify patients with aggressive thyroid cancers [241]. Over-expression of miR-451 and miR-186 in blood plasma also serves as a diagnostic biomarker in gastric cancer screens [242]. Additionally, miRNA expression signatures are predictive of patient response to drug therapy. A signature of 28 miRNAs in blood serum are associated with poor response to sunitinib therapy in renal cell carcinoma patients and a signature of 23 other miRNAs in blood serum are correlated with extended response to sunitinib [243]. Expression signatures of 13 miRNA genes in pre-operative biopsies in rectal cancer patients are predictive of pathological response to chemoradiotherapy [244]. Evaluating their direct role in drug metabolism may help clinicians utilize drug treatments in a patient-specific manner.

1.6.4 miRNA and Drug Metabolism

The role of miRNA regulation of DMEs has not been extensively studied. However, a handful of investigations have identified miRNA genes regulating several members of the human CYP450 family. CYP1B1 activates numerous pro-carcinogens and is highly expressed in cancer tissues. miR-27b has been shown to regulate the expression of CYP1B1 and miR-27b expression levels are inversely correlated with CYP1B1 protein expression levels in breast cancer tissues [245]. Two other CYP’s
important for drug metabolism, CYP3A4 and 2E1, are regulated by miR-27b and miR-378, respectively [246, 247]. CYP24, a critical enzyme in calcium homeostasis and vitamin D signaling, is regulated by miR-125b. CYP24 protein levels and miR-125b expression levels are inversely correlated in breast cancer and it is hypothesized that low miR-125b expression in cancer tissues contributes to over-expression of CYP24 [248]. Additionally, UGT-family member UGT8 is regulated by miR-218 in pancreatic cancer cell lines [249]. Although UGT8 is not a DME, it is still a related family member to the drug-metabolizing UGT1 and UGT2 isoforms and suggests additional UGT family members may be subject to miRNA regulation.

Recent research suggests that the upstream transcription factors responsible for transcriptional activation of CYP, UGT, and other DMEs are also regulated by miRNA genes [250]. miR-27b, implicated in regulation of CYP3A4, also regulates the expression of PPARα in human liver [251]. PXR and HNF4α, involved in the expression of both CYPs and UGTs, are regulated by miR-148a, and miRs-34a and 24, respectively [252, 253]. This provides clear evidence that miRNA genes can interact at two levels of transcriptional regulation of DME expression: 1) at the direct level by interacting with the DME mRNA, and 2) indirectly via regulation of upstream DME transcription regulators. Tables 1.2 and 1.3 summarize the known miRNA interactions with DMEs and many of their transcriptional regulators, respectively. These regulatory networks can become quite complex and will need to be further defined to understand the complete role of miRNA regulation of drug metabolism and help identify potential biomarkers for drug response and/or cancer risk factors.
### Table 1.2: List of Known Human Drug Metabolizing Enzymes with MicroRNA Regulators

<table>
<thead>
<tr>
<th>Enzyme Family:</th>
<th>Enzyme:</th>
<th>Regulating miRNA:</th>
<th>Reference:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytochrome P450 (CYP450s)</td>
<td>CYP2E1</td>
<td>miR-378</td>
<td>[246]</td>
</tr>
<tr>
<td></td>
<td>CYP2C8</td>
<td>miR-103, miR-107</td>
<td>[254]</td>
</tr>
<tr>
<td></td>
<td>CYP3A4</td>
<td>miR-27b</td>
<td>[247]</td>
</tr>
<tr>
<td></td>
<td>CYP1B1</td>
<td>miR-27b</td>
<td>[245]</td>
</tr>
<tr>
<td></td>
<td>CYP24</td>
<td>miR-125b</td>
<td>[248]</td>
</tr>
<tr>
<td></td>
<td>CYP2J2</td>
<td>Let-7b</td>
<td>[255]</td>
</tr>
<tr>
<td></td>
<td>CYP19A1</td>
<td>Let-7f; miR-98, miR-181a</td>
<td>[256]; [257]</td>
</tr>
<tr>
<td></td>
<td>CYP1A1</td>
<td>miR-892a</td>
<td>[258]</td>
</tr>
<tr>
<td>UDP-Glucuronosyltransferases (UGTs)</td>
<td>UGT8</td>
<td>miR-218</td>
<td>[249]</td>
</tr>
<tr>
<td>Sulfotransferases (SULTs)</td>
<td>SULT1A1</td>
<td>miR-631</td>
<td>[259]</td>
</tr>
<tr>
<td>Glutathione-S Transferases (GSTs)</td>
<td>GSTP1</td>
<td>miR-133a; miR-133b; miR-513a-3p</td>
<td>[260]; [261]; [262]</td>
</tr>
<tr>
<td>Aldo-keto Reductases (AKRs)</td>
<td>AKR1C2</td>
<td>miR-193b</td>
<td>[263]</td>
</tr>
<tr>
<td>Enzyme Family</td>
<td>Enzyme</td>
<td>Regulating miRNA</td>
<td>Reference</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>---------</td>
<td>------------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Hepatic Nuclear Factor (HNF)</td>
<td>HNF4α</td>
<td>miR-24, miR-34a; miR-449a</td>
<td>[252]</td>
</tr>
<tr>
<td>Peroxisome Proliferator-Activated Receptor (PPAR)</td>
<td>PPARα</td>
<td>miR-21, miR-27b; miR-506; miR-141</td>
<td>[251], [264], [265]</td>
</tr>
<tr>
<td></td>
<td>PPARγ</td>
<td>miR-130; miR-27a; miR-27b</td>
<td>[266], [267], [268]</td>
</tr>
<tr>
<td></td>
<td>PPARδ</td>
<td>miR-9</td>
<td>[269]</td>
</tr>
<tr>
<td>Ligand-Activated Transcription Factors</td>
<td>PXR</td>
<td>miR-148a</td>
<td>[253]</td>
</tr>
<tr>
<td></td>
<td>RXRα</td>
<td>miR-574-3p</td>
<td>[270]</td>
</tr>
<tr>
<td></td>
<td>VDR</td>
<td>miR-125b; miR-27b</td>
<td>[271], [247]</td>
</tr>
<tr>
<td></td>
<td>FXR</td>
<td>miR-421</td>
<td>[272]</td>
</tr>
<tr>
<td></td>
<td>ERα</td>
<td>Let-7b, Let-7i; miR-22; miR-221, miR-222, miR-18a; miR-18b; miR-193b</td>
<td>[273], [274], [275], [276], [277]</td>
</tr>
<tr>
<td></td>
<td>LXR</td>
<td>miR-613; miR-1, miR-206</td>
<td>[278], [279]</td>
</tr>
<tr>
<td></td>
<td>CAR</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AhR</td>
<td>miR-124; miR-375</td>
<td>[280], [281]</td>
</tr>
<tr>
<td></td>
<td>Nrf2</td>
<td>miR-28; miR-93</td>
<td>[282], [283]</td>
</tr>
</tbody>
</table>
1.7 AIMS and Central Hypothesis

miRNA regulation of the drug-metabolizing UGT enzyme family members has not been explored previously. This literature review has examined the data suggesting that there is an additional level of regulation contributing to the observed interindividual expression variability of UGT isoforms. Additionally, there is consensus that mRNA levels of UGT isoforms do not correlate with UGT protein levels. Considering the broad spectrum of miRNA regulation that has been identified in nearly every field of molecular biology, including phase I and II drug metabolism, investigation into miRNA regulation of the UGT enzymes is warranted. Any future insights into mechanisms governing UGT expression and enzymatic and patient response to drug regimens or cancer risk factors can be useful to clinicians.

The central hypothesis investigated in this dissertation is that miRNA contribute to the expression of the drug-metabolizing UGT family members. This umbrella hypothesis was investigated via two major aims 1) identify novel miRNA genes regulating the expression of the UGT1A gene family; and 2) identify novel miRNA genes regulating the expression of the UGT2B gene family. I also performed preliminary studies examining miRNA genes regulating expression of UGT-regulating transcription factors. The remaining chapters and appendices of this dissertation will outline the specific hypothesis, methods, results, and discussion for each of these aims and conclude with an overall summary of the field to date and discussion of future implications and investigations.
Chapter 2:
MICRORNA 491-3p REGULATION OF UGT1A PROTEIN EXPRESSION IN HUMAN LIVER

Reprinted with permission of the American Society for Pharmacology and Experimental Therapeutics. All rights reserved.
Abstract

The UDP-glucuronosyltransferase (UGT) 1A enzymes are involved in the phase II metabolism of many important endogenous and exogenous compounds. The nine UGT1A isoforms exhibit high inter-individual differences in expression but their epigenetic regulation is not well-understood. The purpose of the present study was to examine microRNA (miRNA) regulation of hepatic UGT1A enzymes and determine if that regulation impacts enzymatic activity. *In silico* analysis identified miRNA 491-3p (miR-491-3p) as a potential regulator of the UGT1A gene family via binding to the shared UGT1A 3’ untranslated region common to all UGT1A enzymes. Transfection of miR-491-3p mimic into HuH-7 cells significantly repressed UGT1A1 (*P*<0.001), UGT1A3 (*P*<0.05), and UGT1A6 (*P*<0.05) mRNA levels. For UGT1A1, this repression correlated with significantly reduced metabolism of raloxifene into raloxifene-6-glucuronide (ral-6-gluc; *P*<0.01) and raloxifene-4’-glucuronide (ral-4’-gluc; *P*<0.01). In HuH-7 cells with repressed miR-491-3p expression, there was a significant increase (~80%; *P*<0.01) in UGT1A1 mRNA and a corresponding increase in glucuronidation of raloxifene into ral-6-gluc (50%, *P*<0.05) and ral-4’-gluc (22%, *P*<0.01). Knockdown of endogenous miR-491-3p in HepG2 cells did not significantly alter UGT1A1 mRNA levels, but did increase the formation of ral-6-gluc (50%, *P*<0.05) and ral-4’-gluc (34%, *P*<0.001). A significant inverse correlation between miR-491-3p expression and both UGT1A3 (*P*<0.05) and UGT1A6 (*P*<0.01) mRNA levels was observed in a panel of normal human liver specimens, with a significant (*P*<0.05) increase in UGT1A3 and UGT1A6 mRNA levels observed in miR-491-3p non-expressing versus expressing liver specimens. These results suggest that miR-491-3p is an important factor in regulating the expression of UGT1A enzymes *in vivo*. 
Introduction

MicroRNA (miRNA) are dynamic epigenetic regulators of gene expression in both normal tissue development as well as tumor biology. miRNA post-transcriptionally repress protein expression primarily by binding to the 3’-untranslated region (UTR) of target mRNA and the inhibition of protein synthesis [183, 284]. Acting as a genetic switch and/or fine-tuner of protein expression, miRNA regulation can account for not only mild (e.g., two-fold) protein repression, but can also cause mRNA destabilization to dramatically reduce protein output [212, 213, 216].

miRNA are short, ~22-nucleotide long, single-stranded RNA species that are gene-encoded and transcribed as a primary transcript (pri-miRNA) by RNA polymerase II. Pri-miRNA form a hairpin loop structure that is recognized and processed in the nucleus and then exported to the cytoplasm as a precursor miRNA (pre-miRNA) consisting of a double-stranded duplex and stem-loop. The precursor is further processed to cleave off the stem-loop structure and the RNA duplex is then recognized by the RNA-induced silencing complex (RISC) consisting of Argonaute-2 and associated proteins, which degrades the passenger strand of the RNA duplex and leaves the mature miRNA [183]. The mature miRNA guides RISC to target mRNA 3’ UTRs (less commonly the mRNA open-reading frame or 5’ UTR), with target recognition primarily driven by the hybridization of the miRNA ‘seed sequence’ to the mRNA target. This ‘seed sequence’ of a target gene is defined as nucleotides 2 to 8 from the 5’ end of the mature miRNA strand and forms perfect complementation to its mRNA target [184].

The role of miRNA in tissue development is continually being investigated, but research has shown miRNA regulation to be a major contributor to tissue and cell homeostasis [285, 286], signaling [195], and oncogenesis [287]. The impact of miRNAs
on drug response has not been studied as extensively. miRNA regulate expression of several human cytochrome P450 (CYP540) phase I metabolic enzymes, including the major drug and hormone metabolizing CYPs 3A4, 2E1, 1B1, and 24 [245-248], but there are few studies on other phase I or II enzymes.

The UDP-glucuronosyltransferase (UGT) phase II metabolic enzyme family primarily consists of two large subfamilies, the UGT1As and 2Bs. The UGT1A family, located on chromosome 2q37, codes for nine functional protein isoforms sharing alternative 1st exons spliced to common exons 2 to 5. The common fifth exon contains a shared 3’ UTR (Fig. 2.1A). The UGT2B family members, located on chromosome 4q13, are expressed as individual genes encoded by six exons each with a unique 3’ UTR [73, 76]. The UGTs are generally highly expressed in liver, the primary organ of metabolism, but are also expressed extra-hepatically in tissues such as colon, prostate, lung, pancreas, kidney, and tissues of the head and neck and other aerodigestive tract tissues [75, 97, 151]. Each UGT1A gene is regulated by an individual promoter, driving tissue-specific expression. UGTs 1A1, 1A3, 1A4, 1A5, 1A6 and 1A9 are the primary hepatic isoforms while UGTs 1A7, 1A8, and 1A10 are exclusively expressed extra-hepatically [76, 96, 97]. The UGTs prime numerous endogenous compounds including bilirubin [106] and steroid hormones [288, 289], as well as exogenous compounds including drugs, chemotherapeutic agents, and carcinogens [51, 76, 110, 116, 119, 139, 145, 146, 290-294], for excretion from the body by catalyzing the conjugation of a sugar moiety (e.g. glucuronide) to its substrate. The exon 1-encoded amino-termini of individual UGT enzymes determines their unique substrate-specificities; however there is a high degree of substrate over-lap between UGT family members [76, 157].

The UGT1A family members exhibit extensive inter-individual variability of expression that contributes to variability in patient response and toxicity [119]. Although
polymorphisms in the UGT genes and altered transcriptional regulation can contribute to altered expression and/or activity of UGT1A enzymes [99, 122, 129, 146], these mechanisms do not explain the observations that UGT1A mRNA and protein levels often do not correlate [96, 123, 124] and that there is a high degree of variability in UGT mRNA and protein expression within different tissues and between different individuals [96, 98]. Such observations suggest an epigenetic mechanism of post-transcriptional regulation may be affecting UGT protein expression. We hypothesized that miRNA regulate expression of UGT1A isoforms and that this may contribute to inter-individual variability in UGT1A expression.

The present study is the first to examine miRNA interactions within the UGT1A family of enzymes. In this study, in silico approaches were used to screen for possible miRNA interactions with the UGT1A family of enzymes and functional studies were then performed to examine the potential role of miRNA on UGT1A gene expression. Evidence is presented demonstrating that miR-491-3p may be an important regulator of hepatic UGT1A expression and activity.

Materials and Methods

Chemicals and Reagents. The pGL3-Promoter and pRL-TK renilla plasmids were obtained from Promega (Madison, WI). All synthesized DNA oligos used for 3’ UTR amplification, site-directed mutagenesis (SDM), and polymerase chain reaction (PCR) analysis were from Integrated DNA Technologies, Inc (Coralville, IA). Lipofectamine 2000 transfection reagent was from Life Technologies (Carlsbad, CA). miRVana miRNA mimic miR-491-3p (#4464066), negative control #1 mimic (#4464058), miRVana miRNA inhibitor miR-491-3p (#4464084), and negative control #1 inhibitor (#4464076) were
purchased from Ambion (Austin, TX). Uridine diphosphate glucuronic acid (UDPGA), raloxifene, alamethicin, and bovine serum albumin (BSA) were obtained from Sigma-Aldrich (St. Louis, MO). Ral-6-glucuronide (ral-6-gluc), ral-4′-glucuronde (ral-4′-gluc), raloxifene-d4, ral-6-gluc-d4, ral-4′-gluc-d4, and epirubicin hydrochloride were purchased from Toronto Research Chemical (Toronto, ON, Canada). Rabbit anti-UGT2B7 and anti-calnexin antibodies were from BD Biosciences (San Jose, CA) and Cell Signaling Technologies (Danvers, MA), respectively. Goat anti-rabbit secondary antibody conjugated to hydrogen peroxidase was from Thermo Scientific (Waltham, MA). All other chemicals used were purchased from Fisher Scientific (Pittsburg, PA) unless otherwise specified.

**Cell lines and Culture Conditions.** Human embryonic kidney cell line 293, human hepatocellular carcinoma cell lines HepG2 and Hep3B, human liver adenocarcinoma cell line SK-HEP-1, human prostate carcinoma cell line LNCaP, human colon adenocarcinoma cell line Caco-2, human breast adenocarcinoma cell line MCF-7, and human lung carcinoma cell line A-549 were purchased from the American Type Culture Collection (ATCC, Manassas, VA). The human hepatocellular carcinoma cell line HuH-7 was a kind gift from Dr. Jianming Hu (Penn State Hershey College of Medicine, Hershey, PA). Hep3B, HEK293, A-549, and HuH-7 cells were cultured in DMEM (Gibco, Carlsbad, CA) supplemented with 10% FBS (Atlanta Biologicals, Lawrenceville, GA) and 1% pen/strep (Gibco). HepG2 and Caco-2 cells were cultured in DMEM supplemented with 10% FBS, 1% pen/strep, and 1% non-essential amino acids (Lonza, Basel, Switzerland). SK-HEP-1 and LNCaP cells were cultured in RPMI 1640 (Gibco) supplemented with 10% FBS, 1% pen/strep, and 1% NEAA. MCF-7 cells were cultured in RPMI 1640 supplemented with 10% FBS and 1% pen/strep. All cells were grown and maintained at 37°C with 5% CO₂.
**Tissues and miRNA Isolation.** Colon and endometrium specimens (n=5 each) were obtained from the tissue bank at Pennsylvania State University College of Medicine while liver specimens and their matching total RNA were obtained from the H. Lee Moffitt Cancer Center Tissue Procurement Facility (n=39). All protocols involving the collection and analysis of tissue specimens from these tissue banks were approved by the Institutional Review Boards at their respective institutions and were in accordance with assurances filed with and approved by the United States Department of Health and Human Services. Normal jejunum tissue specimens were purchased from the Sun Health Research Institute (Sun City, AZ). All tissue samples were isolated and frozen at -70°C within 2 h post-surgery. Colon, liver, jejunum, endometrium and cell line total RNA was extracted using the RNeasy Mini Kit (Qiagen, Hilden, Germany). All total RNA samples from cell lines were subject to on-column DNase digestion during RNA purification (Qiagen). Pooled breast total RNA was purchased from the Biochain Institute (Hayward, CA) while lung, pancreas, larynx, trachea, and kidney RNA was purchased from Clontech (Mountain View, CA) or Agilent (Santa Clara, CA). Small RNA (<200 nt) containing the miRNA fraction was isolated and purified from total RNA using the mirVana miRNA Isolation Kit (Ambion). All RNA concentrations were ascertained using a Nanodrop ND-1000 spectrophotometer (Thermo Scientific) and was eluted and stored in RNase-, DNase-free water in a -80°C freezer.

**miRNA Binding Site Predictions.** The 3’ UTR of the UGT1A enzyme locus was obtained from the UCSC Genome Browser (hg18 assembly). miRNA binding site predictions were obtained using, (i) TargetScan [222], scored with the Total Context+ score as described in Garcia et. al [295], and (ii) miRanda, algorithm v3.0 [220] with the following parameters: Gap Open Penalty, -8.00; Gap Extend, -2.00; Score Threshold, 50.00; Energy Threshold, -20.00 kcal/mol; Scaling Parameter, 4.00.
Quantitative Real-time (qRT)-PCR. cDNAs were synthesized from mRNA using total RNA and the SuperScript First-Strand Synthesis Kit (Invitrogen, Carlsbad, CA). miRNA cDNAs were synthesized using the Taqman MicroRNA Reverse Transcription Kit (Ambion). Taqman gene expression assays (Applied Biosystems, Carlsbad, CA) were used to amplify UGTs 1A1 (Hs02511055_s1), 1A3 (Hs04194492_g1), 1A4 (Hs016555285_s1), 1A5 (Hs01374521_s1), 1A6 (Hs01592477_m1), 1A9 (Hs02516855_gH), RPLPO (Hs99999902_m1), and 2B7 Hs02556232_s1) in HuH-7 and HepG2 cells. Taqman miRNA assays (Ambion) were used to amplify the expression of miR-491-3p (Cat. # 4427974, ID #002360) and RNU6B (Cat #4427975, ID #001093) in all cell lines and tissue samples. PCR reactions were performed in 10 µL reactions in 384-well plates using an ABI 7900HT Sequence Detection System with incubations performed at 50°C for 2 min; 95°C for 10 min; and 40 cycles of 95°C for 15 sec, 60°C for 1 min. Reactions included 2x Universal PCR Master Mix (Applied Biosystems), Taqman gene expression primers or Taqman miRNA primers, and corresponding cDNA according to the manufacturer’s protocol. Each plate was run with a negative control (no DNA template) and all assays were performed in quadruplicate. Gene expression was compared to an endogenous, internal control (RPLPO for mRNA or RNU6B for miRNA) using the $2^{-\Delta\Delta C_T}$ method [296]. C_T values were determined using the SDS 2.4 software (Applied Biosystems) and amplification C_T values higher than 36 cycles were designated as below the limit of detection (B.L.D.). Samples lacking any amplification curves were also designated B.L.D.

Construction of Reporter Plasmids. The luciferase report plasmids used in this study were constructed by inserting the common UGT1A 3’ UTR into the Xbal restriction site located downstream of the luciferase reporter gene in the pGL3-Promoter vector. Briefly, primers (sense, 5’-GCTATCTAGAAAGTGCGGGAGAATAAGGTAA-3’; and
antisense, 5’-GCTATCTAGAGAAACTTGCCCAGCACTTCATAGCT-3’) modified with the XbaI digestion site at both ends were used to amplify the UGT1A subfamily 3’ UTR (corresponding to nucleotides +1618 to +2301 of the human UGT1A1 mRNA, and 683 nucleotides in length) using genomic DNA isolated from the HEK293 cell line. The PCR-amplified region was cloned into the XbaI restriction site of the pGL3-promoter vector using standard protocols. The pGL3-seed deletion plasmid was created by performing SDM using the QuikChange II Site Directed Mutagenesis Kit (Agilent, Santa Clara, CA) and the sense and antisense primers ‘5’-TCATTTTATTCTTATTAAGGAAATAC-TTTAAATTAATCGCCCCAGAGTGCTT-3’ and 5’-AAGCACTCTGGGGCTGATTGATTA-ATTAAAAGTATTTTCCTAATAAGAATAAAATGA-3’, respectively. Nucleotide sequences of all plasmids used in this study were confirmed by DNA sequencing analysis performed at the Pennsylvania State University Nucleic Acid Facility (State College, PA).

**Luciferase Assays.** The pGL3-Promoter vector cloned with the UGT1A 3’ UTR (termed ‘pGL3-UGT1A3’UTR’) was co-transfected with the pRL-TK renilla control vector into HEK293 cells. The day before transfection, HEK293 cells were seeded onto 24-well plates at a density of approximately 50,000 cells/well, and after 24 h, 380 ng of pGL3-UGT1A3’UTR plasmid and 20 ng pRL-TK plasmid were co-transfected together with various concentrations of either scrambled miRNA control or miRNA mimic using Lipofectamine 2000 transfection reagent. HEK293 cells were harvested 24 h after transfection using passive lysis buffer, and luciferase activity was measured with a luminometor (Bio-tek Synergy HT, Winooski, VT) using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI). Renilla luciferase served as the internal control.

**Glucuronidation Assays.** Preparation of cell homogenates for glucuronidation activity assays was performed as previously described [297] after cells were transfected with 50 uL Lipofectamine 2000 in 10 cm dishes for 48 h according to manufacturer’s
protocol. HuH-7 or HepG2 cells were transfected with either 50 nM scrambled miRNA controls, 50 nM miR-491-3p mimic, or 50 nM miR-491-3p inhibitor. Collected cell pellets were freeze-thawed 3x in liquid nitrogen and subjugated to 3x 10-sec pulses using a hand-held Bio-Vortexer (Biospec, Bartlesville, OK) prior to storing of 50 µL aliquots at -80°C. Protein concentrations within the cell homogenates were quantified using the BCA Protein Assay Kit (Pierce Chemical, Rockford, IL) and measured using an Appliskan Luminometer and SkanIT Software v2.3 (Thermo Scientific).

The glucuronidation assays using homogenates from HuH-7 and HepG2 cells were performed essentially as described previously [141, 294]. HuH-7 and HepG2 cell homogenate (100-300 µg) was incubated with 50 µM raloxifene for 90 min, or 500 µM epirubicin hydrochloride for 60 min. Raloxifene glucuronidation reactions were terminated by the addition of 25 µL cold 100% acetonitrile and centrifuged before the drying of the supernatant in a speedvac and reconstitution in 40 µL of 1:1 water:acetonitrile; epirubicin reactions were terminated in the presence of 25 µL of cold 100% acetonitrile, centrifuged for 20 min at 13,000 g and supernatant was collected for analysis on UPLC/MS/MS.

Stock solutions of raloxifene, ral-6-Gluc, and ral-4′-Gluc and their deuterated internal standards were prepared in DMSO. Raloxifene, ral-6-Gluc, and ral-4′-Gluc were combined into a standard stock solution with final concentrations of 50 µg/mL, 100 µg/mL, and 100 µg/mL, respectively. Deuterated internal standards were combined in DMSO to make stock concentrations of 50 µg/mL for raloxifene-d4, ral-6-Gluc-d4, and ral-4′-Gluc-d4. The combined standard solution was then serially diluted in 1:1 water:acetonitrile to make standard working solutions from 50 ng/mL-50 µg/mL for raloxifene, and 1.25 ng/mL-1.28 µg/mL for ral-6-Gluc and ral-4′-Gluc. Standard calibration samples were prepared daily by spiking 1.0 µL of the combined internal
standard stock solution into 99 µL each of the serial-diluted working solutions (500 ng/mL final concentrations for each deuterated internal standard). All solutions were kept at -20°C before use.

Calibration standards as well as formed raloxifene and epirubicin glucuronides were analyzed using a Waters ACQUITY ultra-pressure liquid chromatography-UV detector (UPLC/MS/MS) system (Milford, MA) with a 1.7 µ ACQUITY UPLC BEH C18 analytical column (2.1 mm x 50 mm, Waters, Ireland) in series with a 0.2 µm Waters assay frit filter (2.1 mm). Raloxifene gradient elution conditions were performed as previously described [294]. Epirubicin gradient elution conditions were performed using a flow rate of 0.5 mL/min, starting with 95% Buffer A (0.1% formic acid in water) and 5% Buffer B (0.1% formic acid in acetonitrile) for 30 sec, a subsequent linear gradient to 50% Buffer B over 2.5 min, and then 100% Buffer B maintained over the next 2 min. Raloxifene and epirubicin glucuronides were confirmed by their sensitivity to the treatment of β-glucuronidase.

Quantification of raloxifene, ral-6-gluc, and ral-4'-gluc was performed as previously described [294]. Characterization of epirubicin and epirubicin-glucuronide was conducted using MRM of the transitions m/z 544 to 130 for epirubicin and m/z 720 to 113 for epirubicin-glucuronide. Since an epirubicin-glucuronide standard is not commercially available, quantification was based on the ratio of epirubicin-glucuronide versus total epirubicin after calculating the area under the curve for the epirubicin-glucuronide and epirubicin peaks using a known amount of epirubicin (500 µM) in each reaction as a reference. Data was quantified using the MassLynx™ NT 4.1 software with QuanLynx™ program (Waters). All experiments were performed in triplicate in independent assays.
**Western Blot Analysis.** UGT2B7 protein levels were determined via immunoblotting. HuH-7 and HepG2 protein homogenate was adjusted to equal volumes of loading buffer and heated at 90°C for 10 min. Samples were run at 90 V on a 10% acrylamide gel and transferred to a polyvinylidene difluoride (PVDF) membrane for 2 h at 33 V. PVDF membranes were blocked in 5% milk in TBS with Tween-20 (TBS-T) for 1 h, probed with UGT2B7 or calnexin primary antibody (1:1000 dilution for each) overnight at 4°C, washed 3x in TBS-T, followed by goat anti-rabbit secondary antibody (1:5000 dilution). Protein bands were visualized using the Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare, Pittsburgh, PA) and Hyoblot CL autoradiography film (Denville Scientific, Metuchen, NJ). Calnexin expression levels served as a loading control.

**Statistical Analysis.** Statistical analysis was performed using Graphpad Prism 5 software (La Jolla, CA). For studies involving miR-491-3p miRNA mimic or miR-491-3p inhibitors, the Student’s t-test (two-tailed) was used to compare experimental groups with scrambled miRNA controls. For statistical analysis of expression levels of genes labeled B.L.D., a Ct value of 40 was assigned to generate the necessary $2^{\Delta\Delta C_T}$ values needed for statistical comparison. For the analysis of UGT1A1, 1A3, 1A6, and 1A9 vs. miR-491-3p mRNA expression in human liver specimens, a one-tailed Spearman correlation was used. In this analysis, for those specimens with no miR-491-3p expression, they were assigned a Ct value half that of the lowest miR-491-3p expressing liver specimen. For comparison of UGT1A1, 1A3, 1A6, and 1A9 mRNA levels in miR-491-3p-expressing vs. non-expressing human liver specimens, the one-tailed Mann Whitney t-test was used. $P$-values <0.05 were considered statistically significant.
Results

In silico Analysis of miRNA Binding to UGT1A mRNA. The human UGT1A gene locus consists of nine alternative first exons spliced to common exons 2 to 5, with all nine UGT1A enzymes sharing the same 3' UTR (Fig. 2.1A). The mRNA corresponding to the UGT1A 3'UTR was analyzed by the miRanda (v3.0) and TargetScan algorithms to identify miRNAs that could potentially target UGT1A mRNAs [222, 298, 299]. miRanda is an excellent model algorithm due to its emphasis on stringent base-pairing of the seed region supplemented by downstream binding near the 3' end of the miRNA gene [220]. TargetScan is a similarly effective model algorithm due to its emphasis on contributions from binding position, 3' pairing, and miRNA target site abundance [295]. miRNA 491-3p (miR-491-3p) was the highest scoring prediction by miRanda to bind the UGT1A common 3' UTR, followed by miR-148a and miR-136. TargetScan also identified these three miRNAs as strong candidates to bind the UGT1A 3' UTR, with miR-491-3p having the highest overall total context+ score of the three miRNAs. miR-491-3p targets the common UGT1A 3' UTR at a position 92 to 114 base pairs (bp) downstream of the UGT1A stop codon [termed the miR-491-3p miRNA recognition element (MRE)] (Fig. 2.1B). This predicted MRE contains the canonical 8-nucleotide miRNA "seed" sequence of perfect complementation between miR-491-3p and the UGT1A 3' UTR (Fig. 2.1B, bottom, underlined). The MREs for miRs 148a and miR-136 target regions corresponding to bp 512-533 and 610-633, respectively, downstream of the UGT1A stop codon. Like that observed for miR-491-3p, both contain a 8-nucleotide “seed” sequence of perfect complementation with the UGT1A 3' UTR (Fig. 2.1B, bottom, underlined).

Effect of miR-491-3p on UGT1A 3' UTR-Mediated Luciferase Activity. To determine whether any of the predicted miRNAs described above (miR-491-3p, miR-
148a, and miR-136) could potentially affect UGT1A expression *in vitro*, the UGT1A 3’ UTR was cloned into the luciferase pGL3-promoter vector immediately 3’ of the luciferase open reading frame (Fig. 2.1C). The pGL3-UGT1A3’UTR plasmid was transiently co-transfected into HEK293 cells with scrambled negative control or miR-491-3p, miR-148a, or miR-136 miRNA mimics together with the *renilla* luciferase plasmid (serving as a transfection control). HEK293 cells are an ideal cell line for studying heterologous expression of UGT1A enzymes since they do not exhibit endogenous UGT1A expression [97]. Compared with the scrambled miRNA-transfected control, luciferase activity was significantly repressed at both 1 nM (*P*<0.05) and 2 nM (*P*<0.05) miR-491-3p mimic concentrations, corresponding to a 43 and 48% loss, respectively, of luciferase activity (Fig. 2.1D). To confirm that this inhibition was due to miRNA-491-3p binding to the UGT1A 3’ UTR miR-491-3p MRE, a 4-bp deletion mutation was created within the MRE seed sequence within the pGL3-UGT1A3’UTR (termed ‘pGL3-UGT1A3’UTR/SeedDel; Fig. 2.1C). While no significant difference in luciferase activity was observed between the seed deletion mutant and the negative scrambled miRNA control, a significant (*P*<0.05) 1.5-fold increase in luciferase expression was observed between cells over-expressing the wild-type pGL3-UGT1A3’UTR versus cells over-expressing the mutant pGL3-UGT1A3’UTR/SeedDel plasmids using 2 nM miR-491-3p mimic (Fig. 2.1D). No significant alteration in luciferase activity was observed in co-transfections with either the miR-148a or miR-136 mimics (data not shown).
Figure 2.1. The UGT1A gene locus and potential miR-491-3p miRNA recognition elements (MRE) within the common UGT1A 3' UTR. A, The 9 alternative UGT1A first exons code for unique enzymes. UGTs 1A1, 1A6, and 1A9 (dark boxes) are further highlighted in panel B. B, Alignment of three representative UGT1A mRNAs (1A1, 1A6, and 1A9) to highlight the common 3' UTR region. The miR-491-3p MRE begins +92 nucleotides from the stop codon for all UGT enzymes and the region spans 24 nucleotides in length. The miR-148a (spanning 22 nucleotides) and miR-136 (spanning 24 nucleotides) MREs begin at +512 and +610 nucleotides, respectively, from the stop codon. The predicted hybridization structure between the UGT1A 3' UTR and the miR-491-3p, miR-148a, and miR-136 MREs are enlarged at the bottom. Highlighted above the black bars are the MRE 'seed' sequences. C, Schematic of the pGL3-promoter luciferase vector with the SV40 promoter and the SV40 poly-A tail, including the XbaI 3' UTR restriction digest cloning site. The pGL3-UGT1A 3'UTR luciferase vector contains the first 683 nucleotides of the UGT1A family 3' UTR, encompassing the wild-type miR-491-3p MRE. The pGL3-UGT1A 3'UTR/Seed Deletion vector contains all sequences of the pGL3-UGT1A 3' UTR vector except for a four nucleotide deletion within the miR-491-3p MRE 'seed sequence'. X's indicate deleted base positions. D, Luciferase activity of UGT1A 3' UTR luciferase reporter vectors co-transfected with miR-491-3p mimic or scrambled miRNA control in HEK293 cells. Columns represent mean ± S.E. of three independent experiments. * P<0.05.
miR-491-3p Expression in UGT-Expressing Tissues and Cell Lines. To determine the levels of miR-491-3p expression in cells with known UGT expression, total RNA was screened from a variety of normal human tissues and cell lines by quantitative real-time (qRT)-PCR (Fig. 2.2A). miR-491-3p expression was highest in colon (25.6 ± 5.9) followed by trachea (12.0 ± 4.0) > lung (6.5 ± 0.81) > breast (3.0 ± 1.1) > liver (1.0 ± 0.11; set as reference; Fig. 2.2A). Of the 39 livers analyzed, 12 exhibited no expression of miR-491-3p; of the remaining samples, up to 4.0-fold differences in miRNA-491-3p expression were observed for 27 of the samples (a 7.8-fold difference was observed for the remaining sample; Fig. 2.2B). High miR-491-3p expression was also observed in Hep3B, HuH-7, HepG2, and SK-HEP-1 cells, exhibiting between 6.2 to 15.0-fold higher levels of miRNA-491-3p expression as compared with the lowest expressing individual liver specimen (Fig. 2.2B). Low levels of miR-491-3p expression were observed in the Caco-2, MCF-7, HEK293 and LNCaP cell lines; no expression was detected in jejunum, pancreas, kidney, larynx, and endometrium, or in the A-549 lung carcinoma cell line (data not shown).

UGT1A Expression in miR-491-3p-Expressing Cell Lines. To identify a hepatic cell culture model that could be used to monitor the effects of miR-491-3p on hepatic UGT1A expression, UGT1A expression was quantified by qRT-PCR in both the HuH-7 and HepG2 cell lines. UGT1A1 exhibited the highest level of expression of any hepatically-expressed UGT in both cell lines, with the levels in HuH-7 cells 40-fold higher than that observed in HepG2 cells (P<0.001; Fig. 2.2C). UGT1A1 exhibited >400- and 15-fold higher levels of expression than the next-most well-expressed UGT enzyme, UGT1A6, in HuH-7 and HepG2 cells, respectively. HuH-7 and HepG2 cells also expressed the hepatic UGTs 1A3, 1A4, and 1A9. Neither cell line expressed UGT1A5 (Fig. 2.2C).
Figure 2.2. Tissue and cell line expression of miR-491-3p and UGT1A family members. A, Relative expression levels of miR-491-3p in several UGT-expressing tissues. Expression levels were quantified using qRT-PCR and set relative to the lowest-expressing tissue (ie liver). B, Expression levels of mature miR-491-3p in individual normal livers and in liver cancer cell lines. Expression levels were quantified by qRT-PCR. miR-491-3p expression was adjusted to the RNU6B endogenous control gene in all tissues and cell lines. Expression is shown relative to the lowest overall miR-491-3p-expressing sample (ie liver #13; set to 1.0). C, Comparison of UGT1A mRNA levels between HepG2 and HuH-7 cell lines. mRNA expression was adjusted to the RPLPO endogenous control gene and shown relative to the highest expressing UGT enzyme in both cell lines (UGT1A1 in HuH-7 cells; set to 1.0). The Y-axis is broken into two segments, with a gap between 0.08 and 0.2 to better adjust for the high expression of UGT1A1 in HuH-7 cells. UGT1A1 mRNA levels were >40-fold higher in HuH-7 cells compared to HepG2. B.L.D., below the limit of detection. Columns represent the mean ± S.E. of three independent replicates. *** P<0.001.
Effects of miR-491-3p on UGT1A Expression and Glucuronidation Activity.

To identify which hepatic UGT1A enzyme is regulated by miR-491-3p, HuH-7 cells were transiently transfected with miR-491-3p mimic and UGT1A mRNA levels were assayed by qRT-PCR. miR-491-3p-transfected HuH-7 cells exhibited a significant (P<0.001) 48% decrease in UGT1A1 mRNA levels as compared to scrambled miRNA-transfected control cells (Fig. 2.3A). In addition, UGTs 1A3 (P<0.05) and 1A6 (P<0.05) mRNA were also decreased in the presence of miR-491-3p, with UGT transcript levels reduced by 35% and 27%, respectively. Serving as a negative control, the levels of UGT2B7 mRNA, with its own unique 3’ UTR different from the common UGT1A 3’ UTR and not predicted to bind miR-491-3p, did not significantly change after miR-491-3p transfection (Fig. 2.3A). The expression of UGTs 1A4 and 1A9 were not significantly repressed after transfection of miR-491-3p (data not shown).

Since UGT1A1 was the only UGT1A family member that was relatively well-expressed in the miRNA-491-3p-expressing HuH-7 and HepG2 cell lines and because UGT1A1 exhibited the largest variation in UGT1A expression after miR-491-3p over-expression, the role of miR-491-3p in regulating UGT1A1 expression and activity was examined specifically. Raloxifene, a chemotherapeutic agent used in the prevention of breast cancer in postmenopausal women [300], is exclusively metabolized by the UGT1A family of enzymes, with the major raloxifene glucuronide, ral-6-gluc, formed primarily by UGT1A1 in human liver, while both UGTs 1A1 and 1A9 contribute to hepatic ral-4’-gluc formation [116, 291, 301]. Since UGT1A1 mRNA expression levels are ~469-fold and ~15-fold higher than UGT1A9 mRNA levels in HuH-7 and HepG2 cells respectively (Fig. 2.2C), levels of raloxifene glucuronidation activity can serve as a specific marker for UGT1A1 activity and overall protein expression.
Given the relatively high expression of UGT1A1 in HuH-7 cells as compared to other UGT enzymes (see Fig. 2.2C), the levels of ral-6-gluc and ral-4'-gluc formation were monitored in homogenates from HuH-7 cells transiently transfected with miR-491-3p mimic versus scrambled miRNA-transfected controls as markers of UGT1A1 activity. In HuH-7 cell homogenates, both ral-6-gluc and ral-4'-gluc formation were significantly reduced by 40% ($P<0.01$) and 38%, ($P<0.01$), respectively, after transfection of miR-491-3p mimic as compared to scrambled miRNA-transfected controls (Fig. 2.3B). No change in UGT2B7 protein levels were observed as determined by Western blot analysis, and no difference in UGT2B7 glucuronidation activity was observed against epirubicin, a specific substrate for UGT2B7 [111, 112], in the presence of miR-491-3p mimic as compared to scrambled controls (Fig. 2.3C). Interestingly, no effect on UGT1A1 expression or activity was observed in HepG2 cells after transfection of miR-491-3p mimic as compared to scrambled miRNA-transfected controls (results not shown).
Figure 2.3. Quantitative analysis of UGT1A mRNA and enzymatic activity in HuH-7 cells. A, mRNA levels of UGT1A1, 1A3, 1A6, and 2B7 in miR-491-3p transfected HuH-7 cells. 50 nM of scrambled miRNA or miR-491-3p mimic were transiently transfected into HuH-7 cells for 48 h. mRNA was isolated and levels were compared to scrambled control (set at 1.0) for each gene. Genes were normalized to the RPLPO endogenous control gene using the 2^{-ΔΔct} method. UGT2B7 mRNA levels served as a negative control. Inset, a representative semi-quantitative electrophoresis gel of UGT1A1 mRNA. B, UGT1A1 glucuronidation activity against raloxifene in HuH-7 cells. UPLC/MS/MS was used to quantify the formation of ral-6-gluc (top) and ral-4'-gluc (bottom) from raloxifene substrate in miRNA-transfected HuH-7 homogenates. Ral-6-gluc formation and ral-4’gluc formation were significantly repressed ~40% and...
~38%, respectively. C, UGT2B7 protein levels and glucuronidation activity in the presence of 50 nM miR-491-3p miRNA mimic or scrambled control. Columns represent the mean ± S.E. of three independent experiments. * \( P<0.05; ** P<0.01; *** P<0.001.

To further examine the endogenous effect of miR-491-3p on the expression and downstream activity of UGT1A1, HuH-7 cells were transiently transfected with 100 nM miR-491-3p inhibitor for 48 h. Endogenous miR-491-3p levels were significantly \( P<0.01 \) reduced to below the level of detection in the presence of inhibitor as compared to scrambled miRNA inhibitor control (Fig. 2.4A). UGT1A1 mRNA levels were correspondingly significantly \( P<0.01 \) elevated (~80%) in the presence of inhibitor as compared to scrambled control (Fig. 2.4B). This increase in UGT1A1 mRNA levels corresponded with significant increases in ral-6-gluc (50%; \( P<0.05 \)) and ral-4'-gluc (22%; \( P<0.01 \)) formation for HuH-7 homogenates with repressed miR-491-3p levels as compared to scrambled control inhibitor-treated cell homogenates (Fig. 2.4C). No differences in (i) UGT2B7 mRNA levels (Fig. 2.4B), (ii) UGT2B7 protein levels (Fig. 2.4D), or (iii) glucuronidation activity against epirubicin (Fig. 2.4D) were observed in HuH-7 homogenates with repressed miR-491-3p levels.
Figure 2.4. The effect of miR-491-3p inhibitor on UGT1A1 expression and activity in HuH-7 cells. A, 100 nM of scrambled inhibitor control or miR-491-3p inhibitor were transiently transfected into HuH-7 cells for 48 h. Quantitative analysis of endogenous mature miR-491-3p expression levels were measured following miRNA isolation from transfected HuH-7 cells using qRT-PCR with RNU6B as an endogenous control. B, Quantification of UGT1A1 and UGT2B7 mRNA levels by qRT-PCR in HuH-7 cells in the presence of 100 nM miR-491-3p inhibitor or scrambled control. Quantification was performed with RPLPO serving as the endogenous expression control gene using the $2^{-\Delta\Delta Ct}$ method. Inset, a representative electrophoresis gel of the RT-PCR of UGT1A1 mRNA levels. C, UGT glucuronidation activity against raloxifene in HuH-7 cells with inhibited miR-491-3p. UPLC/MS/MS was used to quantify the formation of ral-6-gluc (top) and ral-4'-gluc (bottom) from raloxifene substrate in miR-491-3p-inhibited HuH-7 cell homogenates. D, UGT2B7 protein levels and glucuronidation activity in the presence of 100 nM miR-491-3p miRNA mimic or scrambled control. Columns represent the mean ± S.E. of three independent experiments. B.L.D., below the limit of detection. * $P<0.05$; ** $P<0.01$.

A similar pattern was observed for the same miRNA inhibitor experiments performed in HepG2 cells. Endogenous miR-491-3p levels in HepG2 cells were significantly ($P<0.001$) repressed in the presence of inhibitor when compared to
scrambled control (Fig. 2.5A). While UGT1A1 mRNA levels were not significantly altered in the presence of repressed miR-491-3p (Fig. 2.5B), miR-491-3p inhibition in HepG2 cells significantly increased ral-6-gluc and ral-4'-gluc formation by nearly 50% (P<0.05) and 34% (P<0.001), respectively (Fig. 2.5C). No difference in, (i) UGT2B7 mRNA levels (Fig. 2.5B), (ii) UGT2B7 protein levels (Fig. 2.5D), or (iii) glucuronidation activity against epirubicin (Fig. 2.5D) were observed in HepG2 homogenates with repressed miR-491-3p levels.
Figure 2.5. The effect of miR-491-3p inhibitor on UGT1A1 expression and activity in HepG2 cells. A, One hundred nM scrambled inhibitor control or miR-491-3p inhibitor were transiently transfected into HepG2 cells for 48 h. Quantitative analysis of endogenous mature miR-491-3p expression levels were measured following miRNA isolation from transfected HepG2 cells using qRT-PCR and the RNU6B endogenous control. B, Quantification of UGT1A1 and UGT2B7 mRNA levels by qRT-PCR in HuH-7 cells in the presence of 100 nM miR-491-3p inhibitor or scrambled inhibitor control. Quantification was performed with RPLPO serving as the endogenous expression control gene using the $2^{-\Delta\Delta CT}$ method. C, UGT glucuronidation activity against raloxifene in HepG2 cells with inhibited miR-491-3p. UPLC/MS/MS was used to quantify the formation of ral-6-gluc (top) and ral-4'-gluc (bottom) from raloxifene substrate in miR-491-3p- inhibited HepG2 cell homogenates. D, UGT2B7 protein levels and glucuronidation activity in the presence of 100 nM miR-491-3p miRNA inhibitor or scrambled inhibitor control. Columns represent the mean ± S.E. of three independent experiments. B.L.D., below the limit of detection. * $P<0.05$; *** $P<0.001$.

Hepatic UGT vs. miR-491-3p expression in human liver specimens. To assess whether there was any functional effect of miR-491-3p expression on UGT1A expression, the expression of UGTs 1A1, 1A3, 1A6, and 1A9 was examined by qRT-PCR in a panel of normal human liver specimens. A significant inverse correlation was observed between UGT1A3 mRNA and miR-491-3p expression levels in 38 liver specimens which expressed UGT1A3 (Spearman $r = -0.296$; $P<0.05$; Fig. 2.6A). A significantly ($P<0.05$) higher level of UGT1A3 expression was observed in miR-491-3p non-expressers as compared to miR-491-3p expressers in these specimens (Fig. 2.6B).

Similar to that observed for UGT1A3, a significant inverse correlation was also observed between UGT1A6 mRNA and miR-491-3p expression in the same liver specimens expressing UGT1A6 mRNA ($n=37$; Spearman $r = -0.487$; $P<0.01$; Fig. 2.6C). Also similar to that observed for UGT1A3, a significantly higher ($P<0.05$) level of UGT1A6 expression was observed in the miR-491-3p non-expressers as compared to expressers (Fig. 2.6D).

No correlation existed between UGT1A1 or UGT1A9 mRNA levels and miR-491-3p expression was observed in the same panel of liver specimens, and no difference in
UGT1A1 or UGT1A9 mRNA levels was observed between miR-491-3p-expressing versus miR-491-3p-non-expressing liver specimens (results not shown).

Figure 2.6. Expression of UGT1A3 and UGT1A6 mRNA versus miR-491-3p in human liver specimens. Levels of UGT1A3 (n=38; panels A and B) and UGT1A6 (n=37; panels C and D) mRNA vs. miR-491-3p were quantified in normal liver samples via qRT-PCR and normalized to RPLPO and RNU6B, respectively. For panels A and B, expression levels are shown relative to the lowest expressing sample (set to 1.0) for each gene, with those specimens with no miR-491-3p expression assigned a Ct value half that of the lowest miR-491-3p expressing liver specimen. A, UGT1A3 mRNA vs. miR-491-3p expression was examined using the Spearman ranking method. B, UGT1A3 mRNA expression levels were examined in liver specimens stratified by expression vs. no expression of miR-491-3p. C, UGT1A6 mRNA vs. miR-491-3p expression was examined using the Spearman ranking method. D, UGT1A6 mRNA expression levels were examined in liver specimens stratified by expression vs. no expression of miR-491-3p. Dots represent the mean ± S.E. of three independent replicates.
Discussion

This is the first study to examine the regulation of the UGT family of enzymes by miRNA. In silico prediction modeling using both miRanda and TargetScan indicated that miR-491-3p was the top candidate to bind to the UGT1A 3' UTR. While miR-491-3p and two other top candidate miRNAs (miR-148a and mir-136) were tested using the luciferase reporter assay, only miRNA-491-3p repressed luciferase activity in vitro. These data were further confirmed in the UGT1A1-expressing Huh-7 cell line, as miR-491-3p over-expression was shown to significantly inhibit UGT1A1 expression and glucuronidation activity, while inhibition of miR-491-3p was shown to significantly induce UGT1A1 expression and glucuronidation activity. In addition, deletion of the seed sequence within the 3' UTR miR-491-3p MRE of the UGT1A1 gene resulted in an elimination of these effects, suggesting specificity of action. These data were consistent with the fact that the mRNA expression of UGTs 1A3 and 1A6 were significantly and inversely correlated with miR-491-3p expression levels in a panel of normal human liver specimens examined in this study. Furthermore, a significant difference in UGT1A3 and UGT1A6 mRNA expression was observed in liver specimens stratified by miR-491-3p-expression versus miR-491-3p-non-expression. Together, these data strongly suggest that miR-491-3p plays a role in the regulation of UGT1A expression and activity and may contribute to the observed inter-individual variability of several UGT1A genes in humans.

While a significant increase in UGT1A1 expression and activity was observed in HepG2 cells after inhibition of miRNA-491-3p, no effect on UGT1A1 expression or activity was observed in HepG2 cells after transfection of miR-491-3p mimic as compared to scrambled miRNA-transfected controls. This contrasts with the differential effect on UGT1A1 expression and subsequent activity observed for both during miR-491-3p over-expression in HuH-7 cells. This difference is likely due to the lower
UGT1A1:miR-491-3p expression ratio observed between HepG2 and HuH-7 cells. This ratio difference is due mainly to the difference in expression of UGT1A1 in the two cell lines – UGT1A1 is expressed in HepG2 cells at levels well below (~40-fold less) that observed in Huh-7 cells, while there is only a five-fold difference in miR-491-3p expression levels between the two cell lines. In HepG2 cells where the UGT1A1:miR-491-3p expression ratio is low, it is likely that further expression of miR-491-3p has little effect on UGT1A1 expression and subsequent activity – it may already be at saturating levels, a situation that is less likely to be observed for HuH-7 cells where the UGT1A1:miR-491-3p expression ratio is significantly higher. When miR-491-3p is inhibited, a significant effect on UGT1A1 expression and subsequent activity would be expected in both cell lines and this was observed in the present study.

While significant inverse correlations were observed between both UGT1A3 and UGT1A6 mRNA levels and miR-491-3p expression in normal human liver specimens, no such association was observed for UGT1A1. Similarly, while a significant difference in UGT1A3 and UGT1A6 mRNA expression was observed in miR-491-3p expressing versus miR-491-3p non-expressing liver specimens in this study, no difference was observed for UGT1A1 expression. Additionally, no association was observed between UGT1A9 mRNA levels and miR-491-3p expression and this is consistent with the fact that UGT1A9 was not shown to be affected by miR-491-3p in vitro. Previous studies have shown that UGTs 1A3 and 1A6 are expressed at levels up to 10- and 3-fold lower, respectively, than UGT1A1 in normal human liver [96, 98]. Similar to that described above for cell lines, the miR-491-3p:UGT1A expression ratio is therefore higher for UGT1A1 than for UGT1A3 and UGT1A6 in human liver, and changes in hepatic miR-491-3p expression may only significantly affect the lower mRNA expression of hepatic UGTs like UGTs 1A3 and 1A6. These results, however, do not rule out the possibility that miR-491-3p regulates UGT1A1 protein expression in human liver. Future studies
are required to examine the impact of miR-491-3p expression on the enzymatic activity of UGT1A1 in liver specimens over-expressing miR-491-3p.

The levels of both UGT1A1 mRNA and enzyme activity were significantly altered in HuH-7 cells whereas in HepG2 cells only UGT1A1 enzyme activity was affected after over-expression of miR491-3p. This may reflect different cell-specific mechanisms of miR-491-3p regulation of UGT1A1 protein expression. It is likely that the alterations in UGT1A1 activity observed in HuH-7 cells are via a reduction in mRNA levels, presumably through mRNA degradation. In HepG2 cells, with no observable reduction in UGT1A1 mRNA levels after miR-491-3p over-expression, the primary mechanism of protein repression could be through translational inhibition or relocation and sequestering of UGT1A1 mRNA into P-bodies, where mRNA can be stored with RISC for later usage or until degraded [302]. Cell-specific miRNA regulation of protein expression is observed for other genes. For example, hepatocyte nuclear factor 4 alpha (HNF4α) is regulated in HepG2 cells by miR-24 and miR-34a. Mechanistically, miR-24 reduces HNF4α protein expression primarily via mRNA degradation; whereas miR-34a primarily regulates HNF4α expression via translational inhibition [252]. In HuH-7 cells, peroxisome proliferator-activated receptor alpha (PPARα) is primarily regulated by mir-21 and miR-27b via translation inhibition [251], while in HepG2 cells PPARα is regulated by miR-141 via reduced mRNA expression levels [265]. In both cases, however, the differential effect observed between cell lines was manifested by two different miRNA species. For UGT1A1, this differential cell line effect was linked to the same miRNA (miR-491-3p).

It is possible that the UGT1A 3’ UTR is not the sole region where miRNA may play a role in UGT1A regulation, as there are other genes which are regulated by miRNA predicted to bind within the 5’ untranslated or coding regions. For example, miR-24
regulates HNF4α protein expression primarily by binding within the mRNA coding region [252]. In the present studies, no effect on UGT1A glucuronidation activity was observed when levels of miR-125a-3p were altered in either HepG2 or Huh7 cells (results not shown). As miR-125a-3p was the miRNA predicted by miRanda to bind within non-3' UTR UGT1A sequences with the most efficacy, this suggests that miRNA regulation of UGT1A expression is primarily via binding to the UGT1A 3' UTR. It is also possible that the prediction programs used in this analysis (miRanda and TargetScan) failed to adequately predict the binding of other miRNA within these UGT1A sequences. miRNA-pull-down techniques using the UGT1A 3' UTR as a probe could be used to better assess this possibility [303].

The expression of hepatic UGT1A4 and UGT1A9 mRNA were not significantly regulated by miR-491-3p over-expression in the present studies. This may be due to UGT1A mRNA secondary structure potentially interfering with miR-491-3p and/or protein binding. We analyzed the mRNA secondary structures for UGTs 1A1, 1A3, 1A4, 1A6, and 1A9 using the RNA Fold software [304] to better assess this possibility (Fig. 2.7). While UGTs 1A3 and 1A4 had the most similar structure, each UGT1A isoform exhibited a unique secondary structure and this may impact miR-491-3p binding to each isoform (Fig. 2.7). mRNA secondary structure can play an important role in miRNA regulation and targeting, as previous reports have identified polymorphisms within mRNA 3' UTRs that alter miRNA regulation due to changes in mRNA secondary structure [305]. Further studies will be necessary to better assess this possibility.
Figure 2.7: UGT1A mRNA secondary structure as assessed by the RNA Fold program. RNA Fold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) was used to predict the mRNA secondary structures for each hepatic UGT1A isoform.

While miR-491-3p was shown to be expressed hepatically, it was also shown to be expressed in several other normal human tissues including colon, lung, trachea, and breast. In fact, the expression of miR-491-3p was higher in these tissues than was observed in liver. Therefore, in addition to playing a potential role in whole-body metabolism via hepatic mechanisms, miR-491-3p could be playing an important role in regulating local metabolism within extra-hepatic tissues. Further studies are required examining the effects of miR-491-3p on UGT1A expression for non-hepatic human tissues.

In summary, this is the first evidence demonstrating that miRNA regulate the phase II UGT metabolic enzymes. Results from both *in vitro* studies in cell lines and
human tissue phenotype data are presented demonstrating that UGTs including 1A3, 1A6, and possibly other UGT1A enzymes including UGT1A1 are regulated by miR-491-3p. Alterations in the expression of UGT1A enzymes by miR-491-3p may be an important mechanism controlling phase II metabolism and may play an important role in modulating drug response in humans.
Chapter 3:
miR-216b Regulation of UGT2B Family Enzymes
Abstract

The UDP-glucuronosyltransferase (UGT) 2B gene family of enzymes are important contributors to the detoxification of a variety of endogenous and exogenous compounds including hormones, bile acids, carcinogens, and chemotherapeutic agents. Identifying novel mechanisms governing their expression is important in understanding patient-specific response to drugs and cancer risk factors. We have identified miR-216b as a novel epigenetic regulator of UGT2B protein expression. Luciferase reporter assays and *in vitro* pull-down assays suggest the presence of a miR-216b functional binding motif within the 3’ UTRs of each of several UGT2B enzymes including UGT2B4, UGT2B7, and UGT2B10, all of which were identified using *in silico* prediction algorithms. miR-216b is expressed in normal human liver as well as extra-hepatic tissues including the pancreas, colon, and kidney. Over-expression of miR-216b mimics in HuH-7 cells significantly repressed UGT2B7 (*P*<0.001) and UGT2B10 (*P*<0.01) mRNA levels, and over-expression of miR-216b in Hep3B cells decreased UGT2B4 (*P*<0.001) and UGT2B10 (*P*<0.01) mRNA levels. UGT2B7 protein levels were significantly repressed in a dose-dependent manner in the presence of increasing miR-216b concentrations in HuH-7 cells and this was correlated with decreasing levels of epirubicin glucuronidation (*P*<0.01). Inhibition of endogenous miR-216b levels increased UGT2B7 mRNA levels in HuH-7 (*P*<0.05) and in Hep3B cells (*P*<0.01) and increased UGT2B7 epirubicin glucuronidation approximately 40% and 25%, respectively (*P*<0.05). This is the first evidence indicating that miRNA genes can regulate members of the UGT2B gene family and that miR-216b regulation of UGT2B proteins may play a role in the metabolism of UGT2B substrates in human liver.
Introduction

The UGT2B gene subfamily is involved in the metabolic clearance of numerous endogenous compounds including bile acids and steroid hormones, as well as exogenous agents including a variety of carcinogens and drugs [306]. UGT2B enzymes are hepatically expressed [98] and identifying novel mechanisms of their regulation may provide insight into UGT2B enzymatic activity and overall metabolic response in humans.

The UGT2B gene family consists of six isoforms: UGTs 2B4, 2B7, 2B10, 2B11, 2B15, 2B17, and 2B28 [73]. All UGT2B isoforms, excluding UGT2B28, are well-expressed in the liver as well as other extra-hepatic tissues involved in metabolism including stomach, colon, small intestine, breast, prostate, and kidney [97, 98, 307] (Jones et al, *DMD, In press*). Several UGT2B isoforms are highly-expressed in extra-hepatic tissues that are locally influenced by their substrates. For instance, UGT2B15 and UGT2B17 have high expression in steroidogenic tissues including the breast, prostrate, ovaries, uterus, and testes [97] and they metabolize several sex steroid hormones, including dihydrotestosterone, androsterone, and epiestradiol [113, 308]. The UGT2B genes contain six coding exons with unique, individual promoters and individual 3’ untranslated regions (UTRs) [73, 157] and appear to have evolutionarily arisen through successive rounds of gene duplication and replication slippage given that the UGT2B gene subfamily exhibits >70% DNA sequence homology with one another [77, 309].

The combined mRNA expression levels of the UGT2B gene family in the liver is close to 70% of all total UGT mRNA and there is a high degree of variability in the hepatic expression of all UGT2B isoforms [96]. Several factors contribute to the
regulation of UGT2B gene expression. For example, UGT2B4 [154] and UGT2B7 [155] can be spliced into alternative mRNA transcripts coding for dominant-negative inactive protein isoforms. Transcriptional regulators also play a role in the expression of UGT2B enzymes. HNF1α can bind to the promoters of all the UGT2B genes to regulate their expression, but the degree to which each gene is regulated by HNF1α is isoform-specific [157]. HNF1α mRNA levels are correlated with UGT2B7 mRNA levels in human liver [310] and HNF4α can bind to the promoters of UGT2B11 and UGT2B15 in hepatocytes. Few studies, however, have concentrated on post-transcriptional mechanisms governing UGT2B gene expression. For example, UGT2B7 protein levels can vary >4-fold [96, 311] and there is no correlation between UGT2B7 mRNA and UGT2B7 protein levels within individual human liver tissue samples [96, 124]. microRNA (miRNA) regulation of UGT2B enzymes may contribute to variability in UGT2B mRNA and protein expression.

miRNA are an endogenous, gene-encoded RNA species that are essential gene regulators in developmental biology [188, 196], including cancer initiation and progression [312], tissue signaling [195], and drug metabolism [246-248, 252, 253, 313]. miRNA are single-stranded RNA of approximately 22 nucleotides in length that, coupled with the RNA induced silencing complex (RISC), bind to target mRNA 3' UTRs and inhibit protein translation, thereby reducing protein expression [183, 184]. miRNA functionally can also decrease target mRNA levels, however this effect is dependent on the concentrations of both the target mRNA and miRNA genes [212-214]. Due to their imperfect binding, miRNA are highly promiscuous and can regulate several different genes at a given time. Reciprocally, some target mRNAs can be targeted by several miRNAs genes at the same time [217, 218]. miRNAs are also involved in the regulation of multiple members of a single pathway including upstream transcriptional regulators, and this complexity contributes to the miRNA role as a fine-tuner of protein expression.
For example, miR-27b regulates the expression of CYP3A4 in human pancreatic cancer cells and also regulates the expression of the vitamin D receptor (VDR) - an upstream transcriptional inducer of CYP3A4 [247].

Previous work in our laboratory identified miR-491-3p as a novel epigenetic miRNA regulator of UGT1A gene expression [313]. miR-491-3p expression levels were shown to be inversely correlated with UGT1A3 and UGT1A6 mRNA levels in normal human liver and contributed to observed interindividual variability of UGT1A gene expression [313]. There have been no previous studies investigated miRNA regulation of the UGT2B family members. In this study, we present evidence that miR-216b regulates the expression of several UGT2B genes.

**Materials and Methods**

**Chemicals and Reagents.** The pGL3-Promoter luciferase and pRL-TK renilla plasmids were from Promega (Madison, WI). All synthesized DNA oligos used for 3’ UTR amplification, site-directed mutagenesis (SDM), and PCR analysis were from Integrated DNA Technologies, Inc (Coralville, IA) or Life Technologies (Carlsbad, CA). Lipofectamine 2000 transfection reagent was from Life Technologies. miRVana miRNA miR-216b mimic (#MC12302), miRVana negative control mimic #1 (#4464058), miRVana miRNA miR-216b inhibitor (#12302), and negative control inhibitor #1 (#4464076) were purchased from Ambion (Austin, TX). The MISSION 3’ UTR Lenti GoClones for UGT2B15 (#HUTR04516) was purchased from SwitchGear Genomics (Menlo Park, CA). For use with the Lenti GoClones, MISSION miRNA mimics for miR-216b (HMI0382) and negative control #1 (HMC0002) were purchased from Sigma-Aldrich (St. Louis, MO). Uridine diphosphate glucuronic acid (UDPGA), alamethicin, and
bovine serum albumin (BSA) were from Sigma-Aldrich (St. Louis, MO). Epirubicin hydrochloride was purchased from Toronto Research Chemicals (Toronto, ON, Canada). Rabbit anti-UGT2B7 and anti-β actin antibodies were from BD Biosciences (San Jose, CA) and Cell Signaling Technologies (Danvers, MA), respectively. Rabbit anti-Ago2 antibody and mouse anti-HuR antibody were purchased from Millipore (Temecula, CA) and Santa Cruz Biotechnology (Dallas, TX), respectively. Goat anti-rabbit and anti-mouse secondary antibodies were from Thermo Scientific (Waltham, MA). Streptavidin M-280 Dynabeads (#11205D) were from Life Technologies. All other chemicals used were purchased from Fisher Scientific (Pittsburgh, PA) unless otherwise specified.

**Cell and Culture Conditions.** Human embryonic kidney cell line 293 (HEK293), human liver adenocarcinoma cell line SK-HEP-1, human hepatocellular carcinoma cell lines HepG2 and Hep3B, colon adenocarcinoma cell line Caco-2, human prostate carcinoma cell line LNCaP, human breast adenocarcinoma cell line MCF-7, and human lung carcinoma cell line A-549 were purchased from the American Type Culture Collection (ATCC, Manassas, VA). The human hepatocellular carcinoma cell line HuH-7 was a kind gift from Dr. Jianming Hu (Penn State Hershey College of Medicine, Hershey, PA). Hep3B, HEK293, A-549, and HuH-7 cells were cultured in DMEM (Gibco, Carlsbad, CA) supplemented with 10% FBS (Atlanta Biologicals, Lawrenceville, GA) and 1% penicillin/streptomycin (Gibco). SK-HEP-1 cells were cultured in RPMI 1640 (Gibco) supplemented with 10% FBS, 1% penicillin/streptomycin, and 1% NEAA. HepG2 and Caco-2 cells were cultured in DMEM supplemented with 10% FBS, 1% penicillin/streptomycin, and 1% non-essential amino acids (Lonza, Basel, Switzerland). MCF-7 cells were cultured in RPMI 1640 supplemented with 10% FBS and 1% penicillin/streptomycin. All cells were grown and maintained in 37°C at 95% air, 5% CO₂.
**Tissues and miRNA Isolation.** Pathologically normal colon and endometrium specimens (n=5 each) were obtained from the tissue bank at Pennsylvania State University College of Medicine while pathologically normal liver specimens and their matching total RNA were obtained from the H. Lee Moffitt Cancer Center Tissue Procurement Facility (n=5). All protocols involving the collection and analysis of tissue specimens from these tissue banks were approved by the Institutional Review Boards at their respective institutions and were in accordance with assurances filed with and approved by the United States Department of Health and Human Services. Normal jejunal tissue specimens were purchased from the Sun Health Research Institute (Sun City, AZ). All tissue samples were isolated and frozen at -70°C within 2 h post-surgery. Colon, liver, jejunal, endometrium and cell line total RNA was extracted using the RNeasy Mini Kit (Qiagen, Hilden, Germany). All total RNA samples from cell lines were subject to on-column DNAse digestion during RNA purification (Qiagen). Pooled breast total RNA was purchased from the Biochain Institute (Hayward, CA) while lung, pancreas, larynx, trachea, and kidney RNA was purchased from Clontech (Mountain View, CA) or Agilent (Santa Clara, CA). Small RNA (<200 nt) containing the miRNA fraction was isolated and purified from total RNA using the mirVana miRNA Isolation Kit (Ambion). All RNA concentrations were ascertained using a Nanodrop ND-1000 spectrophotometer (Thermo Scientific) and were eluted and stored in RNase-, DNase-free water in a -80°C freezer.

**miRNA Binding Site Predictions.** The 3' UTRs of the UGT2B enzymes were obtained from the NCBI database. miRNA binding site predictions were obtained using, (i) TargetScan [222], scored with the Total Context+ score as described in Garcia et. al [295], and (ii) miRanda, algorithm v3.0 [220] with the following parameters: Gap Open
Penalty, -8.00; Gap Extend, -2.00; Score Threshold, 50.00; Energy Threshold, -20.00 kcal/mol; Scaling Parameter, 4.00.

Real-Time Quantitative PCR. cDNAs were synthesized from total RNA using the SuperScript First-Strand Synthesis Kit (Invitrogen, Carlsbad, CA). miRNA cDNAs were synthesized using the Taqman MicroRNA Reverse Transcription Kit (Ambion). Taqman gene expression assays (Applied Biosystems, Carlsbad, CA) were used to amplify UGT2B4 (Hs_02383831_21), UGT2B7 (Hs_02556232_s1), UGT2B10 (Hs_02556282_s1), UGT2B11 (Hs_01894911_gH), UGT2B15 (Hs_00870076_s1), UGT2B17 (Hs_00854486_gH), and RPLPO (Hs99999902_m1). Taqman microRNA assays (Ambion) were used to amplify the expression of miR-216b (Cat. # 4427975, ID #002326) and RNU6B endogenous control (Cat #4427975, ID #001093) in all cell lines and tissue samples. PCR reactions were performed in 10 µL reactions in 384-well plates using an ABI 7900HT Sequence Detection System with incubations performed at 50°C for 2 min; 95°C for 10 min; and 40 cycles of 95°C for 15 sec, 60°C for 1 min. Reactions included 2x Universal PCR Master Mix (Applied Biosystems), Taqman gene expression primers or Taqman miRNA primers, and corresponding cDNA according to manufacturer’s protocol. Each plate was run with a negative control (no DNA template) and in quadruplicate. Gene expression was compared to an endogenous, internal control (RPLPO for mRNA or RNU6B for miRNA) using the 2^{ΔΔC_T} method [296]. C_T values were determined using the SDS 2.4 software (Applied Biosystems) and amplification C_T values higher than 35 cycles were designated as below the limit of detection (B.L.D.). Samples lacking any amplification curves were also designated B.L.D. PCR-analysis of miRNA from biotin pull-down assays were quantified using methods described in the Biotin Pulldown Assays methods section.
**Construction of Reporter Plasmids.** The luciferase reporter plasmids used in this study were constructed by inserting the UGT2B4, UGT2B7, and UGT2B10 3’ UTRs into the XbaI restriction site downstream of the luciferase reporter gene in the pGL3-Promoter vector. Briefly, primers listed in Table 3.1 modified with the XbaI digest site were used to amplify the 3’ UTRs using genomic DNA isolated from the HEK293 cell line. We used two sets of primers to amplify the 3’ UTR of UGT2B7 because of sequence similarities with other UGT2B genes. UGT2B7 “outer primers” amplified the UGT2B7 3’ UTR containing 5’ upstream and 3’ downstream intergenic DNA and the UGT2B7 “inner primers” amplified the UGT2B7 3’ UTR from the PCR product resulting from amplification using the “outer” primers. Each amplified region was cloned into the XbaI restriction site of the pGL3-promoter vector. The miR-216b seed deletion plasmids were created by performing site directed mutagenesis (SDM) using the QuikChange II Site Directed Mutagenesis Kit (Agilent, Santa Clara, CA) and the forward and reverse primers listed in Table 3.1. All plasmids with SNP mutations were created by SDM using primers whose sequences are also listed in Table 3.1. Nucleotide sequences of all plasmids used in this study were confirmed by DNA sequencing analysis performed at the Pennsylvania State University Nucleic Acids Core Facility (State College, PA).

**Luciferase Assay.** The pGL3-Promoter vector cloned with each 3’ UTR was co-transfected with the pRL-TK renilla control vector and miRVana miRNA mimics into HEK293 cells. The day before transfection, HEK293 cells were seeded onto 24-well plates at 50,000 cells/well, and after 24 h were co-transfected with 380 ng of pGL3 plasmid containing each UGT2B 3’ UTR and 20 ng pRL-TK plasmid together with either scrambled miRNA control or various concentrations of miRNA mimic using Lipofectamine 2000 transfection reagent according to manufacturer’s protocol.
<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>UGT2B4 3' UTR For.</td>
<td>5'-GCTATCTAGATTACGTGAGGCTGGGAAAGCTG-3</td>
</tr>
<tr>
<td>UGT2B4 3' UTR Rev.</td>
<td>5'-GCTATCTAGACACAATCTCGCATGAAATGATCC-3</td>
</tr>
<tr>
<td>Outer UGT2B7 3' UTR For.</td>
<td>5'-CCTTCCGGGTGCGACGCACGAC-3</td>
</tr>
<tr>
<td>Outer UGT2B7 3' UTR Rev.</td>
<td>5'-ATAAGGTCTCCAGGATTCAGGAGGGGGG-3</td>
</tr>
<tr>
<td>Inner UGT2B7 3' UTR For.</td>
<td>5'-GCTATCTAGAGGATGGATATATCTGAGATTGAAGCTGGAAAAAC-3</td>
</tr>
<tr>
<td>Inner UGT2B7 3' UTR Rev.</td>
<td>5'-GCTATCTAGATAAGGTGTTTATCTTATTTTTATTTATTTCCG-3</td>
</tr>
<tr>
<td>UGT2B10 3' UTR For.</td>
<td>5'-GCTATCTAGAGGATGGATATATCTGAGATTGAAGCTGGG-3</td>
</tr>
<tr>
<td>UGT2B10 3' UTR Rev.</td>
<td>5'-GCTATCTAGACCAAAGGTGTTTATCTTATTTTTATTTATTTCCG-3</td>
</tr>
<tr>
<td>UGT2B4 216b Del. For.</td>
<td>5'-TTATTACACAAAGAGAGCAGTTTGATACACATCTCTTTTTCTTCCTGTGAT-3</td>
</tr>
<tr>
<td>UGT2B4 216b Del. Rev.</td>
<td>5'-CCAAGAGAGAGAGAGCAGTTTGATACACATCTCTTTTTCTTCCTGTGAT-3</td>
</tr>
<tr>
<td>UGT2B7 216b Del. For.</td>
<td>5'-CCTTTGCAATAAAAAAATTTTTTCTCACCCACCGATGTT-3</td>
</tr>
<tr>
<td>UGT2B7 216b Del. Rev.</td>
<td>5'-TAAACATGGAACCTGGGTTGTAAGCTGGAAAAACAAATTTTTATTTGACCAAGG-3</td>
</tr>
<tr>
<td>UGT2B10 216b Del. For.</td>
<td>5'-CTAACCAGTCAGGTTATTTGTTTCTACCCACCGATGTT-3</td>
</tr>
<tr>
<td>UGT2B10 216b Del. Rev.</td>
<td>5'-CCATGAATGCGGTGTGAAGCTGAAAAAATTTTTCTGACAGAAGCATG-3</td>
</tr>
<tr>
<td>UGT2B10 SNP rs139538667 SDM For.</td>
<td>5'-CAAGAATAATTGTTTTCAGAGGTTCAGCCACGTAAATGGT-3</td>
</tr>
<tr>
<td>UGT2B10 SNP rs139538667 SDM Rev.</td>
<td>5'-CTAACCATTACACTGGAAGCTGAAAAAATTTTTCTGACAGAAGCATG-3</td>
</tr>
<tr>
<td>UGT2B4 Pull-down For.</td>
<td>5'-CCAAGCATTACAATACGACTCACTATAGGGAGATACGTCAGGCTGGAAAGCTG-3</td>
</tr>
<tr>
<td>UGT2B4 Pull-down Rev.</td>
<td>5'-CACAATGCTACGCGAAATGGATCC-3</td>
</tr>
<tr>
<td>UGT2B7 Pull-down For.</td>
<td>5'-CCAAGCTCTAATACGACTCACTATAGGGAGATACGTCAGGCTGGAAAGCTG-3</td>
</tr>
<tr>
<td>UGT2B7 Pull-down Rev.</td>
<td>5'-TAAGGCTTTCTTCTTCTTCTTCTTCGG-3</td>
</tr>
<tr>
<td>UGT2B10 Pull-down For.</td>
<td>5'-CCAAGCTTCTAATACGACTCACTATAGGGAGATACGTCAGGCTGGAAAGCTG-3</td>
</tr>
<tr>
<td>UGT2B10 Pull-down Rev.</td>
<td>5'-CTCAAGTCATGACGACCAATGCGCTCAGAGTGAC-3</td>
</tr>
<tr>
<td>miR-216b Pull-down Real-Time</td>
<td>5'-AAATCTCTGCAGGCAAATGGTGA-3</td>
</tr>
<tr>
<td>SNORD68 Pull-down Real-Time</td>
<td>5'-GATGACATTCTCCGGAATCG-3</td>
</tr>
</tbody>
</table>

**Table 3.1.** List of PCR primers used for UGT2B 3' UTR PCR amplification and cloning, luciferase mutational analysis, and UGT2B 3' UTR pull-down assays. For, Forward; Rev, Reverse; Del, Deletion; SDM, site-directed mutagenesis.
Transfected HEK293 cells were harvested 24 h post-transfection using passive lysis buffer and luciferase activity was measured with a luminometer (Bio-tek Synergy HT, Winooski, VT) using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI). Renilla luciferase served as the internal control.

Biotin Pulldown Assays. The 3’ UTRs of UGT2B4, UGT2B7, and UGT2B10 were amplified from the pGL3-Promoter luciferase vectors containing the respective 3’ UTRs using T7-labeled primers as listed in Table 3.1. The CCL2 control gene 3’ UTR and coding regions used primers as previously described [314]. Gel bands containing each cDNA were isolated and purified using the Qiagen Gel Extraction Kit (Venlo, Netherlands). T7-labeled cDNAs were used as a template to create biotin-labeled RNA using biotin-conjugated CTPs and the T7 MAXIscript IVT Kit (Ambion). Biotin-conjugated CTPs were used in a 1:4 ratio with unconjugated CTPs. Biotin-labeled RNA were subjected to DNase I digestion and purified using NucAway Spin Columns (Ambion). Biotin-labeled RNAs were mixed with HuH-7 cell lysate to confirm the binding of Ago2 or HuR (control) proteins to the UGT2B and CCL2 3’ UTRs and experiments were performed essentially as previously described [315, 316]. Isolation and purification of bound miRNA to the biotin-labeled 3’ UTR RNA transcripts was performed using a phenol-chloroform extraction of the streptavidin beads. Isolated miRNA were reverse transcribed using the qScript cDNA Synthesis Kit (Quanta Biosciences, Gaithersburg, MD). miRNA quantification was performed using PerfeCta SYBR Green Supermix with primers specific for miR-216b, and internal control SNORD68. Primer sequences are listed in Table 3.1. PCR and expression analysis was performed on a BioRad MyiQ2 Two Color Real-Time PCR Detection System and iCycler using the BioRad iQ5 program. Reaction conditions were as follows: an initial incubation at 95°C for 2 min, 40 cycles of
95°C for 10 sec followed by 60°C for 30 sec, and concluding with incubation at 10°C for 2 min.

**Glucuronidation Assays.** Preparation of cell homogenates for glucuronidation activity assays was performed as previously described [297] after cells were transfected with 50 µL Lipofectamine 2000 in 10 cm dishes for 48 h according to manufacturer’s protocol. HuH-7 or Hep3B cells were transfected with either scrambled miRNA controls, various concentrations of miR-216b mimic, or 100 nM miR-216b inhibitor. Collected cell pellets were freeze-thawed 3x in liquid nitrogen and subjugated to 3x 10-sec pulses using a hand-held Bio-Vortexer (Biospec, Bartlesville, OK) prior to storing of 50 µL aliquots at -80°C. Protein concentrations within the cell homogenates were quantified using the BCA Protein Assay Kit (Pierce Chemical, Rockford, IL) and measured using an Appliskan Luminometor and SkanIT Software v2.3 (Thermo Scientific).

The glucuronidation assays using homogenates from HuH-7 and Hep3B cells were performed essentially as described previously [141, 294]. HuH-7 and Hep3B cell homogenate (100-300 µg) was incubated with 500 µM epirubicin hydrochloride for 60 min. Epirubicin reactions were terminated in the presence of 1:1 cold 100% acetonitrile, centrifuged for 20 min at 13,000 g and supernatant was collected for analysis. Epirubicin glucuronides were analyzed using a Waters ACQUITY ultra-pressure liquid chromatography-UV detector (UPLC/MS/MS) system (Milford, MA) with a 1.7 µ ACQUITY UPLC BEH C18 analytical column (2.1 mm x 50 mm, Waters, Ireland) in series with a 0.2 µm Waters assay frit filter (2.1 mm). Epirubicin gradient elution conditions were performed using a flow rate of 0.5 mL/min, starting with 95% Buffer A (0.1% formic acid in water) and 5% Buffer B (0.1% formic acid in acetonitrile) for 30 sec, a subsequent linear gradient to 50% Buffer B over 2.5 min, and then 100% Buffer B maintained over the next 2 min. Epirubicin glucuronides were confirmed by their
sensitivity to the treatment of β-glucuronidase. Characterization of epirubicin and epirubicin-glucuronide was conducted using MRM of the transitions m/z 544 to 130 for epirubicin and m/z 720 to 113 for epirubicin-glucuronide. Since an epirubicin-glucuronide standard is not commercially available, quantification was based on the ratio of epirubicin-glucuronide versus total epirubicin after calculating the area under the curve for the epirubicin-glucuronide and epirubicin peaks using a known amount of epirubicin (500 µM) in each reaction as a reference. Data was quantified using the MassLynxTM NT 4.1 software with QuanLynx™ program (Waters). All experiments were performed in triplicate in independent assays.

**Western Blot Analysis.** UGT2B7, Ago2, HuR, and β-actin protein levels were determined via immunoblotting. Cellular homogenate or pull-down lysate was adjusted to equal volumes of loading buffer and heated at 90°C for 10 min. Samples were run at 90 V on a 12% acrylamide gel and transferred to a polyvinylidene difluoride (PVDF) membrane for 2 h at 33 V. PVDF membranes were blocked in 5% milk in TBS-T for 1 h, probed with primary antibody (1:1000 dilution) overnight at 4°C, washed 3x, followed by incubation with the appropriate secondary antibody (1:5000 dilution). Protein bands were visualized using the Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare, Pittsburgh, PA) and Hyoblot CL autoradiography film (Denville Scientific, Metuchen, NJ). UGT2B7 protein blots were stripped and re-probed for β-actin expression levels which served as a loading control.

**Statistical Analysis.** Statistical analysis was performed using Graphpad Prism 5 software (La Jolla, CA). For studies involving miR-216b miRNA mimic or miR-216b inhibitors, the Student’s t-test (two-tailed) was used to compare experimental groups with scrambled miRNA controls. For statistical analysis of expression levels of genes labeled B.L.D., a Ct value of 40 was assigned to generate the necessary $2^{\Delta \Delta C_T}$ values.
needed for statistical comparison. \( P \)-values <0.05 were considered statistically significant.

**Results**

**miR-216b is predicted to bind to the 3’ UTR of several UGT2B subfamily genes.** The miRanda and TargetScan miRNA prediction algorithms were used to scan the 3’ UTR mRNA sequences of UGT2B4 (NM_021139.2), UGT2B7 (NM_001074), UGT2B10 (NM_001075.4), UGT2B11 (NM_001073.1), UGT2B15 (NM_001076.2), UGT2B17 (NM_001077.2), and UGT2B28 (NM_053039.1) for potential miRNA binding sites. Only those miRNA candidates predicted by both programs were considered for further validation, and miR-216b, a conserved miRNA among vertebrates, was the only miRNA candidate to consistently score within the top three predicted miRNA candidates by both programs for several UGT2B genes including UGTs 2B4, 2B7, 2B10, and 2B28. The only prediction score that fell out of the ‘top three’ for miR-216b was for UGT2B4 predicted by miRanda, where miR-216b was 10\(^{th}\) among predicted miRNA candidates. miR-216b was not predicted to bind to the 3’ UTR of UGTs 2B11, 2B15, or UGT2B17 by either program (not identified within the top 25 miRNA candidates).

The UGT2B4 wild-type mRNA is 2,119 nucleotides (nts) long with its 3’ UTR extended from base pairs 1,634 to 2,119, and it contains one potential miR-216b miRNA recognition element (MRE) within the 3’ UTR. The miR-216b MRE spans 23 nts and begins at base pair 1,699 in the 3’ UTR of UGT2B4 (Fig 3.1A). A canonical 7-nt ‘seed’ sequence is also present in the predicted hybridization between miR-216b and the UGT2B4 mRNA sequence (Fig. 3.1A, underlined).
The UGT2B7 wild-type mRNA is 1,889 nts long and its 3' UTR begins at base pair 1,636. The UGT2B7 miR-216b MRE begins at base pair 1,786 and spans 23 nts (Fig. 3.1B). The 'seed' sequence of the predicted hybridization between the UGT2B7 mRNA and miR-216b is 8 nts long and with perfect complementation to the UGT2B7 3' UTR (Fig. 3.1B, underlined).

Similar to UGTs 2B4 and 2B7, miR-216b is the second highest overall scoring candidate predicted by both algorithms to bind to the 3' UTR of UGT2B10 (Fig. 3.1C). The wild-type UGT2B10 mRNA is 2,756 nts long and contains a single miR-216b MRE within its 3' UTR. The UGT2B10 miR-216b MRE is located beginning at base pair 1,750 and spans 23 nts with a canonical 8 nucleotide 'seed' sequence (Fig. 3.1C, underlined).

The UGT2B28 wild-type mRNA is 1,851 nts long and its 3' UTR begins at base pair 1,592. The UGT2B28 miR-216b MRE begins at base pair 1,794 and spans 23 nts (Fig. 3.1D). The 'seed' sequence of the predicted hybridization between the UGT2B28 mRNA and miR-216b is 8 nts long with perfect complementation to the UGT2B28 3' UTR (Fig. 3.1D, underlined). Few studies have focused on UGT2B28 mRNA expression and mRNA levels are generally expressed at very low levels in the liver [317]. UGT2B28 was, therefore, not investigated with further experimentation and in vitro validation as was done for UGTs 2B4, 2B7, and 2B10.
Figure 3.1. *In silico* prediction and binding of miR-216b to UGT2B isoforms. The miRNA prediction algorithms miRanda and TargetScan were used to identify miRNA binding candidates for all six UGT2B isoform mRNAs. miR-216b is predicted to bind to a 3' UTR MRE located within each of the UGT2B4 (A), UGT2B7 (B), UGT2B10 (C), and UGT2B28 (D) 3' UTR mRNA sequences (black bars). The canonical ‘seed’ sequences of the predicted hybridization between miR-216b and each 3' UTR are located underneath each gene and underlined in black.

**Tissue and cell line expression of miR-216b.** A miRNA gene can only regulate mRNA transcripts that are co-expressed in the same cell or tissue. To determine the expression levels of miR-216b in tissues and cell lines with known UGT2B expression, total RNA was screened from various normal human tissues and cell lines using qRT-PCR (Fig. 3.2). miR-216b expression was highest in the pancreas (2,897 ± 166) followed by kidney (4.60 ± 0.98) > liver (3.28 ± 1.62) > breast (1.62 ± 0.46) > trachea (1.06 ± 0.68) > colon (1.0 ± 0.20; set as reference). High levels of miR-216b were observed in the HuH-7 cell line (34.1 ± 8.82) and some expression was observed in the Hep3B cell line (1.44 ± 0.27) (Fig. 3.2). Very low levels of miR-216b expression were observed in normal lung and the A-549 and HEK293 cell lines; no expression was detected in normal
larynx, jejunum, or endometrium, or in the SK-HEP-1, HepG2, MCF-7 and Caco-2 cell lines (results not shown).

![Graph showing tissue and cell line expression of miR-216b](image)

**Figure 3.2. Tissue and cell line expression of miR-216b.** Expression levels of miR-216b were quantified using qRT-PCR and set relative to the lowest-expressing tissue (i.e. colon). miR-216b expression is adjusted to the RNU6B endogenous control gene using qRT-PCR in the same samples. The y-axis is broken into two segments, with a gap between 45 and 2000 to better adjust for the high expression levels quantified in the pancreas. Columns represent the mean ± S.E. of three independent replicates.

**In vitro validation of miR-216b target predictions.** *In vitro* validation of the *in silico* miR-216b target predictions was performed for the hepatically expressed UGTs 2B4, 2B7, and 2B10. The UGT2B4 3' UTR was cloned downstream of a luciferase reporter vector and transiently transfected into HEK293 cells with increasing concentrations of miR-216b miRNA mimics. We observed a significant dose-dependent repression of luciferase activity in response to increasing concentrations of miR-216b.
from 10 pM to 100 nM \( (P<0.001; \text{Fig. 3.3A, left}) \). Luciferase activity was reduced by up to 45% in the presence of 100 nM miR-216b mimic. In order to determine if the predicted UGT2B4 miR-216b MRE is the functional binding site for the interaction between miR-216b and the UGT2B4 3' UTR, a deletion of the GAGA bases in the UGT2B4 miR-216b MRE ‘seed’ sequence (Fig. 3.1A, underlined) within the UGT2B4 3' UTR luciferase vector was made. Luciferase repression by miR-216b was significantly lost when comparing the mutated seed sequence with the wild-type sequence in the presence of excess miR-216b mimic \( (P<0.001; \text{Fig. 3.3A, right}) \). These results indicate that miR-216b can functionally interact with the UGT2B4 3' UTR and that this interaction is sequence specific.

The UGT2B7 3' UTR was cloned downstream of the luciferase reporter vector, and similar to that observed for UGT2B4, there was a significant dose-dependent repression of luciferase activity in response to increasing concentrations of miR-216b after transfection into HEK293 cells \( (P<0.001; \text{Fig. 3.3B, left}) \). To ascertain if miR-216b is binding to the predicted MRE, the wild-type UGT2B7 3' UTR luciferase vector was mutated by deleting the GAGA bases within the ‘seed’ sequences of the miR-216b MRE. While luciferase expression for the wild-type UGT2B7 3' UTR-containing vector was significantly repressed in the presence of excess miR-216b \( (P<0.05) \), this effect was no longer observed for the mutated miR-216b binding site compared to scrambled control (Fig 3.3B, right). Together, these results indicate that miR-216b can functionally interact with the UGT2B7 3’ UTR and that the interactions are specific to the \textit{in silico} predicted binding sites.

Similar to both UGTs 2B4 and 2B7, the UGT2B10 3' UTR was cloned downstream of the luciferase reporter vector. A significant dose-dependent repression of luciferase occurred with increasing concentrations of miR-216b mimic \( (P<0.01; \text{Fig.} \)
Deletion of the GAGA bases within the ‘seed’ sequence of the miR-216b MRE binding site was created. While luciferase expression for the wild-type UGT2B10 3’ UTR-containing vector was significantly repressed in the presence of excess miR-216b (P < 0.01), this affect was alleviated by the mutated miR-216b binding site (Fig. 3.3C, right). Similar to UGT2B4 and UGT2B7, these results indicate that the interaction of miR-216b within the UGT2B10 3’ UTR is site-specific as predicted in silico.

A functional miR-216b binding site polymorphism in the UGT2B10 3’ UTR.

Analysis of the UGT2B10 3’ UTR mRNA sequence using public data provided by the SNP 500 Cancer Project [318] identified a novel single nucleotide polymorphism (SNP) located within the 3’ UTR of UGT2B10’s mRNA sequence and located within the functional miR-216b MRE. This adenine to guanine transition (rs139538767) is also located within the miR-216b MRE ‘seed’ sequence (Fig. 3.1C, bold, underlined) and has a minor allele frequency of approximately 1%. Using SDM, a polymorphic UGT2B10 3’ UTR luciferase reporter vector was created to ascertain whether this polymorphism confers an effect on the binding of miR-216b to the polymorphic transcript. As shown in the right panel of Figure 3.3C, the guanine variant within the miR-216b MRE ‘seed’ sequence of UGT2B10 eliminated the ability of excess miR-216b to repress luciferase activity as compared to the wild-type UGT2B10 3’ UTR luciferase vector in a similar manner to that observed for the ‘seed’ deletion control (P < 0.05). This mutational analysis indicates that the polymorphic allele has a functional effect on miR-216b repression of UGT2B10 3’ UTR-mediated luciferase activity.
Figure 3.3 UGT2B 3' UTR luciferase activity in the presence of miR-216b mimics. Luciferase activity of the UGT2B4 (A, left), UGT2B7 (B, left), and UGT2B10 (C, left) 3' UTR luciferase reporter vectors co-transfected with increasing concentrations of miR-216b mimic or scrambled miRNA control in HEK293 cells. Luciferase activity was also measured in the presence of 10 nM scrambled or miR-216b mimic after each UGT2B MRE was mutated within the miR-216b 'seed' sequence ('seed' deletion columns; right panels). Luciferase activity of the UGT2B10 3’ UTR variant SNP was also measured (C, right). Columns represent mean ± S.E. of at least three independent experiments and are normalized to scrambled control. *, P<0.05; **, P<0.01; ***, P<0.001.

RISC directly interacts with the 3’ UTRs of UGT2B4, UGT2B7, and UGT2B10. In order to better assess the potential of miR-216b to bind and interact with the UGT2B4, UGT2B7, and UGT2B10 3’ UTRs, an in vitro pull-down of RISC protein member Argonaute 2 (Ago2) and the associated miRNA fraction using biotin-labeled 3’ UTR probes was conducted (Fig. 3.4A). Briefly, biotin-labeled RNA transcripts corresponding to the 3’ UTRs for UGTs 2B4, 2B7, and 2B10 were mixed with cellular lysate prepared from the miR-216b-expressing HuH-7 cell line. Biotin-labeled transcripts were then purified using streptavidin magnetic beads and either, (i) RISC-associated protein components, or (ii) isolation and purification of specific miRNA binding to labeled transcripts, were assessed (Fig. 3.4A).

The 3’ UTRs of UGT2B4, UGT2B7, UGT2B10, and the control gene CCL2, along with the CCL2 coding region control, were amplified by PCR and gel purified. The purified cDNAs served as a template to create single-stranded, biotin-labeled RNAs for use in the in vitro pull-down (Fig. 3.4B). CCL2 was chosen as a negative control transcript because it is not predicted to have miR-216b binding sites within the 3’ UTR or coding region. Ago2, the main protein component of RISC, was found to be bound to the 3’ UTRs of UGT2B4, UGT2B7, UGT2B10, the CCL2 3’ UTR, and the CCL2 coding region (Fig. 3.4C). This indicates miRNA-guided RISC binding to each transcript and the
presence of Ago2 within this complex, typical of miRNA binding. In order to validate the specificity for true interaction between Ago2 and the transcript probes, each probe was also tested for the presence of RNA-binding protein HuR. HuR is an RNA-stabilizing protein which has already been identified to bind the 3' UTR of CCL2 mRNA but not the CCL2 coding region [314]. HuR was shown not to bind to the CCL2 coding region, but does bind to the CCL2 3' UTR as expected, serving as a positive control (Fig. 3.4C). This indicates the specificity of HuR binding to the 3' UTRs for each UGT2B isoform and the CCL2 control, and also indicates that Ago2 binding to the 3' UTR transcripts of the UGT2B genes are not an experimental artifact.

To demonstrate that miR-216b was specifically bound to biotinylated UGT2B 3' UTRs, PCR amplification of miR-216b from the miRNA fractions isolated from pull-down lysates was performed. miR-216b expression was enriched within the UGT2B4 3' UTR probe lysate by >10-fold compared to CCL2 3' UTR negative control, which is not predicted to bind to miR-216b ($P<0.05$; Fig. 3.4D). miR-216b was also enriched by >2.4-fold in the presence of the UGT2B7 3' UTR lysate ($P<0.01$) and by >2.7-fold in the presence of the UGT2B10 3' UTR ($P<0.01$) lysate as compared to the CCL2 3' UTR negative control (Fig. 3.4D). This indicates that in HuH-7 cells, miR-216b is capable of binding and interacting with the 3' UTRs of UGTs 2B4, 2B7, and 2B10 and further validates the in silico prediction analysis.
A

1. Biotin-labeled 3' UTR of Target mRNA
   - Cell lysate/incubate
2. mRNA binding to biotinylated 3' UTR
3. Streptavidin
   - Isolate complex with streptavidin beads
   - Wash and denature
4. Western for RISC, isolate mRNA for PCR

B

- Biotin-labeled RNA

C

- 100 kD
  - Ago2
- 37 kD
  - HuR

D

- HuH-7 miR-216b Enrichment
  - Relative Fold Enrichment
Figure 3.4. RISC and miR-216b association with UGT2B isoform 3’ UTRs. A, Experimental work-flow using biotinylated-3’ UTR probes of UGT2B4, UGT2B7, and UGT2B10 3’ UTRs and CCL2 3’ UTR and CCL2 coding region (C.R.) to pull down either Ago2 to assess RISC binding (C) or specific miR-216b enrichment (D). B, Representative gel electrophoresis of biotin-labeled RNA 3’ UTR probes. C, Representative western blot of Ago2 and HuR binding to each 3’ UTR probe and HuH-7 lysate positive control. D, qRT-PCR analysis of enrichment of miR-216b after pull-down and isolation of the miRNA fraction from the UGT2B4 (left), UGT2B7 (middle), and UGT2B10 (right) 3’ UTR probes. Each gene was normalized to the SNORD68 endogenous small RNA control and set relative to the CCL2 negative control 3’ UTR fraction for each data set. Columns represent the mean ± S.E. of two independent experiments. **, P<0.01.

Over-expression of miR-216b mimics in HuH-7 and Hep3B cells. The relative expression levels of UGT2B4, UGT2B7, and UGT2B10 was examined in the HuH-7 and Hep3B cell lines. HuH-7 and Hep3B cells were chosen because both cell lines endogenously express miR-216b (Fig. 3.2) and endogenous HuH-7 miR-216b has been shown to interact with UGT2B4, UGT2B7, and UGT2B10 3’ UTR probes (Figs. 3.3 and 3.4). UGT2B4 exhibited the highest level of mRNA expression of any UGT2B isoform in both cell lines, with Hep3B UGT2B4 mRNA expression the highest overall in both cell lines (reference set at 1.0; Fig. 3.5). In HuH-7 cells, UGT2B4 expression levels were 63% (±0.04) of Hep3B UGT2B4 mRNA levels; UGT2B7 expression ranged from 0.07 (HepG2) to 0.12 (HuH-7) that observed for UGT2B4 in HepG2 cells, while UGT2B10 expression was low in both cell lines.
Figure 3.5. UGT2B mRNA expression in HuH-7 and Hep3B cells.

UGT2B mRNA expression was quantified in HuH-7 and Hep3B cells using qRT-PCR. mRNA expression was determined using RPLPO as an internal, endogenous control and expressed relative to the highest overall expressing UGT2B mRNA transcript, UGT2B4 in Hep3B cells (set as 1.0 reference). The y-axis contains a gap between 0.15 and 0.50 to better adjust for high UGT2B4 expression levels. Columns represent mean ± S.E. three independent replicates.

In order to determine if miR-216b interacts with endogenous UGT2B mRNAs, 50 nM of miR-216b miRNA mimic was over-expressed in the HuH-7 and Hep3B cell lines. In HuH-7 cells, UGT2B4 mRNA was reduced by approximately 18% by miR-216b over-expression but this was not statistically significant (Fig. 3.6A). In Hep3B cells, UGT2B4 mRNA was significantly reduced by nearly 25% in the presence of miR-216b mimic as compared to scrambled control (P<0.001; Fig. 3.6B). UGT2B7 mRNA levels were similarly reduced in the two cell lines, but this was statistically significant in HuH-7 cells only (~18%; P<0.001; Fig. 3.6A). UGT2B10 mRNA levels were significantly reduced in both HuH-7 (~60%; P<0.01; Fig. 3.6A) and in Hep3B cells (~25%; P<0.05; Fig. 3.6B) in the presence of over-expressed miR-216b.
Figure 3.6. UGT2B mRNA levels in HuH-7 and Hep3B cells with 50 nM miR-216b. mRNA levels of UGT2B4, UGT2B7, and UGT2B10 in HuH-7 (A) and Hep3B (B) cells were quantified using qRT-PCR using the endogenous internal control gene RPLPO. mRNA levels for each gene in the presence of 50 nM miR-216b mimic are shown relative to scrambled control (set as 1.0). Columns represent mean ± S.E.M. of three independent experiments. *, *P* < 0.05; **, **P** < 0.01; ***, **P** < 0.001.

**Effect of inhibition of endogenous miR-216b levels in HuH-7 and Hep3B cells on UGT2B mRNA expression.** In order to further ascertain the endogenous effect of each miRNA on UGT2B mRNA levels, 100 nM of miRNA miR-216b inhibitor was used to transiently inhibit miR-216b function in both cell lines. Endogenous HuH-7 levels of miR-216b were repressed by nearly 100-fold (*P* < 0.05; Fig. 3.7A) with the miR-216b
inhibitor. Hep3B miR-216b levels were repressed by nearly 25-fold as compared to scrambled control \((P<0.01; \text{Fig. 3.7B})\).

The mRNA levels of each UGT2B isoform was quantified using qRT-PCR analysis to measure the effect on mRNA expression when endogenous miR-216b is inhibited. HuH-7 UGT2B4 mRNA levels increased by approximately 43% when miR-216b levels were inhibited \((P<0.001; \text{Fig. 3.7C})\). Significant increases in UGT2B4 mRNA levels were not observed when miR-216b is inhibited in Hep3B cells \(\text{Fig. 3.7D}\).

HuH-7 UGT2B7 mRNA levels were increased by approximately 21% when miR-216b was inhibited \((P<0.05; \text{Fig. 3.7C})\). Similar effects were seen in Hep3B cells with inhibited miR-216b for UGT2B7 mRNA, whose levels increased by approximately 25% \((P<0.01; \text{Fig. 3.7D})\). Interestingly, UGT2B10 mRNA levels were significantly repressed in both cell lines by 40% \((P<0.01)\) and 30% \((P<0.05)\) in HuH-7 and Hep3B cells, respectively, when miR-216b was inhibited \(\text{Fig. 3.7C} \& \text{D}\).
Figure 3.7. UGT2B mRNA levels in HuH-7 and Hep3B cells with inhibited miR-216b. A, Endogenous expression levels of miR-216b were quantified using qRT-PCR in the presence of 100 nM inhibitors for miR-216b in HuH-7 (A) and Hep3B (B) cells. Expression levels were quantified using the RNU6B endogenous internal control gene and each gene is shown relative to scrambled inhibitor control (set as 1.0). The mRNA expression levels of UGT2B4, UGT2B7, and UGT2B10 in HuH-7 (C) and Hep3B (D) cells were quantified using qRT-PCR using the endogenous internal control gene RPLPO. mRNA levels for each gene in the presence of 100 nM miR-216b inhibitor are shown relative to scrambled control (set as 1.0). Columns represent mean ± S.E.M. of three independent experiments. *, P<0.05; **, P<0.01; ***, P<0.001.

Effect of miR-216b over-expression on UGT2B7 protein and activity. To investigate the effect of miR-216b over-expression on UGT2B protein expression and enzymatic activity, UGT2B7 was chosen as the representative UGT2B isoform to be examined because there is a readily available commercial antibody for UGT2B7 protein detection that does not produce high background signal in endogenous cell lines [313]. HuH-7 cells were transiently transfected with 10 nM, 50 nM, or 100 nM miR-216b mimic and UGT2B7 protein levels and activity were assessed. A dose-dependent decrease in UGT2B7 protein expression occurred with increasing concentrations of miR-216b mimic, with UGT2B7 protein levels decreasing by 18, 35, and 47% in the presence of 10, 50, and 100 nM miR-216b, respectively, compared to scrambled control (Fig. 3.8A, top). A similar significant (P<0.01) dose-dependent decrease in epirubicin glucuronidation was observed in assays with 500 µM epirubicin, from 0.33 nmol·min⁻¹·µg⁻¹ (scrambled control) to 0.23 nmol·min⁻¹·µg⁻¹ (71%), 0.14 nmol·min⁻¹·µg⁻¹ (44%) and 0.14 nmol·min⁻¹·µg⁻¹ (31%) at 10, 50, and 100 nM miR-216b mimic (Fig. 3.8A, bottom), respectively. Similarly, we observed a 40% decrease in UGT2B7 protein expression in Hep3B cells transiently transfected with 50 nM miR-216b (Fig. 3.8B, top) and this corresponded to a significant 26% (P<0.05) decrease in epirubicin glucuronidation, from 0.62 nmol·min⁻¹·µg⁻¹ to 0.47 nmol·min⁻¹·µg⁻¹ (Fig. 3.8B, bottom). Epirubicin was used as a substrate marker for
UGT2B7 enzymatic activity because UGT2B7 is the only UGT isoform to glucuronidate epirubicin [111, 112].

**Effect of inhibition of endogenous miR-216b expression on UGT2B7 protein expression and activity.** To examine the impact of endogenous miR-216b expression on the regulation of UGT2B7 in HuH-7 and Hep3B cells, a transient transfection experiment with 100 nM miR-216b inhibitor was performed to significantly inhibit miR-216b endogenous expression in the two cell lines. An increase of nearly 60% in the levels of UGTB7 protein level was observed in miR-216b-inhibited HuH-7 cells (Fig. 3.8C, top) and this corresponded to a significant increase (~80%; P<0.05) in epirubicin glucuronidation in assays containing 500 µM epirubicin, from 0.57 nmol·min⁻¹·µg⁻¹ to 1.03 nmol·min⁻¹·µg⁻¹ (Fig. 3.8C, bottom). Similar results were seen in Hep3B cells, where UGT2B7 protein expression increased by 45% (Fig. 3.8D, top) and epirubicin glucuronidation levels significantly (P<0.05) rose by approximately 50% from 0.54 nmol·min⁻¹·µg⁻¹ to 0.82 nmol·min⁻¹·µg⁻¹ (Fig. 3.8D, bottom).
Figure 3.8. UGT2B7 protein expression and activity in the presence of miR-216b mimic or inhibitor. A, Top, Representative blot of HuH-7 UGT2B7 protein expression in the presence of 10 nM, 50 nM, and 100 nM miR-216b mimics. Protein bands adjusted to β-actin loading control and compared relative to 100 nM scrambled control. Bottom, UGT2B7 glucuronidation activity against epirubicin in HuH-7 cells. UPLC/MS/MS was used to quantify the formation of epirubicin-glucuronide in mimic-transfected cell homogenates. B, Top, Representative blot of Hep3B UGT2B7 protein expression in the presence of 50 nM miR-216b mimics. Bottom, UGT2B7 glucuronidation activity against epirubicin in Hep3B cells with 50 nM miR-216b mimic. C, Top, Representative blot of UGT2B7 protein expression in HuH-7 cells transfected with 100 nM miR-216b inhibitor. Bottom, Representative blot of UGT2B7 glucuronidation activity against epirubicin in HuH-7 cells transfected with miR-216b inhibitor. D, Top, Representative blot of UGT2B7 protein expression in Hep3B cells transfected with 100 nM miR-216b inhibitor. Bottom, UGT2B7 glucuronidation activity against epirubicin in Hep3B cells transfected with miR-216b inhibitor. Columns represent the mean ± S.E. of three independent experiments. * P<0.05; ** P<0.01.
Discussion

This is the first study identifying a miRNA regulator of the UGT2B metabolizing subfamily and experimental evidence is provided supporting the hypothesis that miR-216b regulates several UGT2B genes. The in silico miRNA prediction programs TargetScan [222] and miRanda [220] were used to identify miR-216b binding motifs within the 3’ UTRs of several UGT2B enzymes, including UGT2B4, UGT2B7, and UGT2B10. The positioning of these miR-216b MREs at 65, 150, and 138 base pairs downstream of the stop codon for UGTS 2B4, 2B7, and 2B10, respectively, is likely related to the sequence similarities within the UGT2B gene locus. It has been hypothesized that gene duplication events created the UGT2B gene family, with UGT2B4 being an ancestral gene [77]. Therefore, it is likely that miR-216b binding sites were duplicated and conserved during these events. The changes in distance of the miR-216b binding sites along the 3’ UTR from the stop codon of each isoform may reflect gene-specific elongation or other mutational events.

miR-216b expression levels were quantified in several UGT-expressing tissues including liver, with miR-216b exhibiting the highest overall expression in the pancreas. These data are consistent with data from the only other previous study investigating tissue expression of miR-216b, with Endo et al. observing very high pancreatic expression of miR-216b and also observing hepatic expression of miR-216b [319]. Low miR-216b expression levels were also observed in the prostate in these previous studies [319], a UGT-expressing hormonal tissue that was not tested in the present study. It is therefore possible that miR-216b contributes to both hepatic and extra-hepatic regulation of UGT2B genes in metabolic tissues where UGT2B enzymes and miR-216b are both co-expressed.
In vitro pull-down assays using biotin-labeled 3' UTRs of UGT2B4, UGT2B7, and UGT2B10 were used to assess the binding of RISC complexes. The presence of Ago2 on these transcripts using HuH-7 lysate indicated that miRNA genes expressed in HuH-7 cells can bind and interact with these UGT2B 3’ UTRs. These data are consistent with our over-expression studies in which over-expression of miR-216b mimics in HuH-7 and Hep3B cells can repress the expression of these three UGT2B enzymes. While Ago2 binding to the CCL2 3’ UTR and CCL2 coding region control transcripts was also observed in HuH-7 cell lysates, this suggests there are miRNA genes in HuH-7 cells with complementation to the sequences within the CCL2 3’ UTR and coding region that are guiding Ago2 and RISC to the biotinylated transcripts. HuR binding to the CCL2 3’ UTR but not the CCL2 coding region was consistent with that observed in previous studies [314] and confirmed the specificity of Ago2 binding to the UGT2B 3’ UTR transcripts.

Given that HuR is a RNA-stabilizing protein which binds to AU-rich elements within target mRNA 3’ UTRs in order to stabilize mRNA levels [314], the present study is the first to demonstrate HuR binding to the 3’ UTRs of UGT2B transcripts and also suggests a novel regulatory mechanism governing UGT2B mRNA expression which requires further exploration.

miR-216b is expressed in the liver and may therefore play a role in the regulation of hepatic UGT2B expression. In vitro over-expression of miR-216b in the HuH-7 and Hep3B hepatocellular carcinoma cell lines resulted in decreased expression of the three UGT2B enzymes (UGTs 2B4, 2B7, and 2B10) predicted by in silico models to bind miR-216b, except for UGT2B7 in Hep3B where the decrease was not statistically significant. Interestingly, similar decreases in expression were not observed for several other hepatic UGT2B enzymes after over-expression of miR-216b including UGTs 2B11, 2B15, and 2B17 (results not shown). For UGTs 2B4 and 2B7, these results were
corroborated by *in vitro* studies where endogenous miR-216b was knocked down using a miR-216b inhibitor, resulting in the up-regulation of both UGT2B enzymes in HuH-7 cells and UGT2B7 in Hep3B cells. A contrasting pattern was observed, however, for UGT2B10, where UGT2B10 expression was decreased in the presence of either miR-216b mimics or inhibitors. These results could be an artifact of the low UGT2B10 expression levels in both HuH-7 and Hep3B cells or could be due to the lack of specificity of the inhibitor for miR-216b in this experimental system. These hypotheses require further validation in future studies.

The effect of miR-216b over-expression or inhibition on UGT2B protein expression was examined for UGT2B7 in the present studies. UGT2B7 protein levels were significantly decreased after over-expression of miR-216b. Epirubicin glucuronidation is primarily mediated by UGT2B7 [111, 112] and miR-216b overexpression significantly reduced epirubicin glucuronide formation in both HuH-7 and Hep3B cells. In addition, knockdown of endogenous miR-216b using miRNA inhibitors increased UGT2B7 protein levels and activity in both cell lines. Together, these data indicate that endogenous levels of miR-216b play a role in UGT2B7 functionality and reflects a potential regulatory mechanism in normal human liver and potentially in other UGT2B7-expressing tissues where miR-216b is also expressed, e.g. the pancreas. Interestingly, a larger increase in UGT2B7 protein activity was observed in HuH-7 cells compared with Hep3B cells when endogenous miR-216b was inhibited. This may reflect the fact that endogenous miR-216b expression levels are >30-fold higher in HuH-7 cells compared to Hep3B cells and relief from miR-216b regulation would thus impact UGT2B7 expression to a greater extent in HuH-7 cells. The normal human liver tissue samples used in this study had approximately 2-fold higher endogenous miR-216b expression compared to Hep3B cells. We observed an approximately 25% increase in
UGT2B7 glucuronidation in Hep3B cells with inhibited miR-216b and this may better reflect the endogenous repression of UGT2B7 in normal liver. To assess miR-216b effects on UGT2B4 and UGT2B10 protein, studies similar to those performed for UGT2B7 using UGT2B4- and UGT2B10-specific substrates will be required.

Interestingly, there is a known SNP within the miR-216b MRE ‘seed’ sequence of the UGT2B10 3’ UTR (rs139538767). This polymorphism functions in a similar fashion to the UGT2B10 3’ UTR seed deletion control, resulting in an elimination of the negative regulation of UGT2B10 3’ UTR luciferase expression after miR-216b over-expression. These data suggest that this SNP may cause altered regulation of the polymorphic UGT2B10 allele. The minor allele frequency of this SNP is low with a prevalence of ~1% in the population, but the present study provides the first evidence of a UGT2B 3’ UTR SNP with a direct functional role in aberrant miRNA regulation.

In summary, we provide evidence of a functional miR-216b binding motif within the mRNA of several UGT2B isoforms. UGT2B7 protein levels and overall enzymatic activity are repressed in the presence of over-expressed miR-216b and induced after addition of miR-216b inhibitor and this may have functional consequences on UGT2B7 enzymatic activity in vivo. In addition, a functional SNP was identified in the UGT2B10 3’ UTR which may modify UGT2B10 regulation through a miR-216b-mediated pathway. Together, these data suggest that miR-216b expression may have functional consequences on several UGT2B genes and this regulatory mechanism may contribute to interindividual variability of UGT2B expression in humans.
Chapter 4:
Future Directions and Final Considerations
4.1 Future Directions

The research presented in this dissertation details the first investigations of miRNA regulation of the phase II drug-metabolizing UGT1A and UGT2B enzyme families. The UGT1A and UGT2B families are instrumental in the metabolism of many endogenous and exogenous compounds from the body and participate in patient response to drug regimens [158, 306]. The discovery that miRNA regulate UGT protein expression and enzymatic activity provides an additional layer of epigenetic regulation contributing to interindividual variability of the expression of UGT enzymes. This is important because miRNA regulation of UGT enzymes in the liver may impact clearance of carcinogens and influence drug metabolic rates and this may also contribute to how patient’s respond to clinical treatments.

4.1.1 miR-491-3p and Cancer

These studies have identified miR-491-3p as a regulator of UGT1A expression and activity in human liver cancer cells and may also reflect endogenous regulation in normal human liver. A significant reduction in UGT1A3 and UGT1A6 mRNA expression in normal human liver samples expressing miR-491-3p was observed, which may contribute to variable patient metabolism of UGT1A3 and UGT1A6 substrates. Variable expression of miR-491-3p in the human liver has not been documented before. miR-491-3p expression levels were also observed to be inversely correlated with UGT1A3 and UGT1A6 mRNA in human liver and may serve as a biomarker for UGT1A3 and UGT1A6 hepatic expression. Cancer patients who do have miR-491-3p expression in the liver may have a decreased capacity to metabolize chemotherapeutic agents and other drugs metabolized by these two UGTs and may be predisposed to toxicological issues when seeking treatment.
Interestingly, we observed higher expression of miR-491-3p in four human liver cancer cells as compared with individual normal human livers (Fig. 2.2). This may reflect a mechanism in which miR-491-3p is up-regulated during tumorigenesis and may contribute to altered regulation of UGT1A1 and other UGT1A isoforms in liver cancer cells. Previous reports have shown that UGT1A1 and other UGT1A proteins are down-regulated in cancers of the colon, bladder, breast, and liver [127, 320]. Matched normal and tumor liver tissue from 7 patients with hepatocellular carcinoma were obtained and expression levels of miR-491-3p and UGT1A1 mRNA were quantified to determine if there is an inverse correlation of the two genes between normal and tumor tissue. An inverse association between miR-491-3p and UGT1A1 mRNA was observed in these tissue samples (data not shown) but the data was not significant due to sample size. This preliminary data may potentially support the hypothesis that miR-491-3p and UGT1A1 expression levels are inversely correlated between matched normal and cancerous liver tissue, however, more samples are needed to continue this investigation. If a larger analysis were to show a significant inverse correlation between normal and matched tumor tissues, this would be the first evidence suggesting altered miRNA regulation in cancer has a significant impact on UGT enzyme expression. This could, for example, have an impact on localized chemotherapeutic metabolism within the liver in patients with liver cancers.

miR-491-3p may be a possible therapeutic target designed in tandem with other drugs for cancer patients. Toxicological issues with irinotecan-treatment have been well documented [56, 129, 133, 292] and are reflective of expression levels of irinotecan-metabolizing UGT1A1. Therefore, it would be interesting to determine if targeted knock-down of endogenous liver miR-491-3p expression levels would modestly increase UGT1A1 expression in the liver and help patients with UGT1A1-deficiencies better
metabolize SN-38, the active, yet toxic, metabolite of irinotecan. Already, studies have used anti-miRs to inhibit miRNA genes in tumor tissues. Intravenous delivery of anti-miR-10b in a mouse model system reduced breast cancer metastasis [321] and subcutaneous injection of anti-miR-33a in non-human primates lowered the risk of cardiovascular disease by increasing miR-33a target ABCA1 and raising plasma levels of HDL [322]. Both methods delivered specific miRNA inhibitors to reduce target miRNAs miR-10b and miR-33a and induce downstream protein expression changes.

miRNA have been found to inhabit exosomes and contribute to cellular signaling via exosomal-mediated transfer [195]. Recent experiments in mice using systemic delivery of let-7a and miR-34a mimics packaged within lipid-based particles reduced lung tumors in mice >60% [323, 324]. It would be interesting to package miRNA 491-3p antagonirs into liposomes or nanoparticles specifically targeting the liver to increase downstream UGT1A protein expression. This may be useful in patients who suffer from Type II Crigler-Najjar Syndrome who have low-levels of UGT1A1 due to a genetic mutation within the UGT1A1 gene [130]. Reciprocally, miR-491-3p mimics could be packaged into liposomes and delivered to the intestines to reduce UGT1A protein expression. Drugs such as raloxifene have very low bioavailability due to first-pass glucuronidation in the intestines from UGTs 1A1, 1A8, 1A9, and 1A10 [291]. Co-administration of miR-491-3p mimics within the intestinal tissues could temporarily reduce UGT1A isoform expression and increase raloxifene bioavailability. Preliminary studies in our laboratory have identified that UGT1A8 and UGT1A10 mRNA levels are significantly increased in Caco-2 cells when endogenous miR-491-3p was inhibited (Fig. 4.1). Protein expression analysis and glucuronidation assays using substrates specific for these isoforms are needed to ascertain the endogenous impact of miR-491-3p expression on UGT1A8 and UGT1A10 protein levels in tissues of the GI tract.
Additionally, over-expression analyses using miR-491-3p mimics in intestinal cell lines will help identify the effect on extra-hepatic UGT1A7 expression.

Figure 4.1. Caco-2 UGT1A8 and UGT1A10 mRNA levels. UGT1A8 and UGT1A10 mRNA levels were measured using qRT-PCR in Caco-2 cells transfected with 100 nM miR-491-3p inhibitor for 48 hours. Expression for each gene is set relative to scrambled control (set at 1.0) and quantified using the RPLPO internal, endogenous control. *, P<0.05; ***, P<0.001; student’s t-test. Columns represent mean ± S.E. of three independent experiments.

4.1.2 miR-216b and Cancer

miR-216b regulates the mRNA expression of several UGT2B gene members including UGT2B4, 2B7, and 2B10. We also observed that miR-216b significantly affects UGT2B7 protein and enzymatic expression in several liver cancer cell lines (Chapter 3). Very high expression of both genes was observed in the pancreas (Fig. 3.5), which replicates previous reports that miR-216b is highly expressed in human and rat pancreas [319]. To date, studies outside of our laboratory investigating tissue-specific expression of the UGT2B enzyme families have neglected to look within the pancreas [97, 98, 158, 325]. One study did identify transcriptional binding of HNF4α to the promoters of UGT2B11, 2B15, and 2B17 within pancreatic tissue but mRNA expression levels were
not investigated [162]. Studies in our laboratory have found low levels of expression of UGT2B7 in the pancreas and very low levels of UGT2B10 in the pancreas (Jones et al., In press, DMD). The pancreas is a metabolic and hormonal tissue and UGT activity in this tissue has not been examined previously.

Low UGT2B expression in the pancreas can have very important clinical implications for individuals who smoke. Tobacco smoking has been linked to pancreatic cancer incidence and recent studies have shown that nearly one-third of all cases of pancreatic cancer may be attributed to heavy smoking [326, 327]. This is due to localized exposure of pancreatic cells to TSNAs such as NNAL [326] and possibly nicotine itself, which can constitutively activate pancreatic nicotinic receptors and the downstream JAK2/STAT3 pathway, inducing pancreatic metastasis [328, 329]. UGT2B10, a target of miR-216b regulation, metabolizes several TSNAs including NNAL and nicotine [51, 110, 145, 146]. It is possible that elevated levels of miR-216b in the pancreas reduces UGT2B10 protein expression in the pancreas and affects the metabolism of tobacco carcinogens (e.g. NNAL). Low UGT2B10 activity in pancreatic cells could mean an increased length in time in which these carcinogens are present in this tissue. Studies using pancreatic cells to determine: 1) the endogenous effect of miR-216b regulation of UGT2B10, and 2) the localized metabolism of TSNAs and nicotine in pancreatic cells, should be performed.

It is also possible that the miR-216b:UGT2B10 regulatory network influences the systemic clearance of TSNAs and nicotine in the liver. These studies identified a 37-fold variability in the expression of miR-216b in the liver (data not shown) and this could play a significant role in interindividual response to UGT2B10-mediated metabolic glucuronidation of tobacco carcinogens in the liver of smokers. Higher levels of hepatic miR-216b may predispose smokers to cancer due to decreased metabolism of cigarette
smoke carcinogens and their elimination from the body. Additionally, higher levels of miR-216b, and reciprocally lower levels of UGT2B10, may impact nicotine clearance from the body via the liver. This will be important to investigate because lower levels of UGT2B10 in the liver may be associated with higher levels of nicotine, possibly predisposing individuals to nicotine addiction. However, low levels of UGT2B10 caused by miR-216b expression may have a different effect. Slow clearance of nicotine from the body may cause smokers to consume less nicotine. Studies investigating the UGT2B10^{D67Y} coding polymorphism have linked polymorphic UGT2B10 genotypes with reduced nicotine clearance as well as reduced nicotine intake [146, 330]. Variable miR-216b expression may have influence on UGT2B10 expression and therefor nicotine clearance. Correlation analyses of hepatic miR-216b expression levels and smoking behavior can better address this possibility.

4.1.3 Transcriptional Induction of UGT-Regulating miRNAs

The studies in this dissertation focused on the direct effect of miR-491-3p and miR-216b regulation of UGT1A and 2B protein expression, respectively. However miRNA genes may have an indirect effect on UGT expression via miRNA regulation of transcription factors known to regulate UGT enzymes (preliminary results in Appendix 1).

The transcriptional regulation of miR-491-3p and the miR-216b genes have not been explored here or in the literature. It has been shown that HNF1α and HNF4α can bind and regulate the expression of miRNAs miR-194 [331, 332] and miR-122 [333], respectively, in liver. Using the in silico program Transcription Element Search System (TESS) [334], the immediate and upstream promoters of miR-491-3p and miR-216b were analyzed for potential HNF and other UGT-regulating TF-binding sites. TESS analysis reports multiple binding sites for HNF family members 1α and 1β within the first
2,000 bases immediately 5’ of the miR-491-3p transcriptional start site (data not shown). miR-491-3p’s promoter also contains a predicted PPARα binding site and four potential Cdx-1 binding sites. Cdx-1 is a transcription factor regulating gene expression in tissues of the GI tract including intestine and colon [335]. Future studies should investigate the transcriptional regulation of miR-491-3p in the intestine and/or colon and any possible downstream effects on UGT1A expression as this could have an impact on first-pass metabolism (Fig. 4.2, top). Possible HNF1α and HNF1β regulation of miR-491-3p may also have implications for UGT1A expression in the liver (Fig. 4.2, top).

TESS reports conserved multiple binding sites for HNF1α and HNF3α in the promoter of miR-216b (data not shown). HNF1α regulates expression of UGT2B7 and UGT2B10 [157] and previous work in our laboratory has identified HNF3α as a transcriptional regulator of UGT2B10 expression in liver (Jones et al., unpublished). The TESS predictions require experimental validation but may have preliminarily identified a regulatory feedback loop governing miR-216b expression in the liver. Our results and these in silico predictions suggest miR-216b expression is an important global player in both the direct regulation of UGT2B enzymes but also important UGT-regulating HNF transcriptional factors in the liver (Fig. 4.2, bottom). miRNA regulatory networks and feedback loops are widespread in human tissues and cells [336] and so it will be important to identify any involved with the regulation of UGT enzymes or other drug metabolizing enzyme families.

miR-24 and miR-34a, and potentially miR-216b, regulate HNF4α (see Appendix 1). A feedback loop has been identified in hepatocellular oncogenesis in which over-expressed HNF4α induces miR-124 expression, thereby repressing levels of IL6R and downstream STAT3, an activator of miR-24 expression [287]. In liver cancers HNF4α protein levels are low and this feedback loop contributes to liver tumorigenesis [287].
Low levels of HNF4α in the liver and liver cancers may have an impact on patient response to UGT-mediated drug metabolism. Further investigation will be needed to identify the important of this network in UGT-mediated drug response in liver cancer by studying changes in miR-24, miR-34a, and potentially miR-216b expression during the liver tumorigenesis process and assaying for downstream changes in HNF-expression. It is also possible that miR-24, miR-34a, and miR-216b interact with and regulate expression of other UGT-regulating transcription factors, e.g. LXR, PXR, RXR, AhR, and CAR. If this is the case, that would place miR-24, miR-34a, and miR-216b within a regulatory cycle with possibly profound downstream implications in UGT gene expression.

Figure 4.2. Hypothesized miRNA-transcriptional regulatory networks involved in the regulation of the UGT1A family (Top) and the UGT1A and UGT2B families (Bottom).
4.1.4 Environmental Induction of UGT-Regulating miRNAs

Endogenous bile acids can activate the expression of UGT2B4 [176] and sex steroid hormones induce expression of UGT2B7, UGT2B15, and UGT2B17 [337]. Upstream regulation of miR-491-3p or miR-216b expression via bile acids or hormones has not been explored and may provide an additional epigenetic control governing UGT2B enzymatic function in response to fluctuated levels of endogenous hormones within the body. Additionally, carcinogens or other environmental exposures may have a similar effect. For example, F344 rats subjected to NNK exposure had increased levels of miR-126, a regulator of CYP2A3. CYP2A3 activates the pro-carcinogen NNK and promotes oncogenesis in rat lungs [338]. Another study evaluated miRNA gene expression changes within the bronchial epithelium in twenty current and never-smoking adults, finding 28 miRNAs to be both up-regulated and down-regulated in smokers compared to never smokers [339]. It is possible that carcinogen or chemotherapeutic exposure in the liver may cause fluctuations in UGT-regulating miRNA genes, affecting downstream expression of UGT enzymes and predisposing individuals to tobacco-related or hormone-related cancers. Environmentally-induced changes in miRNA expression may also affect UGT transcriptional regulators, which could have a profound effect on the expression of the entire family at one time. In vivo model systems and in vitro cell lines exposed to tobacco carcinogens or hormones will be needed to preliminarily assess these types of changes and their potential impact on UGT expression.

4.1.5 UGTs and the miRNA Sponge Effect

The UGT gene loci contain 9 individually transcribed pseudo genes; four within the UGT1A family and five within the UGT2B family [66, 73]. These genes are not
protein coding and no discernible function has been identified with these non-coding RNA products. There have been no studies specifically investigating their purpose. Given the sequence homology seen among all of the UGT genes, especially within subfamilies, it is possible that the UGT pseudogenes contain coding and 3' UTR sequence motifs highly similar to their related protein-coding counterparts. If this is the case, the UGT pseudogenes could be exemplifying a phenomena known as the miRNA 'sponge' effect. miRNA sponges are engineered vectors that express RNA sequences with multiple binding motifs and seed sequences which are tandem-aligned and targeting a specific miRNA of interest. The effect is that upon expression, this vector will act as a miRNA 'sponge', binding simultaneously with several copies of a specific miRNA via the numerous binding motifs along the vector and behaving similar to a miRNA inhibitor [340]. There have been two identified natural, endogenous miRNA sponges in humans. One sponge is a pseudogene for PTEN, called PTENP1 [341]. PTENP1 contains a mutated start codon and does not code for a protein, but it does contain a highly similar 3' UTR sequence to PTEN. The PTENP1 pseudogene regulates PTEN expression by soaking up endogenous miRNAs targeting PTEN. PTENP1 is selectively lost in several human cancer and loss of PTENP1 reduces expression of normal PTEN protein [342].

Recently, naturally occurring, circular non-coding RNA cISR-7 found in the human brain contains over 70 miRNA binding sites for let-7. cISR-7 soaks up let-7 miRNA in the brain and thereby increasing let-7 targets [343]. Given the sequence similarity of to the protein-coding UGTs genes, it is plausible that UGT pseudogenes may be exhibiting a regulatory effect on UGT expression by acting as miRNA sponges. In vitro knockdown of pseudogenes present in cell lines will better assess this possibility and help quantify any effects on UGT expression. It is possible that UGT1A
psuedogenes play a role in buffering miR-491-3p regulation of highly expressed UGT1A transcript in the liver, such as UGT1A1. The UGT1A v3 mRNA transcripts contain the endogenous miR-491-3p binding site and could therefore interact with mature miR-491-3p in the liver. Additionally, preliminary sequence analysis of the sequences of the UGT2B24, UGT2B25, and UGT2B26 psuedogenes contain a potential miR-216b binding site (data not shown). Confirmation of these interactions would not only identify a function purpose for UGT psuedogenes, but may also identify another mechanism regulating variable UGT expression in the liver and contributing to patients’ response to drugs.

4.2 Final Considerations

The research presented in this dissertation identifies a novel, epigenetic mechanism that contributes to UGT1A and UGT2B protein expression. Importantly, miRNA regulation provides an additional avenue to explain and explore the extensive interindividual variability with respect to UGT gene expression and function [64, 97, 119, 164]. Predicting and identifying patient responses to cancer and disease risk factors and drug regimens to treat those same ailments is a central pillar of personalized medicine and advanced healthcare and the UGT enzyme family contributes significantly to patient response to drug therapy [129, 132, 137, 153, 344]. miR-491-3p variable expression in the liver is significantly associated with reduced expression of UGT1A isoforms in human liver samples and may play a role in overall hepatic metabolic rates.

It is important to keep in mind that miRNA regulatory pathways and networks are generally modest regulators of gene expression [212, 213, 284] typically attenuating protein expression ~30-40%. However these modest effects can have profound
implications in the regulation of important genetic switches [193, 196], cellular
development and cancer [189, 191, 198, 203], and as discussed in this dissertation, drug
metabolism. It will be imperative to continually evolve our understanding of drug
response and toxicity. Patients rely on the most advanced medicines and medical
procedures to effectively respond to treatments. miRNA regulation of drug metabolism
contributes to this process and identifying these genes and the implications may provide
novel biomarkers and therapeutic targets to enhance patient responses by better
predicting their metabolic capabilities.

The research presented here has contributed to our understanding of UGT gene
expression and contributes to our knowledge of the mechanisms underlying the
significant interindividual variability in UGT gene expression. miRNA regulation of the
UGT gene enzymes may play a role in patient response not only to drug response but
also to cancer risk factors. Future endeavors investigating these hypotheses will only
further our knowledge of how patient’s respond to drug therapies and metabolism
environmental exposures.
Appendix:

Investigations into MicroRNA Regulation of Members of the Hepatic Nuclear Factor Family
Abstract

The hepatic nuclear factor (HNF) gene family codes for liver-enriched transcription factors belonging to the nuclear hormone receptor super family. HNF family members are essential in the normal development of the liver and are master regulators of DME expression, including most members of the UGT family. The purpose of this study is two-fold: to evaluate miRNA regulation of HNF4α and to identify novel miRNA regulators of HNF-family members. Transfection of miR-34a mimic into HepG2 cells significantly decreased UGT1A6 mRNA levels (P<0.05), a transcriptional target of HNF4α. UGT2B15 mRNA levels increased >100-fold in HepG2 cells transfected with miR-34a mimic. In vitro over-expression studies have identified miR-216b as a novel regulator of HNF4α expression. HNF4α luciferase reporters were significantly repressed in the presence of miR-216b mimics in a dose-dependent manner (P<0.001). Transfection of miR-216b mimics in HuH-7 and Hep3B liver cancer cell lines repressed HNF4α protein expression. In silico analysis of related HNF1α and HNF6α 3’ UTR mRNA sequences identified miR-24, miR-34a, and miR-216b as potential regulators of HNF-family expression. Luciferase reporter assays confirmed the significant interaction between miR-24 and the HNF6α 3’ UTR (P<0.01). These preliminary results identify miR-216b as a potential novel regulator of HNF4α expression and miRNA regulation of HNF4α and other HNF-family members has downstream implications in the transcriptional regulation of the UGTs. Further investigation of these findings are required to discern the impact of miR-216b regulation on HNF family members.
Introduction

Investigations discussed in this dissertation have identified several key miRNA regulators of UGT1A (miR-491-3p) and UGT2B (miR-216b) gene expression. However, miRNA genes have also been identified to regulate transcription factors essential in the expression of several DME families, including the UGTs. I investigated miRNA regulation of transcription factors involved in the regulation of UGT enzymes. The hepatic nuclear factor (HNF) transcription factor family consists of nine liver-enriched transcription factors belonging to the nuclear hormone receptor super family: HNF1α, HNF1β, HNF3α, HNF3β, HNF3γ, HNF4α, HNF4γ, HNF6α, and HNF6β. The HNF gene family is essential for the normal growth and development of the liver [345], and is instrumental in the regulation of several DME families in the liver and in other metabolic tissues including the colon and pancreas [157, 162]. This includes members of the UGT family [157, 346, 347]. Specifically, HNF1α can bind to the promoters of UGT1A3, 1A4, 1A7, 1A8, 1A9, 1A10, 2B7, and 2B10 to regulate their transcription [157, 348]. HNF3α has been identified to bind to the promoter of UGT2B10 (Jones et. al, In press, DMD) and HNF4α can bind to the promoters of UGTs 1A1, 1A6, 1A9, 2B11, 2B15 [157, 162, 349].

Takagi et al. identified two miRNA genes, miR-24 and miR-34a, which regulate expression of HNF4α in HepG2 liver cancer cells [252]. Regulation of HNF4α via miRs 24 and 34a affected downstream regulation of CYP7A1, an enzyme involved in the synthesis of bile acids, as well as other HNF4α targets including PEPCK. The authors identified that when miR-24 and 34a are induced via the PCK/MAPK pathway, HNF4α levels are down regulated and in turn downstream CYP7A1 levels are reduced [252]. Ramamoorthy et al. independently confirmed miR-34a as a regulator of HNF4α and additionally identified miR-449a as a novel miRNA regulator of HNF4α expression in
HepG2 cells [350]. To date, none of the other hepatic nuclear factor families have been identified as direct targets of miRNA regulation.

Using *in silico* prediction algorithms, we hypothesized that additional miRNA genes regulate the expression of HNF4α expression in the liver. We have identified miR-216b, a known miRNA regulator of several UGT2B isoforms, as a potential regulator of HNF4α protein expression. Additionally, we hypothesized that the repression of HNF4α protein expression via miR-24 and miR-34a alters the downstream expression of UGT isoforms. This mechanism of UGT transcriptional regulation may also involve miRNA regulation of other HNF family members, including HNF1α and HNF6α. The following experiments outline preliminary studies identifying (i) miR-216b as a novel regulator of HNF4α expression and (ii) miRs 24 and 34a regulate UGT expression *indirectly* via suppression of HNF4α and other HNF-family members.

**Materials and Methods**

**Chemicals and Reagents.** The pGL3-Promoter luciferase and pRL-TK renilla plasmids were from Promega (Madison, WI). The MISSION 3’ UTR Lenti GoClones for UGT1A8 (#HUTR06217), HNF1α (#HUTR12602), and HNF6α (ONECUT1, #HUTR03611) were purchased from SwitchGear Genomics (Menlo Park, CA). All synthesized DNA oligos used for 3’ UTR amplification, SDM, and qRT-PCR analysis were from Integrated DNA Technologies, Inc (Coralville, IA) or Life Technologies (Carlsbad, CA). Lipofectamine 2000 transfection reagent was from Life Technologies.

miRVana miRNA miR-216b mimic (#MC12302), miRVana negative control mimic #1 (#4464058), miRVana miRNA miR-216b inhibitor (#12302), negative control inhibitor #1 (#4464076), miRVana miRNA miR-24 mimic (#10737) and miRVana miRNA miR-34a
mimic (# 11030) were purchased from Ambion (Austin, TX). For use with the Lenti GoClones, MISSION miRNA mimics for miR-216a (HMI0381), miR-216b (HMI0382), miR-491-3p (HMI0606), miR-24 (HMI0412), miR-34a (HMI0508), and negative control #1 (HMC0002) and BSA were purchased from Sigma-Aldrich (St. Louis, MO). Goat anti-HNF4α 1° antibody (sc-6557) and donkey anti-goat secondary antibody (sc-2020) were from Santa Cruz Biotechnology (Dallas, TX). Mouse anti-β-actin 1° antibody and goat anti-mouse secondary antibody were from Sigma and Thermo Scientific (Waltham, MA), respectively. Puromycin was purchased from Sigma Aldrich. All other chemicals used were purchased from Fisher Scientific (Pittsburgh, PA) unless otherwise specified.

**Cell Lines and Culture Conditions.** Human embryonic kidney cell line 293 (HEK293) and human hepatocellular carcinoma cell line HepG2 were purchased from the American Type Culture Collection (ATCC, Manassas, VA). The human hepatocellular carcinoma cell line HuH-7 was a kind gift from Dr. Jianming Hu (Penn State Hershey College of Medicine, Hershey, PA). HEK293 and HuH-7 cells were cultured in DMEM (Gibco, Carlsbad, CA) supplemented with 10% FBS (Atlanta Biologicals, Lawrenceville, GA) and 1% pen/strep (Gibco). HepG2 cells were cultured in DMEM supplemented with 10% FBS, 1% pen/strep, and 1% non-essential amino acids (Lonza, Basel, Switzerland). All cells were grown and maintained in 37°C at 95% air, 5% CO₂.

**miRNA Binding Site Predictions.** The 3’ UTRs of the HNF-family mRNA sequences were obtained from the NCBI Genome database. miRNA binding site predictions were obtained using, (i) TargetScan [222], scored with the Total Context+ score as described in Garcia et. al [295], and (ii) miRanda, algorithm v3.0 [220] with the following parameters: Gap Open Penalty, -8.00; Gap Extend, -2.00; Score Threshold, 50.00; Energy Threshold, -20.00 kcal/mol; Scaling Parameter, 4.00.
**Construction of reporter plasmids.** The pGL3-Promoter vector plasmids used for cloning in this study were constructed by inserting the HNF4α 3’ UTR into the XbaI restriction site downstream of a luciferase reporter gene. Briefly, primers (Forward 5′ – GCTATCTA-GAGTTATCTAGCAAGCCCAGCTGGG – 3′; Reverse 5′ – GCTATCTA-GAGCAACAGGATGCAGACCC – 3′) modified with the XbaI digest site were used to amplify the HNF4α 3’ UTR using cDNA from HepG2 cells. The miR-216b seed deletion plasmids were created via SDM using the QuikChange II Site Directed Mutagenesis Kit (Agilent, Santa Clara, CA) using primers for miR-216b Site #1 (Forward 5′-CTCCCTTCTCCCTAATTAACCTATTTTGTGTTT-TATTCTCCT-3; Reverse 5′-AGGAGAATAAAACAAAAAAACAATGGTTAATAGGGAGGAAG-GGAG-3) and miR-216b Site #2 (Forward 5′-CCACTCCCTCCTAACCTATTGTTACAGAAGC-TGAA-3; Reverse 5′-TTCAGCTTCTGTAATAGGTTAGGAGGGAGTGGA-3). Nucleotide sequences of all plasmids used in this study were confirmed by DNA sequencing analysis performed at the Pennsylvania State University Nucleic Acids Core Facility (State College, PA).

**Generation of Stable Luciferase Cell Lines.** HEK293 cells were transduced with MISSION 3′ UTR Lenti GoClones containing a luciferase reporter vector attached to the 3′ UTR for the individual UGT1A8, HNF1α, and HNF6α genes. HEK293 cells were seeded in 24-well plates at an initial density of 50,000 cells/well and transduced 24 hours later with lentivirus at a multiplicity of infection (MOI) of 5. Media was changed to fresh media without virus 8 hours after transduction and 72 hours later cells were grown under selection of puromycin (0.5 µg/mL) to select for luciferase-expressing cells. HEK293 cells resistant to puromycin selection were grown for two weeks and used for downstream luciferase analysis. Frozen stocks of each cell line were made in fresh media with 10% DMSO and stored in liquid nitrogen.
Luciferase Assays. The pGL3-Promoter vector cloned with the HNF4α 3’ UTR was co-transfected with pRL-TK renilla vector and miRVana miRNA mimics into HEK293 cells. HEK293 cells were seeded at 50,000 cells/well into 24-well plates, and 24 hours following, 380 ng of pGL3 plasmid and 20 ng pRL-TK plasmid were co-transfected together with either scrambled miRNA mimic control or miRVana miRNA mimics using Lipofectamine 2000 according to manufacturer’s protocol. Twenty-four hours post-transfection cells were harvested using passive lysis buffer and luciferase activity was measured with a luminometer (Bio-tek Synergy HT, Winooski, VT) using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI). Renilla luciferase served as the internal control. HEK293 cell lines stably-expressing MISSION Lenti GoClones were seeded in 24-well plates at 50,000 cells/well and 24 hours later each cell line was transfected with either scrambled negative control or MISSION miRNA mimics at varying final concentrations using Lipofectamine 2000. Fresh media was replaced 8 hours later. 24 hours post-transfection, cells were harvested and luciferase activity was measured with a luminometer (Bio-tek Synergy HT, Winooski, VT) using the MISSION LightSwitch Luciferase Assay Kit according to manufacturer’s protocol (SwitchGear Genomics).

Tissues and miRNA Isolation. Liver specimens and their matching total RNA were obtained from the H. Lee Moffitt Cancer Center Tissue Procurement Facility (n=39). All protocols involving the collection and analysis of tissue specimens from these tissue banks were approved by the Institutional Review Boards at their respective institutions and were in accordance with assurances filed with and approved by the United States Department of Health and Human Services. All tissue samples were isolated and frozen at -70°C within 2 h post-surgery. Liver and cell line total RNA was extracted using the RNeasy Mini Kit (Qiagen, Hilden, Germany). All total RNA samples from cell lines were subject to on-column DNAse digestion during RNA purification.
Small RNA (<200 nt) containing the miRNA fraction was isolated and purified from total RNA using the mirVana miRNA Isolation Kit (Ambion). All RNA concentrations were ascertained using a Nanodrop ND-1000 spectrophotometer (Thermo Scientific) and was eluted and stored in RNase-, DNAse-free water in a -80°C freezer.

**Quantitative Real-Time (qRT)-PCR.** cDNAs were synthesized from cell line total RNA samples using the SuperScript First-Strand Synthesis Kit (Invitrogen, Carlsbad, CA). Taqman gene expression assays (Applied Biosystems, Carlsbad, CA) were used to amplify HNF4α (Hs00230853_m1), UGTs 1A1 (Hs02511055_s1), 1A6 (Hs01592477_m1), 2B15 (Hs_00870076_gH), RPLPO endogenous control (Hs99999902_m1), and RLP41 endogenous control (Hs00606029_g1). PCR reactions were performed in 10 µL reactions in 384-well plates using an ABI 7900HT Sequence Detection System with incubations done at 50°C for 2 min; 95°C for 10 min; and 40 cycles of 95°C for 15 sec, 60°C for 1 min. Reactions included 2x Universal PCR Master Mix (Applied Biosystems), Taqman gene expression primers and probes, water, and corresponding cDNA according to manufacturer’s protocol. Each plate was run with negative control (no cDNA template) and all samples were run in quadruplicate. Gene expression was compared to an endogenous, internal control gene (RPLPO or RLP41) using the \(2^{\Delta\Delta C_T}\) method [296]. Ct values were determined using the SDS 2.4 software (Applied Biosystems) and amplification Ct values higher than 35 cycles were designated as BLD. Samples lacking any amplification curves were also designated BLD.

**Western Blot Analysis.** Preparation of cell homogenates for protein analysis was performed as previously described [297]. HuH-7, Hep3B, or HepG2 cells were transfected with either scrambled miRNA controls, miR-24, 34a, or 216b mimic, or miR-217b inhibitor. Collected cell pellets were freeze-thawed 3x in liquid nitrogen and subjected to 3x 10-sec pulses using a hand-held Bio-Vortexer (Biospec, Bartlesville,
OK) prior to storing of 50 µL aliquots at -80°C. Protein concentrations within the cell homogenates were quantified using the BCA Protein Assay Kit (Pierce Chemical, Rockford, IL) and measured using an Appliskan Luminometer and SkanIT Software v2.3 (Thermo Scientific).

HNF4α and β-actin protein levels were determined via immunoblotting. Cellular homogenate was adjusted to equal volumes of loading buffer and heated at 90°C for 10 min. Samples were run at 90 V on a 10% acrylamide gel and transferred to a polyvinylidene difluoride (PVDF) membrane for 2 hrs at 33 V. PVDF membranes were blocked in 5% milk in TBS-T for 1 hr., probed with either HNF4α 1° antibody (1:1000) or β-actin 1° antibody (1:7500) overnight at 4°C, washed 3x in TBS-T, followed by incubation of appropriate second antibody (1:5000) for 1 hr. at room temperature, and washed 3x for 10 min in TBS-T. Protein bands were visualized using the Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare, Pittsburgh, PA) and Hyoblot CL autoradiography film (Denville Scientific, Metuchen, NJ). Blots were stripped and reprobed for β-actin levels which served as a loading control.

**Statistical Analysis.** Statistical analysis was performed using the Graphpad Prism 5 software (La Jolla, CA). The Student’s t-test (2-sided) or One-Way ANOVA analyses were used. For statistical analysis of expression levels that were B.L.D., it was necessary to calculate the level of gene expression using a Ct value of 40 to achieve statistical determination. Only for groups labeled BLD was this performed. Protein densitometry of HNF4α protein bands and β-actin control bands were quantified using Image J software (NIH, Bethesda, MD). Correlation analysis between HNF4α mRNA and miR-216b expression levels used the Spearman nonparametric analysis.
Results

**miR-216b binds to the HNF4α 3' UTR.** The HNF4α mRNA sequence (Fig. A1.1A) contains two binding sites for each miR-24 (blue) and miR-34a (red) in the HNF4α 3' UTR. The HNF4α 3' UTR contains two miR-216b binding sites downstream of the non-functional miR-24 binding sites predicted *in silico* by the miRanda and TargetScan algorithms. These sites are labeled miR-216b binding site #1 and miR-216b binding site #2, respectively (Fig. A1.1A, green). When a luciferase reporter containing the HNF4α 3' UTR is co-expressed with 10 nM miRNA mimics, miR-34a but not negative control miR-24, significantly represses ($P<0.001$) luciferase expression (Fig. A1.1B). This result is expected, as miR-24 does not interact with the HNF4α 3' UTR [252]. miR-216b significantly ($P<0.001$) represses expression of the HNF4α 3' UTR luciferase reporter at 10 nM (Fig. A1.1B). Additionally, a significant dose-dependent response in luciferase expression is seen with increasing concentrations of miR-216b ($P<0.001$; Fig. A1.1C). These results indicate that miR-216b can bind and interact with the HNF4α 3' UTR and confirm the previously reported findings that miR-34a also interacts with the same 3' UTR.

We next wanted to determine which of the miR-216b binding sites (#1 or #2 or both) were the functional binding sites in the HNF4α 3' UTR. Mutant 3' UTR plasmids were created by deleting the four nucleotides GAGA within the miR-216b binding sites corresponding to the miR-216b ‘seed’ sequences. An additional plasmid containing mutations within both binding sites was also created. Wild-type 3' UTR luciferase activity was significantly repressed ($P<0.01$) in the presence of 10 nM miR-216b mimic, serving as a positive control (Fig. A1.1D). Mutation of the miR-216b binding site #1 did not alter luciferase expression in the presence of miR-216b, but luciferase repression was significantly ($P<0.05$) lost when binding site #2 was mutated (Fig. A1.1D). This loss of
repression was also observed when both binding sites are mutated together and compared to wild-type ($P<0.05$; Fig. A1.1D). Taken together, these results indicate that miR-216b functionally binds to binding site #2 within the HNF4α 3' UTR to inhibit luciferase expression.
miR-216b represses HNF4α protein levels *in vitro*. We next investigated the impact of *in vitro* miR-216b over-expression in HuH-7 cells to determine if endogenous HNF4α mRNA and protein levels are repressed. In the presence of 10nM and 50 nM miR-216b mimics, HuH-7 HNF4α protein levels are repressed ~40% but there was no change in HNF4α mRNA levels (Fig. A1.2A). In Hep3B cells, HNF4α mRNA levels were significantly reduced (*P*<0.05) ~27% in the presence of miR-216b mimics and protein levels were repressed nearly two-fold (Fig. A1.2B).
Figure A1.2. Repression of HNF4α protein expression by miR-216b.
A, Transfection of miR-216b mimics into HuH-7 cells for repressed HNF4α protein expression ~40% (compared to scrambled negative control). mRNA levels of HNF4α were not significantly changed. B, Transfection of miR-216b mimic into Hep3B cells repressed HNF4α protein levels ~47% and mRNA levels ~27% D, Knockdown of mir-216b in Hep3B cells did not significantly alter HNF4α mRNA levels. All columns represent ±SEM of three independent experiments. *, P<0.05.

Correlation of miR-216b and HNF4α in human liver. The mRNA levels of HNF4α were quantified using qRT-PCR analysis from RNA extracted from 38 normal human liver tissue samples and correlated with miR-216b expression levels in the same samples. There was no correlation between HNF4α mRNA and mir-216b expression in human liver (Fig. A1.3) indicating mir-216b expression levels are not predictive of HNF4α mRNA expression.
Figure A1.3. Expression of HNF4α mRNA versus miR-216b in human liver samples. Levels of HNF4α mRNA and miR-216b were quantified in 38 individual normal liver samples via qRT-PCR and quantified relative to the lowest-expressing sample of each gene (set to 1.0) and normalized to RPLPO and RNU6B, respectively. Dots represent individual human samples.

miR-24 and miR-34a repress HNF4α protein expression. Previous reports have identified that two miRNA genes, miR-24 and miR-34a, also regulate the expression of HNF4α and the expression of downstream HNF4α target genes, eg, CYP7A1 [252]. We investigated whether down regulation of HNF4α protein expression via miR-24 and miR-34a had an effect on the transcriptional regulation of UGTs regulated by HNF4α; chiefly, UGTs 1A1, 1A6, and UGT2B15. Small-inhibiting RNA (siRNA) targeting HNF4α were transfected into HepG2 cells and HNF4α mRNA and protein levels were assessed. HNF4α mRNA levels were significantly reduced \((P<0.05)\) in the presence of HNF4α siRNA when compared with scrambled control (Fig. A1.4A) and HNF4α protein levels were reduced \(~60\%)\) (Fig. A1.4C) In these cells, reduction of HNF4α protein expression significantly reduced mRNA levels of HNF4α-transcriptional targets UGT1A1 \((P<0.05)\) and 1A6 \((P<0.05)\) and increased UGT2B15 mRNA \((P<0.05)\) (Fig. A1.4D).
HNF4α mRNA levels transfected with 50 nM of either miR-24 or miR-34a mimic were not significantly repressed, although miR-34a reduced HNF4α mRNA levels nearly two-fold (Fig A1.4B). HNF4α protein levels were repressed in the presence of miR-24 and miR-34a mimics by ~70% and ~75%, respectively, and at similar levels to siRNA controls (Fig. A1.4C). In the presence of miR-24 mimic, UGT mRNA levels of UGT1A1, 1A6, and 2B15 were not significantly altered (Fig. A1.4, E). In the presence of miR-34a mimic, UGT mRNA levels of UGT1A6 were significantly repressed by ~50% (P<0.05; Fig. A1.4E). Similar to siRNA controls, UGT2B15 mRNA levels were increased >150-fold (Fig. A1.4E). UGT2B15 mRNA levels increase when HNF4α is knocked-down with siRNA or with over-expression of miR-34a (Fig. A1.4, D and E). To explain the apparent differences between UGT1A6 mRNA levels in siRNA-or-miRNA-transfected HepG2 cells, we hypothesized that miRs 24 and 34a regulate the expression of additional HNF-transcription factors and this may be influencing the changes in UGT1A6 expression. We next sought to determine if other HNF-family members are regulated by miR-24, miR-34a, or miR-216b.
Figure A1.4. miRNA Regulation of HNF4α expression. A, HepG2 cells were transiently transfected with HNF4α siRNA and HNF4α mRNA levels were quantified using qRT-PCR. HNF4α were repressed in the presence of 50 nM siRNA. B, HepG2 cells were transfected with 50 nM miR-24, miR-34a, or scrambled control miRNA mimics. C, HNF4α protein levels were repressed in the presence of HNF4α siRNA when compared with scrambled control. miR-24 and miR-34a over-expression reduced HNF4α protein levels by ~70% and ~75%, respectively. β-actin served as a loading control. D and E, mRNA levels of HNF4α target genes UGT1A1, 1A6, and 2B15 were measured using qRT-PCR in the presence of siRNA targeting HNF4α (D) or miRNA (E). UGT1A1 mRNA were significantly reduced ~2-fold in the presence of siRNA, UGT1A6 mRNA were significantly increased ~1.8 fold, and UGT2B15 levels were drastically increased ~90-fold. miR-34a reduced UGT1A6 mRNA levels ~2-fold. UGT2B15 levels mRNA increased ~150-fold in the presence of over-expressed miR-34a. RLP41 served as an internal control for gene expression analysis. All columns represent ±SEM of two independent experiments. *, P<0.05.
**In silico analysis of miRNA binding to HNF-family mRNAs.** Both the miRanda algorithm [220] and TargetScan algorithm [222] were used to predict the binding of miRNA gene candidates miR-24, miR-34a, and miR-216b to the 3’ UTR mRNA sequences of HNF family members. miR-24 and miR-34a were chosen because of their previously identified role in regulating HNF4α [252] and miR-216b was chosen because of functionally conserved binding sites in the 3’ UTRs of UGT2B4, 2B7, 2B10, and 2B15 (Chapter 3). HNF4α can bind to the promoters of several UGTs, including UGT2B15 [162]. Both programs predict that miR-24, miR-34a, and miR-216b will bind to the 3’ UTR of HNF4α, as already determine, and both predict miR-34a will bind to the HNF1α 3’ UTR (Table A1.1). TargetScan predicts that miR-24 will bind to HNF1α, while miRanda predicts miR-24 binding sites within the 3’ UTRs of HN1α, HNF1β, HNF4γ, and HNF6α. miRanda predicted miR-34a bindings sites with the 3’ UTRs of HNF4γ and a miR-216b binding site in the 3’ UTR of HNF4γ (Table A1.1). Neither miR-24, miR-34a, nor miR-216b have predicted binding sites in the 3’ UTRs of any HNF3 isoform or HNF6β.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Predicted MicroRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNF1α</td>
<td>miR-24 (TargetScan); miR-34a (both)</td>
</tr>
<tr>
<td>HNF1β</td>
<td>miR-24 (miRanda)</td>
</tr>
<tr>
<td>HNF3α</td>
<td>None</td>
</tr>
<tr>
<td>HNF3β</td>
<td>None</td>
</tr>
<tr>
<td>HNF3γ</td>
<td>None</td>
</tr>
<tr>
<td>HNF4α</td>
<td>miR-24 (both); miR-34a (both); miR-216b (both)</td>
</tr>
<tr>
<td>HNF4γ</td>
<td>miR-24 (miRanda); miR-34a (miRanda); miR-216b (miRanda)</td>
</tr>
<tr>
<td>HNF6α</td>
<td>miR-24 (miRanda)</td>
</tr>
<tr>
<td>HNF6β</td>
<td>None</td>
</tr>
</tbody>
</table>

**Table A1.1 Summary of in silico predictions identifying miRNAs targeting HNF 3’ UTRs.** Both miRanda and TargetScan prediction algorithms were used to predict miR-24, miR-34a, and miR-216b binding to the 3’ UTRs of HNF-family member genes.
Regulation of HNF-family 3’ UTRs via miRs-24, 34a, and 216b. We next sought to validate the *in silico* predictions of miRs 24, 34a, and 216b targeting the 3’ UTRs of two of these transcription factors. HNF1α was chosen specifically for its role in the regulation of most UGT isoforms in the liver. Stable cell lines expressing a luciferase reporter gene with attached 3’ UTRS for the UGT1A family, HNF1α, and HNF6α were created using HEK293 cells. The luciferase reporter cell lines were subject to transient transfection of MISSION miRNA mimics for miRs 491-3p, miR-216a, miR-216b, miR-24, or miR-34a to validate the *in silico* binding predictions. HEK293 cell lines over-expressing the luciferase reporter with attached UGT1A 3’ UTR were studied as a control for the experimental system, as it has already been established that miR-491-3p can bind and regulate the UGT1A 3’ UTR (Chapter 2). We first tried over-expressing miR-491-3p mimic with UGT1A 3’ UTR luciferase cell lines at final concentrations of 10 nM and 100 nM. There was significant repression (*P*<0.001) of luciferase reporter activity in the presence of 100 nM miR-491-3p mimic compared to scrambled control, but not at 10 nM (Fig. A1.5A). We next investigated the HNF1α and HNF6α 3’ UTR cell lines using final concentrations of 100 nM and 200 nM miRNA mimics to determine their effect. When the HNF1α 3’ UTR luciferase report cell line was transfected with miR-216a, 216b, 24, and 34a mimics significant repression of luciferase activity was not observed at either concentration (Fig. A1.5B). There was significant repression of luciferase activity attached to the HNF6α 3’ UTR in the presence of miR-24 at 100 nM (*P*<0.01) and 200 nM (*P*<0.05) confirming the miRanda *in silico* prediction (Fig. A1.5C). miR-216a was also tested but did not interact with any of the 3 UTRs (results not shown).
Figure A1.5. Validation of *in silico* miRNA target predictions of HNF-family 3′ UTRs. Over-expressing luciferase reporters with attached 3′ UTRS from the UGT1A gene family (A), HNF1α (B), or HNF6α (C) genes. A, The UGT1A 3′UTR luciferase cell line was transiently transfected with miR-491-3p (positive control) and there was significant repression of luciferase activity in the presence of 100 nM miR-491-3p mimic. B, The HNF1α 3′ UTR luciferase cell line was transfected with miRNA mimics for miRs 216a, 216b, 34a, and 24. There was no observable repression in luciferase activity. C, The HNF6α 3′ UTR luciferase cell line was transfected with the same mimics as in (B). Columns represent ±SEM of four independent experiments. *, P<0.05; **, P<0.01; ***, P<0.001.
Discussion

The findings reported here represent preliminary investigations to determine the extent of miRNA regulation of HNF4α and other related HNF family members. The \textit{in silico} algorithms miRanda [220] and TargetScan [223] predicted miR-216b to bind to the HNF4α 3’ UTR at two distinct sites. This prediction was validated \textit{in vitro} using a luciferase reporter assay as well as in the HuH-7 and Hep3B liver cancer cell lines in which HNF4α protein expression is repressed in the presence of over-expressed miR-216b mimics. No significant correlation between miR-216b expression levels and HNF4α mRNA expression levels were observed, but this result is not surprising because knockdown of endogenous miR-216b expression levels in HuH-7 does not affect HNF4α mRNA levels (Fig. A1.2). This suggests that miR-216b’s primary mechanism of regulating HNF4α protein expression is via translational inhibition as opposed to mRNA cleavage. However, there was significant knockdown of HNF4α mRNA in Hep3B cells, and this may reflect cell-specific mechanisms of miR-216b regulation similar to those observed in the interaction of miR-491-3p with UGT1A1 mRNA (Chapter 2). Additionally, our observations are consistent with previous studies conducted by other laboratories, as miR-449a does not reduce HNF4α mRNA levels in HepG2 cells while still repressing HNF4α protein expression [350].

The interaction between miR-216b and HNF4α could be further tested by knockdown of endogenous miR-216b in HuH-7 and Hep3B cells or primary liver cells to assess the impact on HNF4α mRNA and protein expression. The correlation between miR-216b expression and HNF4α protein expression can also be tested in several other way. For instance, Analysis of miR-216b expression levels can be measured using either qRT-PCR [248, 252] or immunohistochemistry using locked nucleic acid probes specific
for miR-216b. miR-216b expression levels can then be correlated with HNF4α protein expression as measured with specific antibodies measuring HNF4α protein either in western blots or using in situ hybridization [351]. Analyses using locked-nucleic acids and fluorescent antibodies directly measure protein and miRNA together at the same time and offer the most precise method for correlation within tissue or cellular samples [193].

As discussed in Chapter 3, miR-216b can bind to the 3’ UTRs of several UGT2B genes including UGT2B4, 2B7, 2B10, and 2B15 and HNF4α can bind to the promoters of UGTs 1A1, 1A6, 1A9, and 2B15 [157, 162, 349]. This highlights a potential regulatory mechanism involving miR-216b regulation of UGT2B enzymes, as well as one of their transcriptional regulators. Regulatory networks involving miRNA regulation of both transcriptional regulators and their downstream targets has been identified in other pathways. For example, miR-27b regulates both CYP3A4 expression and CYP3A4 transcription regulator VDR [247]. Conserved regulatory networks such as these offer robustness in gene regulation control affecting DME expression. The miR-216b:HNF4α pathway could be a contributing network of regulation to UGT enzyme expression in the liver. HNF4α protein has already been shown to bind to the promoter of UGT2B11 in hepatocytes [162] and this interaction may also be influenced by miR-216b function as evidence in this study. Further investigation of the impact of miR-216b regulation of both HNF4α and HNF4α UGT targets in hepatocytes and other tissues is warranted to identify novel mechanisms regulating UGT expression. It is possible this network may influence the regulation of other metabolic transcriptional targets of HNF4α such as CYP7A1[252].

We investigated whether miR-216b can directly regulate several additional HNF family members related to HNF4α. miR-216b over-expression did not regulate luciferase reporters attached with the 3’ UTRs of HNF1α or HNF6α. This is important to note
because HNF1α regulates the expression of nearly all the UGT1A and 2B enzymes [157, 348] and HNF6α can bind to the promoters of UGTs 1A1, 2B11, and 2B15 in human hepatocytes [162]. HNF4α protein binds to the promoters of HNF1α and HNF6α to regulate their expression in liver [162, 352]. It is possible that miR-216b regulation of HNF4α in the liver may have downstream implications on the expression of HNF4α targets such as HNF1α or HNF6α. HNF4α and HNF1α together are involved with the regulation of nearly all the UGT enzymes, and both proteins can bind cooperatively to the same gene promoter at the same time [162]. Thus, miR-216b repression of HNF4α in liver could have profound downstream effects on the functionality of HNF1α and downstream UGT enzymatic. These questions will need to be investigated. Experiments in our laboratory have identified HNF3α as a promoter regulator of UGT2B10 expression in live cancer cells (Jones and Lazarus, unpublished). We did not test the 3' UTRs for the remaining HNF family members, including HNF3α because of availability of stable luciferase reporter vectors used in this analysis. Interestingly, neither miRanda or TargetScan predicted miR-216b (or miR-24 or miR-34a) to bind to the 3' UTR of HNF3 isoforms. However, this does not rule out the possibility additional miRNA regulators interact with HNF3 isoforms mRNAs.

Investigations in our laboratory confirmed previous reports observing down regulation of HNF4α protein expression in HepG2 cells with miR-24 and miR-34a mimics (Fig. A1.4) [252]. Significant down regulation of UGT1A1 mRNA expression occurred when HNF4α protein was repressed with siRNA, but no change in UGT1A1 mRNA levels occurred in the presence of miR-24 and miR-34a had a modest effect on UGT1A1 mRNA levels. Compared to the siRNA control, it is curious that UGT1A1 mRNA levels did not significantly change when HNF4α protein was repressed by miR-24 and miR-34a. It is possible that UGT1A1 protein levels are still changed via translation inhibition
due to miR-24 and miR-34a function. UGT1A1 protein and activity levels should be measured to ascertain any differences in protein expression when HNF4α is repressed by miRs 24, 34a, and 216b. Regulation of HNF4α protein in HepG2 cells may not have major implications in the regulation of UGT1A1 and several other cell lines and/or hepatocytes could be used to ascertain any effects of miR-24, miR-34a, or miR-216b on UGT1A1 regulation. UGT1A6 mRNA levels were significantly increased in the presence of HNF4α siRNA, but there was significant down-regulation of UGT1A6 in the presence of miR-34a (Fig. A1.4). These are contradictory results because both sets of experiments have reduced expression of HNF4α protein levels. These observations may be explained by the fact that in the presence of reduced HNF4α protein levels, additional regulatory proteins bind and regulate UGT1A6 expression, such as HNF1α. Gel-mobility shift assays and additional promoter binding experiments can be used to better ascertain the promoter conditions of UGT1A6 in the presence/absence of miRNA mimics targeting HNF4α [353, 354].

We did observe an increase in the expression of UGT2B15 mRNA in HepG2 cells in the presence of HNF4α siRNA as well as miR-34a. HNF4α can bind to the promoter of UGT2B15 in hepatocytes as well as pancreatic islet cells [162]. It is unknown at this time whether HNF4α binding to the UGT2B15 promoter acts as a transcriptional repressor, but out data suggests this mechanism due to the increase in UGT2B15 mRNA when HNF4α is repressed via either siRNA or miR-34a. The complete transcriptional regulation of UGT2B15 is unknown and further exploration of this topic will be needed to confirm these preliminary results.

miR-24 and miR-34a may target additional HNF-family members at the same time, thereby affecting and confounding the regulation of downstream UGT targets. In silico prediction of miR-24 and miR-34a gene targets identified miR-24 and miR-34a as
potential regulation of HNF1α. miR-24 is predicted to target HNF6α (Table A1.1). Because miRanda and TargetScan weigh similar criteria differently for miRNA target prediction [220, 295], it is not surprising to see that only miRanda predicted miR-24 to target HNF6α. It is possible that the in silico models may also miss potential miRNA regulators due to the stringent criteria used in the prediction and also owing to the promiscuity of miRNA binding leading to off-target effects. We confirmed that miR-24 regulates the 3’ UTR of HNF6α (Fig. A1.5). We did not observe miR-24 or miR-34a regulation of HNF1α’s 3’ UTR; however, in mice the conserved miR-24 gene regulates the expression of HNF1α in pancreatic cells [355]. Additional experiments are required to confirm these initial findings, including identifying the downstream UGT expression when HNF6α is regulated by miR-24 over-expression. Knockdown of endogenous miR-24 and miR-34a may provide insight into the expression of downstream UGT isoforms, possibly in other tissue types such as the pancreas or colon.

Future studies analyzing changes in global UGT or other DME expression in the presence of over-expressed miR-24 or miR-34a may help ascertain the role of these miRNA and UGT regulation in extra-hepatic metabolic tissues such as the pancreas, colon, kidney, and intestines. HNF4α is important in the development and regulation of metabolic tissues including the pancreas, intestines, and kidney [162, 356, 357]. It would be interesting to identify any changes in UGT expression via miRs 24/34a/216b over-expression or knockdown in cells lines from these extra-hepatic tissues. This is especially important in the pancreas where miR-216b exhibits very high expression and could be having a profound impact on HNF4α protein expression (see Chapter 3). Conserved changes in UGT expression via miRNA regulation of HNF4α in these tissues would provide evidence of the importance of this transcriptional regulatory network governing UGT expression. Comparing and contrasting miR-216b expression with
HNF4α protein levels between liver and the pancreas may elucidate inverse correlations in their expression.

In summary, preliminary evidence suggests that miRNA regulation of HNF4α and other HNF-family members may have implications with respect to downstream expression of UGT isoforms. Results indicate that miR-216b regulates expression of UGT2B mRNA levels, including UGT2B15, via regulation of HNF4α protein expression and suggest that UGT transcriptional regulation may be quite complex.
References


131. Li, Y., et al., *Genetic polymorphisms in the TATA box and upstream phenobarbital-responsive enhancer module of the UGT1A1 promoter have combined effects on UDP-glucuronosyltransferase 1A1 transcription mediated by


163. Wortham, M., et al., *Expression of constitutive androstane receptor, hepatic nuclear factor 4 alpha, and P450 oxidoreductase genes determines interindividual variability in basal expression and activity of a broad scope of

164. Aueviriyavit, S., et al., Hepatocyte nuclear factor 1 alpha and 4 alpha are factors involved in interindividual variability in the expression of UGT1A6 and UGT1A9 but not UGT1A1, UGT1A3 and UGT1A4 mRNA in human livers. Drug Metab Pharmacokinet, 2007. 22(5): p. 391-8.


301. Trdan Lusin, T., J. Trontelj, and A. Mrhar, *Raloxifene glucuronidation in human intestine, kidney, and liver microsomes and in human liver microsomes*


CURRICULUM VITAE

NAME
Douglas F. Dluzen

POSITION TITLE
Graduate Assistant, Genetics Program

EMAIL ADDRESS
ddluzen@hmc.psu.edu

Department of Pharmacology
Penn State College of Medicine

EDUCATION/TRAINING

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Illinois, Urbana-Champaign, IL</td>
<td>B.S.</td>
<td>2004-2007</td>
<td>Molecular and Cellular Biology</td>
</tr>
<tr>
<td>Pennsylvania State University, College of Medicine, Hershey, PA</td>
<td>Ph.D.</td>
<td>2007-2014</td>
<td>Genetics</td>
</tr>
</tbody>
</table>

A. Positions and Honors

Positions:

2008-2014 Graduate Assistant, Laboratory of Dr. Philip Lazarus, Department of Pharmacology, Penn State College of Medicine

Awards and Honors:

2013 Penn State College of Medicine Endowed Scholarship
2012 Sigma Aldrich and SwitchGear Genomics miRNA Target Validation Grant
2012 Penn State College of Medicine Class of 1971 Endowed Scholarship
2011 AstraZeneca Co-Funded Bioscience Ph.D. Studentship Grant
2011 Penn State College of Medicine Endowed Scholarship
2011 Penn State College of Medicine Graduate Alumni Endowed Scholarship
2007 Penn State University Graham Award

B. Publications

1. **Dluzen, DF.**, Sun, D., Salzberg, A., Jones, N., Bushey, RT., Robertson, GP., Lazarus, P. Regulation of UGT1A1 expression and activity by miR-491-3p. (Submitted to *J Pharmacol Exper Ther*)
5. **Dluzen, DF.**, Sun, D., Salzberg, A., Ishmael, FT., Robertson, GP., Lazarus, P. miR-216b Regulation of UGT2B Family Enzymes. (In preparation)

C. Conference Presentations